DT03 Rec'd PCT/PTO 04 AUG 2004

FORM PTO 1390 (REV 10-2003) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: M0025.0311/P311

U.S. APPLICATION NO. (If known, see 37 CFR 1.5): 10/503463

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/GB03/00459	4 February 2003	4 February 2002

TITLE OF INVENTION: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

APPLICANT(S) FOR DO/EO/US: Mark Catchpole

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. [x] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing 35 U.S.C. 371.
3. [x] This is an express request to begin national examination procedures (35 U.S.C. 371 (f)). The submission must include items (5), (6), (9) and (21) indicated below.
4. [x] The US has been elected (Article 31).
5. [x] A copy of the International Application as filed (35 U.S.C. 371 (c)(2))
   a. [x] is attached hereto (required only if not communicated by the International Bureau).
   b. [ ] has been communicated by the International Bureau.
   c. [ ] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [ ] An English language translation of the International Application as filed (35 U.S.C. 371 (c)(2)).
   a. [ ] is attached hereto.
   b. [ ] has been previously submitted under 35 U.S.C. 154(d)(4).
7. [ ] Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   a. [ ] are attached hereto (required only if not communicated by the International Bureau).
   b. [ ] have been communicated by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [ ] have not been made and will not be made.
8. [ ] An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).
9. [ ] An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
10. [ ] An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

Items 11 to 20 below concern document(s) or information included:

11. [x] An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. [ ] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [x] A preliminary amendment.
14. [ ] An Application Data Sheet under 37 CFR 1.76.
15. [ ] A substitute specification.
16. [ ] A power of attorney and/or change of address letter.
17. [ ] A computer-readable form of the sequence listing in accordance with PCT Rule 13*ter*.2 and 37 CFR 1.821 - 1.825.
18. [ ] A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. [ ] A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
20. [x] Other items or information: Intl. Search Report

IPR2023-00033
Apple EX1002 Page 1

U.S. APPLICATION NO. (if known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
10/503463	PCT/GB03/00459	M0025.0311/P311

21. [x] The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) – (5)):

		CALCULATIONS PTO USE ONLY
[ ]	Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO $1080.00	
[x]	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO $920.00	
[ ]	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO $770.00	
[ ]	International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) $730.00	
[ ]	International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) $100.00	

ENTER APPROPRIATE BASIC FEE AMOUNT =					$ 920.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than 30 months from the earliest claimed priority date (37 CFR 1.492 (e)).					$	
CLAIMS	**NUMBER FILED**	**NUMBER EXTRA**	**RATE**			
Total claims	91-20 =	71	x	18.00	$ 1,278.00	
Independent claims	6-3 =	3	x	86.00	$ 258.00	
MULTIPLE DEPENDENT CLAIM(s) (if applicable)			+		$	
TOTAL OF ABOVE CALCULATIONS =					$ 2,456.00	
[x] Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by ½.					$ 1,228.00	
SUBTOTAL =					$ 1,228.00	
Processing fee of **$130.00** for furnishing the English translation later than 30 months from the earliest claimed priority date (37 CFR 1.492 (f)).					$	
TOTAL NATIONAL FEE =					$ 1,228.00	
Fee for recording the enclosed assignment (37 CFR 1.21 (h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +					$	
TOTAL FEES ENCLOSED =					$ 1,228.00	
					Amount to be refunded:	$
					charged:	$

a. [ ] A check in the amount of $ _____ to cover the above fees is enclosed.

b. [ ] Please charge my Deposit Account No. _____ in the amount of $ _____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. [x] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 04-1073. A duplicate copy of this sheet is enclosed.

d. [x] Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
Stephen A. Soffen
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street NW
Washington, DC 20037-1526
(202) 828-4879
CUSTOMER NUMBER: 24998

SIGNATURE: Stephen A. Soffen

NAME

31,063

REGISTRATION NUMBER

DT03 Rec'd PCT/PTO 04 AUG 2004

FORM PTO 1390 (REV 10-2003) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: M0025.0311/P311

U.S. APPLICATION NO. (If known, see 37 CFR 1.5): 10/503463

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/GB03/00459	4 February 2003	4 February 2002

TITLE OF INVENTION: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

APPLICANT(S) FOR DO/EO/US: Mark Catchpole

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. [x] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing 35 U.S.C. 371.
3. [x] This is an express request to begin national examination procedures (35 U.S.C. 371 (f)). The submission must include items (5), (6), (9) and (21) indicated below.
4. [x] The US has been elected (Article 31).
5. [x] A copy of the International Application as filed (35 U.S.C. 371 (c)(2))
   a. [x] is attached hereto (required only if not communicated by the International Bureau).
   b. [ ] has been communicated by the International Bureau.
   c. [ ] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [ ] An English language translation of the International Application as filed (35 U.S.C. 371 (c)(2)).
   a. [ ] is attached hereto.
   b. [ ] has been previously submitted under 35 U.S.C. 154(d)(4).
7. [ ] Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   a. [ ] are attached hereto (required only if not communicated by the International Bureau).
   b. [ ] have been communicated by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [ ] have not been made and will not be made.
8. [ ] An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).
9. [ ] An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
10. [ ] An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

Items 11 to 20 below concern document(s) or information included:

11. [x] An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. [ ] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [x] A preliminary amendment.
14. [ ] An Application Data Sheet under 37 CFR 1.76.
15. [ ] A substitute specification.
16. [ ] A power of attorney and/or change of address letter.
17. [ ] A computer-readable form of the sequence listing in accordance with PCT Rule 13*ter*.2 and 37 CFR 1.821 - 1.825.
18. [ ] A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. [ ] A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
20. [x] Other items or information: Intl. Search Report

U.S. APPLICATION NO. (if known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
10/503463	PCT/GB03/00459	M0025.0311/P311

21. [x] The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) – (5)):

		CALCULATIONS PTO USE ONLY
[ ]	Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO $1080.00	
[x]	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO $920.00	
[ ]	International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO $770.00	
[ ]	International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) $730.00	
[ ]	International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) $100.00	

ENTER APPROPRIATE BASIC FEE AMOUNT =	$ 920.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than 30 months from the earliest claimed priority date (37 CFR 1.492 (e)).	$	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	91-20 =	71	x 18.00	$ 1,278.00	
Independent claims	6-3 =	3	x 86.00	$ 258.00	
MULTIPLE DEPENDENT CLAIM(s) (if applicable)			+	$	
			TOTAL OF ABOVE CALCULATIONS =	$ 2,456.00	
[x] Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by ½.				$ 1,228.00	
			SUBTOTAL =	$ 1,228.00	
Processing fee of **$130.00** for furnishing the English translation later than 30 months from the earliest claimed priority date (37 CFR 1.492 (f)).				$	
			TOTAL NATIONAL FEE =	$ 1,228.00	
Fee for recording the enclosed assignment (37 CFR 1.21 (h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$	
			TOTAL FEES ENCLOSED =	$ 1,228.00	
			Amount to be refunded:	$	
			charged:	$	

a. [ ] A check in the amount of $ _____ to cover the above fees is enclosed.

b. [ ] Please charge my Deposit Account No. _____ in the amount of $ _____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. [x] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 04-1073. A duplicate copy of this sheet is enclosed.

d. [x] Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.495 has not been met, a petition to revive (37 CFR 1.137 (a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
Stephen A. Soffen
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street NW
Washington, DC 20037-1526
(202) 828-4879
CUSTOMER NUMBER: 24998

SIGNATURE: _____

Stephen A. Soffen
NAME

31,063
REGISTRATION NUMBER

SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

The present invention generally relates to a speech recognition circuit which uses parallel processors for processing the input speech data in parallel.

Conventional large vocabulary speech recognition can be divided into two processes: front end processing to generate processed speech parameters such as feature vectors, followed by a search process which attempts to find the most likely set of words spoken from a given vocabulary (lexicon).

The front end processing generally represents no problem for current processing systems. However, for large vocabulary, speaker independent speech recognition, it is the search process that presents the biggest challenge. An article by Deshmukh et al entitled "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition" (IEEE Signal Processing Magazine, September 1999, pages 84 to 107), the content of which is hereby incorporated by reference, discusses the general concepts of large vocabulary speech recognition. As discussed in this paper, one algorithm for performing the search is the Viterbi algorithm. The Viterbi algorithm is a parallel or breadth first search through a transition network of states of Hidden Markov Models. An acoustic model for words in a lexicon are represented as states of Hidden Markov Models. These states represent phones or n phones in a phone model of the words. The search requires the evaluation of possible word matches. It is known that such a search is computationally intensive.

In order to speed up the processing performed during such a search in a speech recognition system, parallel processing has been explored. In an article by M K Ravishankar entitled "Parallel Implementation of Fast Beam Search for Speaker-Independent Continuous Speech Recognition" (Indian Institute of Science, Bangalor, India, July 16, 1993) a multi-threaded implementation of a fast beam search algorithm is disclosed. The multi-threading implementation requires a significant amount of

communication and synchronization among threads. In an MSC project report by R Dujari entitled "Parallel Viterbi Search Algorithm for Speech Recognition" (MIT, February 1992) the parallel processing of input speech parameters is disclosed in which a lexical network is split statically among processors.

It is an object of the present invention to provide an improved circuit which can perform parallel processing of speech parameters.

In accordance with a first embodiment of the present invention, a speech recognition circuit comprises an input port such as input buffer for receiving parameterized speech data such as feature vectors. A lexical memory arrangement is provided which contains lexicon data for word recognition. The lexical data comprises a plurality of lexical tree data structures representing a plurality of lexical trees. Each lexical tree data structure comprises a model of words having common prefix components and an initial component which is unique as an initial component for lexical trees. A plurality of lexical tree processors are connected in parallel to the input port and perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory arrangement. A results memory arrangement is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory arrangement by performing parallel processing of a plurality of lexical tree data structures.

Thus in accordance with this embodiment of the present invention, the processing in order to perform word recognition is distributed across the processors by controlling the processors to perform processing on different lexical trees. The controller controls the processor by the processes to provide for efficient process management by distributing lexical processing to appropriate processors.

The lexical tree data structure can comprise a phone model of words, wherein the components comprise phones. For reduced storage, the lexical tree data structure can comprise a mono phone lexical tree. The mono phone lexical tree can be used to

generate context dependent phone models dynamically. This enables the use of context dependent phone models for matching and hence increased accuracy whilst not increasing memory requirements. Alternatively, the lexical tree data structure can comprise context dependent phone models.

The processing performed by each processor in one embodiment comprises the comparison of the speech parameters with the lexical data, e.g. phone models or data derived from the lexical data (e.g. dynamically generated context dependent phone models) to identify words as a word recognition event and to send information identifying the identified words to the results memory as the processing results. In this embodiment a language model processor arrangement can be provided for providing a language model output for modifying the processing results at a word recognition event by a lexical tree processor. The modification can either take place at each lexical tree processor, or at the language model processing arrangement.

In one embodiment each lexical tree processor determines an output score for words in the processing results at word recognition events. Thus in this embodiment the language model processing arrangement can modify the score using a score for a language model for n preceding words, where n is an integer.

In one embodiment the controller instructs a lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from the results memory. The history data preferably includes an accumulated score for the recognition path. This enables a score to be determined based on the score for the recognition path to accumulate a new score during recognition carried out using the lexical tree data structure. The scores can be output in the processing results to the results memory during the processing of the speech parameters so that the scores can be used for pruning.

In one embodiment of the present invention, each lexical tree processor operates on more than one lexical tree at the same time, e.g. two lexical trees represented by two different lexical tree data structures, or two lexical trees represented by the same data

structure but displaced in time (which can be termed to instances of the same lexical tree).

At word recognition events, the controller determines new lexical tree identifiers for storing in the results memory for words identified in the results memory for respective word events. In order to reduce the processing, the controller can prune the new lexical tree identifiers to reduce the number of lexical trees which are required to be processed. This pruning can be achieved using context dependant n phones to reduce the number of possible next phones. The number can be further reduced by using a language model look ahead technique.

In one embodiment of the present invention, the lexical tree processors are arranged in groups or clusters. The lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory is connected to one of the groups of lexical tree processors and contains part of the lexical data. Thus a group of lexical tree processors and a partial lexical memory form a cluster. Each lexical tree processor is operative to process the speech parameters using a partial lexical memory and the controller controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In another embodiment of the present invention the lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory being connected to one of the lexical tree processors and containing part of the lexical data. Each lexical tree processor processes the speech parameters using a corresponding partial lexical memory and the controller is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In one embodiment of the present invention the lexical memory arrangement stores the lexical tree data structures as Hidden Markov Models and the lexical tree processors are operative to perform the Viterbi search algorithm using each respective lexical tree data structure. Thus in this way, this embodiment of the present invention provides a parallel Viterbi lexical tree search process for speech recognition.

The first aspect of the present invention is a special purpose circuit built for performing the speech recognition search process in which there are a plurality of processors for performing parallel lexical tree processing on individual lexical tree processors.

In another aspect of the present invention a speech recognition circuit comprises an input port such as an input buffer for receiving parameterized speech data such as feature vectors. A plurality of lexical memories are provided which contain in combination complete lexical data for word recognition. Each lexical memory contains part of the complete lexical data. A plurality of processors are provided connected in parallel to the input port for processing the speech parameters in parallel. The processors are arranged in groups in which each group is connected to a corresponding lexical memory to form a cluster. A controller controls each processor to process the speech parameters using partial lexical data read from a corresponding lexical memory. The results of processing the speech parameters are output from the processors as recognition data.

Thus this aspect of the present invention provides a circuit in which speech recognition processing is performed in parallel by groups of processors operating in parallel in which each group accesses a common memory of lexical data. This aspect of the present invention provides the advantage of parallel processing of speech parameters and benefits from a limited segmentation of the lexical data. By providing a plurality of processors in a group with a common memory, flexibility in the processing is provided without being bandwidth limited by the interface to the memory that would occur if only a single memory were used for all processors. The arrangement is more flexible than the parallel processing arrangement in which each processor only has access to its own local memory and requires fewer memory interfaces (i.e. chip pins). Each processor within a group can access the same lexical data as any other processor in the group. The controller can thus control the parallel processing of input speech parameters in a more flexible manner. For example, it allows more than one processor to process input speech parameters using the same lexical data in a lexical memory. This is because the lexical data is segmented into domains which are accessible by multiple processors.

In a preferred embodiment this aspect of the present invention is used in combination with the first aspect of the present invention. In such an arrangement each processor performs lexical tree processing and the lexical data stored in each lexical memory comprises lexical tree data structures which each comprise a model of words having common prefix components and an initial component that is unique.

In preferred embodiments of the second aspect of the present invention, the preferred embodiments of the first aspect of the present invention are incorporated.

Embodiments of the present invention will now be described with reference to the accompanying drawings in which:

Figure 1 is a diagram of a speech data processing circuit for generating parameterized speech data (feature vectors);

Figure 2 is a diagram of a speech recognition circuit in accordance with an embodiment of the present invention;

Figures 3a and 3b are schematic diagrams illustrating lexical tree structures;

Figure 4 is a flow diagram illustrating the process performed by a lexical tree processor to determine a temporary lexical tree score in accordance with an embodiment of the present invention;

Figure 5 is a flow diagram illustrating the process performed by the lexical tree processor for processing the input feature vectors in accordance with an embodiment of the present invention; and

Figure 6 is a flow diagram illustrating the process performed by the controller in accordance with an embodiment of the present invention.

Figure 1 illustrates a typical circuit for the parameterization of input speech data. In this embodiment the parameters generated are speech vectors.

A microphone 1 records speech in an analogue form and this is input through an anti-aliasing filter 2 to an analogue-to-digital converter 3 which samples the speech at 48 kHz at 20 bits per sample. The digitized output signal is normalized (4) to generated a 10 millisecond data frame every 5 milliseconds with 5 milliseconds overlap (5). A pre-emphasis operation 6 is applied to the data followed by a hamming window 7. The data is then fast Fourier transformed (FFT) using a 512 point fast Fourier transform (8) before being filtered by filter bank 9 into 12 frequencies. The energy in the data frame 5 is also recorded (13) as an additional feature and together with the 12 frequency outputs of the filter bank 9, 13 feature vectors (10) are thus produced and these are output as part of the 39 feature vectors 14. First and second derivatives (11 and 12) are taken of the 13 feature vectors 10 to complete the generation of the 39 feature vectors 14.

The arrangement illustrated in Figure 1 is purely given for illustration. The present invention encompasses any means by which speech and data can be parameterized to a suitable form for input to the search process as will be described in more detail hereinafter.

Figure 2 is a schematic diagram of a speech recognition circuit in accordance with an embodiment of the present invention for performing the search process. The parameterized speech data, which in this embodiment comprise feature vectors are input to a feature vector buffer 20. The feature vector buffer 20 is provided to buffer the incoming feature vectors to allow lexical tree processors 21 to read and process the feature vectors in the buffer 20 via a feature vector bus 24. A plurality k of lexical tree processors 21 are arranged in a respective lexical tree processor cluster 22. Each lexical tree processor cluster 22 has an acoustic model memory 23 in which is stored lexical data for use by the lexical tree processors 21 within the lexical tree processor cluster 22. Each lexical tree processor 21 in the lexical tree processor cluster 22 are connected to the acoustic model memory 23 within the lexical tree processor 22. There are N lexical tree processor clusters and thus there are Nk lexical tree processors 21 connected by the

feature vector bus 24 to the feature vector buffer 20. Each lexical tree processor 21 is capable of processing a different lexical tree and thus Nk lexical trees can be processed in parallel. The acoustic model memories 23 store as a whole a complete set of lexical data, i.e. lexical tree data structures for use in the lexical tree processing by the lexical tree processors 21. Each acoustic model memory 23 contains part or a segment of the lexical tree data. Since lexical tree processors 21 in a lexical tree processor cluster 22 access the same acoustic model memory 23, it is possible for more than one lexical tree processor 21 to process the same lexical data. This provides for some degree of flexibility in the controlling of the processing by the lexical tree processors 21. Further, the acoustic model memories 23 need not contain only one copy of the lexical data. It is possible to build in a redundancy in the data to further enhance the flexibility. This avoids any bottleneck in the processing due to the search processing focusing on a small number of lexical trees.

A results memory 25 is provided for storing processing results from the lexical tree processors 21 which are received over the path score and history bus 26. The results memory 25 also stores information on lexical trees to identify which lexical trees are to be processed. A search controller 27 is provided to control the processing performed by the lexical tree processors 21 in dependence upon a program and data stored in program and data memory 28. The search controller reads the path scores and lexical tree identifiers from the results memory and controls the lexical tree processors accordingly. A language model processor 29 is provided which is connected to each lexical tree processor 21 by a language model bus 30. The language model processor 29 accesses a language model memory 31 to read language model data for provision to lexical tree processors 21 in response to language model data requests. External control of the language model memory 31 is provided by a word constrains input. The language model processor 29 determines a score for a word occurring following N previous words using N grams. When a lexical tree processor requires a language model score a request is sent to the language model processor 29 over the language model bus 30 identifying the current word and the N-1 previous words. A language model score for the N gram can be returned to the lexical tree processor 21 for the modification of the score at the end of a branch of lexical tree processing. The lexical tree processor can modify the score in accordance with the language model and output a score to the

results memory 25 for a word at the end of a branch of the lexical tree processing. Thus the results memory stores the results as an ordered list of scores for words together with their histories.

The results memory 25 stores the following data:

1. Initial lexical tree data. This comprises pointers to an initial set of lexical trees. No history data is associated with the initial set of lexical trees. The initial set of lexical trees is predetermined and stored in the results memory 25 based on the most likely initial phones of an utterance. This initial lexical tree data is required to initialize the search process.

2. History data for search results. This comprises a record of a recognition path through the lexical tree recognition process performed by the lexical tree processors 21. The history data includes the current word, the previous N-1 words, the current accumulated score, the phone history (for use in the determination of likely next lexical trees using cross word context dependent tri-phones), and an identifier or pointer to the lexical tree used for identifying the word.

3. Best scores for best paths being processed by each lexical tree processor 21. This information enables the search controller 27 to monitor the processing being performed by lexical tree processors 21 to determine whether a global pruning strategy should be applied in order to reassign processing performed by a lexical tree processor if its best score for its best path is below a threshold or well below the best scores for the paths being processed by other lexical tree processors 21.

4. Temporary lexical tree scores. These comprise tree scores which are determined as temporary scores to prune the next lexical trees to be processed at word ends. The temporary lexical tree scores include lexical tree identifiers or pointers to identify the next lexical trees to be processed. The scores enable the pruning of this list.

5. Pruning threshold. This can be a global threshold value for use in the pruning of the lexical trees globally, or a local threshold value for use by a lexical processor for locally pruning the processing performed by the lexical processor 21.

The acoustic model memory 23 stores a Hidden Markov Model for acoustically modelling words as lexical trees. The acoustic model memory 23 stores a plurality of lexical tree data structures. Each lexical tree data structure comprises an n phone model of a number of words having common prefix phones. The first node of the lexical tree (the root) comprises a common n phone to all words in the lexical tree and uniquely identifies the lexical tree.

Each lexical tree processor 21 includes on-board memory (or local memory) to be used during the lexical tree processing. This working memory has to store all of the parameters currently working on including current scores for all paths being processed within the lexical tree, and previous N-1 words for the lexical tree. The local storage of the previous N-1 words enables the lexical tree processor 21, when a word end is reached along a branch of the lexical tree, to send a request for the language model score for an N gram by sending the identity of the N-1 previous words together with the word identified at the end of the branch.

Figure 3a and 3b schematically illustrate lexical trees which can be processed by the lexical tree processor 21 during the recognition of the two words HARD ROCK. In Figure 3a a previous lexical tree terminated at a branch recognizing the word YOU and terminating with the mono phone uw which is associated with two context dependent tri-phones y-uw+k and y-uw+h. Thus the context dependent tri-phone associated with the last phone in the lexical tree word model for YOU indicates two possible next lexical trees, i.e. the lexical trees beginning with the mono phone k and h and having a left context phone of uw. As can be see in Figure 3a this word end YOU therefore leads to two possible next lexical trees. These two lexical trees are traversed in parallel by two different lexical tree processors accumulating scores for matching of input feature vectors into Hidden Markov Models of the context dependent tri-phones associated with each node in the tree. When the end of the branch is reached, a word end event is reached and a word is recognized. As can be seen in Figure 3a, in this example since

four words are of similar phone length, it is possible for the search strategy based on inputting feature vectors in parallel to simultaneously reach a number of possible word ends. These possible word ends are sent as processing results to the results memory. The results memory stores the accumulated score at the word end together with phone history to identify the last phone and its associated context dependent tri-phones. In this example, considering the branch where the word HARD is recognized, the last phone is d which has associated with it two context dependent tri-phones r-d+l and r-d+r. Thus the search controller 27 can identify next possible lexical trees using the next phone in the context dependent tri-phone. In this case, as can be seen in Figure 3b, the next possible lexical trees begin with the phones r and l and have associated with them context dependent tri-phones d-r+ao and d-l+ao respectively. Thus these are the next lexical trees that require processing following the end node or word event detected in the processing lexical tree 2. Figure 3b thus represents the processing that is required at the end of Figure 3a at the word node for HARD in processing the second lexical tree.

As can be seen in Figure 3b, the two lexical trees are processed in parallel by traversing through the branches to reach word ends by sequentially entering in feature vectors in parallel to lexical tree processors processing the respective lexical trees. When the word end is reached, the accumulated score is output to the results memory as described before. In this case since it is the last word, the important context dependent tri-phone associated with the final phone has silence (sil) as the following phone. The final phone can in fact have 50 context dependent tri-phones associated with it if there are 50 possible next phones (including silence). Figure 3b only illustrates the relevant one for the end of the utterance.

Figure 4 is a flow diagram illustrating the processing performed by a lexical tree processor 21 in order to determine a temporary lexical tree score for a lexical tree. When a word end is identified by a lexical tree processor, the processing results are sent to the results memory 25. The results memory identifies the last phone of the recognized words and thus enables the search controller 27 to identify possible next lexical trees using context dependent tri-phones as illustrated in Figure 3b. Although in theory if there are 50 phones there are 50^2 possible lexical trees, due to the use of the context dependant triphones, only 100-200 lexical trees are likely to be identified as

possible lexical trees. The search controller needs to further prune out the lexical trees which are unlikely to generate likely paths with high scores. In order to do this, instructions are sent to the lexical tree processors to determine a temporary lexical tree score which can be used to prune out lexical trees from the processing which have a low score.

The processing by the lexical tree processor to generate the temporary lexical tree score is shown in Figure 4 and for this process, following initialization, a lexical tree processor 21 awaits receipt of data for a new lexical tree from the search controller 27 (step S2). The data comprises a lexical tree pointer on an identifier to lexical tree data stored in the acoustic model memory 23, i.e. the lexical tree data structure, previous N-1 words (for use in language model score determination by the language model processor 29), and the current path score. The previous N-1 words includes the previously recognized word for which a score has been accumulated.

The lexical tree processor 21 then attempts to read the next feature vector from the feature vector buffer 20 (step S3) and if this is not available, an error occurs (step S4). If the feature vector is available in the feature vector buffer 20, the lexical tree processor 21 reads the feature vector from the buffer 20 (step S5) and evaluates the state transitions for the first lexical tree node using the acoustic model data in the acoustic model memory 23 (step S6). Since the state transitions for the first node will require several feature vectors to complete, a score is determined for a state transition of the first node in the lexical tree (step S7). The lexical tree processor 21 then sends a request to the language model processor 29 for a language model score. The request includes the previous N-1 words and all of the possible words represented by the lexical tree data structure in the acoustic model memory 23 (step S8). The language model processor returns scores for each of the N grams ending in the words represented by the lexical tree. The lexical tree processor thus receives language model scores for the words in the lexical tree and picks the highest score (step S9). Alternatively, the language model processor 29 can select the highest n gram score and return this to the lexical tree processor for the determination of the temporary lexical tree score. The temporary lexical tree score for the lexical tree is then generated using the score determined for the first state transition of the first lexical tree node and the highest language model score

(step S10). The temporary lexical tree score is then sent by the lexical tree processor 21 to the results memory 25 over the bus 26 (step S11). In this mode of processing, the processor then awaits the next data of the new lexical tree from the search controller (step S2).

This process is just one of the processes performed by the lexical tree processor. The main processing performed by the lexical tree processor is the lexical tree evaluation. This process will now be described with reference to the flow diagram of Figure 5.

After the start of the process (step S20) a lexical tree processor awaits receipt of data for a lexical tree from the search controller (step S21). The data comprises a lexical tree pointer to a lexical tree data structure stored in the acoustic model memory 23, the previous N-1 words in the recognition path, and the current accumulated score.

The lexical tree processor then attempts to read the next feature vector from the feature vector buffer 20 (step S22). If this is not available, an error occurs (step s23). If the feature vector is available in the feature vector buffer 20, the feature vector is read from the feature vector buffer 20 by the lexical tree processor 21 (step S24) and the lexical tree processor 21 evaluates the state transitions for each path using the acoustic model data in the acoustic model memory 23 (step S25). Scores for the paths are determined in accordance with the conventional Viterbi search technique and the best score for the best path is sent to the results memory 25 while the path histories are stored locally in the on-board memory of the lexical tree processor 21 (step S26). Pruning is applied to the lexical tree by the lexical tree processor to delete paths in order to keep the breadth of the search manageable (step S27). The pruning applied locally by the lexical tree processor can be purely on the basis of a local threshold, which can be provided by the search controller 27, or it can be determined on a relative basis dependent upon the range of scores for paths being processed within the lexical tree processor 21. If the path has not reached a word end (step S28) the lexical tree processor 21 attempts to read the next feature vector from the feature vector buffer 20 (step S22). If the path reaches a word end (step S28) the score must be modified by the language model score (step S29). There are two ways in which this can be done, in this embodiment the lexical tree processor 21 sends a request to the language model processor 29 for a language model

score. The request includes the current word and the N-1 previous words. The language model processor thus returns the language model score and the language model score is used to modify the current accumulated score at the word end. Alternatively, the lexical tree processor can send the language model processor 29 not just the current word and the N-1 previous words, but also the current accumulated score. The language model processor then determines the language model score using the language model memory 31 and modifies the accumulated score using the language model score. The language model processor 29 can then return the modified score to the lexical tree processor 21 which can then pass it to the results memory 25, or a connection between the language model processor 29 and the results memory 25 (not shown) can enable the language model processor 29 to send the score directly to the results memory 25 for the lexical tree processor 21. In this latter case, the language model processor 29 must also receive a lexical tree pointer to identify the lexical tree for which the score applies.

Assuming in this embodiment that the lexical tree processor calculates the modified score (step S29), the score and history is then sent by the lexical tree processor 21 to the results memory 25 (step S30). The history data sent to the results memory 25 comprises the lexical tree pointer identifying the lexical tree, the modified score at the word end, the phone history identifying at least the last phone to allow for context dependent tri-phone determination of next lexical trees, and the identity of the word identified in the evaluation process.

The lexical tree processor then deletes the path and history data in its on-board memory (step S31) and determines if there are any paths still left to be processed (step S32). If so, the lexical tree processor 21 tries to access the next feature vector available in the feature vector buffer 20 (step S22). If there are no paths left to be processed, a message is sent by the lexical tree processor to the search controller 27 to indicate that the lexical tree has been processed (step S33). In this level of processing the lexical tree processor will then await the next data for a lexical tree from the search controller 27.

The flow diagram of Figure 5 illustrates the processing of a single instance of a lexical tree. The single point in time a lexical tree can start to be processed from a feature

vector, i.e. lexical tree processing starts at a time frame. However, due to a word end returned from lexical tree processing in another lexical tree processor, the lexical tree processor may also be asked during a subsequent time frame to start processing another instance of the lexical tree with a different path history. Thus a lexical tree processor 21 can be processing multiple instances of the lexical tree with different paths having different path histories. In this instance, the pruning applied in step S27 needs to take into account all paths currently being processed in all instances of lexical tree processing by the lexical tree processor 21. Thus although pruning is applied locally, it is applied over more than one instance of the lexical tree.

The operation of the search controller 27 will now be described with reference to the flow diagram of Figure 6.

After initialization (step S40) initial lexical tree data in the results memory 25 is read (step S41). The search controller distributes the initial lexical tree data amongst the lexical tree processors for temporary lexical tree score determination (step S42) as described with reference to Figure 4. Temporary lexical tree scores are returned by the lexical tree processors 21 to the results memory 25 (step S43) and the search controller reads the temporary lexical tree scores in the results memory 25 and prunes the lexical trees in the results memory 25 on the basis of the read temporary lexical scores (step S44). Thus the set of initial lexical trees can be quite large and this requires a large number of temporary lexical tree score determinations to be performed by the lexical tree processors 21. However, since the lexical tree processors only need to determine a score for a single feature vector, this can be performed quite quickly. Thus the number of lexical trees for which full evaluation by the lexical tree processors 21 is required can thus be reduced by the selection of the lexical trees with the highest temporary lexical tree score. The search controller than distributes the lexical tree processing among the lexical tree processors 21 by sending each lexical tree processor data including a lexical tree pointer to a lexical tree data structure in the acoustic model memory 23 in the cluster 22. The search controller 27 then awaits the entry of processing results, i.e. history and scores in the results memory 25 for words (step S46). When history and scores for words are entered in the results memory, the search controller 27 determines next possible lexical trees using cross-word tri-phones (step S47). For example, as

illustrated in Figures 3a and 3b, when the word HARD is entered in the results memory, the search controller identifies the final phone of the word as being d and thus next possible tri-phones can comprise any tri-phones having the left phone d. If there are 50 possible phones, this results in 50^2 possible tri-phones. In practice, only 1 to 200 are likely. However, this may still represent too many for full evaluation and thus the lexical tree data for the next possible lexical trees are distributed amongst the lexical tree processors for temporary lexical tree score determination (step S48). Temporary lexical tree scores are returned to the results memory 25 by the lexical tree processors 21 which execute the process as described with reference to Figure 4 (step S49).

The search controller 27 prunes the list of next lexical trees in the results memory 25 on the basis of the temporary lexical scores (step S50). The search controller then determines if the words in the results memory have been processed and there are no lexical trees left to be processed and whether or not lexical tree processes are finished (step S51). If not, the lexical tree processing is distributed amongst the lexical tree processors (step S45). When the end of the utterance is reached, next possible lexical trees will all have been processed and the results can be output from the results memory 25 (step S52).

It can be seen from the foregoing description that the embodiment of the invention provides for an efficient parallel processing system incorporating a parallel lexical tree processing architecture and a segmented lexicon storage architecture. Each lexical tree processor performs lexical tree evaluation and language model look ahead for evaluation of company scores for lexical trees to assist in pruning of the number of potential lexical trees that require processing. The architecture is ideally suited to synchronous breadth first searching e.g. a Viterbi search (i.e. real time recognition). The breadth first search is performed synchronously by the parallel processors on the input search parameters when the language model look ahead takes place, it is possible for the feature vector buffer 20 to store a small number, e.g. one or two, feature vectors whilst the language model look ahead is performed on the feature vector so that synchronous breadth first searching can be carried out by re-reading the feature vector out of the buffer 20 when it is determined which lexical trees are to be evaluated by the

lexical trees processors 21 after a word end. Thus, the feature vector buffer 20 provides a method of decoupling the search parameterisation circuitry from the search circuit.

If the buffer causes processing to get too far behind, e.g. there are a lot of word end language model look ahead operations to be performed, it is possible for the lexical tree processors to be controlled to temporarily halve their local beamwidths (i.e. more vigorously prune) and thus complete their processing more quickly. This would allow the content of the buffer to be decreased as the processors catch up.

Although the present invention has been described hereinabove with reference to a specific embodiment, it will be apparent to a skilled person in the art that modifications lie within the spirit and scope of the present invention. Although in the embodiment described hereinabove, the acoustic model memories 23 were described as storing parts of the lexical data, it is also possible for more than one copy of the lexical data to be stored in different acoustic model memories 23. Thus, more than one lexical tree processor cluster 22 can have access to the same lexical data. This increases the flexibility and allows the controller to avoid any bottlenecks in processing which, for example, can occur if the highest pass scores all require processing by one or a number of lexical tree data structures residing in one of the acoustic model memories 23.

Although in the embodiments described hereinabove, the memory is partitioned in to a plurality of acoustic model memories 23, the present invention is not limited to the partitioning of the lexical data and a plurality of lexical tree processors can access a single memory containing lexical data.

The lexical data stored can comprise mono phone models of words. When the lexical data is partitioned, these can be distributed evenly across the memories. If there are 50 mono phones, there will be 50 lexical tree data structures having a unique initial model phone data node. Alternatively, the lexical tree data structures can comprise a context dependent n phone word models. If there are 50 mono phones, there are potentially 50^n lexical tree data structures. For tri-phones, this comprises 50^3 (i.e. 125,000). These are distributed so that trees with the same left context tri-phone, are eventually spread across memories when the lexical data is partitioned. Thus in this way the word end

where the left context tri-phone is fixed i.e. the same, the next possible lexical trees will inherently be distributed across the memories, thus ensuring efficient parallel processing of the next lexical trees.

Although the embodiment described hereinabove uses a single results memory accessible by all processors, the present invention encompasses a plurality of results memories. For example, a results memory can be associated with each processor to store the processing results and the controller can read the results in the plurality of results memories in order to control the processors.

Although in the embodiment described above the input speech parameters are described as being feature vectors, the present invention encompasses any type of parameterization of speech data as an input to the circuit. Such parameterization methods are well known in the art.

Although in the embodiment described above there is only one language model processor, more than one can be used and these can be used to operate in parallel for processing scores at word boundaries.

CLAIMS:

1. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

lexical memory means containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input means for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory means;

results memory means connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

control processor means for controlling said lexical tree processors to process lexical trees identified in said results memory means by performing parallel processing on a plurality of said lexical tree data structures.

2. A speech recognition circuit according to claim 1, wherein said lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

3. A speech recognition circuit according to claim 2, wherein said lexical memory means stores each lexical tree data structure as mono phone models of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

4. A speech recognition circuit according to any preceding claim, wherein each lexical tree processor is operative to compare the processed speech parameters with said lexical data or data derived from said lexical data to identify words as a word recognition event and to send information identifying the identified words to said results memory means as the processing results.

5. A speech recognition circuit according to claim 4, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

6. A speech recognition circuit according to claim 5, wherein each lexical tree processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and to output the modified processing result to said results memory means.

7. A speech recognition circuit according to any one of claims 4 to 6, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events.

8. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from said lexical data in said lexical memory means.

9. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to instruct a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory means.

10. A speech recognition circuit according to claim 9, wherein said history data includes an accumulated score for the recognition path.

11. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to pass a lexical tree identifier to a said lexical tree processor, said lexical tree processor is operative to read lexical data for the identified lexical tree from said lexical memory means and to process said speech parameters using said read lexical data.

12. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

13. A speech recognition circuit according to claim 12, wherein said lexical tree processors are operative to prune the processing of the lexical tree in dependence upon the determined scores.

14. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process more than one lexical tree at the same time.

15. A speech recognition circuit according to any one of claims 4 to 7, wherein said control processor means is operative to determine new lexical tree identifiers for storing in said results memory means for words identified in said results memory means.

16. A speech recognition circuit according to claim 15, wherein said control processor means is operative to prune the new lexical tree identifiers.

17. A speech recognition circuit according to claims 16, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

18. A speech recognition circuit according to any preceding claim, wherein said input means is arranged to receive said speech parameters as feature vectors.

19. A speech recognition circuit according to any one of claims 1 to 17, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

20. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are arranged in groups, said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said groups of lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

21. A speech recognition circuit according to any one of claims 1 to 19, wherein said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

22. A speech recognition circuit according to any preceding claim, wherein said lexical memory means stores said lexical tree data structures as Hidden Markov Models, and said lexical tree processors are operative to perform the Viterbi search algorithm using each lexical tree data structure.

23. A speech recognition circuit comprising:

an input buffer for receiving processed speech parameters;

lexical memory containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, the initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory;

results memory connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

a control processor for controlling said lexical tree processors to process lexical trees identified in said results memory by performing parallel processing on a plurality of said lexical tree data structures.

24. A speech recognition circuit according to claim 23, wherein said lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

25. A speech recognition circuit according to claim 24, wherein said lexical memory stores said lexical tree data structure as a mono phone model of words, and each lexical tree processor generates context dependant phone models of words dynamically for use in processing the speech parameters.

26. A speech recognition circuit according to any one of claims 23 to 25, wherein each lexical tree processor compares the processed speech parameters with the lexical data or data derived from the lexical data to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

27. A speech recognition circuit according to claim 26, including a language model processor for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

28. A speech recognition circuit according to claim 27, wherein each lexical tree processor receives said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and outputs the modified processing result to said results memory.

29. A speech recognition circuit according to any one of claims 26 to 28, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events.

30. A speech recognition circuit according to any one of claims 23 to 29, wherein said lexical tree processors process said speech parameters by comparing said speech parameters with said lexical data in said lexical memory.

31. A speech recognition circuit according to any one of claims 23 to 30, wherein said control processor instructs a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory.

32. A speech recognition circuit according to claim 31, wherein said history data includes an accumulated score for the recognition path.

33. A speech recognition circuit according to any one of claims 23 to 32, wherein said control processor passes a lexical tree identifier to a said lexical tree processor, said lexical tree processor reads lexical data for the identified lexical tree from said lexical memory and processes said speech parameters using said read lexical data.

34. A speech recognition circuit according to any one of claims 23 to 33, wherein said lexical tree processors determine and output scores in the processing results during the processing of said speech parameters.

35. A speech recognition circuit according to claim 34, wherein said lexical tree processors prune the processing of the lexical trees in dependence upon the determined scores.

36. A speech recognition circuit according to any one of claims 23 to 35, wherein said lexical tree processors process more than one lexical tree at the same time.

37. A speech recognition circuit according to any one of claims 26 to 29, wherein said control processor determines new lexical tree identifiers for storing in said results memory for words identified in said results memory.

38. A speech recognition circuit according to claim 37, wherein said control processor means is operative to prune the new lexical tree identifiers

39. A speech recognition circuit according to claims 38, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

40. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as feature vectors.

41. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as n phones, where n is an integer.

42. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical tree processors are arranged in groups, said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

43. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory is connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

44. A speech recognition circuit according to any one of claims 23 to 43, wherein said lexical memory stores said lexical tree data structures as Hidden Markov Models, and each lexical tree processor performs the Viterbi search algorithm using a respective said lexical tree data structure.

45. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

a plurality of lexical memory means containing in combination complete lexical data for word recognition, each lexical memory means containing part of said complete lexical data;

a plurality of processors connected in parallel to said input means for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory means;

control processor means for controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

receiving means for receiving the results of the processing of the speech parameters from said processors.

46. A speech recognition circuit according to claim 45, wherein each lexical memory means stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

47. A speech recognition circuit according to claim 46, wherein each lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

48. A speech recognition circuit according to claim 47, wherein each lexical memory means stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

49. A speech recognition circuit according to any one of claims 45 to 48, wherein each processor is operative to compare the processed speech parameters with the lexical data in a corresponding lexical memory means to identify words as a word recognition event and to send information identifying the identified words to said receiving means as the processing results.

50. A speech recognition circuit according to claim 49, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said processor.

51. A speech recognition circuit according to claim 50, wherein each processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said processor and to output the modified processing result to said receiving means.

52. A speech recognition circuit according to any one of claims 49 to 51, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events.

53. A speech recognition circuit according to any one of claims 45 to 52, wherein said processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memory means.

54. A speech recognition circuit according to any one of claims 46 to 48, wherein said control processor means is operative to instruct a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

55. A speech recognition circuit according to claim 54, wherein said history data includes an accumulated score for the recognition path.

56. A speech recognition circuit according to any one of claims 46 to 48, 54, or 55, wherein said control processor means is operative to pass a lexical tree identifier to a said processor, said processor is operative to read lexical data for the identified lexical tree from a corresponding said lexical memory means and to process said speech parameters using said read lexical data.

57. A speech recognition circuit according to any one of claims 45 to 56, wherein said processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

58. A speech recognition circuit according to claim 57, wherein said processors are operative to prune the processing of the speech parameters in dependence upon the determined scores.

59. A speech recognition circuit according to any one of claims 46 to 48, or 54 to 56, wherein said processors are operative to use lexical data corresponding to more than one lexical tree at the same time.

60. A speech recognition circuit according to any one of claims 49 to 52, wherein said control processor means is operative to identify lexical data in said lexical memory means required for processing next words at word recognition events, and to control said processors to process the speech parameters using the identified lexical data.

61. A speech recognition circuit according to claim 60, wherein said control processor means is operative to prune the identified lexical data.

62. A speech recognition circuit according to claims 61, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune the identified lexical data on the basis of the scores for the words.

63. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as feature vectors.

64. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

65. A speech recognition circuit according to any one of claims 45 to 64, wherein said control processor means is operative to control each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory means.

66. A speech recognition circuit according to any one of claims 45 to 65, wherein said receiving means comprises a memory for storing the results of processing.

67. A speech recognition circuit according to any one of claims 45 to 66, wherein said lexical memory means store said lexical data as Hidden Markov Models, and each processor is operative to performs the Viterbi search algorithm using a respective part of said lexical data.

68. A speech recognition circuit comprising:

an input buffer receiving processed speech parameters;

a plurality of lexical memories containing in combination complete lexical data for word recognition, each lexical memory containing part of said complete lexical data;

a plurality of processors connected in parallel to said input buffer for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory;

a control processor controlling each processor to process said speech parameters using partial lexical data read from a respective said lexical memory; and

a results memory storing the results of the processing of the speech parameters from said processors.

69. A speech recognition circuit according to claim 68, wherein each lexical memory stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

70. A speech recognition circuit according to claim 69, wherein each lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

71. A speech recognition circuit according to claim 70, wherein each lexical memory stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

72. A speech recognition circuit according to any one of claims 68 to 71, wherein each processor compares the processed speech parameters with the lexical data in a corresponding lexical memory to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

73. A speech recognition circuit according to claim 72, a including language model processor providing a language model output for modifying the processing results at a word recognition event by a said processor.

74. A speech recognition circuit according to claim 73, wherein each processor receives said language model output to modify the processing result at a word recognition event determined by said processor and outputs the modified processing result to said results memory.

75. A speech recognition circuit according to any one of claims 73 to 74, wherein each processor determines and outputs scores for words in the processing results at word recognition events.

76. A speech recognition circuit according to any one of claims 68 to 75, wherein said processors process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memories.

77. A speech recognition circuit according to any one of claims 69 to 71, wherein said control processor instructs a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

78. A speech recognition circuit according to claim 77, wherein said history data includes an accumulated score for the recognition path.

79. A speech recognition circuit according to any one of claims 69 to 71, 77, or 78, wherein said control processor passes a lexical tree identifier to a said processor, said processor reads lexical data for the identified lexical tree from a corresponding said lexical memory and processes said speech parameters using said read lexical data.

80. A speech recognition circuit according to any one of claims 68 to 79, wherein said processors determine and output scores in the processing results during the processing of said speech parameters.

81. A speech recognition circuit according to claim 80, wherein said processors prune the processing of the speech parameters in dependence upon the determined scores.

82. A speech recognition circuit according to any one of claims 69 to 71, or 77 to 79, wherein said processors use lexical data corresponding to more than one lexical tree at the same time.

83. A speech recognition circuit according to any one of claims 72 to 75, wherein said control processor identifies lexical data in said lexical memories required for processing next words at word recognition events, and controls said processors to process the speech parameters using the identified lexical data.

84. A speech recognition circuit according to claim 83, wherein said control processor prunes the identified lexical data.

85. A speech recognition circuit according to claims 84, wherein each processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes the identified lexical data on the basis of the scores for the words.

86. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as feature vectors.

87. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as n phones, where n is an integer.

88. A speech recognition circuit according to any one of claims 68 to 88, wherein said control processor controls each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory.

89. A speech recognition circuit according to any one of claims 68 to 89, wherein said lexical memories store said lexical data as Hidden Markov Models, and each processor performs the Viterbi search algorithm using a respective part of said lexical data.

90. A speech recognition method comprising:

receiving processed speech parameters;

storing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;;

passing the speech parameters in parallel to a plurality of lexical tree processors to perform parallel lexical tree processing for word recognition by accessing said lexical data;

storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

controlling said lexical tree processors to process lexical trees identified by the stored identifiers by performing parallel processing on a plurality of said lexical tree data structures.

91. A speech recognition method comprising:

receiving processed speech parameters;

storing complete lexical data for word recognition in a plurality of lexical memory means, each lexical memory means containing part of said complete lexical data;

processing the speech parameters in parallel using a plurality of processors connected in parallel, said processors being arranged in groups of processors, each group of processors being connected to a said lexical memory means;

controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

storing the results of the processing of the speech parameters from said processors.

Rec'd PCT/PTO 04 AUG 2004

503463

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2003 (14.08.2003)

PCT

(10) International Publication Number
WO 03/067572 A2

(51) International Patent Classification[7]: G10L 15/00

(21) International Application Number: PCT/GB03/00459

(22) International Filing Date: 4 February 2003 (04.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0202546.8 4 February 2002 (04.02.2002) GB

(71) Applicant *(for all designated States except US)*: **ZENTIAN LIMITED** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **CATCHPOLE, Mark** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(74) Agent: **COLLINS, John, David**; Marks & Clerk, 57-60 Lincolns Inn Fields, London WC2A 3LS (GB).

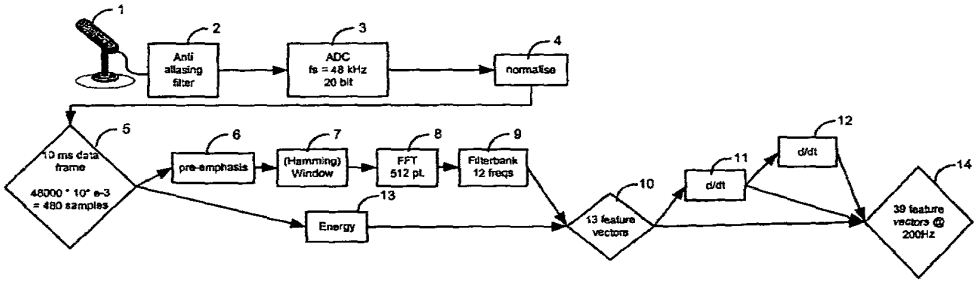
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS



(57) **Abstract:** A speech recognition circuit comprises an input buffer for receiving processed speech parameters. A lexical memory contains lexical data for word recognition. The lexical data comprises a plurality of lexical tree data structures. Each lexical tree data structure comprises a model of words having common prefix components. An initial component of each lexical tree structure is unique. A plurality of lexical tree processors are connected in parallel to the input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory. A results memory is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory by performing parallel processing on a plurality of said lexical tree data structures.

WO 03/067572 A2

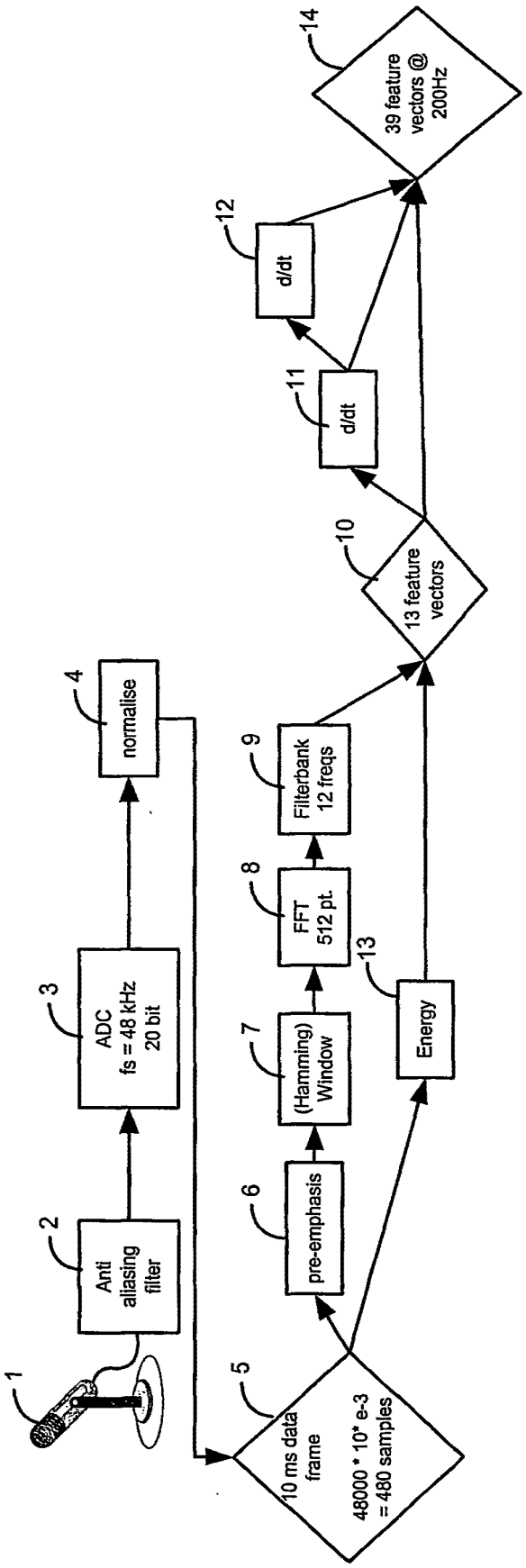

1
2
Anti aliasing filter
3
ADC fs = 48 kHz 20 bit
4
normalise
5
10 ms data frame 48000 * 10* e-3 = 480 samples
6
pre-emphasis
7
(Hamming) Window
8
FFT 512 pt.
9
Filterbank 12 freqs
10
13 feature vectors
11
d/dt
12
d/dt
13
Energy
14
39 feature vectors @ 200Hz

Fig 1

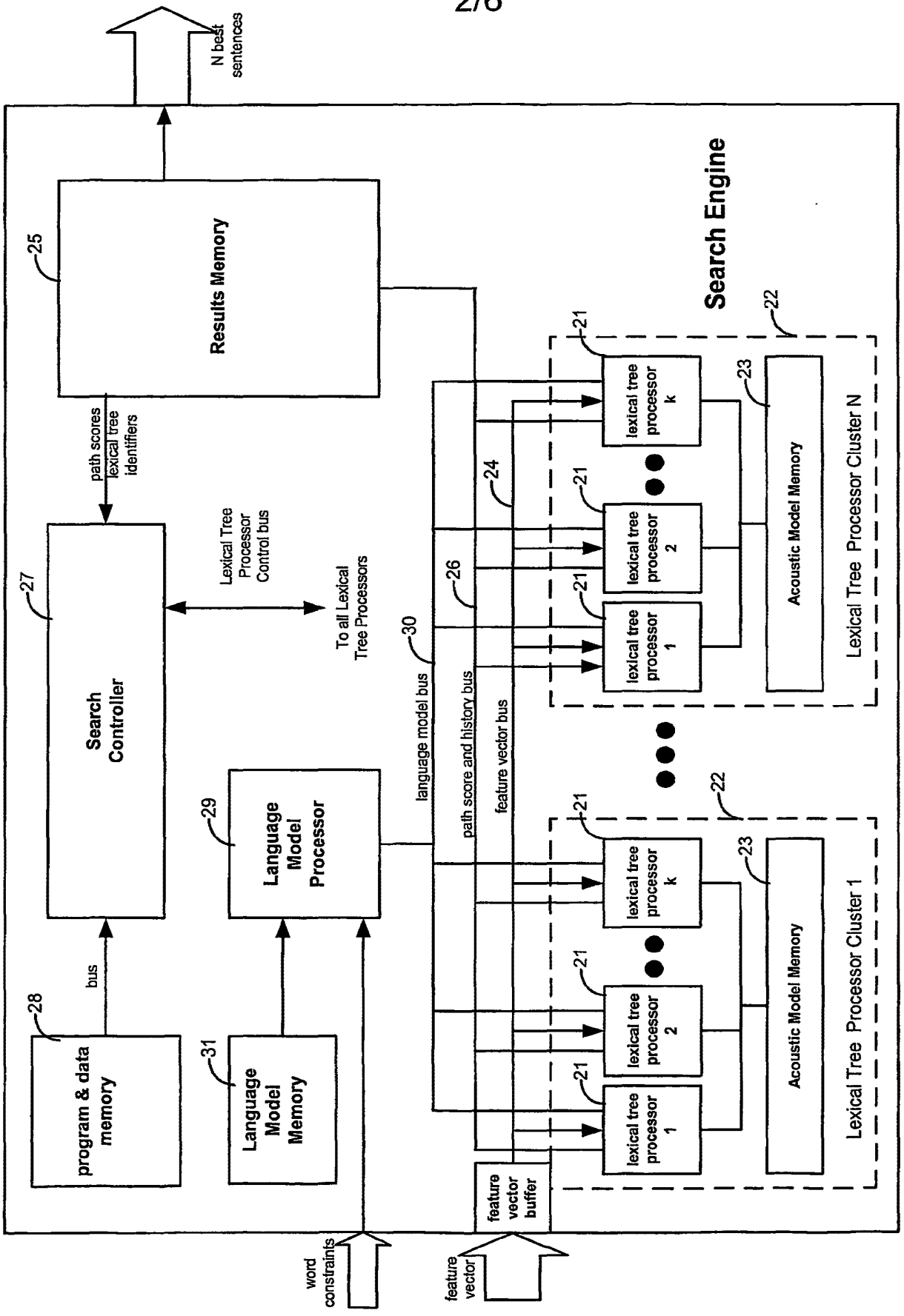

N best sentences
Search Engine
Results Memory
25
path scores
lexical tree identifiers
Search Controller
27
Lexical Tree Processor Control bus
To all Lexical Tree Processors
Language Model Processor
29
language model bus
path score and history bus
feature vector bus
30
24
26
bus
program & data memory
28
Language Model Memory
31
feature vector buffer
word constraints
feature vector
lexical tree processor k
lexical tree processor 2
lexical tree processor 1
21
22
Acoustic Model Memory
23
Lexical Tree Processor Cluster N
Lexical Tree Processor Cluster 1

Fig 2

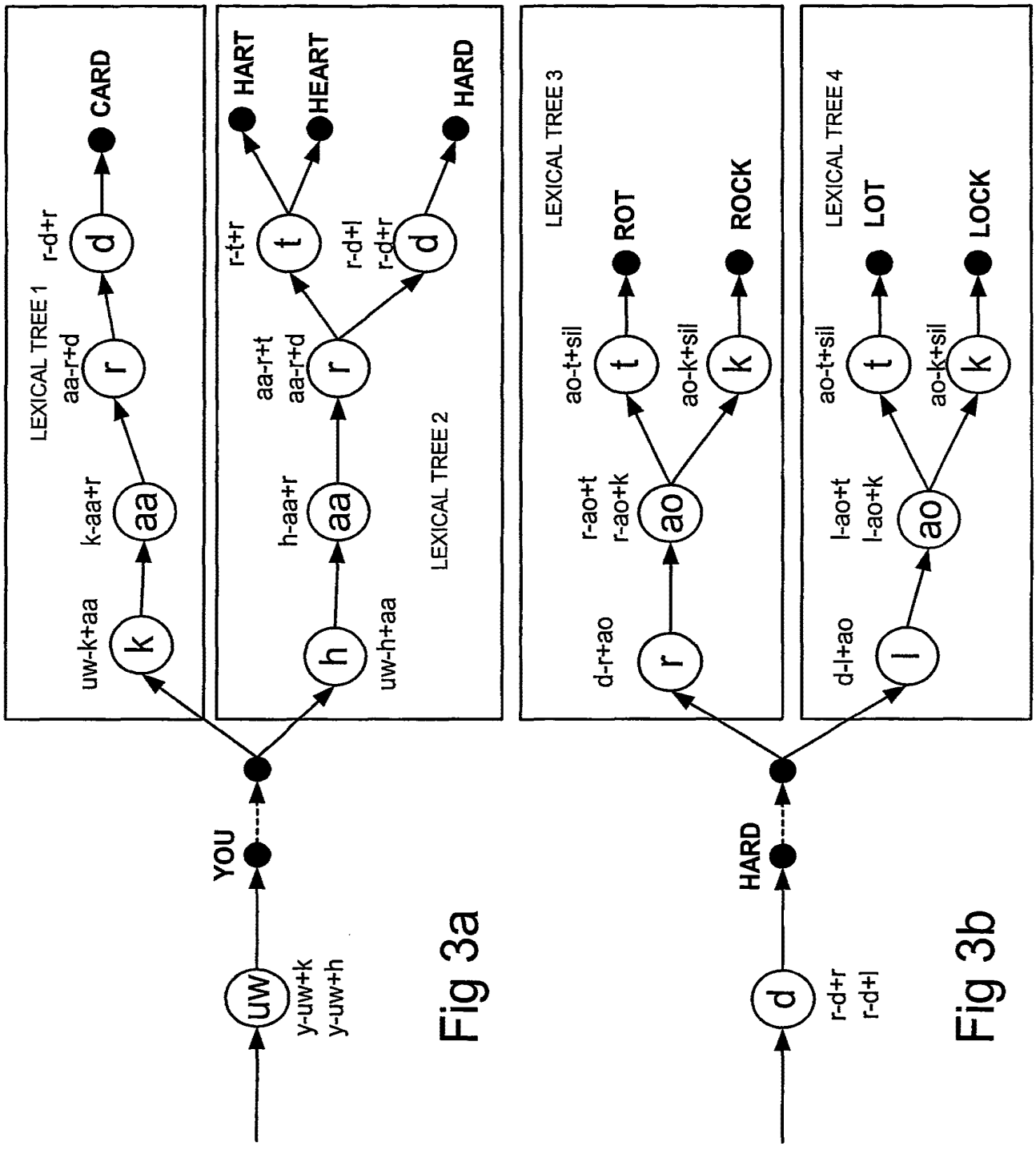

LEXICAL TREE 1
uw-k+aa
k-aa+r
aa-r+d
r-d+r
k
aa
r
d
CARD
LEXICAL TREE 2
uw-h+aa
h-aa+r
aa-r+d
r-t+r
r-d+l
r-d+r
h
aa
r
t
d
HART
HEART
HARD
YOU
uw
y-uw+k
y-uw+h
LEXICAL TREE 3
d-r+ao
r-ao+t
r-ao+k
ao-t+sil
ao-k+sil
r
ao
t
k
ROT
ROCK
LEXICAL TREE 4
d-l+ao
l-ao+t
l-ao+k
ao-t+sil
ao-k+sil
l
ao
t
k
LOT
LOCK
HARD
d
r-d+r
r-d+l

Fig 3a

Fig 3b

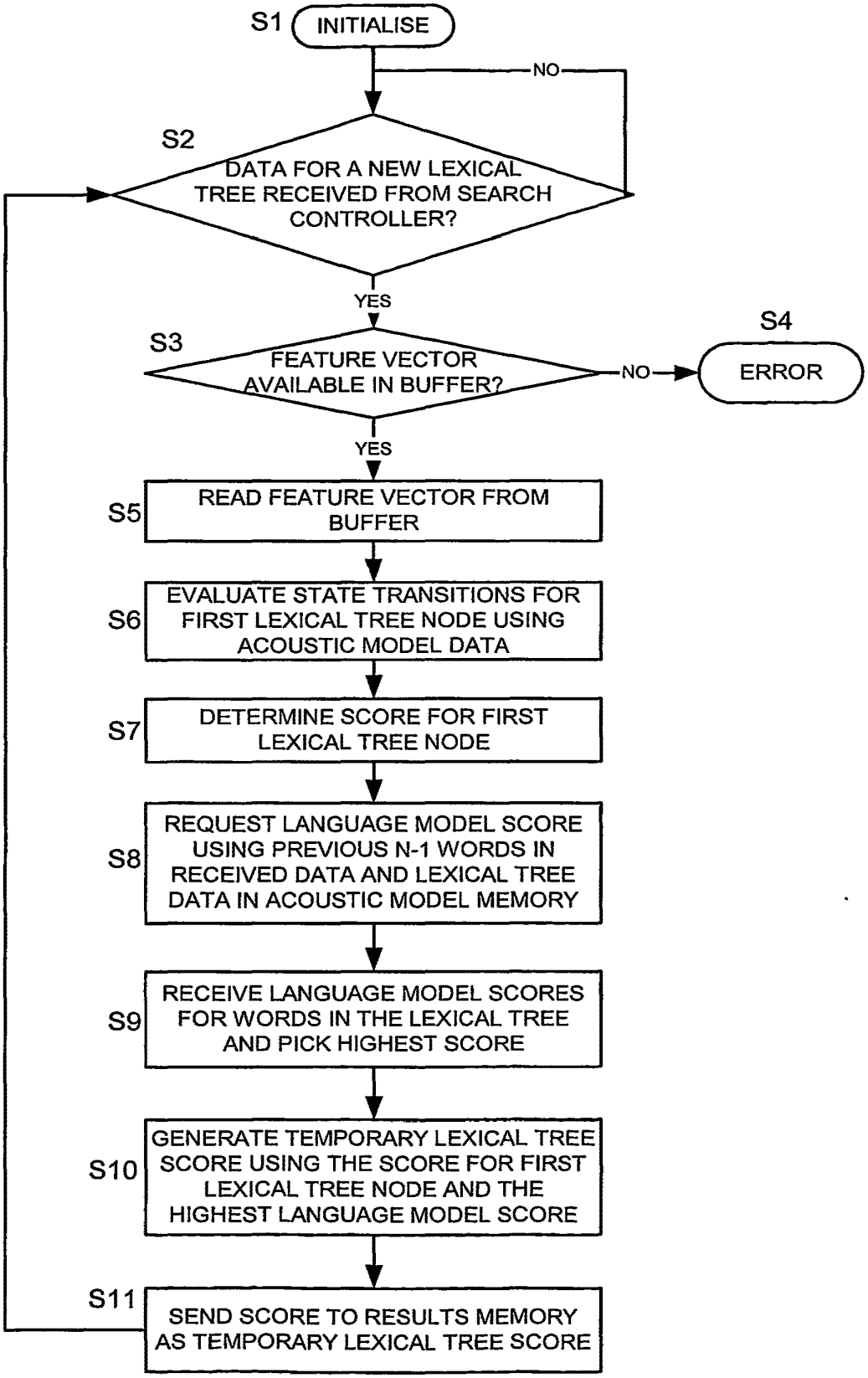

S1
INITIALISE
NO
S2
DATA FOR A NEW LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S3
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S4
ERROR
YES
S5
READ FEATURE VECTOR FROM BUFFER
S6
EVALUATE STATE TRANSITIONS FOR FIRST LEXICAL TREE NODE USING ACOUSTIC MODEL DATA
S7
DETERMINE SCORE FOR FIRST LEXICAL TREE NODE
S8
REQUEST LANGUAGE MODEL SCORE USING PREVIOUS N-1 WORDS IN RECEIVED DATA AND LEXICAL TREE DATA IN ACOUSTIC MODEL MEMORY
S9
RECEIVE LANGUAGE MODEL SCORES FOR WORDS IN THE LEXICAL TREE AND PICK HIGHEST SCORE
S10
GENERATE TEMPORARY LEXICAL TREE SCORE USING THE SCORE FOR FIRST LEXICAL TREE NODE AND THE HIGHEST LANGUAGE MODEL SCORE
S11
SEND SCORE TO RESULTS MEMORY AS TEMPORARY LEXICAL TREE SCORE

Fig 4

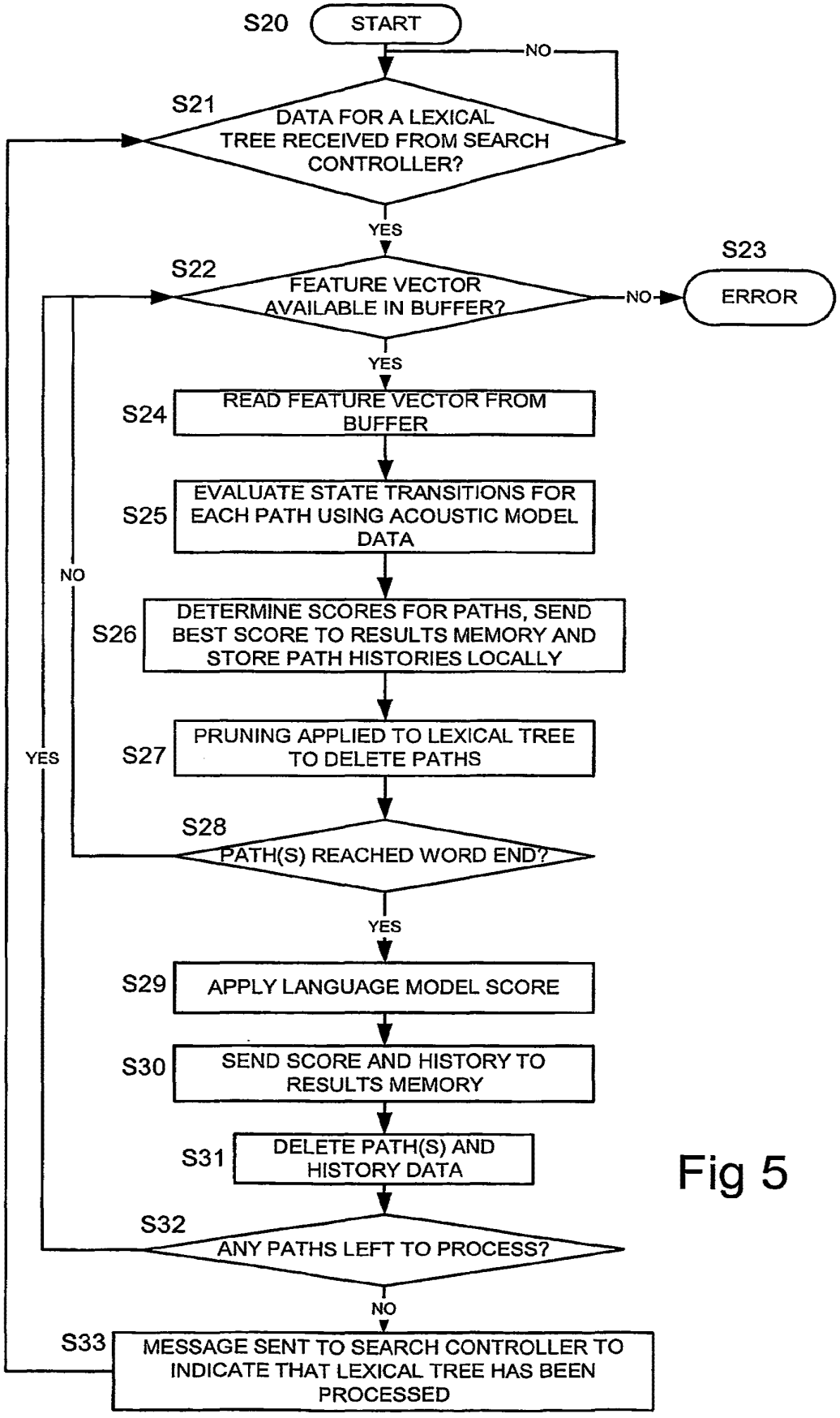

S20
START
NO
S21
DATA FOR A LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S22
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S23
ERROR
YES
S24
READ FEATURE VECTOR FROM BUFFER
S25
EVALUATE STATE TRANSITIONS FOR EACH PATH USING ACOUSTIC MODEL DATA
NO
S26
DETERMINE SCORES FOR PATHS, SEND BEST SCORE TO RESULTS MEMORY AND STORE PATH HISTORIES LOCALLY
S27
PRUNING APPLIED TO LEXICAL TREE TO DELETE PATHS
YES
S28
PATH(S) REACHED WORD END?
YES
S29
APPLY LANGUAGE MODEL SCORE
S30
SEND SCORE AND HISTORY TO RESULTS MEMORY
S31
DELETE PATH(S) AND HISTORY DATA
S32
ANY PATHS LEFT TO PROCESS?
NO
S33
MESSAGE SENT TO SEARCH CONTROLLER TO INDICATE THAT LEXICAL TREE HAS BEEN PROCESSED

Fig 5

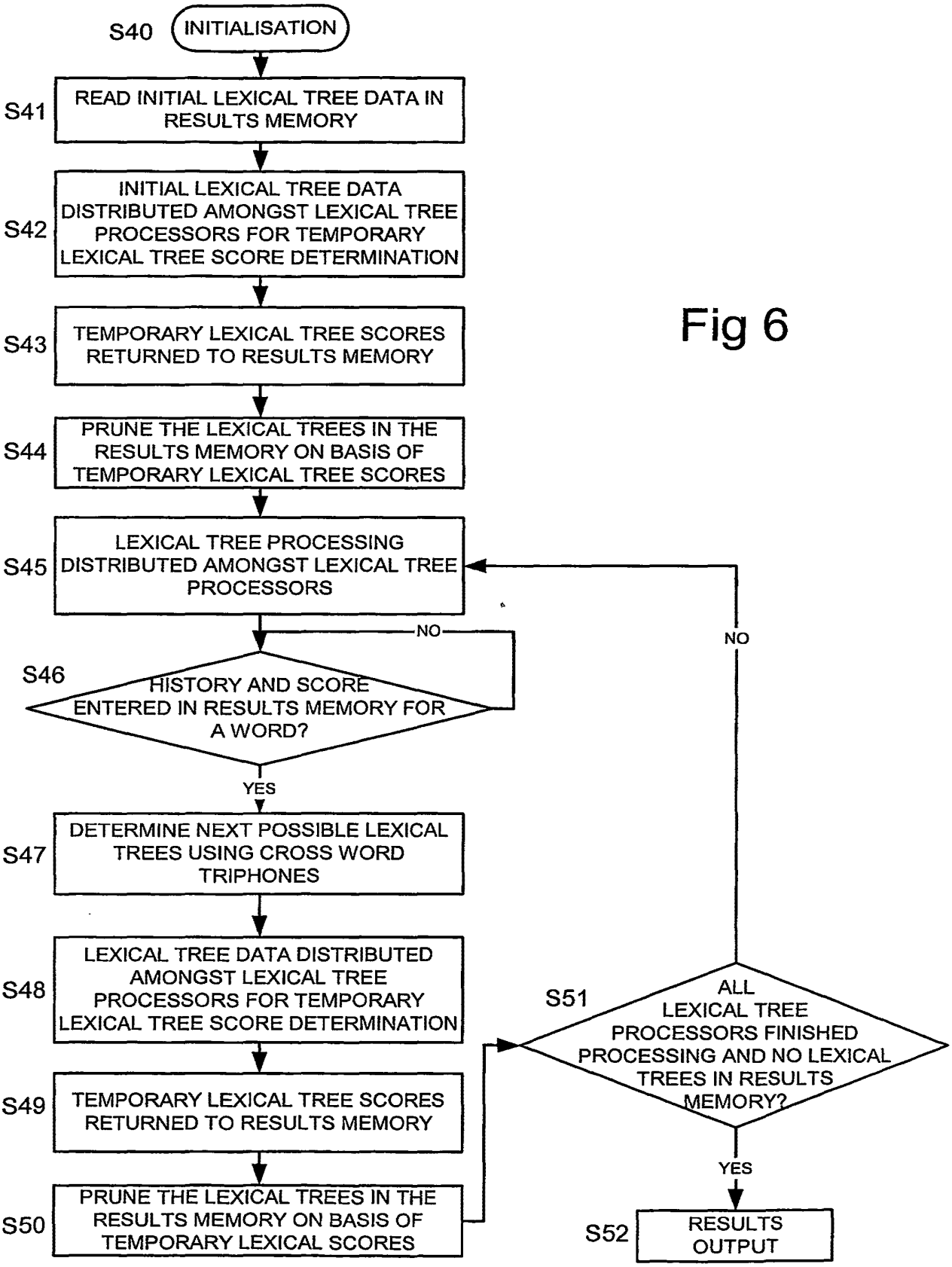

S40
INITIALISATION
S41
READ INITIAL LEXICAL TREE DATA IN RESULTS MEMORY
S42
INITIAL LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S43
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S44
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL TREE SCORES
S45
LEXICAL TREE PROCESSING DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS
S46
HISTORY AND SCORE ENTERED IN RESULTS MEMORY FOR A WORD?
NO
YES
S47
DETERMINE NEXT POSSIBLE LEXICAL TREES USING CROSS WORD TRIPHONES
S48
LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S49
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S50
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL SCORES
S51
ALL LEXICAL TREE PROCESSORS FINISHED PROCESSING AND NO LEXICAL TREES IN RESULTS MEMORY?
NO
YES
S52
RESULTS OUTPUT

Fig 6

PATENT APPLICATION SERIAL NO. _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

08/09/2004 MKAYPAGH 00000147 10503463

01 FC:2613	460.00 OP
02 FC:2615	639.00 OP
03 FC:2614	129.00 OP

PTO-1556
(5/87)

*U.S. Government Printing Office: 2002 — 489-267/69033

PATENT APPLICATION FEE DETERMINATION RECORD
Effective October 1, 2003

Application or Docket Number
10/503463

CLAIMS AS FILED - PART I

	(Column 1)	(Column 2)
TOTAL CLAIMS		
FOR	NUMBER FILED	NUMBER EXTRA
TOTAL CHARGEABLE CLAIMS	91 minus 20=	* 71
INDEPENDENT CLAIMS	6 minus 3 =	* 3
MULTIPLE DEPENDENT CLAIM PRESENT		☐

* If the difference in column 1 is less than zero, enter "0" in column 2

SMALL ENTITY TYPE ☐		OR	OTHER THAN SMALL ENTITY	
RATE	FEE		RATE	FEE
BASIC FEE	$460	OR	BASIC FEE	
XS 9=	$439	OR	XS18=	
X43=	$129	OR	X86=	
+145=		OR	-290=	
TOTAL	$1028	OR	TOTAL	

CLAIMS AS AMENDED - PART II

AMENDMENT A

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM				☐

SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
RATE	ADDITIONAL FEE		RATE	ADDITIONAL FEE
XS 9=		OR	XS18=	
X43=		OR	X86=	
+145=		OR	+290=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

AMENDMENT B

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM				☐

RATE	ADDITIONAL FEE		RATE	ADDITIONAL FEE
XS 9=		OR	XS18=	
X43=		OR	X86=	
+145=		OR	+290=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

AMENDMENT C

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM				☐

RATE	ADDITIONAL FEE		RATE	ADDITIONAL FEE
XS 9=		OR	XS18=	
X43=		OR	X86=	
+145=		OR	+290=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20."
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3."
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

IPR2023-00033
Apple EX1002 Page 46

10/503463
DT11 Rec'd PCT/PTO 04 AUG 2004

Docket No.: M0025.0311/P311
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Mark Catchpole

Application No.: Not Yet Assigned

Art Unit: N/A

Filed: Concurrently Herewith

Examiner: Not Yet Assigned

For: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

PRELIMINARY AMENDMENT

MS Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

INTRODUCTORY COMMENTS

Prior to examination on the merits, please amend the above-identified U.S. patent application as follows:

Amendments to the Claims are reflected in the listing of claims which begins on page 2 of this paper.

Remarks/Arguments begin on page 19 of this paper.

IPR2023-00033
Apple EX1002 Page 47

AMENDMENTS TO THE CLAIMS

1. (original) A speech recognition circuit comprising:

input means for receiving processed speech parameters;

lexical memory means containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input means for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory means;

results memory means connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

control processor means for controlling said lexical tree processors to process lexical trees identified in said results memory means by performing parallel processing on a plurality of said lexical tree data structures.

2. (original) A speech recognition circuit according to claim 1, wherein said lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

3. (original) A speech recognition circuit according to claim 2, wherein said lexical memory means stores each lexical tree data structure as mono phone models of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

IPR2023-00033
Apple EX1002 Page 48

4. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein each lexical tree processor is operative to compare the processed speech parameters with said lexical data or data derived from said lexical data to identify words as a word recognition event and to send information identifying the identified words to said results memory means as the processing results.

5. (original) A speech recognition circuit according to claim 4, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

6. (original) A speech recognition circuit according to claim 5, wherein each lexical tree processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and to output the modified processing result to said results memory means.

7. (currently amended) A speech recognition circuit according to ~~any one of claims 4 to 6~~ claim 4, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events.

8. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said lexical tree processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from said lexical data in said lexical memory means.

9. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said control processor means is operative to instruct a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory means.

IPR2023-00033
Apple EX1002 Page 49

10. (original) A speech recognition circuit according to claim 9, wherein said history data includes an accumulated score for the recognition path.

11. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said control processor means is operative to pass a lexical tree identifier to a said lexical tree processor, said lexical tree processor is operative to read lexical data for the identified lexical tree from said lexical memory means and to process said speech parameters using said read lexical data.

12. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said lexical tree processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

13. (original) A speech recognition circuit according to claim 12, wherein said lexical tree processors are operative to prune the processing of the lexical tree in dependence upon the determined scores.

14. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said lexical tree processors are operative to process more than one lexical tree at the same time.

15. (original) A speech recognition circuit according to ~~any one of claims 4 to 7~~ claim 4, wherein said control processor means is operative to determine new lexical tree identifiers for storing in said results memory means for words identified in said results memory means.

16. (original) A speech recognition circuit according to claim 15, wherein said control processor means is operative to prune the new lexical tree identifiers.

IPR2023-00033
Apple EX1002 Page 50

17. (currently amended) A speech recognition circuit according to ~~claims~~ claim 16, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

18. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said input means is arranged to receive said speech parameters as feature vectors.

19. (currently amended) A speech recognition circuit according to ~~claims 1 to 17~~ claim 1, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

20. (currently amended) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said lexical tree processors are arranged in groups, said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said groups of lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

21. (currently amended) A speech recognition circuit according to ~~any one of claims 1 to 19~~ claim 1, wherein said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a

IPR2023-00033
Apple EX1002 Page 51

lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

22. (original) A speech recognition circuit according to ~~any preceding~~ claim 1, wherein said lexical memory means stores said lexical tree data structures as Hidden Markov Models, and said lexical tree processors are operative to perform the Viterbi search algorithm using each lexical tree data structure.

23. (original) A speech recognition circuit comprising:

an input buffer for receiving processed speech parameters;

lexical memory containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, the initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory;

results memory connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

a control processor for controlling said lexical tree processors to process lexical trees identified in said results memory by performing parallel processing on a plurality of said lexical tree data structures.

24. (original) A speech recognition circuit according to claim 23, wherein said lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

25. (original) A speech recognition circuit according to claim 24, wherein said lexical memory stores said lexical tree data structure as a mono phone model of words, and

IPR2023-00033
Apple EX1002 Page 52

each lexical tree processor generates context dependant phone models of words dynamically for use in processing the speech parameters.

26. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 25~~ claim 23, wherein each lexical tree processor compares the processed speech parameters with the lexical data or data derived from the lexical data to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

27. (original) A speech recognition circuit according to claim 26, including a language model processor for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

28. (original) A speech recognition circuit according to claim 27, wherein each lexical tree processor receives said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and outputs the modified processing result to said results memory.

29. (currently amended) A speech recognition circuit according to ~~any one of claims 26 to 28~~ claim 26, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events.

30. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 29~~ claim 23, wherein said lexical tree processors process said speech parameters by comparing said speech parameters with said lexical data in said lexical memory.

31. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 30~~ claim 23, wherein said control processor instructs a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and

IPR2023-00033
Apple EX1002 Page 53

history data for a recognition path associated with the lexical tree from said results memory.

32. (original) A speech recognition circuit according to claim 31, wherein said history data includes an accumulated score for the recognition path.

33. (original) A speech recognition circuit according to ~~any one of claims 23 to 32~~ claim 23, wherein said control processor passes a lexical tree identifier to a said lexical tree processor, said lexical tree processor reads lexical data for the identified lexical tree from said lexical memory and processes said speech parameters using said read lexical data.

34. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 33~~ claim 23, wherein said lexical tree processors determine and output scores in the processing results during the processing of said speech parameters.

35. (original) A speech recognition circuit according to claim 34, wherein said lexical tree processors prune the processing of the lexical trees in dependence upon the determined scores.

36. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 35~~ claim 23, wherein said lexical tree processors process more than one lexical tree at the same time.

37. (currently amended) A speech recognition circuit according to ~~any one of claims 26 to 29~~ claim 26, wherein said control processor determines new lexical tree identifiers for storing in said results memory for words identified in said results memory.

38. (original) A speech recognition circuit according to claim 37, wherein said control processor means is operative to prune the new lexical tree identifiers

IPR2023-00033
Apple EX1002 Page 54

39. (currently amended) A speech recognition circuit according to ~~claims~~ claim 38, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

40. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 39~~ claim 23, wherein said input buffer receives said speech parameters as feature vectors.

41. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 39~~ claim 23, wherein said input buffer receives said speech parameters as n phones, where n is an integer.

42. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 41~~ claim 23, wherein said lexical tree processors are arranged in groups, said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

43. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 41~~ claim 23, wherein said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory is connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

IPR2023-00033
Apple EX1002 Page 55

44. (currently amended) A speech recognition circuit according to ~~any one of claims 23 to 43~~ claim 23, wherein said lexical memory stores said lexical tree data structures as Hidden Markov Models, and each lexical tree processor performs the Viterbi search algorithm using a respective said lexical tree data structure.

45. (original) A speech recognition circuit comprising:

input means for receiving processed speech parameters;

a plurality of lexical memory means containing in combination complete lexical data for word recognition, each lexical memory means containing part of said complete lexical data;

a plurality of processors connected in parallel to said input means for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory means;

control processor means for controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

receiving means for receiving the results of the processing of the speech parameters from said processors.

46. (original) A speech recognition circuit according to claim 45, wherein each lexical memory means stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

47. (original) A speech recognition circuit according to claim 46, wherein each lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

48. (original) A speech recognition circuit according to claim 47, wherein each lexical memory means stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors are arranged to use said mono phone models to

IPR2023-00033
Apple EX1002 Page 56

generate context dependant phone models of words dynamically for use in processing the speech parameters.

49. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 48~~ claim 45, wherein each processor is operative to compare the processed speech parameters with the lexical data in a corresponding lexical memory means to identify words as a word recognition event and to send information identifying the identified words to said receiving means as the processing results.

50. (original) A speech recognition circuit according to claim 49, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said processor.

51. (original) A speech recognition circuit according to claim 50, wherein each processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said processor and to output the modified processing result to said receiving means.

52. (currently amended) A speech recognition circuit according to ~~any one of claims 49 to 51~~ claim 49, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events.

53. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 52~~ claim 45, wherein said processors are perative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memory means.

54. (currently amended) A speech recognition circuit according to ~~any one of claims 46 to 48~~ claim 46, wherein said control processor means is operative to instruct a

IPR2023-00033
Apple EX1002 Page 57

said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

55. (original) A speech recognition circuit according to claim 54, wherein said history data includes an accumulated score for the recognition path.

56. (currently amended) A speech recognition circuit according to ~~any one of claims 46 to 48, 54 or 55~~ claim 46, wherein said control processor means is operative to pass a lexical tree identifier to a said processor, said processor is operative to read lexical data for the identified lexical tree from a corresponding said lexical memory means and to process said speech parameters using said read lexical data.

57. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 56~~ claim 45, wherein said processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

58. (original) A speech recognition circuit according to claim 57, wherein said processors are operative to prune the processing of the speech parameters in dependence upon the determined scores.

59. (currently amended) A speech recognition circuit according to ~~any one of claims 46 to 48, or 54 to 56~~ claim 46, wherein said processors are operative to use lexical data corresponding to more than one lexical tree at the same time.

60. (currently amended) A speech recognition circuit according to ~~any one of claims 49 to 52~~ claim 49, wherein said control processor means is operative to identify lexical data in said lexical memory means required for processing next words at word recognition events, and to control said processors to process the speech parameters using the identified lexical data.

IPR2023-00033
Apple EX1002 Page 58

61. (original) A speech recognition circuit according to claim 60, wherein said control processor means is operative to prune the identified lexical data.

62. (original) A speech recognition circuit according to claims 61, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune the identified lexical data on the basis of the scores for the words.

63. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 62~~ claim 45, wherein said input means is arranged to receive said speech parameters as feature vectors.

64. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 62~~ claim 45, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

65. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 64~~ claim 45, wherein said control processor means is operative to control each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory means.

66. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 65~~ claim 45, wherein said receiving means comprises a memory for storing the results of processing.

67. (currently amended) A speech recognition circuit according to ~~any one of claims 45 to 66~~ claim 45, wherein said lexical memory means store said lexical data as Hidden Markov Models, and each processor is operative to performs the Viterbi search algorithm using a respective part of said lexical data.

IPR2023-00033
Apple EX1002 Page 59

68. (original) A speech recognition circuit comprising:

an input buffer receiving processed speech parameters;

a plurality of lexical memories containing in combination complete lexical data for word recognition, each lexical memory containing part of said complete lexical data;

a plurality of processors connected in parallel to said input buffer for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory;

a control processor controlling each processor to process said speech parameters using partial lexical data read from a respective said lexical memory; and

a results memory storing the results of the processing of the speech parameters from said processors.

69. (original) A speech recognition circuit according to claim 68, wherein each lexical memory stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

70. (original) A speech recognition circuit according to claim 69, wherein each lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

71. (original) A speech recognition circuit according to claim 70, wherein each lexical memory stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

72. (currently amended) A speech recognition circuit according to ~~any one of claims 68 to 71~~ claim 68, wherein each processor compares the processed speech parameters with the lexical data in a corresponding lexical memory to identify words as a

IPR2023-00033
Apple EX1002 Page 60

word recognition event and sends information identifying the identified words to said results memory as the processing results.

73. (original) A speech recognition circuit according to claim 72, a including language model processor providing a language model output for modifying the processing results at a word recognition event by a said processor.

74. (original) A speech recognition circuit according to claim 73, wherein each processor receives said language model output to modify the processing result at a word recognition event determined by said processor and outputs the modified processing result to said results memory.

75. (currently amended) A speech recognition circuit according to ~~any one of claims 73 to 74~~ claim 73, wherein each processor determines and outputs scores for words in the processing results at word recognition events.

76. (currently amended) A speech recognition circuit according to ~~any one of claims 68 to 75~~ claim 68, wherein said processors process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memories.

77. (currently amended) A speech recognition circuit according to ~~any one of claims 69 to 71~~ claim 69, wherein said control processor instructs a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

78. (original) A speech recognition circuit according to claim 77, wherein said history data includes an accumulated score for the recognition path.

IPR2023-00033
Apple EX1002 Page 61

79. (currently amended) A speech recognition circuit according to ~~any one of claims 69 to 71, 77 or 78~~ claim 69, wherein said control processor passes a lexical tree identifier to a said processor, said processor reads lexical data for the identified lexical tree from a corresponding said lexical memory and processes said speech parameters using said read lexical data.

80. (currently amended) A speech recognition circuit according to ~~any one of claims 68 to 79~~ claim 68, wherein said processors determine and output scores in the processing results during the processing of said speech parameters.

81. (original) A speech recognition circuit according to claim 80, wherein said processors prune the processing of the speech parameters in dependence upon the determined scores.

82. (currently amended) A speech recognition circuit according to ~~any one of claims 69 to 71, or 77~~ claim 69, wherein said processors use lexical data corresponding to more than one lexical tree at the same time.

83. (currently amended) A speech recognition circuit according to ~~any one of claims 72 to 75~~ claim 72, wherein said control processor identifies lexical data in said lexical memories required for processing next words at word recognition events, and controls said processors to process the speech parameters using the identified lexical data.

84. (original) A speech recognition circuit according to claim 83, wherein said control processor prunes the identified lexical data.

85. (original) A speech recognition circuit according to claims 84, wherein each processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes the identified lexical data on the basis of the scores for the words.

IPR2023-00033
Apple EX1002 Page 62

86. (currently amended) A speech recognition circuit according to ~~any one of claims 68 to 85~~ claim 68, wherein said input buffer is arranged to receive said speech parameters as feature vectors.

87. (currently amended) A speech recognition circuit according to ~~any one of claims 68 to 85~~ claim 68, wherein said input buffer is arranged to receive said speech parameters as n phones, where n is an integer.

88. (currently amended) A speech recognition circuit according to ~~any one of claims 68 to 88~~ claim 68, wherein said control processor controls each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory.

89. (currently amended) A speech recognition circuit according to ~~any one of claims 68 to 89~~ claim 68, wherein said lexical memories store said lexical data as Hidden Markov Models, and each processor performs the Viterbi search algorithm using a respective part of said lexical data.

90. (original) A speech recognition method comprising:

receiving processed speech parameters;

storing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;;

passing the speech parameters in parallel to a plurality of lexical tree processors to perform parallel lexical tree processing for word recognition by accessing said lexical data;

storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

IPR2023-00033
Apple EX1002 Page 63

controlling said lexical tree processors to process lexical trees identified by the stored identifiers by performing parallel processing on a plurality of said lexical tree data structures.

91. (original) A speech recognition method comprising:

receiving processed speech parameters;

storing complete lexical data for word recognition in a plurality of lexical memory means, each lexical memory means containing part of said complete lexical data;

processing the speech parameters in parallel using a plurality of processors connected in parallel, said processors being arranged in groups of processors, each group of processors being connected to a said lexical memory means;

controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

storing the results of the processing of the speech parameters from said processors.

IPR2023-00033
Apple EX1002 Page 64

REMARKS

By this Preliminary Amendment, claims 4, 7-9, 11-12, 14-15, 17-22, 26, 29-31, 33-34, 36-37, 39-44, 49, 52-54, 56-57, 59-60, 63-67, 72, 75-77, 79-80, 82-83, and 86-89 has been amended to eliminate multiple dependencies.

Examination on the merits is respectfully requested.

Dated: August 4, 2004

Respectfully submitted,

By _____

Stephen A. Soffen
Registration No.: 31,063
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street NW
Washington, DC 20037-1526
(202) 785-9700
Attorney for Applicant

IPR2023-00033
Apple EX1002 Page 65

10/503463
DT1 Rec'd PCT/PTO 04 AUG 2004

Docket No.: M0025.0311/P311
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Mark Catchpole

Application No.: Not Yet Assigned

Filed: Concurrently Herewith

Art Unit: N/A

For: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

Examiner: Not Yet Assigned

INFORMATION DISCLOSURE STATEMENT (IDS)

U.S. Patent and Trademark Office
220 20th Street S.
Customer Window
Crystal Plaza Two, Lobby, Room 1B03
Arlington, VA 22313-1450

Dear Sir:

Pursuant to 37 CFR 1.56, 1.97 and 1.98, the attention of the Patent and Trademark Office is hereby directed to the references listed on the attached PTO/SB/08. It is respectfully requested that the information be expressly considered during the prosecution of this application, and that the references be made of record therein and appear among the "References Cited" on any patent to issue therefrom.

This Information Disclosure Statement accompanies the new patent application submitted herewith.

Copies of the references on the PTO/SB/08 are not provided.

In accordance with 37 CFR 1.97(g), the filing of this Information Disclosure Statement shall not be construed to mean that a search has been made or that no other material information as defined in 37 CFR 1.56(a) exists. In accordance with 37 CFR

IPR2023-00033
Apple EX1002 Page 66

1.97(h), the filing of this Information Disclosure statement shall not be construed to be an admission that any patent, publication or other information referred to therein is "prior art" for this invention unless specifically designated as such.

It is submitted that the Information Disclosure Statement is in compliance with 37 CFR 1.98 and the Examiner is respectfully requested to consider the listed references.

The Director is hereby authorized to charge any deficiency in the fees filed, asserted to be filed or which should have been filed herewith (or with any paper hereafter filed in this application by this firm) to our Deposit Account No. 04-1073, under Order No. M0025.0311/P311. A duplicate copy of this paper is enclosed.

Dated: August 4, 2004

Respectfully submitted,

By _____

Stephen A. Soffen
Registration No.: 31,063
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street NW
Washington, DC 20037-1526
(202) 785-9700
Attorney for Applicant

IPR2023-00033
Apple EX1002 Page 67

PTO/SB/08a/b (08-03)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449A/B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	Not Yet Assigned
Filing Date	Concurrently Herewith
First Named Inventor	Mark Catchpole
Art Unit	N/A
Examiner Name	Not Yet Assigned
Attorney Docket Number	M0025.0311/P311

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear	T[6]

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	S.H. Chung et al., "A Parallel Phoneme Recognition Algorithm Based on Continuous Hidden Markov Model", Proceedings 13th International Parallel Processing Symposium and 10TH Symposium on Parallel and Distributed Processing, IPPS/SPDP 1999. Proceedings of 13TH International Parallel Processing Symposium and 10TH Symposium on Parallel and Distributed Pro., IEEE Comput. Soc. pp. 453-457 (1999).	
	CB	S.H. Chung et al., "A Parallel Computation Model for Integrated Speech and Natural Language Understanding." IEEE Transactions on Computers, Vol. 42 No. 10, October 1, 1993.	
	CC	N. Deshmukn et al., "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition: Working Toward a Solution to the Decoding Problem"; IEEE Signal Processing Magazine, Vol. 16, No. 5, pp. 84-107 (September 1999).	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1]Applicant's unique citation designation number (optional). [2]Applicant is to place a check mark here if English language Translation is attached.

PCT/GB03/01459
Rec'd PCT/PTC 04 AUG 2004

The Patent Office

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

RECD 04 JUN 2003
WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed J Evans.

Dated 11 April 2003

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

BEST AVAILABLE COPY

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

03FEB02 E673214-9 D03312
P01/7700 0.00-0202546.8

THE PATENT OFFICE
E
04 FEB 2002
LONDON

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1. Your reference GBP84480

2. Patent application number (The Patent Office will fill in this part) **0202546.8** 04 FEB 2002

3 Full name, address and postcode of the or of each applicant (underline all surnames)

Zentian Limited,
78 Cromwell Road
Cambridge
Cambridgeshire CB1 3EG
United Kingdom

Patents ADP number (if you know it) 8318107001

If the applicant is a corporate body, give the country/state of its incorporation United Kingdom

4. Title of the invention SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

5. Name of your agent (if you have one)
"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

Marks & Clerk
57 - 60 Lincolns Inn Fields
London WC2A 3LS

Patents ADP number (if you know it) 18001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country	Priority application No (if you know it)	Date of filing (day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing (day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support ofthis request? (Answer 'Yes' if:
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d)) Yes

Patents Form 1/77

Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	0
Description	18
Claim(s)	15
Abstract	1
Drawing(s)	6 + 6 &r

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent (Patents Form 7/77)	1
Request for preliminary examination and search (Patents Form 9/77)	1
Request for substantive examination (Patents Form 10/77)	
Any other documents (please specify)	

11. I/We request the grant of a patent on the basis of this application.

Signature M+C Date: 4 February 2002

12. Name and daytime telephone number of person to contact in the United Kingdom

GB Patent Filings
0207 400 3000

DUPLICATE

M&C Folio: GBP84480 Document #: 793844

SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

The present invention generally relates to a speech recognition circuit which uses parallel processors for processing the input speech data in parallel.

Conventional large vocabulary speech recognition can be divided into two processes: front end processing to generate processed speech parameters such as feature vectors, followed by a search process which attempts to find the most likely set of words spoken from a given vocabulary (lexicon).

The front end processing generally represents no problem for current processing systems. However, for large vocabulary, speaker independent speech recognition, it is the search process that presents the biggest challenge. An article by Deshmukh et al entitled "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition" (IEEE Signal Processing Magazine, September 1999, pages 84 to 107), the content of which is hereby incorporated by reference, discusses the general concepts of large vocabulary speech recognition. As discussed in this paper, one algorithm for performing the search is the Viterbi algorithm. The Viterbi algorithm is a parallel or breadth first search through a transition network of states of Hidden Markov Models. An acoustic model for words in a lexicon are represented as states of Hidden Markov Models. These states represent phones or n phones in a phone model of the words. The search requires the evaluation of possible word matches. It is known that such a search is computationally intensive.

In order to speed up the processing performed during such a search in a speech recognition system, parallel processing has been explored. In an article by M K Ravishankar entitled "Parallel Implementation of Fast Beam Search for Speaker-Independent Continuous Speech Recognition" (Indian Institute of Science, Bangalor, India, July 16, 1993) a multi-threaded implementation of a fast beam search algorithm is disclosed. The multi-threading implementation requires a significant amount of

communication and synchronization among threads. In an MSC project report by R Dujari entitled "Parallel Viterbi Search Algorithm for Speech Recognition" (MIT, February 1992) the parallel processing of input speech parameters is disclosed in which a lexical network is split statically among processors.

It is an object of the present invention to provide an improved circuit which can perform parallel processing of speech parameters.

In accordance with a first embodiment of the present invention, a speech recognition circuit comprises an input port such as input buffer for receiving parameterized speech data such as feature vectors. A lexical memory arrangement is provided which contains lexicon data for word recognition. The lexical data comprises a plurality of lexical tree data structures representing a plurality of lexical trees. Each lexical tree data structure comprises a model of words having common prefix components and an initial component which is unique as an initial component for lexical trees. A plurality of lexical tree processors are connected in parallel to the input port and perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory arrangement. A results memory arrangement is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory arrangement by performing parallel processing of a plurality of lexical tree data structures.

Thus in accordance with this embodiment of the present invention, the processing in order to perform word recognition is distributed across the processors by controlling the processors to perform processing on different lexical trees. The controller controls the processor by the processes to provide for efficient process management by distributing lexical processing to appropriate processors.

The lexical tree data structure can comprise a phone model of words, wherein the components comprise phones. For reduced storage, the lexical tree data structure can comprise a mono phone lexical tree. The mono phone lexical tree can be used to

generate context dependent phone models dynamically. This enables the use of context dependent phone models for matching and hence increased accuracy whilst not increasing memory requirements. Alternatively, the lexical tree data structure can comprise context dependent phone models.

The processing performed by each processor in one embodiment comprises the comparison of the speech parameters with the lexical data, e.g. phone models or data derived from the lexical data (e.g. dynamically generated context dependent phone models) to identify words as a word recognition event and to send information identifying the identified words to the results memory as the processing results. In this embodiment a language model processor arrangement can be provided for providing a language model output for modifying the processing results at a word recognition event by a lexical tree processor. The modification can either take place at each lexical tree processor, or at the language model processing arrangement.

In one embodiment each lexical tree processor determines an output score for words in the processing results at word recognition events. Thus in this embodiment the language model processing arrangement can modify the score using a score for a language model for n preceding words, where n is an integer.

In one embodiment the controller instructs a lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from the results memory. The history data preferably includes an accumulated score for the recognition path. This enables a score to be determined based on the score for the recognition path to accumulate a new score during recognition carried out using the lexical tree data structure. The scores can be output in the processing results to the results memory during the processing of the speech parameters so that the scores can be used for pruning.

In one embodiment of the present invention, each lexical tree processor operates on more than one lexical tree at the same time, e.g. two lexical trees represented by two different lexical tree data structures, or two lexical trees represented by the same data

structure but displaced in time (which can be termed to instances of the same lexical tree).

At word recognition events, the controller determines new lexical tree identifiers for storing in the results memory for words identified in the results memory for respective word events. In order to reduce the processing, the controller can prune the new lexical tree identifiers to reduce the number of lexical trees which are required to be processed. This pruning can be achieved using context dependant n phones to reduce the number of possible next phones. The number can be further reduced by using a language model look ahead technique.

In one embodiment of the present invention, the lexical tree processors are arranged in groups or clusters. The lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory is connected to one of the groups of lexical tree processors and contains part of the lexical data. Thus a group of lexical tree processors and a partial lexical memory form a cluster. Each lexical tree processor is operative to process the speech parameters using a partial lexical memory and the controller controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In another embodiment of the present invention the lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory being connected to one of the lexical tree processors and containing part of the lexical data. Each lexical tree processor processes the speech parameters using a corresponding partial lexical memory and the controller is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In one embodiment of the present invention the lexical memory arrangement stores the lexical tree data structures as Hidden Markov Models and the lexical tree processors are operative to perform the Viterbi search algorithm using each respective lexical tree data structure. Thus in this way, this embodiment of the present invention provides a parallel Viterbi lexical tree search process for speech recognition.

The first aspect of the present invention is a special purpose circuit built for performing the speech recognition search process in which there are a plurality of processors for performing parallel lexical tree processing on individual lexical tree processors.

In another aspect of the present invention a speech recognition circuit comprises an input port such as an input buffer for receiving parameterized speech data such as feature vectors. A plurality of lexical memories are provided which contain in combination complete lexical data for word recognition. Each lexical memory contains part of the complete lexical data. A plurality of processors are provided connected in parallel to the input port for processing the speech parameters in parallel. The processors are arranged in groups in which each group is connected to a corresponding lexical memory to form a cluster. A controller controls each processor to process the speech parameters using partial lexical data read from a corresponding lexical memory. The results of processing the speech parameters are output from the processors as recognition data.

Thus this aspect of the present invention provides a circuit in which speech recognition processing is performed in parallel by groups of processors operating in parallel in which each group accesses a common memory of lexical data. This aspect of the present invention provides the advantage of parallel processing of speech parameters and benefits from a limited segmentation of the lexical data. By providing a plurality of processors in a group with a common memory, flexibility in the processing is provided without being bandwidth limited by the interface to the memory that would occur if only a single memory were used for all processors. The arrangement is more flexible than the parallel processing arrangement in which each processor only has access to its own local memory and requires fewer memory interfaces (i.e. chip pins). Each processor within a group can access the same lexical data as any other processor in the group. The controller can thus control the parallel processing of input speech parameters in a more flexible manner. For example, it allows more than one processor to process input speech parameters using the same lexical data in a lexical memory. This is because the lexical data is segmented into domains which are accessible by multiple processors.

In a preferred embodiment this aspect of the present invention is used in combination with the first aspect of the present invention. In such an arrangement each processor performs lexical tree processing and the lexical data stored in each lexical memory comprises lexical tree data structures which each comprise a model of words having common prefix components and an initial component that is unique.

In preferred embodiments of the second aspect of the present invention, the preferred embodiments of the first aspect of the present invention are incorporated.

Embodiments of the present invention will now be described with reference to the accompanying drawings in which:

Figure 1 is a diagram of a speech data processing circuit for generating parameterized speech data (feature vectors);

Figure 2 is a diagram of a speech recognition circuit in accordance with an embodiment of the present invention;

Figures 3a and 3b are schematic diagrams illustrating lexical tree structures;

Figure 4 is a flow diagram illustrating the process performed by a lexical tree processor to determine a temporary lexical tree score in accordance with an embodiment of the present invention;

Figure 5 is a flow diagram illustrating the process performed by the lexical tree processor for processing the input feature vectors in accordance with an embodiment of the present invention; and

Figure 6 is a flow diagram illustrating the process performed by the controller in accordance with an embodiment of the present invention.

Figure 1 illustrates a typical circuit for the parameterization of input speech data. In this embodiment the parameters generated are speech vectors.

A microphone 1 records speech in an analogue form and this is input through an anti-aliasing filter 2 to an analogue-to-digital converter 3 which samples the speech at 48 kHz at 20 bits per sample. The digitized output signal is normalized (4) to generated a 10 millisecond data frame every 5 milliseconds with 5 milliseconds overlap (5). A pre-emphasis operation 6 is applied to the data followed by a hamming window 7. The data is then fast Fourier transformed (FFT) using a 512 point fast Fourier transform (8) before being filtered by filter bank 9 into 12 frequencies. The energy in the data frame 5 is also recorded (13) as an additional feature and together with the 12 frequency outputs of the filter bank 9, 13 feature vectors (10) are thus produced and these are output as part of the 39 feature vectors 14. First and second derivatives (11 and 12) are taken of the 13 feature vectors 10 to complete the generation of the 39 feature vectors 14.

The arrangement illustrated in Figure 1 is purely given for illustration. The present invention encompasses any means by which speech and data can be parameterized to a suitable form for input to the search process as will be described in more detail hereinafter.

Figure 2 is a schematic diagram of a speech recognition circuit in accordance with an embodiment of the present invention for performing the search process. The parameterized speech data, which in this embodiment comprise feature vectors are input to a feature vector buffer 20. The feature vector buffer 20 is provided to buffer the incoming feature vectors to allow lexical tree processors 21 to read and process the feature vectors in the buffer 20 via a feature vector bus 24. A plurality k of lexical tree processors 21 are arranged in a respective lexical tree processor cluster 22. Each lexical tree processor cluster 22 has an acoustic model memory 23 in which is stored lexical data for use by the lexical tree processors 21 within the lexical tree processor cluster 22. Each lexical tree processor 21 in the lexical tree processor cluster 22 are connected to the acoustic model memory 23 within the lexical tree processor 22. There are N lexical tree processor clusters and thus there are Nk lexical tree processors 21 connected by the

feature vector bus 24 to the feature vector buffer 20. Each lexical tree processor 21 is capable of processing a different lexical tree and thus Nk lexical trees can be processed in parallel. The acoustic model memories 23 store as a whole a complete set of lexical data, i.e. lexical tree data structures for use in the lexical tree processing by the lexical tree processors 21. Each acoustic model memory 23 contains part or a segment of the lexical tree data. Since lexical tree processors 21 in a lexical tree processor cluster 22 access the same acoustic model memory 23, it is possible for more than one lexical tree processor 21 to process the same lexical data. This provides for some degree of flexibility in the controlling of the processing by the lexical tree processors 21. Further, the acoustic model memories 23 need not contain only one copy of the lexical data. It is possible to build in a redundancy in the data to further enhance the flexibility. This avoids any bottleneck in the processing due to the search processing focusing on a small number of lexical trees.

A results memory 25 is provided for storing processing results from the lexical tree processors 21 which are received over the path score and history bus 26. The results memory 25 also stores information on lexical trees to identify which lexical trees are to be processed. A search controller 27 is provided to control the processing performed by the lexical tree processors 21 in dependence upon a program and data stored in program and data memory 28. The search controller reads the path scores and lexical tree identifiers from the results memory and controls the lexical tree processors accordingly. A language model processor 29 is provided which is connected to each lexical tree processor 21 by a language model bus 30. The language model processor 29 accesses a language model memory 31 to read language model data for provision to lexical tree processors 21 in response to language model data requests. External control of the language model memory 31 is provided by a word constrains input. The language model processor 29 determines a score for a word occurring following N previous words using N grams. When a lexical tree processor requires a language model score a request is sent to the language model processor 29 over the language model bus 30 identifying the current word and the N-1 previous words. A language model score for the N gram can be returned to the lexical tree processor 21 for the modification of the score at the end of a branch of lexical tree processing. The lexical tree processor can modify the score in accordance with the language model and output a score to the

results memory 25 for a word at the end of a branch of the lexical tree processing. Thus the results memory stores the results as an ordered list of scores for words together with their histories.

The results memory 25 stores the following data:

1. Initial lexical tree data. This comprises pointers to an initial set of lexical trees. No history data is associated with the initial set of lexical trees. The initial set of lexical trees is predetermined and stored in the results memory 25 based on the most likely initial phones of an utterance. This initial lexical tree data is required to initialize the search process.

2. History data for search results. This comprises a record of a recognition path through the lexical tree recognition process performed by the lexical tree processors 21. The history data includes the current word, the previous N-1 words, the current accumulated score, the phone history (for use in the determination of likely next lexical trees using cross word context dependent tri-phones), and an identifier or pointer to the lexical tree used for identifying the word.

3. Best scores for best paths being processed by each lexical tree processor 21. This information enables the search controller 27 to monitor the processing being performed by lexical tree processors 21 to determine whether a global pruning strategy should be applied in order to reassign processing performed by a lexical tree processor if its best score for its best path is below a threshold or well below the best scores for the paths being processed by other lexical tree processors 21.

4. Temporary lexical tree scores. These comprise tree scores which are determined as temporary scores to prune the next lexical trees to be processed at word ends. The temporary lexical tree scores include lexical tree identifiers or pointers to identify the next lexical trees to be processed. The scores enable the pruning of this list.

5. Pruning threshold. This can be a global threshold value for use in the pruning of the lexical trees globally, or a local threshold value for use by a lexical processor for locally pruning the processing performed by the lexical processor 21.

The acoustic model memory 23 stores a Hidden Markov Model for acoustically modelling words as lexical trees. The acoustic model memory 23 stores a plurality of lexical tree data structures. Each lexical tree data structure comprises an n phone model of a number of words having common prefix phones. The first node of the lexical tree (the root) comprises a common n phone to all words in the lexical tree and uniquely identifies the lexical tree.

Each lexical tree processor 21 includes on-board memory (or local memory) to be used during the lexical tree processing. This working memory has to store all of the parameters currently working on including current scores for all paths being processed within the lexical tree, and previous N-1 words for the lexical tree. The local storage of the previous N-1 words enables the lexical tree processor 21, when a word end is reached along a branch of the lexical tree, to send a request for the language model score for an N gram by sending the identity of the N-1 previous words together with the word identified at the end of the branch.

Figure 3a and 3b schematically illustrate lexical trees which can be processed by the lexical tree processor 21 during the recognition of the two words HARD ROCK. In Figure 3a a previous lexical tree terminated at a branch recognizing the word YOU and terminating with the mono phone uw which is associated with two context dependent tri-phones y-uw+k and y-uw+h. Thus the context dependent tri-phone associated with the last phone in the lexical tree word model for YOU indicates two possible next lexical trees, i.e. the lexical trees beginning with the mono phone k and h and having a left context phone of uw. As can be see in Figure 3a this word end YOU therefore leads to two possible next lexical trees. These two lexical trees are traversed in parallel by two different lexical tree processors accumulating scores for matching of input feature vectors into Hidden Markov Models of the context dependent tri-phones associated with each node in the tree. When the end of the branch is reached, a word end event is reached and a word is recognized. As can be seen in Figure 3a, in this example since

four words are of similar phone length, it is possible for the search strategy based on inputting feature vectors in parallel to simultaneously reach a number of possible word ends. These possible word ends are sent as processing results to the results memory. The results memory stores the accumulated score at the word end together with phone history to identify the last phone and its associated context dependent tri-phones. In this example, considering the branch where the word HARD is recognized, the last phone is d which has associated with it two context dependent tri-phones r-d+l and r-d+r. Thus the search controller 27 can identify next possible lexical trees using the next phone in the context dependent tri-phone. In this case, as can be seen in Figure 3b, the next possible lexical trees begin with the phones r and l and have associated with them context dependent tri-phones d-r+ao and d-l+ao respectively. Thus these are the next lexical trees that require processing following the end node or word event detected in the processing lexical tree 2. Figure 3b thus represents the processing that is required at the end of Figure 3a at the word node for HARD in processing the second lexical tree.

As can be seen in Figure 3b, the two lexical trees are processed in parallel by traversing through the branches to reach word ends by sequentially entering in feature vectors in parallel to lexical tree processors processing the respective lexical trees. When the word end is reached, the accumulated score is output to the results memory as described before. In this case since it is the last word, the important context dependent tri-phone associated with the final phone has silence (sil) as the following phone. The final phone can in fact have 50 context dependent tri-phones associated with it if there are 50 possible next phones (including silence). Figure 3b only illustrates the relevant one for the end of the utterance.

Figure 4 is a flow diagram illustrating the processing performed by a lexical tree processor 21 in order to determine a temporary lexical tree score for a lexical tree. When a word end is identified by a lexical tree processor, the processing results are sent to the results memory 25. The results memory identifies the last phone of the recognized words and thus enables the search controller 27 to identify possible next lexical trees using context dependent tri-phones as illustrated in Figure 3b. Although in theory if there are 50 phones there are 50^2 possible lexical trees, due to the use of the context dependant triphones, only 100-200 lexical trees are likely to be identified as

possible lexical trees. The search controller needs to further prune out the lexical trees which are unlikely to generate likely paths with high scores. In order to do this, instructions are sent to the lexical tree processors to determine a temporary lexical tree score which can be used to prune out lexical trees from the processing which have a low score.

The processing by the lexical tree processor to generate the temporary lexical tree score is shown in Figure 4 and for this process, following initialization, a lexical tree processor 21 awaits receipt of data for a new lexical tree from the search controller 27 (step S2). The data comprises a lexical tree pointer on an identifier to lexical tree data stored in the acoustic model memory 23, i.e. the lexical tree data structure, previous N-1 words (for use in language model score determination by the language model processor 29), and the current path score. The previous N-1 words includes the previously recognized word for which a score has been accumulated.

The lexical tree processor 21 then attempts to read the next feature vector from the feature vector buffer 20 (step S3) and if this is not available, an error occurs (step S4). If the feature vector is available in the feature vector buffer 20, the lexical tree processor 21 reads the feature vector from the buffer 20 (step S5) and evaluates the state transitions for the first lexical tree node using the acoustic model data in the acoustic model memory 23 (step S6). Since the state transitions for the first node will require several feature vectors to complete, a score is determined for a state transition of the first node in the lexical tree (step S7). The lexical tree processor 21 then sends a request to the language model processor 29 for a language model score. The request includes the previous N-1 words and all of the possible words represented by the lexical tree data structure in the acoustic model memory 23 (step S8). The language model processor returns scores for each of the N grams ending in the words represented by the lexical tree. The lexical tree processor thus receives language model scores for the words in the lexical tree and picks the highest score (step S9). Alternatively, the language model processor 29 can select the highest n gram score and return this to the lexical tree processor for the determination of the temporary lexical tree score. The temporary lexical tree score for the lexical tree is then generated using the score determined for the first state transition of the first lexical tree node and the highest language model score

(step S10). The temporary lexical tree score is then sent by the lexical tree processor 21 to the results memory 25 over the bus 26 (step S11). In this mode of processing, the processor then awaits the next data of the new lexical tree from the search controller (step S2).

This process is just one of the processes performed by the lexical tree processor. The main processing performed by the lexical tree processor is the lexical tree evaluation. This process will now be described with reference to the flow diagram of Figure 5.

After the start of the process (step S20) a lexical tree processor awaits receipt of data for a lexical tree from the search controller (step S21). The data comprises a lexical tree pointer to a lexical tree data structure stored in the acoustic model memory 23, the previous N-1 words in the recognition path, and the current accumulated score.

The lexical tree processor then attempts to read the next feature vector from the feature vector buffer 20 (step S22). If this is not available, an error occurs (step s23). If the feature vector is available in the feature vector buffer 20, the feature vector is read from the feature vector buffer 20 by the lexical tree processor 21 (step S24) and the lexical tree processor 21 evaluates the state transitions for each path using the acoustic model data in the acoustic model memory 23 (step S25). Scores for the paths are determined in accordance with the conventional Viterbi search technique and the best score for the best path is sent to the results memory 25 while the path histories are stored locally in the on-board memory of the lexical tree processor 21 (step S26). Pruning is applied to the lexical tree by the lexical tree processor to delete paths in order to keep the breadth of the search manageable (step S27). The pruning applied locally by the lexical tree processor can be purely on the basis of a local threshold, which can be provided by the search controller 27, or it can be determined on a relative basis dependent upon the range of scores for paths being processed within the lexical tree processor 21. If the path has not reached a word end (step S28) the lexical tree processor 21 attempts to read the next feature vector from the feature vector buffer 20 (step S22). If the path reaches a word end (step S28) the score must be modified by the language model score (step S29). There are two ways in which this can be done, in this embodiment the lexical tree processor 21 sends a request to the language model processor 29 for a language model

score. The request includes the current word and the N-1 previous words. The language model processor thus returns the language model score and the language model score is used to modify the current accumulated score at the word end. Alternatively, the lexical tree processor can send the language model processor 29 not just the current word and the N-1 previous words, but also the current accumulated score. The language model processor then determines the language model score using the language model memory 31 and modifies the accumulated score using the language model score. The language model processor 29 can then return the modified score to the lexical tree processor 21 which can then pass it to the results memory 25, or a connection between the language model processor 29 and the results memory 25 (not shown) can enable the language model processor 29 to send the score directly to the results memory 25 for the lexical tree processor 21. In this latter case, the language model processor 29 must also receive a lexical tree pointer to identify the lexical tree for which the score applies.

Assuming in this embodiment that the lexical tree processor calculates the modified score (step S29), the score and history is then sent by the lexical tree processor 21 to the results memory 25 (step S30). The history data sent to the results memory 25 comprises the lexical tree pointer identifying the lexical tree, the modified score at the word end, the phone history identifying at least the last phone to allow for context dependent tri-phone determination of next lexical trees, and the identity of the word identified in the evaluation process.

The lexical tree processor then deletes the path and history data in its on-board memory (step S31) and determines if there are any paths still left to be processed (step S32). If so, the lexical tree processor 21 tries to access the next feature vector available in the feature vector buffer 20 (step S22). If there are no paths left to be processed, a message is sent by the lexical tree processor to the search controller 27 to indicate that the lexical tree has been processed (step S33). In this level of processing the lexical tree processor will then await the next data for a lexical tree from the search controller 27.

The flow diagram of Figure 5 illustrates the processing of a single instance of a lexical tree. The single point in time a lexical tree can start to be processed from a feature

vector, i.e. lexical tree processing starts at a time frame. However, due to a word end returned from lexical tree processing in another lexical tree processor, the lexical tree processor may also be asked during a subsequent time frame to start processing another instance of the lexical tree with a different path history. Thus a lexical tree processor 21 can be processing multiple instances of the lexical tree with different paths having different path histories. In this instance, the pruning applied in step S27 needs to take into account all paths currently being processed in all instances of lexical tree processing by the lexical tree processor 21. Thus although pruning is applied locally, it is applied over more than one instance of the lexical tree.

The operation of the search controller 27 will now be described with reference to the flow diagram of Figure 6.

After initialization (step S40) initial lexical tree data in the results memory 25 is read (step S41). The search controller distributes the initial lexical tree data amongst the lexical tree processors for temporary lexical tree score determination (step S42) as described with reference to Figure 4. Temporary lexical tree scores are returned by the lexical tree processors 21 to the results memory 25 (step S43) and the search controller reads the temporary lexical tree scores in the results memory 25 and prunes the lexical trees in the results memory 25 on the basis of the read temporary lexical scores (step S44). Thus the set of initial lexical trees can be quite large and this requires a large number of temporary lexical tree score determinations to be performed by the lexical tree processors 21. However, since the lexical tree processors only need to determine a score for a single feature vector, this can be performed quite quickly. Thus the number of lexical trees for which full evaluation by the lexical tree processors 21 is required can thus be reduced by the selection of the lexical trees with the highest temporary lexical tree score. The search controller than distributes the lexical tree processing among the lexical tree processors 21 by sending each lexical tree processor data including a lexical tree pointer to a lexical tree data structure in the acoustic model memory 23 in the cluster 22. The search controller 27 then awaits the entry of processing results, i.e. history and scores in the results memory 25 for words (step S46). When history and scores for words are entered in the results memory, the search controller 27 determines next possible lexical trees using cross-word tri-phones (step S47). For example, as

illustrated in Figures 3a and 3b, when the word HARD is entered in the results memory, the search controller identifies the final phone of the word as being d and thus next possible tri-phones can comprise any tri-phones having the left phone d. If there are 50 possible phones, this results in 50^2 possible tri-phones. In practice, only 1 to 200 are likely. However, this may still represent too many for full evaluation and thus the lexical tree data for the next possible lexical trees are distributed amongst the lexical tree processors for temporary lexical tree score determination (step S48). Temporary lexical tree scores are returned to the results memory 25 by the lexical tree processors 21 which execute the process as described with reference to Figure 4 (step S49).

The search controller 27 prunes the list of next lexical trees in the results memory 25 on the basis of the temporary lexical scores (step S50). The search controller then determines if the words in the results memory have been processed and there are no lexical trees left to be processed and whether or not lexical tree processes are finished (step S51). If not, the lexical tree processing is distributed amongst the lexical tree processors (step S45). When the end of the utterance is reached, next possible lexical trees will all have been processed and the results can be output from the results memory 25 (step S52).

It can be seen from the foregoing description that the embodiment of the invention provides for an efficient parallel processing system incorporating a parallel lexical tree processing architecture and a segmented lexicon storage architecture. Each lexical tree processor performs lexical tree evaluation and language model look ahead for evaluation of company scores for lexical trees to assist in pruning of the number of potential lexical trees that require processing. The architecture is ideally suited to synchronous breadth first searching e.g. a Viterbi search (i.e. real time recognition). The breadth first search is performed synchronously by the parallel processors on the input search parameters when the language model look ahead takes place, it is possible for the feature vector buffer 20 to store a small number, e.g. one or two, feature vectors whilst the language model look ahead is performed on the feature vector so that synchronous breadth first searching can be carried out by re-reading the feature vector out of the buffer 20 when it is determined which lexical trees are to be evaluated by the

lexical trees processors 21 after a word end. Thus, the feature vector buffer 20 provides a method of decoupling the search parameterisation circuitry from the search circuit.

If the buffer causes processing to get too far behind, e.g. there are a lot of word end language model look ahead operations to be performed, it is possible for the lexical tree processors to be controlled to temporarily halve their local beamwidths (i.e. more vigorously prune) and thus complete their processing more quickly. This would allow the content of the buffer to be decreased as the processors catch up.

Although the present invention has been described hereinabove with reference to a specific embodiment, it will be apparent to a skilled person in the art that modifications lie within the spirit and scope of the present invention. Although in the embodiment described hereinabove, the acoustic model memories 23 were described as storing parts of the lexical data, it is also possible for more than one copy of the lexical data to be stored in different acoustic model memories 23. Thus, more than one lexical tree processor cluster 22 can have access to the same lexical data. This increases the flexibility and allows the controller to avoid any bottlenecks in processing which, for example, can occur if the highest pass scores all require processing by one or a number of lexical tree data structures residing in one of the acoustic model memories 23.

Although in the embodiments described hereinabove, the memory is partitioned in to a plurality of acoustic model memories 23, the present invention is not limited to the partitioning of the lexical data and a plurality of lexical tree processors can access a single memory containing lexical data.

The lexical data stored can comprise mono phone models of words. When the lexical data is partitioned, these can be distributed evenly across the memories. If there are 50 mono phones, there will be 50 lexical tree data structures having a unique initial model phone data node. Alternatively, the lexical tree data structures can comprise a context dependent n phone word models. If there are 50 mono phones, there are potentially 50^n lexical tree data structures. For tri-phones, this comprises 50^3 (i.e. 125,000). These are distributed so that trees with the same left context tri-phone, are eventually spread across memories when the lexical data is partitioned. Thus in this way the word end

where the left context tri-phone is fixed i.e. the same, the next possible lexical trees will inherently be distributed across the memories, thus ensuring efficient parallel processing of the next lexical trees.

Although the embodiment described hereinabove uses a single results memory accessible by all processors, the present invention encompasses a plurality of results memories. For example, a results memory can be associated with each processor to store the processing results and the controller can read the results in the plurality of results memories in order to control the processors.

Although in the embodiment described above the input speech parameters are described as being feature vectors, the present invention encompasses any type of parameterization of speech data as an input to the circuit. Such parameterization methods are well known in the art.

Although in the embodiment described above there is only one language model processor, more than one can be used and these can be used to operate in parallel for processing scores at word boundaries.

CLAIMS:

1. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

lexical memory means containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input means for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory means;

results memory means connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

control processor means for controlling said lexical tree processors to process lexical trees identified in said results memory means by performing parallel processing on a plurality of said lexical tree data structures.

2. A speech recognition circuit according to claim 1, wherein said lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

3. A speech recognition circuit according to claim 2, wherein said lexical memory means stores each lexical tree data structure as mono phone models of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

4. A speech recognition circuit according to any preceding claim, wherein each lexical tree processor is operative to compare the processed speech parameters with said lexical data or data derived from said lexical data to identify words as a word recognition event and to send information identifying the identified words to said results memory means as the processing results.

5. A speech recognition circuit according to claim 4, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

6. A speech recognition circuit according to claim 5, wherein each lexical tree processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and to output the modified processing result to said results memory means.

7. A speech recognition circuit according to any one of claims 4 to 6, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events.

8. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from said lexical data in said lexical memory means.

9. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to instruct a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory means.

10. A speech recognition circuit according to claim 9, wherein said history data includes an accumulated score for the recognition path.

11. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to pass a lexical tree identifier to a said lexical tree processor, said lexical tree processor is operative to read lexical data for the identified lexical tree from said lexical memory means and to process said speech parameters using said read lexical data.

12. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

13. A speech recognition circuit according to claim 12, wherein said lexical tree processors are operative to prune the processing of the lexical tree in dependence upon the determined scores.

14. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process more than one lexical tree at the same time.

15. A speech recognition circuit according to any one of claims 4 to 7, wherein said control processor means is operative to determine new lexical tree identifiers for storing in said results memory means for words identified in said results memory means.

16. A speech recognition circuit according to claim 15, wherein said control processor means is operative to prune the new lexical tree identifiers.

17. A speech recognition circuit according to claims 16, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

18. A speech recognition circuit according to any preceding claim, wherein said input means is arranged to receive said speech parameters as feature vectors.

19. A speech recognition circuit according to any one of claims 1 to 17, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

20. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are arranged in groups, said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said groups of lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

21. A speech recognition circuit according to any one of claims 1 to 19, wherein said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

22. A speech recognition circuit according to any preceding claim, wherein said lexical memory means stores said lexical tree data structures as Hidden Markov Models, and said lexical tree processors are operative to perform the Viterbi search algorithm using each lexical tree data structure.

23. A speech recognition circuit comprising:

an input buffer for receiving processed speech parameters;

lexical memory containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, the initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory;

results memory connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

a control processor for controlling said lexical tree processors to process lexical trees identified in said results memory by performing parallel processing on a plurality of said lexical tree data structures.

24. A speech recognition circuit according to claim 23, wherein said lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

25. A speech recognition circuit according to claim 24, wherein said lexical memory stores said lexical tree data structure as a mono phone model of words, and each lexical tree processor generates context dependant phone models of words dynamically for use in processing the speech parameters.

26. A speech recognition circuit according to any one of claims 23 to 25, wherein each lexical tree processor compares the processed speech parameters with the lexical data or data derived from the lexical data to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

27. A speech recognition circuit according to claim 26, including a language model processor for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

28. A speech recognition circuit according to claim 27, wherein each lexical tree processor receives said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and outputs the modified processing result to said results memory.

29. A speech recognition circuit according to any one of claims 26 to 28, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events.

30. A speech recognition circuit according to any one of claims 23 to 29, wherein said lexical tree processors process said speech parameters by comparing said speech parameters with said lexical data in said lexical memory.

31. A speech recognition circuit according to any one of claims 23 to 30, wherein said control processor instructs a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory.

32. A speech recognition circuit according to claim 31, wherein said history data includes an accumulated score for the recognition path.

33. A speech recognition circuit according to any one of claims 23 to 32, wherein said control processor passes a lexical tree identifier to a said lexical tree processor, said lexical tree processor reads lexical data for the identified lexical tree from said lexical memory and processes said speech parameters using said read lexical data.

34. A speech recognition circuit according to any one of claims 23 to 33, wherein said lexical tree processors determine and output scores in the processing results during the processing of said speech parameters.

35. A speech recognition circuit according to claim 34, wherein said lexical tree processors prune the processing of the lexical trees in dependence upon the determined scores.

36. A speech recognition circuit according to any one of claims 23 to 35, wherein said lexical tree processors process more than one lexical tree at the same time.

37. A speech recognition circuit according to any one of claims 26 to 29, wherein said control processor determines new lexical tree identifiers for storing in said results memory for words identified in said results memory.

38. A speech recognition circuit according to claim 37, wherein said control processor means is operative to prune the new lexical tree identifiers

39. A speech recognition circuit according to claims 38, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

40. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as feature vectors.

41. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as n phones, where n is an integer.

42. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical tree processors are arranged in groups, said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

43. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory is connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

44. A speech recognition circuit according to any one of claims 23 to 43, wherein said lexical memory stores said lexical tree data structures as Hidden Markov Models, and each lexical tree processor performs the Viterbi search algorithm using a respective said lexical tree data structure.

45. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

a plurality of lexical memory means containing in combination complete lexical data for word recognition, each lexical memory means containing part of said complete lexical data;

a plurality of processors connected in parallel to said input means for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory means;

control processor means for controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

receiving means for receiving the results of the processing of the speech parameters from said processors.

46. A speech recognition circuit according to claim 45, wherein each lexical memory means stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

47. A speech recognition circuit according to claim 46, wherein each lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

48. A speech recognition circuit according to claim 47, wherein each lexical memory means stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

49. A speech recognition circuit according to any one of claims 45 to 48, wherein each processor is operative to compare the processed speech parameters with the lexical data in a corresponding lexical memory means to identify words as a word recognition event and to send information identifying the identified words to said receiving means as the processing results.

50. A speech recognition circuit according to claim 49, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said processor.

51. A speech recognition circuit according to claim 50, wherein each processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said processor and to output the modified processing result to said receiving means.

52. A speech recognition circuit according to any one of claims 49 to 51, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events.

53. A speech recognition circuit according to any one of claims 45 to 52, wherein said processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memory means.

54. A speech recognition circuit according to any one of claims 46 to 48, wherein said control processor means is operative to instruct a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

55. A speech recognition circuit according to claim 54, wherein said history data includes an accumulated score for the recognition path.

56. A speech recognition circuit according to any one of claims 46 to 48, 54, or 55, wherein said control processor means is operative to pass a lexical tree identifier to a said processor, said processor is operative to read lexical data for the identified lexical tree from a corresponding said lexical memory means and to process said speech parameters using said read lexical data.

57. A speech recognition circuit according to any one of claims 45 to 56, wherein said processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

58. A speech recognition circuit according to claim 57, wherein said processors are operative to prune the processing of the speech parameters in dependence upon the determined scores.

59. A speech recognition circuit according to any one of claims 46 to 48, or 54 to 56, wherein said processors are operative to use lexical data corresponding to more than one lexical tree at the same time.

60. A speech recognition circuit according to any one of claims 49 to 52, wherein said control processor means is operative to identify lexical data in said lexical memory means required for processing next words at word recognition events, and to control said processors to process the speech parameters using the identified lexical data.

61. A speech recognition circuit according to claim 60, wherein said control processor means is operative to prune the identified lexical data.

62. A speech recognition circuit according to claims 61, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune the identified lexical data on the basis of the scores for the words.

63. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as feature vectors.

64. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

65. A speech recognition circuit according to any one of claims 45 to 64, wherein said control processor means is operative to control each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory means.

66. A speech recognition circuit according to any one of claims 45 to 65, wherein said receiving means comprises a memory for storing the results of processing.

67. A speech recognition circuit according to any one of claims 45 to 66, wherein said lexical memory means store said lexical data as Hidden Markov Models, and each processor is operative to performs the Viterbi search algorithm using a respective part of said lexical data.

68. A speech recognition circuit comprising:

an input buffer receiving processed speech parameters;

a plurality of lexical memories containing in combination complete lexical data for word recognition, each lexical memory containing part of said complete lexical data;

a plurality of processors connected in parallel to said input buffer for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory;

a control processor controlling each processor to process said speech parameters using partial lexical data read from a respective said lexical memory; and

a results memory storing the results of the processing of the speech parameters from said processors.

69. A speech recognition circuit according to claim 68, wherein each lexical memory stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

70. A speech recognition circuit according to claim 69, wherein each lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

71. A speech recognition circuit according to claim 70, wherein each lexical memory stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

72. A speech recognition circuit according to any one of claims 68 to 71, wherein each processor compares the processed speech parameters with the lexical data in a corresponding lexical memory to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

73. A speech recognition circuit according to claim 72, a including language model processor providing a language model output for modifying the processing results at a word recognition event by a said processor.

74. A speech recognition circuit according to claim 73, wherein each processor receives said language model output to modify the processing result at a word recognition event determined by said processor and outputs the modified processing result to said results memory.

75. A speech recognition circuit according to any one of claims 73 to 74, wherein each processor determines and outputs scores for words in the processing results at word recognition events.

76. A speech recognition circuit according to any one of claims 68 to 75, wherein said processors process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memories.

77. A speech recognition circuit according to any one of claims 69 to 71, wherein said control processor instructs a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

78. A speech recognition circuit according to claim 77, wherein said history data includes an accumulated score for the recognition path.

79. A speech recognition circuit according to any one of claims 69 to 71, 77, or 78, wherein said control processor passes a lexical tree identifier to a said processor, said processor reads lexical data for the identified lexical tree from a corresponding said lexical memory and processes said speech parameters using said read lexical data.

80. A speech recognition circuit according to any one of claims 68 to 79, wherein said processors determine and output scores in the processing results during the processing of said speech parameters.

81. A speech recognition circuit according to claim 80, wherein said processors prune the processing of the speech parameters in dependence upon the determined scores.

82. A speech recognition circuit according to any one of claims 69 to 71, or 77 to 79, wherein said processors use lexical data corresponding to more than one lexical tree at the same time.

83. A speech recognition circuit according to any one of claims 72 to 75, wherein said control processor identifies lexical data in said lexical memories required for processing next words at word recognition events, and controls said processors to process the speech parameters using the identified lexical data.

84. A speech recognition circuit according to claim 83, wherein said control processor prunes the identified lexical data.

85. A speech recognition circuit according to claims 84, wherein each processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes the identified lexical data on the basis of the scores for the words.

86. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as feature vectors.

87. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as n phones, where n is an integer.

88. A speech recognition circuit according to any one of claims 68 to 88, wherein said control processor controls each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory.

89. A speech recognition circuit according to any one of claims 68 to 89, wherein said lexical memories store said lexical data as Hidden Markov Models, and each processor performs the Viterbi search algorithm using a respective part of said lexical data.

90. A speech recognition method comprising:

receiving processed speech parameters;

storing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;;

passing the speech parameters in parallel to a plurality of lexical tree processors to perform parallel lexical tree processing for word recognition by accessing said lexical data;

storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

controlling said lexical tree processors to process lexical trees identified by the stored identifiers by performing parallel processing on a plurality of said lexical tree data structures.

91. A speech recognition method comprising:

receiving processed speech parameters;

storing complete lexical data for word recognition in a plurality of lexical memory means, each lexical memory means containing part of said complete lexical data;

processing the speech parameters in parallel using a plurality of processors connected in parallel, said processors being arranged in groups of processors, each group of processors being connected to a said lexical memory means;

controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

storing the results of the processing of the speech parameters from said processors.

ABSTRACT:

SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

A speech recognition circuit comprises an input buffer for receiving processed speech parameters. A lexical memory contains lexical data for word recognition. The lexical data comprises a plurality of lexical tree data structures. Each lexical tree data structure comprises a model of words having common prefix components. An initial component of each lexical tree structure is unique. A plurality of lexical tree processors are connected in parallel to the input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory. A results memory is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory by performing parallel processing on a plurality of said lexical tree data structures.

Fig 2

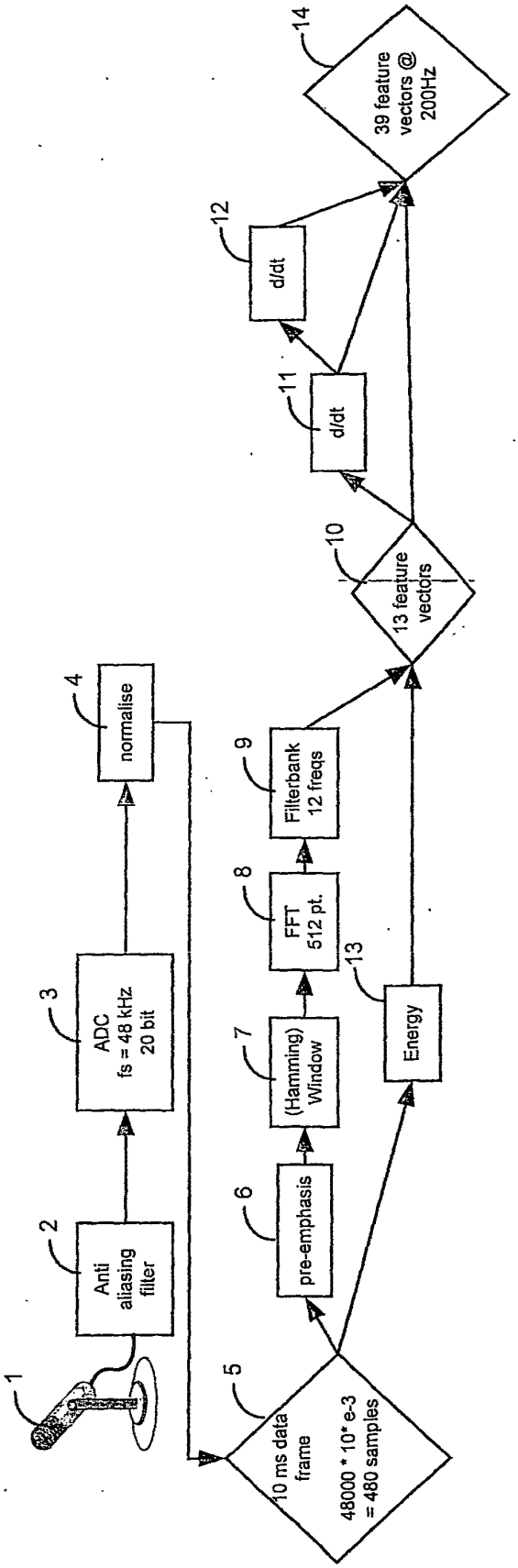

1
2
Anti aliasing filter
3
ADC fs = 48 kHz 20 bit
4
normalise
5
10 ms data frame 48000 * 10* e-3 = 480 samples
6
pre-emphasis
7
(Hamming) Window
8
FFT 512 pt.
9
Filterbank 12 freqs
10
13 feature vectors
11
d/dt
12
d/dt
13
Energy
14
39 feature vectors @ 200Hz

Fig 1

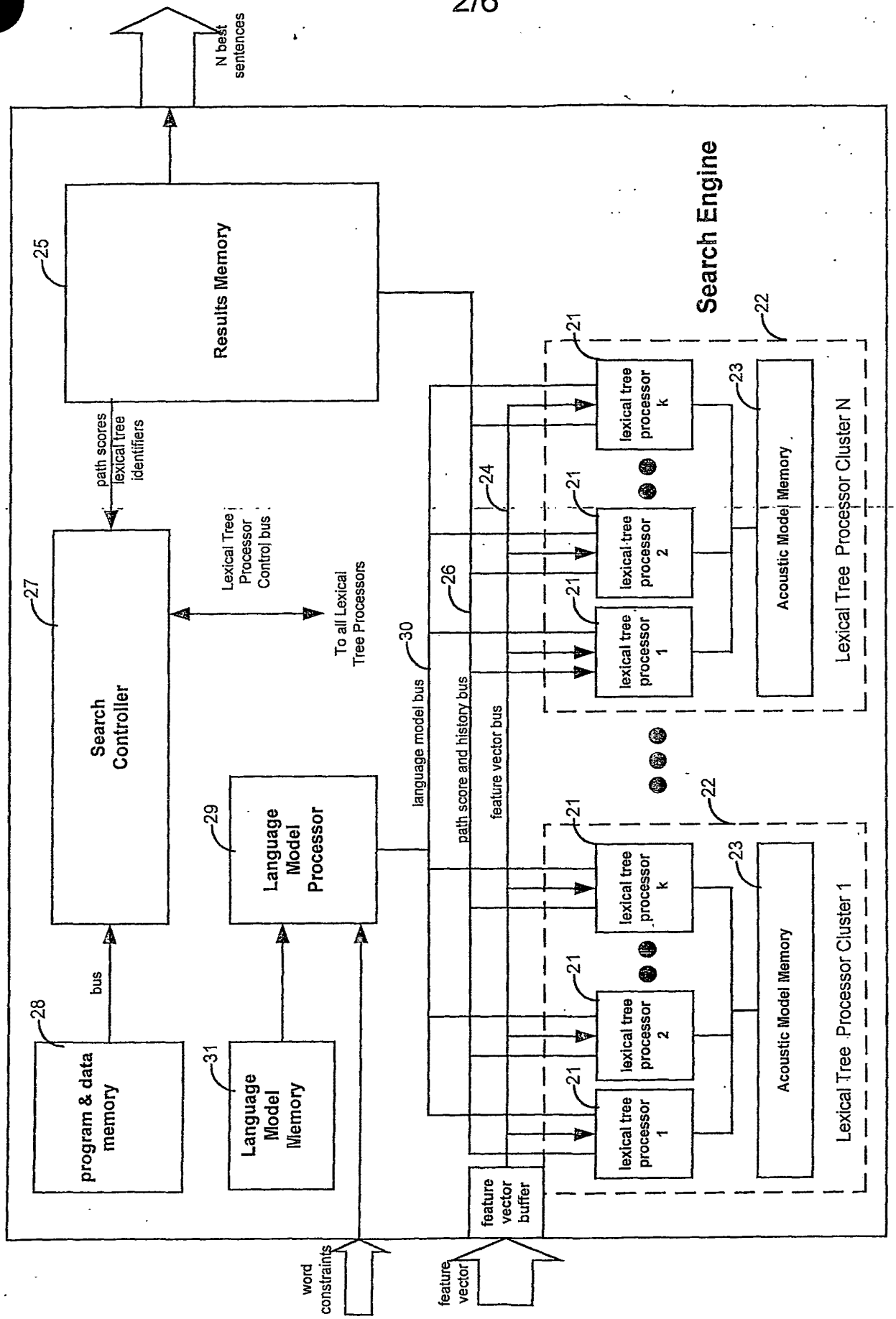

N best sentences
Search Engine
Results Memory
25
path scores
lexical tree identifiers
Search Controller
27
Lexical Tree Processor Control bus
To all Lexical Tree Processors
Language Model Processor
29
language model bus
30
path score and history bus
feature vector bus
26
24
bus
program & data memory
28
Language Model Memory
31
word constraints
feature vector
feature vector buffer
lexical tree processor 1
lexical tree processor 2
lexical tree processor k
21
Acoustic Model Memory
23
22
Lexical Tree Processor Cluster 1
Lexical Tree Processor Cluster N

Fig 2

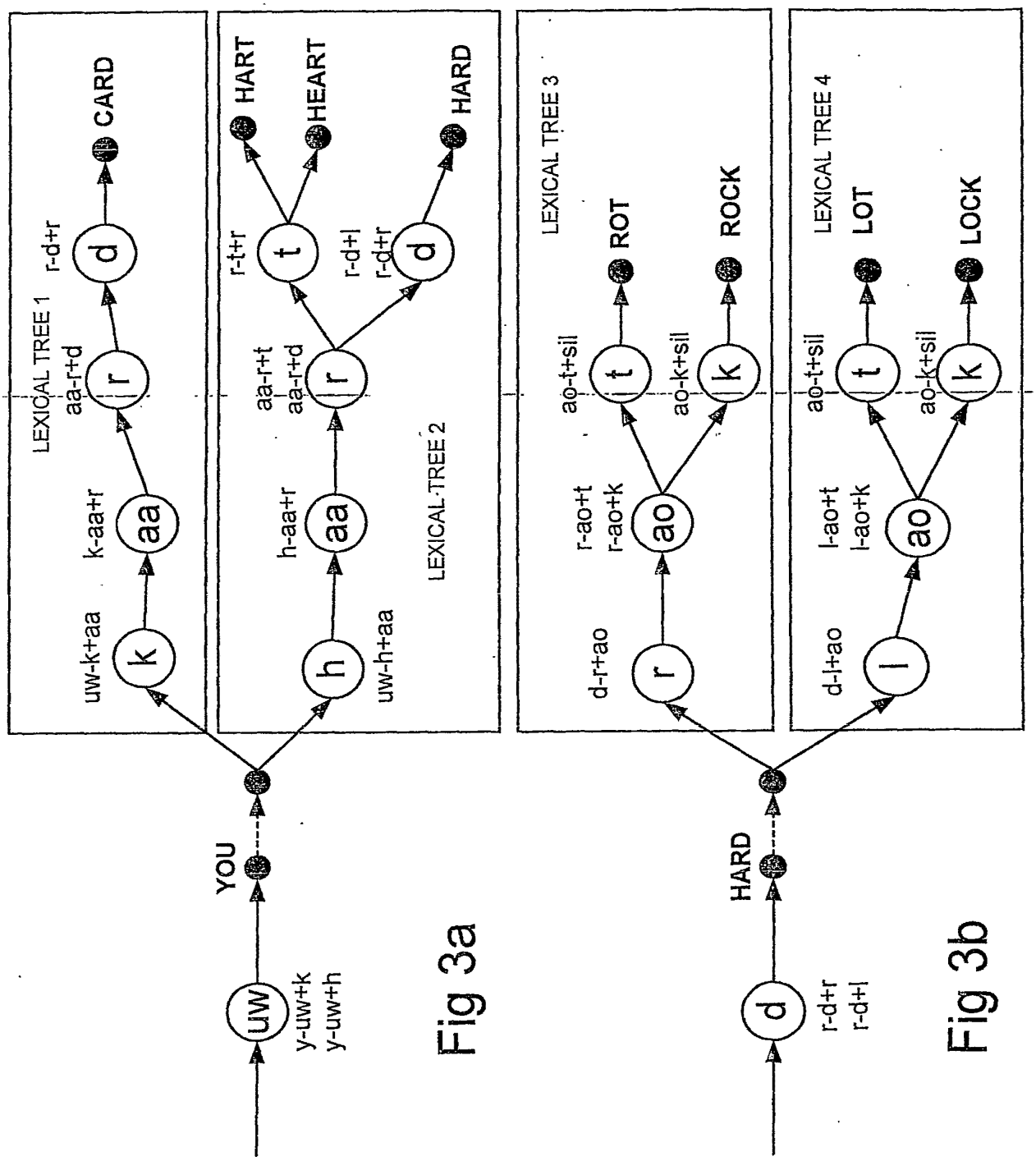

LEXICAL TREE 1
uw-k+aa
k-aa+r
aa-r+d
r-d+r
CARD
LEXICAL-TREE 2
uw-h+aa
h-aa+r
aa-r+t
aa-r+d
r-t+r
r-d+l
r-d+r
HART
HEART
HARD
YOU
y-uw+k
y-uw+h
LEXICAL TREE 3
d-r+ao
r-ao+t
r-ao+k
ao-t+sil
ao-k+sil
ROT
ROCK
LEXICAL TREE 4
d-l+ao
l-ao+t
l-ao+k
ao-t+sil
ao-k+sil
LOT
LOCK
HARD
r-d+r
r-d+l

Fig 3a

Fig 3b

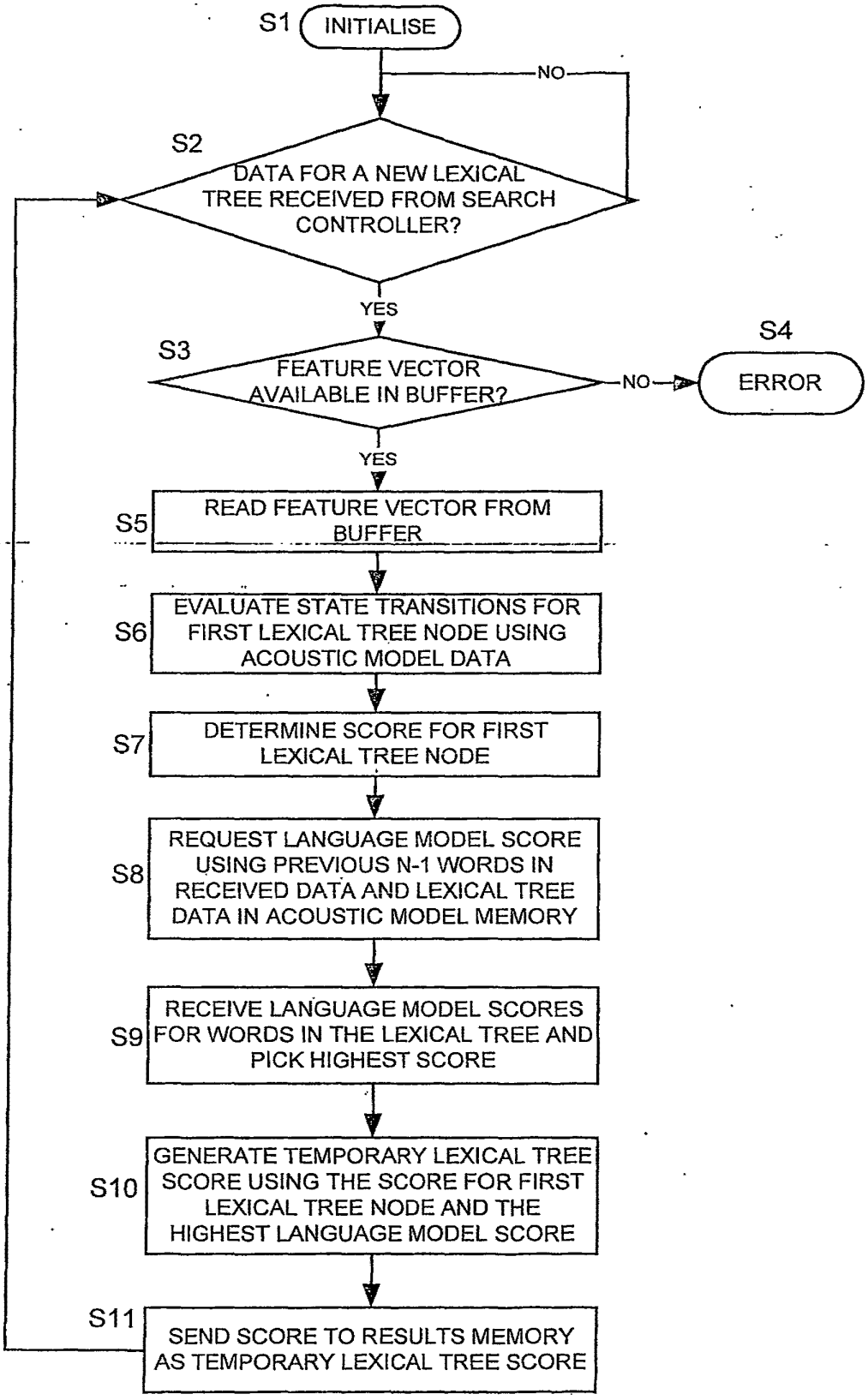

S1
INITIALISE
NO
S2
DATA FOR A NEW LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S3
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S4
ERROR
YES
S5
READ FEATURE VECTOR FROM BUFFER
S6
EVALUATE STATE TRANSITIONS FOR FIRST LEXICAL TREE NODE USING ACOUSTIC MODEL DATA
S7
DETERMINE SCORE FOR FIRST LEXICAL TREE NODE
S8
REQUEST LANGUAGE MODEL SCORE USING PREVIOUS N-1 WORDS IN RECEIVED DATA AND LEXICAL TREE DATA IN ACOUSTIC MODEL MEMORY
S9
RECEIVE LANGUAGE MODEL SCORES FOR WORDS IN THE LEXICAL TREE AND PICK HIGHEST SCORE
S10
GENERATE TEMPORARY LEXICAL TREE SCORE USING THE SCORE FOR FIRST LEXICAL TREE NODE AND THE HIGHEST LANGUAGE MODEL SCORE
S11
SEND SCORE TO RESULTS MEMORY AS TEMPORARY LEXICAL TREE SCORE

Fig 4

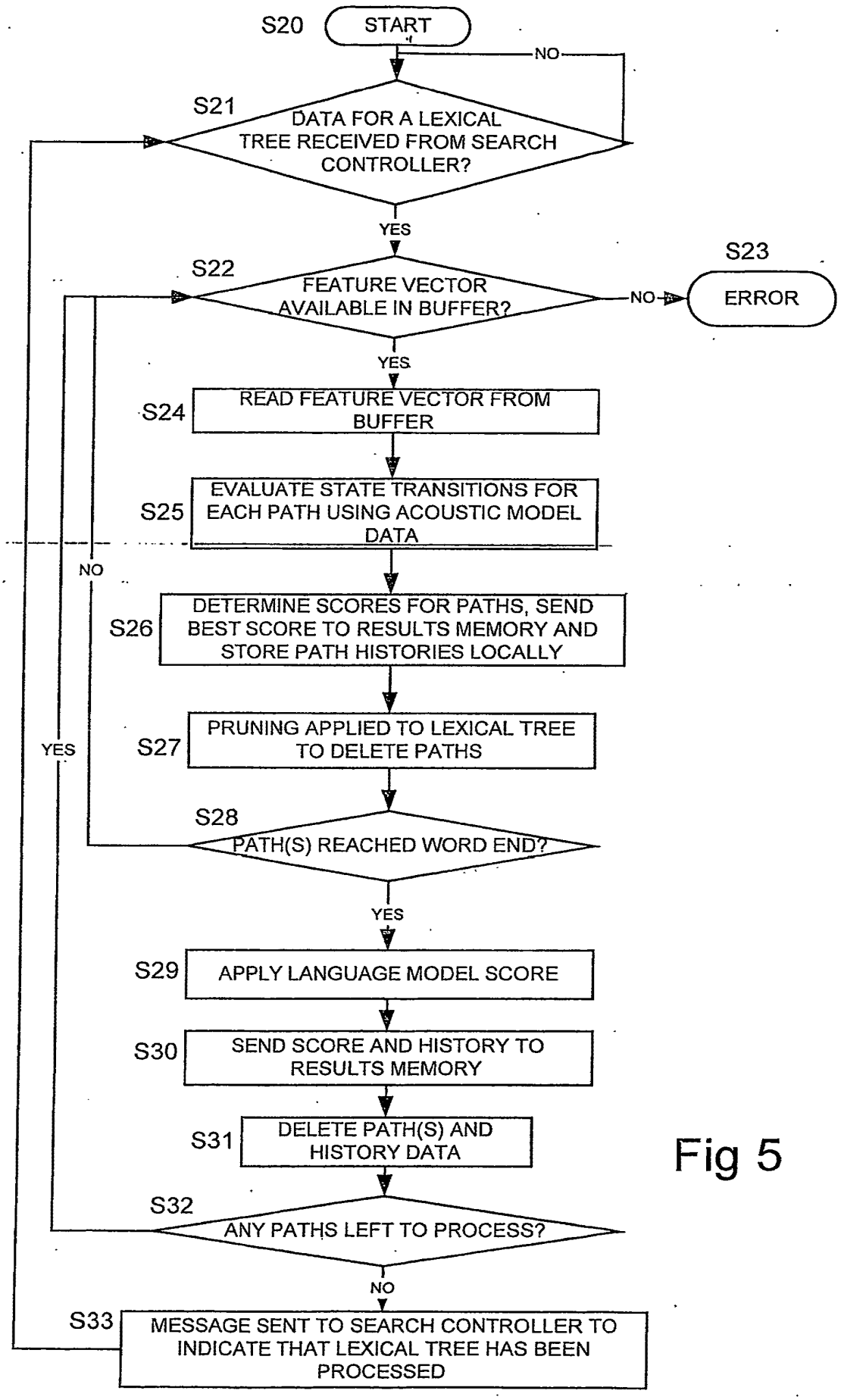

S20
START
NO
S21
DATA FOR A LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S22
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S23
ERROR
YES
S24
READ FEATURE VECTOR FROM BUFFER
S25
EVALUATE STATE TRANSITIONS FOR EACH PATH USING ACOUSTIC MODEL DATA
NO
S26
DETERMINE SCORES FOR PATHS, SEND BEST SCORE TO RESULTS MEMORY AND STORE PATH HISTORIES LOCALLY
S27
PRUNING APPLIED TO LEXICAL TREE TO DELETE PATHS
YES
S28
PATH(S) REACHED WORD END?
YES
S29
APPLY LANGUAGE MODEL SCORE
S30
SEND SCORE AND HISTORY TO RESULTS MEMORY
S31
DELETE PATH(S) AND HISTORY DATA
S32
ANY PATHS LEFT TO PROCESS?
NO
S33
MESSAGE SENT TO SEARCH CONTROLLER TO INDICATE THAT LEXICAL TREE HAS BEEN PROCESSED

Fig 5

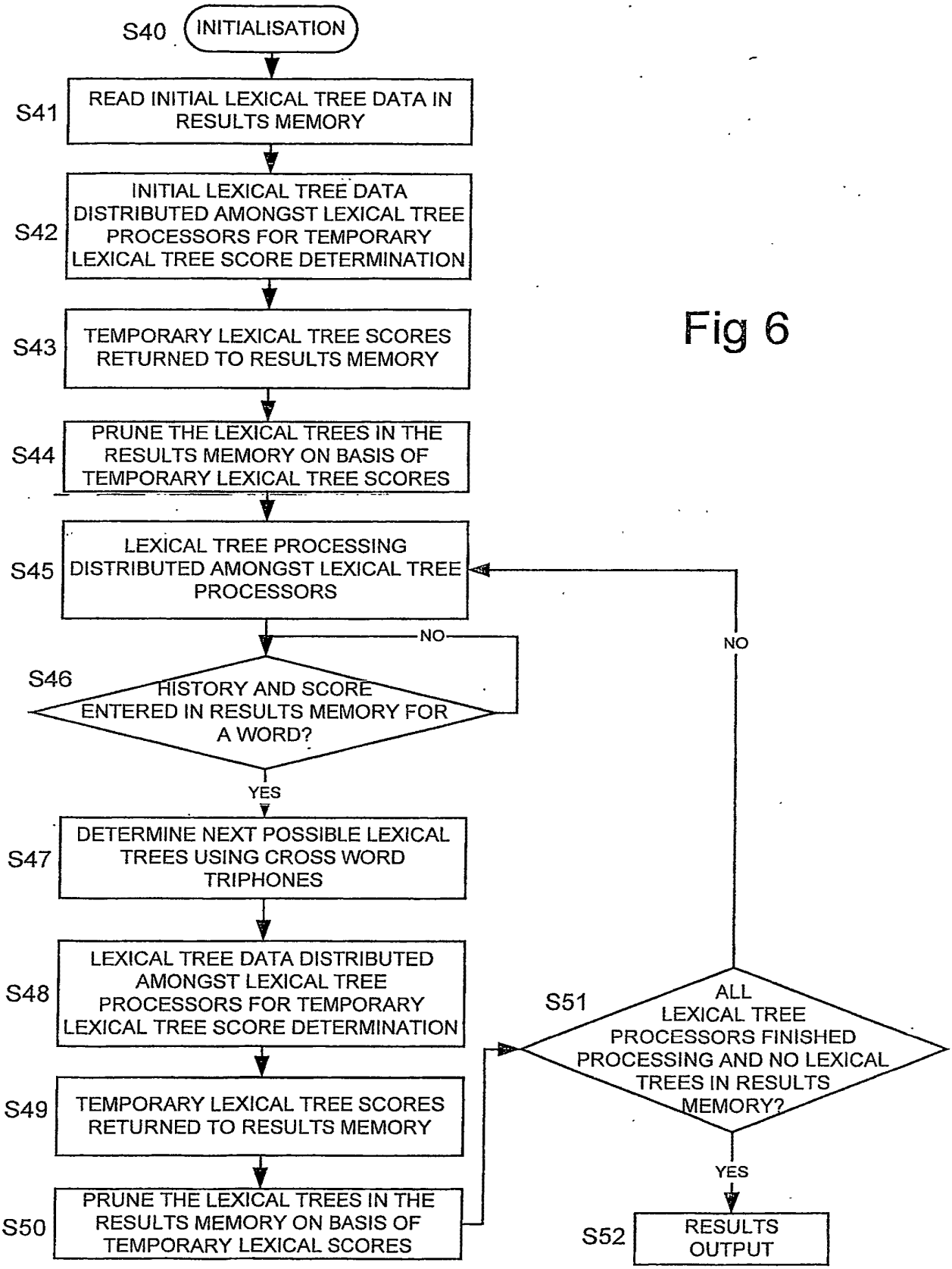

S40
INITIALISATION
S41
READ INITIAL LEXICAL TREE DATA IN RESULTS MEMORY
S42
INITIAL LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S43
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S44
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL TREE SCORES
S45
LEXICAL TREE PROCESSING DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS
S46
HISTORY AND SCORE ENTERED IN RESULTS MEMORY FOR A WORD?
NO
YES
S47
DETERMINE NEXT POSSIBLE LEXICAL TREES USING CROSS WORD TRIPHONES
S48
LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S49
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S50
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL SCORES
S51
ALL LEXICAL TREE PROCESSORS FINISHED PROCESSING AND NO LEXICAL TREES IN RESULTS MEMORY?
NO
YES
S52
RESULTS OUTPUT

Fig 6

THE PATENT OFFICE
11 APR 2003
Received in Patents
International Unit

U.S NATIONAL STAGE WORKSHEET (DO/EO)

U.S. APPL. NO. _____ INTERNATIONAL APPL. _____

APPLICATION FILED BY: 20 MOS., _____ OR 30 MOS., ____ SCREENED BY

INTERNATIONAL APPLICATION PAPERS IN THE APPLICATION FILE:

- ✓ International application
- ✓ Article 19 amendments
- ✓ Priority Document(s) No. 1
- ___ Request Form PCT/RO/101
- ___ PCT/IB/302
- ___ PCT/IB/304
- ___ PCT/IB/306
- ___ PCT/IB/308
- ___ PCT/IB/331
- ___ OTHER PCT/IB/_____
- ___ PCT/IPEA/409 also 416
- ___ 409 annexes to IPER
- ✓ PCT/ISA/210 (Search report)
- ___ Search report References
- ___ Other Papers filed

WIPO PUBLICATION
PUBLICATION NO. *WO* 03/062572
PUBLICATION DATE 7-31-03
PUBLICATION LANG., English
NOT PUBLISHED
_____ U.S. only _____ Requested

RECEIVED FROM THE APPLICANT: (other than checked above)

- ✓ National application basic fee paid
- ✓ Express Processing Requested
- ___ Translation of the International Application
- ✓ Used the IB copy of the IA
- ___ Description
- ✓ Claims 91
- ✓ Drawings 6
- ___ Foreign Language in drawing
- ___ Article 19 Amendments
- ___ Amendment used in application
- ___ Article 34 Amendment
- ___ Amendment used in application
- ___ DNA
- ___ 1194 transaction done
- ___ Sequence listing
- ✓ Preliminary Amendment(s) filed 8/4/04
- ___ second submission
- ✓ Information Disclosure Statement 8/4/04
- ___ second submission
- ✓ Assignment
- ✓ Forward to Assignment Branch
- ___ Substitute Specification
- ___ Small Entity Statement
- type
- ✓ Oath/Declaration (date submitted 5/24/05
- ___ Not executed
- ✓ Executed
- ___ Power of Attorney
- ___ Change of Address
- ___ Sequence Letter

USC Receipt of Request (PTO – 1399 Transmittal Letter)
te Acceptable oath/declaration received
2(e) Date
te complete 35 USC 371 requirements met

DATE NOTICE COMPLETED

)/EO 903 Notice of Acceptance
)/EO 905 Notice of Missing Requirements
)/EO 917 Notice of A defective oath or declaration
)/EO 916 Notice of defective response
)/EO 913 Notice of defective translation
/EO 909 Notification of Abandonment
/EO 920 Notification Nucleotide Sequence:

BEST AVAILABLE COPY

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2003 (14.08.2003)

PCT

(10) International Publication Number
WO 03/067572 A2

(51) International Patent Classification[7]: G10L 15/00

(21) International Application Number: PCT/GB03/00459

(22) International Filing Date: 4 February 2003 (04.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0202546.8 4 February 2002 (04.02.2002) GB

(71) Applicant *(for all designated States except US)*: **ZENTIAN LIMITED** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **CATCHPOLE, Mark** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(74) Agent: **COLLINS, John, David**; Marks & Clerk, 57-60 Lincolns Inn Fields, London WC2A 3LS (GB).

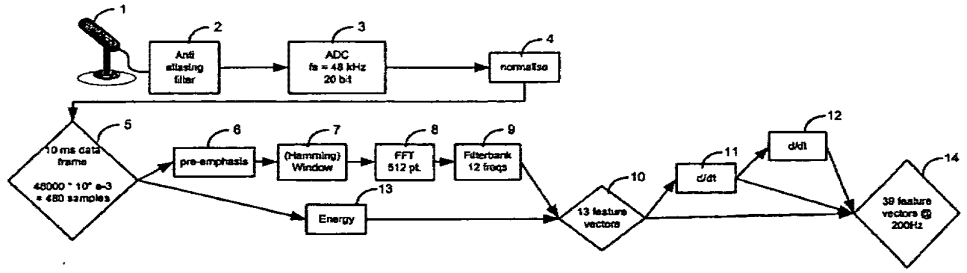
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS



(57) Abstract: A speech recognition circuit comprises an input buffer for receiving processed speech parameters. A lexical memory contains lexical data for word recognition. The lexical data comprises a plurality of lexical tree data structures. Each lexical tree data structure comprises a model of words having common prefix components. An initial component of each lexical tree structure is unique. A plurality of lexical tree processors are connected in parallel to the input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory. A results memory is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory by performing parallel processing on a plurality of said lexical tree data structures.

WO 03/067572 A2

SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

The present invention generally relates to a speech recognition circuit which uses parallel processors for processing the input speech data in parallel.

Conventional large vocabulary speech recognition can be divided into two processes: front end processing to generate processed speech parameters such as feature vectors, followed by a search process which attempts to find the most likely set of words spoken from a given vocabulary (lexicon).

The front end processing generally represents no problem for current processing systems. However, for large vocabulary, speaker independent speech recognition, it is the search process that presents the biggest challenge. An article by Deshmukh et al entitled "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition" (IEEE Signal Processing Magazine, September 1999, pages 84 to 107), the content of which is hereby incorporated by reference, discusses the general concepts of large vocabulary speech recognition. As discussed in this paper, one algorithm for performing the search is the Viterbi algorithm. The Viterbi algorithm is a parallel or breadth first search through a transition network of states of Hidden Markov Models. An acoustic model for words in a lexicon are represented as states of Hidden Markov Models. These states represent phones or n phones in a phone model of the words. The search requires the evaluation of possible word matches. It is known that such a search is computationally intensive.

In order to speed up the processing performed during such a search in a speech recognition system, parallel processing has been explored. In an article by M K Ravishankar entitled "Parallel Implementation of Fast Beam Search for Speaker-Independent Continuous Speech Recognition" (Indian Institute of Science, Bangalor, India, July 16, 1993) a multi-threaded implementation of a fast beam search algorithm is disclosed. The multi-threading implementation requires a significant amount of

communication and synchronization among threads. In an MSC project report by R Dujari entitled "Parallel Viterbi Search Algorithm for Speech Recognition" (MIT, February 1992) the parallel processing of input speech parameters is disclosed in which a lexical network is split statically among processors.

It is an object of the present invention to provide an improved circuit which can perform parallel processing of speech parameters.

In accordance with a first embodiment of the present invention, a speech recognition circuit comprises an input port such as input buffer for receiving parameterized speech data such as feature vectors. A lexical memory arrangement is provided which contains lexicon data for word recognition. The lexical data comprises a plurality of lexical tree data structures representing a plurality of lexical trees. Each lexical tree data structure comprises a model of words having common prefix components and an initial component which is unique as an initial component for lexical trees. A plurality of lexical tree processors are connected in parallel to the input port and perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory arrangement. A results memory arrangement is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory arrangement by performing parallel processing of a plurality of lexical tree data structures.

Thus in accordance with this embodiment of the present invention, the processing in order to perform word recognition is distributed across the processors by controlling the processors to perform processing on different lexical trees. The controller controls the processor by the processes to provide for efficient process management by distributing lexical processing to appropriate processors.

The lexical tree data structure can comprise a phone model of words, wherein the components comprise phones. For reduced storage, the lexical tree data structure can comprise a mono phone lexical tree. The mono phone lexical tree can be used to

generate context dependent phone models dynamically. This enables the use of context dependent phone models for matching and hence increased accuracy whilst not increasing memory requirements. Alternatively, the lexical tree data structure can comprise context dependent phone models.

The processing performed by each processor in one embodiment comprises the comparison of the speech parameters with the lexical data, e.g. phone models or data derived from the lexical data (e.g. dynamically generated context dependent phone models) to identify words as a word recognition event and to send information identifying the identified words to the results memory as the processing results. In this embodiment a language model processor arrangement can be provided for providing a language model output for modifying the processing results at a word recognition event by a lexical tree processor. The modification can either take place at each lexical tree processor, or at the language model processing arrangement.

In one embodiment each lexical tree processor determines an output score for words in the processing results at word recognition events. Thus in this embodiment the language model processing arrangement can modify the score using a score for a language model for n preceding words, where n is an integer.

In one embodiment the controller instructs a lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from the results memory. The history data preferably includes an accumulated score for the recognition path. This enables a score to be determined based on the score for the recognition path to accumulate a new score during recognition carried out using the lexical tree data structure. The scores can be output in the processing results to the results memory during the processing of the speech parameters so that the scores can be used for pruning.

In one embodiment of the present invention, each lexical tree processor operates on more than one lexical tree at the same time, e.g. two lexical trees represented by two different lexical tree data structures, or two lexical trees represented by the same data

structure but displaced in time (which can be termed to instances of the same lexical tree).

At word recognition events, the controller determines new lexical tree identifiers for storing in the results memory for words identified in the results memory for respective word events. In order to reduce the processing, the controller can prune the new lexical tree identifiers to reduce the number of lexical trees which are required to be processed. This pruning can be achieved using context dependant n phones to reduce the number of possible next phones. The number can be further reduced by using a language model look ahead technique.

In one embodiment of the present invention, the lexical tree processors are arranged in groups or clusters. The lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory is connected to one of the groups of lexical tree processors and contains part of the lexical data. Thus a group of lexical tree processors and a partial lexical memory form a cluster. Each lexical tree processor is operative to process the speech parameters using a partial lexical memory and the controller controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In another embodiment of the present invention the lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory being connected to one of the lexical tree processors and containing part of the lexical data. Each lexical tree processor processes the speech parameters using a corresponding partial lexical memory and the controller is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In one embodiment of the present invention the lexical memory arrangement stores the lexical tree data structures as Hidden Markov Models and the lexical tree processors are operative to perform the Viterbi search algorithm using each respective lexical tree data structure. Thus in this way, this embodiment of the present invention provides a parallel Viterbi lexical tree search process for speech recognition.

The first aspect of the present invention is a special purpose circuit built for performing the speech recognition search process in which there are a plurality of processors for performing parallel lexical tree processing on individual lexical tree processors.

In another aspect of the present invention a speech recognition circuit comprises an input port such as an input buffer for receiving parameterized speech data such as feature vectors. A plurality of lexical memories are provided which contain in combination complete lexical data for word recognition. Each lexical memory contains part of the complete lexical data. A plurality of processors are provided connected in parallel to the input port for processing the speech parameters in parallel. The processors are arranged in groups in which each group is connected to a corresponding lexical memory to form a cluster. A controller controls each processor to process the speech parameters using partial lexical data read from a corresponding lexical memory. The results of processing the speech parameters are output from the processors as recognition data.

Thus this aspect of the present invention provides a circuit in which speech recognition processing is performed in parallel by groups of processors operating in parallel in which each group accesses a common memory of lexical data. This aspect of the present invention provides the advantage of parallel processing of speech parameters and benefits from a limited segmentation of the lexical data. By providing a plurality of processors in a group with a common memory, flexibility in the processing is provided without being bandwidth limited by the interface to the memory that would occur if only a single memory were used for all processors. The arrangement is more flexible than the parallel processing arrangement in which each processor only has access to its own local memory and requires fewer memory interfaces (i.e. chip pins). Each processor within a group can access the same lexical data as any other processor in the group. The controller can thus control the parallel processing of input speech parameters in a more flexible manner. For example, it allows more than one processor to process input speech parameters using the same lexical data in a lexical memory. This is because the lexical data is segmented into domains which are accessible by multiple processors.

In a preferred embodiment this aspect of the present invention is used in combination with the first aspect of the present invention. In such an arrangement each processor performs lexical tree processing and the lexical data stored in each lexical memory comprises lexical tree data structures which each comprise a model of words having common prefix components and an initial component that is unique.

In preferred embodiments of the second aspect of the present invention, the preferred embodiments of the first aspect of the present invention are incorporated.

Embodiments of the present invention will now be described with reference to the accompanying drawings in which:

Figure 1 is a diagram of a speech data processing circuit for generating parameterized speech data (feature vectors);

Figure 2 is a diagram of a speech recognition circuit in accordance with an embodiment of the present invention;

Figures 3a and 3b are schematic diagrams illustrating lexical tree structures;

Figure 4 is a flow diagram illustrating the process performed by a lexical tree processor to determine a temporary lexical tree score in accordance with an embodiment of the present invention;

Figure 5 is a flow diagram illustrating the process performed by the lexical tree processor for processing the input feature vectors in accordance with an embodiment of the present invention; and

Figure 6 is a flow diagram illustrating the process performed by the controller in accordance with an embodiment of the present invention.

Figure 1 illustrates a typical circuit for the parameterization of input speech data. In this embodiment the parameters generated are speech vectors.

A microphone 1 records speech in an analogue form and this is input through an anti-aliasing filter 2 to an analogue-to-digital converter 3 which samples the speech at 48 kHz at 20 bits per sample. The digitized output signal is normalized (4) to generated a 10 millisecond data frame every 5 milliseconds with 5 milliseconds overlap (5). A pre-emphasis operation 6 is applied to the data followed by a hamming window 7. The data is then fast Fourier transformed (FFT) using a 512 point fast Fourier transform (8) before being filtered by filter bank 9 into 12 frequencies. The energy in the data frame 5 is also recorded (13) as an additional feature and together with the 12 frequency outputs of the filter bank 9, 13 feature vectors (10) are thus produced and these are output as part of the 39 feature vectors 14. First and second derivatives (11 and 12) are taken of the 13 feature vectors 10 to complete the generation of the 39 feature vectors 14.

The arrangement illustrated in Figure 1 is purely given for illustration. The present invention encompasses any means by which speech and data can be parameterized to a suitable form for input to the search process as will be described in more detail hereinafter.

Figure 2 is a schematic diagram of a speech recognition circuit in accordance with an embodiment of the present invention for performing the search process. The parameterized speech data, which in this embodiment comprise feature vectors are input to a feature vector buffer 20. The feature vector buffer 20 is provided to buffer the incoming feature vectors to allow lexical tree processors 21 to read and process the feature vectors in the buffer 20 via a feature vector bus 24. A plurality k of lexical tree processors 21 are arranged in a respective lexical tree processor cluster 22. Each lexical tree processor cluster 22 has an acoustic model memory 23 in which is stored lexical data for use by the lexical tree processors 21 within the lexical tree processor cluster 22. Each lexical tree processor 21 in the lexical tree processor cluster 22 are connected to the acoustic model memory 23 within the lexical tree processor 22. There are N lexical tree processor clusters and thus there are Nk lexical tree processors 21 connected by the

feature vector bus 24 to the feature vector buffer 20. Each lexical tree processor 21 is capable of processing a different lexical tree and thus Nk lexical trees can be processed in parallel. The acoustic model memories 23 store as a whole a complete set of lexical data, i.e. lexical tree data structures for use in the lexical tree processing by the lexical tree processors 21. Each acoustic model memory 23 contains part or a segment of the lexical tree data. Since lexical tree processors 21 in a lexical tree processor cluster 22 access the same acoustic model memory 23, it is possible for more than one lexical tree processor 21 to process the same lexical data. This provides for some degree of flexibility in the controlling of the processing by the lexical tree processors 21. Further, the acoustic model memories 23 need not contain only one copy of the lexical data. It is possible to build in a redundancy in the data to further enhance the flexibility. This avoids any bottleneck in the processing due to the search processing focusing on a small number of lexical trees.

A results memory 25 is provided for storing processing results from the lexical tree processors 21 which are received over the path score and history bus 26. The results memory 25 also stores information on lexical trees to identify which lexical trees are to be processed. A search controller 27 is provided to control the processing performed by the lexical tree processors 21 in dependence upon a program and data stored in program and data memory 28. The search controller reads the path scores and lexical tree identifiers from the results memory and controls the lexical tree processors accordingly. A language model processor 29 is provided which is connected to each lexical tree processor 21 by a language model bus 30. The language model processor 29 accesses a language model memory 31 to read language model data for provision to lexical tree processors 21 in response to language model data requests. External control of the language model memory 31 is provided by a word constrains input. The language model processor 29 determines a score for a word occurring following N previous words using N grams. When a lexical tree processor requires a language model score a request is sent to the language model processor 29 over the language model bus 30 identifying the current word and the N-1 previous words. A language model score for the N gram can be returned to the lexical tree processor 21 for the modification of the score at the end of a branch of lexical tree processing. The lexical tree processor can modify the score in accordance with the language model and output a score to the

results memory 25 for a word at the end of a branch of the lexical tree processing. Thus the results memory stores the results as an ordered list of scores for words together with their histories.

The results memory 25 stores the following data:

1. Initial lexical tree data. This comprises pointers to an initial set of lexical trees. No history data is associated with the initial set of lexical trees. The initial set of lexical trees is predetermined and stored in the results memory 25 based on the most likely initial phones of an utterance. This initial lexical tree data is required to initialize the search process.

2. History data for search results. This comprises a record of a recognition path through the lexical tree recognition process performed by the lexical tree processors 21. The history data includes the current word, the previous N-1 words, the current accumulated score, the phone history (for use in the determination of likely next lexical trees using cross word context dependent tri-phones), and an identifier or pointer to the lexical tree used for identifying the word.

3. Best scores for best paths being processed by each lexical tree processor 21. This information enables the search controller 27 to monitor the processing being performed by lexical tree processors 21 to determine whether a global pruning strategy should be applied in order to reassign processing performed by a lexical tree processor if its best score for its best path is below a threshold or well below the best scores for the paths being processed by other lexical tree processors 21.

4. Temporary lexical tree scores. These comprise tree scores which are determined as temporary scores to prune the next lexical trees to be processed at word ends. The temporary lexical tree scores include lexical tree identifiers or pointers to identify the next lexical trees to be processed. The scores enable the pruning of this list.

5. Pruning threshold. This can be a global threshold value for use in the pruning of the lexical trees globally, or a local threshold value for use by a lexical processor for locally pruning the processing performed by the lexical processor 21.

The acoustic model memory 23 stores a Hidden Markov Model for acoustically modelling words as lexical trees. The acoustic model memory 23 stores a plurality of lexical tree data structures. Each lexical tree data structure comprises an n phone model of a number of words having common prefix phones. The first node of the lexical tree (the root) comprises a common n phone to all words in the lexical tree and uniquely identifies the lexical tree.

Each lexical tree processor 21 includes on-board memory (or local memory) to be used during the lexical tree processing. This working memory has to store all of the parameters currently working on including current scores for all paths being processed within the lexical tree, and previous N-1 words for the lexical tree. The local storage of the previous N-1 words enables the lexical tree processor 21, when a word end is reached along a branch of the lexical tree, to send a request for the language model score for an N gram by sending the identity of the N-1 previous words together with the word identified at the end of the branch.

Figure 3a and 3b schematically illustrate lexical trees which can be processed by the lexical tree processor 21 during the recognition of the two words HARD ROCK. In Figure 3a a previous lexical tree terminated at a branch recognizing the word YOU and terminating with the mono phone uw which is associated with two context dependent tri-phones y-uw+k and y-uw+h. Thus the context dependent tri-phone associated with the last phone in the lexical tree word model for YOU indicates two possible next lexical trees, i.e. the lexical trees beginning with the mono phone k and h and having a left context phone of uw. As can be see in Figure 3a this word end YOU therefore leads to two possible next lexical trees. These two lexical trees are traversed in parallel by two different lexical tree processors accumulating scores for matching of input feature vectors into Hidden Markov Models of the context dependent tri-phones associated with each node in the tree. When the end of the branch is reached, a word end event is reached and a word is recognized. As can be seen in Figure 3a, in this example since

four words are of similar phone length, it is possible for the search strategy based on inputting feature vectors in parallel to simultaneously reach a number of possible word ends. These possible word ends are sent as processing results to the results memory. The results memory stores the accumulated score at the word end together with phone history to identify the last phone and its associated context dependent tri-phones. In this example, considering the branch where the word HARD is recognized, the last phone is d which has associated with it two context dependent tri-phones r-d+l and r-d+r. Thus the search controller 27 can identify next possible lexical trees using the next phone in the context dependent tri-phone. In this case, as can be seen in Figure 3b, the next possible lexical trees begin with the phones r and l and have associated with them context dependent tri-phones d-r+ao and d-l+ao respectively. Thus these are the next lexical trees that require processing following the end node or word event detected in the processing lexical tree 2. Figure 3b thus represents the processing that is required at the end of Figure 3a at the word node for HARD in processing the second lexical tree.

As can be seen in Figure 3b, the two lexical trees are processed in parallel by traversing through the branches to reach word ends by sequentially entering in feature vectors in parallel to lexical tree processors processing the respective lexical trees. When the word end is reached, the accumulated score is output to the results memory as described before. In this case since it is the last word, the important context dependent tri-phone associated with the final phone has silence (sil) as the following phone. The final phone can in fact have 50 context dependent tri-phones associated with it if there are 50 possible next phones (including silence). Figure 3b only illustrates the relevant one for the end of the utterance.

Figure 4 is a flow diagram illustrating the processing performed by a lexical tree processor 21 in order to determine a temporary lexical tree score for a lexical tree. When a word end is identified by a lexical tree processor, the processing results are sent to the results memory 25. The results memory identifies the last phone of the recognized words and thus enables the search controller 27 to identify possible next lexical trees using context dependent tri-phones as illustrated in Figure 3b. Although in theory if there are 50 phones there are 50^2 possible lexical trees, due to the use of the context dependant triphones, only 100-200 lexical trees are likely to be identified as

possible lexical trees. The search controller needs to further prune out the lexical trees which are unlikely to generate likely paths with high scores. In order to do this, instructions are sent to the lexical tree processors to determine a temporary lexical tree score which can be used to prune out lexical trees from the processing which have a low score.

The processing by the lexical tree processor to generate the temporary lexical tree score is shown in Figure 4 and for this process, following initialization, a lexical tree processor 21 awaits receipt of data for a new lexical tree from the search controller 27 (step S2). The data comprises a lexical tree pointer on an identifier to lexical tree data stored in the acoustic model memory 23, i.e. the lexical tree data structure, previous N-1 words (for use in language model score determination by the language model processor 29), and the current path score. The previous N-1 words includes the previously recognized word for which a score has been accumulated.

The lexical tree processor 21 then attempts to read the next feature vector from the feature vector buffer 20 (step S3) and if this is not available, an error occurs (step S4). If the feature vector is available in the feature vector buffer 20, the lexical tree processor 21 reads the feature vector from the buffer 20 (step S5) and evaluates the state transitions for the first lexical tree node using the acoustic model data in the acoustic model memory 23 (step S6). Since the state transitions for the first node will require several feature vectors to complete, a score is determined for a state transition of the first node in the lexical tree (step S7). The lexical tree processor 21 then sends a request to the language model processor 29 for a language model score. The request includes the previous N-1 words and all of the possible words represented by the lexical tree data structure in the acoustic model memory 23 (step S8). The language model processor returns scores for each of the N grams ending in the words represented by the lexical tree. The lexical tree processor thus receives language model scores for the words in the lexical tree and picks the highest score (step S9). Alternatively, the language model processor 29 can select the highest n gram score and return this to the lexical tree processor for the determination of the temporary lexical tree score. The temporary lexical tree score for the lexical tree is then generated using the score determined for the first state transition of the first lexical tree node and the highest language model score

(step S10). The temporary lexical tree score is then sent by the lexical tree processor 21 to the results memory 25 over the bus 26 (step S11). In this mode of processing, the processor then awaits the next data of the new lexical tree from the search controller (step S2).

This process is just one of the processes performed by the lexical tree processor. The main processing performed by the lexical tree processor is the lexical tree evaluation. This process will now be described with reference to the flow diagram of Figure 5.

After the start of the process (step S20) a lexical tree processor awaits receipt of data for a lexical tree from the search controller (step S21). The data comprises a lexical tree pointer to a lexical tree data structure stored in the acoustic model memory 23, the previous N-1 words in the recognition path, and the current accumulated score.

The lexical tree processor then attempts to read the next feature vector from the feature vector buffer 20 (step S22). If this is not available, an error occurs (step s23). If the feature vector is available in the feature vector buffer 20, the feature vector is read from the feature vector buffer 20 by the lexical tree processor 21 (step S24) and the lexical tree processor 21 evaluates the state transitions for each path using the acoustic model data in the acoustic model memory 23 (step S25). Scores for the paths are determined in accordance with the conventional Viterbi search technique and the best score for the best path is sent to the results memory 25 while the path histories are stored locally in the on-board memory of the lexical tree processor 21 (step S26). Pruning is applied to the lexical tree by the lexical tree processor to delete paths in order to keep the breadth of the search manageable (step S27). The pruning applied locally by the lexical tree processor can be purely on the basis of a local threshold, which can be provided by the search controller 27, or it can be determined on a relative basis dependent upon the range of scores for paths being processed within the lexical tree processor 21. If the path has not reached a word end (step S28) the lexical tree processor 21 attempts to read the next feature vector from the feature vector buffer 20 (step S22). If the path reaches a word end (step S28) the score must be modified by the language model score (step S29). There are two ways in which this can be done, in this embodiment the lexical tree processor 21 sends a request to the language model processor 29 for a language model

score. The request includes the current word and the N-1 previous words. The language model processor thus returns the language model score and the language model score is used to modify the current accumulated score at the word end. Alternatively, the lexical tree processor can send the language model processor 29 not just the current word and the N-1 previous words, but also the current accumulated score. The language model processor then determines the language model score using the language model memory 31 and modifies the accumulated score using the language model score. The language model processor 29 can then return the modified score to the lexical tree processor 21 which can then pass it to the results memory 25, or a connection between the language model processor 29 and the results memory 25 (not shown) can enable the language model processor 29 to send the score directly to the results memory 25 for the lexical tree processor 21. In this latter case, the language model processor 29 must also receive a lexical tree pointer to identify the lexical tree for which the score applies.

Assuming in this embodiment that the lexical tree processor calculates the modified score (step S29), the score and history is then sent by the lexical tree processor 21 to the results memory 25 (step S30). The history data sent to the results memory 25 comprises the lexical tree pointer identifying the lexical tree, the modified score at the word end, the phone history identifying at least the last phone to allow for context dependent tri-phone determination of next lexical trees, and the identity of the word identified in the evaluation process.

The lexical tree processor then deletes the path and history data in its on-board memory (step S31) and determines if there are any paths still left to be processed (step S32). If so, the lexical tree processor 21 tries to access the next feature vector available in the feature vector buffer 20 (step S22). If there are no paths left to be processed, a message is sent by the lexical tree processor to the search controller 27 to indicate that the lexical tree has been processed (step S33). In this level of processing the lexical tree processor will then await the next data for a lexical tree from the search controller 27.

The flow diagram of Figure 5 illustrates the processing of a single instance of a lexical tree. The single point in time a lexical tree can start to be processed from a feature

vector, i.e. lexical tree processing starts at a time frame. However, due to a word end returned from lexical tree processing in another lexical tree processor, the lexical tree processor may also be asked during a subsequent time frame to start processing another instance of the lexical tree with a different path history. Thus a lexical tree processor 21 can be processing multiple instances of the lexical tree with different paths having different path histories. In this instance, the pruning applied in step S27 needs to take into account all paths currently being processed in all instances of lexical tree processing by the lexical tree processor 21. Thus although pruning is applied locally, it is applied over more than one instance of the lexical tree.

The operation of the search controller 27 will now be described with reference to the flow diagram of Figure 6.

After initialization (step S40) initial lexical tree data in the results memory 25 is read (step S41). The search controller distributes the initial lexical tree data amongst the lexical tree processors for temporary lexical tree score determination (step S42) as described with reference to Figure 4. Temporary lexical tree scores are returned by the lexical tree processors 21 to the results memory 25 (step S43) and the search controller reads the temporary lexical tree scores in the results memory 25 and prunes the lexical trees in the results memory 25 on the basis of the read temporary lexical scores (step S44). Thus the set of initial lexical trees can be quite large and this requires a large number of temporary lexical tree score determinations to be performed by the lexical tree processors 21. However, since the lexical tree processors only need to determine a score for a single feature vector, this can be performed quite quickly. Thus the number of lexical trees for which full evaluation by the lexical tree processors 21 is required can thus be reduced by the selection of the lexical trees with the highest temporary lexical tree score. The search controller than distributes the lexical tree processing among the lexical tree processors 21 by sending each lexical tree processor data including a lexical tree pointer to a lexical tree data structure in the acoustic model memory 23 in the cluster 22. The search controller 27 then awaits the entry of processing results, i.e. history and scores in the results memory 25 for words (step S46). When history and scores for words are entered in the results memory, the search controller 27 determines next possible lexical trees using cross-word tri-phones (step S47). For example, as

illustrated in Figures 3a and 3b, when the word HARD is entered in the results memory, the search controller identifies the final phone of the word as being d and thus next possible tri-phones can comprise any tri-phones having the left phone d. If there are 50 possible phones, this results in 50^2 possible tri-phones. In practice, only 1 to 200 are likely. However, this may still represent too many for full evaluation and thus the lexical tree data for the next possible lexical trees are distributed amongst the lexical tree processors for temporary lexical tree score determination (step S48). Temporary lexical tree scores are returned to the results memory 25 by the lexical tree processors 21 which execute the process as described with reference to Figure 4 (step S49).

The search controller 27 prunes the list of next lexical trees in the results memory 25 on the basis of the temporary lexical scores (step S50). The search controller then determines if the words in the results memory have been processed and there are no lexical trees left to be processed and whether or not lexical tree processes are finished (step S51). If not, the lexical tree processing is distributed amongst the lexical tree processors (step S45). When the end of the utterance is reached, next possible lexical trees will all have been processed and the results can be output from the results memory 25 (step S52).

It can be seen from the foregoing description that the embodiment of the invention provides for an efficient parallel processing system incorporating a parallel lexical tree processing architecture and a segmented lexicon storage architecture. Each lexical tree processor performs lexical tree evaluation and language model look ahead for evaluation of company scores for lexical trees to assist in pruning of the number of potential lexical trees that require processing. The architecture is ideally suited to synchronous breadth first searching e.g. a Viterbi search (i.e. real time recognition). The breadth first search is performed synchronously by the parallel processors on the input search parameters when the language model look ahead takes place, it is possible for the feature vector buffer 20 to store a small number, e.g. one or two, feature vectors whilst the language model look ahead is performed on the feature vector so that synchronous breadth first searching can be carried out by re-reading the feature vector out of the buffer 20 when it is determined which lexical trees are to be evaluated by the

lexical trees processors 21 after a word end. Thus, the feature vector buffer 20 provides a method of decoupling the search parameterisation circuitry from the search circuit.

If the buffer causes processing to get too far behind, e.g. there are a lot of word end language model look ahead operations to be performed, it is possible for the lexical tree processors to be controlled to temporarily halve their local beamwidths (i.e. more vigorously prune) and thus complete their processing more quickly. This would allow the content of the buffer to be decreased as the processors catch up.

Although the present invention has been described hereinabove with reference to a specific embodiment, it will be apparent to a skilled person in the art that modifications lie within the spirit and scope of the present invention. Although in the embodiment described hereinabove, the acoustic model memories 23 were described as storing parts of the lexical data, it is also possible for more than one copy of the lexical data to be stored in different acoustic model memories 23. Thus, more than one lexical tree processor cluster 22 can have access to the same lexical data. This increases the flexibility and allows the controller to avoid any bottlenecks in processing which, for example, can occur if the highest pass scores all require processing by one or a number of lexical tree data structures residing in one of the acoustic model memories 23.

Although in the embodiments described hereinabove, the memory is partitioned in to a plurality of acoustic model memories 23, the present invention is not limited to the partitioning of the lexical data and a plurality of lexical tree processors can access a single memory containing lexical data.

The lexical data stored can comprise mono phone models of words. When the lexical data is partitioned, these can be distributed evenly across the memories. If there are 50 mono phones, there will be 50 lexical tree data structures having a unique initial model phone data node. Alternatively, the lexical tree data structures can comprise a context dependent n phone word models. If there are 50 mono phones, there are potentially 50^n lexical tree data structures. For tri-phones, this comprises 50^3 (i.e. 125,000). These are distributed so that trees with the same left context tri-phone, are eventually spread across memories when the lexical data is partitioned. Thus in this way the word end

where the left context tri-phone is fixed i.e. the same, the next possible lexical trees will inherently be distributed across the memories, thus ensuring efficient parallel processing of the next lexical trees.

Although the embodiment described hereinabove uses a single results memory accessible by all processors, the present invention encompasses a plurality of results memories. For example, a results memory can be associated with each processor to store the processing results and the controller can read the results in the plurality of results memories in order to control the processors.

Although in the embodiment described above the input speech parameters are described as being feature vectors, the present invention encompasses any type of parameterization of speech data as an input to the circuit. Such parameterization methods are well known in the art.

Although in the embodiment described above there is only one language model processor, more than one can be used and these can be used to operate in parallel for processing scores at word boundaries.

CLAIMS:

1. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

lexical memory means containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input means for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory means;

results memory means connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

control processor means for controlling said lexical tree processors to process lexical trees identified in said results memory means by performing parallel processing on a plurality of said lexical tree data structures.

2. A speech recognition circuit according to claim 1, wherein said lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

3. A speech recognition circuit according to claim 2, wherein said lexical memory means stores each lexical tree data structure as mono phone models of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

4. A speech recognition circuit according to any preceding claim, wherein each lexical tree processor is operative to compare the processed speech parameters with said lexical data or data derived from said lexical data to identify words as a word recognition event and to send information identifying the identified words to said results memory means as the processing results.

5. A speech recognition circuit according to claim 4, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

6. A speech recognition circuit according to claim 5, wherein each lexical tree processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and to output the modified processing result to said results memory means.

7. A speech recognition circuit according to any one of claims 4 to 6, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events.

8. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from said lexical data in said lexical memory means.

9. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to instruct a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory means.

10. A speech recognition circuit according to claim 9, wherein said history data includes an accumulated score for the recognition path.

11. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to pass a lexical tree identifier to a said lexical tree processor, said lexical tree processor is operative to read lexical data for the identified lexical tree from said lexical memory means and to process said speech parameters using said read lexical data.

12. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

13. A speech recognition circuit according to claim 12, wherein said lexical tree processors are operative to prune the processing of the lexical tree in dependence upon the determined scores.

14. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process more than one lexical tree at the same time.

15. A speech recognition circuit according to any one of claims 4 to 7, wherein said control processor means is operative to determine new lexical tree identifiers for storing in said results memory means for words identified in said results memory means.

16. A speech recognition circuit according to claim 15, wherein said control processor means is operative to prune the new lexical tree identifiers.

17. A speech recognition circuit according to claims 16, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

18. A speech recognition circuit according to any preceding claim, wherein said input means is arranged to receive said speech parameters as feature vectors.

19. A speech recognition circuit according to any one of claims 1 to 17, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

20. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are arranged in groups, said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said groups of lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

21. A speech recognition circuit according to any one of claims 1 to 19, wherein said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

22. A speech recognition circuit according to any preceding claim, wherein said lexical memory means stores said lexical tree data structures as Hidden Markov Models, and said lexical tree processors are operative to perform the Viterbi search algorithm using each lexical tree data structure.

23. A speech recognition circuit comprising:

an input buffer for receiving processed speech parameters;

lexical memory containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, the initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory;

results memory connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

a control processor for controlling said lexical tree processors to process lexical trees identified in said results memory by performing parallel processing on a plurality of said lexical tree data structures.

24. A speech recognition circuit according to claim 23, wherein said lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

25. A speech recognition circuit according to claim 24, wherein said lexical memory stores said lexical tree data structure as a mono phone model of words, and each lexical tree processor generates context dependant phone models of words dynamically for use in processing the speech parameters.

26. A speech recognition circuit according to any one of claims 23 to 25, wherein each lexical tree processor compares the processed speech parameters with the lexical data or data derived from the lexical data to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

27. A speech recognition circuit according to claim 26, including a language model processor for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

28. A speech recognition circuit according to claim 27, wherein each lexical tree processor receives said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and outputs the modified processing result to said results memory.

29. A speech recognition circuit according to any one of claims 26 to 28, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events.

30. A speech recognition circuit according to any one of claims 23 to 29, wherein said lexical tree processors process said speech parameters by comparing said speech parameters with said lexical data in said lexical memory.

31. A speech recognition circuit according to any one of claims 23 to 30, wherein said control processor instructs a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory.

32. A speech recognition circuit according to claim 31, wherein said history data includes an accumulated score for the recognition path.

33. A speech recognition circuit according to any one of claims 23 to 32, wherein said control processor passes a lexical tree identifier to a said lexical tree processor, said lexical tree processor reads lexical data for the identified lexical tree from said lexical memory and processes said speech parameters using said read lexical data.

34. A speech recognition circuit according to any one of claims 23 to 33, wherein said lexical tree processors determine and output scores in the processing results during the processing of said speech parameters.

35. A speech recognition circuit according to claim 34, wherein said lexical tree processors prune the processing of the lexical trees in dependence upon the determined scores.

36. A speech recognition circuit according to any one of claims 23 to 35, wherein said lexical tree processors process more than one lexical tree at the same time.

37. A speech recognition circuit according to any one of claims 26 to 29, wherein said control processor determines new lexical tree identifiers for storing in said results memory for words identified in said results memory.

38. A speech recognition circuit according to claim 37, wherein said control processor means is operative to prune the new lexical tree identifiers

39. A speech recognition circuit according to claims 38, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

40. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as feature vectors.

41. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as n phones, where n is an integer.

42. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical tree processors are arranged in groups, said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

43. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory is connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

44. A speech recognition circuit according to any one of claims 23 to 43, wherein said lexical memory stores said lexical tree data structures as Hidden Markov Models, and each lexical tree processor performs the Viterbi search algorithm using a respective said lexical tree data structure.

45. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

a plurality of lexical memory means containing in combination complete lexical data for word recognition, each lexical memory means containing part of said complete lexical data;

a plurality of processors connected in parallel to said input means for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory means;

control processor means for controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

receiving means for receiving the results of the processing of the speech parameters from said processors.

46. A speech recognition circuit according to claim 45, wherein each lexical memory means stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

47. A speech recognition circuit according to claim 46, wherein each lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

48. A speech recognition circuit according to claim 47, wherein each lexical memory means stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

49. A speech recognition circuit according to any one of claims 45 to 48, wherein each processor is operative to compare the processed speech parameters with the lexical data in a corresponding lexical memory means to identify words as a word recognition event and to send information identifying the identified words to said receiving means as the processing results.

50. A speech recognition circuit according to claim 49, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said processor.

51. A speech recognition circuit according to claim 50, wherein each processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said processor and to output the modified processing result to said receiving means.

52. A speech recognition circuit according to any one of claims 49 to 51, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events.

53. A speech recognition circuit according to any one of claims 45 to 52, wherein said processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memory means.

54. A speech recognition circuit according to any one of claims 46 to 48, wherein said control processor means is operative to instruct a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

55. A speech recognition circuit according to claim 54, wherein said history data includes an accumulated score for the recognition path.

56. A speech recognition circuit according to any one of claims 46 to 48, 54, or 55, wherein said control processor means is operative to pass a lexical tree identifier to a said processor, said processor is operative to read lexical data for the identified lexical tree from a corresponding said lexical memory means and to process said speech parameters using said read lexical data.

57. A speech recognition circuit according to any one of claims 45 to 56, wherein said processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

58. A speech recognition circuit according to claim 57, wherein said processors are operative to prune the processing of the speech parameters in dependence upon the determined scores.

59. A speech recognition circuit according to any one of claims 46 to 48, or 54 to 56, wherein said processors are operative to use lexical data corresponding to more than one lexical tree at the same time.

60. A speech recognition circuit according to any one of claims 49 to 52, wherein said control processor means is operative to identify lexical data in said lexical memory means required for processing next words at word recognition events, and to control said processors to process the speech parameters using the identified lexical data.

61. A speech recognition circuit according to claim 60, wherein said control processor means is operative to prune the identified lexical data.

62. A speech recognition circuit according to claims 61, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune the identified lexical data on the basis of the scores for the words.

63. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as feature vectors.

64. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

65. A speech recognition circuit according to any one of claims 45 to 64, wherein said control processor means is operative to control each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory means.

66. A speech recognition circuit according to any one of claims 45 to 65, wherein said receiving means comprises a memory for storing the results of processing.

67. A speech recognition circuit according to any one of claims 45 to 66, wherein said lexical memory means store said lexical data as Hidden Markov Models, and each processor is operative to performs the Viterbi search algorithm using a respective part of said lexical data.

68. A speech recognition circuit comprising:

an input buffer receiving processed speech parameters;

a plurality of lexical memories containing in combination complete lexical data for word recognition, each lexical memory containing part of said complete lexical data;

a plurality of processors connected in parallel to said input buffer for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory;

a control processor controlling each processor to process said speech parameters using partial lexical data read from a respective said lexical memory; and

a results memory storing the results of the processing of the speech parameters from said processors.

69. A speech recognition circuit according to claim 68, wherein each lexical memory stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

70. A speech recognition circuit according to claim 69, wherein each lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

71. A speech recognition circuit according to claim 70, wherein each lexical memory stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

72. A speech recognition circuit according to any one of claims 68 to 71, wherein each processor compares the processed speech parameters with the lexical data in a corresponding lexical memory to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

73. A speech recognition circuit according to claim 72, a including language model processor providing a language model output for modifying the processing results at a word recognition event by a said processor.

74. A speech recognition circuit according to claim 73, wherein each processor receives said language model output to modify the processing result at a word recognition event determined by said processor and outputs the modified processing result to said results memory.

75. A speech recognition circuit according to any one of claims 73 to 74, wherein each processor determines and outputs scores for words in the processing results at word recognition events.

76. A speech recognition circuit according to any one of claims 68 to 75, wherein said processors process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memories.

77. A speech recognition circuit according to any one of claims 69 to 71, wherein said control processor instructs a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

78. A speech recognition circuit according to claim 77, wherein said history data includes an accumulated score for the recognition path.

79. A speech recognition circuit according to any one of claims 69 to 71, 77, or 78, wherein said control processor passes a lexical tree identifier to a said processor, said processor reads lexical data for the identified lexical tree from a corresponding said lexical memory and processes said speech parameters using said read lexical data.

80. A speech recognition circuit according to any one of claims 68 to 79, wherein said processors determine and output scores in the processing results during the processing of said speech parameters.

81. A speech recognition circuit according to claim 80, wherein said processors prune the processing of the speech parameters in dependence upon the determined scores.

82. A speech recognition circuit according to any one of claims 69 to 71, or 77 to 79, wherein said processors use lexical data corresponding to more than one lexical tree at the same time.

83. A speech recognition circuit according to any one of claims 72 to 75, wherein said control processor identifies lexical data in said lexical memories required for processing next words at word recognition events, and controls said processors to process the speech parameters using the identified lexical data.

84. A speech recognition circuit according to claim 83, wherein said control processor prunes the identified lexical data.

85. A speech recognition circuit according to claims 84, wherein each processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes the identified lexical data on the basis of the scores for the words.

86. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as feature vectors.

87. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as n phones, where n is an integer.

88. A speech recognition circuit according to any one of claims 68 to 88, wherein said control processor controls each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory.

89. A speech recognition circuit according to any one of claims 68 to 89, wherein said lexical memories store said lexical data as Hidden Markov Models, and each processor performs the Viterbi search algorithm using a respective part of said lexical data.

90. A speech recognition method comprising:

receiving processed speech parameters;

storing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;;

passing the speech parameters in parallel to a plurality of lexical tree processors to perform parallel lexical tree processing for word recognition by accessing said lexical data;

storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

controlling said lexical tree processors to process lexical trees identified by the stored identifiers by performing parallel processing on a plurality of said lexical tree data structures.

91. A speech recognition method comprising:

receiving processed speech parameters;

storing complete lexical data for word recognition in a plurality of lexical memory means, each lexical memory means containing part of said complete lexical data;

processing the speech parameters in parallel using a plurality of processors connected in parallel, said processors being arranged in groups of processors, each group of processors being connected to a said lexical memory means;

controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

storing the results of the processing of the speech parameters from said processors.

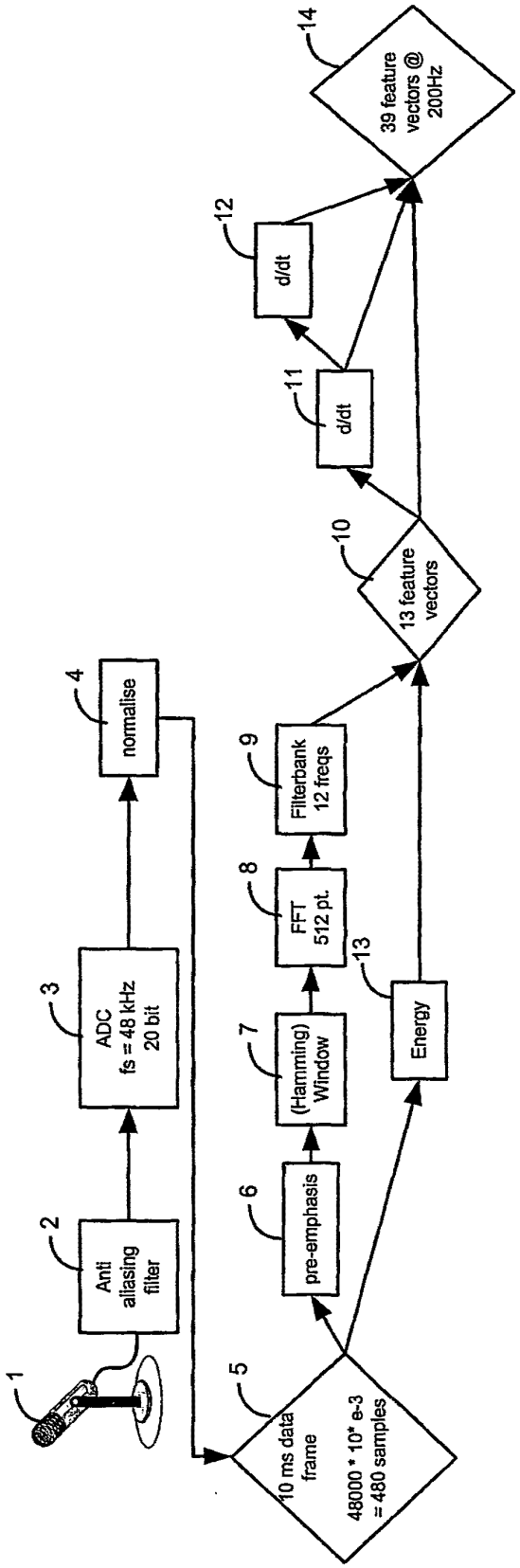

1
2
Anti aliasing filter
3
ADC fs = 48 kHz 20 bit
4
normalise
5
10 ms data frame 48000 * 10* e-3 = 480 samples
6
pre-emphasis
7
(Hamming) Window
8
FFT 512 pt.
9
Filterbank 12 freqs
10
13 feature vectors
11
d/dt
12
d/dt
13
Energy
14
39 feature vectors @ 200Hz

Fig 1

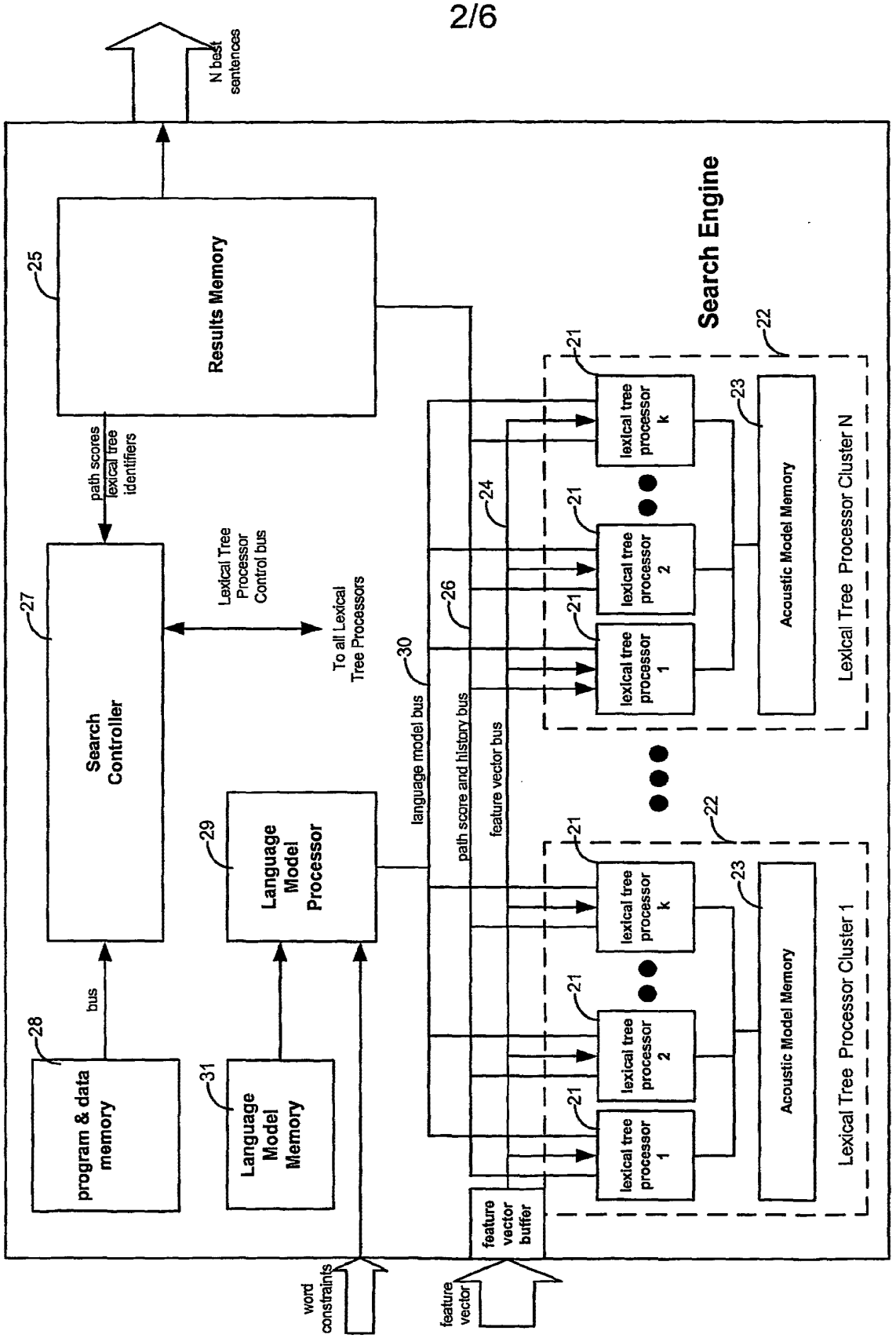

N best sentences
Search Engine
Results Memory
25
path scores lexical tree identifiers
Search Controller
27
Lexical Tree Processor Control bus
To all Lexical Tree Processors
Language Model Processor
29
language model bus
30
path score and history bus
feature vector bus
26
24
program & data memory
28
bus
Language Model Memory
31
word constraints
feature vector
feature vector buffer
lexical tree processor k
lexical tree processor 2
lexical tree processor 1
21
22
Acoustic Model Memory
23
Lexical Tree Processor Cluster N
Lexical Tree Processor Cluster 1

Fig 2

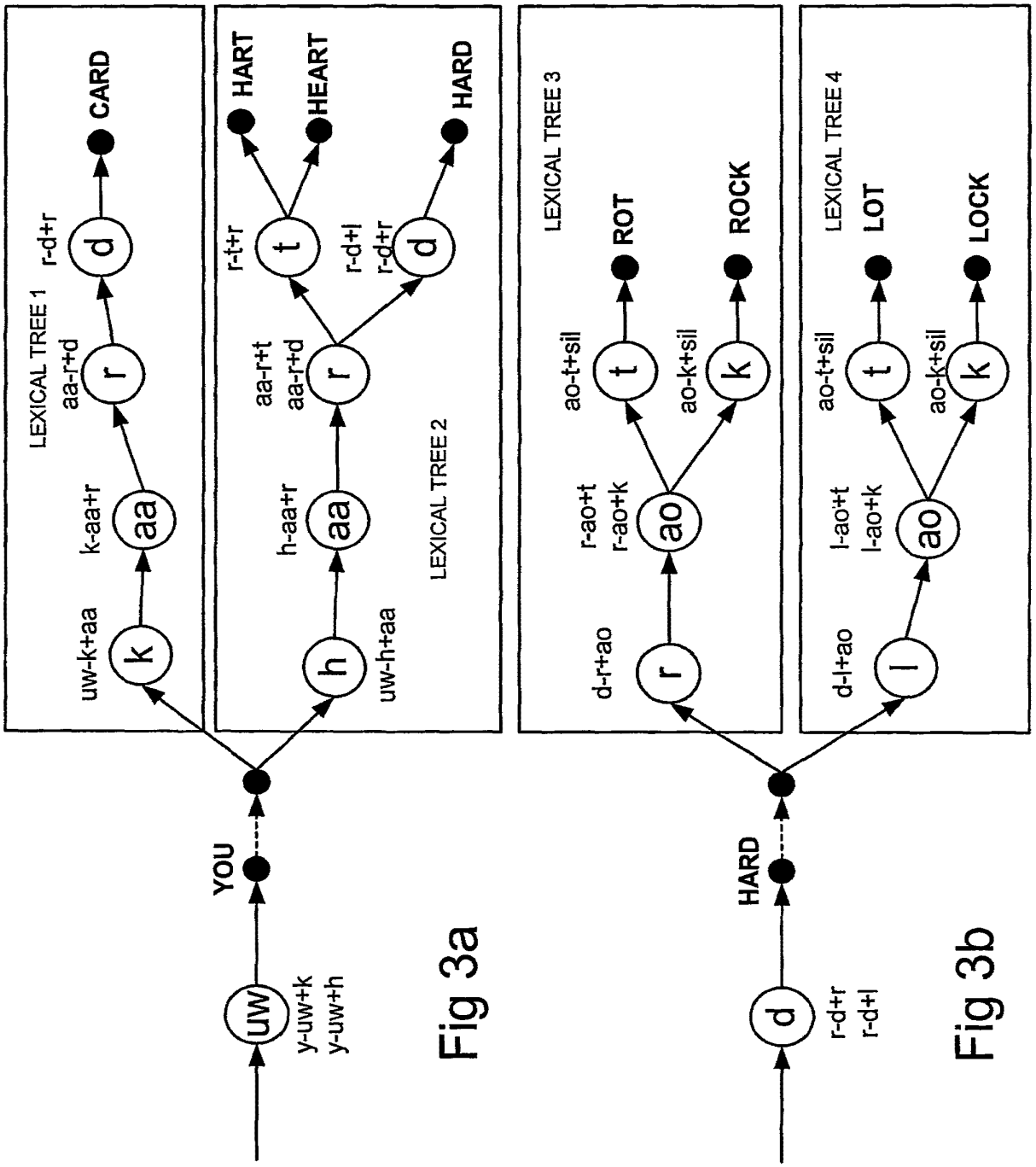

LEXICAL TREE 1
uw-k+aa
k-aa+r
aa-r+d
r-d+r
k
aa
r
d
CARD
LEXICAL TREE 2
uw-h+aa
h-aa+r
aa-r+t
aa-r+d
r-t+r
r-d+l
r-d+r
h
aa
r
t
d
HART
HEART
HARD
you
y-uw+k
y-uw+h
uw

Fig 3a

LEXICAL TREE 3
d-r+ao
r-ao+t
r-ao+k
ao-t+sil
ao-k+sil
r
ao
t
k
ROT
ROCK
LEXICAL TREE 4
d-l+ao
l-ao+t
l-ao+k
ao-t+sil
ao-k+sil
l
ao
t
k
LOT
LOCK
HARD
r-d+r
r-d+l
d

Fig 3b

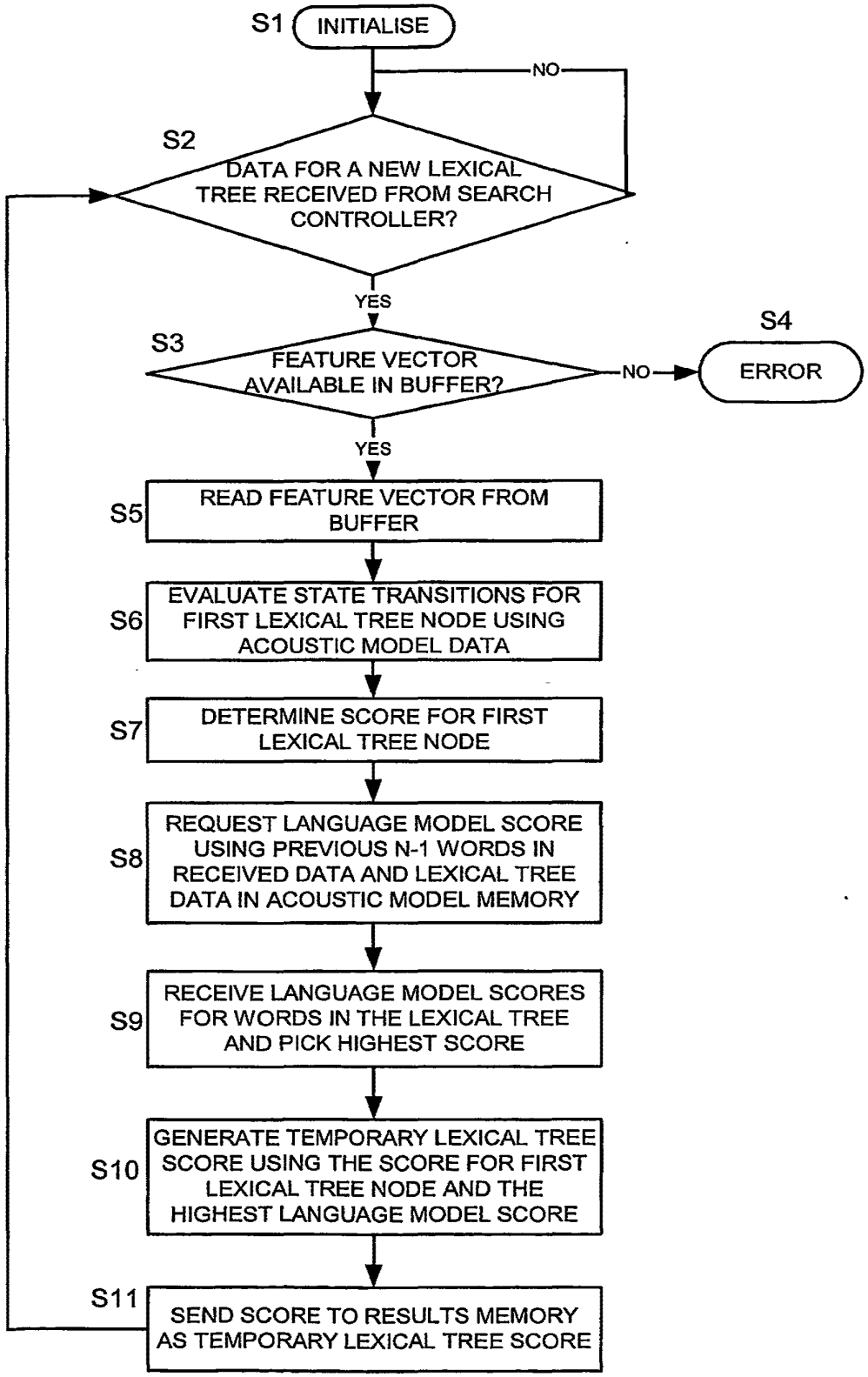

S1
INITIALISE
NO
S2
DATA FOR A NEW LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S3
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S4
ERROR
YES
S5
READ FEATURE VECTOR FROM BUFFER
S6
EVALUATE STATE TRANSITIONS FOR FIRST LEXICAL TREE NODE USING ACOUSTIC MODEL DATA
S7
DETERMINE SCORE FOR FIRST LEXICAL TREE NODE
S8
REQUEST LANGUAGE MODEL SCORE USING PREVIOUS N-1 WORDS IN RECEIVED DATA AND LEXICAL TREE DATA IN ACOUSTIC MODEL MEMORY
S9
RECEIVE LANGUAGE MODEL SCORES FOR WORDS IN THE LEXICAL TREE AND PICK HIGHEST SCORE
S10
GENERATE TEMPORARY LEXICAL TREE SCORE USING THE SCORE FOR FIRST LEXICAL TREE NODE AND THE HIGHEST LANGUAGE MODEL SCORE
S11
SEND SCORE TO RESULTS MEMORY AS TEMPORARY LEXICAL TREE SCORE

Fig 4

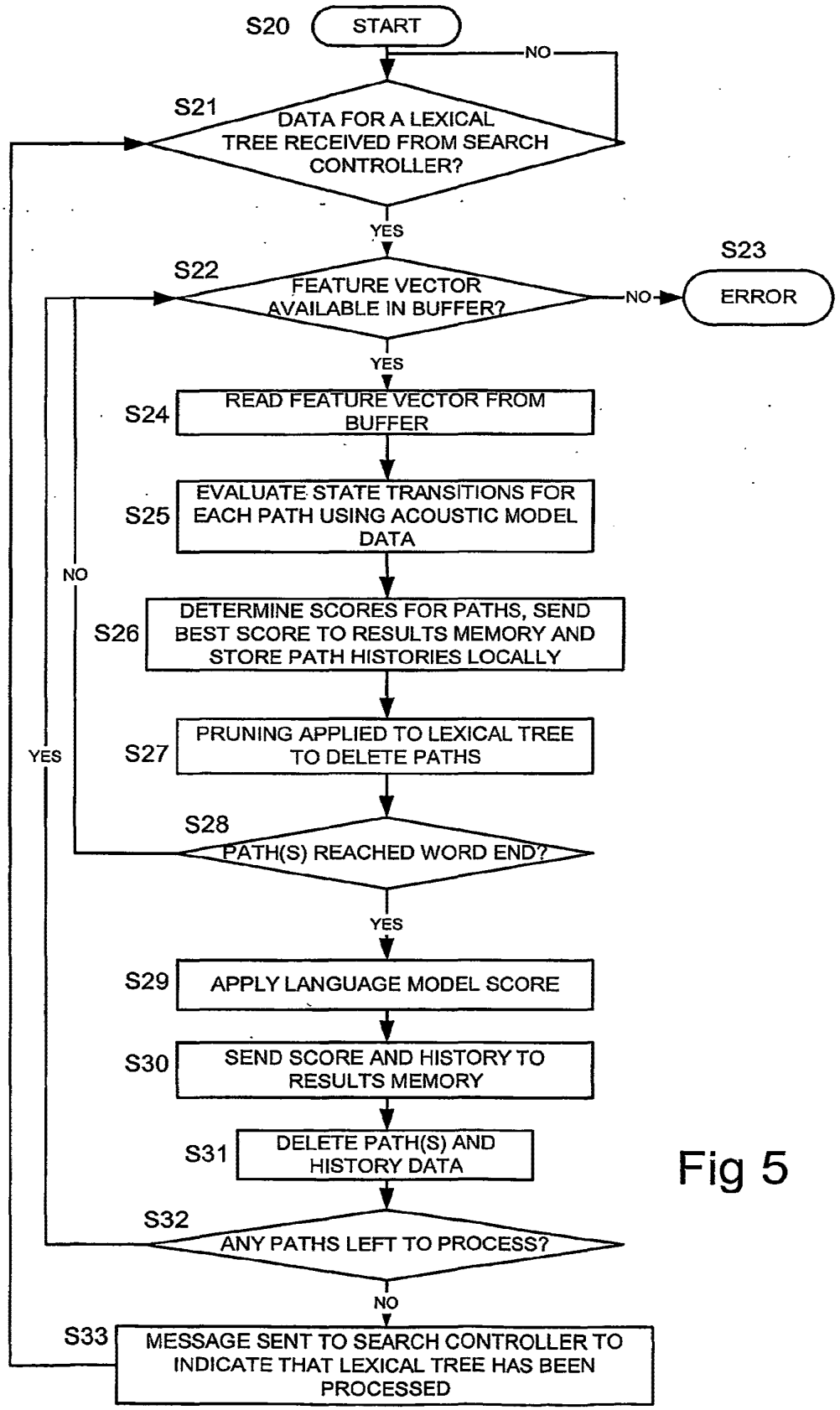

S20
START
NO
S21
DATA FOR A LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S22
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S23
ERROR
YES
S24
READ FEATURE VECTOR FROM BUFFER
S25
EVALUATE STATE TRANSITIONS FOR EACH PATH USING ACOUSTIC MODEL DATA
NO
S26
DETERMINE SCORES FOR PATHS, SEND BEST SCORE TO RESULTS MEMORY AND STORE PATH HISTORIES LOCALLY
S27
PRUNING APPLIED TO LEXICAL TREE TO DELETE PATHS
YES
S28
PATH(S) REACHED WORD END?
YES
S29
APPLY LANGUAGE MODEL SCORE
S30
SEND SCORE AND HISTORY TO RESULTS MEMORY
S31
DELETE PATH(S) AND HISTORY DATA
S32
ANY PATHS LEFT TO PROCESS?
NO
S33
MESSAGE SENT TO SEARCH CONTROLLER TO INDICATE THAT LEXICAL TREE HAS BEEN PROCESSED

Fig 5

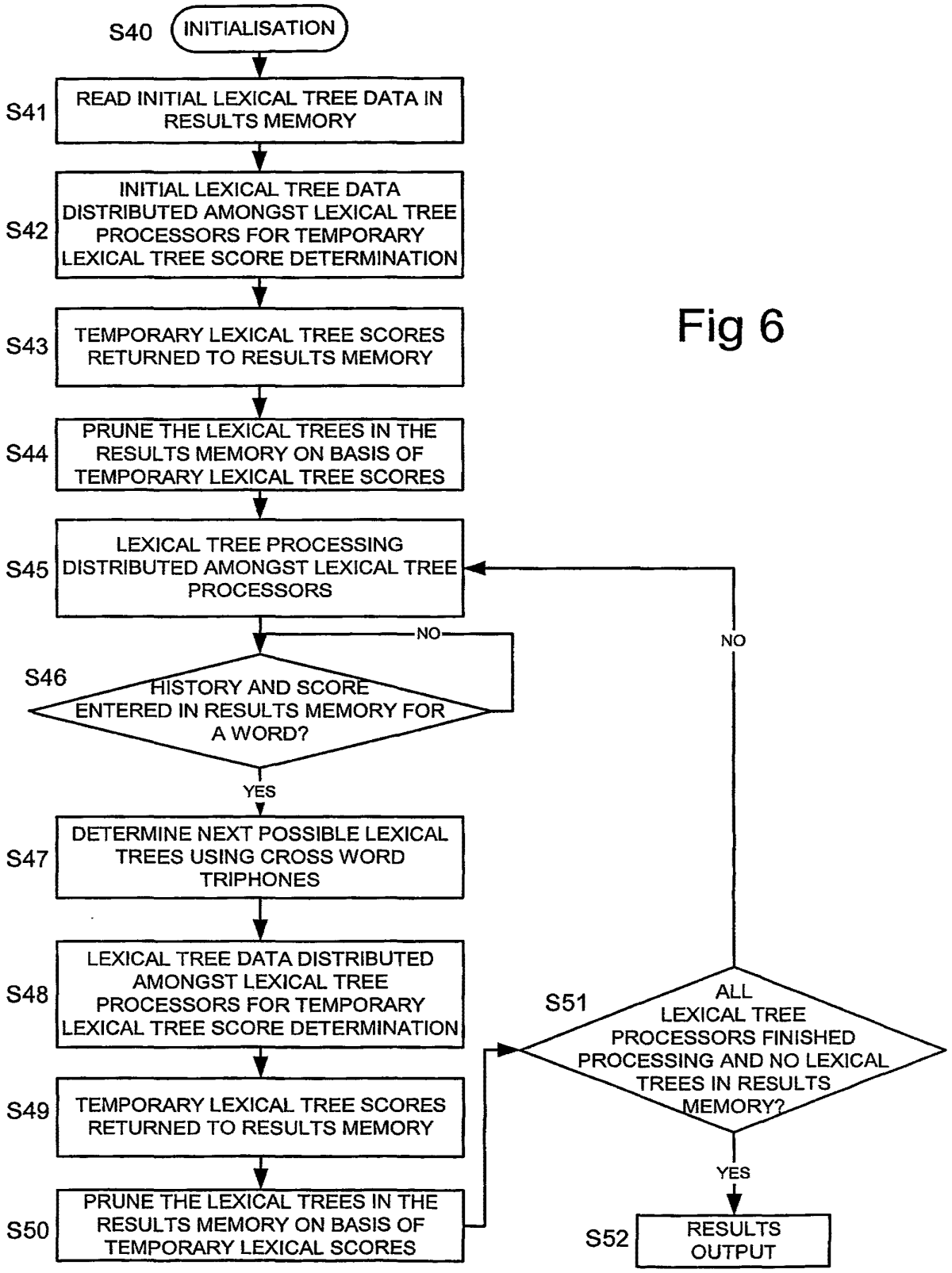

S40
INITIALISATION
S41
READ INITIAL LEXICAL TREE DATA IN RESULTS MEMORY
S42
INITIAL LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S43
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S44
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL TREE SCORES
S45
LEXICAL TREE PROCESSING DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS
S46
HISTORY AND SCORE ENTERED IN RESULTS MEMORY FOR A WORD?
NO
YES
S47
DETERMINE NEXT POSSIBLE LEXICAL TREES USING CROSS WORD TRIPHONES
S48
LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S49
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S50
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL SCORES
S51
ALL LEXICAL TREE PROCESSORS FINISHED PROCESSING AND NO LEXICAL TREES IN RESULTS MEMORY?
NO
YES
S52
RESULTS OUTPUT

Fig 6

Rec'd PCT/PTC 04 AUG 2004

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2003 (14.08.2003)

PCT

(10) International Publication Number
WO 03/067572 A3

(51) International Patent Classification[7]: G10L 15/28, 15/18

(21) International Application Number: PCT/GB03/00459

(22) International Filing Date: 4 February 2003 (04.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0202546.8 4 February 2002 (04.02.2002) GB

(71) Applicant *(for all designated States except US)*: **ZENTIAN LIMITED** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **CATCHPOLE, Mark** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(74) Agent: **COLLINS, John, David**; Marks & Clerk, 57-60 Lincolns Inn Fields, London WC2A 3LS (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

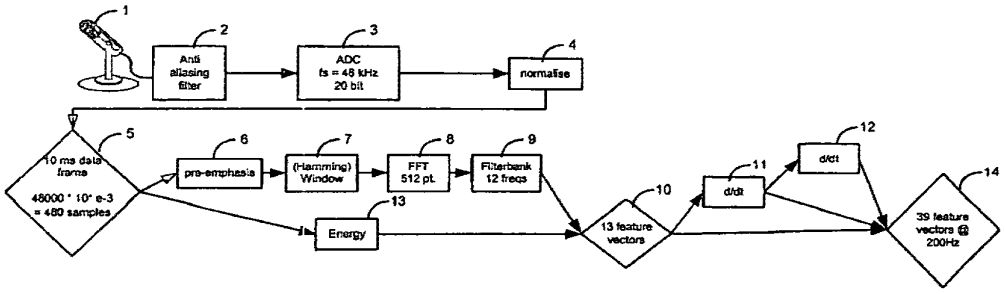
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
20 November 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS



(57) Abstract: A speech recognition circuit comprises an input buffer for receiving processed speech parameters. A lexical memory contains lexical data for word recognition. The lexical data comprises a plurality of lexical tree data structures. Each lexical tree data structure comprises a model of words having common prefix components. An initial component of each lexical tree structure is unique. A plurality of lexical tree processors are connected in parallel to the input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory. A results memory is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory by performing parallel processing on a plurality of said lexical tree data structures.

WO 03/067572 A3

INTERNATIONAL SEARCH REPORT

PCT/GB 03/00459

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G10L15/28 G10L15/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, INSPEC, IBM-TDB, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SANG-HWA CHUNG ET AL: "A parallel phoneme recognition algorithm based on continuous hidden Markov model" PROCEEDINGS 13TH INTERNATIONAL PARALLEL PROCESSING SYMPOSIUM AND 10TH SYMPOSIUM ON PARALLEL AND DISTRIBUTED PROCESSING. IPPS/SPDP 1999, PROCEEDINGS OF 13TH INTERNATIONAL PARALLEL PROCESSING SYMPOSIUM AND 10TH SYMPOSIUM ON PARALLEL AND DISTRIBUTED PRO, pages 453-457, XP010330371 1999, Los Alamitos, CA, USA, IEEE Comput. Soc, USA ISBN: 0-7695-0143-5 abstract	1,23,45, 68,90,91
A	page 455, left-hand column, paragraph 2 -page 456, left-hand column, paragraph 1	2-22, 24-44, 46-67, 69-89

-/--

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

° Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 August 2003	14/08/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Greiser, N

Form PCT/ISA/210 (second sheet) (July 1992)

IPR2023-00033
Apple EX1002 Page 155

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	CHUNG S-H ET AL: "A PARALLEL COMPUTATIONAL MODEL FOR INTEGRATED SPEECH AND NATURAL LANGUAGE UNDERSTANDING" IEEE TRANSACTIONS ON COMPUTERS, IEEE INC. NEW YORK, US, vol. 42, no. 10, 1 October 1993 (1993-10-01), pages 1171-1183, XP000418436 ISSN: 0018-9340 abstract	1,23,45, 68,90,91
A	DESHMUKH N ET AL: "Hierarchical search for large-vocabulary conversational speech recognition: working toward a solution to the decoding problem" IEEE SIGNAL PROCESSING MAGAZINE, SEPT. 1999, IEEE, USA, vol. 16, no. 5, pages 84-107, XP002249037 ISSN: 1053-5888 abstract page 99, left-hand column, paragraph 4 -page 100, right-hand column, paragraph 1; figures 17,18 page 86, left-hand column, paragraph 2 -page 87, left-hand column, paragraph 1; figures 2-4	1,23,45, 68,90,91

IPR2023-00033
Apple EX1002 Page 156

PATENT APPLICATION FEE DETERMINATION RECORD
Effective October 1, 2003

Application or Docket Number: 10/503463

CLAIMS AS FILED - PART I

	(Column 1)	(Column 2)
TOTAL CLAIMS		
FOR	NUMBER FILED	NUMBER EXTRA
TOTAL CHARGEABLE CLAIMS	91 minus 20=	* 71
INDEPENDENT CLAIMS	6 minus 3 =	* 3
MULTIPLE DEPENDENT CLAIM PRESENT		☐

* If the difference in column 1 is less than zero, enter "0" in column 2

SMALL ENTITY TYPE ☐		OR	OTHER THAN SMALL ENTITY	
RATE	FEE		RATE	FEE
BASIC FEE	$460	OR	BASIC FEE	
XS 9=	$39	OR	XS16=	
X43=	$129	OR	X86=	
+145=		OR	+290=	
TOTAL	$628	OR	TOTAL	

CLAIMS AS AMENDED - PART II

8-24-04

AMENDMENT A	(Column 1)		(Column 2)	(Column 3)
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA
Total	* 91	Minus	** 91	= 0
Independent	* 10	Minus	*** 10	= 0
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM				☐

SMALL ENTITY		OR	OTHER THAN SMALL ENTITY	
RATE	ADDITIONAL FEE		RATE	ADDITIONAL FEE
XS 9=	0	OR	XS18=	
X43=	0	OR	X86=	
+145=	0	OR	+290=	
TOTAL ADDIT. FEE	0	OR	TOTAL ADDIT. FEE	

AMENDMENT B	(Column 1)		(Column 2)	(Column 3)
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM				☐

RATE	ADDITIONAL FEE		RATE	ADDITIONAL FEE
XS 9=		OR	XS18=	
X43=		OR	X86=	
+145=		OR	+290=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

AMENDMENT C	(Column 1)		(Column 2)	(Column 3)
	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	PRESENT EXTRA
Total	*	Minus	**	=
Independent	*	Minus	***	=
FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM				☐

RATE	ADDITIONAL FEE		RATE	ADDITIONAL FEE
XS 9=		OR	XS18=	
X43=		OR	X86=	
+145=		OR	+290=	
TOTAL ADDIT. FEE		OR	TOTAL ADDIT. FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20."
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3."
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

IPR2023-00033
Apple EX1002 Page 157

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

U.S. APPLICATION NUMBER NO.	FIRST NAMED APPLICANT	ATTY. DOCKET NO.
10/503,463	Mark Catchpole	M0025.0311/P311

INTERNATIONAL APPLICATION NO.	
PCT/GB03/00459	
I.A. FILING DATE	PRIORITY DATE
02/04/2003	02/04/2002

24998
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street, NW
Washington, DC 20037

CONFIRMATION NO. 8875

371 FORMALITIES LETTER

OC000000015454716

Date Mailed: 04/01/2005

NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 U.S.C. 371 IN THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

The following items have been submitted by the applicant or the IB to the United States Patent and Trademark Office as a Designated / Elected Office (37 CFR 1.495).

- Indication of Small Entity Status
- Copy of the International Application filed on 08/04/2004
- Copy of the International Search Report filed on 08/04/2004
- Preliminary Amendments filed on 08/04/2004
- Information Disclosure Statements filed on 08/04/2004
- Request for Immediate Examination filed on 08/04/2004
- U.S. Basic National Fees filed on 08/04/2004
- Priority Documents filed on 08/04/2004

The following items **MUST** be furnished within the period set forth below in order to complete the requirements for acceptance under 35 U.S.C. 371:

- Oath or declaration of the inventors, in compliance with 37 CFR 1.497(a) and (b), identifying the application by the International application number and international filing date.

- **$65** Surcharge for providing the oath or declaration later than 30 months from the priority date (37 CFR 1.492(e)) is required.

SUMMARY OF FEES DUE:

Total additional fees required for this application is **$65** for a Small Entity:

- **$65** Late oath or declaration Surcharge.

ALL OF THE ITEMS SET FORTH ABOVE MUST BE SUBMITTED WITHIN TWO (2) MONTHS FROM THE

DATE OF THIS NOTICE OR BY 32 MONTHS FROM THE PRIORITY DATE FOR THE APPLICATION, WHICHEVER IS LATER. FAILURE TO PROPERLY RESPOND WILL RESULT IN ABANDONMENT.

The time period set above may be extended by filing a petition and fee for extension of time under the provisions of 37 CFR 1.136(a).

Applicant is reminded that any communications to the United States Patent and Trademark Office must be mailed to the address given in the heading and include the U.S. application no. shown above (37 CFR 1.5)

*A copy of this notice **MUST** be returned with the response.*

INDIA L EVANS

Telephone: (703) 308-9140 EXT 212

PART 2 - OFFICE COPY

U.S. APPLICATION NUMBER NO.	INTERNATIONAL APPLICATION NO.	ATTY. DOCKET NO.
10/503,463	PCT/GB03/00459	M0025.0311/P311

FORM PCT/DO/EO/905 (371 Formalities Notice)

PTO/SB/21 (09-04)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	10/503,463-Conf. #8875
Filing Date	August 4, 2004
First Named Inventor	Mark Catchpole
Art Unit	N/A
Examiner Name	Not Yet Assigned
Total Number of Pages in This Submission	
Attorney Docket Number	M0025.0311/P311

ENCLOSURES (*Check all that apply*)

- [x] Fee Transmittal Form
- [ ] Fee Attached
- [ ] Amendment/Reply
- [ ] After Final
- [ ] Affidavits/declaration(s)
- [ ] Extension of Time Request
- [ ] Express Abandonment Request
- [ ] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [x] Reply to Missing Parts/ Incomplete Application
- [ ] Reply to Missing Parts under 37 CFR 1.52 or 1.53
- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) _____
- [ ] Landscape Table on CD
- [ ] After Allowance Communication to TC
- [ ] Appeal Communication to Board of Appeals and Interferences
- [ ] Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- [ ] Proprietary Information
- [ ] Status Letter
- [x] Other Enclosure(s) (please Identify below):

Combined Declaration and Power of Attorney
Assignment and Recordation Form Cover Sheet
Copy of Notice

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP		
Signature			
Printed name	Stephen A. Soffen		
Date	May 24, 2005	Reg. No.	31,063

#5

OIPE JC25
MAY 24 2005
PATENT & TRADEMARK OFFICE

JC10 Rec'd PCT/PTO 24 MAY 2005

PTO/SB/17 (12-04v2)
Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Effective on 12/08/2004.
Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818).

FEE TRANSMITTAL
For FY 2005

[X] Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT ($) 105.00

Complete if Known	
Application Number	10/503,463-Conf. #8875
Filing Date	August 4, 2004
First Named Inventor	Mark Catchpole
Examiner Name	Not Yet Assigned
Art Unit	N/A
Attorney Docket No.	M0025.0311/P311

METHOD OF PAYMENT (check all that apply)

[ ] Check [X] Credit Card [ ] Money Order [ ] None [ ] Other (please identify): _____

[X] Deposit Account Deposit Account Number: 04-1073 Deposit Account Name: Dickstein Shapiro Morin & Oshinsky LLP

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

[X] Charge fee(s) indicated below
[ ] Charge fee(s) indicated below, **except for the filing fee**
[X] Charge any additional fee(s) or underpayment of fee(s) under 37 CFR 1.16 and 1.17
[X] Credit any overpayments

FEE CALCULATION

1. BASIC FILING, SEARCH, AND EXAMINATION FEES

	FILING FEES		SEARCH FEES		EXAMINATION FEES		
		Small Entity		Small Entity		Small Entity	
Application Type	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fees Paid ($)
Utility	300	150	500	250	200	100	
Design	200	100	100	50	130	65	
Plant	200	100	300	150	160	80	
Reissue	300	150	500	250	600	300	
Provisional	200	100	0	0	0	0	

2. EXCESS CLAIM FEES

Fee Description	Fee ($)	Small Entity Fee ($)
Each claim over 20 (including Reissues)	50	25
Each independent claim over 3 (including Reissues)	200	100
Multiple dependent claims	360	180

Total Claims		Extra Claims	Fee ($)		Fee Paid ($)
_____	- =	_____	x _____	=	_____

Multiple Dependent Claims	
Fee ($)	Fee Paid ($)
_____	_____

Indep. Claims		Extra Claims	Fee ($)		Fee Paid ($)
_____	- =	_____	x _____	=	_____

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper (excluding electronically filed sequence or computer listings under 37 CFR 1.52(e)), the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets		Extra Sheets		Number of each additional 50 or fraction thereof		Fee ($)		Fee Paid ($)
_____	- 100 =	_____	/50	_____	(round up to a whole number) x	_____	=	_____

4. OTHER FEE(S)

	Fees Paid ($)
Non-English Specification, $130 fee (no small entity discount)	
Other (e.g., late filing surcharge): 2051 Surcharge-Late oath or declaration	65.00
8021 Recording each patent assignment, agreement or ...	40.00

SUBMITTED BY

Signature	Stephen A. Soffen	Registration No. (Attorney/Agent)	31,063	Telephone	(202) 828-4879
Name (Print/Type)	Stephen A. Soffen			Date	May 24, 2005

05/27/2005 GFREY1 00000075 10503463
01 FC:2617 65.00 OP

IPR2023-00033
Apple EX1002 Page 161

Docket No.: M0025.0311/P311
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Mark Catchpole

Application No.: 10/503,463

Filed: August 4, 2004

For: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

Confirmation No.: 8875

Art Unit: N/A

Examiner: Not Yet Assigned

RESPONSE TO NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 U.S.C. 371 IN THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

MS Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Notification of Missing Requirements under 35 U.S.C. 371 in the United States Designated/Elected Office (DO/EO/US) – Filing Date Granted mailed April 1, 2005, Applicant respectfully submits a Combined Declaration and Power of Attorney, an Assignment and Recordation Form Cover Sheet, and Part 2 Copy of Notice.

Please charge our Credit Card in the amount of $105.00 covering the fee set forth in 37 CFR 1.16(f) and 1.21(h). Credit Card Payment Form SB-2038, with a signature from an authorized cardholder, is enclosed. The Director is hereby authorized to charge any deficiency in the fees filed, asserted to be filed or which should have been filed herewith (or with any paper hereafter filed in this application by this firm) to our Deposit

IPR2023-00033
Apple EX1002 Page 162

Account No. 04-1073, under Order No. M0025.0311/P311. A duplicate copy of this paper is enclosed.

Dated: May 24, 2005

Respectfully submitted,

By _____

Stephen A. Soffen
Registration No.: 31,063
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street NW
Washington, DC 20037-1526
(202) 785-9700
Attorney for Applicant

IPR2023-00033
Apple EX1002 Page 163

Rec'd PCT/PTO 24 MAY 2005

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

U.S. APPLICATION NUMBER NO.	FIRST NAMED APPLICANT	ATTY. DOCKET NO.
10/503,463	Mark Catchpole	M0025.0311/P311

INTERNATIONAL APPLICATION NO.	
PCT/GB03/00459	
I.A. FILING DATE	PRIORITY DATE
02/04/2003	02/04/2002

24998
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street, NW
Washington, DC 20037

CONFIRMATION NO. 8875
371 FORMALITIES LETTER

OC000000015454716

Date Mailed: 04/01/2005

NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 U.S.C. 371 IN THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

The following items have been submitted by the applicant or the IB to the United States Patent and Trademark Office as a Designated / Elected Office (37 CFR 1.495).

- Indication of Small Entity Status
- Copy of the International Application filed on 08/04/2004
- Copy of the International Search Report filed on 08/04/2004
- Preliminary Amendments filed on 08/04/2004
- Information Disclosure Statements filed on 08/04/2004
- Request for Immediate Examination filed on 08/04/2004
- U.S. Basic National Fees filed on 08/04/2004
- Priority Documents filed on 08/04/2004

RECEIVED
APR - 8 2005
M.S. 4/8/05
SAS/ Formalities
TW 4/11/05

The following items **MUST** be furnished within the period set forth below in order to complete the requirements for acceptance under 35 U.S.C. 371:

- Oath or declaration of the inventors, in compliance with 37 CFR 1.497(a) and (b), identifying the application by the International application number and international filing date.

- **$65** Surcharge for providing the oath or declaration later than 30 months from the priority date (37 CFR 1.492(e)) is required.

SUMMARY OF FEES DUE:

Total additional fees required for this application is **$65** for a Small Entity:

- **$65** Late oath or declaration Surcharge.

ALL OF THE ITEMS SET FORTH ABOVE MUST BE SUBMITTED WITHIN TWO (2) MONTHS FROM THE

DATE OF THIS NOTICE OR BY 32 MONTHS FROM THE PRIORITY DATE FOR THE APPLICATION, WHICHEVER IS LATER. FAILURE TO PROPERLY RESPOND WILL RESULT IN ABANDONMENT.

The time period set above may be extended by filing a petition and fee for extension of time under the provisions of 37 CFR 1.136(a).

Applicant is reminded that any communications to the United States Patent and Trademark Office must be mailed to the address given in the heading and include the U.S. application no. shown above (37 CFR 1.5)

*A copy of this notice **MUST** be returned with the response.*

INDIA L EVANS

Telephone: (703) 308-9140 EXT 212

PART 1 - ATTORNEY/APPLICANT COPY

U.S. APPLICATION NUMBER NO.	INTERNATIONAL APPLICATION NO.	ATTY. DOCKET NO.
10/503,463	PCT/GB03/00459	M0025.0311/P311

FORM PCT/DO/EO/905 (371 Formalities Notice)

PTO/SB/21 (09-04)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	10/503,463-Conf. #8875
Filing Date	August 4, 2004
First Named Inventor	Mark Catchpole
Art Unit	N/A
Examiner Name	Not Yet Assigned
Total Number of Pages in This Submission	
Attorney Docket Number	M0025.0311/P311

ENCLOSURES (*Check all that apply*)

- [x] Fee Transmittal Form
- [ ] Fee Attached
- [ ] Amendment/Reply
- [ ] After Final
- [ ] Affidavits/declaration(s)
- [ ] Extension of Time Request
- [ ] Express Abandonment Request
- [ ] Information Disclosure Statement
- [ ] Certified Copy of Priority Document(s)
- [x] Reply to Missing Parts/ Incomplete Application
- [ ] Reply to Missing Parts under 37 CFR 1.52 or 1.53
- [ ] Drawing(s)
- [ ] Licensing-related Papers
- [ ] Petition
- [ ] Petition to Convert to a Provisional Application
- [ ] Power of Attorney, Revocation Change of Correspondence Address
- [ ] Terminal Disclaimer
- [ ] Request for Refund
- [ ] CD, Number of CD(s) _____
- [ ] Landscape Table on CD
- [ ] After Allowance Communication to TC
- [ ] Appeal Communication to Board of Appeals and Interferences
- [ ] Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- [ ] Proprietary Information
- [ ] Status Letter
- [x] Other Enclosure(s) (please Identify below):

Combined Declaration and Power of Attorney
Assignment and Recordation Form Cover Sheet
Copy of Notice

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP		
Signature			
Printed name	Stephen A. Soffen		
Date	May 24, 2005	Reg. No.	31,063

#5

OIPE JC25 MAY 24 2005 PATENT & TRADEMARK OFFICE

JC10 Rec'd PCT/PTO 24 MAY 2005

PTO/SB/17 (12-04v2)
Approved for use through 7/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Effective on 12/08/2004.
Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818).

FEE TRANSMITTAL
For FY 2005

[X] Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT ($) 105.00

Complete if Known	
Application Number	10/503,463-Conf. #8875
Filing Date	August 4, 2004
First Named Inventor	Mark Catchpole
Examiner Name	Not Yet Assigned
Art Unit	N/A
Attorney Docket No.	M0025.0311/P311

METHOD OF PAYMENT (check all that apply)

[ ] Check [X] Credit Card [ ] Money Order [ ] None [ ] Other (please identify): _____

[X] Deposit Account Deposit Account Number: 04-1073 Deposit Account Name: Dickstein Shapiro Morin & Oshinsky LLP

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

[X] Charge fee(s) indicated below
[ ] Charge fee(s) indicated below, **except for the filing fee**
[X] Charge any additional fee(s) or underpayment of fee(s) under 37 CFR 1.16 and 1.17
[X] Credit any overpayments

FEE CALCULATION

1. BASIC FILING, SEARCH, AND EXAMINATION FEES

	FILING FEES		SEARCH FEES		EXAMINATION FEES		
		Small Entity		Small Entity		Small Entity	
Application Type	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fee ($)	Fees Paid ($)
Utility	300	150	500	250	200	100	
Design	200	100	100	50	130	65	
Plant	200	100	300	150	160	80	
Reissue	300	150	500	250	600	300	
Provisional	200	100	0	0	0	0	

2. EXCESS CLAIM FEES

Fee Description	Fee ($)	Small Entity Fee ($)
Each claim over 20 (including Reissues)	50	25
Each independent claim over 3 (including Reissues)	200	100
Multiple dependent claims	360	180

Total Claims		Extra Claims	Fee ($)		Fee Paid ($)
_____	- =	_____	x _____	=	_____

Multiple Dependent Claims	
Fee ($)	Fee Paid ($)
_____	_____

Indep. Claims		Extra Claims	Fee ($)		Fee Paid ($)
_____	- =	_____	x _____	=	_____

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper (excluding electronically filed sequence or computer listings under 37 CFR 1.52(e)), the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets		Extra Sheets		Number of each additional 50 or fraction thereof		Fee ($)		Fee Paid ($)
_____	- 100 =	_____	/50	_____	(round up to a whole number) x	_____	=	_____

4. OTHER FEE(S)

	Fees Paid ($)
Non-English Specification, $130 fee (no small entity discount)	
Other (e.g., late filing surcharge): 2051 Surcharge-Late oath or declaration	65.00
8021 Recording each patent assignment, agreement or ...	40.00

SUBMITTED BY

Signature	Stephen A. Soffen	Registration No. (Attorney/Agent)	31,063	Telephone	(202) 828-4879
Name (Print/Type)	Stephen A. Soffen			Date	May 24, 2005

05/27/2005 GFREY1 00000075 10503463
01 FC:2617 65.00 OP

IPR2023-00033
Apple EX1002 Page 167

Docket No.: M0025.0311/P311
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Mark Catchpole

Application No.: 10/503,463

Filed: August 4, 2004

For: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

Confirmation No.: 8875

Art Unit: N/A

Examiner: Not Yet Assigned

RESPONSE TO NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 U.S.C. 371 IN THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

MS Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Notification of Missing Requirements under 35 U.S.C. 371 in the United States Designated/Elected Office (DO/EO/US) – Filing Date Granted mailed April 1, 2005, Applicant respectfully submits a Combined Declaration and Power of Attorney, an Assignment and Recordation Form Cover Sheet, and Part 2 Copy of Notice.

Please charge our Credit Card in the amount of $105.00 covering the fee set forth in 37 CFR 1.16(f) and 1.21(h). Credit Card Payment Form SB-2038, with a signature from an authorized cardholder, is enclosed. The Director is hereby authorized to charge any deficiency in the fees filed, asserted to be filed or which should have been filed herewith (or with any paper hereafter filed in this application by this firm) to our Deposit

IPR2023-00033
Apple EX1002 Page 168

Account No. 04-1073, under Order No. M0025.0311/P311. A duplicate copy of this paper is enclosed.

Dated: May 24, 2005

Respectfully submitted,

By _____

Stephen A. Soffen
Registration No.: 31,063
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street NW
Washington, DC 20037-1526
(202) 785-9700
Attorney for Applicant

IPR2023-00033
Apple EX1002 Page 169

Rec'd PCT/PTO 24 MAY 2005

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

U.S. APPLICATION NUMBER NO.	FIRST NAMED APPLICANT	ATTY. DOCKET NO.
10/503,463	Mark Catchpole	M0025.0311/P311

INTERNATIONAL APPLICATION NO.	
PCT/GB03/00459	
I.A. FILING DATE	PRIORITY DATE
02/04/2003	02/04/2002

24998
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street, NW
Washington, DC 20037

CONFIRMATION NO. 8875

371 FORMALITIES LETTER

OC000000015454716

Date Mailed: 04/01/2005

NOTIFICATION OF MISSING REQUIREMENTS UNDER 35 U.S.C. 371 IN THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

The following items have been submitted by the applicant or the IB to the United States Patent and Trademark Office as a Designated / Elected Office (37 CFR 1.495).

- Indication of Small Entity Status
- Copy of the International Application filed on 08/04/2004
- Copy of the International Search Report filed on 08/04/2004
- Preliminary Amendments filed on 08/04/2004
- Information Disclosure Statements filed on 08/04/2004
- Request for Immediate Examination filed on 08/04/2004
- U.S. Basic National Fees filed on 08/04/2004
- Priority Documents filed on 08/04/2004

RECEIVED
APR - 8 2005
M.S. 4/8/05
SAS/ Formalities
JW 4/11/05

The following items **MUST** be furnished within the period set forth below in order to complete the requirements for acceptance under 35 U.S.C. 371:

- Oath or declaration of the inventors, in compliance with 37 CFR 1.497(a) and (b), identifying the application by the International application number and international filing date.

- **$65** Surcharge for providing the oath or declaration later than 30 months from the priority date (37 CFR 1.492(e)) is required.

SUMMARY OF FEES DUE:

Total additional fees required for this application is **$65** for a Small Entity:

- **$65** Late oath or declaration Surcharge.

ALL OF THE ITEMS SET FORTH ABOVE MUST BE SUBMITTED WITHIN TWO (2) MONTHS FROM THE

DATE OF THIS NOTICE OR BY 32 MONTHS FROM THE PRIORITY DATE FOR THE APPLICATION, WHICHEVER IS LATER. FAILURE TO PROPERLY RESPOND WILL RESULT IN ABANDONMENT.

The time period set above may be extended by filing a petition and fee for extension of time under the provisions of 37 CFR 1.136(a).

Applicant is reminded that any communications to the United States Patent and Trademark Office must be mailed to the address given in the heading and include the U.S. application no. shown above (37 CFR 1.5)

*A copy of this notice **MUST** be returned with the response.*

INDIA L EVANS

Telephone: (703) 308-9140 EXT 212

PART 1 - ATTORNEY/APPLICANT COPY

U.S. APPLICATION NUMBER NO.	INTERNATIONAL APPLICATION NO.	ATTY. DOCKET NO.
10/503,463	PCT/GB03/00459	M0025.0311/P311

FORM PCT/DO/EO/905 (371 Formalities Notice)

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

Bib Data Sheet

CONFIRMATION NO. 8875

SERIAL NUMBER	FILING OR 371(c) DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
10/503,463	05/24/2005 RULE	704	2654	M0025.0311/P311

APPLICANTS

Mark Catchpole, Cambridgeshire, GBN, UNITED KINGDOM;

**** CONTINUING DATA ***************************

This application is a 371 of PCT/GB03/00459 02/04/2003

**** FOREIGN APPLICATIONS *********************

UNITED KINGDOM 0202546.8 02/04/2002

** SMALL ENTITY **

Foreign Priority claimed	STATE OR COUNTRY	SHEETS DRAWING	TOTAL CLAIMS	INDEPENDENT CLAIMS
☐ yes ☐ no	GBN	6	91	6
35 USC 119 (a-d) conditions met ☐ yes ☐ no ☐ Met after Allowance				
Verified and Acknowledged Examiner's Signature Initials				

ADDRESS

24998

TITLE

Speech recognition circuit using parallel processors

FILING FEE RECEIVED		
1293	FEES: Authority has been given in Paper No. _____ to charge/credit DEPOSIT ACCOUNT No. _____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

Rec'd PCT/PTC 24 MAY 2005
10/544480

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

DECLARATION FOR PATENT APPLICATION

As the below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor of the subject matter which is claimed and for which a patent is sought on the invention entitled:

SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

the specification of which was filed on February 4, 2003 as PCT Application No. PCT/GB03/00459 .

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by an amendment, if any, specifically referred to herein.

I acknowledge the duty to disclose all information known to me that is material to patentability in accordance with Title 37, Code of Federal Regulations, § 1.56.

FOREIGN PRIORITY CLAIM

I hereby claim foreign priority benefits under Title 35, United States Code § 119(a)-(d) of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

☐ no such foreign applications have been filed

☒ such foreign application have been filed as follows:

EARLIEST FOREIGN APPLICATION(S), IF ANY FILED WITHIN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO THIS U.S. APPLICATION

Application Number	Country	Date of Filing	Priority Claimed Under 35 USC 119
0202546.8	GB	February 4, 2002	x Yes No ___
			___ Yes No ___
			___ Yes No ___

ALL FOREIGN APPLICATION(S), IF ANY FILED MORE THAN 12 MONTHS (6 MONTHS FOR DESIGN) PRIOR TO THIS U.S. APPLICATION

Application Number	Country	Date of Filing

CLAIM FOR BENEFIT OF EARLIER U.S. PROVISIONAL APPLICATIONS

I hereby claim priority benefits under Title 35, United States Code §119(e), of any United States provisional patent application(s) listed below:

[x] no such U.S. provisional applications have been filed.

[ ] such U.S. provisional application have been filed as follows:

Application Number	Date of Filing	Priority Claimed Under 35 USC 119
		___ Yes No ___
		___ Yes No ___
		___ Yes No ___

CLAIM FOR BENEFIT OF EARLIER U.S./PCT APPLICATION(S)

I hereby claim the benefit under Title 35, United States Code, §120 of the United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose all information that is material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56 which became available to me between the

filing date of the prior application and the national or PCT international filing date of this application:

[x] no such U.S./PCT applications have been filed.

[ ] such U.S./PCT application have been filed as follows:

Application Number	Relationship	Parent Application	Date of Filing

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

I hereby appoint:

All practitioners at Customer Number 24998

all of **Dickstein Shapiro Morin & Oshinsky LLP**, 2101 L Street NW, Washington, DC 20037-1526, jointly, and each of them severally, my attorneys at law/patent agent(s), with full power of substitution, delegation and revocation, to prosecute this application, to make alterations and amendments therein, to receive the patent, and to transact all business in the U. S. Patent and Trademark Office connected therewith.

Please mail all correspondence to Stephen A. Soffen, whose address is:

Dickstein Shapiro Morin & Oshinsky LLP
2101 L Street NW
Washington, DC 20037-1526

Please direct telephone calls to: Stephen A. Soffen at (202) 828-4879.

Please direct facsimiles to: (202) 887-0689

US P84480
Rec'd PCT/PTC 24 MAY 2005

1-00

Full name of sole or first inventor Mark Catchpole
Sole or first inventor's signature X Mark Catchpole.
Date X 02/09/04
Residence Cambridgeshire, United Kingdom GBN
Citizenship United Kingdom
Mailing Address ~~Ash House~~ St John's Innovation Centre ~~Breckenwood Road~~ Cowley Road ~~Fulbourn~~, Cambridge CB4 0WS CAMBRIDGESHIRE; ~~CB1 5DQ~~; UNITED KINGDOM

Full name of second inventor, if any	
Second inventor's signature	Date
Residence	
Citizenship	
Mailing Address	

Full name of third inventor, if any	
Third inventor's signature	Date
Residence	
Citizenship	
Mailing Address	

Full name of fourth inventor, if any	
Fourth inventor's signature	Date
Residence	
Citizenship	
Mailing Address	

IPR2023-00033
Apple EX1002 Page 176

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

U.S. APPLICATION NUMBER NO.	FIRST NAMED APPLICANT	ATTY. DOCKET NO.
10/503,463	Mark Catchpole	M0025.0311/P311

INTERNATIONAL APPLICATION NO.	
PCT/GB03/00459	
I.A. FILING DATE	PRIORITY DATE
02/04/2003	02/04/2002

24998
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street, NW
Washington, DC 20037

CONFIRMATION NO. 8875

371 ACCEPTANCE LETTER

OC000000016508451

Date Mailed: 07/15/2005

NOTICE OF ACCEPTANCE OF APPLICATION UNDER 35 U.S.C 371 AND 37 CFR 1.495

The applicant is hereby advised that the United States Patent and Trademark Office in its capacity as a Designated / Elected Office (37 CFR 1.495), has determined that the above identified international application has met the requirements of 35 U.S.C. 371, and is ACCEPTED for national patentability examination in the United States Patent and Trademark Office.

The United States Application Number assigned to the application is shown above and the relevant dates are:

05/24/2005	05/24/2005
DATE OF RECEIPT OF 35 U.S.C. 371(c)(1), (c)(2) and (c)(4) REQUIREMENTS	DATE OF COMPLETION OF ALL 35 U.S.C. 371 REQUIREMENTS

A Filing Receipt (PTO-103X) will be issued for the present application in due course. **THE DATE APPEARING ON THE FILING RECEIPT AS THE " FILING DATE" IS THE DATE ON WHICH THE LAST OF THE 35 U.S.C. 371 (c)(1), (c)(2) and (c)(4) REQUIREMENTS HAS BEEN RECEIVED IN THE OFFICE. THIS DATE IS SHOWN ABOVE.** *The filing date of the above identified application is the international filing date of the international application (Article 11(3) and 35 U.S.C. 363).* Once the Filing Receipt has been received, send all correspondence to the Group Art Unit designated thereon.

The following items have been received:

- Indication of Small Entity Status
- Copy of the International Application filed on 08/04/2004
- Copy of the International Search Report filed on 08/04/2004
- Preliminary Amendments filed on 08/04/2004
- Information Disclosure Statements filed on 08/04/2004
- Oath or Declaration filed on 05/24/2005
- Request for Immediate Examination filed on 08/04/2004
- U.S. Basic National Fees filed on 08/04/2004
- Assignment filed on 05/24/2005
- Priority Documents filed on 08/04/2004

Applicant is reminded that any communications to the United States Patent and Trademark Office must be mailed to the address given in the heading and include the U.S. application no. shown above (37 CFR 1.5)

INDIA L EVANS
Telephone: (703) 308-9140 EXT 212

PART 3 - OFFICE COPY

FORM PCT/DO/EO/903 (371 Acceptance Notice)

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/503,463	05/24/2005	Mark Catchpole	M0025.0311/P311	8875

24998 7590 10/29/2008

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403

EXAMINER	
ABEBE, DANIEL DEMELASH	
ART UNIT	**PAPER NUMBER**
2626	
MAIL DATE	**DELIVERY MODE**
10/29/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/503,463	CATCHPOLE, MARK	
	Examiner	Art Unit	
	Daniel D. Abebe	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-91* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☒ Claim(s) *1-44 and 68-91* is/are allowed.

6)☒ Claim(s) *45-47,49-58 and 60-67* is/are rejected.

7)☒ Claim(s) *48 and 59* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

Information Disclosure Statement

The information disclosure statement filed on 8/04/2004 fails to comply with 37 CFR 1.98(a)(2), which requires a legible copy of each cited foreign patent document; each non-patent literature publication or that portion which caused it to be listed; and all other information or that portion which caused it to be listed. It has been placed in the application file, but the information referred to therein has not been considered.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 45-47, 49-58 and 60-67 are rejected under 35 U.S.C. 103(a) as being unpatentable over Phillips (2001/0011218) in view of Young et al (7,120,582).

As to claim 45, Phillips teaches a speech recognition circuit comprising:

Input means for receiving speech parameters;

memory containing acoustic unit for word recognition where the acoustic unit comprises plurality of acoustic models including phonemes;

plurality of parallel processors connected to the input means for processing the speech parameter in parallel, where each processor is connected to the acoustic model in parallel;

means for controlling the processor to process the speech parameter against the acoustic units according to the states of model; and

receiving means for receiving the result of the processing the speech parameter from the plurality of the processors (Fig.1; Par.0015-0020).

It is noted that Phillips doesn't explicitly teach where the acoustic model used by the parallel processor for matching the speech parameter is lexical model. in short Phillips doesn't explicitly teach lexical model. Young, however, teaches a speech recognition circuit, comprising input means for receiving speech parameters; and

memory means containing a plurality of lexical data for processing the input speech parameter (Fig.6; abstract). It would have been obvious to one of ordinary skill in the art at the time of applicant's invention to use lexical model in Phillips teaching as the acoustic unit model since lexical models are well known to be a more efficient model for classifying an input speech parameter.

As to claims 46-47, Young teaches a speech recognition circuit where each lexical memory means stores lexical data (including phones, phonemes.. etc) as a tree structure where each tree structure comprises a model of words having common prefix component as well as a language model for providing context for the traversed tree of an active vocabulary (abstract; Figs.14-15).

As to claim 49, Phillip teaches where each processor is operative to compare the processed speech parameter with the acoustic data in the memory to identify words as a word recognition event and send the result to the user (Fig.1; Par.0017).

As to claims 50-52, Phillip teaches a language model means for constraining the processing result by the acoustic model and where each processor generates likelihood score (Par.0004, 0007).

As to claims 53-54, Young teaches processing the speech parameter using lexical tree structure model including word list with a corresponding propagated score (Figs.8,14-15; Col.16, lines 53-65).

As to claim 55, Young teaches accumulated score for the node (Col.12, line 60-Col.13, line 10; Fig.11).

As to claims 56-57 and 60, Young teaches where the lexical processor is operative to identify active vocabulary represented as lexical tree structure and process to identify the input speech parameter in accordance to the score (Fig.12, 27).

As to claims 58 and 61-62, Young teaches control module and pruning means for discarding the nodes processing according to their score (Col.16, lines 35-50; Fig.2).

As to claims 63-64, Phillips teaches input means arranged to receive the speech parameters as vectors and phonemes (Fig.1).

As to claim 65, Young teaches where in the control processor controls to process a lexical tree corresponding to lexical data (Fig.2, 5-6, 12, 27).

As to claim 66, Phillips teaches memory means (Fig.1).

As to claim 67, Young teaches HMM model and performing viterbi search on the lexical data (Fig.12-27).

Claim Objections

Claim 48 is objected to because of the following informalities: Claim 48 recites the limitation "said lexical tree processor" in line 3. There is insufficient antecedent basis for this limitation in the claim. Appropriate correction is required.

Allowable Subject Matter

Claims 1-44 and 68-91 are allowed. The following is a statement of reasons for the indication of allowable subject matter the claims are allowed because the prior arts of record alone or in combination fail to teach the claimed speech recognition circuit comprising a plurality of lexical tree processors connected in parallel to the input means for processing a plurality of lexical trees at the same time.

Claim 59 is objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The following is a statement of reasons for the indication of allowable subject matter: the claim is allowable because the prior arts of record fail to teach where more than one lexical tree is processed at the same time by a pluralities of parallel.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Daniel D. Abebe whose telephone number is 571-272-7615. The examiner can normally be reached on monday-friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Daniel D Abebe/
Primary Examiner, Art Unit 2626

Notice of References Cited	Application/Control No. 10/503,463	Applicant(s)/Patent Under Reexamination CATCHPOLE, MARK	
	Examiner Daniel D. Abebe	Art Unit 2626	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-7,120,582	10-2006	Young et al.	704/255
*	B	US-5,832,428	11-1998	Chow et al.	704/254
*	C	US-5,995,930	11-1999	Hab-Umbach et al.	704/254
*	D	US-2001/0011218	08-2001	Phillips et al.	704/256
*	E	US-6,047,283	04-2000	Braun, Bernhard	707/3
*	F	US-7,035,802	04-2006	Rigazio et al.	704/256
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Application Number	Application/Control No.	Applicant(s)/Patent under Reexamination	
	10/503,463	CATCHPOLE, MARK	
	Examiner	Art Unit	
	Daniel D. Abebe	2626	

Search Notes	Application/Control No.	Applicant(s)/Patent under Reexamination	
	10/503,463	CATCHPOLE, MARK	
	Examiner	Art Unit	
	Daniel D. Abebe	2626	

SEARCHED			
Class	Subclass	Date	Examiner
704	235, 242, 243, 256	10/20/2008	DA

INTERFERENCE SEARCHED			
Class	Subclass	Date	Examiner

SEARCH NOTES (INCLUDING SEARCH STRATEGY)		
	DATE	EXMR
WEST NPL	10/20/2008	DA

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 8875

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
10/503,463	05/24/2005	704	2626	M0025.0311/P311

APPLICANTS

Mark Catchpole, Cambridgeshire, GBN, UNITED KINGDOM;

**** CONTINUING DATA ***************************

This application is a 371 of PCT/GB03/00459 02/04/2003

**** FOREIGN APPLICATIONS **************************

UNITED KINGDOM 0202546.8 02/04/2002

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED **** ** SMALL ENTITY **

Foreign Priority claimed	35 USC 119(a-d) conditions met	Verified and Acknowledged	Met after Allowance	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
☑ Yes ☐ No	☑ Yes ☐ No	/DANIEL DEMELASH ABEBE/ Examiner's Signature	☐ Initials	GBN	6	91	6

ADDRESS

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403
UNITED STATES

TITLE

Speech recognition circuit using parallel processors

FILING FEE RECEIVED		
1293	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

10/503463
DT1 Rec'd PCT/PTO 04 AUG 2004

Docket No.: M0025.0311/P311
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Mark Catchpole

Application No.: Not Yet Assigned

Filed: Concurrently Herewith

Art Unit: N/A

For: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

Examiner: Not Yet Assigned

INFORMATION DISCLOSURE STATEMENT (IDS)

U.S. Patent and Trademark Office
220 20th Street S.
Customer Window
Crystal Plaza Two, Lobby, Room 1B03
Arlington, VA 22313-1450

Dear Sir:

Pursuant to 37 CFR 1.56, 1.97 and 1.98, the attention of the Patent and Trademark Office is hereby directed to the references listed on the attached PTO/SB/08. It is respectfully requested that the information be expressly considered during the prosecution of this application, and that the references be made of record therein and appear among the "References Cited" on any patent to issue therefrom.

This Information Disclosure Statement accompanies the new patent application submitted herewith.

Copies of the references on the PTO/SB/08 are not provided.

In accordance with 37 CFR 1.97(g), the filing of this Information Disclosure Statement shall not be construed to mean that a search has been made or that no other material information as defined in 37 CFR 1.56(a) exists. In accordance with 37 CFR

IPR2023-00033
Apple EX1002 Page 190

1.97(h), the filing of this Information Disclosure statement shall not be construed to be an admission that any patent, publication or other information referred to therein is "prior art" for this invention unless specifically designated as such.

It is submitted that the Information Disclosure Statement is in compliance with 37 CFR 1.98 and the Examiner is respectfully requested to consider the listed references.

The Director is hereby authorized to charge any deficiency in the fees filed, asserted to be filed or which should have been filed herewith (or with any paper hereafter filed in this application by this firm) to our Deposit Account No. 04-1073, under Order No. M0025.0311/P311. A duplicate copy of this paper is enclosed.

Dated: August 4, 2004

Respectfully submitted,

By _____
Stephen A. Soffen
Registration No.: 31,063
DICKSTEIN SHAPIRO MORIN & OSHINSKY LLP
2101 L Street NW
Washington, DC 20037-1526
(202) 785-9700
Attorney for Applicant

IPR2023-00033
Apple EX1002 Page 191

PTO/SB/08a/b (08-03)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449A/B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 1

Complete if Known	
Application Number	Not Yet Assigned
Filing Date	Concurrently Herewith
First Named Inventor	Mark Catchpole
Art Unit	N/A
Examiner Name	Not Yet Assigned
Attorney Docket Number	M0025.0311/P311

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear	T[6]

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	S.H. Chung et al., "A Parallel Phoneme Recognition Algorithm Based on Continuous Hidden Markov Model", Proceedings 13th International Parallel Processing Symposium and 10TH Symposium on Parallel and Distributed Processing, IPPS/SPDP 1999. Proceedings of 13TH International Parallel Processing Symposium and 10TH Symposium on Parallel and Distributed Pro., IEEE Comput. Soc. pp. 453-457 (1999).	
	CB	S.H. Chung et al., "A Parallel Computation Model for Integrated Speech and Natural Language Understanding." IEEE Transactions on Computers, Vol. 42 No. 10, October 1, 1993.	
	CC	N. Deshmukn et al., "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition: Working Toward a Solution to the Decoding Problem", IEEE Signal Processing Magazine, Vol. 16, No. 5, pp. 84-107 (September 1999).	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1]Applicant's unique citation designation number (optional). [2]Applicant is to place a check mark here if English language Translation is attached.

/Daniel Abebe/ (10/25/2008)

WEST Search History for Application 10503463

Creation Date: 2008102513:47

Query	DB	Op.	Plur.	Thes.	Date
((lexical near2 tree) and ((lexical near3 process$5) and 704/$5)) AND @pd>20081021	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
((lexical near2 tree) and ((lexical near3 process$5) and 704/$5))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel) and (((lexical near2 tree) and ((lexical near3 process$5) and 704/$5)))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 lexcial near4 proces$4) and (lexical tree)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 lexical near4 proces$4) and (lexical tree)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 lexical near4 proces$4)	PGPB, USPT, USOC, EPAB, JPAB,	ADJ	YES		10-22-2008

IPR2023-00033
Apple EX1002 Page 193

	DWPI, TDBD				
(parallel near5 parameter near4 proces$4) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 parameter near4 proces$4) and (704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(pluralit near4 proces$4) and (704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(pluralit$4 near4 proces$4) and lexical and (704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
tree structure and ((pluralit$4 near4 proces$4) and lexical and (704/$5))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel adj2 model) and lexical and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near2 match$6) and lexical and 704/$5	PGPB, USPT,	ADJ	YES		10-22-2008

	USOC, EPAB, JPAB, DWPI, TDBD				
7120582.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(7120582.pn.) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near4 tree near4 search$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(pluralit$5 near4 tree near4 search$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near4 tree near4 search$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near4 tree near4 process$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI,	ADJ	YES		10-22-2008

	TDBD				
((parallel near4 tree near4 search$4) and 704/$5) or ((parallel near4 tree near4 process$4) and 704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
lexical adj3 processors!	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel adj3 processors!) and hmm and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel adj3 processors!) and hmm	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(k near lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(k near3 lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near2 pars$4) and (tree structure) and lexical	PGPB, USPT, USOC,	ADJ	YES		10-22-2008

	EPAB, JPAB, DWPI, TDBD				
7035802.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
7035802.pn. or 7120582.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
5995930.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn.) or 7120582.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
lexic$4 and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008

tree$4 and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(process$5) and ((pluralit$4 near4 proces$4) and lexical and (704/$5))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
((process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)) AND @pd>20081022	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
((process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
lexical and (((process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
20010011218.pn. or 6047283.pn. or 7120582.pn.	PGPB, USPT, USOC, EPAB,	ADJ	YES		10-22-2008

	JPAB, DWPI, TDBD				
(20010011218.pn. or 7120582.pn.) AND @pd>20081023	PGPB, USPT	ADJ	YES		10-23-2008
(20010011218.pn. or 7120582.pn.) and score	PGPB, USPT	ADJ	YES		10-23-2008
(history) and ((20010011218.pn. or 7120582.pn.) and score)	PGPB, USPT	ADJ	YES		10-23-2008
(20010011218.pn. or 7120582.pn.) and (stor$6)	PGPB, USPT	ADJ	YES		10-23-2008
(20010011218.pn. or 7120582.pn.) and (propagat$6)	PGPB, USPT	ADJ	YES		10-23-2008
prun$4 and ((20010011218.pn. or 7120582.pn.) and (stor$6))	PGPB, USPT	ADJ	YES		10-23-2008
(multiple or pluralit$4 or more than one) near4 (tree) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008
(704/$5) and ((multiple or pluralit$4 or more than one) near4 (tree) and lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008
((multiple or pluralit$4 or more than one) near4 (tree) near5 (process$4)) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008
(704/$5) and (((multiple or pluralit$4 or more than one) near4 (tree) near5 (process$4)) and lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008

((20010011218.pn. or 7120582.pn.) and (stor$6)) AND @pd>20081023	PGPB, USPT	ADJ	YES		10-24-2008
(7120582.pn. or 20010011218.pn.) and (control$6)	PGPB, USPT	ADJ	YES		10-24-2008
(prun$4 and ((20010011218.pn. or 7120582.pn.) and (stor$6))) AND @pd>20081023	PGPB, USPT	ADJ	YES		10-24-2008
704/$5	USPT	ADJ	YES		10-24-2008
(7120582.pn. or 20010011218.pn.)	PGPB, USPT	ADJ	YES		10-24-2008
lexical tree near4 parallel	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-24-2008
(704/$5) and lexical tree	USPT	ADJ	YES		10-24-2008
704/$5	USPT	ADJ	YES		10-24-2008
((7120582.pn. or 20010011218.pn.)) AND @pd>20081024	PGPB, USPT	ADJ	YES		10-25-2008
((7120582.pn. or 20010011218.pn.))	PGPB, USPT	ADJ	YES		10-25-2008
control$5 and (((7120582.pn. or 20010011218.pn.)))	PGPB, USPT	ADJ	YES		10-25-2008

Docket No.: M0025.0311/P311
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Mark Catchpole

Application No.: 10/503,463 | Confirmation No.: 8875

Filed: May 24, 2005 | Art Unit: 2626

For: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS | Examiner: D. D. Abebe

AMENDMENT IN RESPONSE TO NON-FINAL OFFICE ACTION

MS Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

INTRODUCTORY COMMENTS

In response to the Office Action dated October 29, 2008, please amend the above-identified U.S. patent application as follows:

Amendments to the Claims are reflected in the listing of claims which begins on page 2 of this paper.

Remarks/Arguments begin on page 15 of this paper.

IPR2023-00033
Apple EX1002 Page 201

AMENDMENTS TO THE CLAIMS

1. (Original) A speech recognition circuit comprising:

input means for receiving processed speech parameters;

lexical memory means containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input means for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory means;

results memory means connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

control processor means for controlling said lexical tree processors to process lexical trees identified in said results memory means by performing parallel processing on a plurality of said lexical tree data structures.

2. (Original) A speech recognition circuit according to claim 1, wherein said lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

3. (Original) A speech recognition circuit according to claim 2, wherein said lexical memory means stores each lexical tree data structure as mono phone models of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

IPR2023-00033
Apple EX1002 Page 202

4. (Previously Presented) A speech recognition circuit according to claim 1, wherein each lexical tree processor is operative to compare the processed speech parameters with said lexical data or data derived from said lexical data to identify words as a word recognition event and to send information identifying the identified words to said results memory means as the processing results.

5. (Original) A speech recognition circuit according to claim 4, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

6. (Original) A speech recognition circuit according to claim 5, wherein each lexical tree processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and to output the modified processing result to said results memory means.

7. (Previously Presented) A speech recognition circuit according to claim 4, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events.

8. (Previously Presented) A speech recognition circuit according to claim 1, wherein said lexical tree processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from said lexical data in said lexical memory means.

9. (Previously Presented) A speech recognition circuit according to claim 1, wherein said control processor means is operative to instruct a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory means.

10. (Original) A speech recognition circuit according to claim 9, wherein said history data includes an accumulated score for the recognition path.

11. (Previously Presented) A speech recognition circuit according to claim 1, wherein said control processor means is operative to pass a lexical tree identifier to a said lexical tree processor, said lexical tree processor is operative to read lexical data for the identified lexical tree from said lexical memory means and to process said speech parameters using said read lexical data.

12. (Previously Presented) A speech recognition circuit according to claim 1, wherein said lexical tree processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

13. (Original) A speech recognition circuit according to claim 12, wherein said lexical tree processors are operative to prune the processing of the lexical tree in dependence upon the determined scores.

14. (Previously Presented) A speech recognition circuit according to claim 1, wherein said lexical tree processors are operative to process more than one lexical tree at the same time.

15. (Previously Presented) A speech recognition circuit according to claim 4, wherein said control processor means is operative to determine new lexical tree identifiers for storing in said results memory means for words identified in said results memory means.

16. (Original) A speech recognition circuit according to claim 15, wherein said control processor means is operative to prune the new lexical tree identifiers.

17. (Previously Presented) A speech recognition circuit according to claim 16, wherein each lexical tree processor is operative to determine and output scores for words in the processing

results at word recognition events, and said control processor means is operative to prune said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

18. (Previously Presented) A speech recognition circuit according to claim 1, wherein said input means is arranged to receive said speech parameters as feature vectors.

19. (Previously Presented) A speech recognition circuit according to claim 1, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

20. (Previously Presented) A speech recognition circuit according to claim 1, wherein said lexical tree processors are arranged in groups, said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said groups of lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

21. (Previously Presented) A speech recognition circuit according to claim 1, wherein said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

22. (Previously Presented) A speech recognition circuit according to claim 1, wherein said lexical memory means stores said lexical tree data structures as Hidden Markov Models, and said lexical tree processors are operative to perform the Viterbi search algorithm using each lexical tree data structure.

23. (Original) A speech recognition circuit comprising:

an input buffer for receiving processed speech parameters;

lexical memory containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, the initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory;

results memory connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

a control processor for controlling said lexical tree processors to process lexical trees identified in said results memory by performing parallel processing on a plurality of said lexical tree data structures.

24. (Original) A speech recognition circuit according to claim 23, wherein said lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

25. (Original) A speech recognition circuit according to claim 24, wherein said lexical memory stores said lexical tree data structure as a mono phone model of words, and each lexical tree processor generates context dependant phone models of words dynamically for use in processing the speech parameters.

26. (Previously Presented) A speech recognition circuit according to claim 23, wherein each lexical tree processor compares the processed speech parameters with the lexical data or data derived from the lexical data to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

27. (Original) A speech recognition circuit according to claim 26, including a language model processor for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

28. (Original) A speech recognition circuit according to claim 27, wherein each lexical tree processor receives said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and outputs the modified processing result to said results memory.

29. (Previously Presented) A speech recognition circuit according to claim 26, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events.

30. (Previously Presented) A speech recognition circuit according to claim 23, wherein said lexical tree processors process said speech parameters by comparing said speech parameters with said lexical data in said lexical memory.

31. (Previously Presented) A speech recognition circuit according to claim 23, wherein said control processor instructs a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory.

32. (Original) A speech recognition circuit according to claim 31, wherein said history data includes an accumulated score for the recognition path.

33. (Previously Presented) A speech recognition circuit according to claim 23, wherein said control processor passes a lexical tree identifier to a said lexical tree processor, said lexical tree processor reads lexical data for the identified lexical tree from said lexical memory and processes said speech parameters using said read lexical data.

34. (Previously Presented) A speech recognition circuit according to claim 23, wherein said lexical tree processors determine and output scores in the processing results during the processing of said speech parameters.

35. (Original) A speech recognition circuit according to claim 34, wherein said lexical tree processors prune the processing of the lexical trees in dependence upon the determined scores.

36. (Previously Presented) A speech recognition circuit according to claim 23, wherein said lexical tree processors process more than one lexical tree at the same time.

37. (Previously Presented) A speech recognition circuit according to claim 26, wherein said control processor determines new lexical tree identifiers for storing in said results memory for words identified in said results memory.

38. (Original) A speech recognition circuit according to claim 37, wherein said control processor means is operative to prune the new lexical tree identifiers

39. (Previously Presented) A speech recognition circuit according to claim 38, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

40. (Previously Presented) A speech recognition circuit according to claim 23, wherein said input buffer receives said speech parameters as feature vectors.

41. (Previously Presented) A speech recognition circuit according to claim 23, wherein said input buffer receives said speech parameters as n phones, where n is an integer.

42. (Previously Presented) A speech recognition circuit according to claim 23, wherein said lexical tree processors are arranged in groups, said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

43. (Previously Presented) A speech recognition circuit according to claim 23, wherein said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory is connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

44. (Previously Presented) A speech recognition circuit according to claim 23, wherein said lexical memory stores said lexical tree data structures as Hidden Markov Models, and each lexical tree processor performs the Viterbi search algorithm using a respective said lexical tree data structure.

45-58. (Canceled)

59. (Currently Amended) A speech recognition circuit comprising: ~~according to claim 46,~~

input means for receiving processed speech parameters;

a plurality of lexical memory means containing in combination complete lexical data for word recognition, each lexical memory means containing part of said complete lexical data;

a plurality of processors connected in parallel to said input means for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory means;

IPR2023-00033
Apple EX1002 Page 209

control processor means for controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

receiving means for receiving the results of the processing of the speech parameters from said processors, wherein

each lexical memory means stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique, and

wherein said processors are operative to use lexical data corresponding to more than one lexical tree at the same time.

60-67. (Canceled)

68. (Original) A speech recognition circuit comprising:

an input buffer receiving processed speech parameters;

a plurality of lexical memories containing in combination complete lexical data for word recognition, each lexical memory containing part of said complete lexical data;

a plurality of processors connected in parallel to said input buffer for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory;

a control processor controlling each processor to process said speech parameters using partial lexical data read from a respective said lexical memory; and

a results memory storing the results of the processing of the speech parameters from said processors.

69. (Original) A speech recognition circuit according to claim 68, wherein each lexical memory stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

70. (Original) A speech recognition circuit according to claim 69, wherein each lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

71. (Original) A speech recognition circuit according to claim 70, wherein each lexical memory stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

72. (Previously Presented) A speech recognition circuit according to claim 68, wherein each processor compares the processed speech parameters with the lexical data in a corresponding lexical memory to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

73. (Original) A speech recognition circuit according to claim 72, a including language model processor providing a language model output for modifying the processing results at a word recognition event by a said processor.

74. (Original) A speech recognition circuit according to claim 73, wherein each processor receives said language model output to modify the processing result at a word recognition event determined by said processor and outputs the modified processing result to said results memory.

75. (Previously Presented) A speech recognition circuit according to claim 73, wherein each processor determines and outputs scores for words in the processing results at word recognition events.

76. (Previously Presented) A speech recognition circuit according to claim 68, wherein said processors process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memories.

77. (Previously Presented) A speech recognition circuit according to claim 69, wherein said control processor instructs a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

78. (Original) A speech recognition circuit according to claim 77, wherein said history data includes an accumulated score for the recognition path.

79. (Previously Presented) A speech recognition circuit according to claim 69, wherein said control processor passes a lexical tree identifier to a said processor, said processor reads lexical data for the identified lexical tree from a corresponding said lexical memory and processes said speech parameters using said read lexical data.

80. (Previously Presented) A speech recognition circuit according to claim 68, wherein said processors determine and output scores in the processing results during the processing of said speech parameters.

81. (Original) A speech recognition circuit according to claim 80, wherein said processors prune the processing of the speech parameters in dependence upon the determined scores.

82. (Previously Presented) A speech recognition circuit according to claim 69, wherein said processors use lexical data corresponding to more than one lexical tree at the same time.

83. (Previously Presented) A speech recognition circuit according to claim 72, wherein said control processor identifies lexical data in said lexical memories required for processing next words at word recognition events, and controls said processors to process the speech parameters using the identified lexical data.

84. (Original) A speech recognition circuit according to claim 83, wherein said control processor prunes the identified lexical data.

85. (Original) A speech recognition circuit according to claims 84, wherein each processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes the identified lexical data on the basis of the scores for the words.

86. (Previously Presented) A speech recognition circuit according to claim 68, wherein said input buffer is arranged to receive said speech parameters as feature vectors.

87. (Previously Presented) A speech recognition circuit according to claim 68, wherein said input buffer is arranged to receive said speech parameters as n phones, where n is an integer.

88. (Previously Presented) A speech recognition circuit according to claim 68, wherein said control processor controls each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory.

89. (Previously Presented) A speech recognition circuit according to claim 68, wherein said lexical memories store said lexical data as Hidden Markov Models, and each processor performs the Viterbi search algorithm using a respective part of said lexical data.

90. (Original) A speech recognition method comprising:

receiving processed speech parameters;

storing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;;

passing the speech parameters in parallel to a plurality of lexical tree processors to perform parallel lexical tree processing for word recognition by accessing said lexical data;

storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

controlling said lexical tree processors to process lexical trees identified by the stored identifiers by performing parallel processing on a plurality of said lexical tree data structures.

91. (Original) A speech recognition method comprising:

receiving processed speech parameters;

storing complete lexical data for word recognition in a plurality of lexical memory means, each lexical memory means containing part of said complete lexical data;

processing the speech parameters in parallel using a plurality of processors connected in parallel, said processors being arranged in groups of processors, each group of processors being connected to a said lexical memory means;

controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

storing the results of the processing of the speech parameters from said processors.

REMARKS

Claims 1-44, 59, and 68-91 are pending in this application. Applicant acknowledges with appreciation the indication of allowability of claims 1-44 and 68-91. Claim 59 has been amended. Claims 45-58 and 60-67 have been canceled. In view of the amendments to the claim 59 and the remarks below, Applicant respectfully requests that the rejections be withdrawn and that the claims be allowed.

Claim 59 stands objected to as being dependent upon a rejected base claim, but Examiner indicates that it would be allowable if rewritten in independent form including all the limitations of the base claim and any intervening claims. Claim 59 has been rewritten, as suggested by Examiner, to include all limitations of base claim 45 and intervening claim 46. Applicant respectfully requests that the objection be withdrawn and the claim be allowed.

In view of the above amendment, Applicant believes the pending application is in condition for allowance. If there are any additional charges in connection with this filing or any subsequent filings (including but not limited to issue fees), the Examiner is respectfully requested and authorized to charge Deposit Account No. 04-1073 therefor under Order No. M0025.0311/P311.

Dated: January 29, 2009

Respectfully submitted,

By MARK THRONSON #33,082

Stephen A. Soffen
Registration No.: 31,063
Jonathan L. Falkler
Registration No.: 62,115
DICKSTEIN SHAPIRO LLP
1825 Eye Street, NW
Washington, DC 20006-5403
(202) 420-2200
Attorneys for Applicant

Electronic Acknowledgement Receipt

EFS ID:	4696696
Application Number:	10503463
International Application Number:	
Confirmation Number:	8875
Title of Invention:	Speech recognition circuit using parallel processors
First Named Inventor/Applicant Name:	Mark Catchpole
Customer Number:	24998
Filer:	Mark J. Thronson/Nance Pitzer
Filer Authorized By:	Mark J. Thronson
Attorney Docket Number:	M0025.0311/P311
Receipt Date:	29-JAN-2009
Filing Date:	24-MAY-2005
Time Stamp:	12:49:59
Application Type:	U.S. National Stage under 35 USC 371

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		M00250311.pdf	654854 853a5c86f33c4cef92205b6501d33213b41c43e0	yes	15

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Amendment/Req. Reconsideration-After Non-Final Reject	1	1
	Claims	2	14
	Applicant Arguments/Remarks Made in an Amendment	15	15
Warnings:			
Information:			
	Total Files Size (in bytes):	654854	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (07-06)
Approved for use through 1/31/2007. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD

Substitute for Form PTO-875

Application or Docket Number	Filing Date	☐ To be Mailed
10/503,463	05/24/2005	

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY ☒ RATE ($)	SMALL ENTITY ☒ FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A			N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A			N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A			N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =		OR	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =			X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $250 ($125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).						
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))							
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT 01/29/2009	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE ($)	SMALL ENTITY ADDITIONAL FEE ($)	OR	OTHER THAN SMALL ENTITY RATE ($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 69	Minus	** 91	= 0	X $26 =	0	OR	X $ =	
Independent (37 CFR 1.16(h))	* 6	Minus	***6	= 0	X $110 =	0	OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE	0	OR	TOTAL ADD'L FEE	

AMENDMENT	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)	OR	RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =		OR	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =		OR	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))									
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
					TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.

** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".

*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

Legal Instrument Examiner:
/BRENDA MURPHY/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Docket No.: M0025.0311/P311
(PATENT)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:
Mark Catchpole

Application No.: 10/503,463 Confirmation No.: 8875

Filed: May 24, 2005 Art Unit: 2626

For: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS Examiner: D. D. Abebe

INFORMATION DISCLOSURE STATEMENT (IDS)

MS Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Madam:

Pursuant to 37 CFR 1.56, 1.97 and 1.98, the attention of the Patent and Trademark Office is hereby directed to the references listed on the attached PTO/SB/08. It is respectfully requested that the information be expressly considered during the prosecution of this application, and that the references be made of record therein and appear among the "References Cited" on any patent to issue therefrom.

The references listed as No.'s AA, BA, CA, and CB on the attached PTO/SB/08 form were cited in the enclosed International Search Report dated February 24, 2006. Reference No. BC was cited in the enclosed GB Search Report dated December 5, 2003. Reference No. BB was cited in the enclosed GB Search Report dated October 1, 2002. The reference No.'s CC-CE were cited in the International Search Report previously submitted on August 4, 2004. All other references were cited in application serial No. 11/662,704, which is listed as US Publication No. 2008/0255839 A1 on the attached PTO/SB/08 form.

IPR2023-00033
Apple EX1002 Page 219

In accordance with 37 CFR 1.98(a)(2)(ii), Applicant has not submitted copies of U.S. patents and U.S. patent applications. Applicant submits herewith copies of foreign patents and non-patent literature in accordance with 37 CFR 1.98(a)(2).

In accordance with 37 CFR 1.97(g), the filing of this Information Disclosure Statement shall not be construed to mean that a search has been made or that no other material information as defined in 37 CFR 1.56(a) exists. In accordance with 37 CFR 1.97(h), the filing of this Information Disclosure Statement shall not be construed to be an admission that any patent, publication or other information referred to therein is "prior art" for this invention unless specifically designated as such.

It is submitted that the Information Disclosure Statement is in compliance with 37 CFR 1.98 and the Examiner is respectfully requested to consider the listed references.

Please charge our Credit Card in the amount of $180.00 covering the fee set forth in 37 CFR 1.17(p). Credit Card Payment Form SB-2038, with a signature from an authorized cardholder, is enclosed. The Director is hereby authorized to charge any deficiency in the fees filed, asserted to be filed or which should have been filed herewith (or with any paper hereafter filed in this application by this firm) to our Deposit Account No. 04-1073, under Order No. M0025.0311/P311.

Dated: February 3, 2009

Respectfully submitted,

By _____

Stephen A. Soffen
Registration No.: 31,063
Jonathan L. Falkler
Registration No.: 62,115
DICKSTEIN SHAPIRO LLP
1825 Eye Street, NW
Washington, DC 20006-5403
(202) 420-2200
Attorneys for Applicant

Used in Lieu of PTO/SB/08A/B
(Based on PTO 10-07 version)

Substitute for form 1449/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 2

Complete if Known	
Application Number	10/503,463-Conf. #8875
Filing Date	May 24, 2005
First Named Inventor	Mark Catchpole
Art Unit	2626
Examiner Name	D. D. Abebe
Attorney Docket Number	M0025.0311/P311

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA	5,881,312	03-09-1999	Dulong	
	AB	7,174,037 B2	02-06-2007	Arnone et al.	
	AC	5,293,213	03-08-1994	Klein et al.	
	AD	2003/0178584 A1	09-25-2003	Arnone et al.	
	AE	2008/0255839 A1	10-16-2008	Larri et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] (*if known*)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]
	BA	WO 03/067572 A2	08-14-2003			
	BB	GB 2 331 392 A	05-19-1999			
	BC	GB 2 112 194 A	07-13-1983			
	BD	GB 2 380 920 A	04-16-2003			
	BE	GB 2 347 835 A	09-13-2000			
	BF	WO 00/50859	08-31-2000			

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	S. Glinski et al., "Spoken Language Recognition on a DSP Array Processor", IEEE Transactions on Parallel and Distributed Systems, July 5, 1994, No. 7, New York, USA, pp. 697-703.	
	CB	S. Chatterjee et al., "Connected Speech Recognition on a Multiple Processor Pipeline", ICASSP 89, May 23, 1989, Glasgow, UK, pp. 774-777.	
	CC	S-H. Chung et al., "A Parallel Phoneme Recognition Algorithm Based on Continuous Hidden Markov Model", Proceedings of 13th International Parallel Processing Symposium and 10th Symposium on Parallel and Distributed Processing, 1999, Los Alamitos, CA, USA, pp. 453-457.	
	CD	S-H. Chung et al., "A Parallel Computational Model for Integrated Speech and Natural Language Understanding", IEEE Transactions on Computers, New York, USA, Vol. 42, No. 10, October 1, 1993, pp. 1171-1183.	

Examiner Signature		Date Considered	

Used in Lieu of PTO/SB/08A/B
(Based on PTO 10-07 version)

Substitute for form 1449/PTO **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** *(Use as many sheets as necessary)*				*Complete if Known*	
				Application Number	10/503,463-Conf. #8875
				Filing Date	May 24, 2005
				First Named Inventor	Mark Catchpole
				Art Unit	2626
				Examiner Name	D. D. Abebe
Sheet	2	of	2	Attorney Docket Number	M0025.0311/P311

	CE	N. Deshmukh et al., "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition: Working Toward a Solution for the Decoding Problem", IEEE Signal Processing Magazine, Vol. 16, No. 5, September 1999, pp. 84-107.	
	CF	R.M. Woodward et al., " Tissue Classification Using Terahertz Pulsed Imaging", Proceedings of the SPIE - The International Society for Optical Engineering SPIE-INT. SOC. OPT. ENG. USA, Vol. 5318, No.1, 2004, pp. 23-33. Abstract also attached.	
	CG	V.P. Wallace et al., "Biomedical Applications of THz Imaging", Microwave Symposium Digest, 2004, IEEE MTT-S International Fort Worth, TX, USA, June 6-11, 2004, Piscataway, NJ, USA, IEEE, Vol. 3, 6 June 2004 (2004 06-06), pp. 1579-1581	
	CH	V.P. Wallace et al., "Biomedical Applications of Terahertz Pulse Imaging", Second Joint EMBS-BMES Conference 2002. Conference Proceedings. 24th Annual International Conference of the Engineering in Medicinal and Biology Society. Annual Fall Meeting of the Biomedical Engineering Society, Houston, TX, Oct. 23-26, 2002, Annual Vol. 1 of 3. Conf. 24, 23 October 2002 (2002-10-23), pp. 2333-2334.	
	CI	A.J. Fitzgerald et al., "Terahertz Imaging of Breast Cancer, a Feasibility Study", Conference Digest of the 2004 Joint 29th International Conference on Infrared and Millimeter Waves and 12th International Conference on Terahertz Electronics IEEE Piscataway, NJ, USA, 2004, pp. 823-824.	
	CJ	B.E. Cole et al., "Terahertz Imaging and Spectroscopy of Human Skin, In-vivo", Proceedings of SPIE, Vol. 4276, pp. 1-10.	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1]Applicant's unique citation designation number (optional). [2]Applicant is to place a check mark here if English language Translation is attached.

Examiner Signature		Date Considered	

IPR2023-00033
Apple EX1002 Page 222

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2003 (14.08.2003)

PCT

(10) International Publication Number
WO 03/067572 A2

(51) International Patent Classification[7]: G10L 15/00

(21) International Application Number: PCT/GB03/00459

(22) International Filing Date: 4 February 2003 (04.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0202546.8 4 February 2002 (04.02.2002) GB

(71) Applicant *(for all designated States except US)*: **ZENTIAN LIMITED** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **CATCHPOLE, Mark** [GB/GB]; Ash House, Breckenwood Road, Fulbourn, Cambridge, Cambridgeshire CB1 5DQ (GB).

(74) Agent: **COLLINS, John, David**; Marks & Clerk, 57-60 Lincolns Inn Fields, London WC2A 3LS (GB).

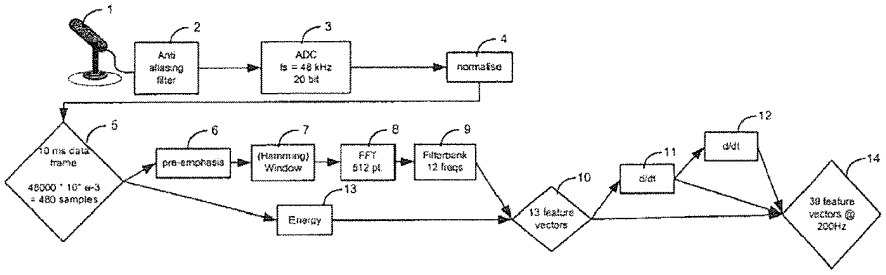
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS



(57) Abstract: A speech recognition circuit comprises an input buffer for receiving processed speech parameters. A lexical memory contains lexical data for word recognition. The lexical data comprises a plurality of lexical tree data structures. Each lexical tree data structure comprises a model of words having common prefix components. An initial component of each lexical tree structure is unique. A plurality of lexical tree processors are connected in parallel to the input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory. A results memory is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory by performing parallel processing on a plurality of said lexical tree data structures.

WO 03/067572 A2

SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

The present invention generally relates to a speech recognition circuit which uses parallel processors for processing the input speech data in parallel.

Conventional large vocabulary speech recognition can be divided into two processes: front end processing to generate processed speech parameters such as feature vectors, followed by a search process which attempts to find the most likely set of words spoken from a given vocabulary (lexicon).

The front end processing generally represents no problem for current processing systems. However, for large vocabulary, speaker independent speech recognition, it is the search process that presents the biggest challenge. An article by Deshmukh et al entitled "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition" (IEEE Signal Processing Magazine, September 1999, pages 84 to 107), the content of which is hereby incorporated by reference, discusses the general concepts of large vocabulary speech recognition. As discussed in this paper, one algorithm for performing the search is the Viterbi algorithm. The Viterbi algorithm is a parallel or breadth first search through a transition network of states of Hidden Markov Models. An acoustic model for words in a lexicon are represented as states of Hidden Markov Models. These states represent phones or n phones in a phone model of the words. The search requires the evaluation of possible word matches. It is known that such a search is computationally intensive.

In order to speed up the processing performed during such a search in a speech recognition system, parallel processing has been explored. In an article by M K Ravishankar entitled "Parallel Implementation of Fast Beam Search for Speaker-Independent Continuous Speech Recognition" (Indian Institute of Science, Bangalor, India, July 16, 1993) a multi-threaded implementation of a fast beam search algorithm is disclosed. The multi-threading implementation requires a significant amount of

communication and synchronization among threads. In an MSC project report by R Dujari entitled "Parallel Viterbi Search Algorithm for Speech Recognition" (MIT, February 1992) the parallel processing of input speech parameters is disclosed in which a lexical network is split statically among processors.

It is an object of the present invention to provide an improved circuit which can perform parallel processing of speech parameters.

In accordance with a first embodiment of the present invention, a speech recognition circuit comprises an input port such as input buffer for receiving parameterized speech data such as feature vectors. A lexical memory arrangement is provided which contains lexicon data for word recognition. The lexical data comprises a plurality of lexical tree data structures representing a plurality of lexical trees. Each lexical tree data structure comprises a model of words having common prefix components and an initial component which is unique as an initial component for lexical trees. A plurality of lexical tree processors are connected in parallel to the input port and perform parallel lexical tree processing for word recognition by accessing the lexical data in the lexical memory arrangement. A results memory arrangement is connected to the lexical tree processors for storing processing results from the lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by the lexical tree processors. A controller controls the lexical tree processors to process lexical trees identified in the results memory arrangement by performing parallel processing of a plurality of lexical tree data structures.

Thus in accordance with this embodiment of the present invention, the processing in order to perform word recognition is distributed across the processors by controlling the processors to perform processing on different lexical trees. The controller controls the processor by the processes to provide for efficient process management by distributing lexical processing to appropriate processors.

The lexical tree data structure can comprise a phone model of words, wherein the components comprise phones. For reduced storage, the lexical tree data structure can comprise a mono phone lexical tree. The mono phone lexical tree can be used to

generate context dependent phone models dynamically. This enables the use of context dependent phone models for matching and hence increased accuracy whilst not increasing memory requirements. Alternatively, the lexical tree data structure can comprise context dependent phone models.

The processing performed by each processor in one embodiment comprises the comparison of the speech parameters with the lexical data, e.g. phone models or data derived from the lexical data (e.g. dynamically generated context dependent phone models) to identify words as a word recognition event and to send information identifying the identified words to the results memory as the processing results. In this embodiment a language model processor arrangement can be provided for providing a language model output for modifying the processing results at a word recognition event by a lexical tree processor. The modification can either take place at each lexical tree processor, or at the language model processing arrangement.

In one embodiment each lexical tree processor determines an output score for words in the processing results at word recognition events. Thus in this embodiment the language model processing arrangement can modify the score using a score for a language model for n preceding words, where n is an integer.

In one embodiment the controller instructs a lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from the results memory. The history data preferably includes an accumulated score for the recognition path. This enables a score to be determined based on the score for the recognition path to accumulate a new score during recognition carried out using the lexical tree data structure. The scores can be output in the processing results to the results memory during the processing of the speech parameters so that the scores can be used for pruning.

In one embodiment of the present invention, each lexical tree processor operates on more than one lexical tree at the same time, e.g. two lexical trees represented by two different lexical tree data structures, or two lexical trees represented by the same data

structure but displaced in time (which can be termed to instances of the same lexical tree).

At word recognition events, the controller determines new lexical tree identifiers for storing in the results memory for words identified in the results memory for respective word events. In order to reduce the processing, the controller can prune the new lexical tree identifiers to reduce the number of lexical trees which are required to be processed. This pruning can be achieved using context dependant n phones to reduce the number of possible next phones. The number can be further reduced by using a language model look ahead technique.

In one embodiment of the present invention, the lexical tree processors are arranged in groups or clusters. The lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory is connected to one of the groups of lexical tree processors and contains part of the lexical data. Thus a group of lexical tree processors and a partial lexical memory form a cluster. Each lexical tree processor is operative to process the speech parameters using a partial lexical memory and the controller controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In another embodiment of the present invention the lexical memory arrangement comprises a plurality of partial lexical memories. Each partial lexical memory being connected to one of the lexical tree processors and containing part of the lexical data. Each lexical tree processor processes the speech parameters using a corresponding partial lexical memory and the controller is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

In one embodiment of the present invention the lexical memory arrangement stores the lexical tree data structures as Hidden Markov Models and the lexical tree processors are operative to perform the Viterbi search algorithm using each respective lexical tree data structure. Thus in this way, this embodiment of the present invention provides a parallel Viterbi lexical tree search process for speech recognition.

The first aspect of the present invention is a special purpose circuit built for performing the speech recognition search process in which there are a plurality of processors for performing parallel lexical tree processing on individual lexical tree processors.

In another aspect of the present invention a speech recognition circuit comprises an input port such as an input buffer for receiving parameterized speech data such as feature vectors. A plurality of lexical memories are provided which contain in combination complete lexical data for word recognition. Each lexical memory contains part of the complete lexical data. A plurality of processors are provided connected in parallel to the input port for processing the speech parameters in parallel. The processors are arranged in groups in which each group is connected to a corresponding lexical memory to form a cluster. A controller controls each processor to process the speech parameters using partial lexical data read from a corresponding lexical memory. The results of processing the speech parameters are output from the processors as recognition data.

Thus this aspect of the present invention provides a circuit in which speech recognition processing is performed in parallel by groups of processors operating in parallel in which each group accesses a common memory of lexical data. This aspect of the present invention provides the advantage of parallel processing of speech parameters and benefits from a limited segmentation of the lexical data. By providing a plurality of processors in a group with a common memory, flexibility in the processing is provided without being bandwidth limited by the interface to the memory that would occur if only a single memory were used for all processors. The arrangement is more flexible than the parallel processing arrangement in which each processor only has access to its own local memory and requires fewer memory interfaces (i.e. chip pins). Each processor within a group can access the same lexical data as any other processor in the group. The controller can thus control the parallel processing of input speech parameters in a more flexible manner. For example, it allows more than one processor to process input speech parameters using the same lexical data in a lexical memory. This is because the lexical data is segmented into domains which are accessible by multiple processors.

In a preferred embodiment this aspect of the present invention is used in combination with the first aspect of the present invention. In such an arrangement each processor performs lexical tree processing and the lexical data stored in each lexical memory comprises lexical tree data structures which each comprise a model of words having common prefix components and an initial component that is unique.

In preferred embodiments of the second aspect of the present invention, the preferred embodiments of the first aspect of the present invention are incorporated.

Embodiments of the present invention will now be described with reference to the accompanying drawings in which:

Figure 1 is a diagram of a speech data processing circuit for generating parameterized speech data (feature vectors);

Figure 2 is a diagram of a speech recognition circuit in accordance with an embodiment of the present invention;

Figures 3a and 3b are schematic diagrams illustrating lexical tree structures;

Figure 4 is a flow diagram illustrating the process performed by a lexical tree processor to determine a temporary lexical tree score in accordance with an embodiment of the present invention;

Figure 5 is a flow diagram illustrating the process performed by the lexical tree processor for processing the input feature vectors in accordance with an embodiment of the present invention; and

Figure 6 is a flow diagram illustrating the process performed by the controller in accordance with an embodiment of the present invention.

Figure 1 illustrates a typical circuit for the parameterization of input speech data. In this embodiment the parameters generated are speech vectors.

A microphone 1 records speech in an analogue form and this is input through an anti-aliasing filter 2 to an analogue-to-digital converter 3 which samples the speech at 48 kHz at 20 bits per sample. The digitized output signal is normalized (4) to generated a 10 millisecond data frame every 5 milliseconds with 5 milliseconds overlap (5). A pre-emphasis operation 6 is applied to the data followed by a hamming window 7. The data is then fast Fourier transformed (FFT) using a 512 point fast Fourier transform (8) before being filtered by filter bank 9 into 12 frequencies. The energy in the data frame 5 is also recorded (13) as an additional feature and together with the 12 frequency outputs of the filter bank 9, 13 feature vectors (10) are thus produced and these are output as part of the 39 feature vectors 14. First and second derivatives (11 and 12) are taken of the 13 feature vectors 10 to complete the generation of the 39 feature vectors 14.

The arrangement illustrated in Figure 1 is purely given for illustration. The present invention encompasses any means by which speech and data can be parameterized to a suitable form for input to the search process as will be described in more detail hereinafter.

Figure 2 is a schematic diagram of a speech recognition circuit in accordance with an embodiment of the present invention for performing the search process. The parameterized speech data, which in this embodiment comprise feature vectors are input to a feature vector buffer 20. The feature vector buffer 20 is provided to buffer the incoming feature vectors to allow lexical tree processors 21 to read and process the feature vectors in the buffer 20 via a feature vector bus 24. A plurality k of lexical tree processors 21 are arranged in a respective lexical tree processor cluster 22. Each lexical tree processor cluster 22 has an acoustic model memory 23 in which is stored lexical data for use by the lexical tree processors 21 within the lexical tree processor cluster 22. Each lexical tree processor 21 in the lexical tree processor cluster 22 are connected to the acoustic model memory 23 within the lexical tree processor 22. There are N lexical tree processor clusters and thus there are Nk lexical tree processors 21 connected by the

feature vector bus 24 to the feature vector buffer 20. Each lexical tree processor 21 is capable of processing a different lexical tree and thus Nk lexical trees can be processed in parallel. The acoustic model memories 23 store as a whole a complete set of lexical data, i.e. lexical tree data structures for use in the lexical tree processing by the lexical tree processors 21. Each acoustic model memory 23 contains part or a segment of the lexical tree data. Since lexical tree processors 21 in a lexical tree processor cluster 22 access the same acoustic model memory 23, it is possible for more than one lexical tree processor 21 to process the same lexical data. This provides for some degree of flexibility in the controlling of the processing by the lexical tree processors 21. Further, the acoustic model memories 23 need not contain only one copy of the lexical data. It is possible to build in a redundancy in the data to further enhance the flexibility. This avoids any bottleneck in the processing due to the search processing focusing on a small number of lexical trees.

A results memory 25 is provided for storing processing results from the lexical tree processors 21 which are received over the path score and history bus 26. The results memory 25 also stores information on lexical trees to identify which lexical trees are to be processed. A search controller 27 is provided to control the processing performed by the lexical tree processors 21 in dependence upon a program and data stored in program and data memory 28. The search controller reads the path scores and lexical tree identifiers from the results memory and controls the lexical tree processors accordingly. A language model processor 29 is provided which is connected to each lexical tree processor 21 by a language model bus 30. The language model processor 29 accesses a language model memory 31 to read language model data for provision to lexical tree processors 21 in response to language model data requests. External control of the language model memory 31 is provided by a word constrains input. The language model processor 29 determines a score for a word occurring following N previous words using N grams. When a lexical tree processor requires a language model score a request is sent to the language model processor 29 over the language model bus 30 identifying the current word and the N-1 previous words. A language model score for the N gram can be returned to the lexical tree processor 21 for the modification of the score at the end of a branch of lexical tree processing. The lexical tree processor can modify the score in accordance with the language model and output a score to the

results memory 25 for a word at the end of a branch of the lexical tree processing. Thus the results memory stores the results as an ordered list of scores for words together with their histories.

The results memory 25 stores the following data:

1. Initial lexical tree data. This comprises pointers to an initial set of lexical trees. No history data is associated with the initial set of lexical trees. The initial set of lexical trees is predetermined and stored in the results memory 25 based on the most likely initial phones of an utterance. This initial lexical tree data is required to initialize the search process.

2. History data for search results. This comprises a record of a recognition path through the lexical tree recognition process performed by the lexical tree processors 21. The history data includes the current word, the previous N-1 words, the current accumulated score, the phone history (for use in the determination of likely next lexical trees using cross word context dependent tri-phones), and an identifier or pointer to the lexical tree used for identifying the word.

3. Best scores for best paths being processed by each lexical tree processor 21. This information enables the search controller 27 to monitor the processing being performed by lexical tree processors 21 to determine whether a global pruning strategy should be applied in order to reassign processing performed by a lexical tree processor if its best score for its best path is below a threshold or well below the best scores for the paths being processed by other lexical tree processors 21.

4. Temporary lexical tree scores. These comprise tree scores which are determined as temporary scores to prune the next lexical trees to be processed at word ends. The temporary lexical tree scores include lexical tree identifiers or pointers to identify the next lexical trees to be processed. The scores enable the pruning of this list.

5. Pruning threshold. This can be a global threshold value for use in the pruning of the lexical trees globally, or a local threshold value for use by a lexical processor for locally pruning the processing performed by the lexical processor 21.

The acoustic model memory 23 stores a Hidden Markov Model for acoustically modelling words as lexical trees. The acoustic model memory 23 stores a plurality of lexical tree data structures. Each lexical tree data structure comprises an n phone model of a number of words having common prefix phones. The first node of the lexical tree (the root) comprises a common n phone to all words in the lexical tree and uniquely identifies the lexical tree.

Each lexical tree processor 21 includes on-board memory (or local memory) to be used during the lexical tree processing. This working memory has to store all of the parameters currently working on including current scores for all paths being processed within the lexical tree, and previous N-1 words for the lexical tree. The local storage of the previous N-1 words enables the lexical tree processor 21, when a word end is reached along a branch of the lexical tree, to send a request for the language model score for an N gram by sending the identity of the N-1 previous words together with the word identified at the end of the branch.

Figure 3a and 3b schematically illustrate lexical trees which can be processed by the lexical tree processor 21 during the recognition of the two words HARD ROCK. In Figure 3a a previous lexical tree terminated at a branch recognizing the word YOU and terminating with the mono phone uw which is associated with two context dependent tri-phones y-uw+k and y-uw+h. Thus the context dependent tri-phone associated with the last phone in the lexical tree word model for YOU indicates two possible next lexical trees, i.e. the lexical trees beginning with the mono phone k and h and having a left context phone of uw. As can be see in Figure 3a this word end YOU therefore leads to two possible next lexical trees. These two lexical trees are traversed in parallel by two different lexical tree processors accumulating scores for matching of input feature vectors into Hidden Markov Models of the context dependent tri-phones associated with each node in the tree. When the end of the branch is reached, a word end event is reached and a word is recognized. As can be seen in Figure 3a, in this example since

four words are of similar phone length, it is possible for the search strategy based on inputting feature vectors in parallel to simultaneously reach a number of possible word ends. These possible word ends are sent as processing results to the results memory. The results memory stores the accumulated score at the word end together with phone history to identify the last phone and its associated context dependent tri-phones. In this example, considering the branch where the word HARD is recognized, the last phone is d which has associated with it two context dependent tri-phones r-d+l and r-d+r. Thus the search controller 27 can identify next possible lexical trees using the next phone in the context dependent tri-phone. In this case, as can be seen in Figure 3b, the next possible lexical trees begin with the phones r and l and have associated with them context dependent tri-phones d-r+ao and d-l+ao respectively. Thus these are the next lexical trees that require processing following the end node or word event detected in the processing lexical tree 2. Figure 3b thus represents the processing that is required at the end of Figure 3a at the word node for HARD in processing the second lexical tree.

As can be seen in Figure 3b, the two lexical trees are processed in parallel by traversing through the branches to reach word ends by sequentially entering in feature vectors in parallel to lexical tree processors processing the respective lexical trees. When the word end is reached, the accumulated score is output to the results memory as described before. In this case since it is the last word, the important context dependent tri-phone associated with the final phone has silence (sil) as the following phone. The final phone can in fact have 50 context dependent tri-phones associated with it if there are 50 possible next phones (including silence). Figure 3b only illustrates the relevant one for the end of the utterance.

Figure 4 is a flow diagram illustrating the processing performed by a lexical tree processor 21 in order to determine a temporary lexical tree score for a lexical tree. When a word end is identified by a lexical tree processor, the processing results are sent to the results memory 25. The results memory identifies the last phone of the recognized words and thus enables the search controller 27 to identify possible next lexical trees using context dependent tri-phones as illustrated in Figure 3b. Although in theory if there are 50 phones there are 50^2 possible lexical trees, due to the use of the context dependant triphones, only 100-200 lexical trees are likely to be identified as

possible lexical trees. The search controller needs to further prune out the lexical trees which are unlikely to generate likely paths with high scores. In order to do this, instructions are sent to the lexical tree processors to determine a temporary lexical tree score which can be used to prune out lexical trees from the processing which have a low score.

The processing by the lexical tree processor to generate the temporary lexical tree score is shown in Figure 4 and for this process, following initialization, a lexical tree processor 21 awaits receipt of data for a new lexical tree from the search controller 27 (step S2). The data comprises a lexical tree pointer on an identifier to lexical tree data stored in the acoustic model memory 23, i.e. the lexical tree data structure, previous N-1 words (for use in language model score determination by the language model processor 29), and the current path score. The previous N-1 words includes the previously recognized word for which a score has been accumulated.

The lexical tree processor 21 then attempts to read the next feature vector from the feature vector buffer 20 (step S3) and if this is not available, an error occurs (step S4). If the feature vector is available in the feature vector buffer 20, the lexical tree processor 21 reads the feature vector from the buffer 20 (step S5) and evaluates the state transitions for the first lexical tree node using the acoustic model data in the acoustic model memory 23 (step S6). Since the state transitions for the first node will require several feature vectors to complete, a score is determined for a state transition of the first node in the lexical tree (step S7). The lexical tree processor 21 then sends a request to the language model processor 29 for a language model score. The request includes the previous N-1 words and all of the possible words represented by the lexical tree data structure in the acoustic model memory 23 (step S8). The language model processor returns scores for each of the N grams ending in the words represented by the lexical tree. The lexical tree processor thus receives language model scores for the words in the lexical tree and picks the highest score (step S9). Alternatively, the language model processor 29 can select the highest n gram score and return this to the lexical tree processor for the determination of the temporary lexical tree score. The temporary lexical tree score for the lexical tree is then generated using the score determined for the first state transition of the first lexical tree node and the highest language model score

(step S10). The temporary lexical tree score is then sent by the lexical tree processor 21 to the results memory 25 over the bus 26 (step S11). In this mode of processing, the processor then awaits the next data of the new lexical tree from the search controller (step S2).

This process is just one of the processes performed by the lexical tree processor. The main processing performed by the lexical tree processor is the lexical tree evaluation. This process will now be described with reference to the flow diagram of Figure 5.

After the start of the process (step S20) a lexical tree processor awaits receipt of data for a lexical tree from the search controller (step S21). The data comprises a lexical tree pointer to a lexical tree data structure stored in the acoustic model memory 23, the previous N-1 words in the recognition path, and the current accumulated score.

The lexical tree processor then attempts to read the next feature vector from the feature vector buffer 20 (step S22). If this is not available, an error occurs (step s23). If the feature vector is available in the feature vector buffer 20, the feature vector is read from the feature vector buffer 20 by the lexical tree processor 21 (step S24) and the lexical tree processor 21 evaluates the state transitions for each path using the acoustic model data in the acoustic model memory 23 (step S25). Scores for the paths are determined in accordance with the conventional Viterbi search technique and the best score for the best path is sent to the results memory 25 while the path histories are stored locally in the on-board memory of the lexical tree processor 21 (step S26). Pruning is applied to the lexical tree by the lexical tree processor to delete paths in order to keep the breadth of the search manageable (step S27). The pruning applied locally by the lexical tree processor can be purely on the basis of a local threshold, which can be provided by the search controller 27, or it can be determined on a relative basis dependent upon the range of scores for paths being processed within the lexical tree processor 21. If the path has not reached a word end (step S28) the lexical tree processor 21 attempts to read the next feature vector from the feature vector buffer 20 (step S22). If the path reaches a word end (step S28) the score must be modified by the language model score (step S29). There are two ways in which this can be done, in this embodiment the lexical tree processor 21 sends a request to the language model processor 29 for a language model

score. The request includes the current word and the N-1 previous words. The language model processor thus returns the language model score and the language model score is used to modify the current accumulated score at the word end. Alternatively, the lexical tree processor can send the language model processor 29 not just the current word and the N-1 previous words, but also the current accumulated score. The language model processor then determines the language model score using the language model memory 31 and modifies the accumulated score using the language model score. The language model processor 29 can then return the modified score to the lexical tree processor 21 which can then pass it to the results memory 25, or a connection between the language model processor 29 and the results memory 25 (not shown) can enable the language model processor 29 to send the score directly to the results memory 25 for the lexical tree processor 21. In this latter case, the language model processor 29 must also receive a lexical tree pointer to identify the lexical tree for which the score applies.

Assuming in this embodiment that the lexical tree processor calculates the modified score (step S29), the score and history is then sent by the lexical tree processor 21 to the results memory 25 (step S30). The history data sent to the results memory 25 comprises the lexical tree pointer identifying the lexical tree, the modified score at the word end, the phone history identifying at least the last phone to allow for context dependent tri-phone determination of next lexical trees, and the identity of the word identified in the evaluation process.

The lexical tree processor then deletes the path and history data in its on-board memory (step S31) and determines if there are any paths still left to be processed (step S32). If so, the lexical tree processor 21 tries to access the next feature vector available in the feature vector buffer 20 (step S22). If there are no paths left to be processed, a message is sent by the lexical tree processor to the search controller 27 to indicate that the lexical tree has been processed (step S33). In this level of processing the lexical tree processor will then await the next data for a lexical tree from the search controller 27.

The flow diagram of Figure 5 illustrates the processing of a single instance of a lexical tree. The single point in time a lexical tree can start to be processed from a feature

vector, i.e. lexical tree processing starts at a time frame. However, due to a word end returned from lexical tree processing in another lexical tree processor, the lexical tree processor may also be asked during a subsequent time frame to start processing another instance of the lexical tree with a different path history. Thus a lexical tree processor 21 can be processing multiple instances of the lexical tree with different paths having different path histories. In this instance, the pruning applied in step S27 needs to take into account all paths currently being processed in all instances of lexical tree processing by the lexical tree processor 21. Thus although pruning is applied locally, it is applied over more than one instance of the lexical tree.

The operation of the search controller 27 will now be described with reference to the flow diagram of Figure 6.

After initialization (step S40) initial lexical tree data in the results memory 25 is read (step S41). The search controller distributes the initial lexical tree data amongst the lexical tree processors for temporary lexical tree score determination (step S42) as described with reference to Figure 4. Temporary lexical tree scores are returned by the lexical tree processors 21 to the results memory 25 (step S43) and the search controller reads the temporary lexical tree scores in the results memory 25 and prunes the lexical trees in the results memory 25 on the basis of the read temporary lexical scores (step S44). Thus the set of initial lexical trees can be quite large and this requires a large number of temporary lexical tree score determinations to be performed by the lexical tree processors 21. However, since the lexical tree processors only need to determine a score for a single feature vector, this can be performed quite quickly. Thus the number of lexical trees for which full evaluation by the lexical tree processors 21 is required can thus be reduced by the selection of the lexical trees with the highest temporary lexical tree score. The search controller than distributes the lexical tree processing among the lexical tree processors 21 by sending each lexical tree processor data including a lexical tree pointer to a lexical tree data structure in the acoustic model memory 23 in the cluster 22. The search controller 27 then awaits the entry of processing results, i.e. history and scores in the results memory 25 for words (step S46). When history and scores for words are entered in the results memory, the search controller 27 determines next possible lexical trees using cross-word tri-phones (step S47). For example, as

illustrated in Figures 3a and 3b, when the word HARD is entered in the results memory, the search controller identifies the final phone of the word as being d and thus next possible tri-phones can comprise any tri-phones having the left phone d. If there are 50 possible phones, this results in 50^2 possible tri-phones. In practice, only 1 to 200 are likely. However, this may still represent too many for full evaluation and thus the lexical tree data for the next possible lexical trees are distributed amongst the lexical tree processors for temporary lexical tree score determination (step S48). Temporary lexical tree scores are returned to the results memory 25 by the lexical tree processors 21 which execute the process as described with reference to Figure 4 (step S49).

The search controller 27 prunes the list of next lexical trees in the results memory 25 on the basis of the temporary lexical scores (step S50). The search controller then determines if the words in the results memory have been processed and there are no lexical trees left to be processed and whether or not lexical tree processes are finished (step S51). If not, the lexical tree processing is distributed amongst the lexical tree processors (step S45). When the end of the utterance is reached, next possible lexical trees will all have been processed and the results can be output from the results memory 25 (step S52).

It can be seen from the foregoing description that the embodiment of the invention provides for an efficient parallel processing system incorporating a parallel lexical tree processing architecture and a segmented lexicon storage architecture. Each lexical tree processor performs lexical tree evaluation and language model look ahead for evaluation of company scores for lexical trees to assist in pruning of the number of potential lexical trees that require processing. The architecture is ideally suited to synchronous breadth first searching e.g. a Viterbi search (i.e. real time recognition). The breadth first search is performed synchronously by the parallel processors on the input search parameters when the language model look ahead takes place, it is possible for the feature vector buffer 20 to store a small number, e.g. one or two, feature vectors whilst the language model look ahead is performed on the feature vector so that synchronous breadth first searching can be carried out by re-reading the feature vector out of the buffer 20 when it is determined which lexical trees are to be evaluated by the

lexical trees processors 21 after a word end. Thus, the feature vector buffer 20 provides a method of decoupling the search parameterisation circuitry from the search circuit.

If the buffer causes processing to get too far behind, e.g. there are a lot of word end language model look ahead operations to be performed, it is possible for the lexical tree processors to be controlled to temporarily halve their local beamwidths (i.e. more vigorously prune) and thus complete their processing more quickly. This would allow the content of the buffer to be decreased as the processors catch up.

Although the present invention has been described hereinabove with reference to a specific embodiment, it will be apparent to a skilled person in the art that modifications lie within the spirit and scope of the present invention. Although in the embodiment described hereinabove, the acoustic model memories 23 were described as storing parts of the lexical data, it is also possible for more than one copy of the lexical data to be stored in different acoustic model memories 23. Thus, more than one lexical tree processor cluster 22 can have access to the same lexical data. This increases the flexibility and allows the controller to avoid any bottlenecks in processing which, for example, can occur if the highest pass scores all require processing by one or a number of lexical tree data structures residing in one of the acoustic model memories 23.

Although in the embodiments described hereinabove, the memory is partitioned in to a plurality of acoustic model memories 23, the present invention is not limited to the partitioning of the lexical data and a plurality of lexical tree processors can access a single memory containing lexical data.

The lexical data stored can comprise mono phone models of words. When the lexical data is partitioned, these can be distributed evenly across the memories. If there are 50 mono phones, there will be 50 lexical tree data structures having a unique initial model phone data node. Alternatively, the lexical tree data structures can comprise a context dependent n phone word models. If there are 50 mono phones, there are potentially 50^n lexical tree data structures. For tri-phones, this comprises 50^3 (i.e. 125,000). These are distributed so that trees with the same left context tri-phone, are eventually spread across memories when the lexical data is partitioned. Thus in this way the word end

where the left context tri-phone is fixed i.e. the same, the next possible lexical trees will inherently be distributed across the memories, thus ensuring efficient parallel processing of the next lexical trees.

Although the embodiment described hereinabove uses a single results memory accessible by all processors, the present invention encompasses a plurality of results memories. For example, a results memory can be associated with each processor to store the processing results and the controller can read the results in the plurality of results memories in order to control the processors.

Although in the embodiment described above the input speech parameters are described as being feature vectors, the present invention encompasses any type of parameterization of speech data as an input to the circuit. Such parameterization methods are well known in the art.

Although in the embodiment described above there is only one language model processor, more than one can be used and these can be used to operate in parallel for processing scores at word boundaries.

CLAIMS:

1. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

lexical memory means containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input means for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory means;

results memory means connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

control processor means for controlling said lexical tree processors to process lexical trees identified in said results memory means by performing parallel processing on a plurality of said lexical tree data structures.

2. A speech recognition circuit according to claim 1, wherein said lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

3. A speech recognition circuit according to claim 2, wherein said lexical memory means stores each lexical tree data structure as mono phone models of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

4. A speech recognition circuit according to any preceding claim, wherein each lexical tree processor is operative to compare the processed speech parameters with said lexical data or data derived from said lexical data to identify words as a word recognition event and to send information identifying the identified words to said results memory means as the processing results.

5. A speech recognition circuit according to claim 4, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

6. A speech recognition circuit according to claim 5, wherein each lexical tree processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and to output the modified processing result to said results memory means.

7. A speech recognition circuit according to any one of claims 4 to 6, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events.

8. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from said lexical data in said lexical memory means.

9. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to instruct a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory means.

10. A speech recognition circuit according to claim 9, wherein said history data includes an accumulated score for the recognition path.

11. A speech recognition circuit according to any preceding claim, wherein said control processor means is operative to pass a lexical tree identifier to a said lexical tree processor, said lexical tree processor is operative to read lexical data for the identified lexical tree from said lexical memory means and to process said speech parameters using said read lexical data.

12. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

13. A speech recognition circuit according to claim 12, wherein said lexical tree processors are operative to prune the processing of the lexical tree in dependence upon the determined scores.

14. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are operative to process more than one lexical tree at the same time.

15. A speech recognition circuit according to any one of claims 4 to 7, wherein said control processor means is operative to determine new lexical tree identifiers for storing in said results memory means for words identified in said results memory means.

16. A speech recognition circuit according to claim 15, wherein said control processor means is operative to prune the new lexical tree identifiers.

17. A speech recognition circuit according to claims 16, wherein each lexical tree processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

18. A speech recognition circuit according to any preceding claim, wherein said input means is arranged to receive said speech parameters as feature vectors.

19. A speech recognition circuit according to any one of claims 1 to 17, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

20. A speech recognition circuit according to any preceding claim, wherein said lexical tree processors are arranged in groups, said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said groups of lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

21. A speech recognition circuit according to any one of claims 1 to 19, wherein said lexical memory means comprises a plurality of partial lexical memory means, each partial lexical memory means being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor is operative to process said speech parameters using a said partial lexical memory means, and said control processor means is operative to control each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory means.

22. A speech recognition circuit according to any preceding claim, wherein said lexical memory means stores said lexical tree data structures as Hidden Markov Models, and said lexical tree processors are operative to perform the Viterbi search algorithm using each lexical tree data structure.

23. A speech recognition circuit comprising:

an input buffer for receiving processed speech parameters;

lexical memory containing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, the initial component of each lexical tree data structure being unique;

a plurality of lexical tree processors connected in parallel to said input buffer for processing the speech parameters in parallel to perform parallel lexical tree processing for word recognition by accessing said lexical data in said lexical memory;

results memory connected to said lexical tree processors for storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

a control processor for controlling said lexical tree processors to process lexical trees identified in said results memory by performing parallel processing on a plurality of said lexical tree data structures.

24. A speech recognition circuit according to claim 23, wherein said lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

25. A speech recognition circuit according to claim 24, wherein said lexical memory stores said lexical tree data structure as a mono phone model of words, and each lexical tree processor generates context dependant phone models of words dynamically for use in processing the speech parameters.

26. A speech recognition circuit according to any one of claims 23 to 25, wherein each lexical tree processor compares the processed speech parameters with the lexical data or data derived from the lexical data to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

27. A speech recognition circuit according to claim 26, including a language model processor for providing a language model output for modifying the processing results at a word recognition event by a said lexical tree processor.

28. A speech recognition circuit according to claim 27, wherein each lexical tree processor receives said language model output to modify the processing result at a word recognition event determined by said lexical tree processor and outputs the modified processing result to said results memory.

29. A speech recognition circuit according to any one of claims 26 to 28, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events.

30. A speech recognition circuit according to any one of claims 23 to 29, wherein said lexical tree processors process said speech parameters by comparing said speech parameters with said lexical data in said lexical memory.

31. A speech recognition circuit according to any one of claims 23 to 30, wherein said control processor instructs a said lexical tree processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree from said results memory.

32. A speech recognition circuit according to claim 31, wherein said history data includes an accumulated score for the recognition path.

33. A speech recognition circuit according to any one of claims 23 to 32, wherein said control processor passes a lexical tree identifier to a said lexical tree processor, said lexical tree processor reads lexical data for the identified lexical tree from said lexical memory and processes said speech parameters using said read lexical data.

34. A speech recognition circuit according to any one of claims 23 to 33, wherein said lexical tree processors determine and output scores in the processing results during the processing of said speech parameters.

35. A speech recognition circuit according to claim 34, wherein said lexical tree processors prune the processing of the lexical trees in dependence upon the determined scores.

36. A speech recognition circuit according to any one of claims 23 to 35, wherein said lexical tree processors process more than one lexical tree at the same time.

37. A speech recognition circuit according to any one of claims 26 to 29, wherein said control processor determines new lexical tree identifiers for storing in said results memory for words identified in said results memory.

38. A speech recognition circuit according to claim 37, wherein said control processor means is operative to prune the new lexical tree identifiers

39. A speech recognition circuit according to claims 38, wherein each lexical tree processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes said new lexical tree identifiers on the basis of the scores for the words stored in said results memory means.

40. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as feature vectors.

41. A speech recognition circuit according to any one of claims 23 to 39, wherein said input buffer receives said speech parameters as n phones, where n is an integer.

42. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical tree processors are arranged in groups, said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory being connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

43. A speech recognition circuit according to any one of claims 23 to 41, wherein said lexical memory comprises a plurality of partial lexical memories, each partial lexical memory is connected to one of said lexical tree processors and containing part of the lexical data, each lexical tree processor processes said speech parameters using a said partial lexical memory, and said control processor controls each lexical tree processor to process a lexical tree corresponding to partial lexical data in a corresponding partial lexical memory.

44. A speech recognition circuit according to any one of claims 23 to 43, wherein said lexical memory stores said lexical tree data structures as Hidden Markov Models, and each lexical tree processor performs the Viterbi search algorithm using a respective said lexical tree data structure.

45. A speech recognition circuit comprising:

input means for receiving processed speech parameters;

a plurality of lexical memory means containing in combination complete lexical data for word recognition, each lexical memory means containing part of said complete lexical data;

a plurality of processors connected in parallel to said input means for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory means;

control processor means for controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

receiving means for receiving the results of the processing of the speech parameters from said processors.

46. A speech recognition circuit according to claim 45, wherein each lexical memory means stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

47. A speech recognition circuit according to claim 46, wherein each lexical memory means stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

48. A speech recognition circuit according to claim 47, wherein each lexical memory means stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors are arranged to use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

49. A speech recognition circuit according to any one of claims 45 to 48, wherein each processor is operative to compare the processed speech parameters with the lexical data in a corresponding lexical memory means to identify words as a word recognition event and to send information identifying the identified words to said receiving means as the processing results.

50. A speech recognition circuit according to claim 49, including language model means for providing a language model output for modifying the processing results at a word recognition event by a said processor.

51. A speech recognition circuit according to claim 50, wherein each processor is operative to receive said language model output to modify the processing result at a word recognition event determined by said processor and to output the modified processing result to said receiving means.

52. A speech recognition circuit according to any one of claims 49 to 51, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events.

53. A speech recognition circuit according to any one of claims 45 to 52, wherein said processors are operative to process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memory means.

54. A speech recognition circuit according to any one of claims 46 to 48, wherein said control processor means is operative to instruct a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

55. A speech recognition circuit according to claim 54, wherein said history data includes an accumulated score for the recognition path.

56. A speech recognition circuit according to any one of claims 46 to 48, 54, or 55, wherein said control processor means is operative to pass a lexical tree identifier to a said processor, said processor is operative to read lexical data for the identified lexical tree from a corresponding said lexical memory means and to process said speech parameters using said read lexical data.

57. A speech recognition circuit according to any one of claims 45 to 56, wherein said processors are operative to determine and output scores in the processing results during the processing of said speech parameters.

58. A speech recognition circuit according to claim 57, wherein said processors are operative to prune the processing of the speech parameters in dependence upon the determined scores.

59. A speech recognition circuit according to any one of claims 46 to 48, or 54 to 56, wherein said processors are operative to use lexical data corresponding to more than one lexical tree at the same time.

60. A speech recognition circuit according to any one of claims 49 to 52, wherein said control processor means is operative to identify lexical data in said lexical memory means required for processing next words at word recognition events, and to control said processors to process the speech parameters using the identified lexical data.

61. A speech recognition circuit according to claim 60, wherein said control processor means is operative to prune the identified lexical data.

62. A speech recognition circuit according to claims 61, wherein each processor is operative to determine and output scores for words in the processing results at word recognition events, and said control processor means is operative to prune the identified lexical data on the basis of the scores for the words.

63. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as feature vectors.

64. A speech recognition circuit according to any one of claims 45 to 62, wherein said input means is arranged to receive said speech parameters as n phones, where n is an integer.

65. A speech recognition circuit according to any one of claims 45 to 64, wherein said control processor means is operative to control each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory means.

66. A speech recognition circuit according to any one of claims 45 to 65, wherein said receiving means comprises a memory for storing the results of processing.

67. A speech recognition circuit according to any one of claims 45 to 66, wherein said lexical memory means store said lexical data as Hidden Markov Models, and each processor is operative to performs the Viterbi search algorithm using a respective part of said lexical data.

68. A speech recognition circuit comprising:

an input buffer receiving processed speech parameters;

a plurality of lexical memories containing in combination complete lexical data for word recognition, each lexical memory containing part of said complete lexical data;

a plurality of processors connected in parallel to said input buffer for processing the speech parameters in parallel, said processors being arranged in groups of processors, each group of processors being connected to a lexical memory;

a control processor controlling each processor to process said speech parameters using partial lexical data read from a respective said lexical memory; and

a results memory storing the results of the processing of the speech parameters from said processors.

69. A speech recognition circuit according to claim 68, wherein each lexical memory stores lexical data as at least one lexical tree data structure, each lexical tree data structure comprises a model of words having common prefix components, and an initial component of each lexical tree data structure is unique.

70. A speech recognition circuit according to claim 69, wherein each lexical memory stores said lexical tree data structure as an n phone model of words, where n is an integer, and said components comprise n phones.

71. A speech recognition circuit according to claim 70, wherein each lexical memory stores said lexical tree data structure as a mono phone model of words, and said lexical tree processors use said mono phone models to generate context dependant phone models of words dynamically for use in processing the speech parameters.

72. A speech recognition circuit according to any one of claims 68 to 71, wherein each processor compares the processed speech parameters with the lexical data in a corresponding lexical memory to identify words as a word recognition event and sends information identifying the identified words to said results memory as the processing results.

73. A speech recognition circuit according to claim 72, a including language model processor providing a language model output for modifying the processing results at a word recognition event by a said processor.

74. A speech recognition circuit according to claim 73, wherein each processor receives said language model output to modify the processing result at a word recognition event determined by said processor and outputs the modified processing result to said results memory.

75. A speech recognition circuit according to any one of claims 73 to 74, wherein each processor determines and outputs scores for words in the processing results at word recognition events.

76. A speech recognition circuit according to any one of claims 68 to 75, wherein said processors process said speech parameters by comparing said speech parameters with said lexical data or data derived from lexical data in respective said lexical memories.

77. A speech recognition circuit according to any one of claims 69 to 71, wherein said control processor instructs a said processor to process a lexical tree by passing a lexical tree identifier for the lexical tree and history data for a recognition path associated with the lexical tree to said processor.

78. A speech recognition circuit according to claim 77, wherein said history data includes an accumulated score for the recognition path.

79. A speech recognition circuit according to any one of claims 69 to 71, 77, or 78, wherein said control processor passes a lexical tree identifier to a said processor, said processor reads lexical data for the identified lexical tree from a corresponding said lexical memory and processes said speech parameters using said read lexical data.

80. A speech recognition circuit according to any one of claims 68 to 79, wherein said processors determine and output scores in the processing results during the processing of said speech parameters.

81. A speech recognition circuit according to claim 80, wherein said processors prune the processing of the speech parameters in dependence upon the determined scores.

82. A speech recognition circuit according to any one of claims 69 to 71, or 77 to 79, wherein said processors use lexical data corresponding to more than one lexical tree at the same time.

83. A speech recognition circuit according to any one of claims 72 to 75, wherein said control processor identifies lexical data in said lexical memories required for processing next words at word recognition events, and controls said processors to process the speech parameters using the identified lexical data.

84. A speech recognition circuit according to claim 83, wherein said control processor prunes the identified lexical data.

85. A speech recognition circuit according to claims 84, wherein each processor determines and outputs scores for words in the processing results at word recognition events, and said control processor prunes the identified lexical data on the basis of the scores for the words.

86. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as feature vectors.

87. A speech recognition circuit according to any one of claims 68 to 85, wherein said input buffer is arranged to receive said speech parameters as n phones, where n is an integer.

88. A speech recognition circuit according to any one of claims 68 to 88, wherein said control processor controls each processor to process a lexical tree corresponding to partial lexical data in a corresponding said lexical memory.

89. A speech recognition circuit according to any one of claims 68 to 89, wherein said lexical memories store said lexical data as Hidden Markov Models, and each processor performs the Viterbi search algorithm using a respective part of said lexical data.

90. A speech recognition method comprising:

receiving processed speech parameters;

storing lexical data for word recognition, said lexical data comprising a plurality of lexical tree data structures, each lexical tree data structure comprising a model of words having common prefix components, an initial component of each lexical tree data structure being unique;;

passing the speech parameters in parallel to a plurality of lexical tree processors to perform parallel lexical tree processing for word recognition by accessing said lexical data;

storing processing results from said lexical tree processors and lexical tree identifiers to identify lexical trees to be processed by said lexical tree processors; and

controlling said lexical tree processors to process lexical trees identified by the stored identifiers by performing parallel processing on a plurality of said lexical tree data structures.

91. A speech recognition method comprising:

receiving processed speech parameters;

storing complete lexical data for word recognition in a plurality of lexical memory means, each lexical memory means containing part of said complete lexical data;

processing the speech parameters in parallel using a plurality of processors connected in parallel, said processors being arranged in groups of processors, each group of processors being connected to a said lexical memory means;

controlling each processor to process said speech parameters using partial lexical data read from a said lexical memory means; and

storing the results of the processing of the speech parameters from said processors.

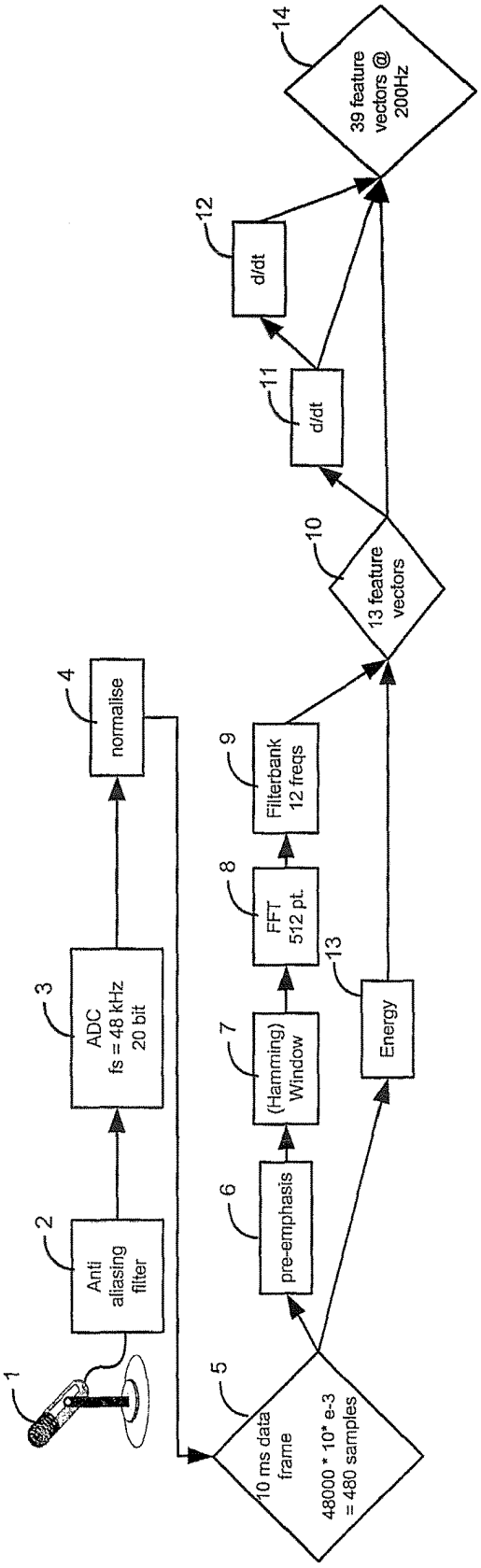

1
2
Anti aliasing filter
3
ADC fs = 48 kHz 20 bit
4
normalise
5
10 ms data frame 48000 * 10* e-3 = 480 samples
6
pre-emphasis
7
(Hamming) Window
8
FFT 512 pt.
9
Filterbank 12 freqs
10
13 feature vectors
11
d/dt
12
d/dt
13
Energy
14
39 feature vectors @ 200Hz

Fig 1

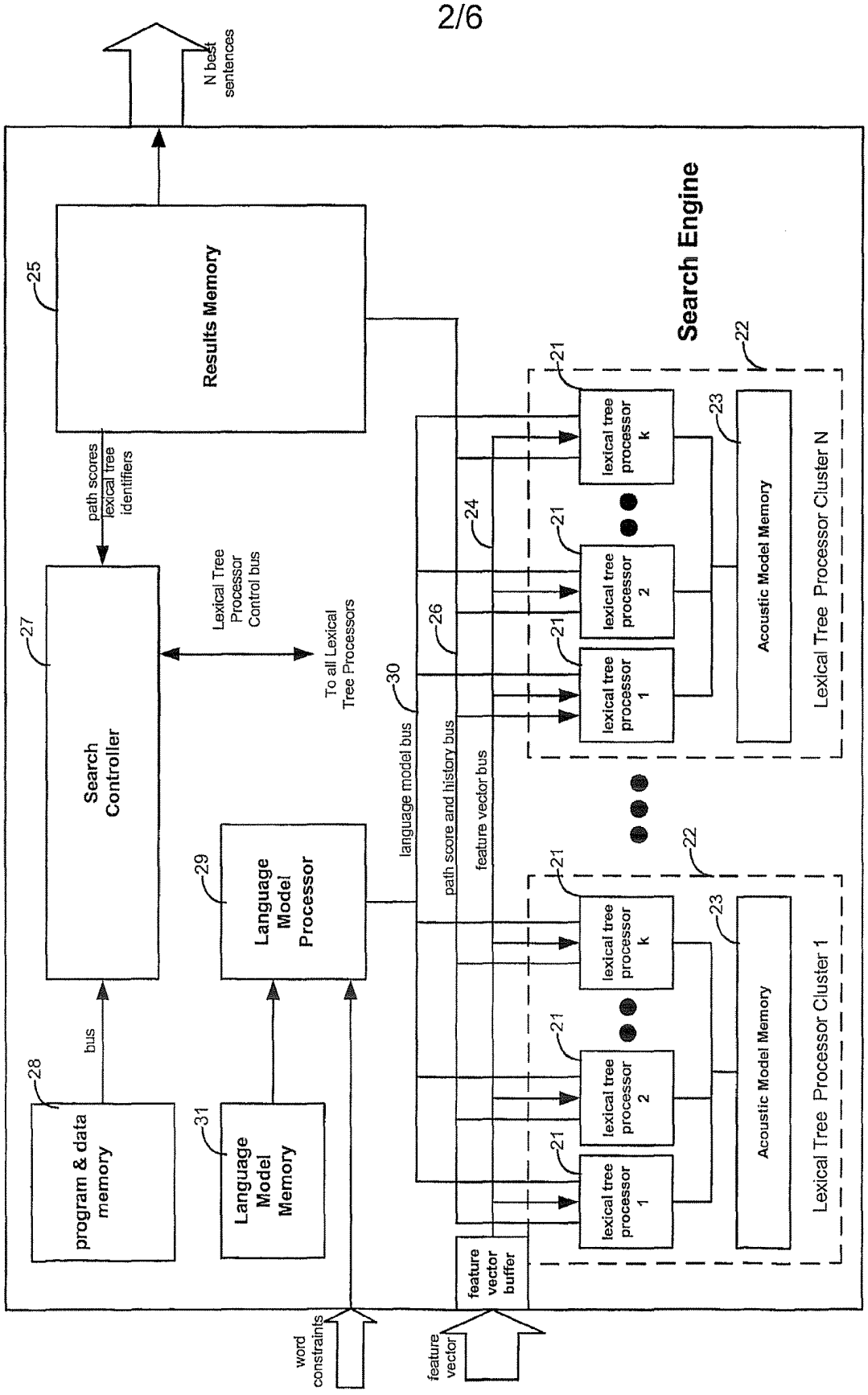

N best sentences
Search Engine
Results Memory
25
path scores lexical tree identifiers
Search Controller
27
Lexical Tree Processor Control bus
To all Lexical Tree Processors
Language Model Processor
29
Language Model Memory
31
program & data memory
28
bus
language model bus
30
path score and history bus
feature vector bus
24
26
feature vector buffer
word constraints
feature vector
lexical tree processor 1
lexical tree processor 2
lexical tree processor k
21
Acoustic Model Memory
23
Lexical Tree Processor Cluster 1
Lexical Tree Processor Cluster N
22

Fig 2

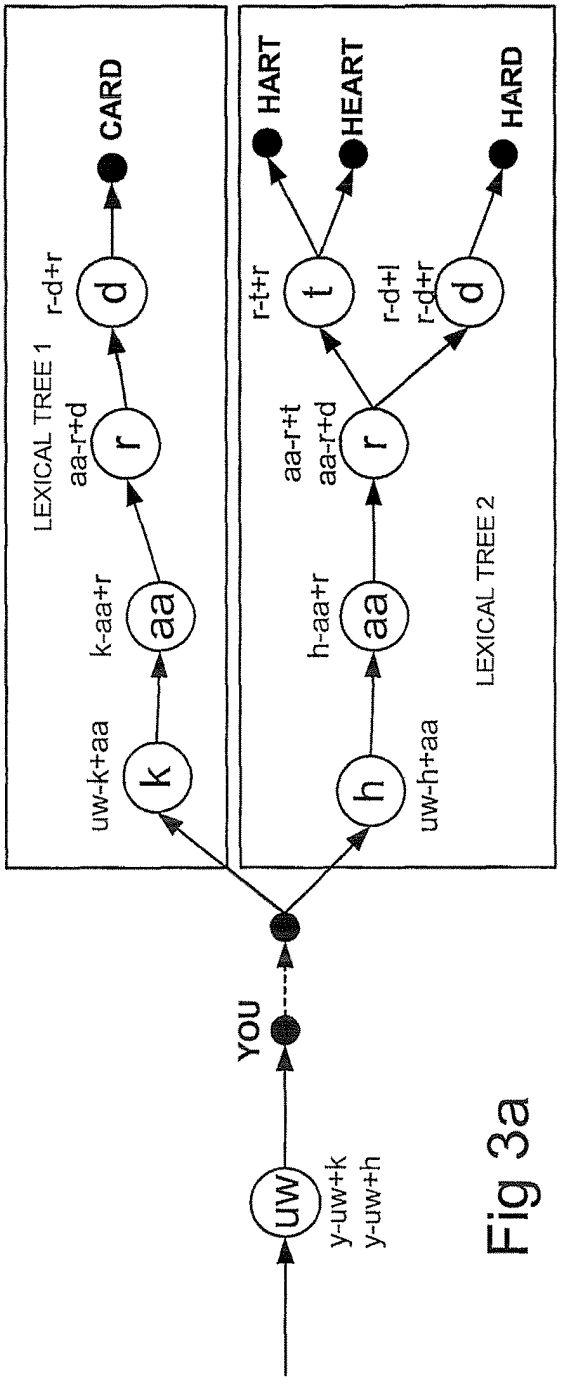

LEXICAL TREE 1
uw-k+aa
k-aa+r
aa-r+d
r-d+r
k
aa
r
d
CARD
LEXICAL TREE 2
uw-h+aa
h-aa+r
aa-r+t
aa-r+d
r-t+r
r-d+l
r-d+r
h
aa
r
t
d
HART
HEART
HARD
YOU
uw
y-uw+k
y-uw+h

Fig 3a

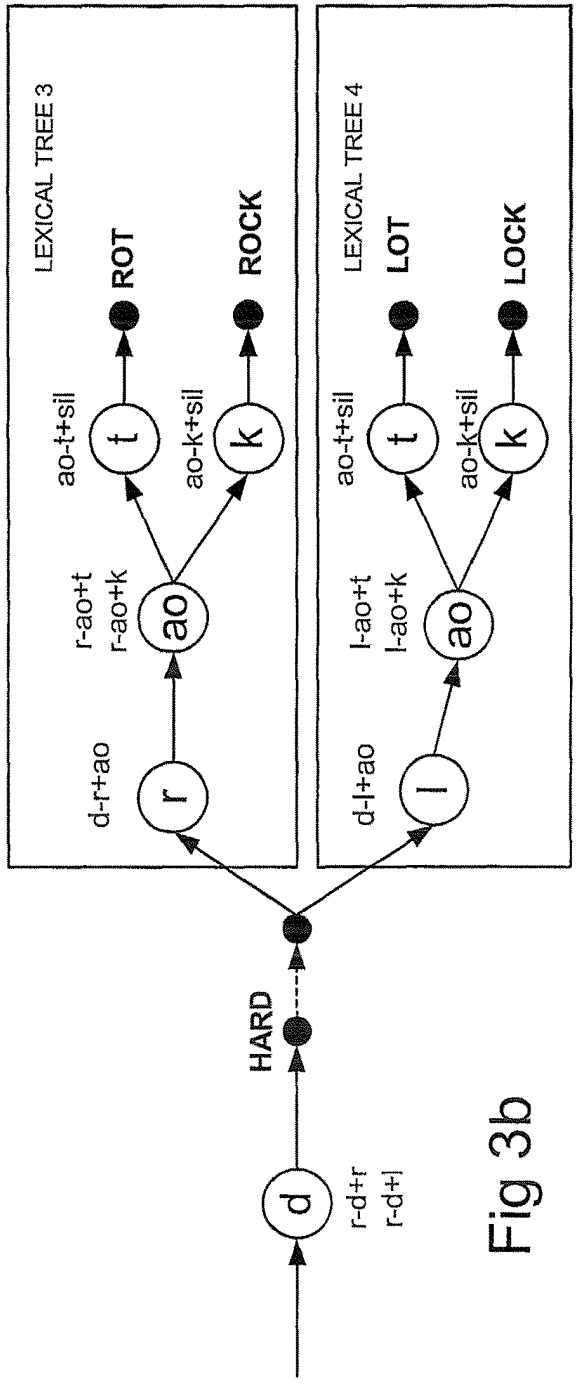

LEXICAL TREE 3
d-r+ao
r-ao+t
r-ao+k
ao-t+sil
ao-k+sil
r
ao
t
k
ROT
ROCK
LEXICAL TREE 4
d-l+ao
l-ao+t
l-ao+k
ao-t+sil
ao-k+sil
l
ao
t
k
LOT
LOCK
HARD
d
r-d+r
r-d+l

Fig 3b

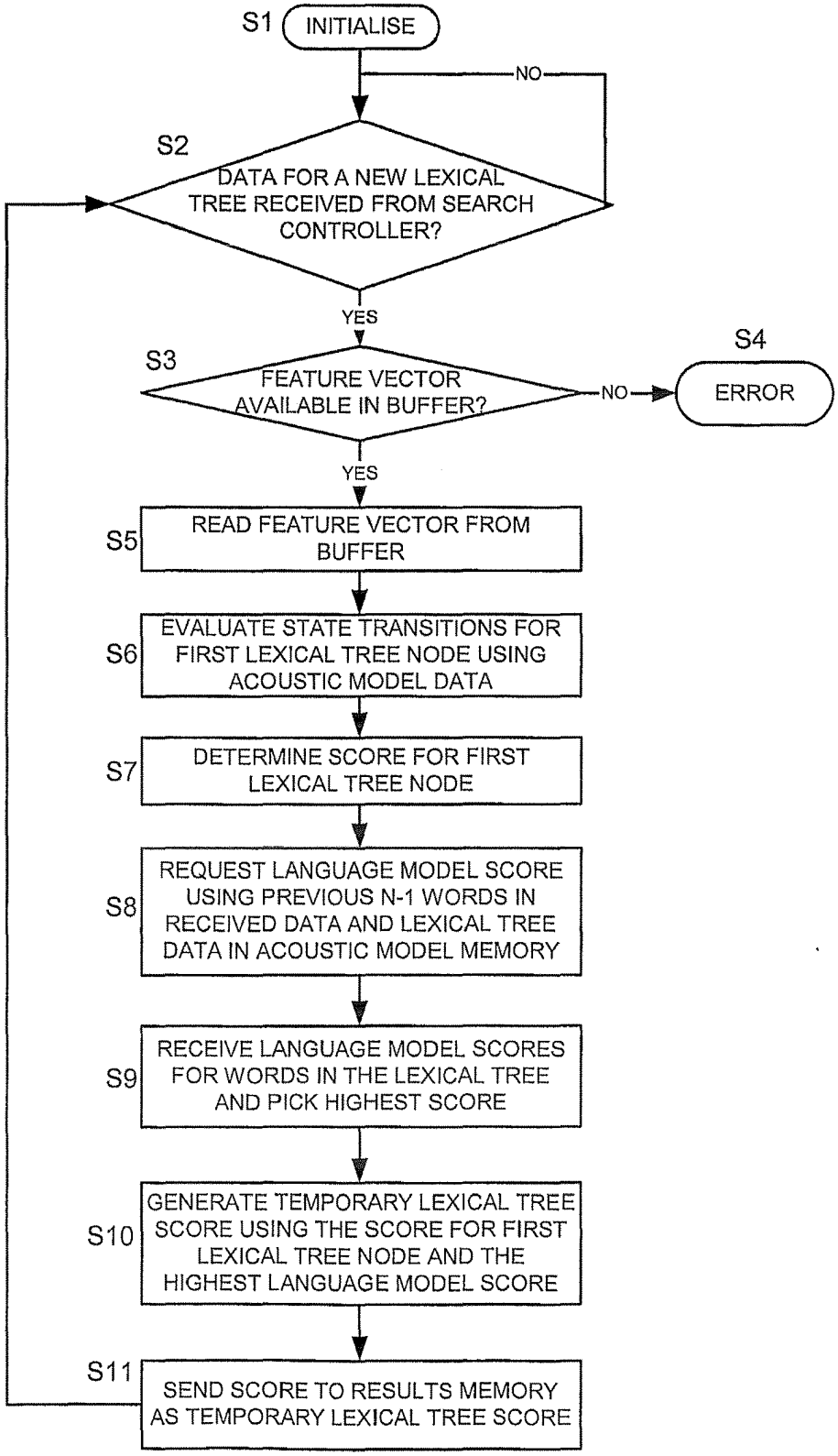

S1
INITIALISE
NO
S2
DATA FOR A NEW LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S3
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S4
ERROR
YES
S5
READ FEATURE VECTOR FROM BUFFER
S6
EVALUATE STATE TRANSITIONS FOR FIRST LEXICAL TREE NODE USING ACOUSTIC MODEL DATA
S7
DETERMINE SCORE FOR FIRST LEXICAL TREE NODE
S8
REQUEST LANGUAGE MODEL SCORE USING PREVIOUS N-1 WORDS IN RECEIVED DATA AND LEXICAL TREE DATA IN ACOUSTIC MODEL MEMORY
S9
RECEIVE LANGUAGE MODEL SCORES FOR WORDS IN THE LEXICAL TREE AND PICK HIGHEST SCORE
S10
GENERATE TEMPORARY LEXICAL TREE SCORE USING THE SCORE FOR FIRST LEXICAL TREE NODE AND THE HIGHEST LANGUAGE MODEL SCORE
S11
SEND SCORE TO RESULTS MEMORY AS TEMPORARY LEXICAL TREE SCORE

Fig 4

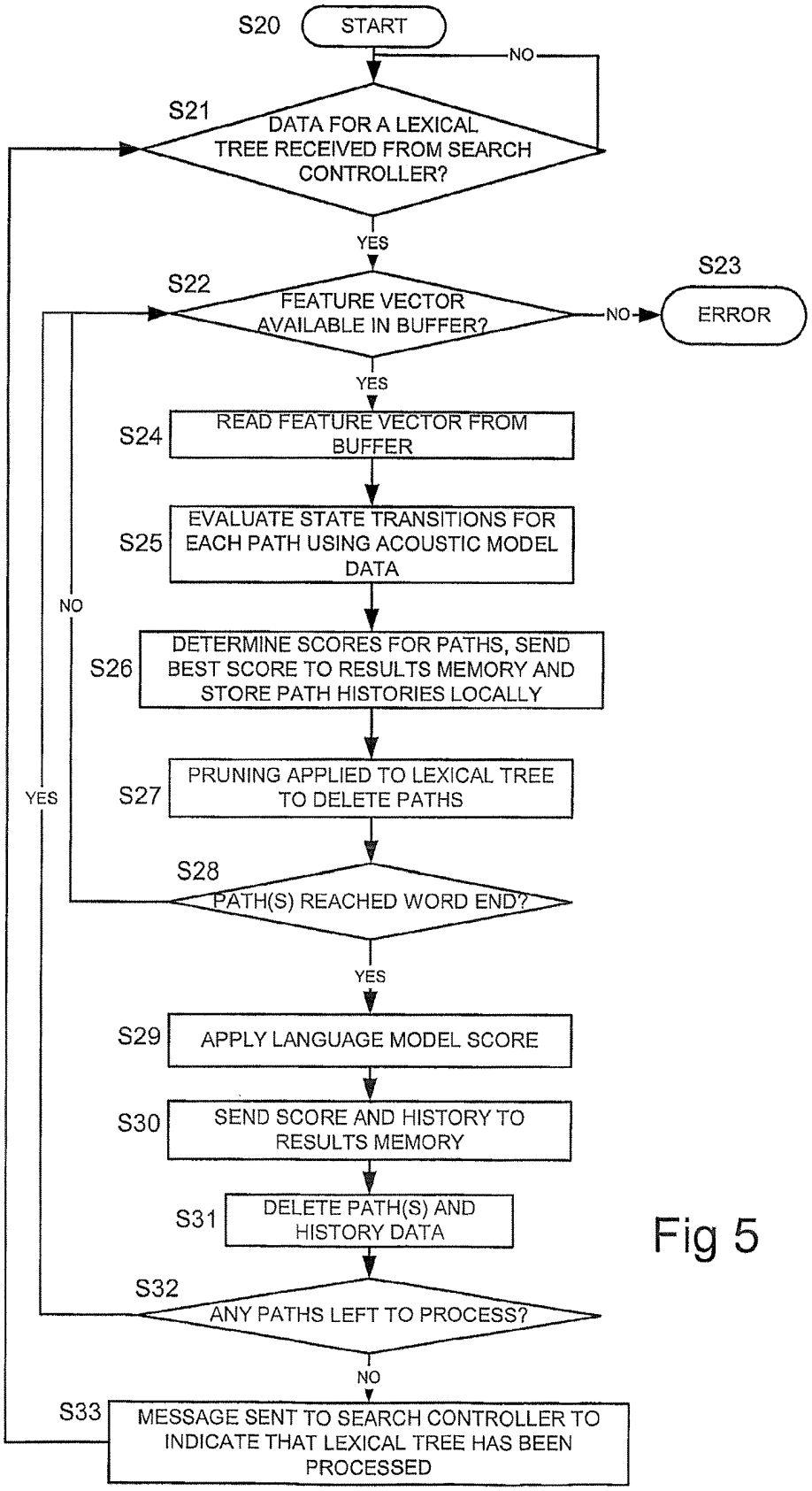

S20
START
NO
S21
DATA FOR A LEXICAL TREE RECEIVED FROM SEARCH CONTROLLER?
YES
S22
FEATURE VECTOR AVAILABLE IN BUFFER?
NO
S23
ERROR
YES
S24
READ FEATURE VECTOR FROM BUFFER
S25
EVALUATE STATE TRANSITIONS FOR EACH PATH USING ACOUSTIC MODEL DATA
NO
S26
DETERMINE SCORES FOR PATHS, SEND BEST SCORE TO RESULTS MEMORY AND STORE PATH HISTORIES LOCALLY
S27
PRUNING APPLIED TO LEXICAL TREE TO DELETE PATHS
YES
S28
PATH(S) REACHED WORD END?
YES
S29
APPLY LANGUAGE MODEL SCORE
S30
SEND SCORE AND HISTORY TO RESULTS MEMORY
S31
DELETE PATH(S) AND HISTORY DATA
S32
ANY PATHS LEFT TO PROCESS?
NO
S33
MESSAGE SENT TO SEARCH CONTROLLER TO INDICATE THAT LEXICAL TREE HAS BEEN PROCESSED

Fig 5

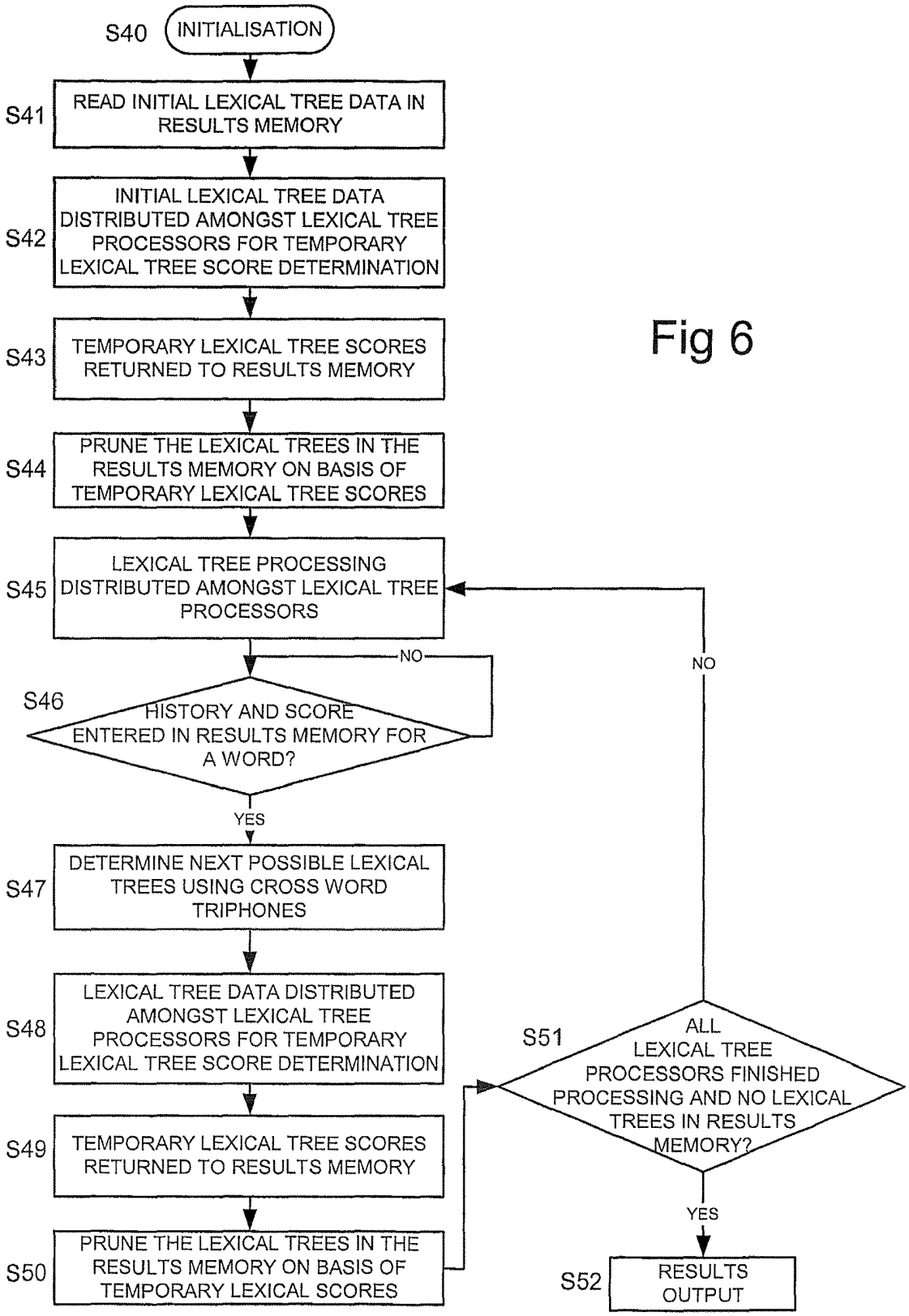

S40
INITIALISATION
S41
READ INITIAL LEXICAL TREE DATA IN RESULTS MEMORY
S42
INITIAL LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S43
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S44
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL TREE SCORES
S45
LEXICAL TREE PROCESSING DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS
S46
HISTORY AND SCORE ENTERED IN RESULTS MEMORY FOR A WORD?
NO
YES
S47
DETERMINE NEXT POSSIBLE LEXICAL TREES USING CROSS WORD TRIPHONES
S48
LEXICAL TREE DATA DISTRIBUTED AMONGST LEXICAL TREE PROCESSORS FOR TEMPORARY LEXICAL TREE SCORE DETERMINATION
S49
TEMPORARY LEXICAL TREE SCORES RETURNED TO RESULTS MEMORY
S50
PRUNE THE LEXICAL TREES IN THE RESULTS MEMORY ON BASIS OF TEMPORARY LEXICAL SCORES
S51
ALL LEXICAL TREE PROCESSORS FINISHED PROCESSING AND NO LEXICAL TREES IN RESULTS MEMORY?
NO
YES
S52
RESULTS OUTPUT

Fig 6

(12) UK Patent Application (19) GB (11) 2 331 392 (13) A

(43) Date of A Publication 19.05.1999

(21) Application No 9820182.5

(22) Date of Filing 17.09.1998

(30) Priority Data
(31) 08950621 (32) 15.10.1997 (33) US

(51) INT CL[6]
G10L 3/00

(52) UK CL (Edition Q)
G4R RRP R1F

(56) Documents Cited
EP 0533491 A2

(58) Field of Search
UK CL (Edition Q) G4R RHA RHB RRL RRM RRP
INT CL[6] G10L 3/00 5/00 5/06 9/00 9/18
Online:WPI

(71) Applicant(s)
International Business Machines Corporation
(Incorporated in USA - New York)
Armonk, New York 10504, United States of America

(72) Inventor(s)
Satyanarayana Dharanipragada
Ellen Marie Eide
Salim Estephan Roukos

(74) Agent and/or Address for Service
C J Ling
IBM United Kingdom Limited, Intellectual Property Department, Mail Point 110, Hursley Park, WINCHESTER, Hampshire, SO21 2JN, United Kingdom

(54) Abstract Title
Fast vocabulary-independent spotting of words in speech

(57) A fast vocabulary independent method for spotting words in speech utilizes a preprocessing step and a coarse-to-detailed search strategy for spotting a word/phone sequence in speech. The preprocessing includes a Viterbi-beam phone level decoding 14 using a tree-based phone language model. The coarse search 20 matches phone-ngrams to identify regions of speech as putative word hits, and the detailed search 22 performs an acoustic match at the putative hits with a model of the given word included in the vocabulary of the recognizer.

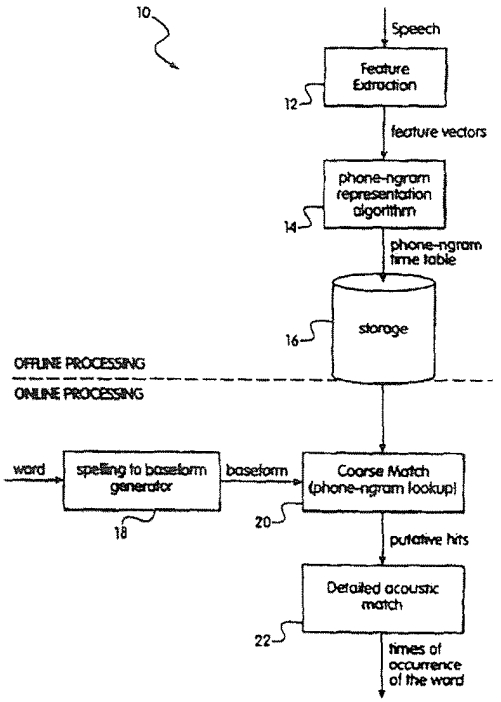


Fig. 1

GB 2 331 392 A

BNSDOCID: <GB___2331392A__I_>

2331392

A FAST VOCABULARY INDEPENDENT METHOD AND APPARATUS FOR SPOTTING WORDS IN SPEECH

Field of the Invention

The invention relates to speech recognition. More particularly, it relates to a method and apparatus for vocabulary independent wordspotting in speech.

Background of the Invention

In practical applications of wordspotting in speech recognition, such as, for example, audio-indexing, voice mail retrieval, spoken message retrieval and audio-browsing, it is necessary to have the ability to process large amounts of speech at speeds many times faster than real-time. The present wordspotting techniques are generally composed of three types:

1. A large vocabulary continuous speech recognizer which is used to produce N best transcriptions of the speech. From the N-best lists, the a posteriori probability of the word in question, is estimated. If this probability exceeds a user-defined threshold, the given word is deemed present. *M Weintraub, "LVCSR Log-Likelihood Ratio Scoring For Keyword Spotting", ICASSP 1995, Vol 1, pp 297-300.*

2. Building a detailed acoustic model for background speech, and using it in parallel with a detailed model for the given word to compute the *a posteriori* probability of the word. If this probability exceeds a user-defined threshold, the given word is deemed present. *J. R. Rohlicek, W. Russel, S. Roukos, H. Gish. "Word Spotting", ICASSP 1989, pp 627-630.*

3. Speech is pre-processed and stored as a phone lattice by running a modified Viterbi decoder on null-grammar phone network. The presence of a given word is determined by conducting a dynamic programming search on the phone lattice. *D. A. James, S. J. Young. "A Fast Lattice-Based Approach To Vocabulary Independent Wordspotting", ICASSP 1994, pp 3377-380.*

IPR2023-00033
Apple EX1002 Page 264

All of the above methods have their shortcomings. In the first method, wordspotting essentially reduces to searching through text. Consequently, retrieval is fast. However, words that do not appear in the vocabulary of the speech recognizer cannot be spotted. The second method has no limitations on the words that can be searched for, but is very slow since it requires re-running the wordspotter every time a new word is specified. The third method has both the flexibility of being able to search for any word and speed of retrieval. However, it relies heavily on phone recognition accuracy which is often very poor.

Disclosure of the Invention

The method according to an illustrative embodiment proposes a three-step procedure using a preprocessing step and a two-stage search strategy. In the preprocessing step, the speech waveform is converted into a representation consisting of a table of phone-ngrams with the times at which the phone-ngrams occur with a high likelihood. This representation allows searching through the speech very efficiently. The subsequent two stage search consists of a first coarse match step where a phone-ngram lookup is implemented to narrow down the time intervals where the word was likely to have been uttered, and then a detailed acoustic match at these time intervals to finally decide more accurately whether the word was actually uttered in that time interval.

The proposed method for wordspotting has both speed of retrieval and the flexibility of being able to search for any word. In order to obtain an increased speed of retrieval, the regions of speech where the word could have been uttered are narrowed down through the coarse match. By imposing a strong language model in the Viterbi phone decoder in a preprocessing state, and by running the detailed acoustic match with context dependent acoustic models at the regions of speech determined by the coarse match, a higher degree of accuracy is obtained.

The method according to an illustrative embodiment of the invention can be used in applications such as voice mail retrieval, and browsing and searching audio/video content. In these applications, the user types a search query, and a first determination is made as to the words that are out of the vocabulary of the speech recognizer used in transcribing the speech content. For each of these words, the phonetic baseforms are determined using either a large dictionary of baseforms or a spelling to baseform generator. The method of the invention is then employed to

determine the locations of these words in the audio/video database. These locations and the number of occurrences are then used in the scoring functions of a search engine (e.g., information retrieval algorithm) to retrieve the segments of audio/video that are relevant to the query.

Brief Description of the Drawings

The invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

FIG. 1 is a high level block diagram of the wordspotting apparatus according to an illustrative embodiment of the invention;

FIG. 2 is a block diagram schematically representing the preprocessing stage;

FIG. 3 is a flow chart diagram of the preprocessing method according to an illustrative embodiment of the invention; and

FIG. 4 is a representation of the filler model and network used in the detailed acoustic match stage of the presented method.

Detailed Description of the Invention

Referring to FIG. 1, the wordspotting apparatus 10 of the invention is shown. The apparatus includes an offline preprocessing stage that consists of converting the speech waveform into a table consisting of times of occurrence of phone-ngrams (usually triphones) in the given speech. The phone-ngram time-table is obtained by performing a feature extraction 12 to the input speech waveform to produce feature vectors f_i, which are then processed by a phone n-gram representation algorithm 14. The algorithm performs a time-synchronous Viterbi-beam search which uses a phone language model organized as a tree to constrain the dynamic programming search. Algorithm 14 produces the phone-ngram time table which is stored in a storage device 16. Storage device 16 can be any suitably known storage device, such as, for example, a random access memory, a magnetic disk or tape, an optical disk, etc.

The online processing stage performs a coarse match 20 of the phone-ngram time table previously stored. The coarse match is a phone-ngram lookup step, and receives a baseform from a spelling to

baseform generator 18. The putative hits determined in the coarse match step are then processed in a detailed acoustic match step 22. The detailed acoustic match will provide the different times of occurrence of an identified word.

Referring to FIGS. 2 and 3, the phone-ngram algorithm 14 arranges the phonetic baseforms of the words in a large vocabulary speech recognizer in the form of a tree 24 (i.e., the fast match tree). The tree representation is then converted into a graph capable of representing arbitrary sequences of words by adding arcs which go from each leaf back to the root node. Transition probabilities are assigned to each arc as 1/N where N is the number of arcs leaving the source node. A self-loop probability of 0.5 is also assumed for each node. This graph is used to constrain a dynamic programming search. This conversion is known in the art as trellis construction and can be found disclosed in U. S. Patent application serial number 08/798,011.

Phones attached to each arc of the fast match tree are represented by a context independent single state Hidden Markov Model (HMM) to reduce computation. In order to maintain the desired three-frame minimum duration, transitions only every third frame are allowed. When a single state topology is used, each node of the fast match tree directly corresponds to a state in the trellis. The trellis is then input into the preprocessing stage 26.

FIG. 3 shows the processing performed in the phone n-gram representation algorithm 14. The $s_j(t)$ is calculated for each phone j (step 32) from the feature vectors f_t. Where

$$S_j(t) = p\,(f_{3t}|j)p(f_{3t-1}|j)p(f_{3t-2}|j) \qquad (1)$$

and t = 0, 1, ...T/3. The probabilities $p(f_k|j)$ are computed as the maximum probability of all leaves of the phone, thus providing context-independence. At t=0, all start nodes are placed in the active list A where score (i) = 1.0 $\forall i \in A$ (step 34). A modified Viterbi search is used to find the best path through the trellis, where nodes in time correspond to every third frame of utterance. *(See, G. D. Forney, Jr. "The Viterbi Algorithm," Proc. IEEE, vol 61, pp 268-278, 1973).* Thus, at step 36 a determination is made as to whether t = T/3. If it does, the phone-ngram is stored in the storage device 16. If $t \neq T/3$, the following steps 38-44 are performed.

At every third frame, the scores of all the active nodes are updated and the maximum score, M, is computed (step 38). Thus, at step 38, the following determinations are made: score (i) * = $s_{\phi_i}(t)$ $\forall i \in A$, where ϕ_i is the phone-id at node i, and M = $\max_{i \in A}$ score(i). In step 40, for each i ε A for which score (i) > M/Beam, a list of all successor nodes succ(i) is determined. Here M is determined in step 38 and Beam is a user defined parameter. Next for each j ε succ(i), score (j) is updated as score (j) = score (i) * $t_{i,j}$, where $t_{i,j}$ is the transition probability between node i to node j. All the successors succ(i) are placed in the next list N. At step 42, the phone ngrams produced by the top K nodes in A are determined.

Once step 42 is complete, t is incremented by one, i.e., t = t + 1, A = N, and the N list is emptied. At this point, the determination as to whether t = T/3 is performed again at step 36. The entire process is repeated until there are no more phone ngrams produced by the nodes.

Each node in the trellis represents a sub-word or a phone sequence which is to be determined by traversing the tree from the root node to that node. The sub-word or phone sequence can also be described by the triphones it generates. The active nodes thus signify the top triphones that best describe that local region of speech. These triphones are then stored. To reduce storage, the time axis is discretized into T second intervals, where T is a user defined parameter, and all active triphones are binned into these intervals. Thus, a table of triphones versus their times of occurrence in the given speech, is generated, e.g. if the time axis is binned in 1 second intervals and if triphones N EH DX appeared in the top K nodes in the time interval [21,22] and [31,32], EH DX AX appeared in the top K nodes in the time interval [21,22] and [51,52] etc, then the table appears as:

triphone	time	(secs)
N EH DX	21	31
EH DX AX	21	51
DX AX N	31	40

Coarse Search for Putative Hits

The first step of the online processing is the coarse match step 20 (FIG. 1). Given a keyword, its phonetic-transcription (baseform) is determined using either a large dictionary of baseforms or a spelling to

IPR2023-00033
Apple EX1002 Page 268

baseform generation mechanism 18. From the baseform, all the triphones the baseform generates can be determined. Let N_p be the number of these triphones. For example, if the phonetic baseform for the word "Netanyahu" is N EH DX AX N Y AA HU W then $N_p = 7$, where the 7 triphones are N EH DX, EH DX AX, DX AX N, A XN Y, N YA A, Y AA H, AA HU W. For each of these triphones we look-up, from the phone-ngram table generated in the preprocessing step, the time where the triphone was found in the speech. When more than a certain user-defined fraction, r, of the total number of triphones are found in a given time interval, that time-interval is classified as a putative hit. The fraction r controls the number of false-alarms and detection rate. By reducing r the number of false alarms as well as the detection rate increases. When r is increased the number of false alarms and detection rate is reduced. Thus, by varying r we get a Receiver operating characteristic (ROC) curve for the wordspotter at the coarse search stage.

Detailed Match at the Putative Hits

In order to reduce the number of false alarms a more detailed acoustic match 22 can be performed at the regions classified as putative-hits in the previous coarse match stage. During this step, a detailed model using a 'right and left within word context' is built for the given word. With this model of the word included in the vocabulary, a modified Viterbi-search or an A* search is conducted at the putative hits. In the modified Viterbi search scheme, all words other than the given word represent the background speech or filler model. The *a posteriori* probability of the end state of the given word is then computed. If this probability exceeds a user-defined threshold, it is classified as a hit. Once again, this threshold controls the number of false alarms and the detection rate.

In order to speed up the detailed acoustic match 22, a simpler two step approach is employed. First a network consisting of "mumble" or "filler" models and all the alternative baseforms is constructed as shown in FIG. 4. For each putative hit, the best path through this network is computed using a Viterbi algorithm with traceback and the start and end times, t_s, t_e, for the word are determined. Next, the filler model and each of the alternative baseforms are scored between the start and end-time determined in the previous step, and the duration normalized log-likelihood ratio (DNLLR) is computer as follows:

IPR2023-00033
Apple EX1002 Page 269

$$DNLLR = \frac{logScore(S_e, t_e) - logScore(S_o, t_s)}{t_e - t_s},$$

where LogScore(s,t) is the viterbi score of state s at time t and S_o, S_e are start and exit states respectively of the word HMM. All putative hits are ranked on the best DNLLR and the top N putative hits are classified as hits.

All steps described in the illustrative embodiments of the invention can be performed on a general purpose computer being programmed accordingly.

The presented embodiment has been described with respect to a single state context independent HMM to model a phone in the preprocessing stage. However, since computational expense is not a primary issue for preprocessing, a more detailed three-state context dependent HMM for a phone can be used. This would improve phone recognition accuracy at the expense of more computation and would also reduce the number of false alarms in the coarse match stage.

As an alternative to the two stage search for the word, the speech can be pre-processed in the same way as before but the results can be scored as a phone lattice. For retrieval, a dynamic programming match of the given word against the lattice is performed.

CLAIMS

1. A fast vocabulary independent method for spotting words in speech comprising the steps of:

converting (14) a speech waveform into a table of phone-ngrams;

performing (20) a coarse match to narrow time intervals of speech occurrence; and

performing (22) a detailed acoustic match at the narrowed time intervals.

2. A method as claimed in claim 1, wherein said step of converting (14) further comprises the steps of:

generating (12) feature vectors representative of input speech;

generating (14) a phone-ngram time table from the feature vectors; and

storing (16) the phone-ngram time table in a memory device.

3. A method as claimed in any preceding claim, wherein said step of performing (20) a coarse match further comprises the steps of:

determining a phonetic transcription (baseform) of a given keyword;

determining all triphones generated by the phonetic transcription;

identifying the time interval where the triphone was found in the speech; and

classifying the identified time interval as a putative hit.

4. A method as claimed in any preceding claim, wherein said step of performing (22) a detailed acoustic match further comprises the steps of:

building a detailed model of a given word;

conducting a modified Viterbi search at the putative hits;

computing an *a posteriori* probability of the end state of the given word; and

outputting the times of occurrence of the given word.

5. A method as claimed in claim 2, wherein said step of generating (14) a phone-ngram time table further comprises the steps of:

constructing a phone language model from a phonetic baseforms vocabulary; and

performing a time synchronous Viterbi-Beam search using the phone language model.

6. A method as claimed in claim 5, wherein said step of performing a time synchronous Viterbi-Beam search further comprises the steps of:

computing (32) the maximum probability of all leaves of a phone;

placing (34) all start nodes in an active list "A";

determining the best path through the constructed phone language model;

updating the scores of all active nodes at every third frame;

computing (38) the maximum score "M";

determining (40) a list of all successor nodes;

placing (40) the successor nodes in a next list "N"; and

determining (42) the phone-ngrams produced by the nodes for the top K scores in A.

7. A method as claimed in claim 3, wherein said step of identifying the time interval further comprises the step of looking up the time interval on the generated phone-ngram time table.

8. A method as claimed in claim 3, further comprising the step of identifying a time interval where a user defined fraction of the total number of triphones are found.

9. A method as claimed in claim 1, wherein said step of converting (14) a speech waveform is performed offline as a preprocessing step, and said steps of performing (20) a coarse match and performing (22) a detailed acoustic match are performed online during processing.

10. A method as claimed in claim 1 wherein said step of converting is with the times at which an utterance of speech occurs with a high likelihood;

said step of performing a coarse match is by implementing a phone-ngram lookup; and

said step of performing a detailed acoustic match is to finally decide whether a word was actually uttered in that time interval.

11. A fast vocabulary independent method for spotting words in speech for use in voice mail retrieval systems and browsing and searching audio/video content, the method comprising the steps of:

receiving a search query from a user;

determining the words that are out of a vocabulary of a speech recognizer used in transcribing the speech content;

determining the phonetic baseforms for each of the previously determined words;

identifying the locations of these words in an audio/video database using a method as claimed in claim 1 or claim 2; and

retrieving segments of audio/video that are relevant to the received query.

12. A method as claimed in claim 11, wherein said step of determining the phonetic baseforms is performed using one of a group consisting of a large dictionary of baseforms and a spelling to baseform generator.

The Patent Office

INVESTOR IN PEOPLE

Application No:	GB 9820182.5	**Examiner:**	John Donaldson
Claims searched:	1 to 12	**Date of search:**	11 March 1999

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.Q): G4R(RHA, RHB, RRL, RRM, RRP)

Int Cl (Ed.6): G10L 3/00, 5/00, 5/06, 9/00, 9/18

Other: Online - WPI

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	EP 0533491 A2 (XEROX), see abstract	-

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

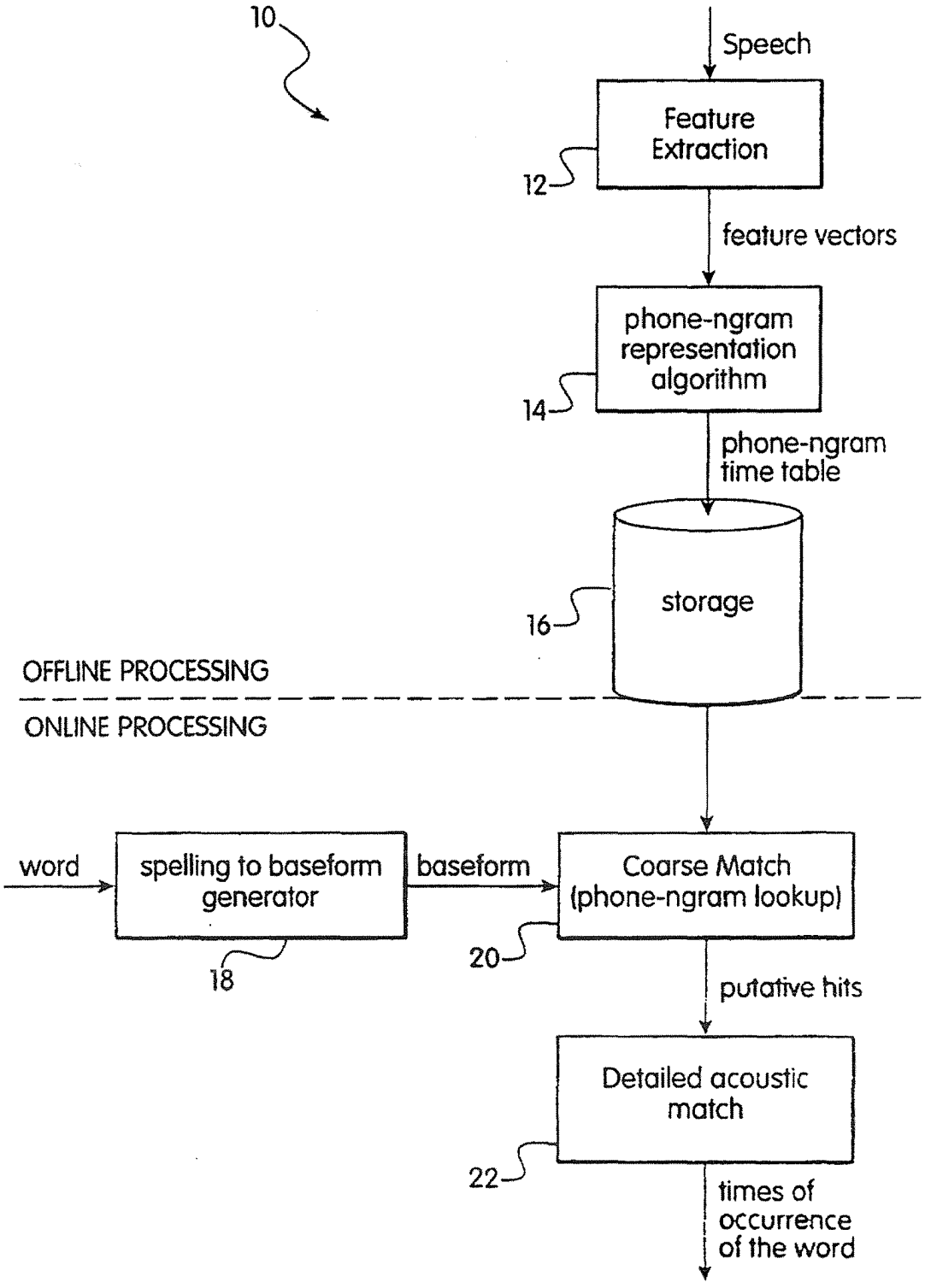

10
Speech
Feature Extraction
12
feature vectors
phone-ngram representation algorithm
14
phone-ngram time table
storage
16
OFFLINE PROCESSING
ONLINE PROCESSING
word
spelling to baseform generator
18
baseform
Coarse Match (phone-ngram lookup)
20
putative hits
Detailed acoustic match
22
times of occurrence of the word

Fig. 1

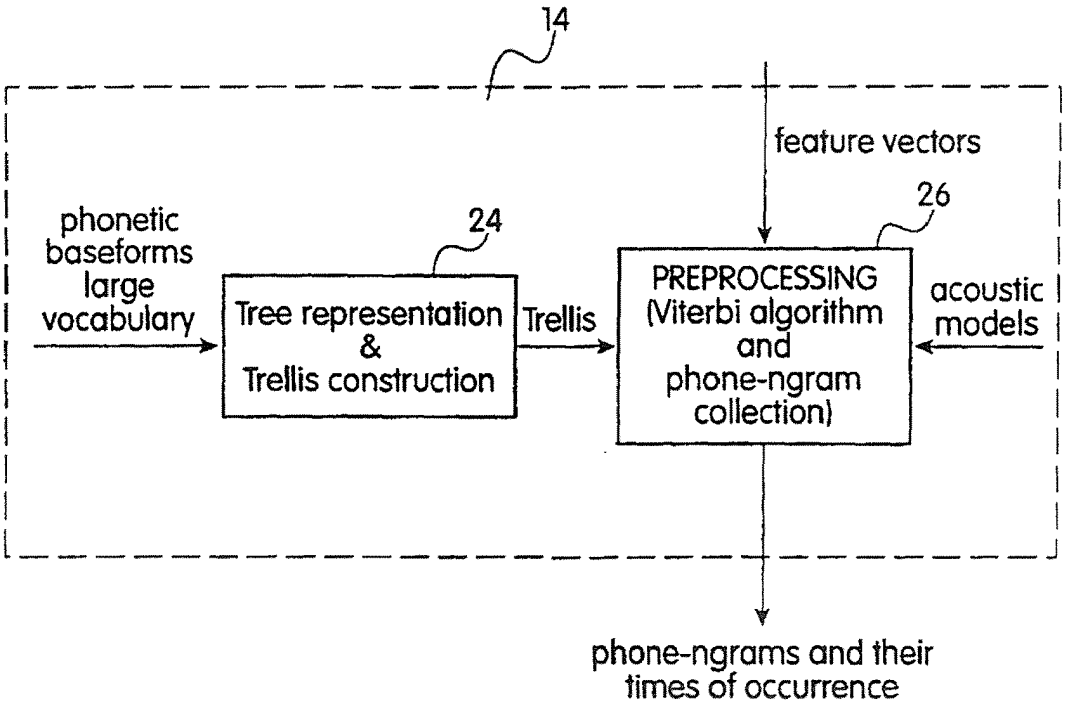


Fig. 2

BNSDOCID: <GB___2331392A__I_>

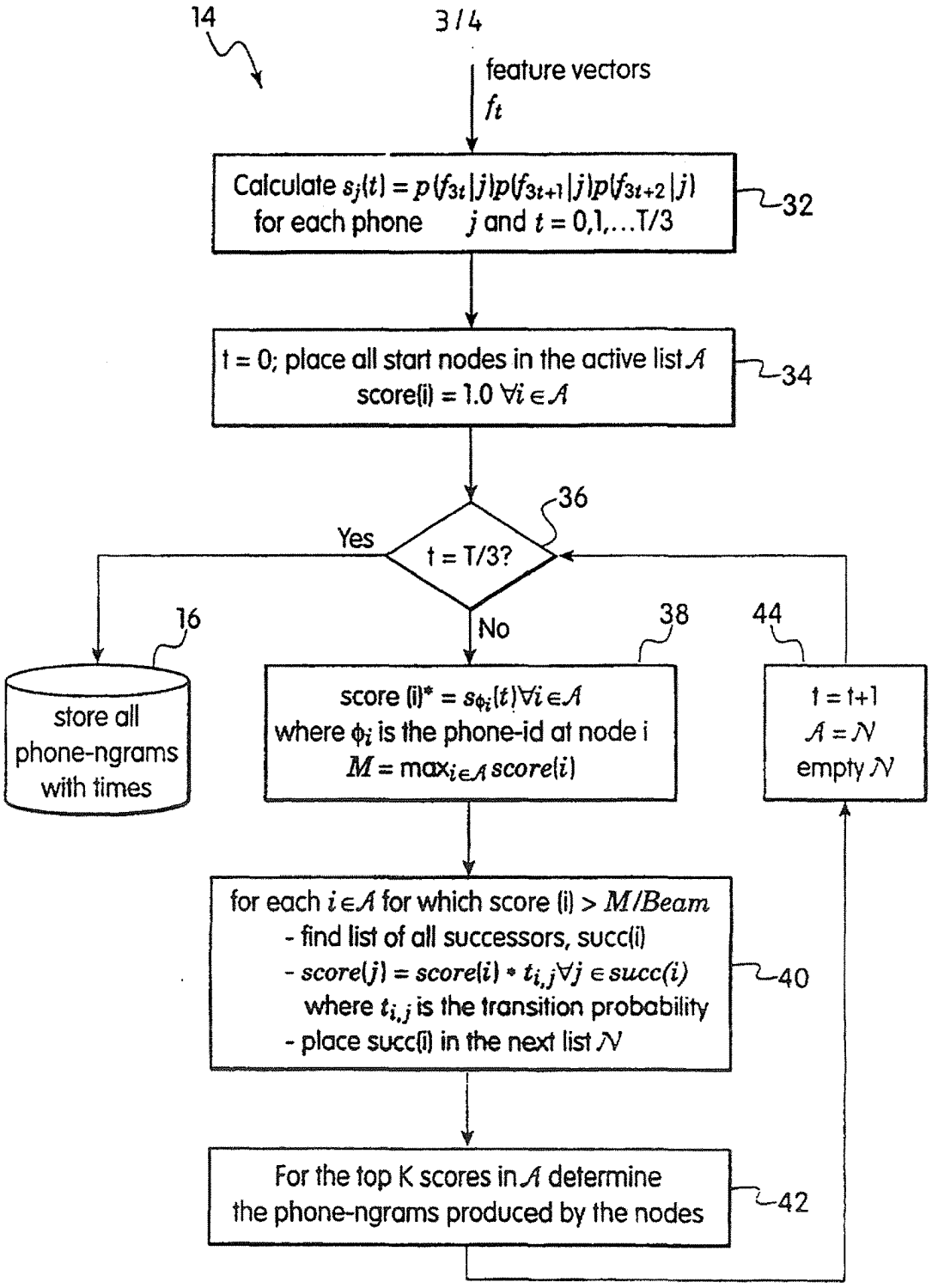
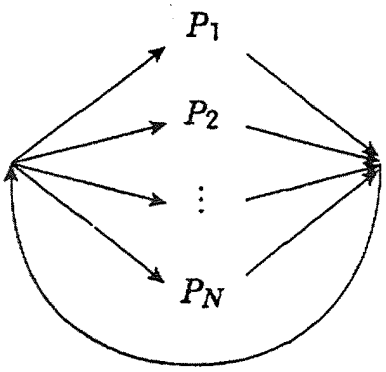


Fig. 3

BNSDOCID: <GB __2331392A__I_>



Filler or mumble model

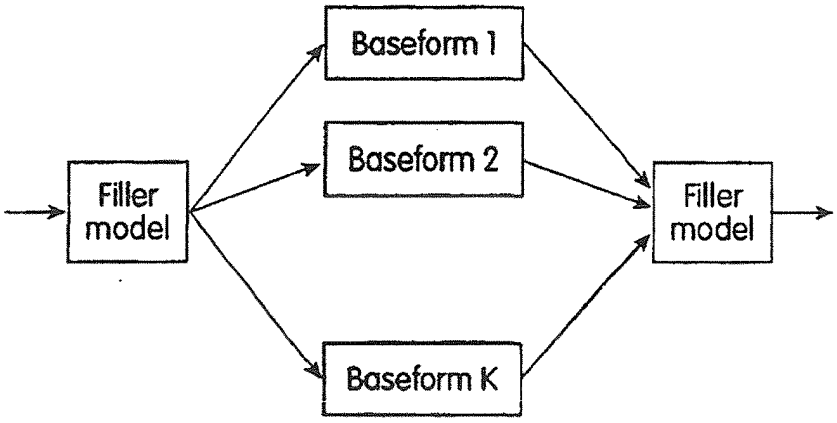


Fig. 4

(12) UK Patent Application (19) GB (11) 2 112 194 A

(21) Application No **8231240**
(22) Date of filing **2 Nov 1982**
(30) Priority data
(31) **8135939**
(32) **27 Nov 1981**
(33) **United Kingdom (GB)**
(43) Application published
13 Jul 1983
(51) **INT CL³**
G06K 9/68
(52) Domestic classification
G4R 10E 10X 11C 11D 12A 1F 1X 2F 4A3 5X 6B 7B 7F 8A 9B 9C 9D
U1S 1279 1288 1878 2271 2272 G4R
(56) Documents cited
None
(58) Field of search
G4R
(71) Applicant
National Research Development Corporation,
(Great Britain),
P.O. Box 236,
Kingsgate House,
66—74 Victoria Street,
London,
SW1E 6SL
(72) Inventors
Igor Aleksander,
Thomas John Stonham,
Bruce Armour Wilkie
(74) Agent and/or address for service
C. Hasler,
National Research Development Corporation,
101 Newington Causeway,
London,
SE1 6BU

(54) Recognition apparatus

(57) In image recognition systems of the type employing a frame store having groups (N-tuples) of locations mapped to storage forming discriminators trained to respond to different images, a problem arises in that either a large number of separate discriminator stores is required or recognition times are long. In the present invention the discriminator storage comprises p discrete stores 41, 42, 43 having address terminals (A) divided into two groups: one group (N) directly mapped to respective portions of the frame store and the other group (E and F) being connected in parallel and used to count through a number of N-tuple groups (each of size p) and/or a number of discriminators. Such a storage arrangement allows many different combinations of number of discriminator stores and recognition time, making possible an optimum combination for a particular application.

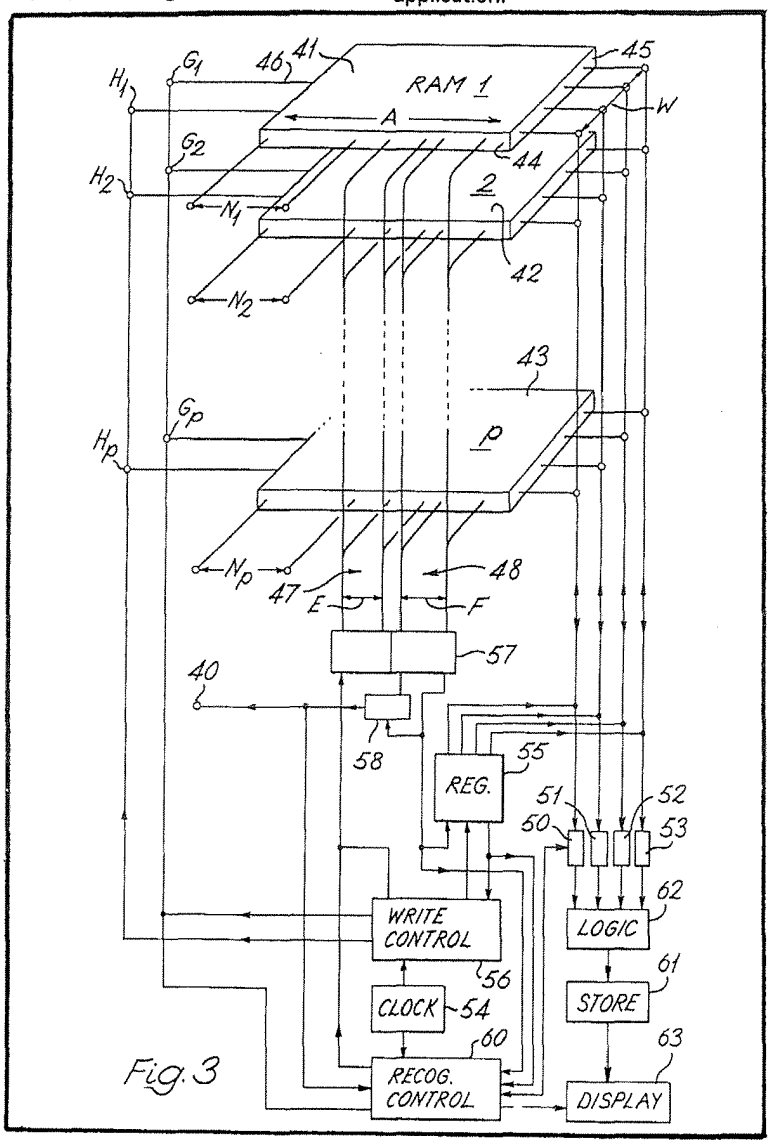


Fig. 3

GB 2 112 194 A

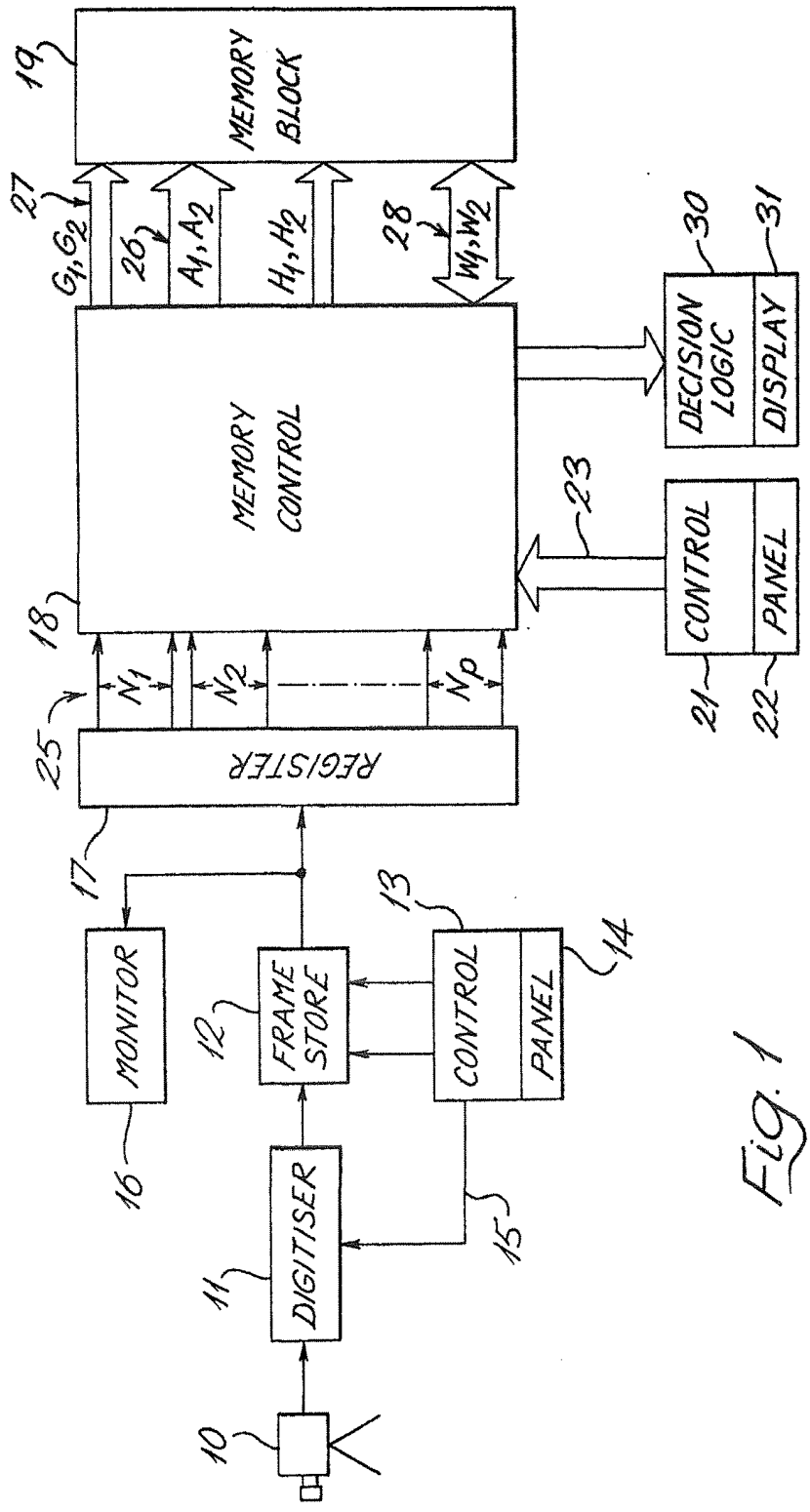

MEMORY BLOCK
19
27
G_1, G_2
26
A_1, A_2
H_1, H_2
28
W_1, W_2
MEMORY CONTROL
18
DECISION LOGIC
30
DISPLAY
31
CONTROL
PANEL
21
22
23
N_1
N_2
N_p
25
REGISTER
17
MONITOR
16
FRAME STORE
12
CONTROL
PANEL
13
14
DIGITISER
11
15
10

Fig. 1

2112194

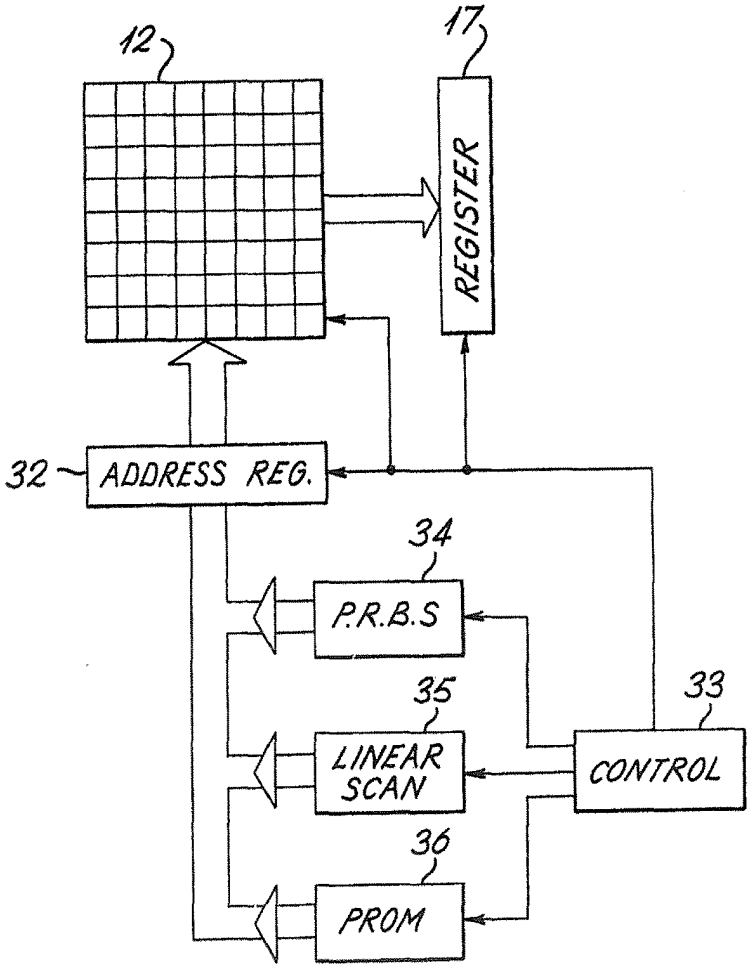


Fig. 2

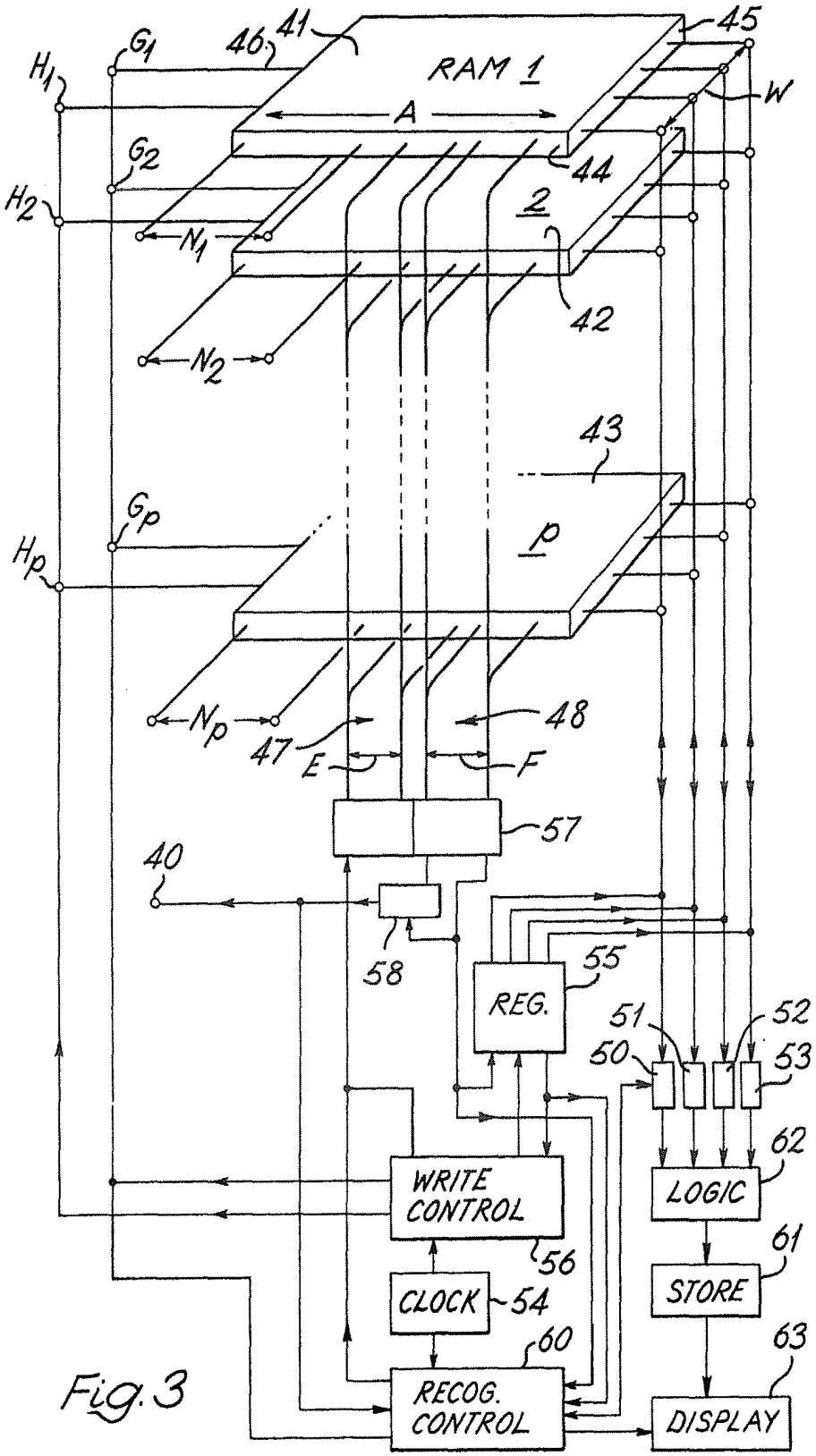

RAM 1
A
W
G1
H1
G2
H2
N1
N2
Gp
Hp
Np
1
2
p
41
42
43
44
45
46
47
48
E
F
40
57
58
55
REG.
51
52
53
50
62
LOGIC
61
STORE
63
DISPLAY
WRITE CONTROL
56
CLOCK
54
60
RECOG. CONTROL

Fig. 3

SPECIFICATION

Recognition apparatus

The present invention relates to an image recognition apparatus. An object or pattern to be recognised may for example be viewed by an optical system in forming the image. The apparatus usually requires two phases for operation: firstly a learning phase in which a plurality of patterns relating to those which are to be recognised are presented for example to the optical system and secondly a recognition phase in which patterns to be recognised are presented. However, in apparatus dedicated to a particular application, the learning phase may be omitted when the apparatus can be built to incorporate the results of the learning phase.

Pattern recognition systems employing random access memories (RAMs) in which the memory is divided into sections each requiring N address lines are known, these sections being known in this specification amd claims as N-tuples. Groups of N-tuples are called discriminators and in general there is one discriminator for each class of pattern to be recognised. For example if there are three patterns to be recognised then there will be three discriminators each containing a number of N-tuples and in general there is a value of N for which the capability of the apparatus to distinguish between patterns is highest.

The image to be recognised is held in a buffer store and each element in the store is connected to an address line of one RAM in each discriminator.

These concepts are described in the paper "Guide to pattern recognition using random-access memories" by I. Aleksander and T. J. Stonham, in Computers and Digital Techniques, February 1979, Vol. 2, No. 1.

According to the present invention there is provided apparatus for recognising images, comprising an input store having a plurality of storage locations allowing data representative of elements of an image which is to be recognised to be stored at locations related to the positions of the elements, a plurality of output stores each having a number of address terminals which allow every storage location in that store to be addressed, the locations of the output store either containing, or being capable of containing, data related to images which are to be recognised, the address terminals being divided into at least a lower significance group and a higher significance group, each group of terminals containing at least one terminal, mapping means with input connected to the input store and a plurality of output terminals divided into groups with each group corresponding to a respective output store and at least one of the terminals of that group connected to an address terminal of the lower significance group of the corresponding store, addressing means connected to the terminals of the higher significance groups for cyclically addressing groups of storage locations, and discriminator means responsive to output signals from groups of output-store storage locations, when containing data and when addressed by way of the mapping means and by the addressing means during recognition, to provide a measure of recognition.

An advantage of the present invention is that a parallel/series structure is provided for the output stores which form the storage for the discriminators. In a totally parallel storage structure such as is shown in the above mentioned paper a large number of discrete stores each with individual wiring are required and in a totally series storage structure where there is one bus-addressed store for use by all discriminators, times for recognition are comparatively long. The present arrangement allows structures between the totally parallel and totally serial so making possible an optimum arrangement for a particular application.

Each output store may be a single integrated-circuit but is usually a board carrying a number of such circuits. If the apparatus is to be able to carry out both learning and recognition the integrated circuits are conveniently random access memories (RAMs) but if only the recognition phase is to be carried out the integrated circuits may be read only memories (ROMs). Each integrated circuit may store either a bit or a byte at each storage location.

The mapping means may include means for generating a series of addresses for the input store and addressing the input store with the addresses generated, and a register for storing successive sequences of bits from the input store, the said output terminals of the mapping means being output terminals of respective stages of the register.

Each output store has at least one output terminal at which a signal appears corresponding to the contents of a location in that store which is addressed, and the discriminator means may include one response means for, and connected to, each said output-store output terminal, and operating means for receiving signals from the response means, and clearing the response means, at the end of each cycle of the addressing means and operating upon the signals so received and any previously received signals relating to the same image in providing the measure of recognition.

Apparatus according to the invention may include write means for storing data in the output stores when during a learning phase data corresponding to an image is present in the input store, the addressing means being constructed to generate at least one change-image signal in each cycle, each such signal being generated after a group of output stores has been addressed, and indicating that a change of image is required in the learning phase.

Each output store may have a number of write terminals equal to the number of bits which can be stored at each location, the application of a predetermined signal to a write terminal allowing

IPR2023-00033
Apple EX1002 Page 283

one state of a binary signal to be stored at a location addressed at the same time. The apparatus may then include means for sequentially applying the said predetermined signal to groups of the said write terminals, each group containing corresponding write terminals of every output store, and means for generating a change-image signal each time the said predetermined signal is applied to a different group of the said write terminals.

By using the apparatus in both learning and recognition phases with a type of pattern to be recognised and varying both the mapping and the N-tuple size it is possible to find a combination of these variables which provides improved recognition.

Preferably the means for mapping includes a random or pseudo random generator to determine the mapping. Having once determined the mapping, the mapping means preserves this mapping during learning and recognition relating to one pattern or type of pattern but different mappings may be used in order to determine which mapping is useful for a particular class of pattern.

Certain embodiments of the invention will now be described by way of example with reference to the accompanying drawings, in which:—

Figure 1 is a block diagram including apparatus according to the invention,

Figure 2 is a block diagram of part of the control circuit 13 of Figure 1, and

Figure 3 is a block diagram of the set of RAM elements 19 of Figure 1 and circuits for controlling the RAM elements.

In Figure 1 patterns for learning or recognition are viewed by a television camera 10, the output of which is converted to digital form in a digitiser circuit 11 and stored in an input store in the form of a frame store 12. The frame store and digitiser circuit are under the control of a circuit 13 having a user control panel 14. The camera of course, allows an area of the pattern which is to be recognised to be selected and the camera controls allows the magnification of the pattern to be adjusted. Controls of the circuit 13 enable a "window" (that is a selected area of camera output) to be set up for storage in the frame store 12. A picture corresponding to the contents of the frame store 12 is displayed on the monitor 16 to enable the operator to see the part of the pattern which has been selected by the camera and the use of the controls. Since the above described part of Figure 1 uses known equipment it is not described further.

The contents of the frame store 12 are read out to a register 17 in which the data are marshalled. As is explained below read-out is carried out according to a predetermined mapping between the frame store 12 and the register 17. Mapping may be, for example, pseudo random, linear, or special purpose and is controlled by the circuit 13. If selectable mapping is provided it may be under the control of a panel 14.

The contents of the register 17 are transferred by way of a memory control circuit 18 a plurality of output stores in the form of a set of RAM memory elements in a block 19.

An optional control circuit 21 and operator's control panel 22 may be used to set up the memory control circuit 18 by way of a channel 23.

In Figure 1 the contents of the register 17 can be regarded as divided according to the selected N-tuples as indicated at 25 by the designations $N_1, N_2 \ldots N_p$. The selection of memory locations in the block 19 is according to the selected N-tuples and is effected by way of a channel 26 where groups of address lines are indicated as A_1, $A_2 \ldots$.

In the learning phase device select terminals 27 ($G_1, G_2 \ldots$) and write-enable terminals ($H_1, H_2 \ldots$) are activated by the control circuit 18 so that as each pattern is presented to the TV camera 10 the N-tuples of a respective discriminator block in the block 19 are activated and storage locations in respective N-tuples are written with binary ones in accordance with the addresses presented from the pattern by way of the mapping. Writing is carried out in a conventional manner using input/output lines 28 ($W_1, W_2 \ldots$). In this way each discriminator learns one pattern presented.

In the recognition phase the mapping between the register 17 and the N-tuples in the block 19 is the same as in the learning phase but now the RAM write-enable lines are not activated and responses corresponding to the contents of the storage locations in the block 19, making up the discriminators, are passed by way of the input/output lines 28 to the control circuit 18 where they are decoded according to discriminators and applied by way of a channel 24 to decision logic 30 which records the responses of each discriminator and provides signals for a display panel 31. Such responses may simply be the total number of binary ones read out of the N-tuples making up a discriminator when a pattern is presented to the camera 10 and the display may indicate the pattern with which the discriminator with the highest response was trained.

Mapping control will now be described in more detail with reference to Figure 2. The frame store 12 will generally be of the type which stores, for example, eight grey levels for each picture element (pixel). Each of the bits representing a pixel has its own serial number and the store is addressed by entering the serial number of the required bit in an address register 32. Sequences of addresses cause bits to be shifted into the register 17 and when pN successive addresses have been read into the register 32, p N-tuples fill the register 17. Each N-tuple is thus made up as required of a section of bits governed by the sequence of addresses used. Mapping between the frame store 12 and the register 17 and transfer of data is under the control of a circuit 33 which may itself be controlled from the panel 14. A number of different mapping schemes may be

IPR2023-00033
Apple EX1002 Page 284

adopted and circuits for three schemes are shown in Figure 2:—

a pseudo random binary sequence (P.R.B.S.) generator 34 (such generators are known),

a linear scan generator 35, and

a programmable ROM containing the addresses of special purpose mappings.

Additional mapping schemes may be provided but in a system for carrying out recognition only a single ROM equivalent to the programmable ROM 36 is usually provided. It will be clear that if, for example, P.R.B.S.s are to be used then the control circuit 33 will enable the generator 34 and provide clock pulses to read bytes one at a time from the sequence into the address register 32 where the appropriate bits will be addressed in turn. Instead the linear scan generator 35 may be actuated, when each bit is read out in turn along a row or a column or the equivalent of a diagonal in a frame store 12. Equally if special purpose mapping is required words are read from the programmable ROM 36 in sequence under the control of the circuit 33 into the register 32.

A P.R.B.S. generator provides sequences in response to a "seed" (a number passed to the generator) and each time the same seed is used the same sequence is generated. Since once a mapping has been chosen it is used for training all discriminators for a set of images and must be used for recognition dependant on that training, the control circuit 33 is constructed to store "seeds" and pass them to the generator 34 as required.

Where it is required to change the size of N-tuples, the register 17 may simply be regarded as being divided into a different number of sections, although as will be explained arrangements for dealing with the output connections from the register 17 are changed. It may sometimes be desirable to change the mapping when the size of N-tuple is changed and this can be carried out, for example, by changing the programmable ROM 36 or its contents.

In another mapping arrangement the frame store 12 can be regarded as an array of storage locations in which each location has a rectangular coordinate address comprising an X coordinate and a y coordinate. These addresses are under the control of a mapping control circuit which is part of the control circuit 13. The output of the mapping control circuit is then in the form of a sequence of X and Y addresses passed to decode circuits.

The block 19 which will now be described in more detail with reference to Figure 3 comprises p RAMs of which three 41, 42 and 43 are shown. Each RAM takes the form of a board carrying a number of RAM integrated circuits. Alternatively a single integrated circuit may be used instead of each board and its attached integrated circuits. Each board has A address terminals (shown on the face 44 of the RAM 41) which enable every location in the RAMs on the board to be addressed. The address terminals of each board are divided into three groups: a first group which is connected to a respective group of locations corresponding to an N-tuple in the register 17, a group of wires 47 (E) for selecting an N-tuple in the learning or recognition phase and a group of wires 48 (F) for selecting discriminators.

In general each location is capable of storing a word comprising W bits but in Figure 3 for ease of explanation W equals 4 and the input/output connections for the board 40 are shown on a face 45. Each input/output wire is connected to a respective one of four counters 50 to 53 which are used in discriminators and to the output terminals of a register 55 which is used in the learning phase.

The boards may instead have separate input and output wires and in this case the register 55 will be connected to the input wires and the counters 50 to 53 will be connected to the output wires.

Each of the p boards has a device enable terminal G and a write enable terminal H (G and H having suffixes relating to the board).

In the learning phase, the p stages of the register 17 apply signals determined by a teaching pattern at the camera 10 and thus an image held in the frame store 12 to the address terminals of the boards in N_1 to N_p respective N-tuples.

Write control logic 56 is then used to reset the register 55, to set a counter 57 to zero, to enable all the device select terminals G and all the write-enable terminals H of the boards, and to prepare to pass clock pulses from a clock pulse generator 54 to the counter 57. The stages of lower significance E of the counter 57 are connected to the wires 47 and the higher significance stages F are connected to the wires 48. One of the wires W receives a logical one from the register 55 and the counter 57 stands at zero. In this state one section of each RAM board is selected corresponding to the address in the counter 57 and RAM elements corresponding to the addresses applied at the terminals N_1 to N_p are selected and written with a binary one by way of the wire selected by the register 55. In this way a first group of N-tuples corresponding to part of the first discriminator is written.

When the first clock pulse occurs the register 17 is loaded with another p N-tuples but the training pattern is not changed. The counter 57 is incremented by one so that another p locations in the RAM boards are addressed. By continuing this process 2^Exp N-tuples are written to correspond to the first discriminator but when the section E of the register 57 overflows as indicated by a pulse at a terminal 40 from a monostable circuit 58, a new pattern is presented to the camera 10 and while the 2^E states of the least significant end of the counter 57 are cycled the N-tuples of the second discriminator are written.

This process continues until 2^F discriminators have been written when an overflow from the most significant end of the counter 57 causes the shift register to activate the next of the input/output wires W. At this stage the

IPR2023-00033
Apple EX1002 Page 285

monostable 58 again produces a pulse to indicate that a new pattern should be presented to the camera 10. The process continues until each stage of the register 55 has been marked in turn, and then a signal is sent to the write control circuit 56 to indicate that all discriminators have been written. By this time the N-tuples corresponding to 2^ExW discriminators have been written using the same number of teaching patterns, and the apparatus is ready for the recognition phase.

Some learning procedures require the register 55 to mark two of the wires W at the same time, for example where an image presented contains two patterns, each corresponding to a different discriminator. Means for entering specific arrangements of bits into the the register 55 may therefore be provided.

In the recognition phase the pattern to be recognised is presented to the camera 10 and the discriminators are all cycled and their responses compared. The discriminator which, for example, provides the highest response to the pattern presented is indicated and by correlating this discriminator with the corresponding pattern (that is the pattern presented to this discriminator during the learning phase), the pattern currently presented is recognised as being most nearly like the said corresponding pattern. Other recognition schemes may be used, for example the discriminators may be used to provide a not recognised signal if a certain minimum response is not achieved or if two discriminators give the same response. Various schemes for providing recognition signals from discriminator responses are known.

A recognition control circuit 60 is used in the recognition phase which first sets the counters 50 to 53 and 57 to zero and empties a store 61. In addition the device select terminals G of the RAM boards are enabled but the write enable terminals are not. Thus when addressed each RAM board outputs along the four wires W the byte stored at the location addressed. The first section (corresponding to p N-tuples) of the pattern to be recognised is now read from the frame store 12 into the register 17 addressing the terminals N_1 to N_p of the RAM boards. Since the counter 57 is in its lowest state the first group of N-tuples corresponding to the four discriminators are addressed and where a binary one has been stored in the learning phase the counters 50 to 53 are incremented. Thus at this time the counters 50 to 53 correspond to the said four discriminators. The recognition control circuit 60 now passes a clock pulse to the counter 57 and at the same time the next p N-tuples of the pattern are loaded from the frame store 12 into the register 17. In this way the next group of N-tuples of each of the said four discriminators are addressed and the counters 50 to 53 are again incremented where binary ones were stored in the learning phase. A degree of recognition is thus indicated. The process continues until all the 2^E groups of p N-tuples of the pattern have been read into the register 17 and used to address the RAM boards. At this stage the section E of the counter 57 overflows and the monostable circuit 58 provides a pulse to the recognition control circuit 60 indicating that four discriminators have been cycled. The contents of the counters 50 to 53 are transferred to the logic circuit 62 where, in this example, their magnitudes are compared and then stored in a stack in the store 61 with the highest count and the corresponding discriminator number stored in the most significant position.

The recognition process described above is repeated until the section E of the counter 57 again overflows and the counters 50 to 53 are cleared when the logic 62 compares the contents of these counters with the contents of the store 61 to arrange the contents of the next four discriminators and the four previously stored in ascending order of magnitude. When the section F of the store 57 overflows it indicates that the recognition process is complete when after the last contents of the counters 50 to 53 have been transferred to the store 61, the discriminator number corresponding to the highest discriminator score is read into a display 63 indicating by correlation with a training pattern which of the training patterns is recognised as being most likely to be the same, or related to the pattern which is to be recognised.

If the apparatus is to be used only in the recognition phase the RAM memory comprising the p boards is replaced by a ROM memory, and the register 55 and the write control circuit 56 are omitted. Each of the p boards carries ROMs which have been set up to contain the same data as the corresponding RAMs in the arrangement described after the RAMs have experienced the learning phase. Thus the contents of the ROMs are specific to a particular set of teaching patterns and a particular mapping between the frame store 12 and the register 17.

The circuits 56 and 58 to 62 can conveniently be replaced by a microprocessor. If preferred the microprocessor can also provide a replacement for the counters 50 to 53 and 57, and the register 55.

The advantages of the parallel/serial memory structure of Figure 3 are now apparent:

(a) there are 2^ExN N-tuples although only p RAM boards are provided;

(b) in the recognition procedure, the number of passes to process one pattern to be recognised is 2^{E+F} instead of 2^A in an entirely serial arrangement; and

(c) the number of discriminators is 2^FxW.

Hence an optimum size of N-tuple and number of passes (equivalent to recognition time) can be achieved for many different applications using the parallel/series memory structure.

The control circuit 21 of Figure 1 need only be provided in apparatus where it is required to vary the size of the N-tuples. For example N-tuple size can be increased by allocating one or more of the wires in the group E for this purpose. If for

instance the N-tuple size is to be increased by two then two wires in the group E are so allocated and in each N-tuple from the register 17 two stages are gated to two of the wires in the group E while only one of the RAM boards receives a device enable signal. When the next N-tuple in the register 17 is processed two of its outputs are gated to the two wires of the group E and the device enable signal is applied only to the next RAM board. To obtain complete flexibility in the size of the N-tuples the number of wires in each of the groups N_1 to N_p may be reduced to one.

When an optimum parallel/series structure has been decided, for example for a particular application, further apparatus can be constructed in which the RAM boards are replaced by ROM boards (containing data provided as described above) and the register 55 and the write control circuit 56 are omitted.

The parallel/series memory structure lends itself to a modular approach which can easily be varied for different applications. A standard rack can be provided which will accommodate the maximum number of RAM or ROM boards required and the main items for individual construction are then the number of address wires A selected for direct connection to the register 17 and the number of wires for connection in the two groups E and F.

While a specific form of the invention has been described it will be realised that it can be put into practice in many other ways, for example the manual control panels 14 and 22 may be replaced by an interface for a computer, and the optical input can be in many forms and can be replaced by other forms of input such as a touch panel, a graphical input, and an audio transducer.

Uses envisaged for the invention include provision of a large number of recognition terminals in premises which require to recognise such items as credit cards, bank notes or people's faces or voices. Other uses include recognising parts in factory assembly operations and sorting such parts. Quality control of fruit and other products may be carried out by using the recognition apparatus to recognise satisfactory items.

Claims

1. Apparatus for recognising images, comprising an input store having a plurality of storage locations allowing data representative of elements of an image which is to be recognised to be stored at locations related to the positions of the elements, a plurality of output stores each having a number of address terminals which allow every storage location in that store to be addressed, the locations of the output store either containing, or being capable of containing, data related to images which are to be recognised, the address terminals being divided into at least a lower significance group and a higher significance group, each group of terminals containing at least one terminal, mapping means with input connected to the input store and a plurality of output terminals divided into groups with each group corresponding to a respective output store and at least one of the terminals of that group connected to an address terminal of the lower significance group of the corresponding store, addressing means connected to the terminals of the higher significance groups for cyclically addressing groups of storage locations, and discriminator means responsive to output signals from groups of output-store storage locations, when containing data and when addressed by way of the mapping means and by the addressing means during recognition, to provide a measure of recognition.

2. Apparatus according to Claim 1 wherein each location of the input store is capable of storing a word of at least one bit and the mapping means includes means for generating a series of addresses for the input store and addressing the input store with the addresses generated, and a register for storing successive sequences of bits from the input store, the said output terminals of the mapping means being output terminals of respective stages of the register.

3. Apparatus according to Claim 1 or 2 wherein all the mapping means output terminals of each group are connected to respective address terminals of the lower significance group of the corresponding output store.

4. Apparatus according to any preceding claim wherein each output store has at least one output terminal at which a signal appears corresponding to the contents of a location in that store which is addressed, the output terminals of the output stores being in groups of corresponding terminals, and the discriminator means includes one response means for each group of output-store output terminals, and connected to the output terminals in that group, and operating means for receiving signals from the response means and clearing the response means, at the end of each cycle of the addressing means and operating upon the signals so received and any previously received signals relating to the same image in providing the measure of recognition.

5. Apparatus according to Claim 4 wherein the addressing means is constructed to generate at least one read signal in each cycle, each such signal being generated after a group of output stores has been addressed, and the operating means is constructed to receive signals from the response means and clear the response means in response to each read signal, and to operate on the signals so received and any previously received signals relating to the same image in providing the measure of recognition.

6. Apparatus according to Claim 4 or 5 wherein each response means total the number of binary ones it receives before being cleared, and the operating means operates on the response received from each response means by comparing it with other responses.

7. Apparatus according to any preceding claim including write means for storing data in the

output stores when during a learning phase data corresponding to an image is present in the input store, the addressing means being constructed to generate at least one change-image signal in each cycle, each such signal being generated after a group of output stores has been addressed, and indicating that a change of image is required in the learning phase.

8. Apparatus according to Claim 7 wherein each output store has a number of write terminals equal to the number of bits which can be stored at each location, the application of a predetermined signal to a write terminal allowing one state of binary signal to be stored at a location addressed at the same time, and the apparatus includes means for sequentially applying the said predetermined signal to groups of the said write terminals, each group containing corresponding write terminals of every output store, and means for generating a change-image signal each time the said predetermined signal is applied to a different group of the said write terminals.

9. Apparatus according to any of Claims 1 to 6 wherein the output stores comprise a plurality of read only memories.

10. Apparatus according to any preceding claim comprising a video camera, means for digitising at least a portion of the signal from the camera and transferring the result to the input store.

11. Apparatus for recognising images substantially as hereinbefore described with reference to Figure 1, Figure 2 or Figure 3.

Printed for Her Majesty's Stationery Office by the Courier Press, Leamington Spa, 1983. Published by the Patent Office, 25 Southampton Buildings, London, WC2A 1AY, from which copies may be obtained

(12) UK Patent Application (19) GB (11) 2 380 920 (13) A

(43) Date of A Publication 16.04.2003

(21) Application No 0201614.5

(22) Date of Filing 24.01.2002

(30) Priority Data
(31) 0122052 (32) 12.09.2001 (33) GB

(71) Applicant(s)
TeraView Limited
(Incorporated in the United Kingdom)
302/304 Cambridge Science Park,
Milton Road, CAMBRIDGE,
Cambridgeshire, CB4 0WG,
United Kingdom

(72) Inventor(s)
Vincent Patrick Wallace
Ruth Mary Woodward
Donald Dominic Arnone
Bryan Edward Cole

(74) Agent and/or Address for Service
Marks & Clerk
57-60 Lincoln's Inn Fields, LONDON,
WC2A 3LS, United Kingdom

(51) INT CL[7]
G01N 21/49 // A61B 6/00 , G01N 21/35

(52) UK CL (Edition V)
H5R RAT RTQ

(56) Documents Cited
GB 2360186 A GB 2347835 A
EP 0864857 A1 EP 0841548 A1
INSPEC abstract Number 6705308 & C M Cisela et al, "Biomedical Applications of terahertz pulse imaging", Proceedings of the SPIE - International Society for Optical Engineering, 2000, SPIE, pp73-81, Vol 3934.
INSPEC abstract Number 6665442 & D D Arnone et al, "Applications of terahertz technology to medical imaging", Proceedings of the SPIE - International Society for Optical Engineering, 1999, SPIE, pp209-219, Vol 3828.

(58) Field of Search
UK CL (Edition V) H5R
INT CL[7] A61B, G01J, G01N
Other: Online: EPODOC, WPI, PAJ.

(54) Abstract Title
Terahertz radiation imaging

(57) A THz radiation imaging apparatus and method where the sample to be imaged is irradiated with a pulse of electromagnetic radiation with a plurality of frequencies in the range 25 GHz to 100 THz. A parameter relating to radiation reflected or transmitted by the sample is determined. A calculation is made of the parameter value at a first time relative to the parameter value at a second time. The second time may coincide with a physical feature of the parameter waveform (e.g. a maxima or minima), or the time interval between the first and second times may be constant. An image is generated by plotting the calculated relative value for different points of the sample. The relative calculation may be subtraction, division, multiplication or addition of the parameters at the two times. The parameter may be the amplitude of the reflected or transmitted radiation, or may be filtered (e.g using reference values, Fourier transform).

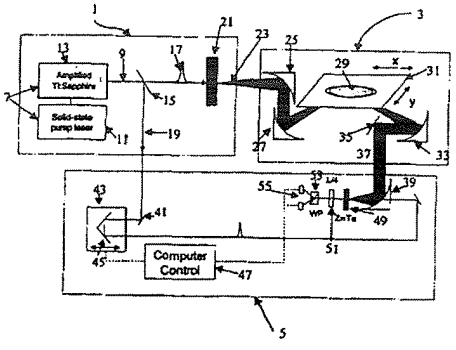


Figure 1

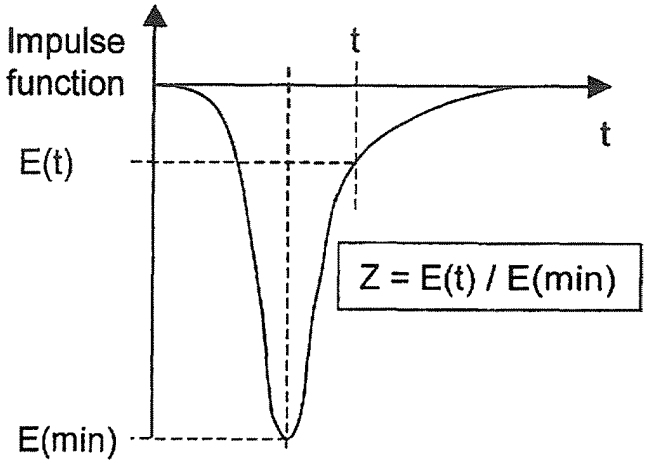


Figure 5

GB 2 380 920 A

IPR2023-00033
Apple EX1002 Page 289

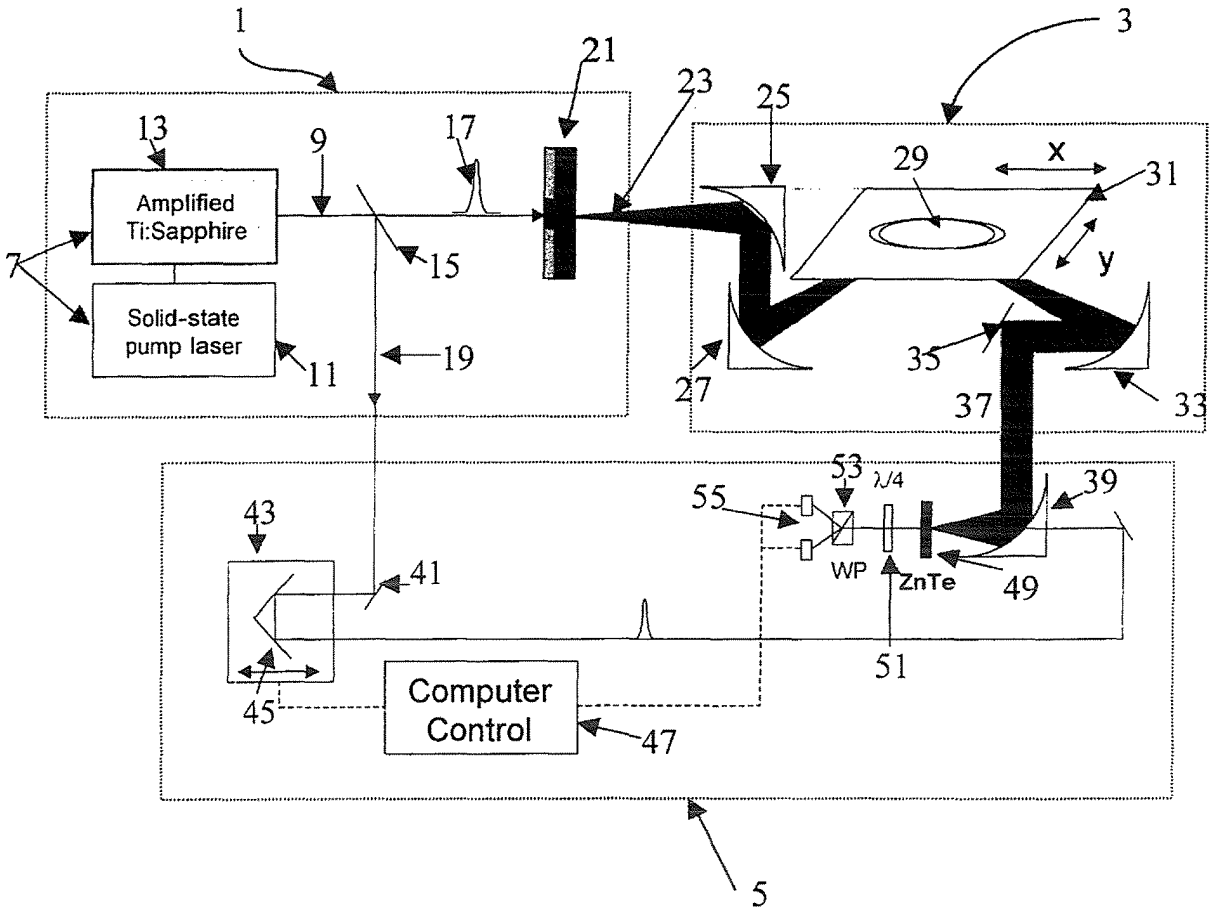

1
3
21
23
25
13
9
17
29
x
31
Amplified
Ti:Sapphire
7
15
y
Solid-state
pump laser
11
19
27
35
37
33
53
λ/4
39
55
43
41
WP
ZnTe
49
51
Computer
Control
45
47
5

Figure 1

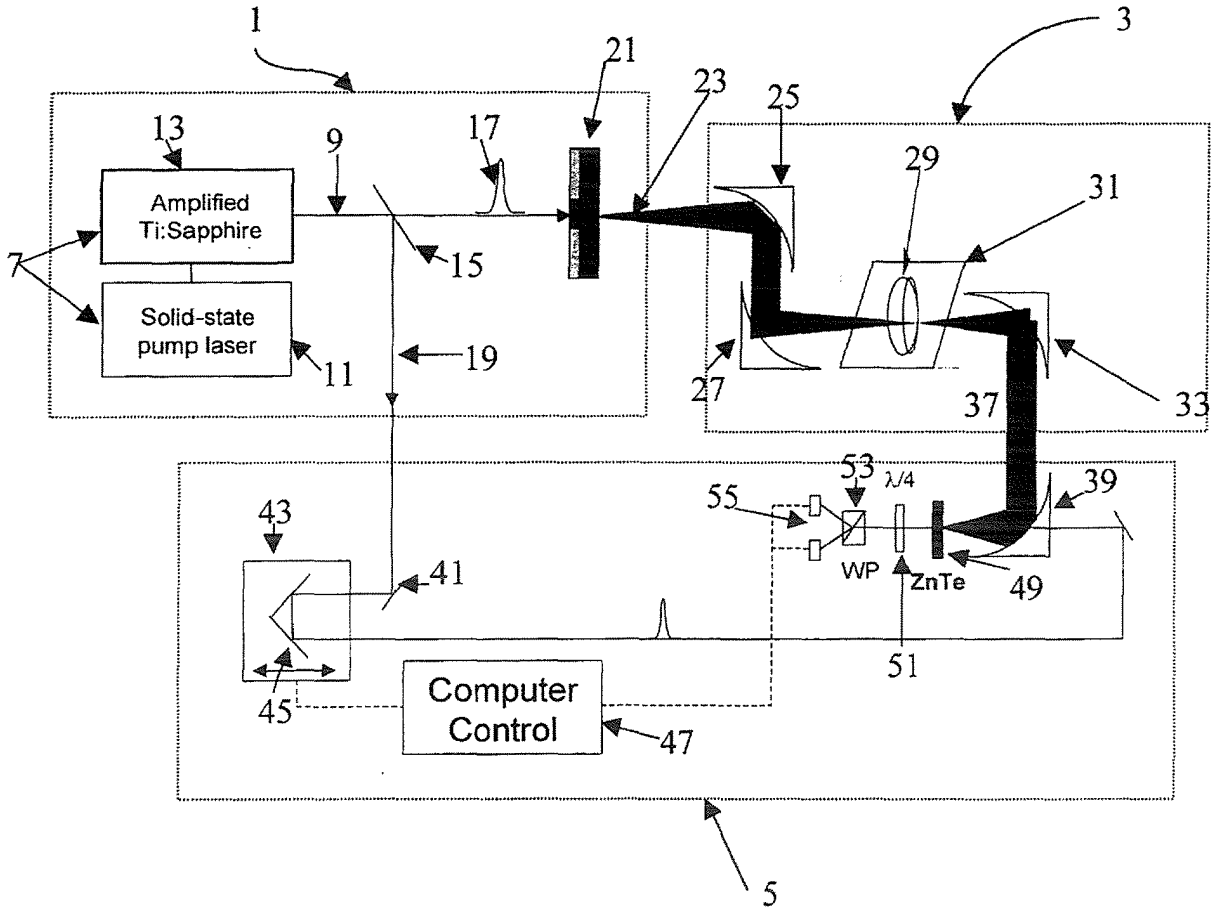

1
3
21
23
25
13
9
17
29
31
Amplified
Ti:Sapphire
7
15
Solid-state
pump laser
11
19
27
37
33
53
λ/4
55
39
43
41
WP
ZnTe
49
51
Computer
Control
45
47
5

Figure 2

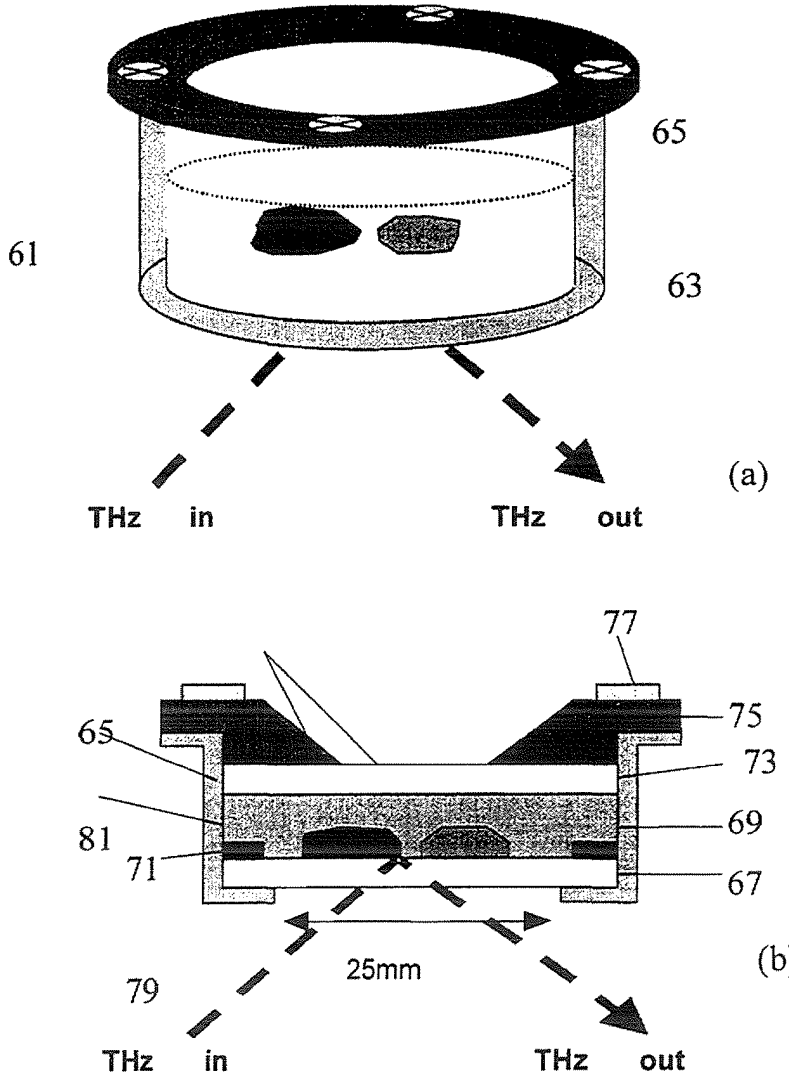

61
65
63
THz in
THz out
(a)
77
75
65
73
81
69
71
67
25mm
79
THz in
THz out
(b)

Figure 3

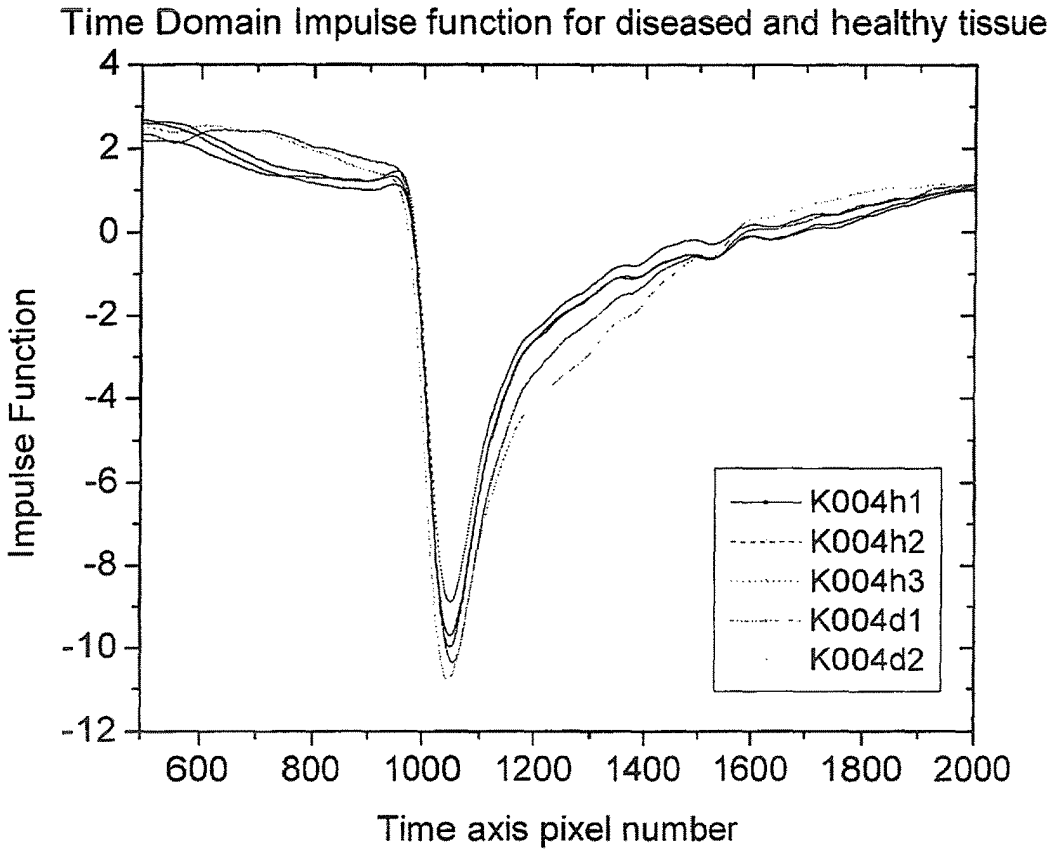

Time Domain Impulse function for diseased and healthy tissue
Impulse Function
Time axis pixel number
4
2
0
-2
-4
-6
-8
-10
-12
600
800
1000
1200
1400
1600
1800
2000
K004h1
K004h2
K004h3
K004d1
K004d2

Figure 4

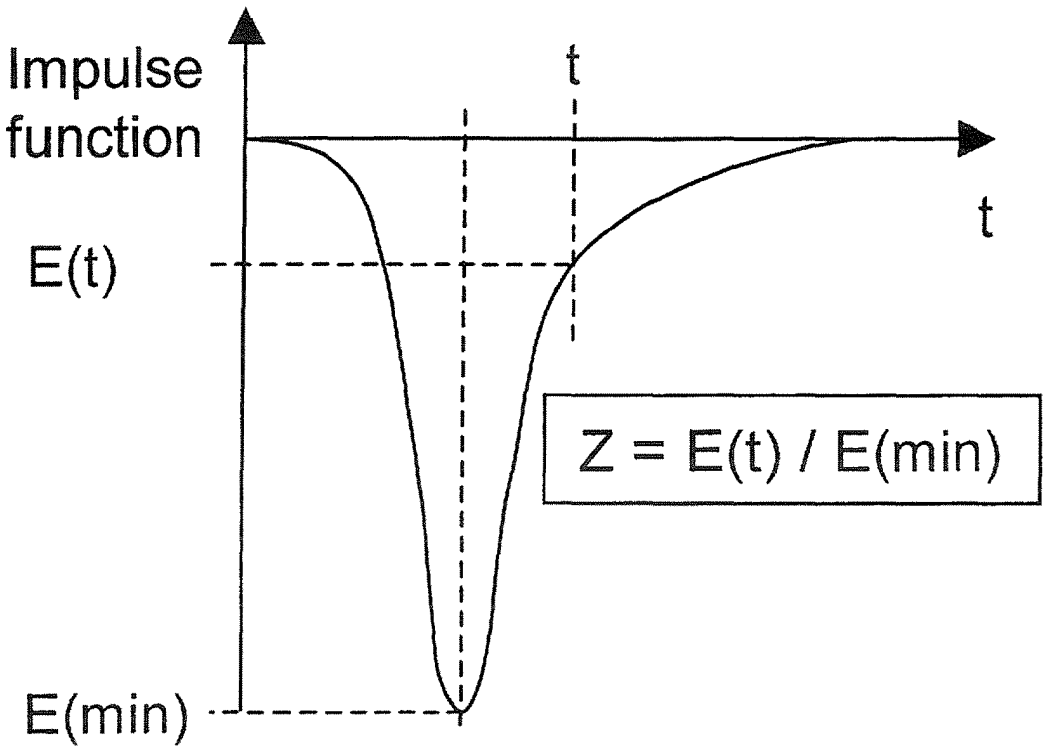


Figure 5

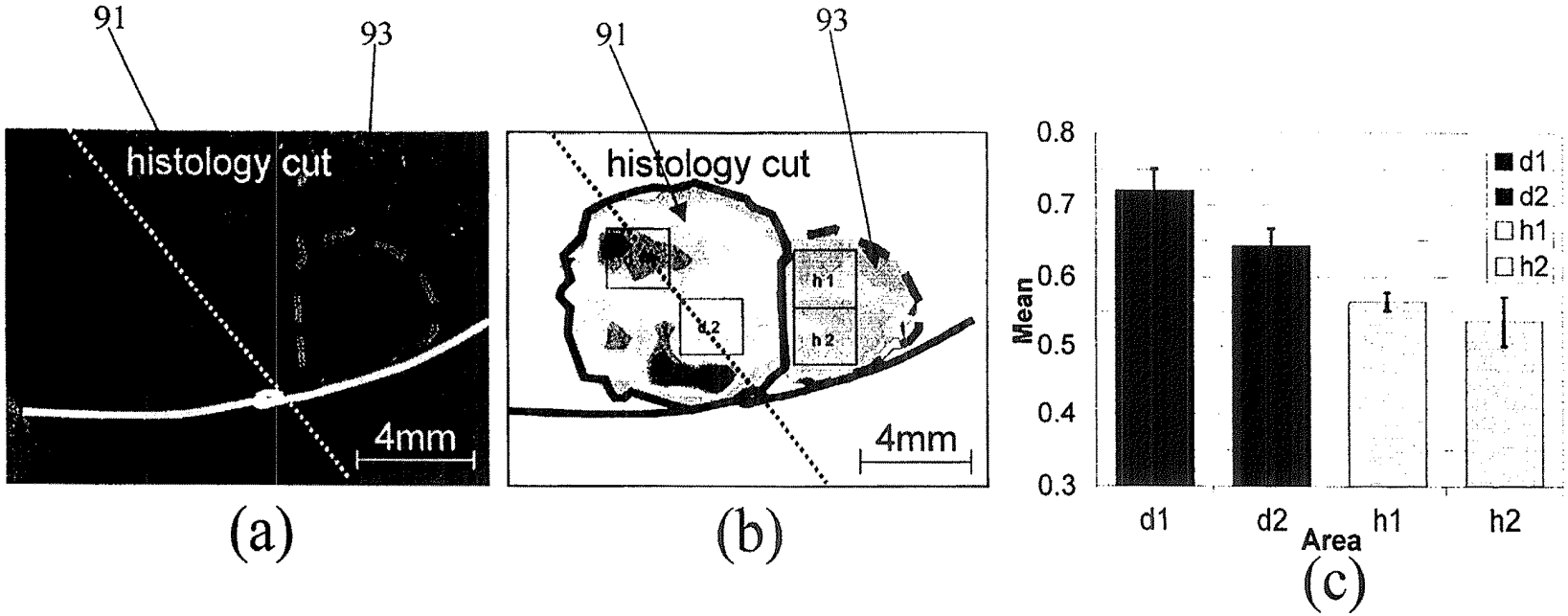


Figure 6

IPR2023-00033
Apple EX1002 Page 295

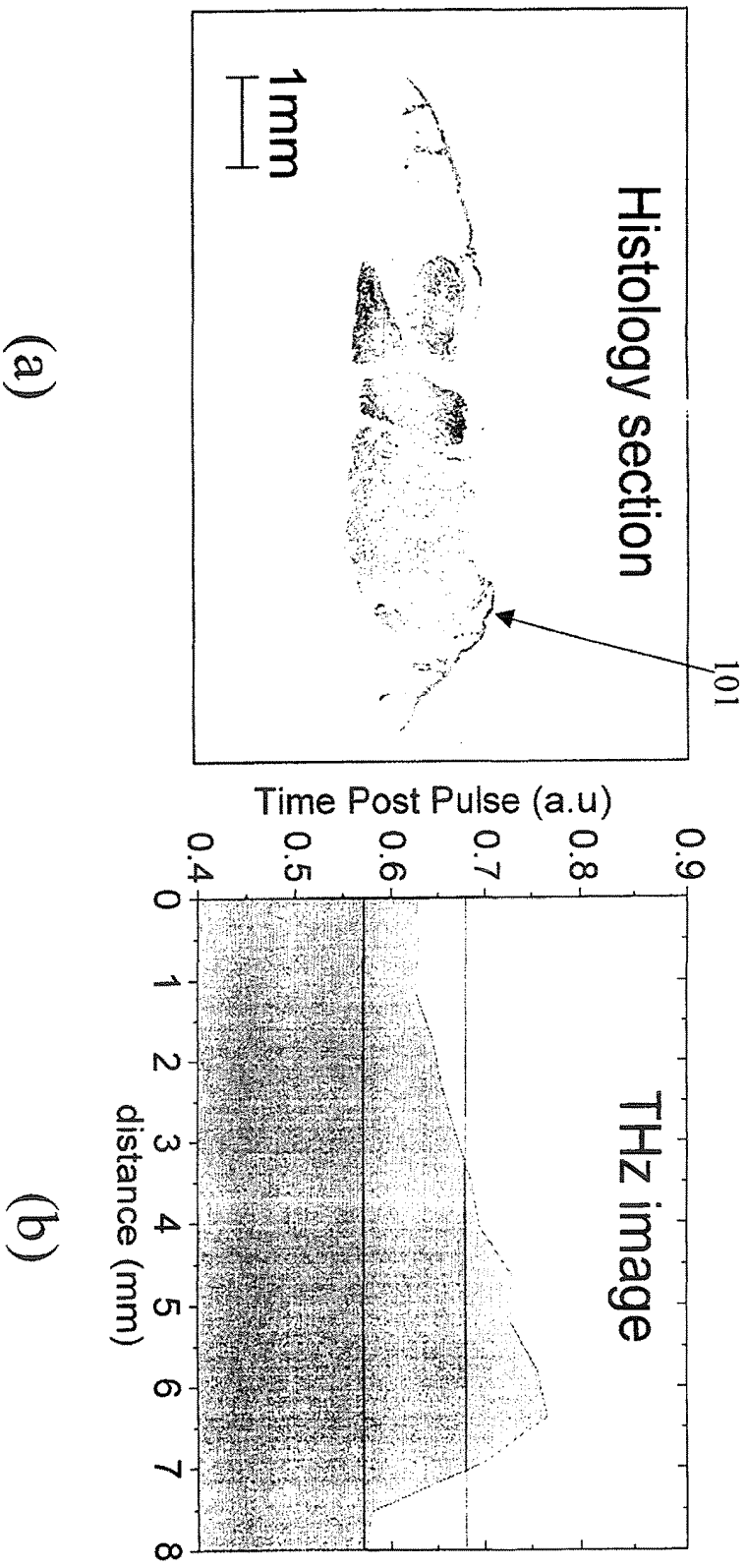

(a)
Histology section
1mm
101
(b)
THz image
Time Post Pulse (a.u)
0.4
0.5
0.6
0.7
0.8
0.9
distance (mm)
0
1
2
3
4
5
6
7
8

Figure 7

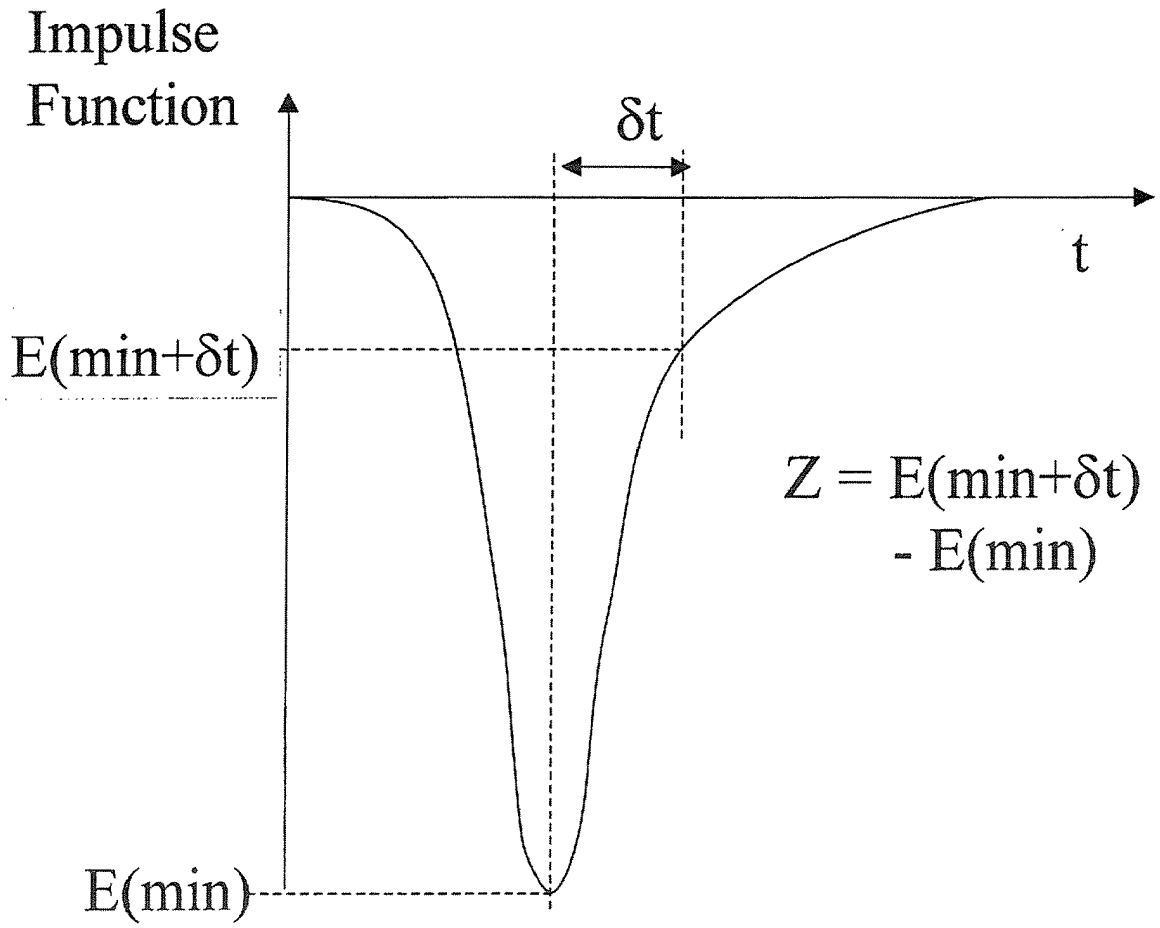


Figure 8

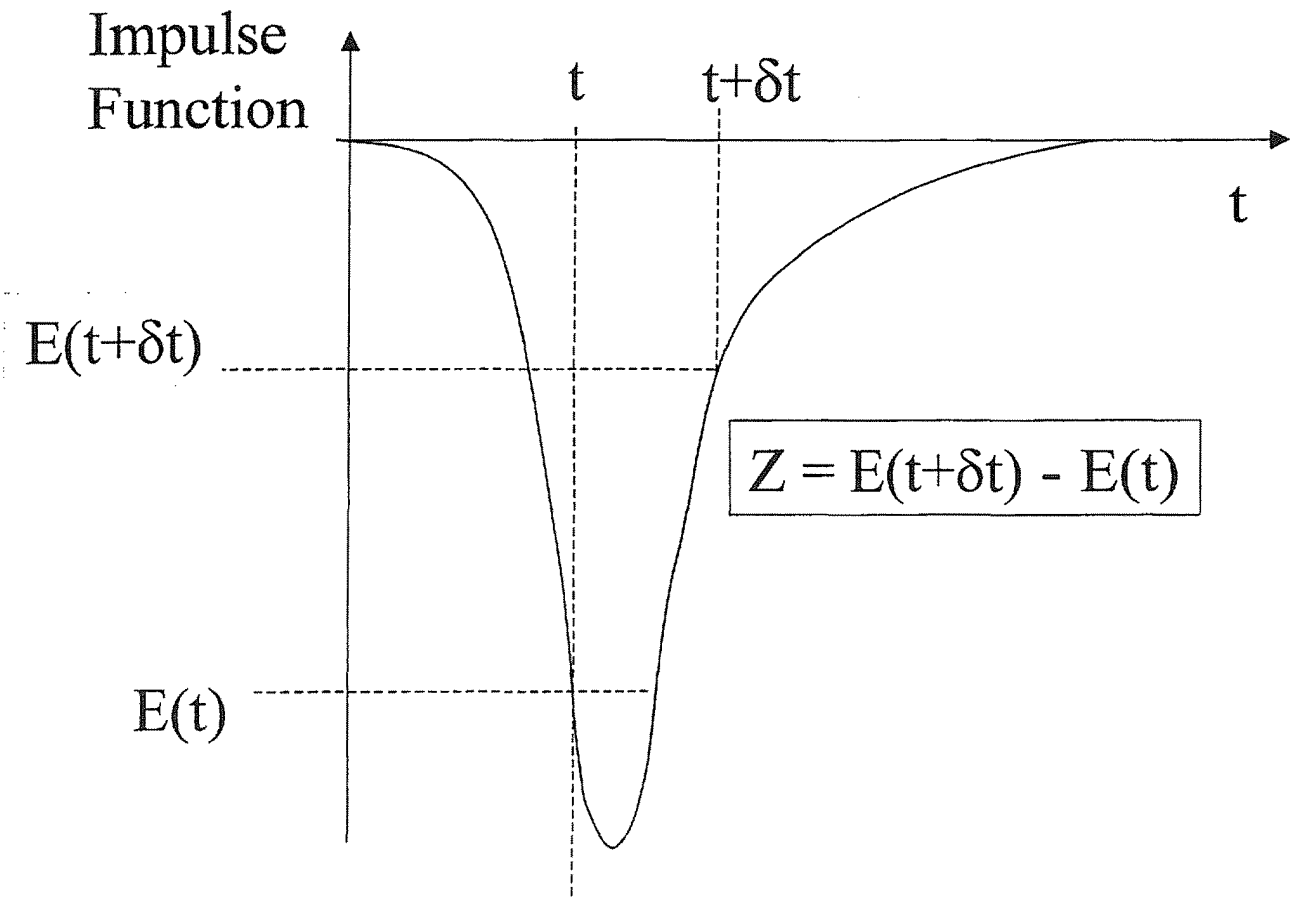


Figure 9

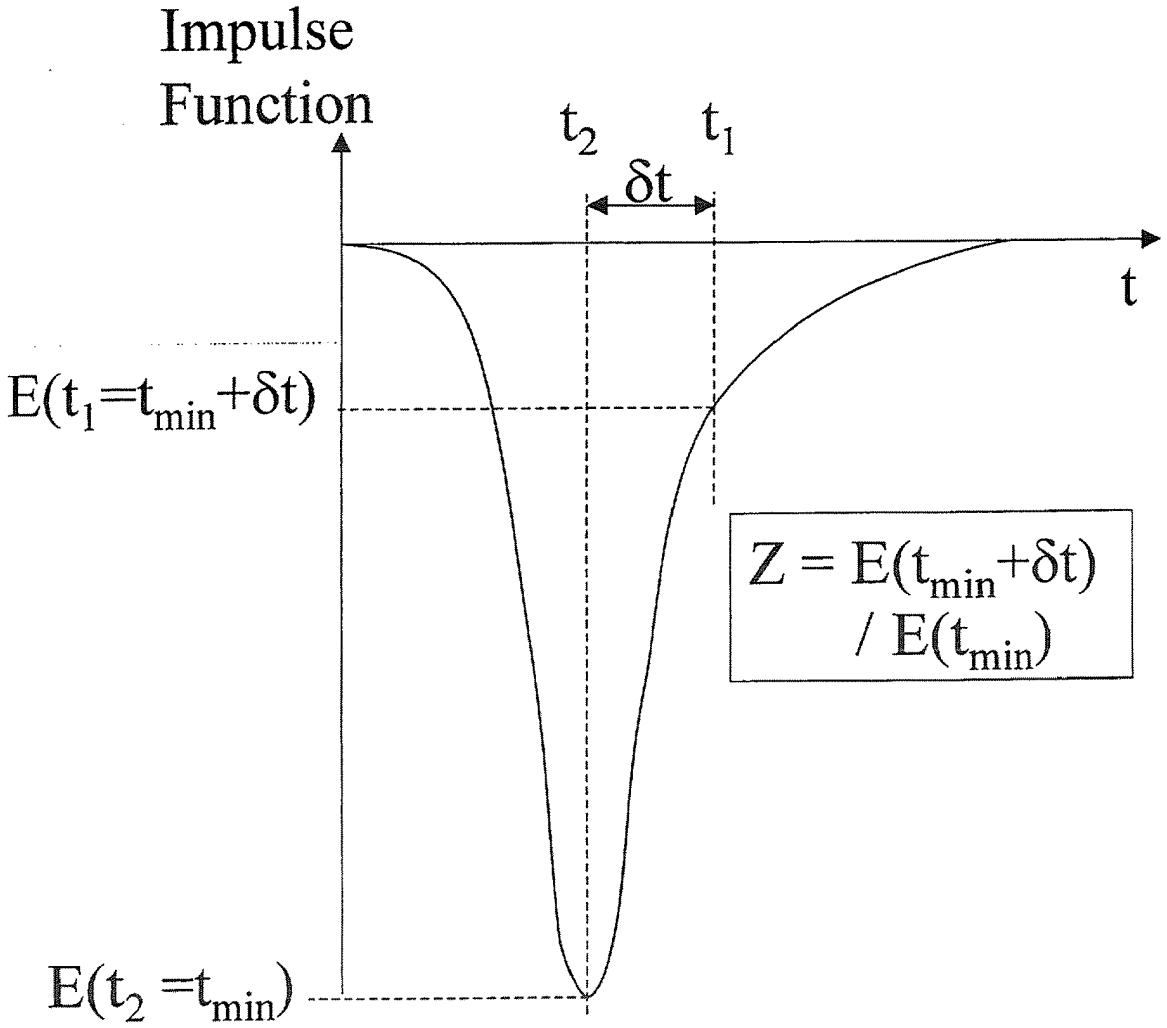

Impulse
Function
t_2
t_1
δt
t
$E(t_1=t_{min}+\delta t)$
$Z = E(t_{min}+\delta t) / E(t_{min})$
$E(t_2=t_{min})$

Figure 10

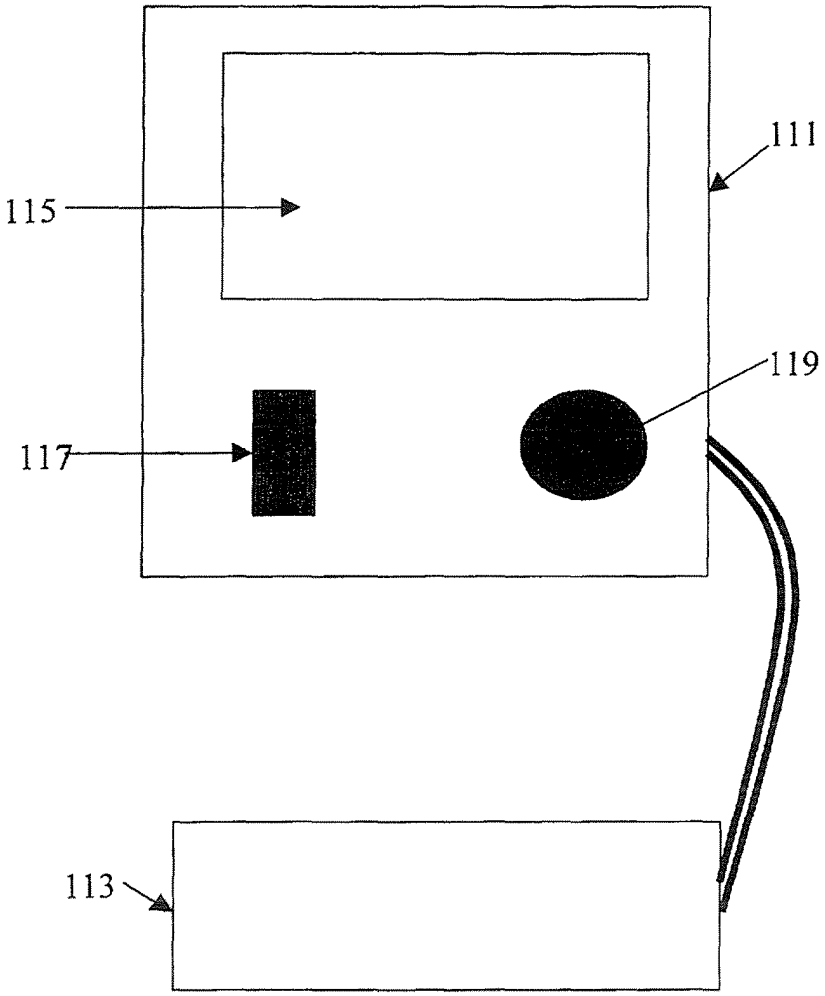


Figure 11

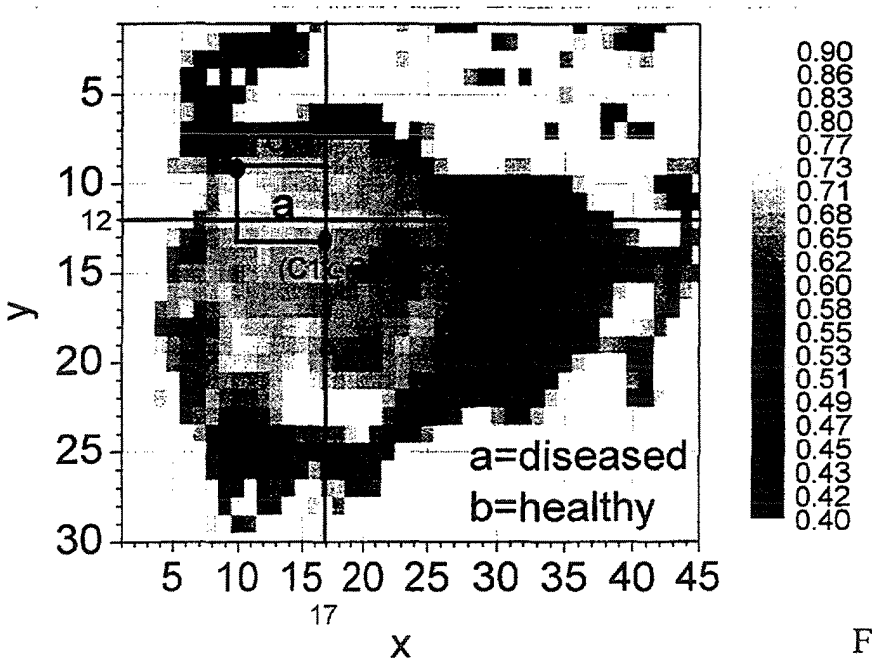


Figure 12

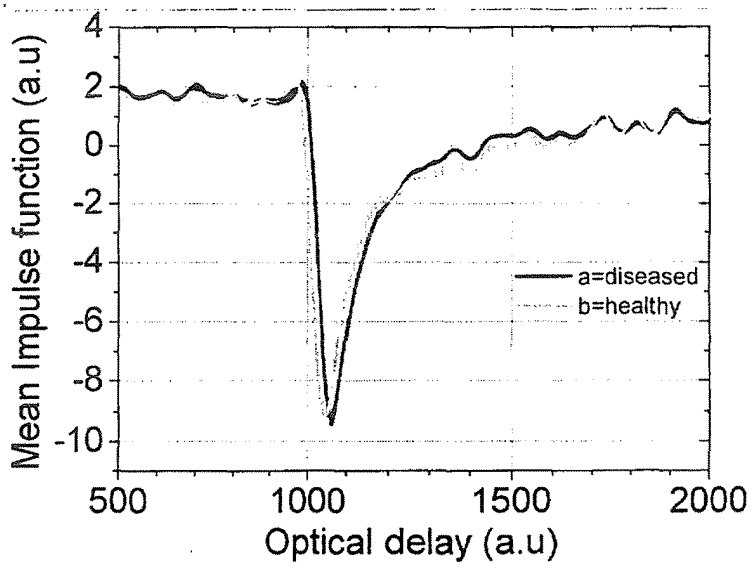


Figure 13

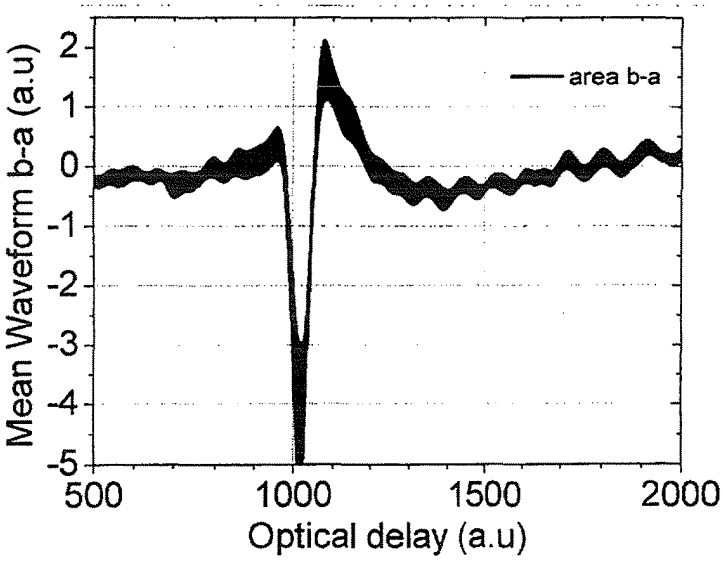


Figure 14

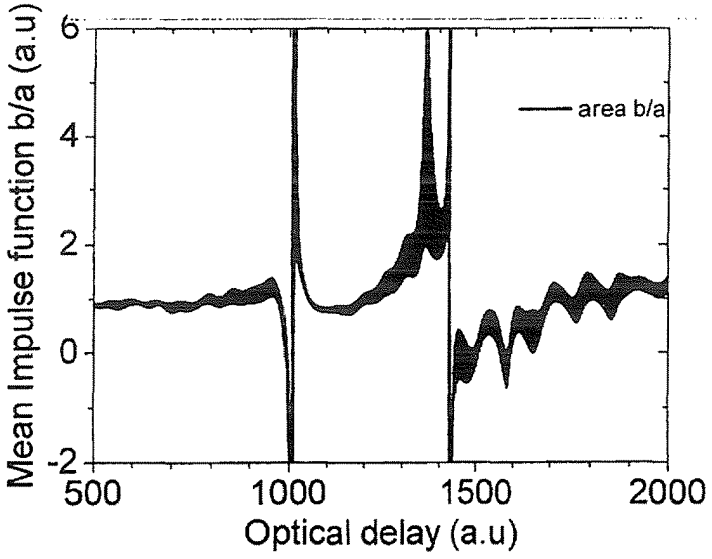


Figure 15

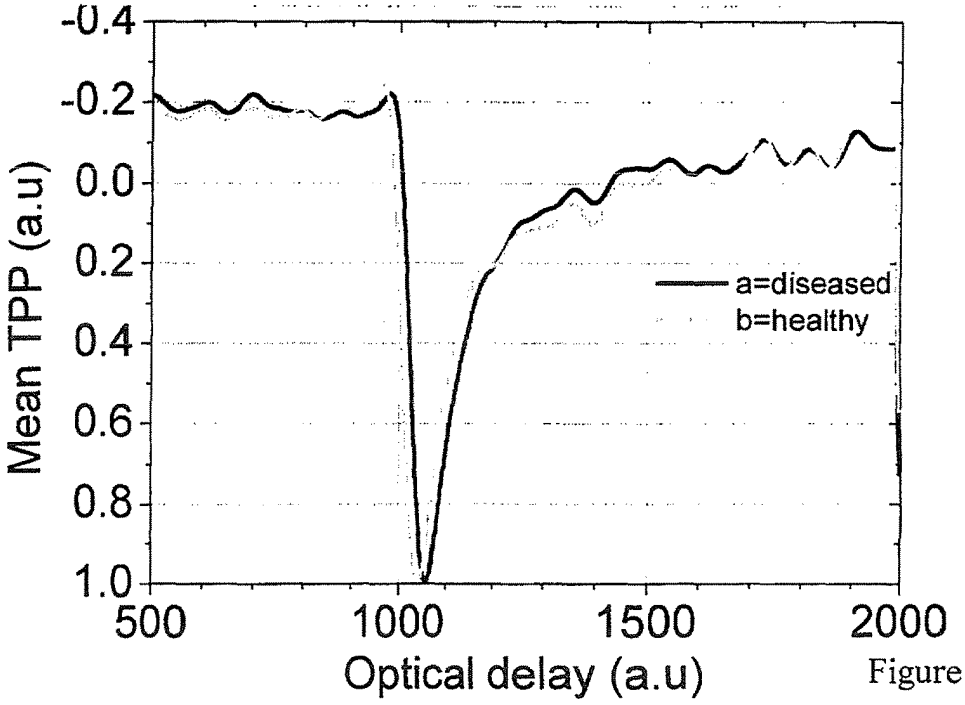


Figure 16

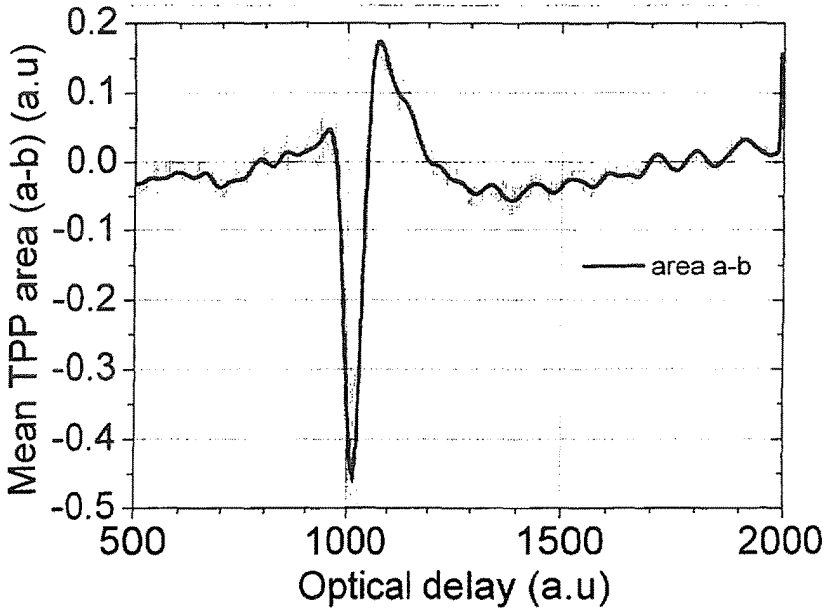


Figure 17

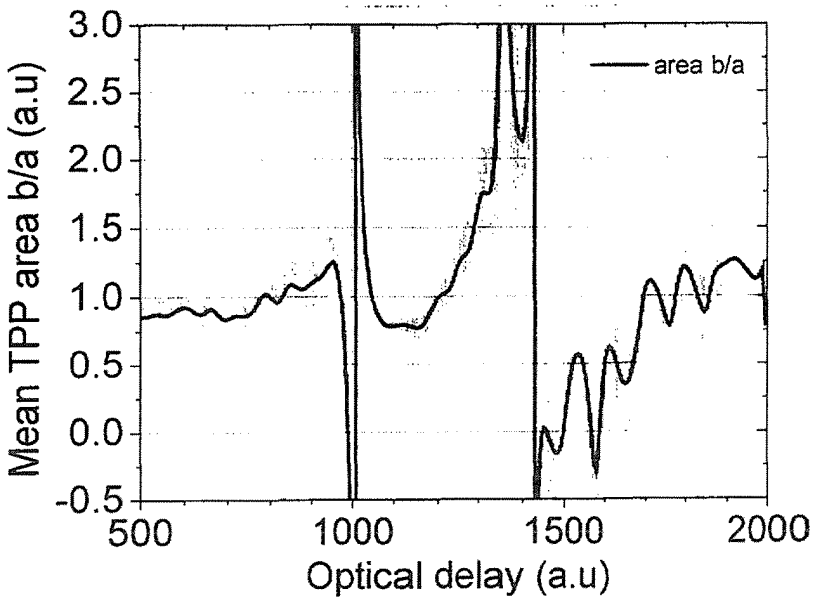
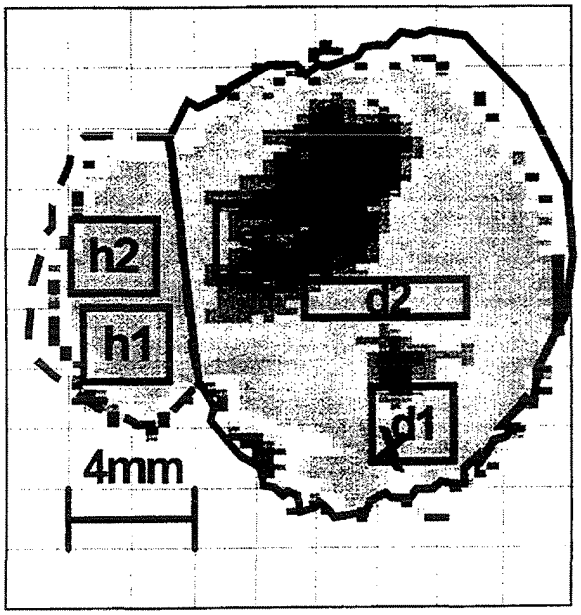

Mean TPP area b/a (a.u)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
500
1000
1500
2000
Optical delay (a.u)
area b/a

Figure 18



E_{min}

Figure 19

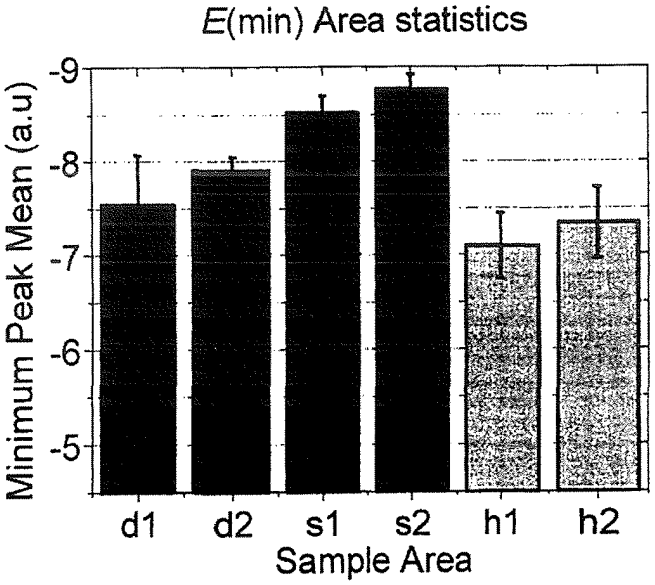


Figure 20

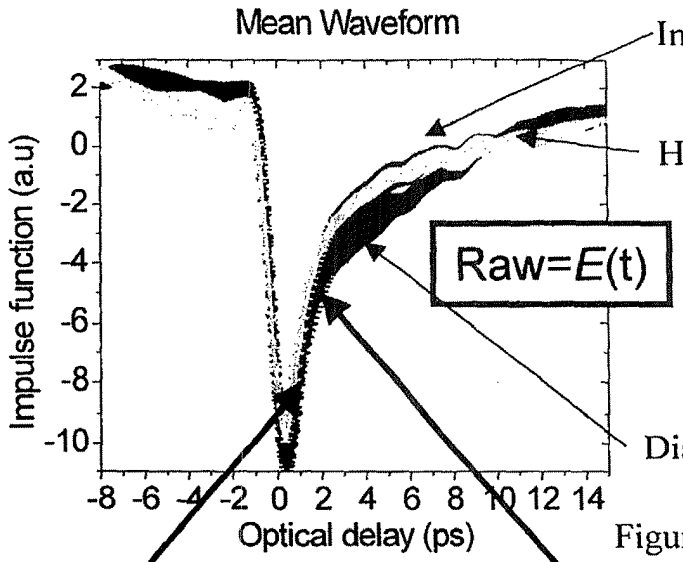


Figure 21

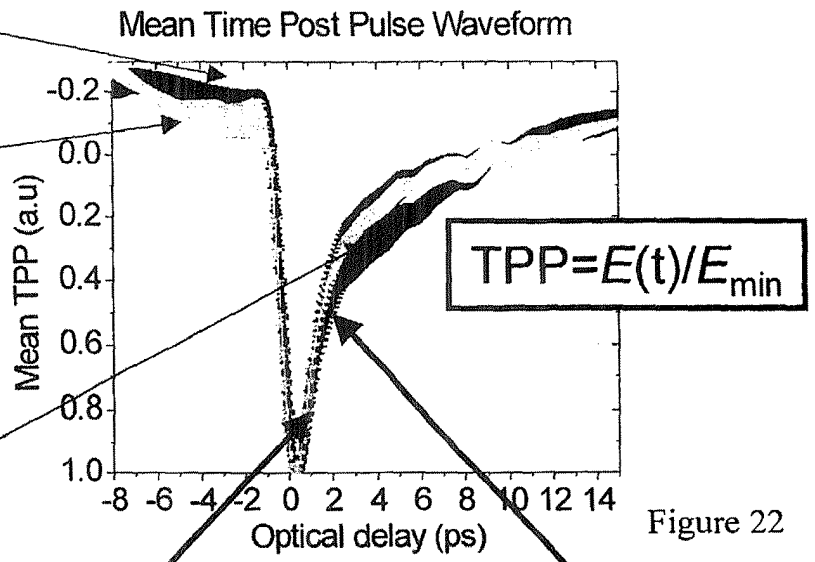


Figure 22

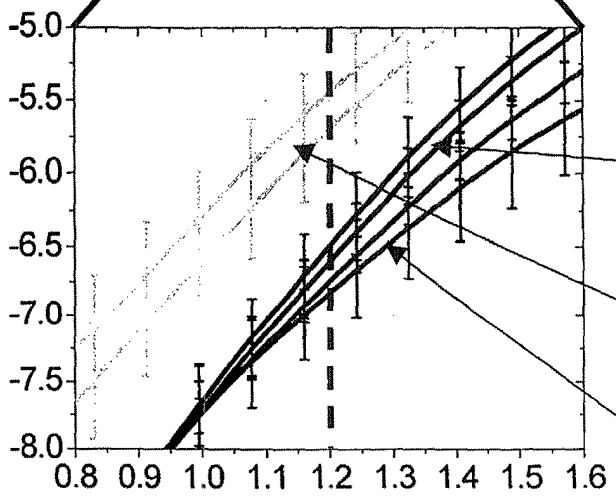


Figure 23

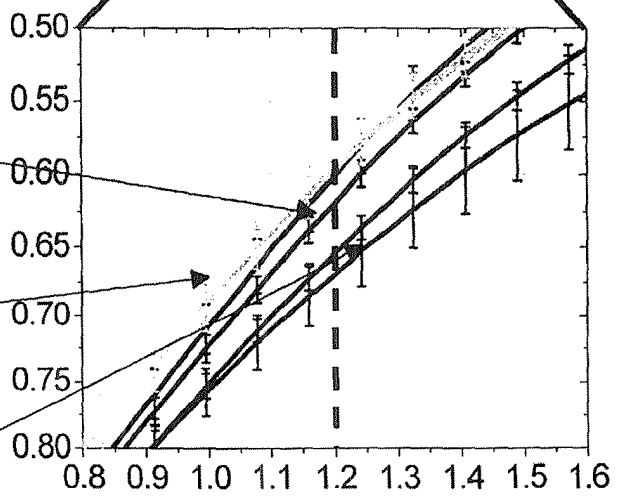


Figure 24

IPR2023-00033
Apple EX1002 Page 306

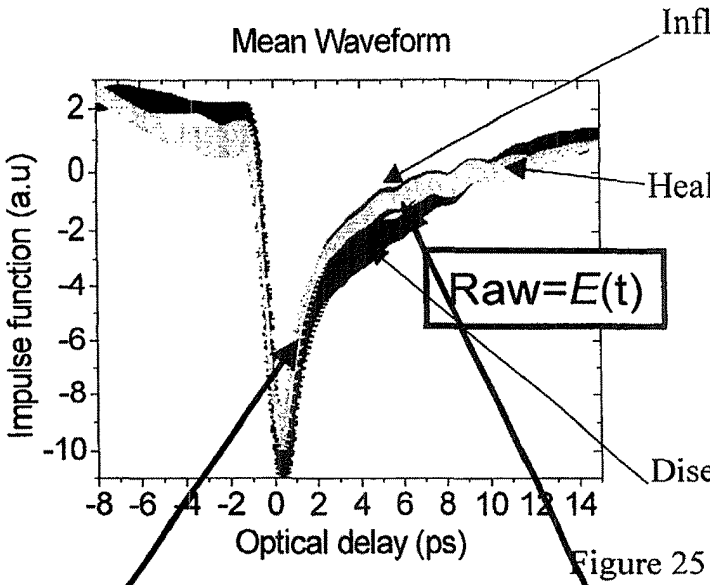


Figure 25

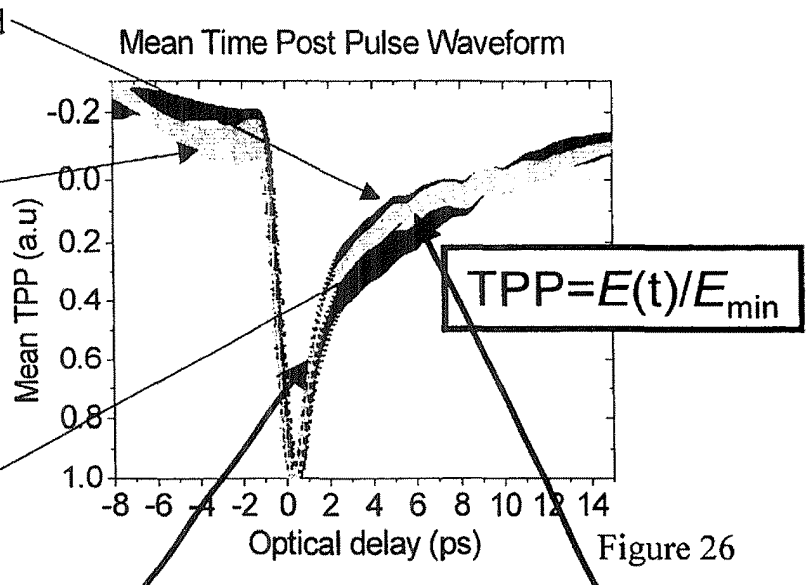


Figure 26

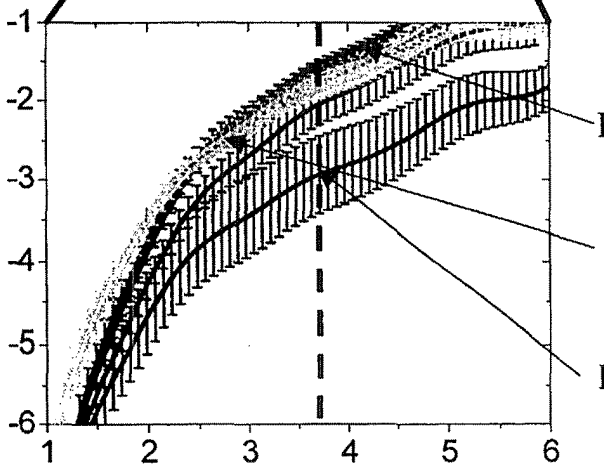


Figure 27

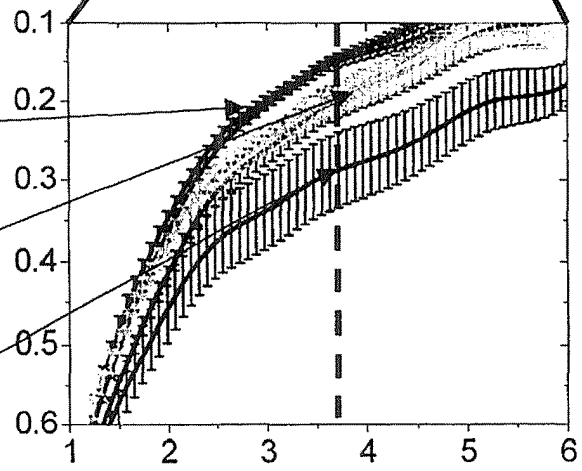


Figure 28

IPR2023-00033
Apple EX1002 Page 307

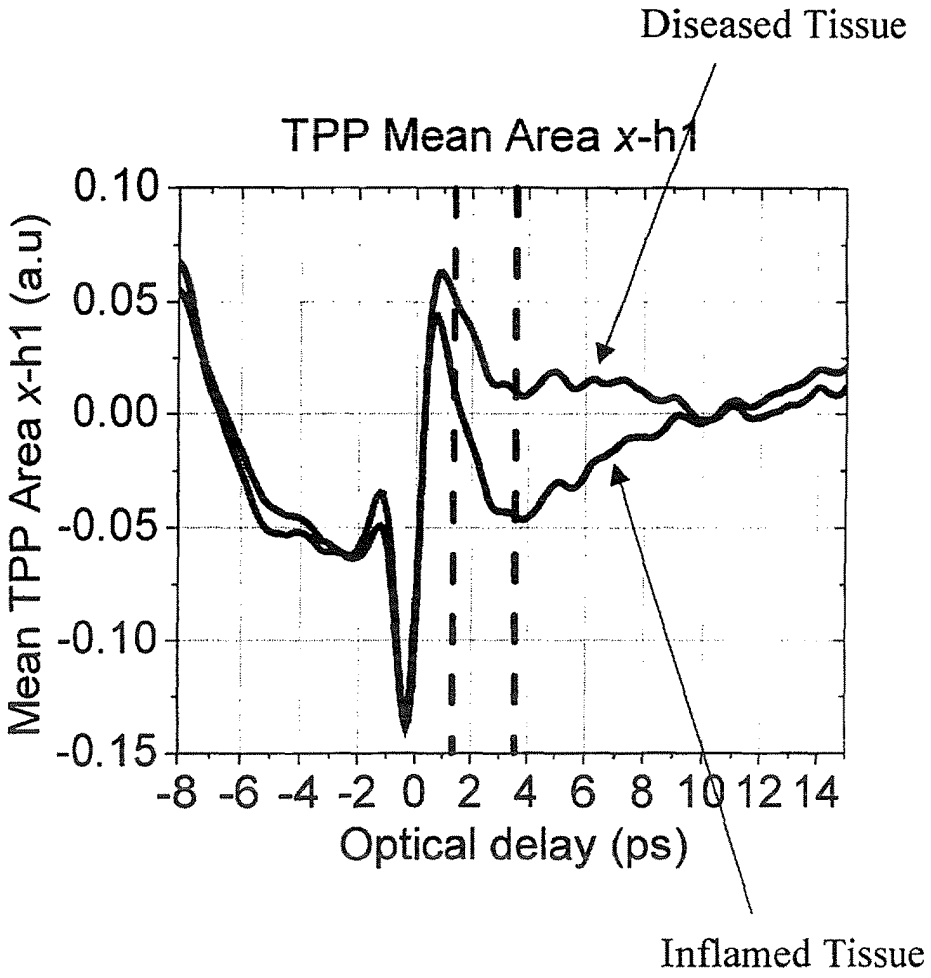
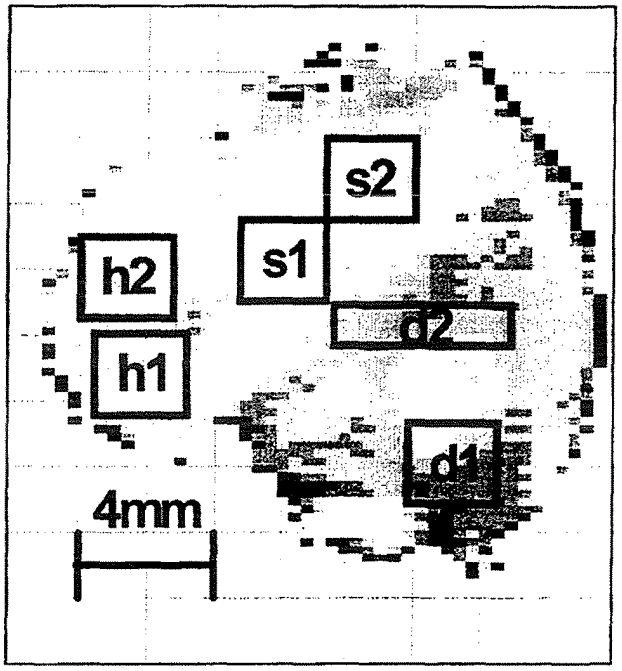
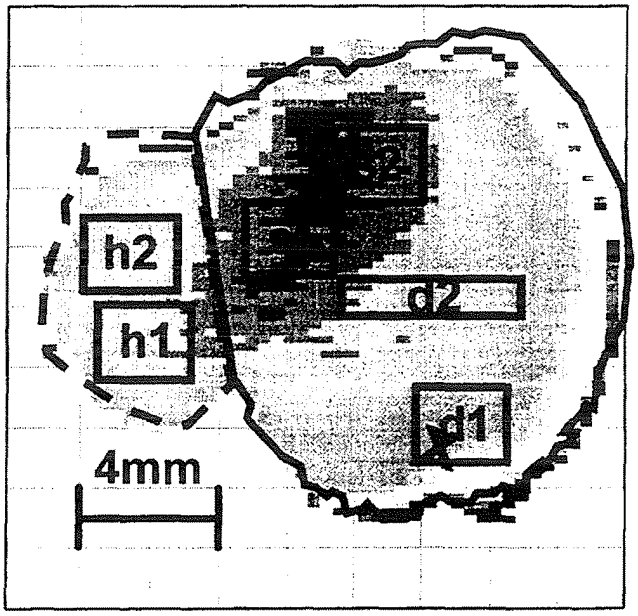

Diseased Tissue
TPP Mean Area x-h1
Mean TPP Area x-h1 (a.u)
0.10
0.05
0.00
-0.05
-0.10
-0.15
-8 -6 -4 -2 0 2 4 6 8 10 12 14
Optical delay (ps)
Inflamed Tissue

Figure 29



t = 1.2ps

Figure 30



t = 3.7ps

Figure 31

2380920

Imaging Apparatus and Method

The present invention relates to the field of imaging samples with radiation in the infra-red (IR) and Terahertz frequency range and specifically using radiation in the higher Gigahertz (GHz) and the Terahertz (THz) frequency ranges. In this field, all such radiation is colloquially referred to as THz radiation, particularly that in the range from 25GHz to 100THz, more particularly that in the range of 50GHz to 84THz, especially that in the range from 100GHz to 50THz.

Such radiation is non-ionising and, as a result, it is particularly of use in medical applications. In medical imaging, the radiation is generally reflected from or transmitted through the patient.

Components of the sample being imaged will have a frequency dependent absorption coefficient and refractive index, thus each component of a sample subjected to radiation will leave its own characteristic fingerprint in the detected radiation. Thus, researchers have attempted to image samples using a plurality of frequencies to create an image from spectral information.

Measurements have been made using both frequency domain techniques, (where the amplitude of each frequency components is analysed), and time domain techniques, (where the radiation is analysed as a function of the delay time introduced by the sample into the path of the radiation). Time domain imaging is described in earlier patent GB 2 347 835.

For imaging, in the time domain, a time domain spectra is obtained for each pixel of the sample. It is then necessary to obtain a single parameter from each spectra to plot for each pixel. Previous attempts at producing images from such spectra have used the amplitude of the highest maximum or lowest minimum.

It is an object of the present invention to provide a method and apparatus which derives a parameter from time domain spectra to allow an image with enhanced contrast to be generated. The image can be a two dimensional image of an area of a sample, or it may be a profile of a line passing through the sample.

Thus, in a first aspect, the present invention provides a method of imaging a sample, the method comprising:

(a) irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) determining a first parameter at least related to the amplitude of the radiation which is either reflected from and/or transmitted by the sample, in the time domain;

(c) calculating the value of the first parameter at a first time value relative to the value of the first parameter at a second time value, wherein the second time value coincides with a physical feature of the first parameter with respect to time; and

(d) generating an image by plotting the value calculated in step (c) for different points of the sample.

As previously described, the image may be just a profile through the sample or a full 2D image of an area of the sample or even a 3D image of a volume of the sample.

The inventors have found that the method of the present invention has produced surprisingly good results for identifying the extent of tumours and particularly Basal Cell Carcinoma (BCC) which is the most common form of skin cancer in caucasians. BCCs seldom metatise, but they can be locally invasive and with 80% occurring on the head and neck, their effects on the patient can be cosmetically traumatic.

The value of the first parameter at a first time value is calculated relative to the value of the first parameter of the second time value. The second time value is chosen to coincide with a physical feature of the dataset with respect to time such as a minima, maxima, zero-crossing point, point of inversion, discontinuity etc.

The second time value preferably corresponds to a predetermined maxima or minima of the first parameter with respect to time.

For example, the value of the first parameter at the first time value, i.e. $E(t_1)$ may be divided or multiplied by the value of the first parameter at the second time value i.e. $E(t_2)$. Alternatively, these two values may be added or subtracted from one another. Some further numerical relationship may also be implemented.

It has been found for imaging tumours and especially BCCs that good contrast can be obtained if Z is plotted where Z:

$$Z = \frac{E(t_1)}{E(t_2)}$$

Preferably, t_1 is chosen as one of the time values ascending or descending to or from a local maxima or minima where t_2 is chosen to coincide with this local maxima or minima.

In a particularly preferred configuration, the first time value is chosen from the subsequent time values ascending from a local minima, where t_2 is chosen to coincide with that minima.

When constructing the image, it is necessary to derive Z for each pixel of the sample of interest. t_2 is tied to a physical feature of the time domain plot of the first parameter which may move in time from point to point. For example, t_2 may coincide with a minima which shifts slightly from point to point. The first time value t_1 may be fixed in time or may be located at a fixed time interval δt from t_2 such that t_1 moves with t_2 from point to point of the sample.

The first time value may be chosen to coincide with the signal received from a point of interest within the sample. For example, if the sample is measured using reflection geometry, the first value may coincide with the time in which the reflected pulse takes the reach the part of interest in the sample, e.g. if the sample is a BCC

sample, the time which it takes to be reflected from the tumour. If t_2 is chosen to coincide with a main minima obtained using reflection geometry, δt may be chosen as approximately the time which the radiation takes to travel from the sample surface to the tumour and back to the surface. In practice, due to variations in the composition of the sample, t_1 will be determined experimentally.

As the contrast can be optimised by careful selection of the first time value, preferably, the method further comprises the step of determining the first parameter for a plurality of first time values and calculating these values of the first parameter relative to the value of the first parameter at the second time value. The method then further comprising the step of generating an image of the sample for each of the said first time values.

Thus, by viewing each of the images, an optimum first time value can be determined. Alternatively, each image may be analysed by means of a computer in order to derive the best value of t_1 using computer means.

If t_1 is fixed for each point in a single image, then the image may be optimised by finding the optimum value of the fixed t_1. If t_1 is a located fixed time interval δt from t_2 for each point in the image, then the image may optimised by finding the optimum value of the time interval, δt.

Good contrast may also be obtained by careful choice of the fixed time interval between the first and second time values regardless of whether or not the second time value corresponds to a physical feature of the time domain data set. Thus, in a second aspect, the present invention provides a method of imaging a sample, the method comprising:

(a) irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) determining a first parameter related to the amplitude of the radiation, which is either reflected from and/or transmitted by the sample, in the time domain;

(c) calculating the value of the first parameter at a first time value relative to the value of the first parameter at a second time value; and

(d) generating an image by plotting the value calculated in step (c) for different points of the sample, wherein the time interval between the first and second time values is constant for all points used to generate the image.

The value of the first parameter for the first time value with respect to the second time value may be calculated using any of the methods suggested in relation to the first aspect.

As enhanced contrast may be obtained by measuring the time domain signal at just two points, the image acquisition time may be substantially reduced. However, in some cases, it may be necessary to still acquire the data for more than two time values in order to derive the first parameter.

The first parameter may be the amplitude of the radiation itself. However, preferably, the detected amplitude is processed in order to obtain a first parameter.

When imaging non-rigid samples, it is preferable if a flat surface is provided for receiving the radiation. This is typically provided by a member such as a window which is transparent to the irradiating radiation. The sample is then pressed against this window to provide a flat analysis surface. However, the window itself can cause its own problems. This is because the window, has two interfaces, a front interface upon which the irradiating radiation is directly incident and a back surface which abuts the sample. Both of these surfaces will reflect radiation given rise to at least two unwanted signals in the detected reflected radiation. We refer to at least two reflections because multiple internal reflections may occur within the window giving even further unwanted effects.

Preferably, if the sample is irradiated through such a member, steps are taken in order to remove this so-called "baseline" signal. Thus, if step (a) comprises a step of irradiating the sample through a member which is transparent to the irradiating radiation and which abuts the sample, the method preferably further comprises the step of

obtaining a baseline signal by irradiating a member in the absence of the sample and detecting the amplitude of the reflected radiation. The first parameter is then determined by subtracting the baseline signal from the detected amplitude of the reflected radiation from both the member and the sample.

The baseline signal can be subtracted from the amplitude of the detected reflected radiation in the time domain or the frequency domain. The frequency domain signal is created by Fourier transforming the time domain signal.

In addition to or as an alternative, it may also be desirable to obtain a reference signal which is obtained by replacing the sample with the reference object of known reflectance or transmittance and measuring the amplitude of radiation reflected from the reference object (if the sample is to be measured in reflectance mode) or the transmittance of the object (if the sample is to be measured in transmission mode). The first parameter is then derived by dividing the detected amplitude of the radiation from the sample with the reference signal in the frequency domain.

Where both a reference signal and a baseline signal are obtained, the baseline signal is subtracted from both the sample signal and the reference signal and the baseline subtracted sample signal is then divided by the baseline subtracted reference signal. This division is performed in the frequency domain.

It is possible to interchange the reference and baseline signals. For example, the baseline signal may be obtained by replacing the sample with an object of known reflectance or transmittance, for example, a Quartz window and the reference signal may be obtained in the absence of a sample or a reference object.

Preferably, the amplitude of the detected radiation or the amplitude of the detected radiation which has been baseline subtracted and/or divided by a reference signal is then filtered. More preferably, this is achieved by multiplying the amplitude which may have been baseline subtracted and/or divided by a reference signal in the by the complex Fourier transform of a filter function F(t), where F(t) is a non-zero function

whose integral between time limits t_a and t_b is zero, and where t_a and t_b are chosen to encompass the part of interest of the reflected or transmitted pulse of radiation.

To clarify, radiation is supplied to the sample in the form of a pulse. The pulse will have a finite length. As the pulse passes through free space and through the sample, the length of the pulse may change. It is important to set t_a and t_b so that integration is performed over the part of the pulse of interest.

Considering the case of transmission, if the sample has the same properties as free space, then the radiation will pass through the sample without any delay. However, if the sample introduces a time delay into the beam, then t_a will be set to a negative value which has a magnitude is equal to or larger than the expected value of this time delay. T_b will generally, be set to the positive value of t_a.

Similarly, during reflection, t_a will be set to a negative value which has a magnitude is equal to or larger than the expected value of this time delay introduced by the pulse being reflected from the deepest point of interest in the sample. t_b will generally, be set to the positive value of t_a.

Generally, the method of the above aspect of the present invention will be performed using the above described apparatus where the radiation is measured using a reference beam and wherein a scanning delay line is introduced into the path of the reference beam or irradiating beam in order to measure the phase change introduced by the sample. In this situation, t_a and t_b will be the negative and positive limits of the scanning delay line. These may be set to the duration of the pulse or possibly a shorter time range.

Preferably function F(t) comprises a Gaussian component. Generally, t_a and t_b are symmetric about 0, preferably F(t) is also symmetric about zero.

As the signal is usually digitally sampled, strictly a summation is performed as opposed to an integral.

A particularly preferable form of F(t) is provided by:

$$F(t)=\frac{2}{\pi}\left\{\frac{e^{-2\left(\frac{t}{\alpha}\right)^2}}{\alpha}-\frac{e^{-2\left(\frac{t}{\beta}\right)^2}}{\beta}\right\}$$

where α and β are constants.

Preferable α is substantially equal to the shortest pulse length of the beam of pulsed radiation and β is set to be much longer than the pulse length, typically 5 to 100 times the pulse length. However, both of these values will generally be optimised by the operator.

If β is greater than or comparable to the time which it takes the radiation to penetrate the sample to the point of interest then, F(t) can take the simplified form:

$$F(t)=\frac{2}{\pi}\left\{\frac{e^{-2\left(\frac{t}{\alpha}\right)^2}}{\alpha}-\frac{1}{T}\right\}$$

where α is a constant which is substantially equal to the shortest pulse length of the beam and T is substantially equal to the time which it takes the beam of radiation to penetrate to the deepest point of interest in the sample.

A sample signal which has been baseline subtracted, divided by a reference signal and filtered is referred to as the impulse function. Preferably, the first parameter is the impulse function.

As the present invention is concerned with producing an image, the actual value of the impulse function at a pixel does not have to have a strong numerical correlation with a physical parameter of the sample. However, it is important that the impulse function is capable of showing variations in the composition of the sample.

The first time value may be determined by:

(i) selecting n regions of said sample, where n is an integer of at least 2;

(ii) producing n time domain spectra of a second parameter which is at least related to the amplitude of the radiation reflected from and/or transmitted by the sample. where the nth spectra is obtained by averaging the second parameter across the nth area and plotting this averaged value, as it varies with delay time;

(iii) determining the first time value by comparing at least two of the n spectra derived in step (ii).

Preferably, each spectra is normalised to the value of the second parameter at a time value which coincides with a physical feature of the second parameter with respect to time, prior to comparing the spectra. A physical feature of the spectra of the second parameter may be a minimum, maximum, point of inflection etc. Preferably, the same physical feature is chosen which identifies the second time value.

The spectra may be compared by performing a mathematical operation on at least two of the spectra. For example, the spectra may be divided by or subtracted from one another.

The selected regions may be lines or areas of the sample. They may be continuous lines or areas or discontinuous lines or areas. The regions may be of the same size of different sizes and may comprise any number of pixels.

Preferably, the regions selected by first generating a preliminary image of an area of the sample by plotting a preliminary imaging parameter which is at least related to the amplitude of the reflection reflected from and/or transmitted by the sample.

The preliminary imaging parameter may be any parameter which can be used to generate an image of the sample, for example, it may be the value of the first parameter and/or a second parameter, at a minima, maxima or point of inflection, etc. Alternatively, it may be the value of the first parameter or second parameter at an

arbitrary time value or a time value which has previously been used to show contrast in the sample or similar samples.

In the simplest case, just two areas are selected. However, the method may also be used to optimise contrast between 3 or more different regions. To achieve this, three regions are chosen, the value of the preliminary imaging parameter differing between said three selected regions, the first time value being chosen to coincide with a region where there is a difference between the three spectra.

The first parameter and second parameter may be the same or different. The second parameter may be preferably chosen from any of the parameters previously described in relation to the first parameter.

Regions of the sample may be chosen at random or, alternatively, regions of the sample may be chosen where some contrast has already been observed. In this example, the n areas are chosen such that the level of the preliminary imaging parameter within each region is substantially constant and the level of the preliminary imaging parameter is different between at least two of the regions.

Preferably, a standard deviation of about 30% in the primary imaging parameter is allowed within a selected region, more preferably 20%, even more preferably 10%.

The areas may be selected randomly by a person. Alternatively, the area selection may be automated. For example, a processor may be used to select n regions at random or sequentially and then compare the spectra from these regions in order to optimise the time value. The processor is preferably configured to select each area such that the preliminary imaging parameter is substantially constant within each area as previously described. The processor may further be configured to vary the size of the selected region in order to select regions which fall within the above criteria.

The above method of deriving a first time value may also be advantageously used with the method described in GB 2 360 186. Thus, in a third aspect, the present invention provides a method of imaging a sample, the method comprising:

(a) irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) determining a first parameter at least related to the amplitude of the radiation, which is either reflected from and/or transmitted by the sample, in the time domain; and

(c) generating an image of the sample by plotting the value of the first parameter at a first time value, said first time value being determined by:

(i) selecting n regions of said sample, where n is an integer of at least 2,

(ii) producing n time domain spectra of a second parameter which is at least related to the amplitude of the radiation reflected from and/or transmitted by the sample, where the nth spectra is obtained by averaging the second parameter across the nth area and plotting this averaged value, as it varies with delay time;

(iii) determining the first time value by comparing at least two of the n spectra derived in step (ii).

In a fourth aspect, the present invention provides imaging apparatus comprising:

a source for irradiating a sample with a pulse of electro-magnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

a detector for detecting the amplitude of radiation which is either reflected from or transmitted by the sample;

means for determining a first parameter which is related to the amplitude of the reflected and/or transmitted radiation in the time domain;

calculating means for calculating the value of the first parameter at a first time value relative to the value of the first parameter at a second time value, wherein the second time value coincides with a physical feature of the dataset of the first parameter with respect to time; and

means for generating an image by plotting the value calculated by the calculating means for different points of the sample.

In a fifth aspect , the present invention provides an apparatus for imaging a sample, the apparatus comprising:

a source for irradiating a sample with a pulse of electro-magnetic radiation, said pulse having a plurality of frequencies in the range from 25GHz to 100THz;

detector means for detecting the amplitude of radiation reflected from and/or transmitted by the sample;

means for determining a first parameter related to the amplitude of the radiation in the time domain;

calculating means for calculating the value of the first parameter at a first time value with respect to the value of the first parameter at a second time value, and

imaging means for generating an image by plotting the value calculated by the calculating means for different parts of the sample, wherein the time interval between the first and second time intervals is constant for all points used to generate the image.

The detector may be a direct detector of THz radiation or an apparatus according to any of the above three aspects of the invention, the detector may be a direct detector of THz radiation or it may be of the type which converts THz radiation into an easily readable signal.

For example, the detector may comprise a non-linear crystal which is configured such that upon irradiation of a probe beam and a THz beam, the polarisation of the probe beam is rotated. The probe beam can be of a frequency which can be easily measured (for example near infra-red). Typical crystals which exhibit this effect, the so-called "AC Pockels" effect are GaAs, GaSe, $NH_4H_2PO_4$, ADP, KH_2PO_4, KH_2ASO_4, Quartz, $AlPO_4$, ZnO, CdS, GaP, $BaTiO_3$, $LiTaO_3$, $LiNbO_3$, Te, Se, ZnTe, ZnSe, $Ba_2NaNb_5O_{15}$, $AgAsS_3$, proustite, CdSe, $CdGeAs_2$, $AgGaSe_2$, $AgSbS_3$, ZnS, organic crystals such as DAST (4-N-methylstilbazolium. This type of detection mechanism is generally referred to as 'Electro-optic sampling' or EOS.

Alternatively, the detector could be a so-called photoconducting detector. Here, the detector comprises a photoconductive material such as low temperature grown GaAs, Arsenic implanted GaAs or radiation damaged Si on Sapphire. A pair of electrodes, for example in a bow-tie configuration or in a transmission line configuration are provided on a surface of the photoconductive material. When the photoconductive material is irradiated by the reflected radiation and also, the probe

beam, a current is generated between the two electrodes. The magnitude of this photovoltage current is an indication of the magnitude of the THz signal.

Although it is possible to generate THz radiation directly, the most effective THz generation can be achieved by converting a pump beam into a THz beam. To do this, the source comprises a frequency conversion member and a source of a pump beam.

There are many possible options for the frequency conversion member. For example, the frequency conversion member may comprise a non-linear member, which is configured to emit a beam of emitted radiation in response to irradiation by a pump beam. Preferably, the pump beam comprises at least two frequency components, (or two pump beams having different frequencies are used), the non-linear member can be configured to emit an emitted beam having a frequency which is the difference of the at least two frequencies of the pump beam or beams. Typical non-linear members are: GaAs or Si based semiconductors. More preferably, a crystalline structure is used. The following are further examples of possible materials :

$NH_4H_2PO_4$, ADP, KH_2PO_4, KH_2ASO_4, Quartz, $AlPO_4$, ZnO, CdS, GaP, $BaTiO_3$, $LiTaO_3$, $LiNbO_3$, Te, Se, ZnTe, ZnSe, $Ba_2NaNb_5O_{15}$, $AgAsS_3$, proustite, CdSe, $CdGeAs_2$, $AgGaSe_2$, $AgSbS_3$, ZnS, GaSe or organic crystals such as DAST (4-N-methylstilbazolium).

In order to produce an emitted beam having a frequency in the THz regime, preferably the at least two frequencies of the pump beam or beams are in the near infra-red regime. Typically, frequencies between $0.1x10^{12}$Hz and $5x10^{14}$Hz are used.

Alternatively the frequency conversion member is a photoconducting emitter, such an emitter comprises a photoconductive material such as low temperature grown or arsenic implanted GaAs or radiation damaged Si or Sapphire.

Electrodes which may be of any shape such as a dipole arrangement, a double dipole arrangement, a bow-tie arrangement or transmission line arrangement are

provided on the surface of the photoconductive material. At least two electrodes are provided. Upon application of a bias between the electrodes and irradiation of a pump beam(s) having at least two different frequency components, a beam of radiation is emitted having a frequency different to that of the at least two frequency components of the pump beam or beams.

When a pulse having a plurality of frequencies passes via a sample to a detector, the various frequencies will not arrive at the detector at the same time due to the frequency dependent response of the sample. A time domain signal can be established by measuring the amplitude of the detected radiation with respect to time. In order to achieve this, it is preferable if a scanning delay line is inserted into either the path of the probe or pump beam. The delay line can be configured to scan over the whole length of the pulse.

As mentioned with respect to the method of the present invention, the first time value can be selected in order to optimise contrast in the image. The first time value may be fixed with respect to zero across an image, or it may be fixed with respect to the second time value.

Therefore, preferably, the apparatus further comprises means for generating a plurality of images corresponding to different first time values, the apparatus also comprising means for scanning through said images by varying said first time value in order to determine a first time value which allows the best contrast.

The apparatus preferably further comprises:

(i) means for selecting n regions of said sample, where n is an integer of at least 2;

(ii) means for producing n time domain spectra of a second parameter which is at least related to the amplitude of the radiation reflected from and/or transmitted by the sample, where the nth spectra is obtained by averaging the second parameter across the nth area and plotting this averaged value as it varies with delay time; and

(iii) means determining the first time value by comparing at least two of the n spectra.

In a sixth aspect, the present invention provides:

(a) means for irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) means for determining a first parameter, at least related to the amplitude of the radiation, which is either reflected from and/or transmitted by the sample, in the time domain; and

(c) means for generating an image of the sample by using the value of the first parameter at a first time value,

said apparatus being configured to determine said first time value by:

(i) selecting n regions of said sample, where n is an integer of at least 2;

(ii) producing n spectra in the time domain, where the nth spectra is obtained by plotting the second parameter, averaged across the nth area, against delay time;

(iii) determining the first time value by comparing at least two of the n spectra.

The means for selecting regions may be automatic means. For example, a program may be executed to select two or more regions of the sample at random or two or more regions of the sample which have the largest difference between their mean colour levels, the standard deviation of these areas may be calculated to ensure that the selected area does not cover an area of high contrast.

Alternatively, the apparatus may be provided with means which allow an operator to select two or more areas. For example, the apparatus may be configured to support a cursor which the operator may use to select particular areas of the image.

The present invention will now be described with reference to the preferred non-limiting embodiments in which:

Figure 1 is a schematic of a reflection imaging system which may be used in accordance with an embodiment of the present invention;

Figure 2 is a is a schematic of a transmission imaging system which may be used in accordance with an embodiment of the present invention;

Figure 3a is a perspective view of a sample holder and figure 3b is a plan view of a sample holder which may be used in accordance with an embodiment of the present invention;

Figure 4 illustrates a plurality of deconvoled impulse functions for a sample of a skin tumour;

Figure 5 is a schematic illustrating the analysis of an impulse function in accordance with an embodiment of the present invention;

Figure 6a illustrates a visible image of a skin tumour, figure 6b illustrates a corresponding Terahertz image of the same sample produced in accordance with an embodiment of the present invention and figure 6c is a plot of the mean value of the THz signal of the image of figure 6b plotted for four pixels of the image;

Figure 7a is a vertical histology section of a skin tumour and figure 7b is a corresponding section through a Terahertz image obtained in accordance with an embodiment of the present invention;

Figure 8 is a schematic illustrating the analysis of an impulse function in accordance with a further embodiment of the present invention;

Figure 9 is a schematic illustrating the analysis of an impulse function in accordance with a yet further embodiment of the present invention;

Figure 10 is a schematic illustrating the analysis of an impulse function in accordance with an even further embodiment of the present invention;

Figure 11 is an apparatus which can optimise the contrast of a THz image in accordance with an embodiment of the present invention;

Figure 12 is an image of a tumour generated using E(t)/E(min) as explained with reference to Figure 5;

Figure 13 illustrates two time domain spectra averaged over areas "a" and "b" of Figure 12;

Figure 14 illustrates a spectra derived from the difference between the two spectra of Figure 13;

Figure 15 illustrates a spectra derived by dividing one spectra of Figure 13 by the other spectra of Figure 13;

Figure 16 illustrates a plot of two time domain spectra derived by plotting E(t)/E(min), the spectra are similar to those of Figure 13 but have been normalised to the value E(min);

Figure 17 illustrates a time domain spectra derived from the difference of the two spectra of Figure 16;

Figure 18 illustrates a time domain spectra derived from dividing one spectra of Figure 16 by the other spectra of Figure 16;

Figure 19 illustrates an image of a tumour derived by plotting the magnitude of the main minima of the impulse function for each pixel in the x and y direction;

Figure 20 illustrates the mean value of the minimum view averaged over each of the areas s1, s2, d1, d2, h1 and h2 indicated in Figure 19;

Figure 21 illustrates a plot of the mean impulse function in the time domain for areas s1, s2, d1, d2, h1 and h2 of Figure 19;

Figure 22 illustrates a plot similar to that of Figure 21 where the spectra have been normalised to the value E(min);

Figure 23 illustrates a detail of the plot of Figure 21;

Figure 24 illustrates a detail of the plot of Figure 22;

Figure 25 is a plot identical to that of Figure 21 but repeated to allow easy comparison;

Figure 26 is a plot identical to that of Figure 22 but repeated to allow easy comparison;

Figure 27 illustrates a detail of the plot of Figure 25;

Figure 26 illustrates a detail of the plot of Figure 26;

Figure 29 illustrates a spectra of two time domain spectra, the upper spectra being produced by subtracting the spectra obtained from area h1 in Figure 19 from the spectra obtained from the diseased tissue areas d1 and d2; the lower plot is a spectra obtained by subtracting a mean spectra obtained from healthy tissue area h1 of Figure 19 from inflamed tissue areas s1 and s2;

Figure 30 illustrates an enhanced image obtained by plotting the parameter E(t)/E(min) for t = 1.2 picoseconds; and

Figure 31 illustrates an image obtained by plotting E(t)/E(min) for t = 3.7 picoseconds.

Figure 1 shows a basic THz reflection imaging system. The system can be split into three main sections: a generation section 1, an imaging section 3 and a detection section 5.

A THz beam of radiation is produced in generating section 1. Generation section 1 comprises an optically amplified Ti:Sapphire laser 7 to produce a beam 9 of 250 fs pulses at 800 nm.

The optically amplified laser 7 comprises a solid state pump laser 11 and an amplified Ti:Sapphire section 13. Beam 9 is divided by 50:50 beam splitter 15. Beam splitter 15 divides the beam into a pump pulse 17 which is used to generate THz radiation and a probe pulse 19 which is used in the detection of the radiation reflected from the sample.

THz pulses are generated by optical excitation and charge acceleration of biased GaAs stripline antenna 21 which outputs THz pulses 23 which have a bandwidth of 0.1 THz to 3 THz. The emitted beam 23 is then passed into imaging section 3.

The irradiating THz beam 23 is then directed via first off-axis parabolic mirror 25 onto second off-axis parabolic mirror 27 and then focussed to a 400μm spot on sample 29. Sample 29 is mounted on motorised stage 31. The sample mount 31 is computer controlled and can be configured to move the sample such that each section (pixel) of the sample in question can be imaged.

The reflected radiation is then collected via third off-axis parabolic mirror 33 and reflected via planar mirror 35 into the detection system 5. In the detection system, the reflected radiation 37 is reflected off-axis parabolic mirror 39. Off-axis parabolic mirror 39 has a hole through it which allows the reflected beam 37 to be combined with the probe beam 19 obtained by beam splitter 15.

Prior to combining the probe beam 19 with the reflected radiation 37 at off-axis parabolic mirror 39, the probe beam 19 is reflected off planar mirror 41 into scanning delay line 43. Scanning delay line comprises cuboid mirror 45 which moves back and forth under computer control 47 in order to vary the pulse length of the probe beam 19 with respect to the pump beam 17 and hence the irradiating and reflected radiation.

Thus, the relative phase of the probe beam 19 can be varied with respect to the pump pulse 19.

The combined probe beam and reflected radiation 49 is then collimated and focused onto ZnTe detection crystal 49. The THz radiation is detected using the linear electro-optic Pockel's effect.

The linearly polarised probe beam 19 has its polarisation oriented such that it has components along both the ordinary and extra-ordinary axis of detection crystal 49. Each of the axis has distinct refractive indices n_0 and n_t along the ordinary and extra-ordinary axis of the crystal 49 respectively. In the absence of reflected radiation 37, the linearly polarised probe beam 19 passes through the detection crystal with a negligible change in its polarisation.

The applicant wishes to clarify that the angle through which the polarisation is rotated is negligible. However, the linearly polarised beam can become slightly elliptical. This effect is compensated for by a variable retardation waveplate, for example, quarter waveplate 51.

The beam emitted by detection crystal 49 is converted into circularly polarised beam by quarter waveplate 51. The beam is then split into two linearly polarised beams by Wollaston prism 53 (or an equivalent device for separating orthogonal polarisation components) which directs the two orthogonal components of the polarised beam onto a balance photo diode assembly 55. The balance photo diode signal is adjusted using waveplate 51 so that the difference in outputs between the two diodes is zero when only the probe beam 19 is incident on the detection crystal 49.

However, if the detector also detects reflected radiation 37, the angle through which the polarisation is rotated by is not negligible. This is because the THz electric field modifies the refractive index of the probe beam 19 along one of the axis n_e, n_o. This results in the physical field emitted by detection crystal 49 being elliptical. Therefore, the polarisation components separated by prism 53 are not equal and the

difference in voltage between the output diodes of balance photo-diode assembly 55 gives a detection voltage.

The rotation in the polarisation is obscured if the phase of the reflected radiation 37 reaching the detection crystal 49 is different to the phase of the probe beam 19. As the sample 29 will introduce a different phase change for different frequency components of the irradiating radiation 23, the delay line 43 is swept under control of computer 47 in order to sweep the phase of the probe beam 19 with respect to the reflected radiation 37 and hence detect the delay time of each of the frequency components of reflected radiation 37. This produces a so-called time-domain signal.

Computer control 47 also is used to control the movement of stage 31 in order to ensure that a full data set is collected for each pixel of interest of sample 29.

The above system illustrates is configured for reflection measurements. However, it will be appreciated that those skilled in the art that the system could also be adapted for a transmission method. Such an apparatus is shown in Figure 2. To avoid unnecessary repetition, like reference numerals will be used to denote like features. In imaging section 3, first 25 and second 27 parabolic mirrors are configured to focus radiation onto sample 29. The third parabolic mirror 23 then collects the transmitted radiation and directs it into detection section 5. The generation 1 and detection 5 sections are the same as those for the reflection apparatus of Figure 1.

The method of the present invention has been found to be of particular use in imaging Basal Cell Carcinoma. Figure 3 illustrates a preferred mounting method for a sample of this type. Figure 3a shows a perspective view of the sample holder whereas Figure 3b shows a cross section of the sample holder.

Figures 3a and 3b show two samples 61 and 63 located in sample holder 65. The sample holder 65 is based on a modified liquid cell of the type used for Fourier transform infra-red spectrometry. The liquid cell is cylindrical in shape and comprises a hollow cylindrical container 66, a 2mm thick Quartz plate 67 with a diameter of 25mm,

is placed at the base of container 66. The THz radiation irradiates the sample through Quartz window 67.

As water is a strong absorber of radiation in the THz regime, excess fluid is removed from the sample using lint free paper. The samples were mounted on the Quartz window 67 using tweezers and the top surface of the skin was placed on Quartz window 67. Sponge 69 is then placed on top of the sample so that the samples are sandwiched between the sponge 69 and Quartz window 67. The retaining ring 71 is placed on top of Quartz window 67 in order to keep the window level and also to maintain the position of the samples 61 and 63.

A polythene window 63 is provided on the opposing side of the sponge to the Quartz window and in contact with the sponge. The polythene window is held in place by retainer plate 75 which is screwed into the sample holder 65. This construction provides an airtight holder for the sample which prevents the sample from drying out during imaging.

Figure 4 illustrates a time domain impulse function for diseased and healthy tissue. Figure 4 shows the impulsed function along the y axis plotted against time delay (arbitrary units) along the x axis. Five traces are shown in Figure 4 for five different pixels.

For each pixel, the entire THz waveform is recorded and averaged over four readings to reduce any fluctuations arising from the laser. The signal to noise ratio is approximately 1,000 to 1.

The impulse function is derived from the amplitude of the reflected radiation. In order to derive the input function for a particular pixel, two further measurements are made in addition to a measurement of the sample.

First, a baseline signal is measured. In the reflection measurement, it is important to try to reduce the effects of spurious reflections. The apparatus of Figure 3 uses a Quartz window 67 through which the samples 61 and 63 are illuminated. Quartz

window 67 has a lower interface 79 which will cause spurious reflections and also an upper interface 81 which will cause spurious reflections.

It is important to try to eliminate the effects of these two interfaces from the measurements. Therefore, prior to imaging the sample, a second Quartz window (not shown) is placed in the sample holder instead of sample 61 and 63. The amplitude of the detected THz signal is measured for each pixel so that a baseline signal is obtained through each pixel. Variations in Quartz window 67 and also variations in the incident angle of the irradiating radiation may cause artificial variations of the measurements of the sample. Therefore, the baseline is obtained for each pixel.

The baseline measurement was averaged twenty times to obtain baseline signal B(t).

A reference spectrum is also measured. This is measured by using the sample holder with just Quartz window 67 on its own. This reference signal will be represented in the time domain as **R(t)**.

First, in order to obtain an impulse function for one waveform, the baseline signal **B(t)** is subtracted from the measured time domain waveform **S(t)**. This operation can be formed in either the time domain or frequency domain.

This baseline subtracted waveform should provide data where the spurious reflections due to the lower 79 and upper 81 interfaces of Quartz window 67 have been eliminated.

Next, the baseline subtracted waveform is complex Fourier transformed to give:

$$\mathbf{S'(\nu)-B'(\nu)}$$

In order to use the reference signal, the baseline signal is again subtracted from the reference waveform **R(t)**. This subtraction operation can be performed in either the

time domain or the frequency domain. The result is then complex Fourier transformed to give:

$$\mathbf{R'(\nu)-B'(\nu)}$$

The baseline subtracted measured data is then divided by the baseline subtracted reference signal in the frequency domain to give:

$$\frac{S'(\nu)-B'(\nu)}{R'(\nu)-B'(\nu)}$$

The data is then filtered. In general, a filter function will be described by **F(t)** and is complex Fourier transform will be represented by a **F'(ν)**.

A filter function preferably is used because the THz pulse system can generate and detect pulses comprising frequencies over some finite range, typically from less than 100GHz to over 3Thz. There is a high frequency limit above which the THz signal falls below the noise level of the detection system.

Similarly, the THz signal level falls below the noise level at low frequencies. Thus, there is a need to remove the high and low frequency noise. A particularly preferable function for achieving this is:

$$F(t)=\left|\frac{2}{\pi}\left\{\frac{e^{-2\left(\frac{t}{\alpha}\right)^2}}{\alpha}-\frac{e^{-2\left(\frac{t}{\beta}\right)^2}}{\beta}\right\}\right|$$

The parameters α and β are selected to control the high and low frequency roll-off of the function. α is set approximately the shortest THz pulse length (half cycle) obtainable within the THz system. β is set to be much longer than the THz pulse. In

operation, the two parameters are optimised manually by the operator to obtain the best compromise between bandwidth and noise.

As the above function comprises two Gaussian functions with similar areas but opposite signs, the above function ensures that the integral of the filter function for all time is zero.

If the value of β is comparable to or greater than the total time-delay scan range, then an alternative function can be used:

$$F(t)=\left|\frac{2}{\pi}\left\{\frac{e^{-2\left(\frac{t}{\alpha}\right)^2}}{\alpha}-\frac{1}{T}\right\}\right|$$

where T represents the total range of delay times used i.e. $T=T_{max}-T_{min}$. This ensures that the overall integral from T_{min} to T_{max} is always zero.

The filter function is multiplied by the complex Fourier transform of the signal which is being corrected by the baseline and the reference.

Once the impulse function has been derived as illustrated in Figure 4, the image is then generated using the analysis described with reference to Figure 5.

Figure 5 shows a schematic impulse function for just one pixel. In order to generate an image of the sample, it is important to derive a meaningful parameter from this trace which can be plotted for each pixel. By a meaningful parameter, we mean a parameter which can be used to indicate and emphasise variations in the sample.

The inventors have surprisingly found that the following method of deriving a parameter provides an image with a remarkable sharp contrast for illustrating a number of variations in the sample but has found to be of particular use for indicating the presence of Basal Cell Carcinoma.

The value of the impulse function at the minima is measured. Also, the value of the impulse function at a time t. In this example, the time t is chosen from the times which correspond to ascending values of the first parameter after the minima.

A parameter Z is derived by dividing the first parameter at a time t by the value of the first parameter at the minima i.e.

$$Z = \frac{E(t)}{E(\min)}$$

To generate an image, each value of Z is assigned a colour and a full colour map is generated by plotting z on the z axis against x and y which define the pixel positions. As the contrast varies with the chosen value of t, t is chosen in order to give the best contrast. "t" is fixed for the image.

Figure 6a illustrates a visual image of a non-recurrent tumour present on a nasal tip. The tumour is 7mm in diameter. The tumour 91 is shown on the left of the image with healthy tissue 93 shown on the front of the image. A histology cut is indicated on both diagrams 6a and 6b with a dotted line. The importance of this will be described with reference to Figure 7.

The false colour THz image 6b shows a strong contrast between the diseased and healthy tissue. The healthy tissue 93 shows a fairly uniform greyscale contrast throughout. The strong contrast between the diseased and healthy tissue which is illustrated in Figure 6b is not present in the visible image where it is difficult to differentiate between the two.

Four areas, d1, d2, h1 and h2 corresponding to two areas of diseased tissue and two areas of healthy tissue are indicated on Figure 6b. Figure 6c illustrates a plot of the main value of the impulse function for each of these areas. It can be seen that the mean value for the two diseased areas is much higher than that of the mean value for the two

healthy areas. The standard derivation is used to illustrate an error bar for each of the regions. It can be seen that even in the worse case scenario, it is possible to clearly distinguish the diseased areas d1 and d2 from healthy areas h1 and h2.

A histological analysis of the sample was then performed. A vertical histology section was taken through the sample illustrated in Figures 6a and 6b along the dotted line. The results of which are indicated in Figures 7a and 7b.

In the histology section shown in Figure 7a, the epidermis is shown as a thick line 101 of histology section 7a. It is noted that the epidermis is missing in the centre of the section. This is because the tissue is preferentially scraped prior to excision. As a result of increased vascularisation around the tumour site and the changing composition of the skin structure within this region, the epidermis is more delicate and susceptible to preferential bleeding in the surrounding healthy tissue. This effect is utilised by the surgeons in assisting identifying the location of the tumour. The tumour had not broken through to the surface before preferential scrapping.

Tumour shrinkage is due to desification, a result of the histological preparation process. The tumour is highly cellular and has a fibrous casing. The epidermis is pushed to one side by the tumour and as a result, thickened.

Turning to the THz image of Figure 7b, the upper line indicated on the plot is the mean value for the diseased tissue. The lower line is the mean value for the healthy tissue. The THz image is taken in the same direction as that of the histology section. In other words, it is just a slice through the greyscale image of Figure 6b taken along the dotted line. Towards the right hand side of the trace, diseased tissue is seen which corresponds to the histology section. Towards the left hand side of the trace, the healthy tissue is seen as shown in the histology section.

In the analysis technique shown in Figure 8, the impulse function is derived in the same manner as that described with reference to Figure 5. The parameter Z is then derived by subtracting the value of the first parameter at the minima from the value of

the first parameter at an arbitrary time t which occurs at a time δt after the minima i.e. $Z = E(\min+\delta t) - E(\min)$.

The minima may move in time from point to point across the sample. t is defined relative to E(min), thus t moves with the minima. To optimise the image, the parameter δt is optimised as δt is fixed for each image.

Figure 9 shows a variation on the analysis technique described with reference to Figure 8. Again, the impulse function is derived in the same manner as described with reference to Figure 5.

The parameter Z is calculated by subtracting the value of the first parameter at a time t_2 from the value of the first parameter at time t_1 where $t_1 = t_2 + \delta t$. t_2 does not have to coincide with the minima. As shown in Figure 9, t_1 and t_2 may lie on opposing sides of the minima or they may lie on the same side.

Again, as described with reference to Figure 8, since t_1 is defined relative to t_2, δt is optimised.

Figure 10 shows a yet further technique for deriving Z. Again, as described with reference to Figure 5, the impulse function is derived as a function of time. Z is then calculated by dividing the value of the first parameter at a time t which equals $t_{min} + \delta t$ divided by the value of the first parameter at the minima. In other words:

$$Z = \frac{E(\min+\delta t)}{E(\min)}$$

This technique is similar to that described in relation to Figure 5. However, in the analysis technique of Figure 5, t_1 is fixed across a single image. If the minima moves (and hence t_2 moves from pixel to pixel) t_1 still remains fixed. In the analysis technique described with reference to Figure 10, t_1 is defined in relation to t_2, thus t_1 moves with t_2 and δt is fixed for a particular image.

As mentioned with reference to Figures 5, 8, 9 and 10, careful selection of t enhance the contrast. Generally, the operator of such equipment will be able to scan through a number of images at different t in order to obtain the best contrast. Such an apparatus for performing this is illustrated in Figure 11.

Figure 11 shows a schematic view of such a system. The system comprises a controller 111, a THz imaging system 113. The THz imaging system is of the type described with reference to either of Figures 1 and 2. To avoid unnecessary repetition, details of it will not be repeated here and the whole system will be represented by box 113. Controller 111 has a screen 115 which the operator can use to view a false colour THz image of the sample under test as explained with reference to Figures 4 and 5 and as illustrated in Figure 6b. The controller 111 has two modes of operation. A set-up mode and a fixed imaging mode. The controller switches between these two modes using switch 117.

In the set-up mode, the full THz spectrum and hence impulse function is derived for each pixel of the sample. The image is then derived by calculating Z for a given t. t may be a fixed value for all pixels or may be fixed from a second time value which may correspond to a physical feature of the data set. The user can then turn dialler 119 and watch the contrast of the image displayed on screen 115 change as t or δt is varied. The operator can thus choose an optimum value of t or δt.

Alternatively, the controller itself may be provided with its own means for determining the optimum contrast as a function of t. For example, the controller can calculate the mean value of Z for each pixel and choose a value of t where the variation in Z is at its largest. Alternatively, more complex statistical techniques may be employed where only representative values of the impulse function are considered in determining the best contrast.

Once the operator or controller itself 111 has determined the best contrast, the operator can then switch 117 to switch the controller 111 onto fixed mode. In this fixed mode, the controller will instruct imaging apparatus 113 to only take data at the minima

and at the preferred value of t. This technique substantially reduces the acquisition time for the image. Once t has been determined for a sample, it can be used for that whole sample. It has also been found that the cell value of t can be used as samples of the same type. For example, for samples which are Basal Cell Carcinoma, the operator would expect to use the same value of t.

Controller 111 can be pre-programmed with t values which are known to provide optimum contrast for certain samples. Alternatively, controller 111 may be provided with a look-up table such that the operator can enter the type of sample which is to imaged and the controller can automatically select the correct t or δt value.

Figure 12 illustrates an image of a skin tumour obtained by dividing the magnitude of the impulse function at a time t = 1098 (arbitrary units) with the impulse function at the minima of the electric field. This is explained in detail with reference to Figure 5. Specifically, the value of E(t)/E(min) for t = 1098 is plotted for each pixel to produce the image.

This figure will now be used to explain how the image may be optimised. Two areas "a" and "b" are selected from the image. Area "a" is believed to represent diseased tissue whereas area "b" is believed to represent healthy tissue.

The mean impulse function for both areas "a" and "b" is plotted for each time point as shown in Figure 13. Formally, for area "a", the function :

$$\overline{E_a}(t)$$

is plotted where

$$\overline{E_a}(t) = \frac{\sum_{i}^{n} E_i(t)}{n}$$

E being the impulse function and n being the number of pixels in area "a". The mean impulse function for area "b" is calculated in the same manner and plotted.

The width of the curves represents the standard deviation since the waveform is calculated over an area.

These two waveforms are then compared in order to determine the time values where a plot of E(t) will produce an image with the best contrast between areas "a" and "b".

Figure 14 illustrates the case where the spectra for the diseased .issue is subtracted from the waveform for the healthy tissue. It can be seen that the largest difference between these two waveforms occurs close to t=1100. Thus, indicating that an image produced using this value of "t" as the first time value will produce good contrast. For example, the image may be produced by plotting E(t), E(t)/E(min), E(t)-E(min) etc.

Figure 15 illustrates the case where the two waveforms are compared by dividing the waveform for healthy tissue with that from diseased tissue. This comparison method also allows a value of "t" to be determined which can be used as the first time value.

In Figure 13 the mean impulse function was derived for two areas and used to determine an optimum value of "t". In Figure 16, the mean impulse functions of Figure 13 are normalised to the value E_{min}, formally:

$$\frac{\overline{E_a}(t)}{E_{\min}}$$

is plotted, where:

$$\overline{E_a}(t) = \frac{\sum_{i}^{n} E_i(t)}{n}$$

E being the impulse function and n being the number of pixels in area a. The normalised mean impulse function for area 'b is calculated in the same manner and plotted.

Figure 17 illustrates the spectra of Figure 16 subtracted from one another whereas Figure 18 illustrates the spectra from Figure 16 divided by one another.

In the same manner as described with reference to Figures 14 and 15, values of "t" may be derived from the spectra of Figures 17 and 18 which produce an image with good contrast.

The present invention may also be used to compare three different types of tissue within the sample and different t values may be selected to identify different types of tissue.

In Figure 19, six areas are selected, h1 and h2 which represent healthy tissue, s1 and s2 which represent inflamed tissue and d1 and d2 which represent diseased tissue.

The image of Figure 19 is generated by plotting the minima of the impulse function for each pixel in the x and y directions i.e. Emin is plotted. The minimum mean impulse function for each of these areas, averaged across each area, is then plotted and the graph shown in Figure 20. The vertical error bars represent the standard deviation.

Figure 21 illustrates the mean impulse function E(t) for healthy tissue (calculated from areas h1, h2), inflamed tissue (calculated from areas s1 and s2) and diseased tissue (calculated from areas d1 and d2) derived from the sample illustrated in figure 19. Figure 22 illustrates the spectra of figure 21 which has been normalised to

the value of E_{min} i.e E(t)/E(min). The width of the spectra represents the standard deviation.

It can be seen by viewing figures 20 and 21 side by side that the shape of the waveforms change between a plot of mean E(t) and mean $E(t)/E_{min}$. The changes between these two waveforms may be used to distinguish between different tissue types.

Figures 23 and 24, illustrate in detail the region between t = 0.8 ps and t=1.6 ps for figures 21 and 22 respectively. In figure 23, it can be seen that there is little difference between the curves for the diseased and inflamed tissue which just start to deviate from one another at about t = 1.1 ps. However, the spectra which represents healthy tissue is clearly separated from the spectra from both diseased and inflamed tissue over the whole of the selected time range.

In figure 24, the spectra from the healthy and inflamed tissue is seen to follow a similar trajectory whereas the spectra from the diseased tissue follows a lower path which is well separated from the other two spectra.

The time value of 1.2ps is clearly marked on both figures 23 and 24. Generating an image by plotting E(t=1.2ps) can thus be used to clearly distinguish between healthy tissue on the one hand and diseased and inflamed tissue on the other. Generating an image by plotting $E(t=1.2ps)/E_{min}$ can be used to clearly distinguish between diseased tissue on the one hand and health and inflamed tissue on the other.

By choosing a different time value, it is possible to distinguish between diseased and inflamed tissue. Figures 25 and 26 are identical to figures 21 and 22, but are repeated in order to allow comparison with figures 27 and 28. Figures 27 and 28 show a detail of the plots of figures 25 and 26 between 1ps and 6 ps.

In figure 27, the diseased tissue trace lies well below that of the traces for healthy tissue and inflamed tissue which follow a similar path. However, in figure 28, where $E(t)/E_{min}$ is plotted, a clear distinction can be seen between the three traces, the

diseases tissue trace being the lowest trace, the healthy tissue trace being the middle trace and the upper trace corresponding to inflamed tissue.

Thus, an image formed by plotting E(t=3.7 ps) can be used to distinguish between diseased tissue on the one hand and healthy and inflamed tissue on the other, whereas an image formed by plotting E(t)/E(min) where t = 3.7 ps can be used to distinguish between diseased tissue, healthy tissue and inflamed tissue.

When it is desirable to distinguish between three different tissue types, it is necessary to chose a time value which emphasises the difference between all three tissue types and not just two tissue types.

Figure 29 illustrates two spectra. The upper spectra is produced by subtracting the spectra formed by calculating $E(t)/E_{min}$ for healthy tissue from the spectra formed by calculating $E(t)/E_{min}$ for diseased tissue. The lower spectra is produced by subtracting the spectra formed by calculating $E(t)/E_{min}$ for healthy tissue from the spectra formed by calculating $E(t)/E_{min}$ for inflamed tissue.

In order to produce an image which clearly shows contrast between healthy and diseased tissue, a time value should be chosen where the magnitude of the upper spectra is greatest. To produce an image which clearly shows contrast between healthy and inflamed tissue, a time value should be chosen where the magnitude of the lower spectra is greatest. However, in order to distinguish between diseased and inflamed tissue, it is necessary to choose points where the difference between the two spectra is also large.

For example, in figure 29, both traces show a large negative value at about 0.7 ps. However, although taking this value of t will produce an image where there is good contrast between healthy tissue and diseased tissue and between healthy tissue and inflamed tissue, such an image is unlikely to show contrast between diseased tissue and inflamed tissue. However, producing an image of $E(t)/E_{min}$ for t=3.7ps or t=1.2 ps will show good contrast between all three tissue types.

Figures 30 and 31 illustrate images produced by plotting $E(t)/E_{min}$ for each pixel for t = 1.2 ps (figure 30) and t=3.7 ps(figure 31).

CLAIMS:

1. A method of imaging a sample, the method comprising:

(a) irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) determining a first parameter at least related to the amplitude of the radiation, which is either reflected from and/or transmitted by the sample, in the time domain;

(c) calculating the value of the first parameter at a first time value relative to the value of the first parameter at a second time value, wherein the second time value coincides with a physical feature of the first parameter with respect to time; and

(d) generating an image by plotting the value calculated in step (c) for different points of the sample.

2. A method according to claim 1, wherein said physical feature is a predetermined maxima or minima of the first parameter with respect to time.

3. A method according to either of claims 1 or 2, wherein step (c) comprises the step of dividing the value of the first parameter at the first time value with the value of the first parameter at the second time value.

4. A method according to either of claims 1 or 2, wherein step (c) comprises the step of subtracting the value of the first parameter at the second time value from the value of the first parameter at the first time value.

5. A method according to any preceding claim, wherein in step (c), the first time value is one of the values, ascending or descending, to or from, a local minima or maxima and said second time value coincides with said local minima or maxima.

6. A method according to claim 5, wherein step (c) comprises the step of choosing the second value to correspond to a predetermined minima of the first parameter and the first time value is chosen from the subsequent time values ascending from the said minima.

7. A method according to any preceding claim, wherein the first time value is fixed in time for each point of the sample used to create a single image.

8. A method according to any preceding claim, wherein the first time value is a fixed time interval from the second time value for each part of the sample used to create a single image.

9. A method according to claim 7, further comprising the step of determining the first parameter for a plurality of first time values relative to the first parameter for the second time value for each point, the method further comprising the step of generating a plurality of images of the sample corresponding to each of the said first time values.

10. A method according to claim 9, further comprising the step of selecting the optimum first time value by selecting the image which provides the best contrast.

11. A method according to claim 8, further comprising the step of determining the first parameter for a plurality of different time intervals relative to the second time value for each part, the method further comprising the step of generating a plurality of images of the sample corresponding to each of different time intervals.

12. A method according to claim 11, further comprising the step of selecting the optimum time interval by selecting the image which provides the best contrast.

13. A method of imaging a sample, the method comprising:

(a) irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) determining a first parameter at least related to the amplitude of the radiation, which is either reflected from and/or transmitted by the sample, in the time domain;

(c) calculating the value of the first parameter at a first time value relative to the value of the first parameter at a second time value; and

(d) generating an image by plotting the value calculated in step (c) for different points of the sample, wherein the time interval between the first and second time values is constant for all points used to generate the image.

14. A method of imaging a sample, the method comprising:

(a) irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) determining a first parameter at least related to the amplitude of the radiation, which is either reflected from and/or transmitted by the sample, in the time domain; and

(c) generating an image of the sample by plotting the value of the first parameter at a first time value, said first time value being determined by:

(i) selecting n regions of said sample, where n is an integer of at least 2,

(ii) producing n time domain spectra of a second parameter which is at least related to the amplitude of the radiation reflected from and/or transmitted by the sample, where the nth spectra is obtained by averaging the second parameter across the nth area and plotting this averaged value, as it varies with delay time;

(iii) determining the first time value by comparing at least two of the n spectra derived in step (ii).

15. A method according to any of claims 1 to 13, further comprising the steps of:

(i) selecting n regions of said sample, where n is an integer of at least 2;

(ii) producing n time domain spectra of a second parameter which is at least related to the amplitude of the radiation reflected from and/or transmitted by the sample, where the nth spectra is obtained by averaging the second parameter across the nth area and plotting this averaged value, as it varies with delay time;

(iii) determining the first time value by comparing at least two of the n spectra derived in step (ii).

16. A method according to either of claims 14 or 15, wherein step (ii) comprises generating a preliminary image of an area of the sample by plotting a preliminary imaging parameter which is at least related to the amplitude of the radiation reflected

from and/or transmitted by the sample; and selecting said n regions from said preliminary image.

17. A method according to claim 16, wherein said n regions are selected such that said n regions being selected such that the level of the preliminary imaging parameter within each region is substantially constant and the level of the preliminary imaging parameter is different between at least two of the said regions.

18. A method according to either of claims 16 to 17, wherein the preliminary imaging parameter is the value of the first and/or second parameter at a time which corresponds to a physical feature of the time domain spectra of the first and/or second parameter.

19. A method according to any of claims 14 to 18, wherein each spectra is normalised to the value of the third parameter at a time value which coincides with a physical feature of the second parameter with respect to time, prior to comparing the spectra.

20. A method according to any of claims 14 to 19, wherein step (v) comprises performing a mathematical operation to combine at least two of the spectra

21. A method according to claim 20, wherein the difference between at least two of the spectra is calculated

22. A method according to claim 20, wherein at least one of the spectra is divided by another one of said n spectra

23. A method according to any of claims 14 to 22, wherein n=2.

24. A method according to any of claims 14 to 22, where n = 3 and the value of the preliminary imaging parameter differs between said three selected regions, the first time value being chosen to coincide with a region where there is a difference between the three spectra.

25. A method according to any of claims 14 to 24, wherein the second parameter and the first parameter are the same.

26. A method according to any preceding claim, wherein the first parameter is the amplitude of the radiation.

27. A method according to any preceding claim, wherein step (a) comprises:

irradiating the sample through a member which is transparent to the irradiating radiation and which abuts the sample,

the method further comprising the step of:

obtaining a baseline signal by irradiating the member in the absence of the sample and detecting the amplitude of the reflected radiation, and

wherein in step (b):

the first parameter is determined by subtracting the baseline signal from the detected amplitude of the radiation reflected from both the member and the sample.

28. A method according to any of claims 1 to 25, wherein step (a) comprises:

irradiating the sample through a member which is transparent to the irradiating radiation and which abuts the sample,

the method further comprises:

obtaining a baseline signal by irradiating the member in the absence of the sample and detecting the amplitude of the reflected radiation, and

replacing the sample with a reference object of known reflectance and measuring a reference signal by measuring the amplitude of radiation reflected from the reference object,

wherein in step (b) comprises:

determining the first parameter by subtracting the baseline signal from the detected radiation and dividing the result by the reference signal in the frequency domain.

29. A method according to either of claims 27 or 28, wherein the baseline signal is subtracted from the detected reflected radiation in the time domain.

30. A method according to any of claims 1 to 25, further comprising the step of:

replacing the sample with a reference object of known reflectance and measuring a reference signal by measuring the amplitude of radiation reflected from the reference object, and

wherein step (b) comprises:

determining the first parameter by dividing the detected amplitude with the reference signal in the frequency domain.

31. A method according to any of claims 1 to 25, wherein step (b) further comprises:

determining the first parameter by multiplying the amplitude of the detected radiation in the frequency domain by the complex Fourier Transform of function F(t), where F(t) is a non-zero function whose integral between time limits t_a and t_b is zero, and where t_a and t_b are chosen to encompass the part of interest of the reflected or transmitted radiation pulse.

32. A method according to claim 27, wherein step (b) comprises determining the first parameter by subtracting the baseline signal from the detected radiation from the sample and multiplying the baseline subtracted sample signal in the frequency domain by the complex Fourier transform of the function F(t), where F(t) is a non-zero function whose integral between time limits t_a and t_b is zero, and where t_a and t_b are chosen to encompass the part of interest of the reflected or transmitted pulse.

33. A method according to claim 28, wherein step (b) comprises determining the first parameter by subtracting the baseline signal from the amplitude of the detected radiation and dividing the result by the reference signal in the frequency domain and multiplying by the complex Fourier transform of the function F(t) in the frequency domain, where F(t) is a non-zero function whose integral between time limits t_a and t_b is zero, and where t_a and t_b are chosen to encompass the part of interest of the reflected or transmitted pulse.

34. A method according to claim 30, wherein step (b) comprises determining the first parameter by dividing the amplitude of the detected radiation from the sample by the reference signal in the frequency domain and multiplying the result by the complex Fourier transform of F(t) in the frequency domain, where F(t) is a non-zero function whose integral between time limits t_a and t_b is zero, and where t_a and t_b are chosen to encompass the part of interest of the reflected or transmitted pulse.

35. A method according to any of claims 31 to 34, where

$$F(t)=\left|\frac{2}{\pi}\left\{\frac{e^{-2\left(\frac{t}{\alpha}\right)^2}}{\alpha}-\frac{e^{-2\left(\frac{t}{\beta}\right)^2}}{\beta}\right\}\right|$$

and α and β are constants.

36. A method according to clam 31, wherein α is substantially equal to the shortest pulse length of the beam of pulsed radiation and β is set to be longer than the pulse length.

37. A method according to any of claims 30 to 33, where:

$$F(t)=\left|\frac{2}{\pi}\left\{\frac{e^{-2\left(\frac{t}{\alpha}\right)^2}}{\alpha}-\frac{1}{T}\right\}\right|$$

and wherein α is a constant which is substantially equal to the shortest pulse length of the beam and T is substantially equal to the time which it takes the beam of radiation to penetrate to the deepest point of interest in the sample.

38. An apparatus for imaging a sample, the apparatus comprising:

a source for irradiating a sample with a pulse of electro-magnetic radiation, said pulse having a plurality of frequencies in the range from 25GHz to 100Thz;

detector means for detecting the amplitude of radiation reflected from and/or transmitted by the sample;

means for determining a first parameter at least related to the amplitude of the radiation in the time domain;

calculating means for calculating the value of the first parameter at a first time value with respect to the value of the first parameter at a second time value, wherein the second time value coincides with a physical feature of the dataset of the first parameter with respect to time, and

imaging means for generating an image by plotting the value calculated by the calculating means for different parts of the sample.

39. An apparatus for imaging a sample, the apparatus comprising:

a source for irradiating a sample with a pulse of electro-magnetic radiation, said pulse having a plurality of frequencies in the range from 25GHz to 100THz;

detector means for detecting the amplitude of radiation reflected from and/or transmitted by the sample;

means for determining a first parameter at least related to the amplitude of the radiation in the time domain;

calculating means for calculating the value of the first parameter at a first time value with respect to the value of the first parameter at a second time value, and

imaging means for generating an image by plotting the value calculated by the calculating means for different parts of the sample, wherein the time interval between the first and second time intervals is constant for all points used to generate the image.

40. An apparatus according to either of claims 38 or 39, wherein the calculating means is configured to calculate the value of the first parameter for the first time value with respect to the second time value for a plurality of different first time values for each point of the sample used to generate the image.

41. An apparatus according to claim 40, further comprising means to produce a plurality of images for different first time values.

42. An apparatus according to claim 41, wherein the apparatus comprises means to scan through the plurality of images.

43. An apparatus according to any of claims 38 to 42, wherein the first time value is constant for all points used to generate a single image.

44. An apparatus according to any of claims 38 to 43, wherein the time interval between the first time value and the second time value is constant for all points used to generate a single image.

45. An apparatus according to claim 43, further comprising means to optimise contrast in the image by selecting the first time value which provides the best contrast.

46. An apparatus according to claim 44, further comprising means to optimise contrast in an image by selecting the fixed time interval which provides the best contrast.

47. An apparatus for imaging a sample, the apparatus comprising:

(a) means for irradiating a sample with a pulse of electromagnetic radiation, said pulse having a plurality of frequencies in the range from 25 GHz to 100 THz;

(b) means for determining a first parameter, at least related to the amplitude of the radiation, which is either reflected from and/or transmitted by the sample, in the time domain; and

(c) means for generating an image of the sample by using the value of the first parameter at a first time value,

said apparatus being configured to determine said first time value by:

(i) selecting n regions of said sample, where n is an integer of at least 2;

(ii) producing n spectra in the time domain, where the nth spectra is obtained by plotting the second parameter, averaged across the nth area, against delay time;

(iii) determining the first time value by comparing at least two of the n spectra.

48. An apparatus according to any of claims 36 to 44, further comprising :

(i) means for selecting n regions of said sample, where n is an integer of at least 2;

(ii) means for producing n time domain spectra of a second parameter which is at least related to the amplitude of the radiation reflected from and/or transmitted by the sample, where the nth spectra is obtained by averaging the second parameter across the nth area and plotting this averaged value as it varies with delay time; and

(iii) means determining the first time value by comparing at least two of the n spectra.

49. A method as substantially hereinbefore described with reference to any of the accompanying drawings.

50. An apparatus as substantially hereinbefore described with reference to any of the accompanying drawings.

INVESTOR IN PEOPLE

Application No:	GB 0201614.5	**Examiner:**	Eleanor Thurston
Claims searched:	1-13 & 26-46	**Date of search:**	31 January 2003

Patents Act 1977 : Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance	
X, P	1-3, 26 & 38.	GB 2360186 A	(TERAPROBE LTD.) see abstract, pages 18 & 19 and figures 14 & 15.
X	1, 2, 26 & 38.	GB 2347835 A	(TOSHIBA) see abstract, pages 3 & 17 and figure 1.
A		EP 0864857 A1	(LUCENT)
A		EP 0841548 A1	(LUCENT)
A		INSPEC abstract Number 6705308 & C M Cisela et al, "Biomedical Applications of terahertz pulse imaging", Proceedings of the SPIE - International Society for Optical Engineering, 2000, SPIE, pp73-81, Vol 3934.	
A		INSPEC abstract Number 6665442 & D D Arnone et al, "Applications of terahertz technology to medical imaging", Proceedings of the SPIE - International Society for Optical Engineering, 1999, SPIE, pp209-219, Vol 3828.	

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC[V]:

H5R

Worldwide search of patent documents classified in the following areas of the IPC[7] :

A61B; G01N; G01J.

The following online and other databases have been used in the preparation of this search report :

IPR2023-00033
Apple EX1002 Page 355







INVESTOR IN PEOPLE

Application No:	GB 0201614.5	47	**Examiner:**	Eleanor Thurston
Claims searched:	1-13 & 26-46		**Date of search:**	31 January 2003

EPODOC, WPI, PAJ.

IPR2023-00033
Apple EX1002 Page 356

(12) UK Patent Application (19) GB (11) 2 347 835 (13) A

(43) Date of A Publication 13.09.2000

(21) Application No 9913087.4

(22) Date of Filing 04.06.1999

(30) Priority Data
(31) 9904166 (32) 23.02.1999 (33) GB

(71) Applicant(s)
Toshiba Research Europe Limited
(Incorporated in the United Kingdom)
Cambridge Research Laboratory,
260 Cambridge Science Park, Milton Road,
CAMBRIDGE, CB4 4WE, United Kingdom

(72) Inventor(s)
Donald Dominic Arnone
Craig Michael Ciesla

(74) Agent and/or Address for Service
Marks & Clerk
57-60 Lincoln's Inn Fields, LONDON, WC2A 3LS,
United Kingdom

(51) INT CL[7]
G01N 21/49

(52) UK CL (Edition R)
H5R RAD RAM RAT RTQ
G1A AAMA ACDX AG17 AR6 AT20 AT21

(56) Documents Cited
EP 0841548 A2 EP 0828162 A2 US 5710430 A
US 5623145 A

(58) Field of Search
UK CL (Edition Q) G1A AAMA ACDL ACDX , H5R RAT RTQ
INT CL[6] A61B 5/00 , G01N 21/47 21/49
Online: WPI, EPODOC, JAPIO

(54) Abstract Title
Terahertz radiation imaging

(57) In an apparatus and method for imaging a sample. The sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz. An area of the sample is subdivided into a two dimensional array of pixels, and radiation from each pixel is detected over a plurality of frequencies and an image is generated from the radiation detected using a frequency or a selection of frequencies from the plurality of frequencies in the pulsed electro-magnetic radiation.

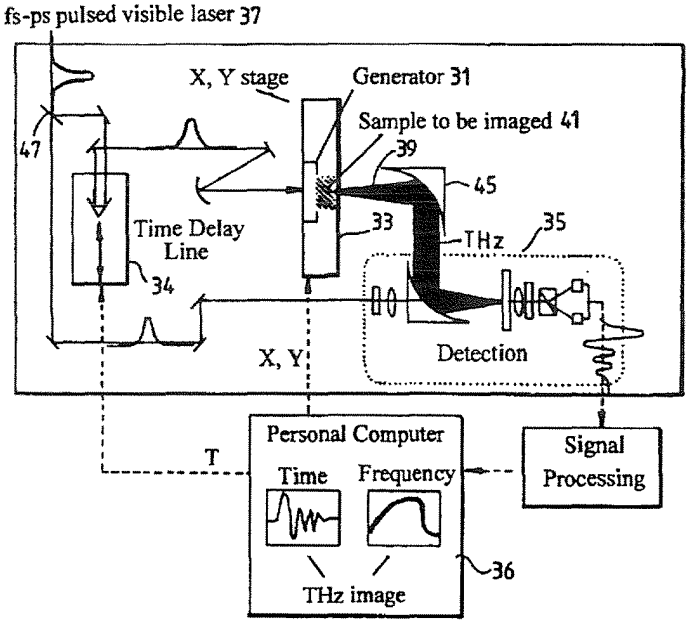


Fig. 17

GB 2 347 835

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

IPR2023-00033
Apple EX1002 Page 357

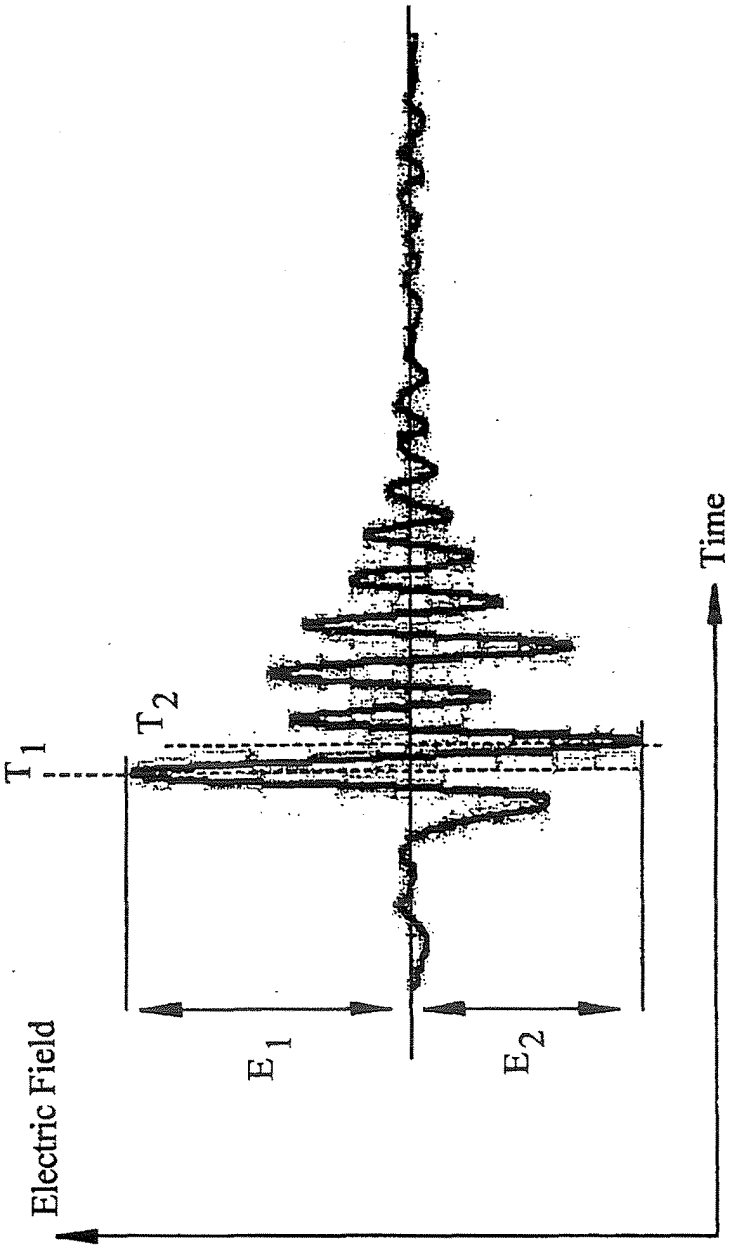


Fig. 1

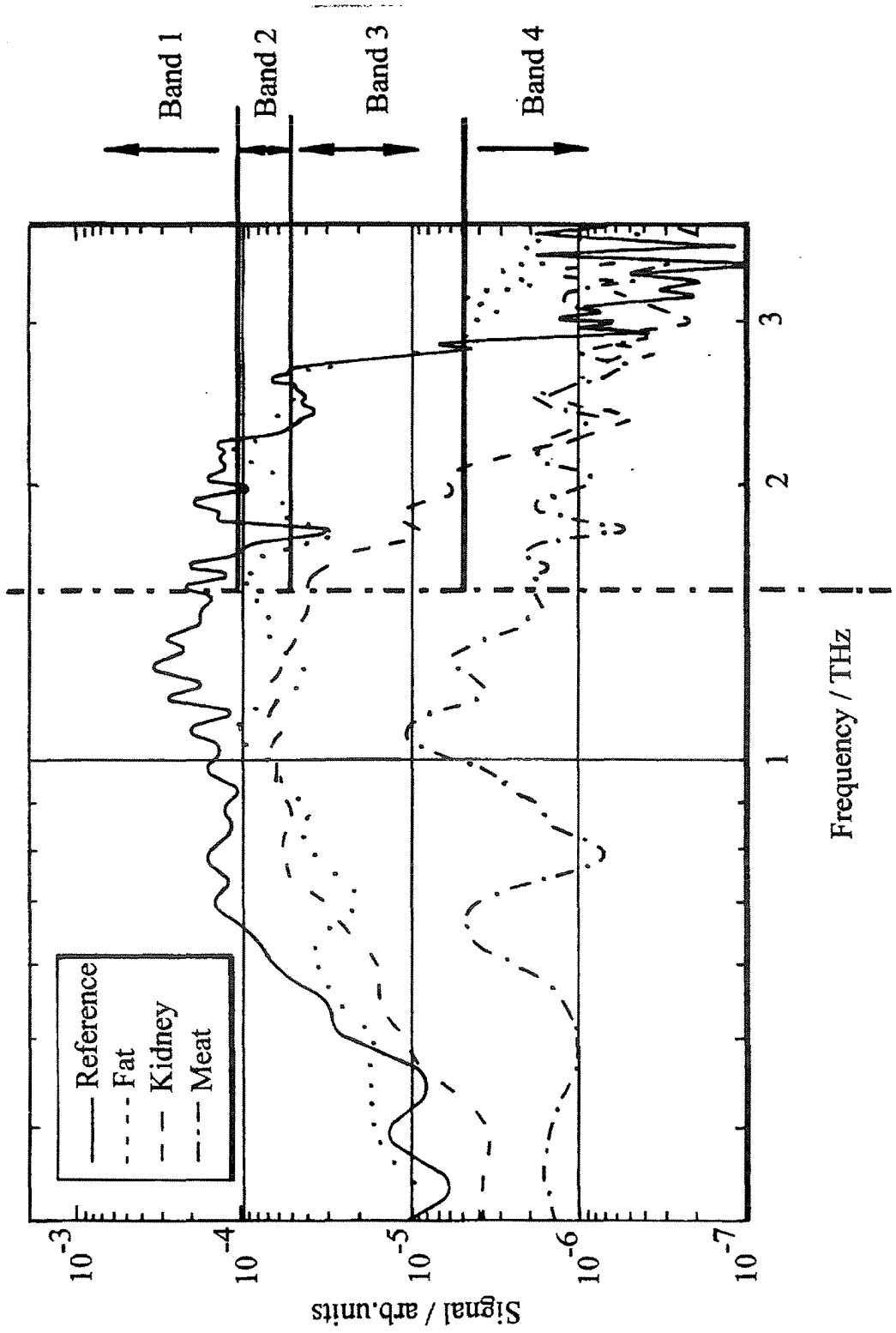

Band 1
Band 2
Band 3
Band 4
Reference
Fat
Kidney
Meat
Signal / arb.units
Frequency / THz

Fig. 2

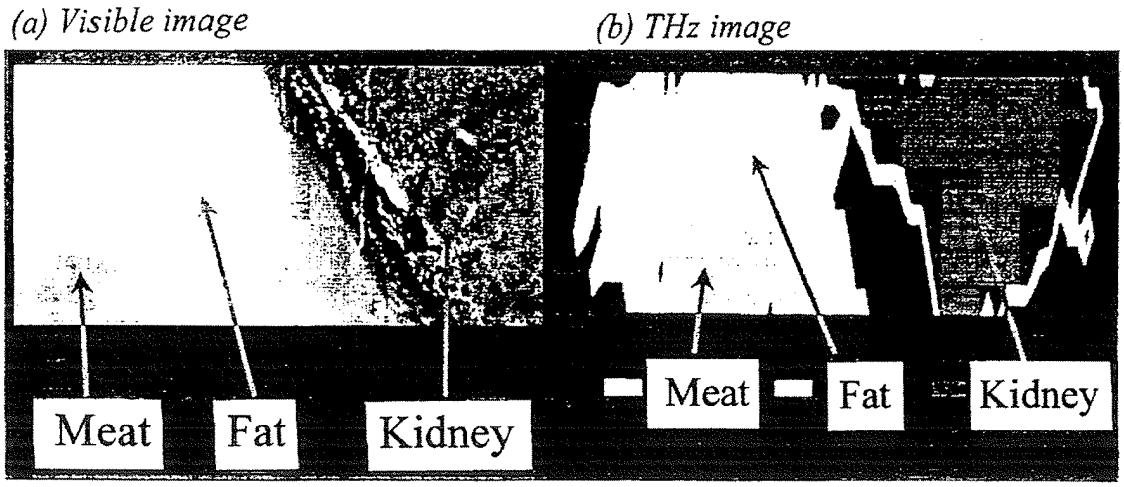


Fig. 3

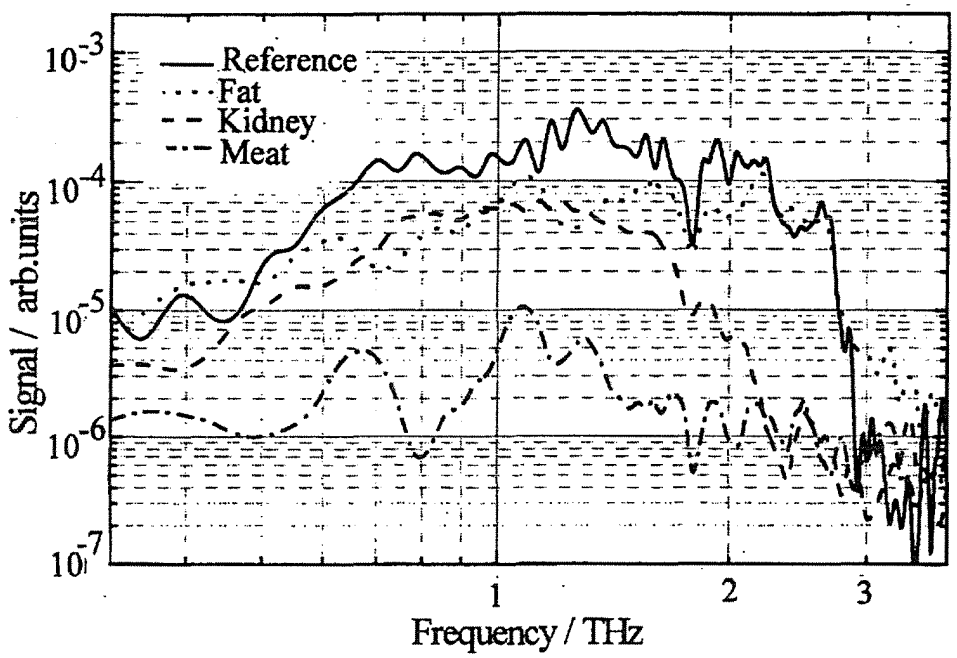

Reference
Fat
Kidney
Meat
Signal / arb.units
Frequency / THz

Fig. 4

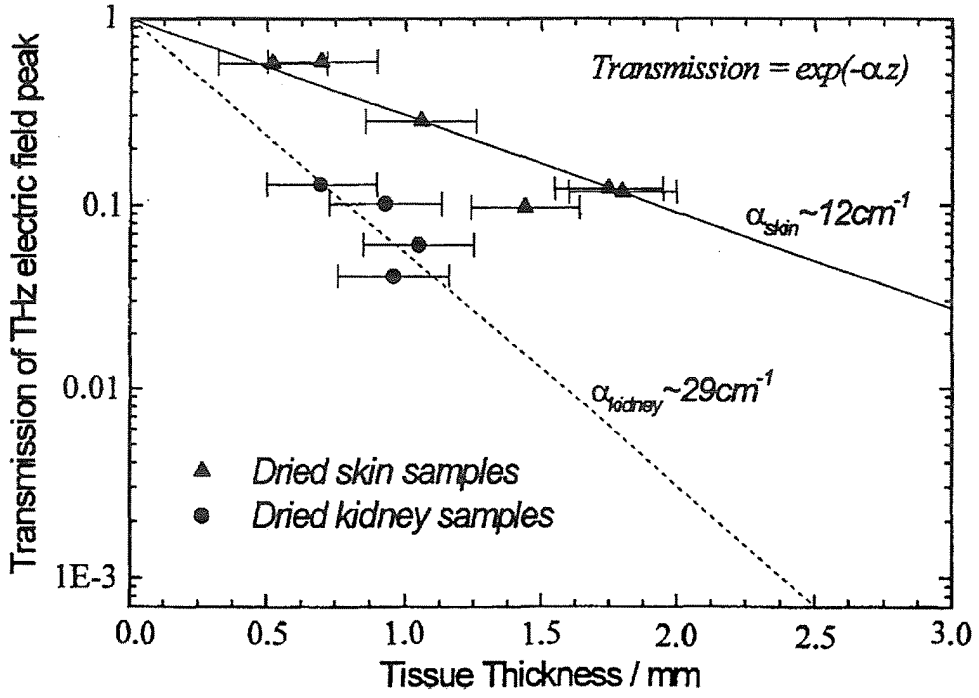

Transmission = exp(-αz)
α_skin ~12cm^-1
α_kidney ~29cm^-1
Dried skin samples
Dried kidney samples
Transmission of THz electric field peak
Tissue Thickness / mm

Fig. 5

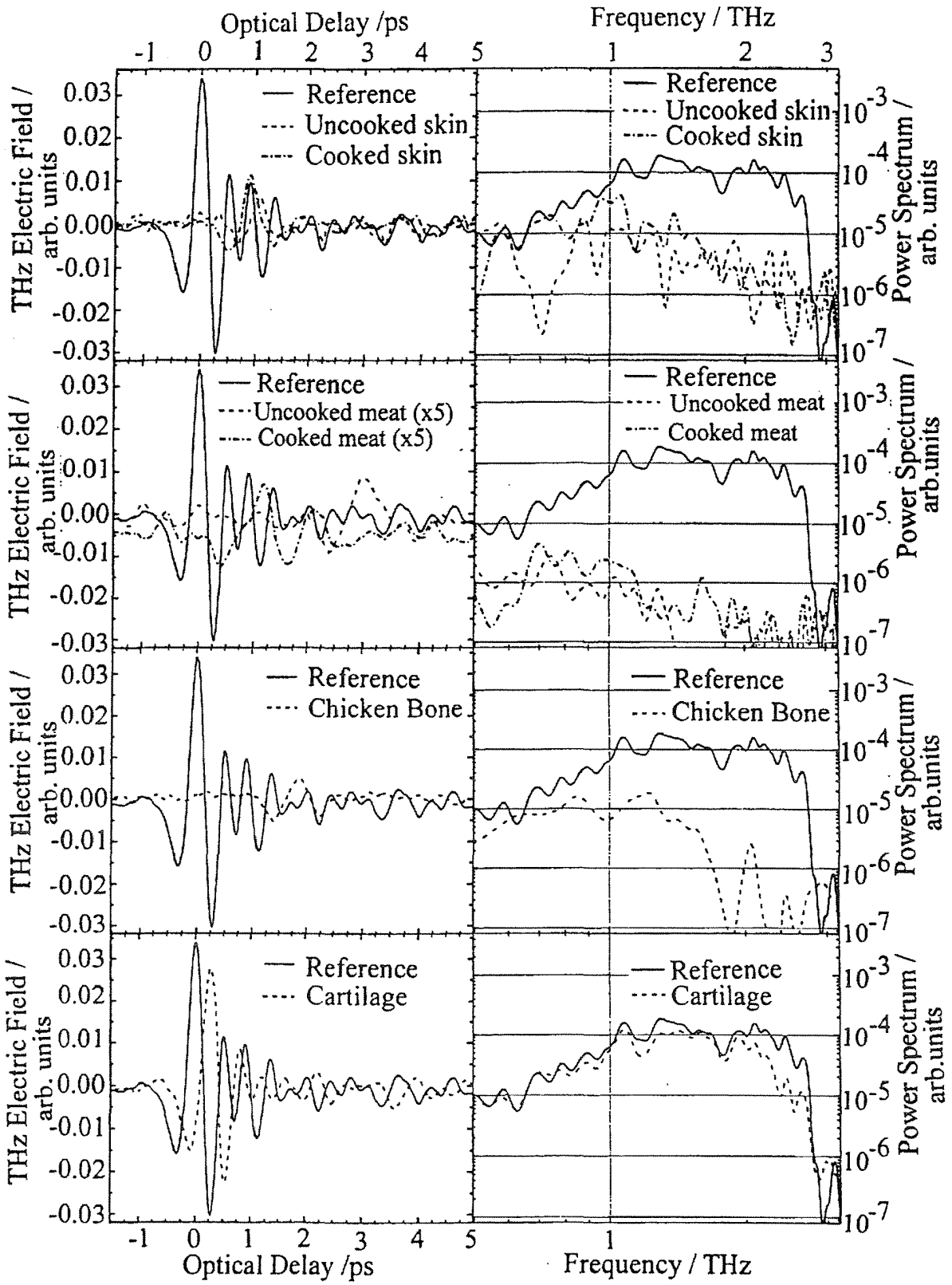

Optical Delay /ps
Frequency / THz
THz Electric Field / arb. units
Power Spectrum / arb. units
Reference
Uncooked skin
Cooked skin
Uncooked meat (x5)
Cooked meat (x5)
Uncooked meat
Cooked meat
Chicken Bone
Cartilage

Fig. 6

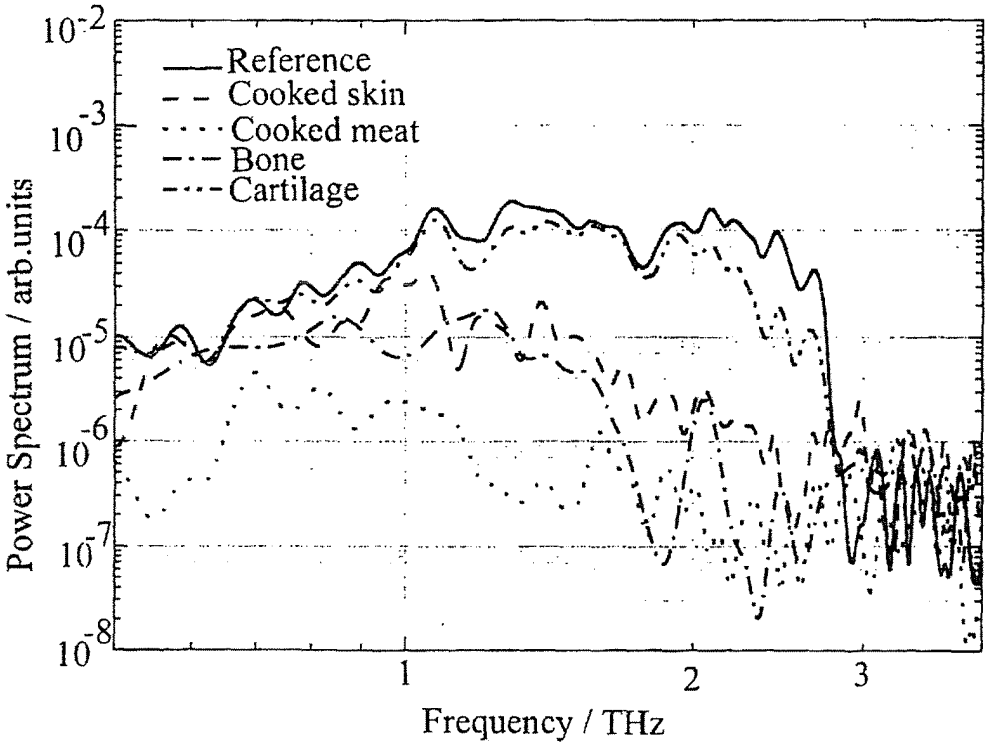

Reference
Cooked skin
Cooked meat
Bone
Cartilage
Power Spectrum / arb.units
Frequency / THz
10-2
10-3
10-4
10-5
10-6
10-7
10-8
1
2
3

Fig. 7

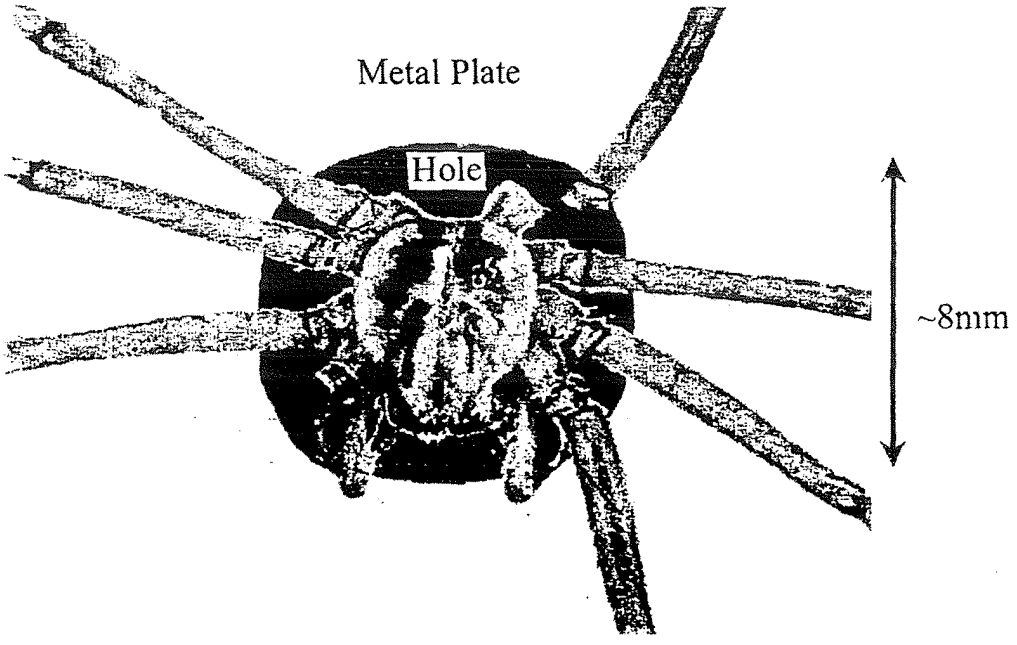

Metal Plate
Hole
~8mm

Fig. 8

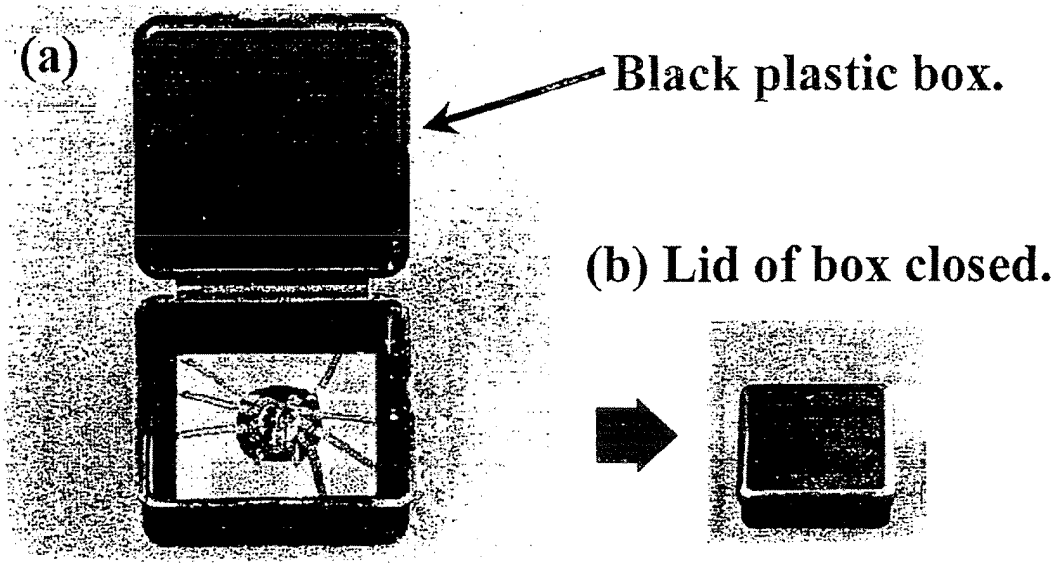


Fig. 9

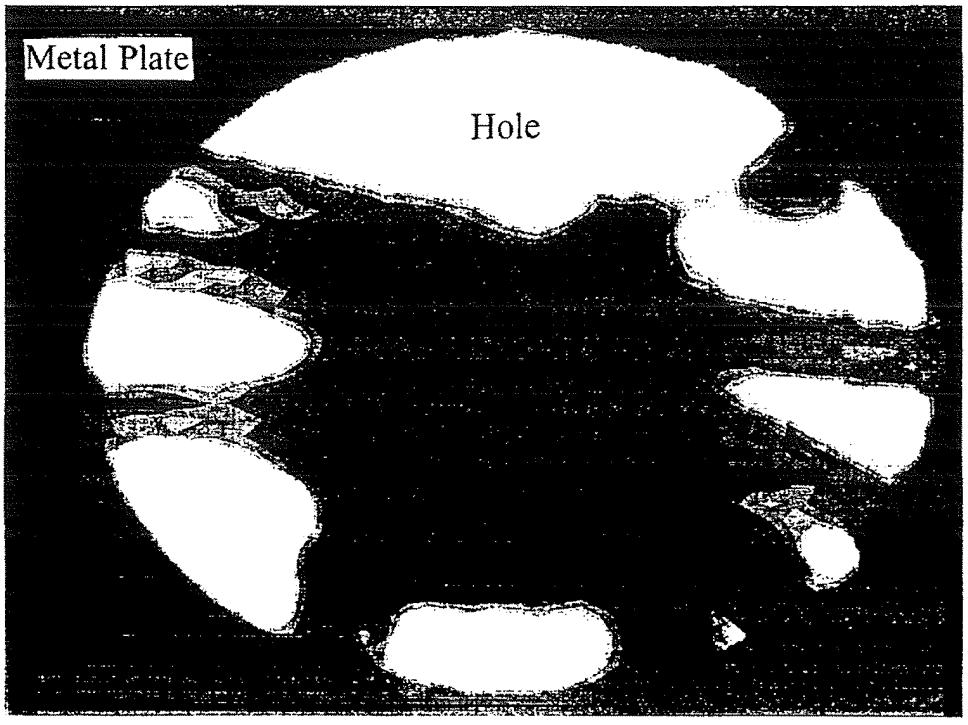


Fig. 10

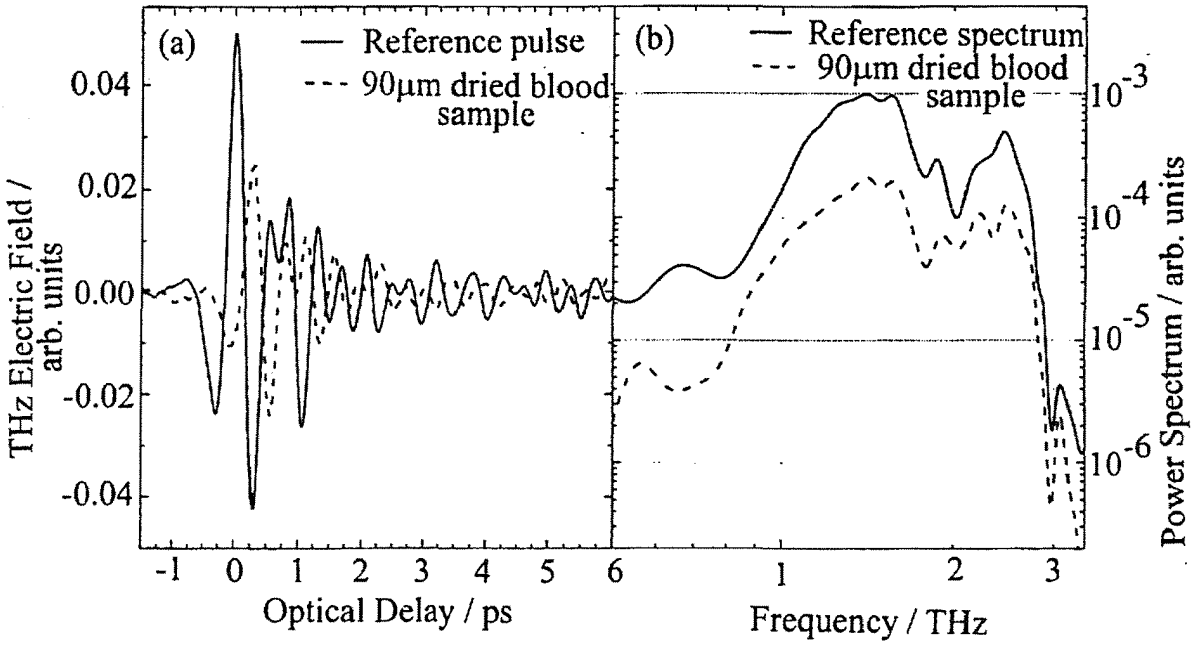

(a)
Reference pulse
90μm dried blood sample
(b)
Reference spectrum
90μm dried blood sample
THz Electric Field / arb. units
Optical Delay / ps
Frequency / THz
Power Spectrum / arb. units

Fig. 11

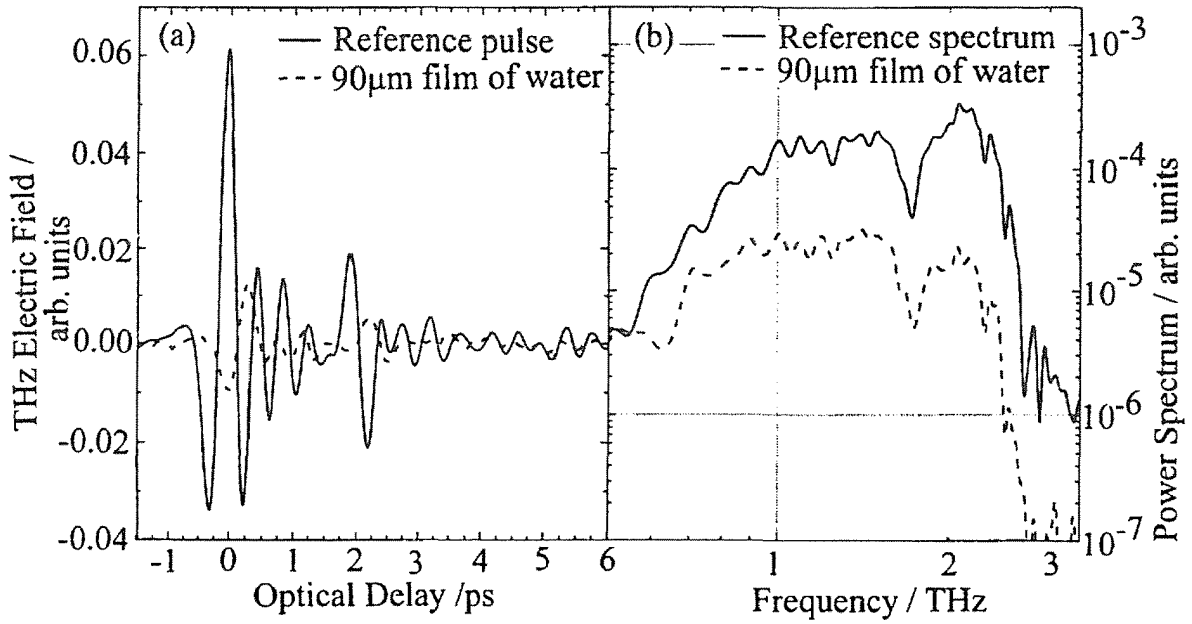

(a)
Reference pulse
90μm film of water
(b)
Reference spectrum
90μm film of water
THz Electric Field / arb. units
Optical Delay /ps
Frequency / THz
Power Spectrum / arb. units

Fig. 12

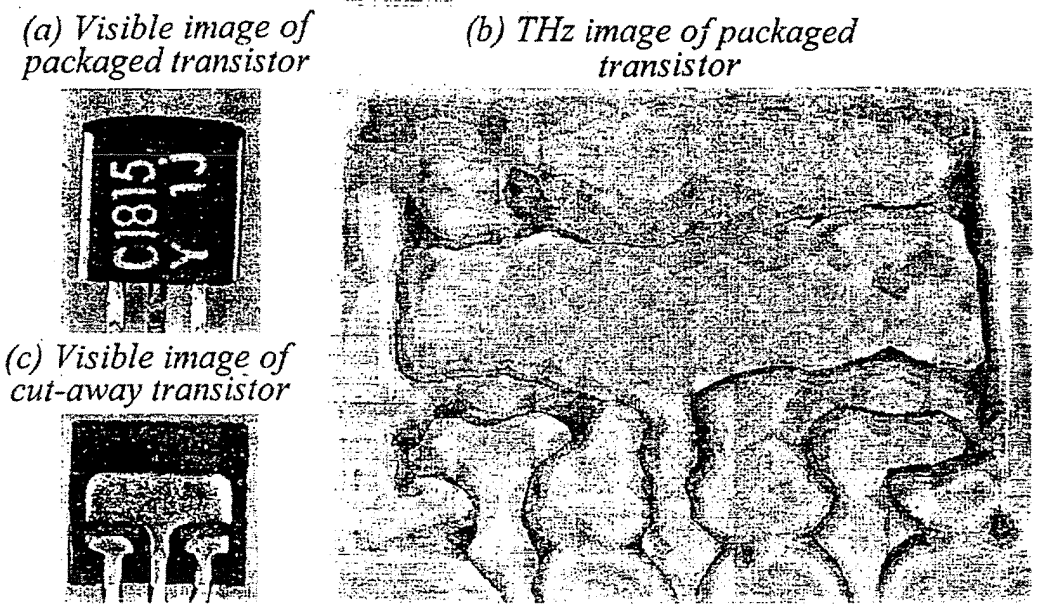

(a) Visible image of packaged transistor
(b) THz image of packaged transistor
(c) Visible image of cut-away transistor
C1815
Y
1J

Fig. 13

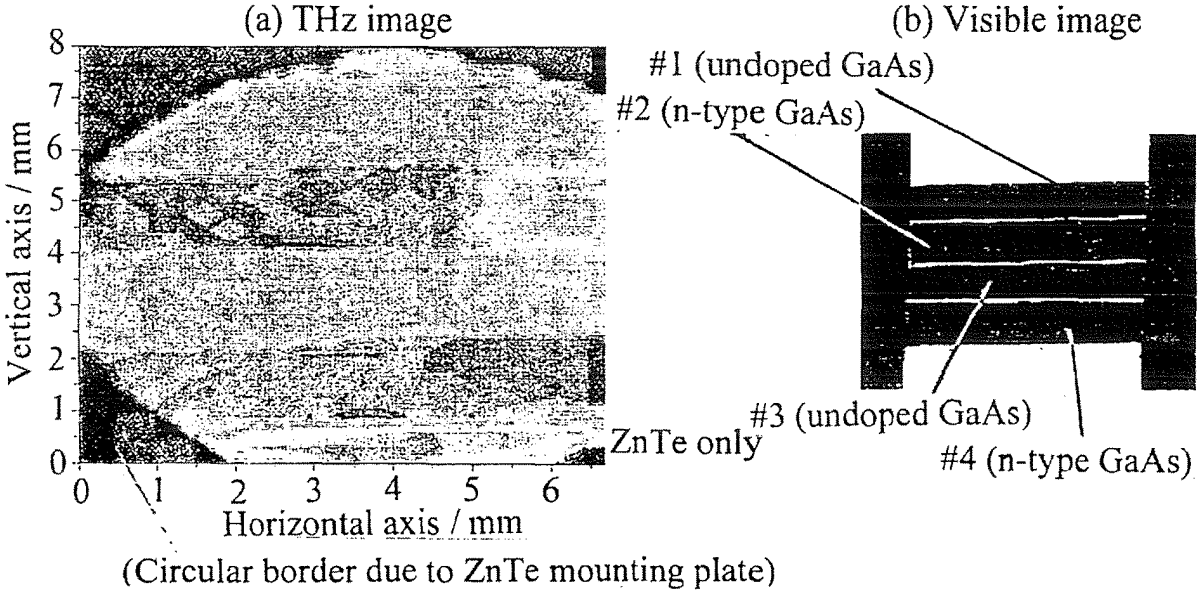

(a) THz image
(b) Visible image
#1 (undoped GaAs)
#2 (n-type GaAs)
#3 (undoped GaAs)
#4 (n-type GaAs)
ZnTe only
Vertical axis / mm
Horizontal axis / mm
0
1
2
3
4
5
6
7
8
(Circular border due to ZnTe mounting plate)

Fig. 14

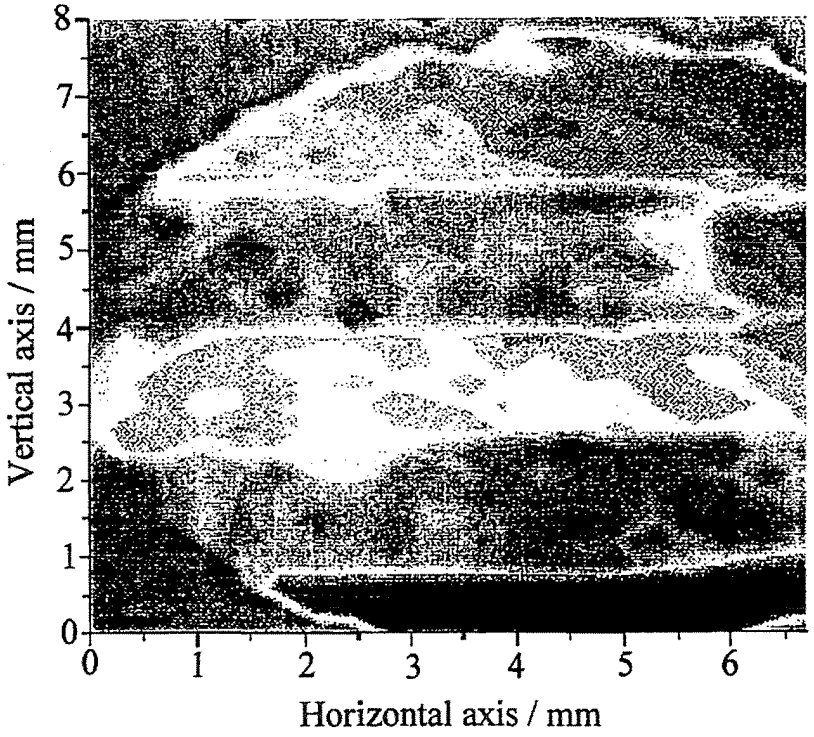

Vertical axis / mm
Horizontal axis / mm
0
1
2
3
4
5
6
7
8

Fig. 15

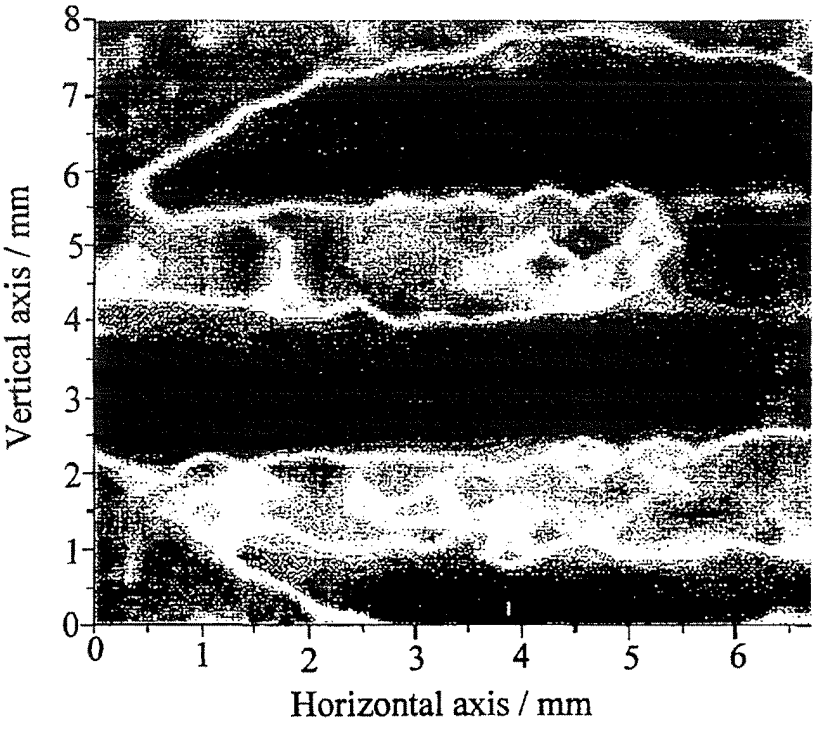

Vertical axis / mm
Horizontal axis / mm
0
1
2
3
4
5
6
7
8

Fig. 16

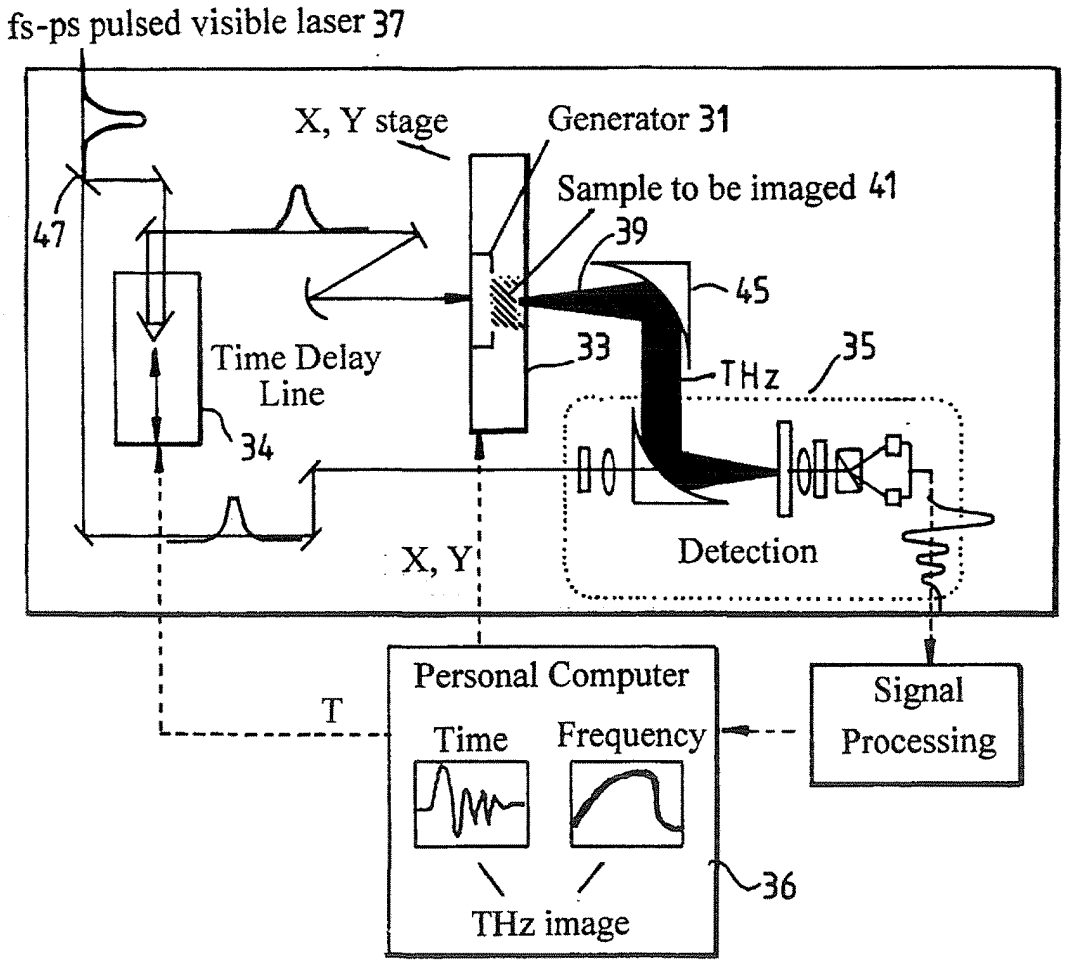

fs-ps pulsed visible laser 37
X, Y stage
Generator 31
Sample to be imaged 41
39
47
45
Time Delay Line
33
THz
35
34
X, Y
Detection
Personal Computer
T
Time
Frequency
Signal Processing
36
THz image

Fig. 17

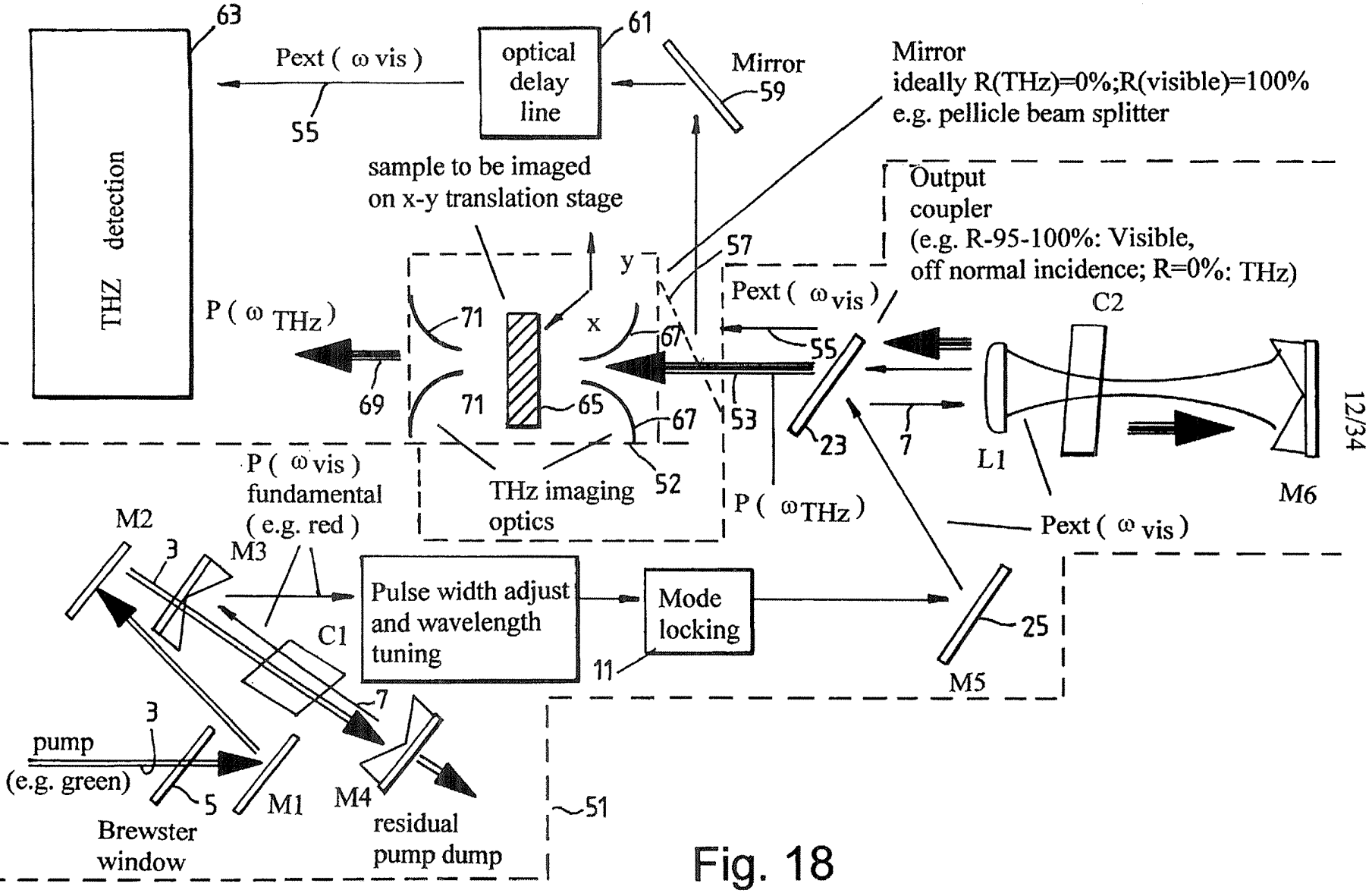


Fig. 18

Apple EX1002 Page 369 IPR2023-00033

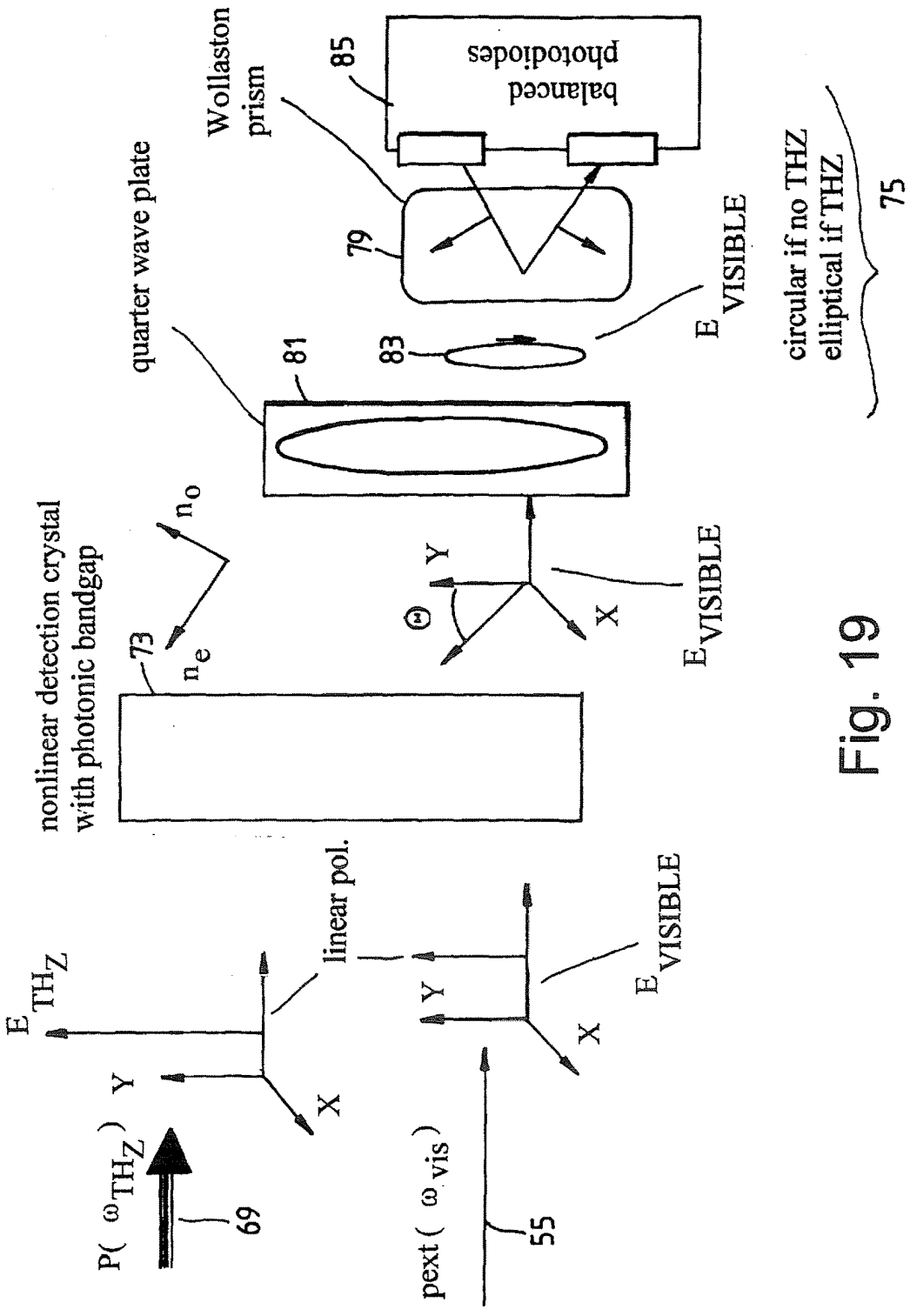

P(ω THZ) Y
69
E THZ
X
linear pol.
pext (ω vis)
55
Y
X
E VISIBLE
nonlinear detection crystal
with photonic bandgap
73
n e
n o
quarter wave plate
81
Θ
Y
X
E VISIBLE
83
E VISIBLE
Wollaston
prism
79
85
balanced
photodiodes
circular if no THZ
elliptical if THZ
75

Fig. 19

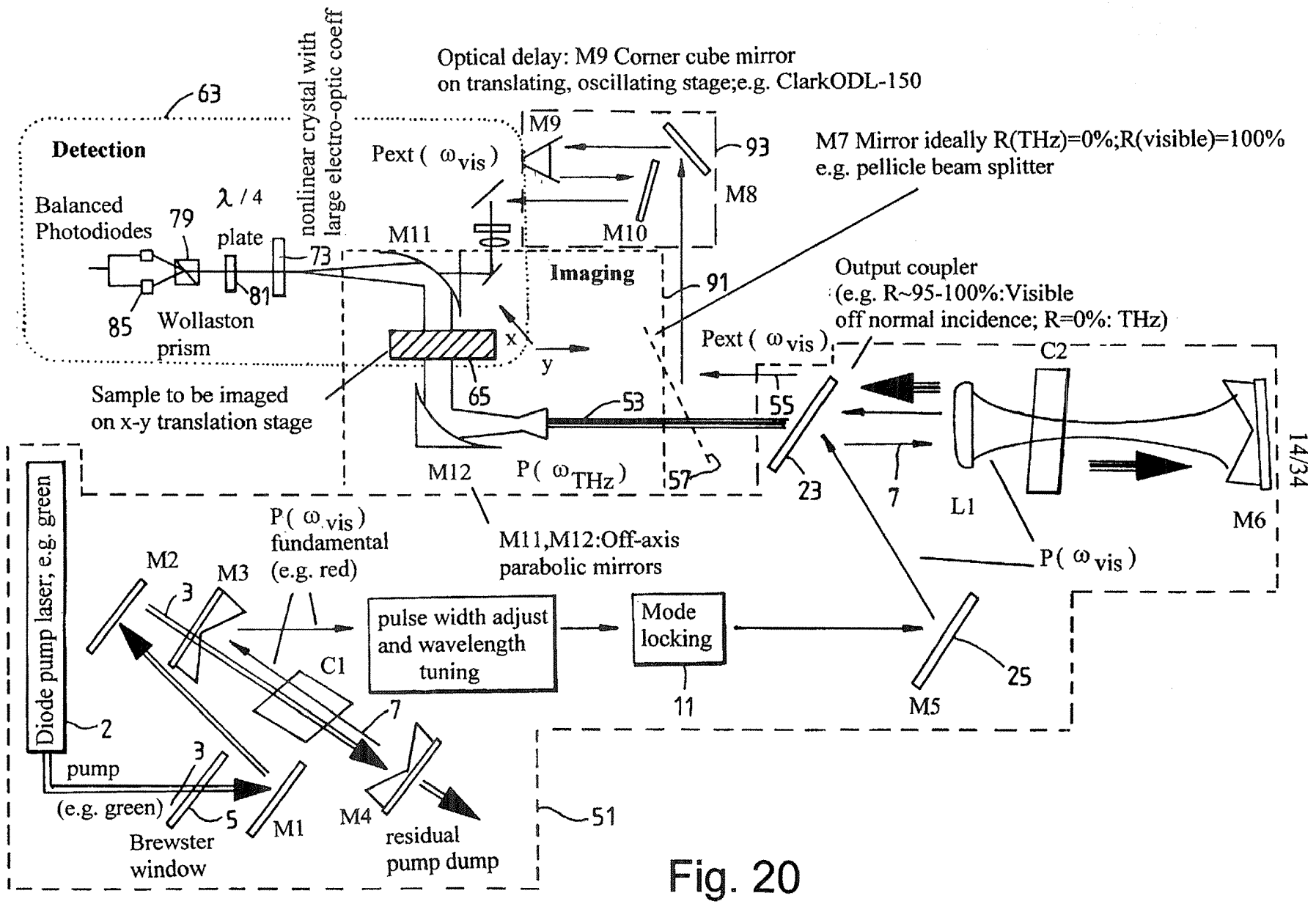


Fig. 20

IPR2023-00033
Apple EX1002 Page 371

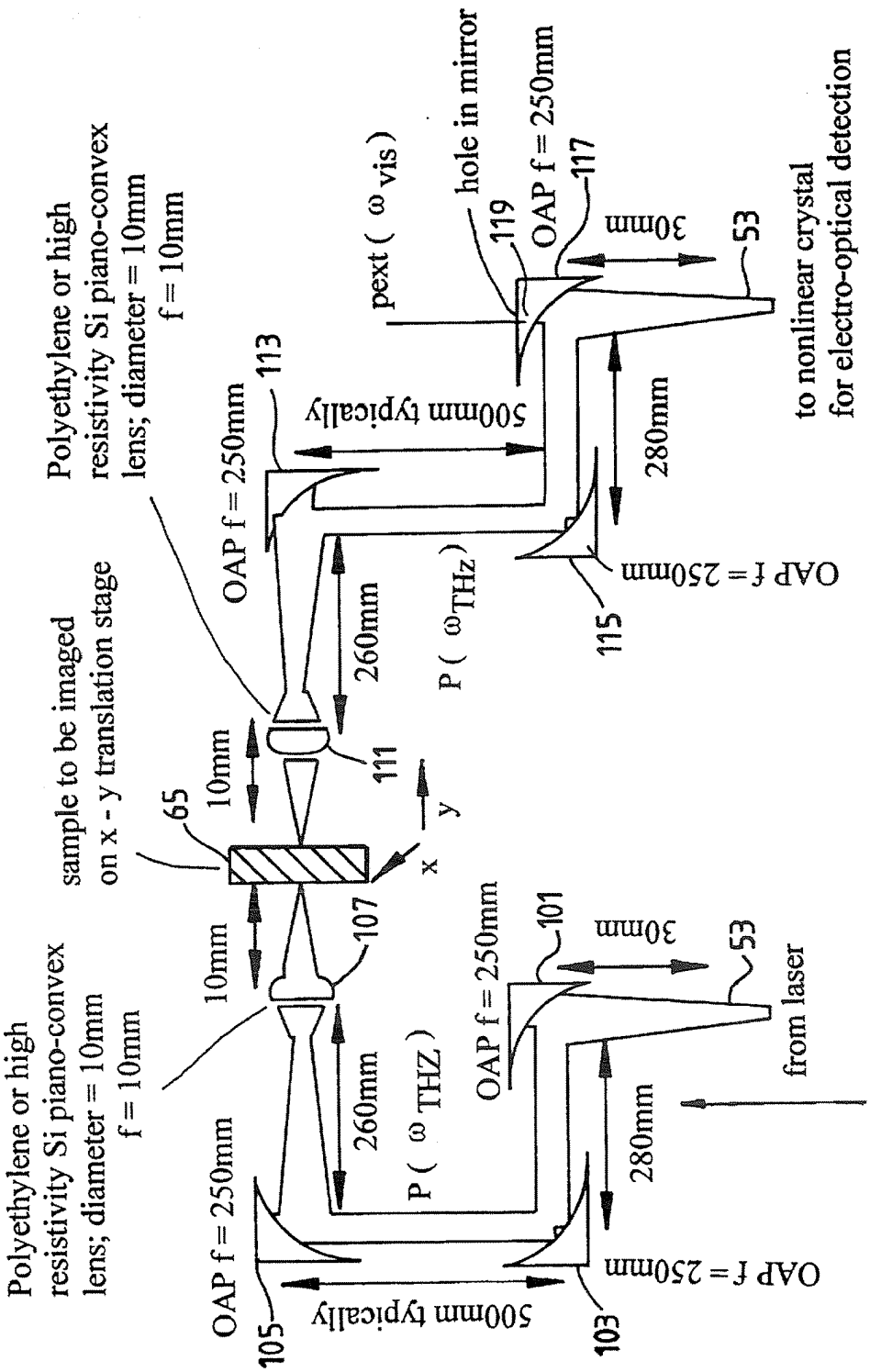

Polyethylene or high resistivity Si piano-convex lens; diameter = 10mm f = 10mm
sample to be imaged on x - y translation stage
Polyethylene or high resistivity Si piano-convex lens; diameter = 10mm f = 10mm
65
10mm
10mm
107
111
x
y
OAP f = 250mm
105
500mm typically
260mm
P (ω THZ)
OAP f = 250mm
101
30mm
53
from laser
280mm
OAP f = 250mm
103
OAP f = 250mm
113
500mm typically
260mm
P (ωTHz)
pext (ω vis)
hole in mirror
119
OAP f = 250mm
117
30mm
53
280mm
OAP f = 250mm
115
to nonlinear crystal for electro-optical detection

Fig. 21

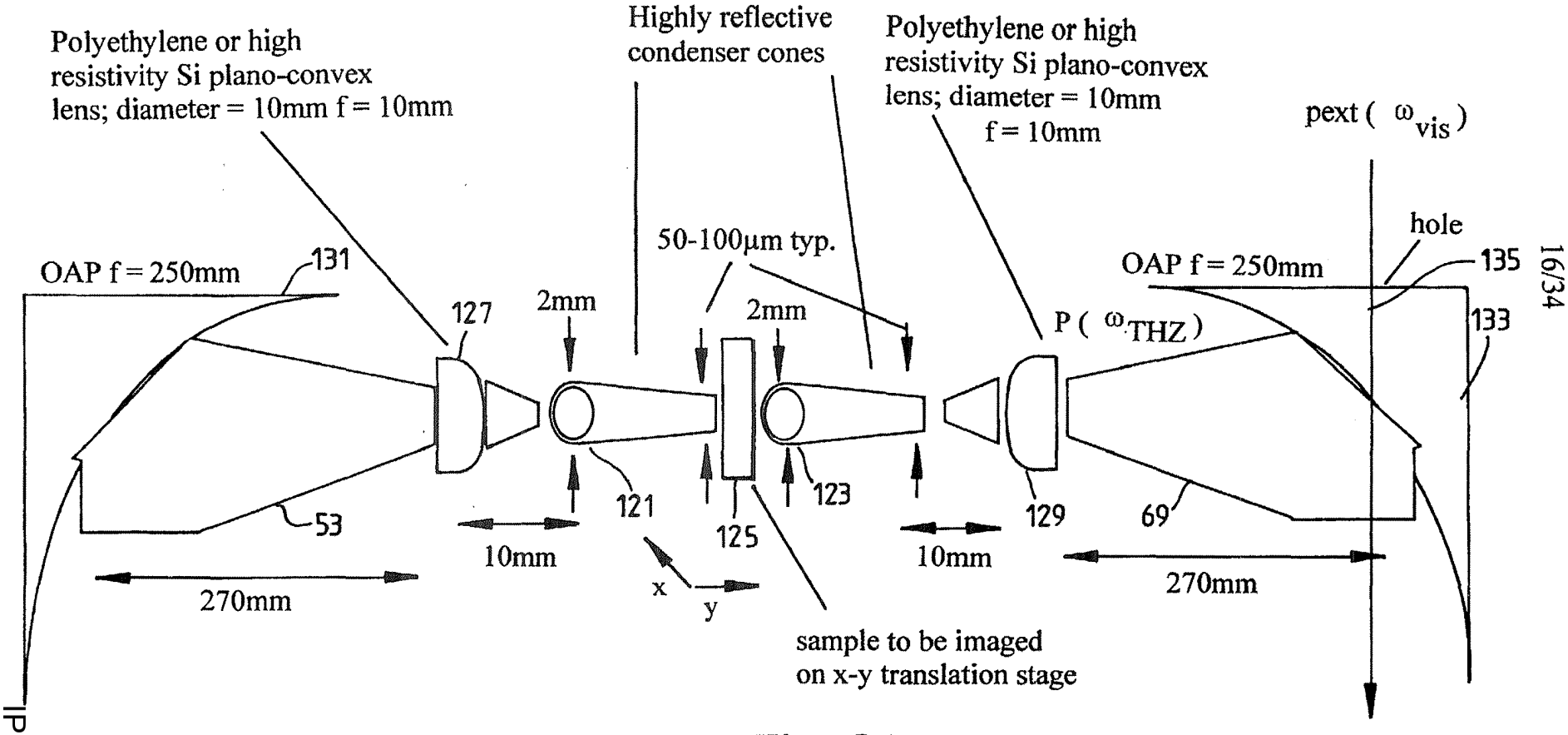
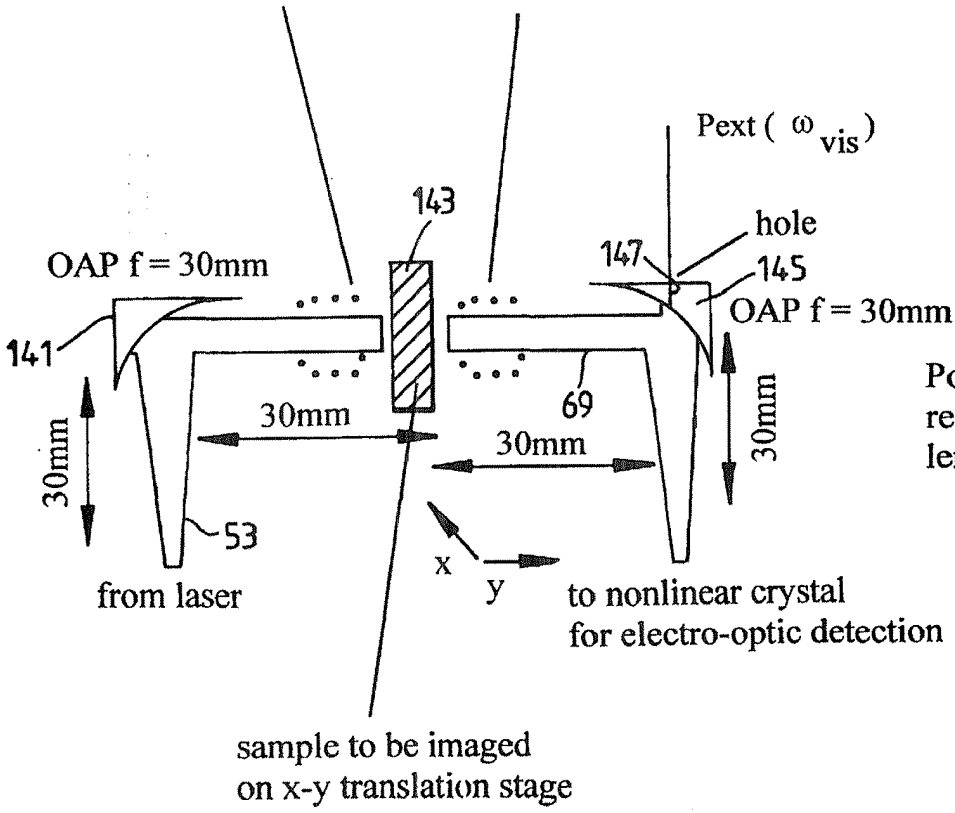


Fig. 22

IPR2023-00033
Apple EX1002 Page 373

Optional: Lens and condenser cones similar to those in Figure 10
for enhanced spatial resolution in x-y plane of object.



Polyethylene or high
restivity Si plano-convex
lens; diameter = 10mm, f = 10mm

Fig. 23

IPR2023-00033
Apple EX1002 Page 374

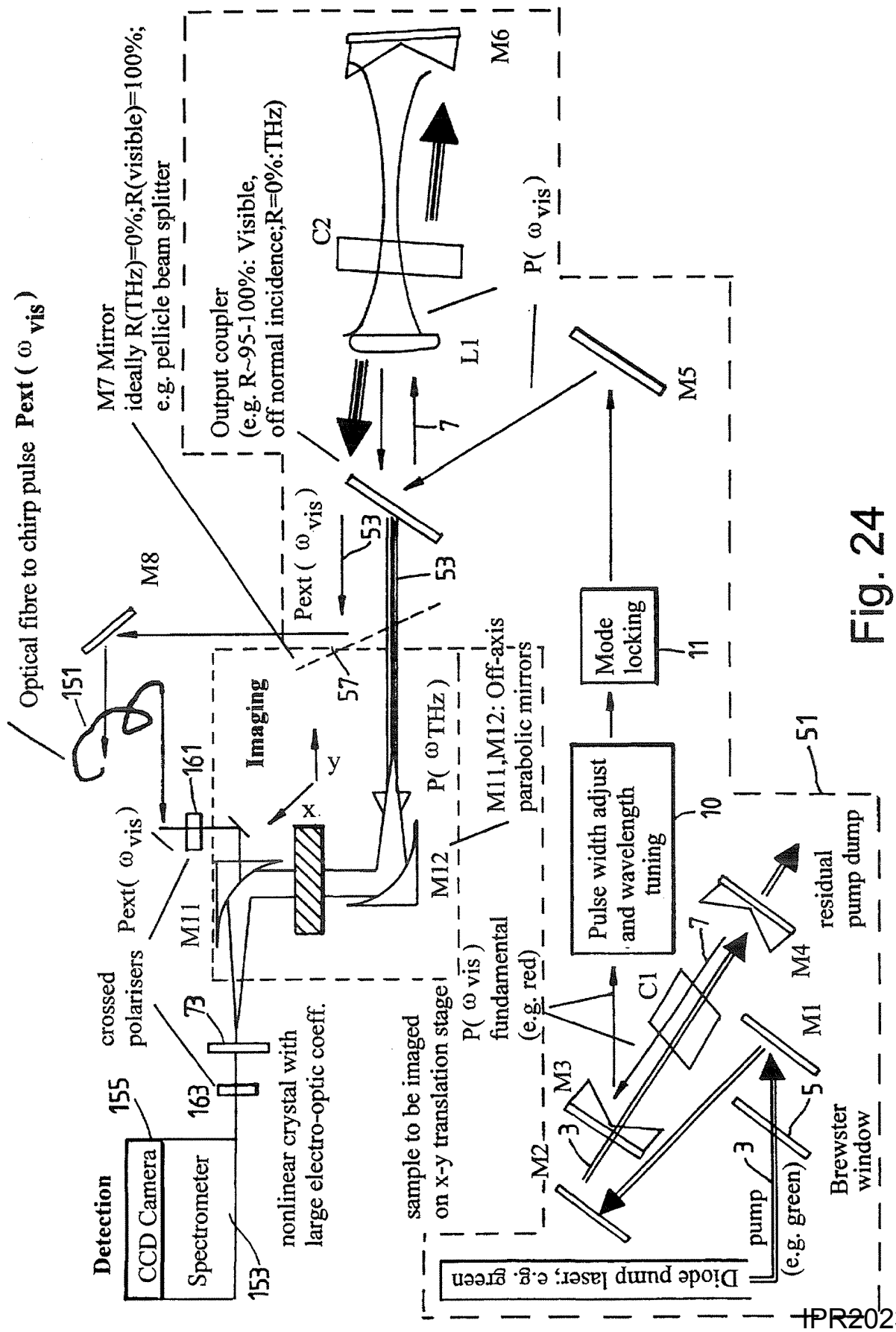

Optical fibre to chirp pulse Pext (ω vis)
M7 Mirror
ideally R(THz)=0%;R(visible)=100%;
e.g. pellicle beam splitter
Output coupler
(e.g. R~95-100%: Visible,
off normal incidence;R=0%:THz)
C2
M6
P(ω vis)
L1
M5
M8
Pext (ω vis)
53
53
Imaging
57
M11,M12: Off-axis
parabolic mirrors
P(ω THz)
Mode
locking
11
Pulse width adjust
and wavelength
tuning
10
51
Pext(ω vis)
M11
M12
crossed
polarisers
73
163
155
Detection
CCD Camera
Spectrometer
153
nonlinear crystal with
large electro-optic coeff.
sample to be imaged
on x-y translation stage
P(ω vis)
fundamental
(e.g. red)
C1
M2
M3
M4
M1
residual
pump dump
Brewster
window
pump
(e.g. green)
Diode pump laser; e.g. green

Fig. 24

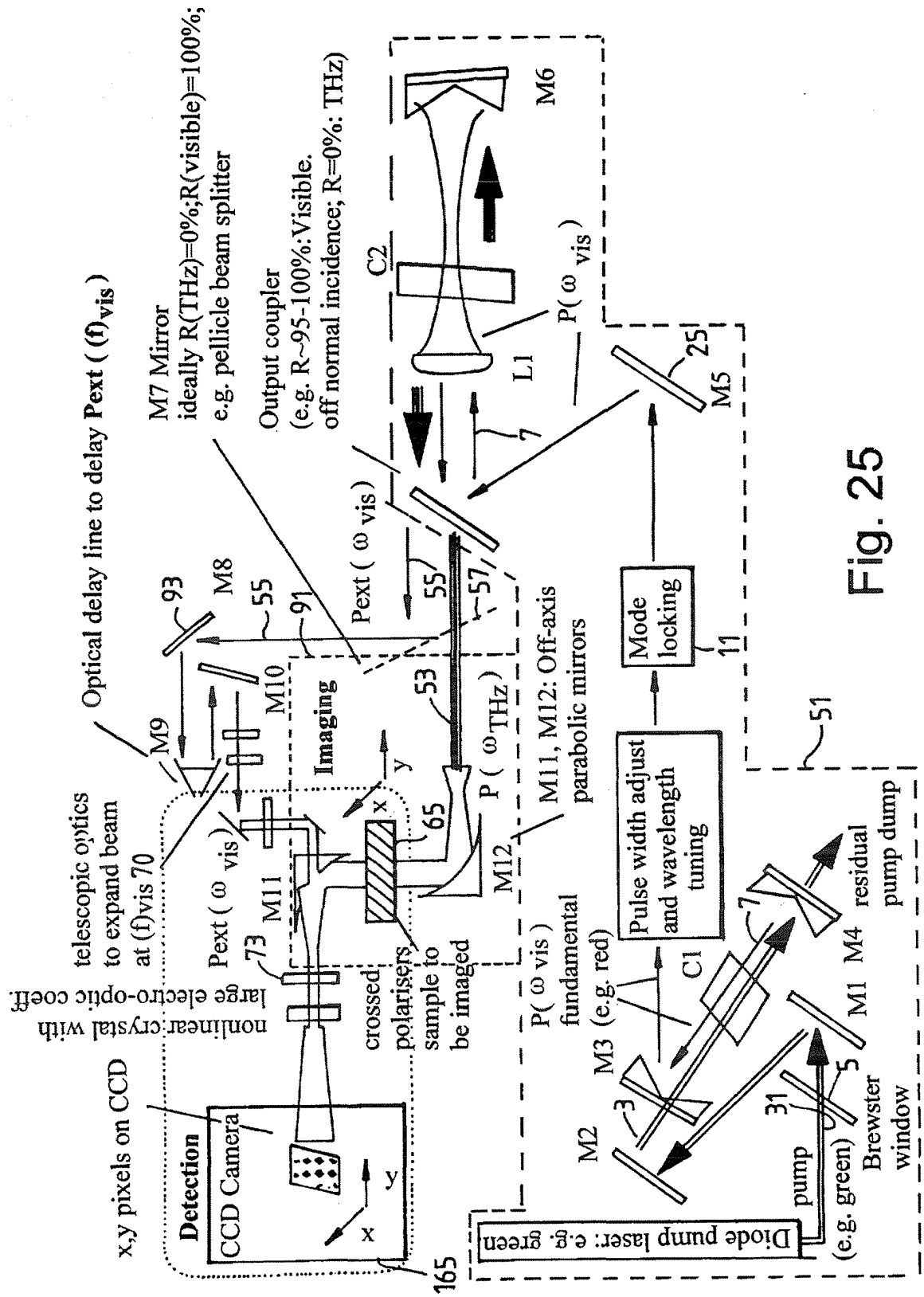

Optical delay line to delay Pext ((f)vis)
M7 Mirror
ideally R(THz)=0%;R(visible)=100%;
e.g. pellicle beam splitter
Output coupler
(e.g. R~95-100%:Visible.
off normal incidence; R=0%: THz)
C2
M6
L1
P(ω vis)
25
M5
7
Pext (ω vis)
55
57
93
M8
55
91
M10
Imaging
53
Mode locking
11
M9
telescopic optics
to expand beam
at (f)vis 70
P (ωTHz)
M11, M12: Off-axis
parabolic mirrors
51
Pext (ω vis)
73
M11
65
M12
nonlinear crystal with
large electro-optic coeff.
crossed
polarisers
sample to
be imaged
P(ω vis)
fundamental
(e.g. red)
Pulse width adjust
and wavelength
tuning
C1
7
M4
residual
pump dump
x,y pixels on CCD
Detection
CCD Camera
M3
M1
M2
3
31
5
pump
(e.g. green)
Brewster
window
Diode pump laser: e.g. green
165

Fig. 25

: All-Optical THz Imaging System with photodiode detection

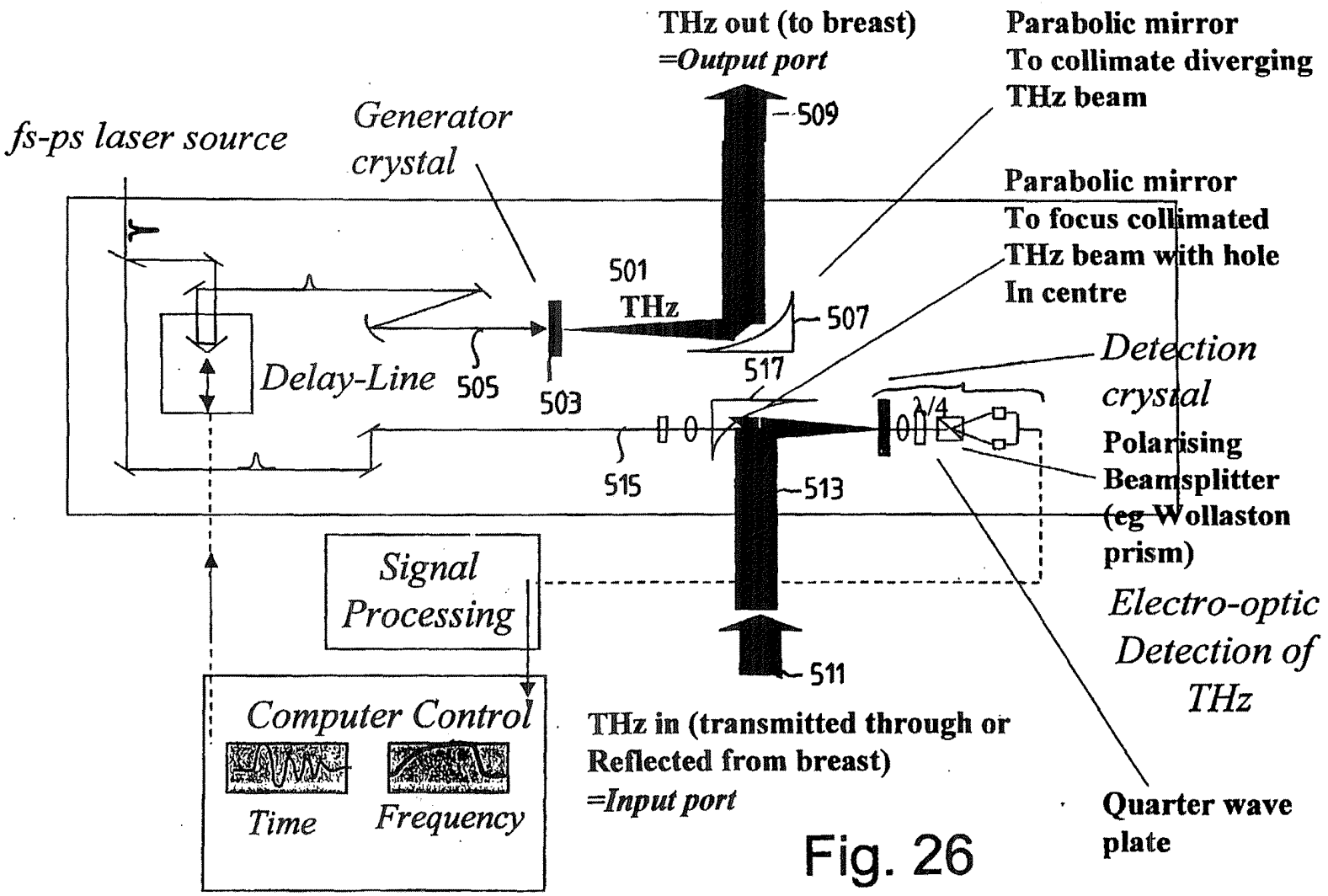


Fig. 26

IPR2023-00033
Apple EX1002 Page 377

: All-Optical THz Imaging System with CCD detection

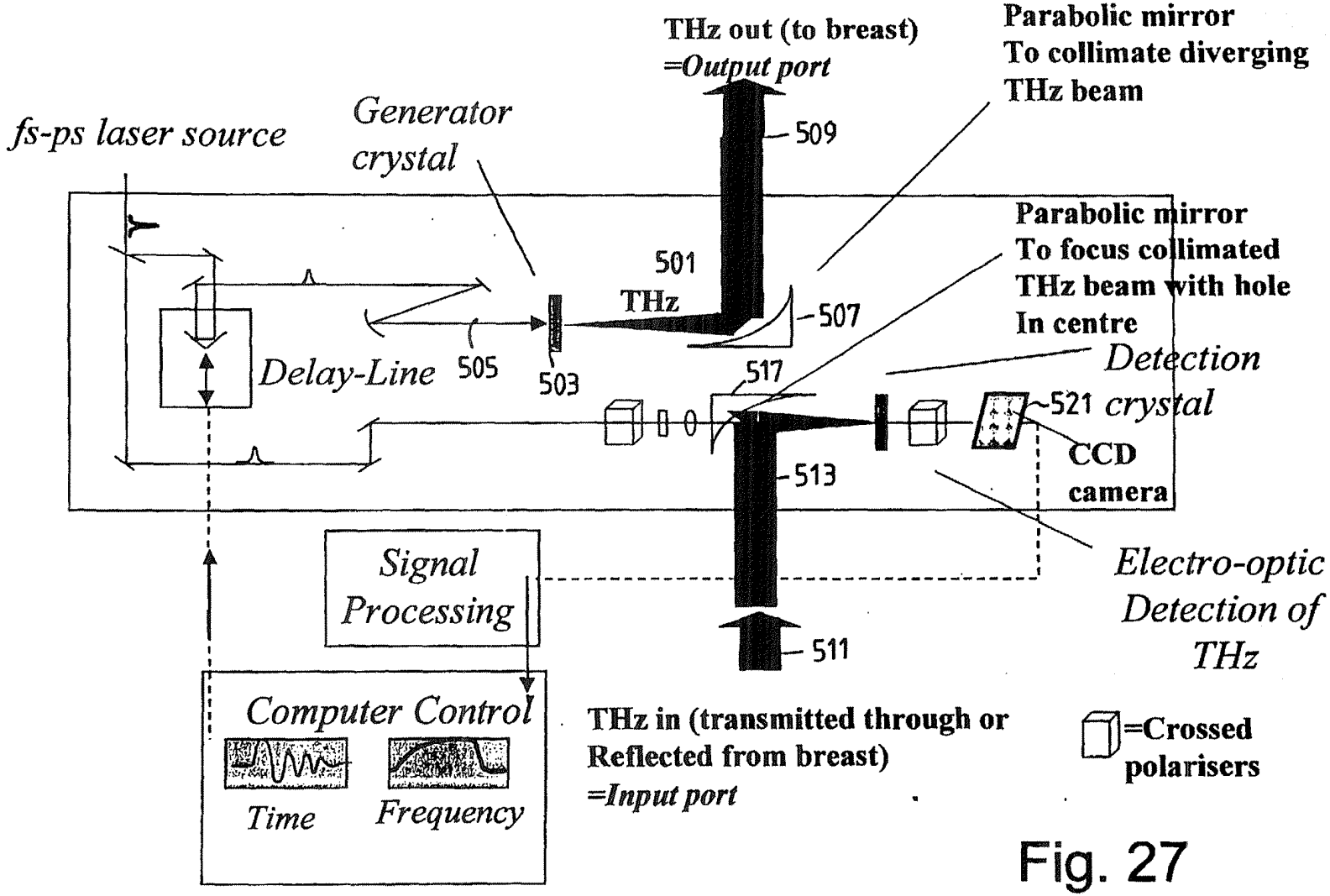


Fig. 27

IPR2023-00033
Apple EX1002 Page 378

: Sample port for THz breast screening
(for use with all-optical systems in Figures 2 and 3)

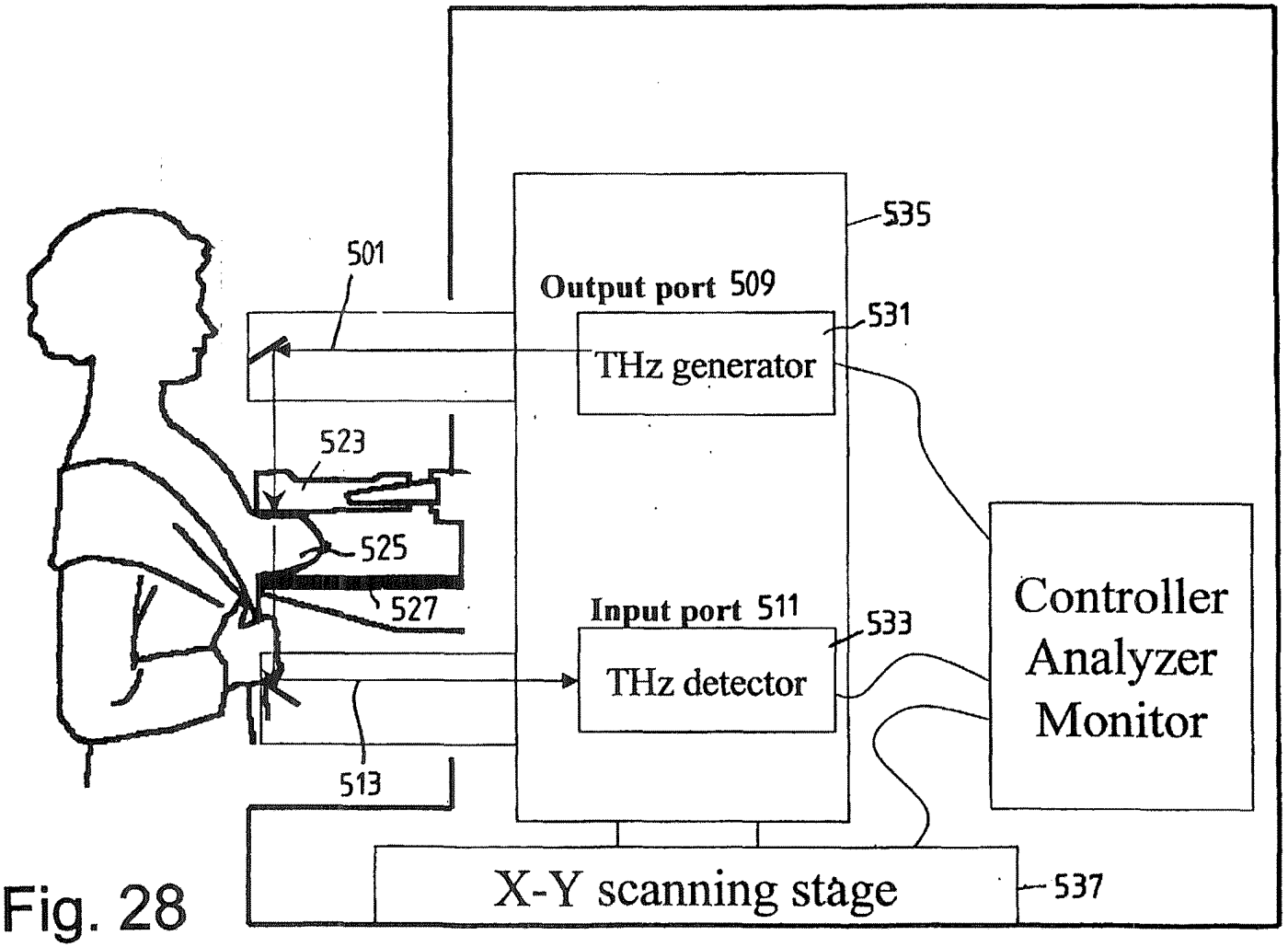


Fig. 28

IPR2023-00033
Apple EX1002 Page 379

: Sample port for THz breast screening
(for use with all-optical systems in Figures 2 and 3)

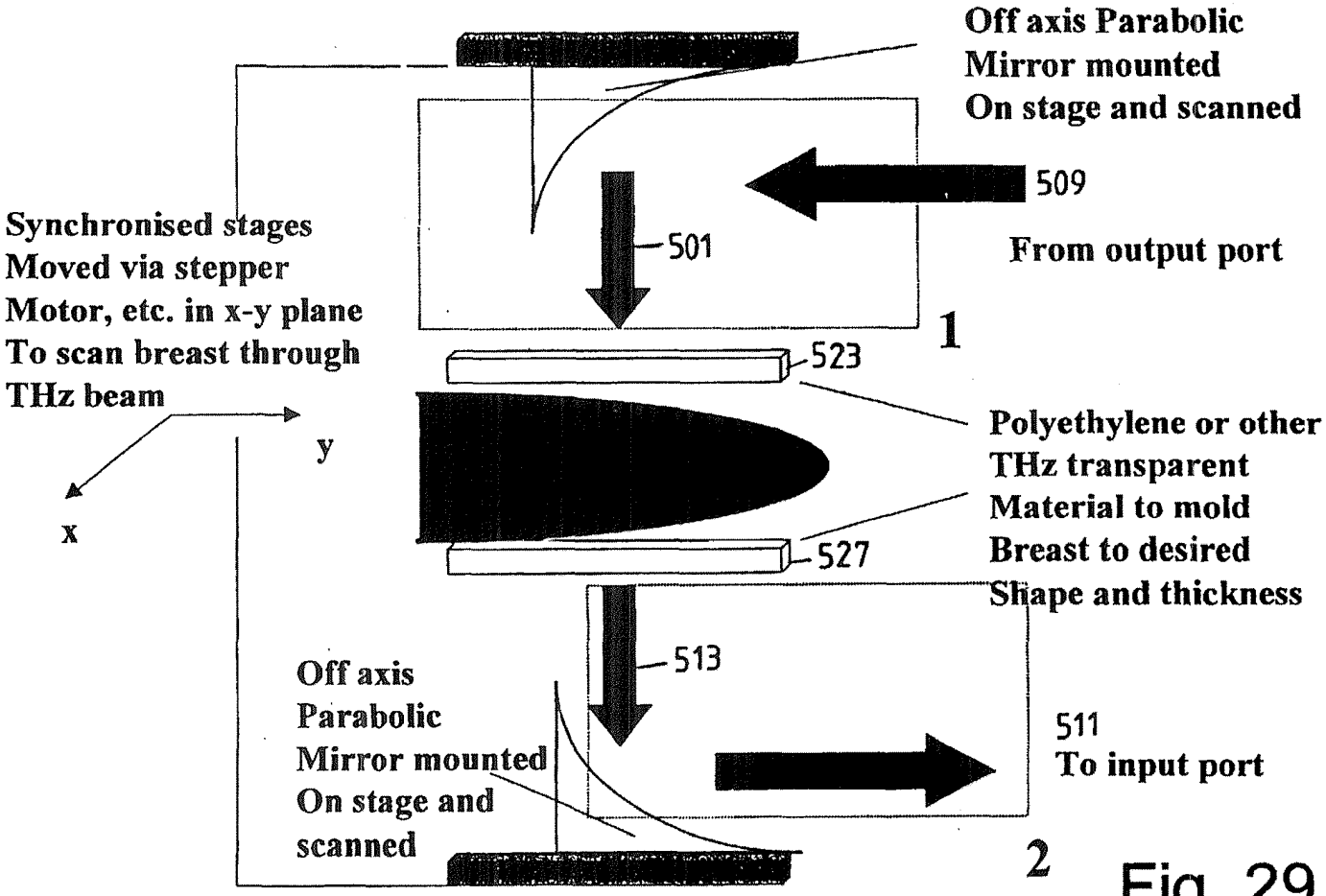
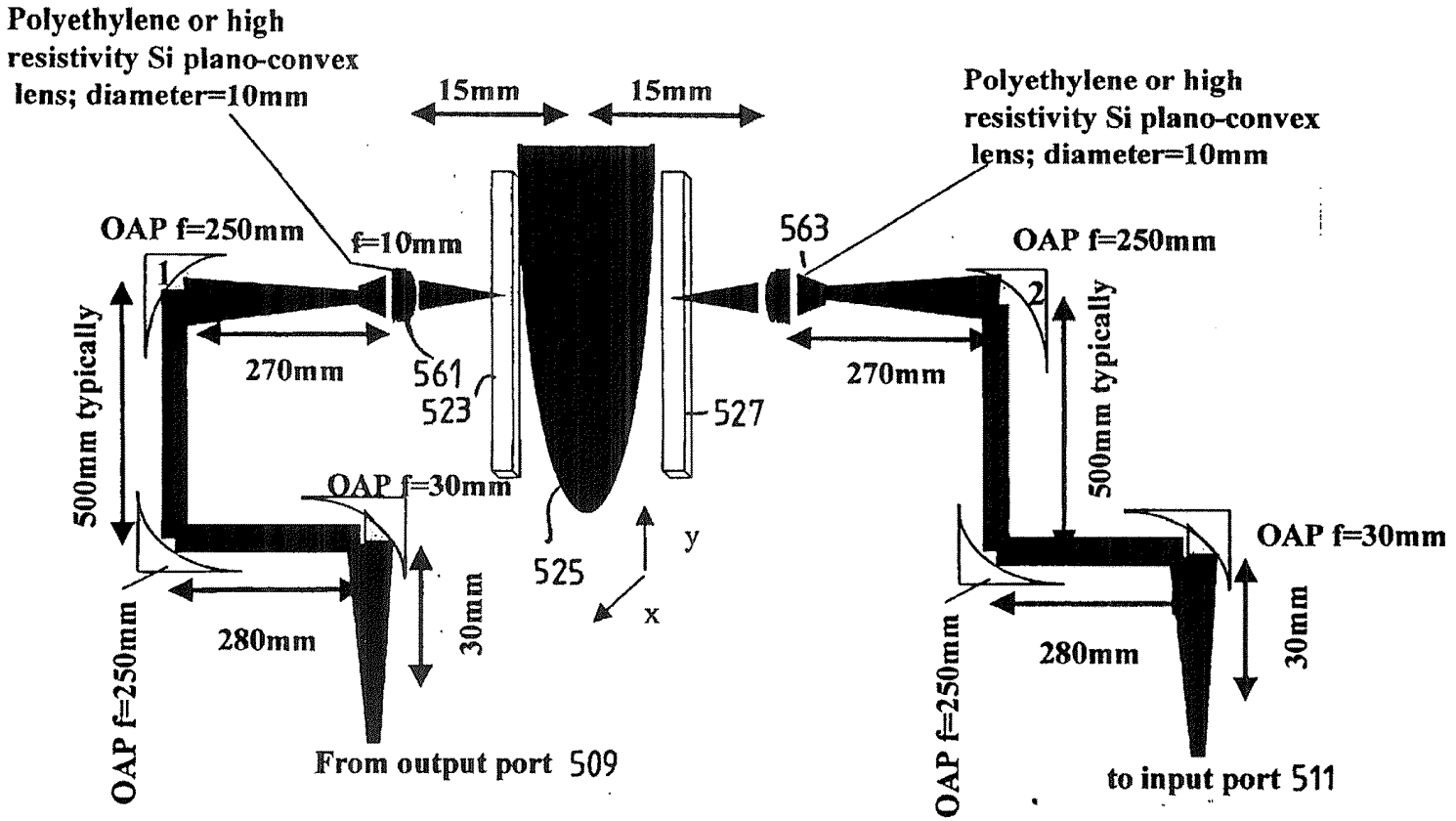


Fig. 29

IPR2023-00033
Apple EX1002 Page 380

: Possible alternative configuration of coupling optics used in Figure 5 (regions denoted 1. And 2. In dotted boxes).



All mirrors are off-axis parabolics (OAPs).
OAP's designated by 1. And 2. Are on synchronised X-Y stepper motor stages

Fig. 30

IPR2023-00033
Apple EX1002 Page 381

: Possible alternative configuration of coupling optics used in Figure 5 (regions denoted 1. And 2. In dotted boxes).

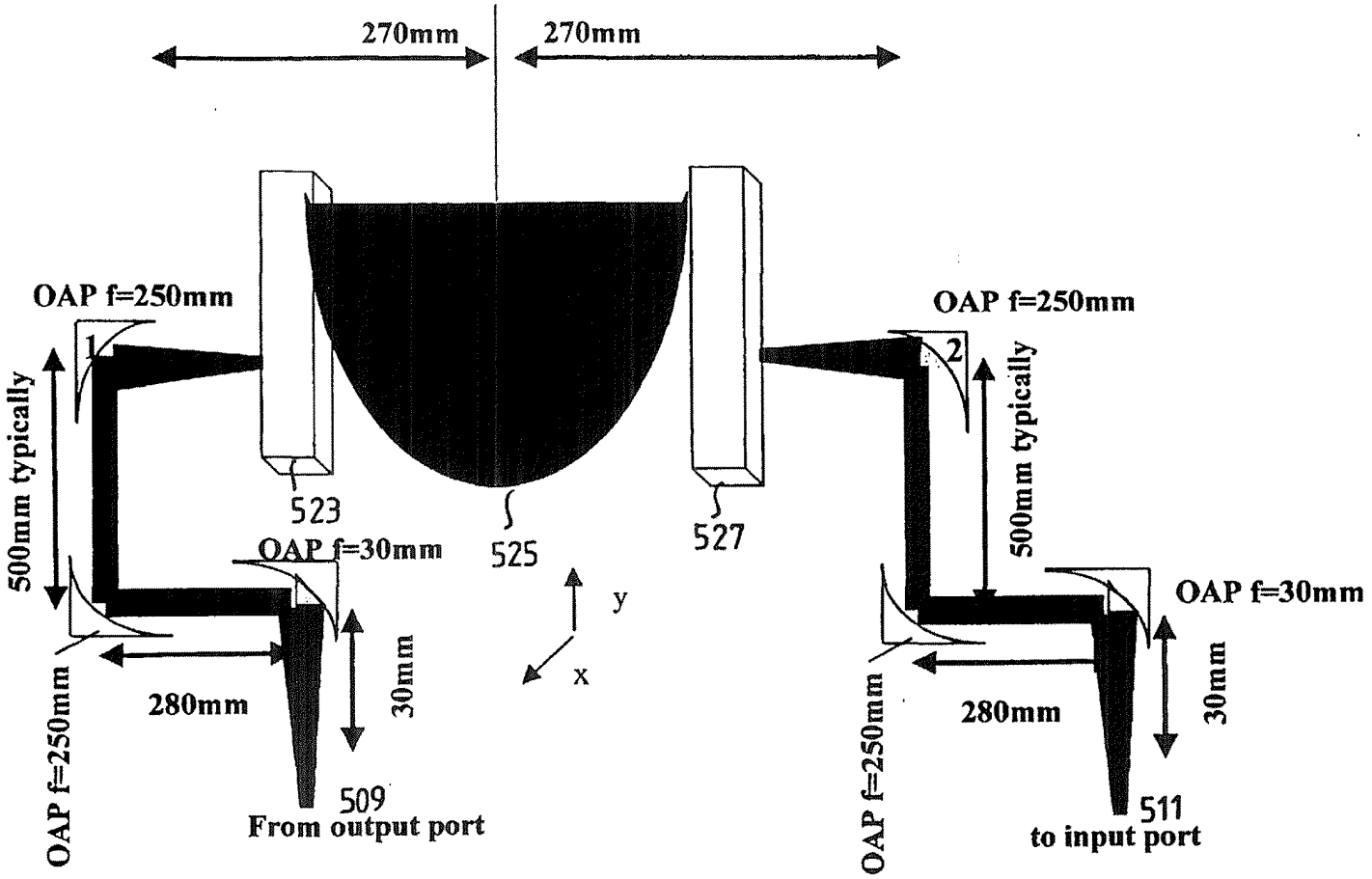


Fig. 31

All mirrors are off-axis parabolics (OAPs).
OAP's designated by 1. And 2. Are on synchronised X-Y stepper motor stages

IPR2023-00033
Apple EX1002 Page 382

: Sample port for THz breast screening
(for use with all-optical systems in Figures 2 and 3) in reflection mode

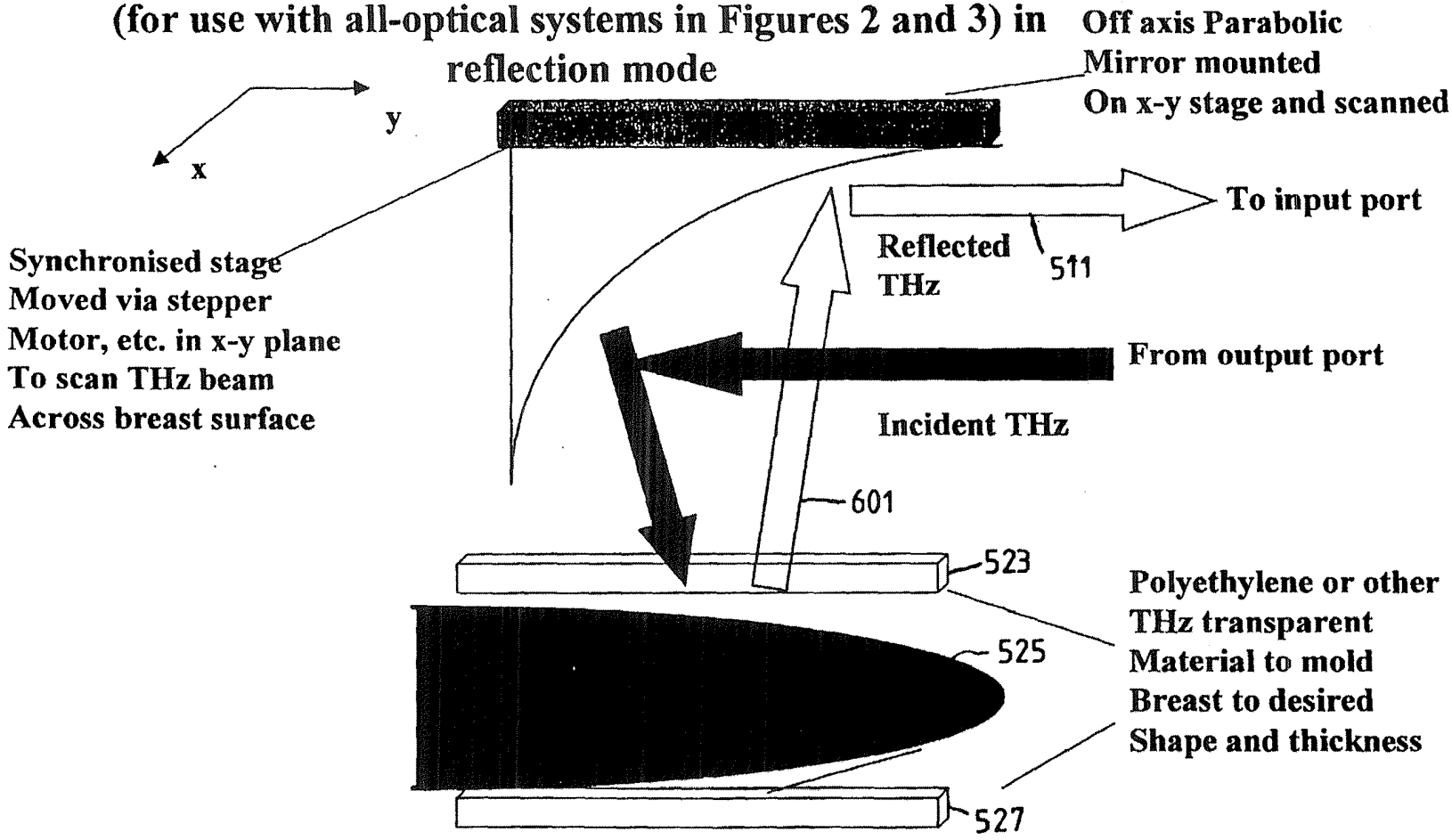


Fig. 32

All mirrors are off-axis parabolics (OAPs).

IPR2023-00033
Apple EX1002 Page 383

: Sample port for THz breast screening
(for use with ail-optical systems in Figures 2 and 3) in reflection mode

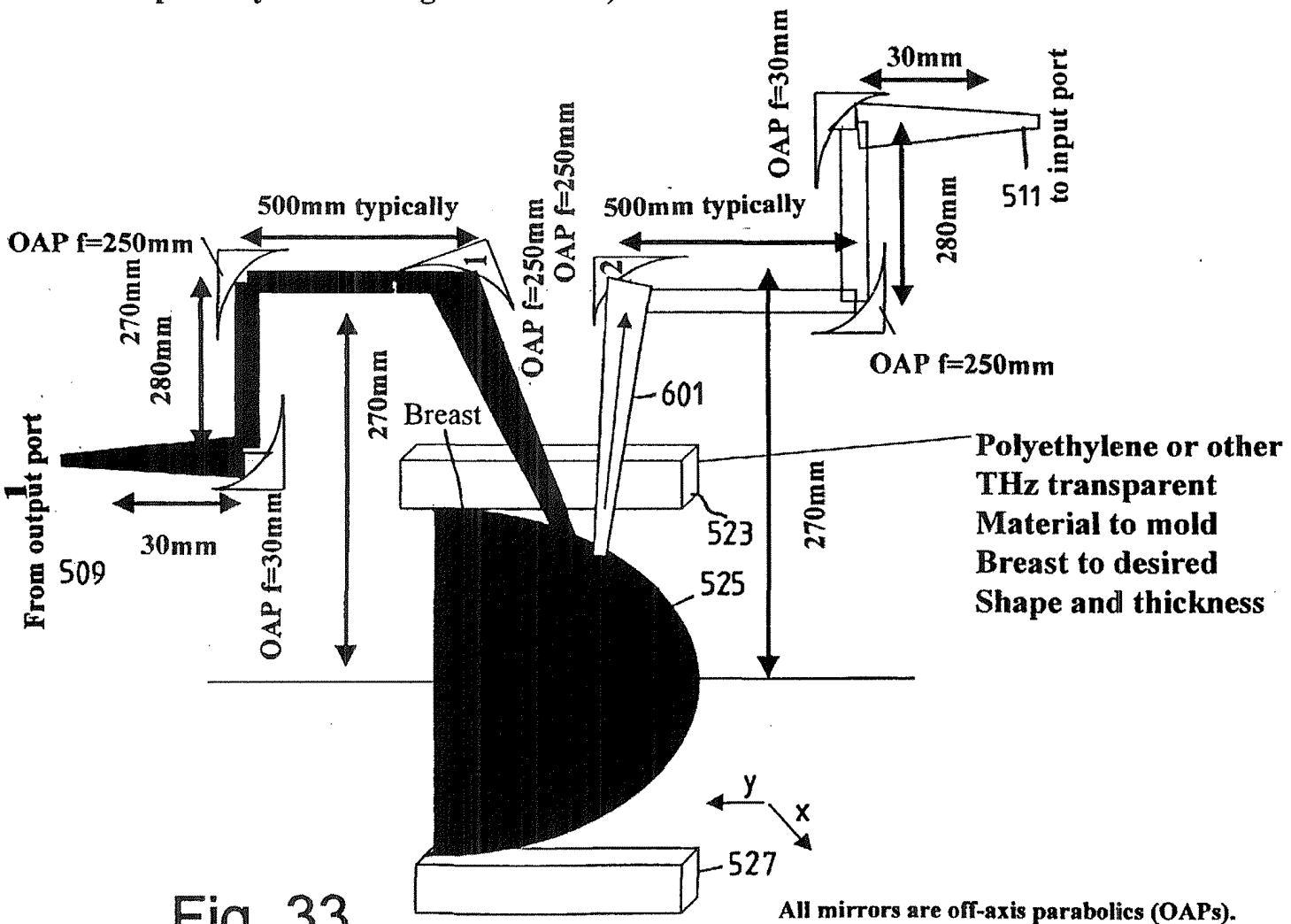


Fig. 33

All mirrors are off-axis parabolics (OAPs).
OAP's designated by 1. And 2.
Are on synchronised X-Y stepper motor stages.

IPR2023-00033
Apple EX1002 Page 384

: All-Optical THz Imaging System with photodiode detection for two input port measurements

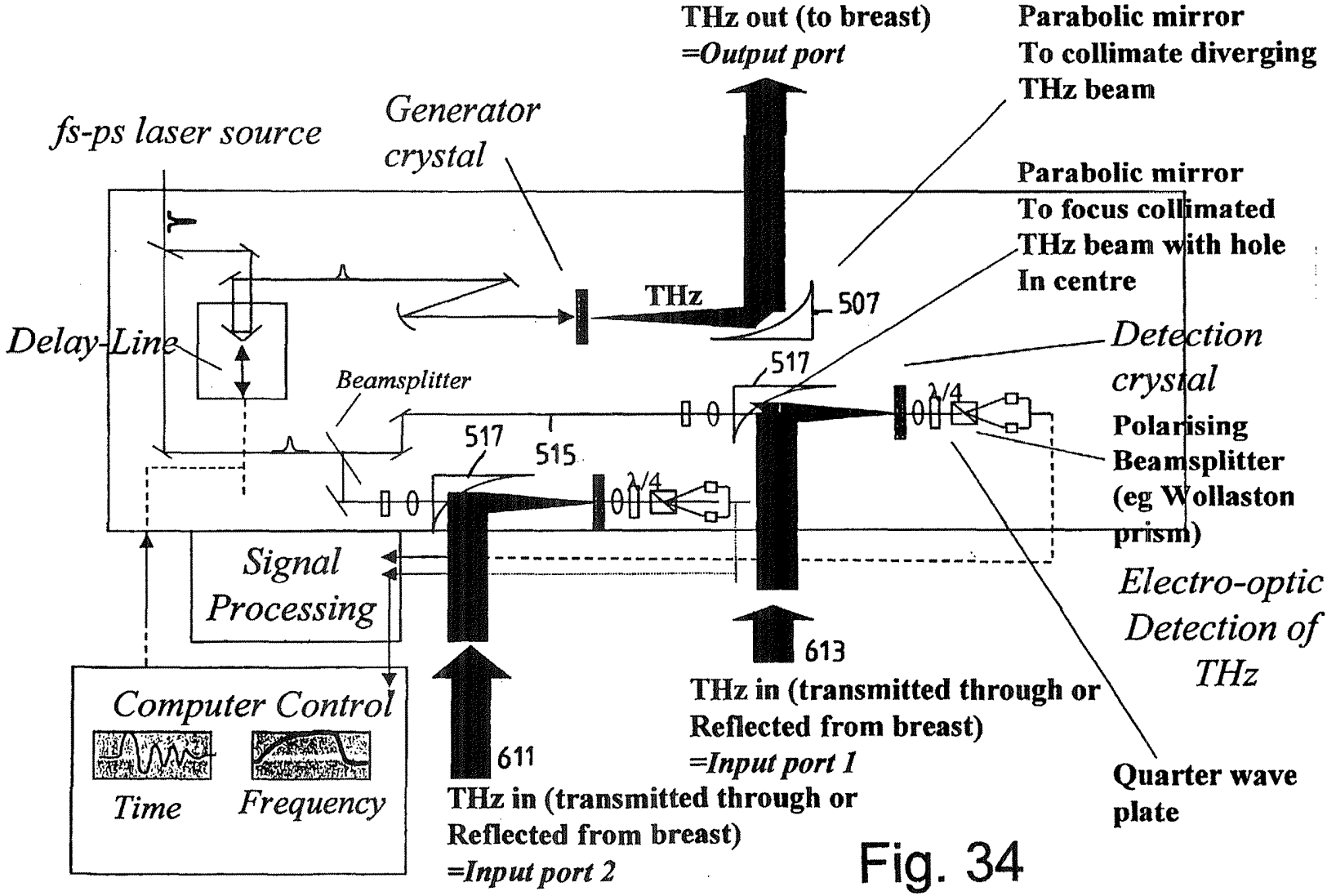


Fig. 34

IPR2023-00033
Apple EX1002 Page 385

: All-Optical THz Imaging System with photodiode detection for two input port measurements

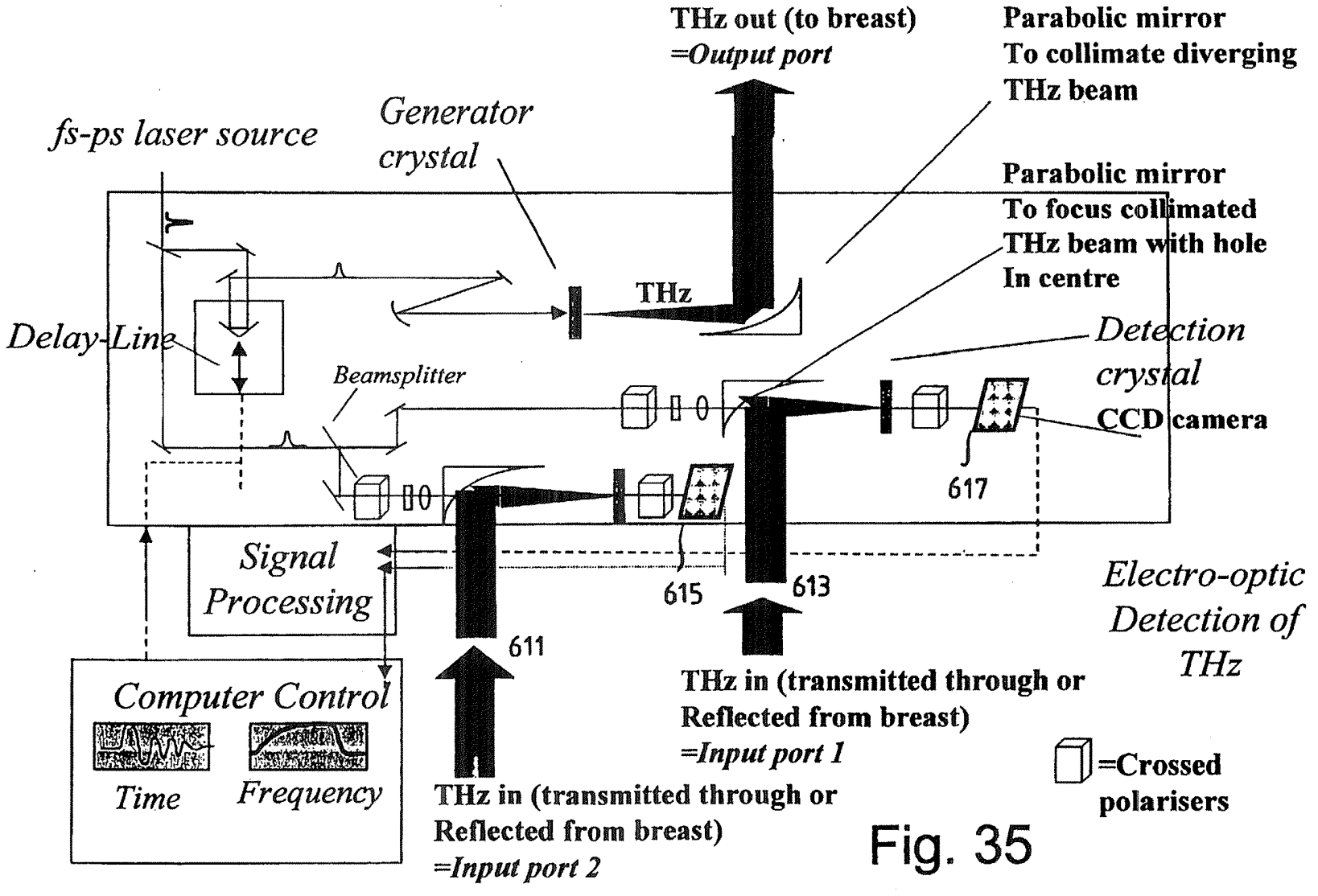


Fig. 35

IPR2023-00033
Apple EX1002 Page 386

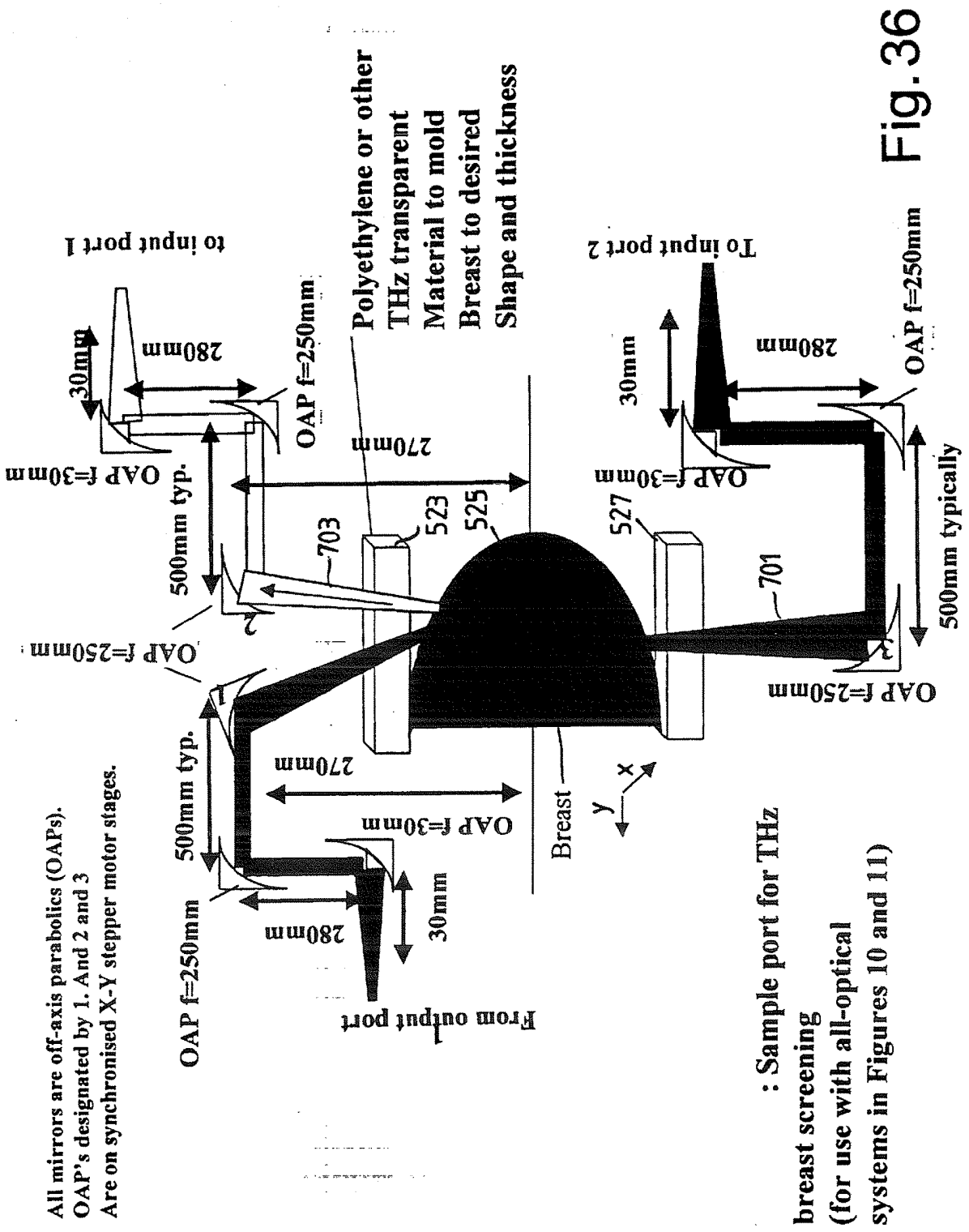

All mirrors are off-axis parabolics (OAPs).
OAP's designated by 1. And 2 and 3
Are on synchronised X-Y stepper motor stages.
: Sample port for THz
breast screening
(for use with all-optical
systems in Figures 10 and 11)
From output port
OAP f=250mm
280mm
30mm
OAP f=30mm
270mm
500mm typ.
OAP f=250mm
Breast
Polyethylene or other
THz transparent
Material to mold
Breast to desired
Shape and thickness
703
523
525
527
701
To input port 2
500mm typically
to input port 1

Fig. 36

: Sample port for osteoporosis screening of bones in foot (for use with all-optical systems in Figures 2 and 3) in reflection mode

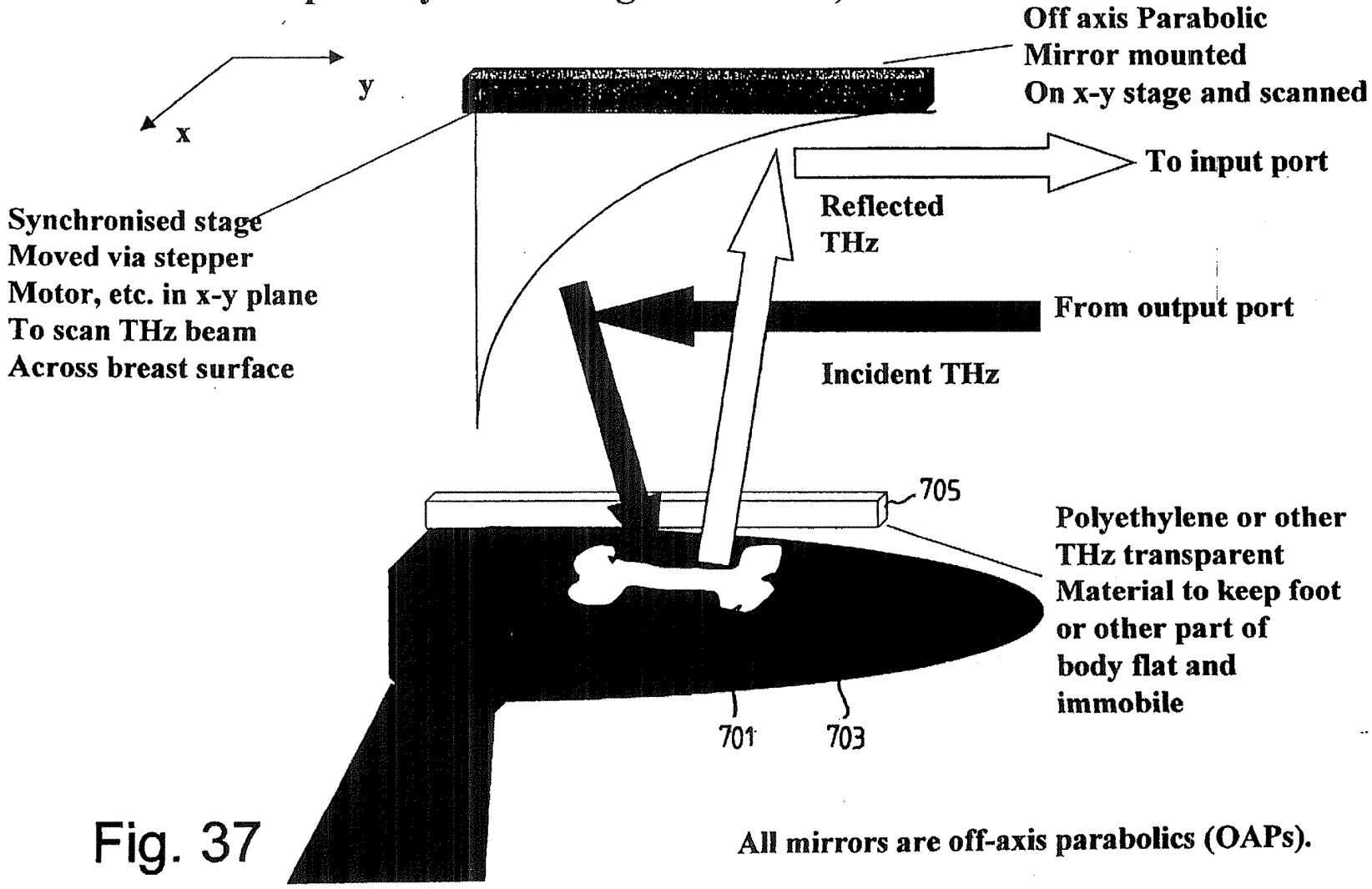


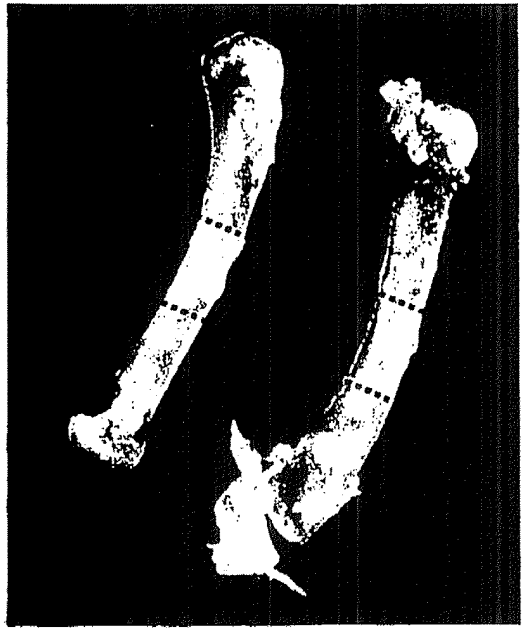
Fig. 37

All mirrors are off-axis parabolics (OAPs).

IPR2023-00033
Apple EX1002 Page 388

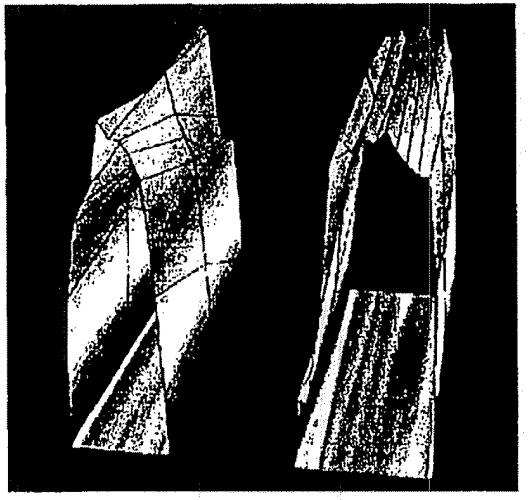
: THz transmission through animal bone

Example of 2 bones - visible image



(a)

3D THz image through 2 bone sample



(b)

Fig. 38

IPR2023-00033
Apple EX1002 Page 389

: Medical Imaging using Terahertz light in a human tooth

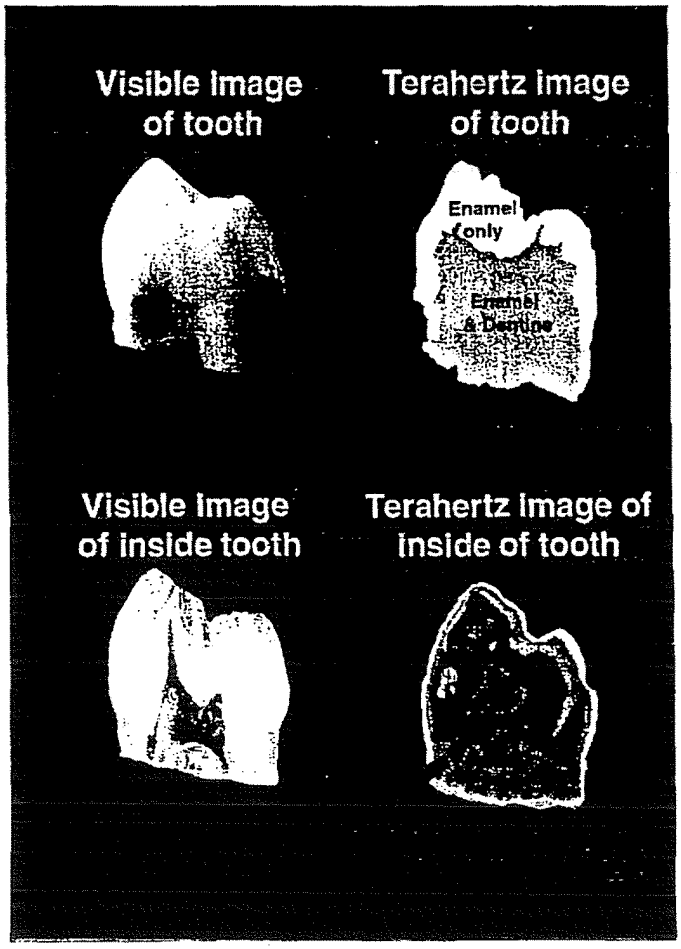
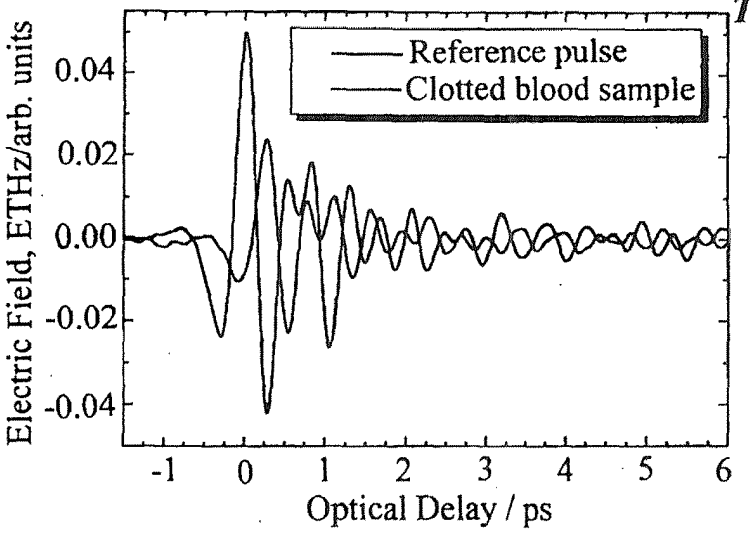


Fig. 39

: *THz imaging of different types of animal tissue: clotted blood*



THz time domain pulse measured through thin layer (90μm) of clotted blood

THz spectrum measured through thin layer (90μm) of clotted blood

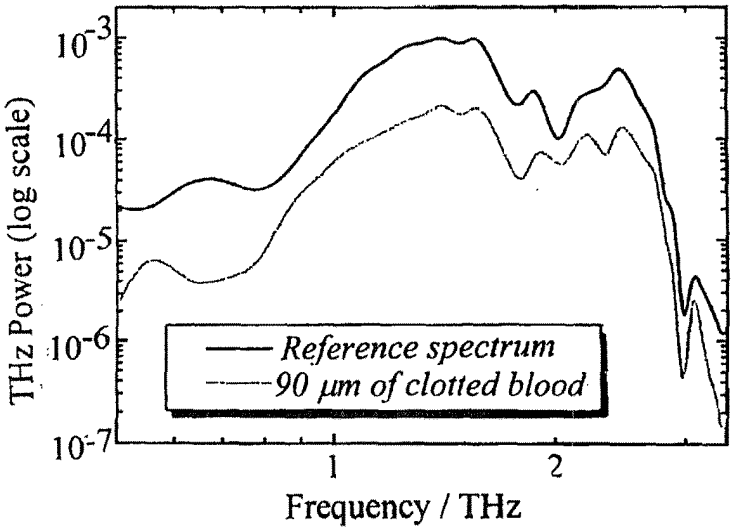


Fig. 40

IPR2023-00033
Apple EX1002 Page 391

Radiation Imaging

The present invention relates to the field of imaging samples with radiation in the infra-red (IR) frequency range. More specifically, the present invention relates to apparatus and methods for improving contrast in images obtained using electromagnetic radiation in the higher Gigahertz (GHz) and the Terahertz (THz) frequency ranges. However, in such imaging technology, all such radiation is colloquially referred to as THz radiation, especially that in the range from 200GHz to 84Thz.

Recently, there has been much interest in using THz radiation to look at a wide variety of samples using a range of methods. THz radiation has been used for both imaging samples and obtaining spectra. Recently, work by Mittleman *et al*, IEEE Journal of Selected Topics in Quantum Electronics, Vol. 2, No. 3, September 1996, page 679 to 692 illustrates the use of using THz radiation to image various objects such as a flame, a leaf, a moulded piece of plastic and semiconductors.

THz radiation penetrates most dry, non metallic and non polar objects like plastics, paper, cardboard and non polar organic substances. Therefore, THz radiation can be used instead of x-rays to look inside boxes, cases etc. THz has lower energy, non-ionising photons than X-rays, hence, the health risks of using THz radiation are expected to be vastly reduced compared to those using conventional X-rays.

The use of THz imaging for medical purposes has also been suggested in the above referenced paper. However, it has previously been believed that strong-water absorption prevents the use of THz in many bio-medical research areas. Previously, a large amount of THz imaging has been analysed on the basis of obtaining contrasts between strongly water absorbing regions and non strongly water absorbing regions.

The present invention addresses the above problems and is concerned with methods for enhancing image contrasts in such THz images to allow THz to be used to not only image contrast based on water absorption but to allow a fine contrast to be shown in THz imaging over a range of different samples.

In a first aspect, the present invention provides a method of imaging a sample, the method comprising the steps of:

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image from the radiation detected in step (b) using a frequency or a selection of frequencies from the plurality of frequencies in the incident pulsed electro-magnetic radiation.

Preferably, the sample is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz; more preferably from 500 GHz to 10 THz.

The detected radiation may be analysed for a single frequency, or, it may be analysed over a selected frequency range. A selected frequency range is taken to be a frequency range which is typically, less than a third of the total frequency range of the pulsed e-m radiation used to irradiate the sample. More preferably, the selected frequency range is less than 10% of the total frequency range of the pulsed e-m radiation used to irradiate the sample.

For example, it has previously been mentioned that water is a strong absorber of THz. There are 'windows' in the water absorption spectra from 50 GHz to 500 GHz, from 30 THz to 45 THz and from 57 THz to 84 THz. If the sample is a irradiated with a range of frequencies from 50 Ghz to 84 THz, it may be preferable to generate the image using one or more of the following selected frequency ranges: 50 GHz to 500 GHz, 30 THz to 45 THz and 57 THz to 84 THz. The image may be generated by integrating over the selected frequency range.

Thus, the present invention allows an image to be created from single frequency or from a selected frequency range. Also, the present invention allows a plurality of images to be derived from a plurality of frequencies or allows a single image to be derived from data from two or more distinct frequencies.

The present invention can either be used to image a sample by detecting radiation transmitted through the sample or reflected from the sample.

The image or images generated by the present invention my be displayed in a number of ways. For example, the method of the first aspect of the present invention may further comprise the step of displaying a sequence of images generated in step (c) for a plurality of different frequencies.

The image generated in step (c) may be scanable through a continuum of frequencies. Alternatively, the image generated in step (c) can be stepped through a plurality of discrete frequencies.

For many imaging contrast techniques discussed herein, a reference signal is required. Ideally, the reference signal is obtained from THz radiation which has not been passed through or reflected from the area of the sample which is to be imaged.

The reference signal may be obtained from a fraction of the electro-magnetic pulsed radiation which has not been passed through the sample, alternatively, the

reference signal may be obtained by passing the pulsed electro-magnetic radiation through a different part of the sample. For example, the present invention may be used to image tumours tissue, the area of the sample which is to be imaged would the tumour, the reference signal could be obtained by passing the radiation through a healthy part of the tissue. The reference signal may also be measured when the sample is absent. For example, the reference signal could be measured before the sample is positioned in the path of the radiation beam or the reference signal could be measured after the sample has been imaged.

In general, the present invention will be performed using imaging apparatus which is configured to detect temporal data at each pixel. Preferably, the data is Fourier transformed to give the complex THz electric field in the frequency domain $E(\omega)$.

The image can be obtained in a number of ways from the complex THz electric field $E(\omega)$, e.g.:

(i) The power spectrum $P_{sample}(\omega)$ of the sample and the power spectrum $P_{ref}(\omega)$ of the reference signal may be calculated. The image could then be generated by plotting the difference between the two Power spectrums for a given frequency for each pixel at a selected frequency over integrated over a selected frequency range.

(ii) The power spectrum P_{sample} of the sample and the reference power spectrum P_{ref} may be divided to give the transmittance. The transmittance may then be plotted for each pixel at a selected frequency over integrated over a selected frequency range.

(iii) The frequency dependent absorption coefficient $\alpha(\omega)$ may be calculated from the complex electric field $E(\omega)$ and plotted for each pixel at a selected frequency over integrated over a selected frequency range.

(iv) The frequency dependent refractive index $\eta(\omega)$ may also be calculated from the complex electric field and plotted for each pixel at a selected frequency over integrated over a selected frequency range.

The detected temporal electric field contains both phase and amplitude information which give a complete description of the complex dielectric constant of the medium in the beam path. The sample to be characterised is inserted into the beam and the shape of the pulses that have propagated through the sample or have been reflected from the sample are compared with the reference temporal profile acquired without the sample. The ratio of the complex electric field $E(\omega)$ and the reference signal $E_{ref}(\omega)$ is calculated to give the complex response function of the sample, $S(\omega)$. In the most simple case, the complex response function is given by:

$$S(\omega) = \frac{E(\omega)}{E_{ref}(\omega)} \propto \exp\left(\frac{i\omega d}{c}(\eta(\omega) - 1)\right) \exp(-\alpha(\omega) d) \qquad (1)$$

where d is the sample thickness, c is the velocity of light in vacuum, η is the refractive index and α is the absorption coefficient. The experimental absorption coefficient $\alpha(\omega)$ and the refractive index $\eta(\omega)$ may then be easily extracted from the magnitude $M(\omega)$ and the phase $\phi(\omega)$ of $S(\omega)$, respectively, according to

$$\alpha(\omega) = -\frac{1}{d}\ln(M(\omega)) \qquad (2)$$

$$\eta(\omega) = 1 + \left(\frac{c}{\omega d}\right)\phi(\omega) \qquad (3)$$

Additional terms may be included in equations (1) to (3) to account for reflections at dielectric interfaces of a sample, thus allowing accurate analysis of multilayered samples.

These parameters are simply related to the complex dielectric function $\varepsilon(\omega)$ of the sample

$$\varepsilon(\omega) = (\eta(\omega) + ik(\omega))^2 = (\eta(\omega) + i\alpha(\omega)c / 2\omega)^2 \qquad (4)$$

The data derived as discussed in (i) to (iv) above, may be directly plotted either as a colour or a grey scale image where the colour or shade of grey of each pixel represents a given magnitude.

Instead of a single frequency, a selected frequency range could be chosen and the result and data of (i) to (iv) integrated over that range. The integrated data could then be plotted.

It may also be preferred to subdivide the magnitude of the data process in accordance with any of (i) to (iv) above into various bands. For example, all data below a certain value could be assigned the value 0, all data in the next magnitude range could be assigned the value 1, etc. These ranges may have equal widths in magnitude or they may have different widths. Different widths may be preferable to enhance contrast e.g. to emphasise contrast in regions of the sample where there is little variation in the sample absorption of THz.

Preferably, the present invention uses two or more frequencies. The data from say two frequencies is processed in accordance with any of (i) to (iv) above. The data is then banded as described for a single frequency above.

The data may be split into two hands, one assigned the value "0" and the other "1". The data from both frequencies can then be added together using a rule such as a Boolean algebraic expression e.g. AND, OR, NOT, NAND, XOR, etc.

Of course, the present invention also allows images to be compared from two different frequencies. This may be particularly useful to identify a substance where the absorption to THz changes over a certain frequency range.

Other methods are also possible for such THz imaging. It is known to plot the maximum or minimum of the electric field. Indeed, this has been done by Mittleman in the earlier reference paper. However, far greater contrast is achieved by plotting the peak to peak signal, i.e. the distance between the minimum and maxima of the electric field. Therefore, in a second aspect, the present invention provides a method for imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with a pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image from the radiation detected in step (b) by calculating the difference between the magnitude of a maxima and a minima of the radiation detected in step (b) for each pixel.

Preferably, the main peak minima and maxima are chosen i.e. the highest and lowest parts of the temporal trace.

More preferably, the sample is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz;

It is also known to derive temporal information from the data. For example, the time difference between the maxima of the electric field which is passed through the sample and the field which is not passed through or reflected from the area of the sample, which is to imaged, can be plotted. However, there has been no suggestion of plotting the temporal difference between the maximum and the minimum of the electric field. This provides data on the frequency dependent absorption. Therefore, in a third

aspect, the present invention provides a method for imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image from the radiation detected in step (b) by calculating the difference between the temporal positions of a maxima and a minima of the radiation detected in step (b) for each pixel.

Further aspects of the present invention provide apparatus for effecting the methods, respectively, of the first, second and third aspects of the present invention, each respectively comprising means for carrying-out each of the steps of those methods.

Preferably, the main peak minima and maxima are chosen i.e. the highest and lowest parts of the temporal trace.

More preferably, the sample is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz;

To obtain even better contrast, the temporal difference between the maximum and the minimum field may also multiplied by a value of the field at that pixel. For example, the minimum, maxima or difference between the minima and maxima.

As mentioned above, the method of the present invention is suitable for medical imaging and is ideally suited to probing certain tissue abnormalities and conditions. It is also useful for imaging items inside containers, for investigating the internal structure

and composition of foods, as well as authentication of documents, labels, banknotes or other printed matter and for analogous security applications. Some particular medical applications will now be explained in more detail.

Cancerous tissue:

Normal tissue contains a large number of mature cells of uniform shape and size. Each cell is characterised by a nucleus of uniform size.

Benign neoplasm:

Benign neoplasm involve cellular proliferation of adult or mature cells growing slowly in an orderly manner in a capsule. These tumours do not invade surrounding tissue but may cause harm through pressure on vital structures within an enclosed structure such as the skull.

Malignant cells:

A malignant cell is one in which the basic structure and activity have become deranged in a manner that is unknown and from a cause or causes that are still poorly understood. The malignant cells lose the normal specialised function of the normal cell or may take on new characteristics and functions. A characteristic of malignant cells that can be observed through a microscope is loss of differentiation, or loss of likeness to the original cell (parent tissue) from which the tumour growth originated. This loss of differentiation is called anaplasia, and its extent is a determining factor in the extent of malignancy of the tumour. Anaplasia is one of the most reliable indicators of malignancy. It is seen only in cancers and does not appear in benign neoplasms.

Other characteristics of malignant cells that can be seen through a microscope are the presence of nuclei of various sizes, many of which contain unusually large

amounts of chromatin, and the presence of mitotic figures (cells in the provision of division), which denotes rapid and disorderly division of cells. The proportion of cells actively proliferating in malignant tumours is generally greater than that of normal cells. Malignant cells have no enclosing capsule; thus they invade adjacent or surrounding tissue.

Spread of Cancer:

It has been calculated in general that a tumour mass will double in size 30 times before it is 1cm in size, when there is a chance for clinical detection by conventional means. The rate of growth of a malignant neoplasm determines its capacity to spread. Cancer may spread by direct extension, by gravitational metastasis, or by metastatic spread.

Direct extension or invasion of neighbouring tissue produces the typical local effects of ulcerating, bulky, hemorhagic masses or indurative, fibrosing lesions with tissue fixation, distortion of structure, and pitting of the skill seen in some breast cancer. Infection may accompany this spread.

Gravitional metastasis involves erosion of cancer cells in the body cavities and their dropping onto the serous membrane lining in the cavity.

Metastatic spread occurs when the cancer cells invade vascular or lymphatic channels and travel to distant parts of the body where implantation occurs. In metastatic spread, there is almost always a high degree of histologic, cytologic, and functional similarity between the primary cancer and these metastases. Consequently, the type of cell and probable site of the primary tumour can be identified from the morphology of the metastasies. In addition, the mestastes usually mimic the primary tumour in the formation of cell products and secretions.

Anaplasia should result in the modification of the transmission, absorption, and reflection properties of tissue 50GHz-40THz frequency range which can be studied by the method of the present invention. The loss of differentiation and additional ulcerating, bulky, hemorhagic masses or indurative, fibrosing lesions with tissue fixation, distortion of structure should modify the following quantities which can be imaged by a method of the present invention.

Refractive index change:

For a given thickness, the probability of interaction between radiation and tissue is dependent on the density of the material. The appearance of fibrous material and additional (mutated) cells in a tumorous region suggests that the density of the tissue in this region will be modified relative to that of healthy tissue, leading to a density change. This in turn should modify the refractive index of the medium (either at a fixed frequency ω or over a range of frequencies) relative to that in normal tissue. These changes in the refractive index can be detected by either 1) simply, by shifts in the time domain of a transmitted or reflected THz pulse or 2) least squares fits to frequency domain information, which will yield the refractive index at each frequency $\eta(\omega)$. In the cancerous regions, we expect increases in the refractive index due to increases in the density of cells in that region. However, exclusion of water in the cancer will also modify the refractive index. Due to either effect, we would anticipate using refractive index determination at different pixels in the image, using either 1), 2) or a combination of the two, as an important contrast mechanism.

Thickness change: Changes in thickness d between abnormal and normal tissue will lead to changes in either

a) the absorption in a given region $a(\omega)d$ or

b) the time delay of the pulse in that region $\eta(\omega)d$.

Both a) and b) can be determined using electric field and time domain spectra, respectively, forming the basis of a contrast mechanism. Because malignant cancer

grows rapidly relative to the tissue which it invades, it is likely that a malignant tumour and the normal tissue adjacent to it will differ appreciably in thickness, which can be determined by the methods of the present invention using a) and/or b) above.

Absorption coefficient change: For a given thickness, the probability of interaction between radiation and tissue is dependent on the density of the material. The appearance of fiborous material and additional (mutated) cells in a tumourous region suggests that the density of the tissue in this region will be modified relative to that of healthy tissue, leading to a density change between normal and cancerous tissue. The linear attenuation coefficient $a(\omega)$ (usually quoted in units of cm^{-1}) is dependent on the density of the material. Thus, if a tumour represents a larger amount of tissue in the same volume, the linear attenuation coefficient $a(\omega)$ will increases proportionally to the increase in density. For a given thickness, this change will manifest itself as increased absorption in the tumour relative to the normal tissue.

Convolved with such density-induced changes in $a(\omega)$ will be variations in $\alpha(\omega)$ due to changes in the composition of the tissue itself. Changes in chemical composition are the most obvious reason for this. For example, the presence of ulcerating, bulky, hemorhagic masses in tumours implies a change in chemical composition relative to healthy tissue. In certain types of malignant tumours such as cytologic anaplasia, there is increased or altered nucleic acid synthesis in growing tissue, which implies a change in chemical content. Moreover, simpler changes such as the exclusion or inclusion of additional water in a tumour will change its chemical composition relative to that of healthy tissue.

$\alpha(\omega)$ may also change in the THz range due to increased disorder in the system. Thus, even if there is no change in the type of molecules in tumourous vs normal tissue, the tumour is likely to be more disordered in terms of the arrangements between cells. A similar phenomenon occurs in liquids, where increasing disorder (or lack of crystallisation) leads to increased absorption. It is therefore probable that $\alpha(\omega)$ will be larger in randomly arranged tumour tissue than in more "regularly arranged" tissues.

II.1.d. Use of TPI and THz TDS to detect increased vascularity around tumors and cancerous regions. As noted in the text above, the increased number of blood vessels necessary to feed a tumor can be a first indication of the appearance of cancer. Therefore, any element of TP1 or THz-TDS which detects increased blood flow or the increased presence of blood in a certain region may be useful for the detection of cancer.

The present invention will now be described by way of example and with reference to the following non-limiting embodiments in which:

Figure 1 shows a typical THz time domain data obtained from imaging a sample;

Figure 2 shows a power spectrum of a THz trace originally measured in the time domain;

Figure 3a is an image of a pork sample at visible light wavelengths and figure 3b is the same pork sample at THz frequencies;

Figure 4 shows three power spectra for different pixels of the image of figure 3b;

Figure 5 shows data for absorption coefficient for dried pork skin and dried pork kidney samples;

Figure 6 shows THz electric field data and time domain and power spectrum frequency domain for different types of chicken tissue;

Figure 7 shows power spectra obtained from the data of Figure 6 plotted on a single axis;

Figure 8 shows a visible image of a spider mounted over a hole in the metal plate;

Figure 9 shows a visible image of the spider in a box with the box open (Figure 9a) and the box closed (Figure 9b);

Figure 10 shows the THz image of the spider in the box of Figure 9b;

Figure 11 shows the THz electric field and power spectrum derived from a blood sample;

Figure 12 shows the THz electric field and power spectra derived from a water sample;

Figure 13a shows a visible image of a commercial package transistor, Figure 13b shows a THz image of Figure 13a and Figure 13c shows a cut away image of the transistor of Figure 13a;

Figure 14a shows the THz image of doped and undoped GaAs, Figure 14b shows the same sample using visible radiation;

Figure 15 shows a monochromatic image of the semiconductor of Figures 13 and 14;

Figure 16 shows a monochromatic image similar to Figure 15 but at a different frequency;

Figure 17 is a schematic diagram of a THz imaging system;

Figure 18 shows a variation of imaging system of figure 17;

Figure 19 shows a schematic of a detection section which can be used with the imaging system of Figure 18;

Figure 20 shows the imaging system of Figure 18 with the detection section of Figure 19;

Figure 21 shows a variation on the imaging section of Figure 20;

Figure 22 shows a variation on the imaging section of Figure 20;

Figure 23 shows a variation on the imaging section of Figure 20;

Figure 24 shows a further example of an imaging; and

Figure 25 shows another example of an imaging system.

Figure 26 shows a THz imaging system configured for detecting breast cancer;

Figure 27 shows a variation on the system of Figure 26;

Figure 28 shows a schematic view of breast screening apparatus using the detection systems of Figures 26 and/or Figure 27;

Figure 29 shows in detail the sample receiving area of the breast screening apparatus of Figure 28;

Figure 30 shows an alternative configuration of the optics of the breast screening apparatus;

Figure 31 shows yet another configuration of the optics of the breast screening apparatus;

Figure 32 shows a variation on the sample receiving area of the breast screening apparatus;

Figure 33 shows a further variation on the sample receiving area and optics of the breast screening apparatus;

Figure 34 shows a further variation on the optics of the breast screening apparatus;

Figure 35 shows yet a further variation on the optics of the breast screening apparatus;

Figure 36 shows a further variation on the optics and sample receiving area of the breast screening apparatus;

Figure 37 shows apparatus for studying bone;

Figure 38 shows visible and THz images of chicken bone;

Figure 39 shows THz and visible images of teeth; and

Figure 40 shows THz images of clotted blood.

The sample is imaged by irradiating with pulsed electro-magnetic radiation in 100GHz to 20THz. The THz radiation is either passed through the sample or reflected from the sample. Figure 1 shows a typical trace of detected electric field over time detected from the sample for a single pixel. If there is a variation in the sample composition from pixel to pixel, then this will show up in the THz trace for each pixel.

As can be seen from Figure 1, there is a vast amount of information obtained for each pixel therefore, the information needs to be processed to provide a meaningful image. There are four main parameters which can be derived from this time dependent signal. These are: the maxima of the electric field (E_1), the minima of the electric field (E_2), the temporal position of the maxima of the electric field (T_1) and the temporal position of the minima of the electric field (T_2).

To obtain a THz image, one or more of these parameters may be plotted directly. However, further contrast may be obtained by plotting the following functions:

1. The peak to peak height i.e. E_1 and E_2;

2. The difference in the temporal spacing between a position of the maxima E_1 and the position of the minima E_2 i.e. T_1-T_2. This parameter is particularly useful as it gives information on the frequency dependent absorption of the sample.

3. The product of the temporal peak position with one or more electric field parameters. For example:

$$(T_2-T_1) \times E_1$$

$$(T_2-T_1) \times E_2$$

$$T_2-T_1 \times (E_1+E_2)$$

Further information may also be obtained from Figure 1 by performing a frequency analysis. Figure 2 shows a power spectrum for a typical trace. The power spectrum is obtained in the conventional way from Fourier transforming the temporal data (i.e. the type of data shown in Figure 1). The magnitude of the signal is plotted logarithmically on the y-axis and the frequency is plotted on the x-axis. The data for Figure 2 was

derived from a sample which had regions of fat, kidney and meat. This will be described later in more detail with reference to Figure 3. A reference signal is also shown which is the power spectrum of radiation which has not passed through the sample. By viewing the data in this form, the spacing between the data (along the y axis) for the three tissue types is seen to vary with frequency. Therefore, by choosing an appropriate frequency (f_{select}) an image with strong contrast may be obtained.

Also, to distinguish certain tissue types it is often advantageous to plot the image at two separate frequencies for comparison. This is particularly useful when there is little contrast between the tissue types at various frequencies. For example, taking the data of Figure 2, it is seen that the power spectrum corresponding to kidney tissue and fat tissue has an almost identical value at 1THz. Therefore, in an image generated at this frequency, both fat and kidney would look very similar. However, looking at the trace at around 2.5THz, it is seen that the kidney data has a similar value to that of the meat data. Therefore, at an image taken at 2.5THz there is very little contrast between the meat and the kidney data. However, comparing the image taken at 1THz with the image taken at 2.5THz, allows the three tissue types to be equally distinguished. Two or more images at different frequencies can also be compared by a mathematical operation. For example, the data taken at 2.5THz could be subtracted from the data taken at 1THz to clearly show the characteristic signal due to absorption by kidney tissue. Other mathematical operations are also possible such as multiplication, addition, and division etc.

To further enhance contrast, the data can be subdivided into a series of bands based on magnitude. In Figure 2, four bands are shown. Each of these bands can be assigned a single numerical value and this data can then be plotted. In the data of Figure 2, it can be seen by assigning a single value to band 2 and a single integer value to band 2 and a single integer value to band 3 where substantially enhanced contrasts between the fat tissue and the kidney tissue. This analysis can further be extended to just two amplitude bands. In one amplitude band, the amplitude is assigned a value of zero and in the second amplitude band the amplitude is assigned a value of one. This

IPR2023-00033
Apple EX1002 Page 409

analysis allows data at two different frequencies, f_1 and f_2 to be compared using certain rules. Such a rule could be:

Signal (f_1)	Signal (f_2)	New Value
0	0	1
1	0	2
0	1	3
1	1	4

Also, certain Boolean algebraic expressions could be used, for example, AND, NAND, XOR, OR, NOR. For an AND comparison, the following table would be used:

Signal (f_1)	Signal (f_2)	New Value
0	0	0
1	0	0
0	1	0
1	1	1

Instead of the power spectrum, the frequency dependent absorption coefficient $a(\omega)$ or the frequency dependent refractive index $\eta(\omega)$ could be analysed in the same way as the power spectrum.

Figures 3a and 3b show images obtained from a pork sample.

The pork sample consists of a variety of different pork tissue types, meat, fat and kidney. Each tissue type is approximately 1mm thick, and the area measured using THz is approximately 13x10mm^2 with a 500μm step size. The temporal scan at each pixel was 10ps long, giving a frequency resolution of 100GHz, and each pixel took approximately 1 minute to measure. The sample was mounted on cellulose nitrate film and a 1mm thick polythene window, before being placed behind a THz radiation source (<110>ZnTe; 1mm thick, 20x25mm^2). Details of the apparatus used to obtain the images will be described later with reference to Figures 18 to 26.

At each pixel, the tissue type was identified using both the change in the magnitude and temporal position of the electric field maximum. All tissue types were found to transmit THz pulses, with the absorption increasing in the order: fat, meat, and kidney. The THz image shown in Figure 3(b) is a panchromatic THz image since the analysis of the absorption was not frequency specific.

Comparing the visible and THz images, the different tissue types have been successfully imaged and identified at THz frequencies. Due to the simplistic nature of this analysis there is an increased error along the boundary of each tissue region due to some scattering, which results in tissue types being incorrectly identified in these regions. Figure 4 shows power spectra for each tissue type . The data was taken at 3 different pixels. Clear differences in the absorption characteristics for each tissue type are shown.

Figure 5 shows data of the absorption coefficient for skin and kidney using the attenuation of the peak of the THz electric field. All samples were dried to minimise water absorption effects in the data. The data illustrates that the absorption coefficients for the two tissue types are different, which may be used to identify the tissues in an image; the thickness can be found from the temporal shift. Thus, the maximum tissue thickness that can be measured assuming a signal to noise ration (SNR) of 10^4:1. For this example, these values are 3.0mm for kidney and 7.5mm for skin. With the addition

of an amplifier to boost visible, and hence THz, powers 10^5:1 - 10^6:1 can be achieved, which allows several cm of tissue to be probed in theory.

Figure 6 shows data from six types of chicken tissue: bone, cartilage, skin (cooked and uncooked), meat (cooked and uncooked). All tissue types were found to transmit THz by varying amounts. The meat has the strongest absorption of all the tissue types, with the electric field results multiplied by five for clarity. The frequency domain data (NH5) shows a clear difference in the nature of the absorption for each type of tissue.

Figure 7 shows the data from Figure 6 plotted on a single axis for comparison.

THz can be used to image the contents of objects that are otherwise opaque at particularly visible wavelengths, much in the same way as x-rays are currently used. However, due to the non-ionising nature of THz plus the low average powers, THz is inherently safe. For example, common packaging materials such as plastics, paper and cardboard have been found to be transparent at THz frequencies. Figure 8 shows a visible image of a spider mounted over a hole in a metal plate. The spider is placed in a plastic box for imaging as shown in Figure 9a. At visible wavelengths it was not possible to observe the contents of the plastic box with the lid closed, Figure 9b.

The spider and the box were imaged over a 10x10 mm region with a 200μm step size. Analysing the image using the change in peak of the THz electric field, a THz image was formed of the contents of the box, Figure 10. The light areas correspond to regions of highest THz transmission. The spider is clearly resolved inside the box, with fine detail, such as a break in one of legs (top right corner) resolvable. Closer inspection of the Figure 10 also revealed some weak internal structure inside the spider.

Of considerable interest to medical applications would be the ability to measure the properties of blood using a non-invasive method on a human subject, i.e. a

"contactless" in vivo method of assessing the constituents of the subject's blood. Thus, a technique that combines imaging and spectroscopic capabilities to locate and identify, respectively, different tissue types and their constituents would be of enormous medical and commercial value. Also, for imaging purposes blood is (partially) transmitting at THz frequencies. A sample of dried blood (black pudding) was taken and a thin layer spread on a piece of cellulose nitrate film. The transmission was measured at a single point on the sample, Figure 11a, with the peak electric field decreasing by 55%. In the frequency domain, Figure 11b, the absorption is found to be strongest between 1.00THz and 1.90THz, with a strong absorption around 0.66THz.

One of the most important materials that must be considered for biological applications of THz is liquid water. Water is known to have a number of strong absorption bands in the infrared and far-infrared/THz regions due to the polar nature of the molecule, and ultimately it may be absorption due to liquid water in biological samples that limits the applications of THz technology. To quantify the strength of water absorption a thin (90μm) layer of liquid water was measured. This layer was formed on the surface of a piece of cellulose nitrate film. The time-domain results for the sample and a reference (dry cellulose nitrate) are shown in Figure 12a, and the frequency domain results in Figure 12b. From the time-domain results alone the absorption is clearly stronger than in the blood sample. Considering the frequency domain results, the absorption is found to increase with increasing frequency; at frequencies above 2.5THz the repeatability of the spectra is less reliable and therefore only the frequency range 0.6THz to 2.5THz should be examined. It is therefore important for biological applications that the frequency range of THz spectroscopy is extended to higher and lower frequencies where absorption windows are present. It should be noted, however, that the ability of THz to probe the skin, blood, and water thicknesses noted above suggests that it is possible to use the present invention.

In Figures 14 a to c, a commercially available, discrete transistor was taken which has the active elements encapsulated in a black casing that is opaque at visible wavelengths, Figure 13a. Fortunately, because the black casing is partially transparent

to THz, the internal structure of the device can be imaged. The THz image in Figure 13b was constructed using 150μm steps over a 5x6mm^2 region centred on the transistor. Where there is only plastic, THz pulses are transmitted, but where there are metal tracks of the three leads, no THz is transmitted. To confirm the accuracy of the THz image, an identical transistor was cut open and the same lead arrangement was found, as recorded in the visible photograph of Figure 13c. Additional information on the depth of the leads below the outer surface could be obtained using reflection-based imaging. Scanning the THz beam across the front of the sample, strong reflections were recorded in regions with the metal tracks, with the temporal delay of the pulses giving the distance of the tracks below the surface.

A simple image of the distribution of the doping in a semiconductor can be formed using the change in the peak of a transmitted THz pulse, due to the approximately linear increase of THz free-carrier absorption with carrier density. Figure 14b shows a visible image of a sample consisting of two strips of undoped GaAs and two strips of n-type (n~1x10^{18} cm^{-3}) with the strips arranged alternately side by side. Each strip was 10x1x0.5mm. An area of approximately 8x8mm^2 was imaged using THz, with a 200μm step size. While at visible wavelengths there was no means of identifying the GaAs, the THz image clearly shows a difference in the absorption, Figure 14a. Here the peak of the THz field is plotted for each pixel. The highest transmission (at the bottom of the image) is where there was no GaAs present. Moving across the GaAs strips there is variations in the absorption due to the different dopings, with absorption strongest in the n-type strips. Strip #3 also shows a step in the absorption at the boundary of the etch through the n-type doping.

Figures 15 and 16 show frequency dependent (monochromatic) images derived from the THz data of the semiconductor sample. Instead of plotting the peak of the field, power spectrum for each pixel is found using a fast Fourier transform and the power at a given frequency selected. Figures 15 and 16 show the images for f=2.38Thz and f=890Ghz respectively. At 2.75 THz the image is similar to Figure 14a, with the difference in the absorption between the two type of GaAs clearly resolvable.

IPR2023-00033
Apple EX1002 Page 414

However, for the 890GHz image there is no difference in the transmission for the undoped and areas with no sample, with both appearing as the same shade of grey indicating that no absorption in the GaAs at this frequency.

Figure 17 shows a basic THz imaging system. The system can be simplified into three main sections, a generator 31, an imaging section 33 and a detection section 35. THz radiation is generated in the generating section 31 by using a THz emitter which is supplied by a visible pulsed laser 37.

A THz beam 39 is emitted from generation section 31 and is directed onto sample 41 of the imaging section 33. The THz beam 39 is then directed via further optics 45 into the detection section 35. The system of Figure 17 is an example of a near field imaging system, where the sample to be imaged is placed immediately behind the THz source.

The detection section reads the information carried in the detected THz signal via a visible light signal and AC Pockels effect. The visible light is ideally obtained from laser 37 via beam splitter 47. A time delay is added to the THz pulse via time delay line 34. The system (e.g. the control of the sample 41 movement, the time delay 34 and the detected signal processing) is controlled by computer 36.

Details of the AC pockels effect will be described with reference to Figure 19.

Figure 18 shows another imaging system. Here, for simplicity, details of the detection part of the system will be omitted. These will be described with reference to Figure 19.

The THz generation section is indicated by the components within box 51. The imaging system requires both a visible light pulse and a THz pulse to be emitted from the generation section 51. Therefore, the output coupler 23 should not 100% reflective

to visible radiation to allow some visible radiation to be emitted from the THz generation section 51.

The emitted THz beam 53 and visible 55 from the generation system are incident on beam splitter 57. This beam splitter 57 allows transmission of the THz beam 53 but reflects the visible light beam 55 onto mirror 59 which reflects the beam 55 into optical delay line 61. The delayed beam 55 is then inputted into the THz detection unit 63.

The THz beam 53 is directed into the imaging section 52 and onto sample 65 via THz imaging optics 67. The sample 65 is located on a motorised X-Y translation stage (not shown) so that the whole sample 65 can be imaged. (The x-y plane is orthogonal to the beam axis). The THz radiation 69 carrying the imaging information from the sample is reflected into the THz detection system 63 via THz imaging optics 71.

The presence of visible radiation 55 as well as THz radiation 69 allows for imaging and electro-optic detection to be performed inside a single nitrogen-purged unit.

The sample 65 is mounted on a X-Y motorised translation stage (not shown) which is controlled by a PC computer (not shown). Each section (pixel) of the object may then be imaged. To improve the spatial resolution of the technique, off-axis parabolic mirrors, condenser cones, and lenses may be used to focus the beam to a diffraction limit spot. By mounting the sample in the near field of a condenser cone, the diffraction limit may be overcome and spatial resolution of about 50 μm may be achieved. The imaging system can function with or without such objects depending on the nature of the object to be imaged and the nature of the detection circuit.

Figure 19 shows the detection system in detail. The THz beam 69 carrying the imaging information and a visible light beam 55 are inputted into the THz detection system. The visible light beam 55 acts as a reference beam which is incident on the

detection crystal 73. The reference beam 55 is linearly polarised and the polarisation is orientated such that it has components along both the ordinary and extraordinary axis of the detection crystal 73. Each of the axes has distinct refractive indices n_o and n_e along the ordinary and extraordinary axis of the crystal 73 respectively. In the absence of a second (THz) radiation beam 69, the linearly polarised reference beam 55 passes through the detection crystal 73 with negligible change in its polarisation.

The applicant wishes to clarify that the angle Θ through which the polarisation is rotated by is negligible. However, the linearly polarised beam can become slightly elliptical. This effect is compensated for by a variable retardation waveplate, e.g. a quarter waveplate 81.

The emitted beam 77 is converted into a circularly polarised beam 83 using quarter wave plate 81. This is then split into two linearly polarised beams by a Wollaston Prism 79 (or equivalent device for separating orthogonal polarisation components) which directs the two orthogonal components of the polarised beam onto a balanced photodiode 85. The balanced photodiode signal is adjusted using wave plate 81 such that the difference in outputs between the two diodes is zero.

However, if the detector 73 also detects a secondary beam 69 (in this case a beam with a frequency in the THz range) as well as a reference beam, the angle Θ through which the polarisation is rotated by is not negligible. This is because the THz electric field modifies the refractive index of the visible (fundamental) radiation along one of the axes n_e, n_o. This results in the visible field after the detector 73 being elliptical and hence the polarisation components separated by the prism 79 are not equal. The difference in the voltage between the output diodes gives a detection voltage.

The reference beam 55 and the THz beam 69 should stay in phase as they pass through the crystal 73. Otherwise the polarisation rotation Θ is obscured. Therefore, the detection crystal 73 has phase matching means to produce a clear signal.

All of the items shown in Figure 20 sit on an optical bread board of dimensions 36 inches by 36 inches. The only external units required are a power supply for the diode laser and a cooling unit for the generation section 51.

The imaging section 91, has a motorised stage which is movable in the x-y plane, i.e. along two orthogonal axis which are perpendicular to the incident beam of THz radiation.

The imaging section 91 has two mirrors M11 and M12. Mirror M12 directs the THz beam 53 onto the sample 65. Mirror M11 is positioned to reflect the THz radiation transmitted through sample 65 onto the detection crystal 73. Mirrors M11 and M12 are off axis parabolic (OAP) mirrors. Such mirrors are configured so that the phase difference between the incident and reflected beams is the same at all points on the mirror. The parameters resulting in an off axis parabolic surface are characterised by the focal length of the mirror.

An optical delay section 93 is also shown. The visible light beam emitted from the generating section is reflected by beamsplitter 57 into delay section 93. The delay section 93 has a corner cube mirror M9 which is moveable along the beam axis. The beam is directed onto corner cube mirror M9 via mirror M8. The beam is reflected off corner cube mirror M9 onto mirror M10. Corner cube mirror M9 is oscillated back and forth along the beam direction with an oscillation frequency of several 10s of Hz. This increases or decreases the path length of the visible beam 55 as required. A Clark ODL-150 system may be used to drive the mirror, this is capable of delays of 150ps. The emitted beam is then combined with the emitted THz beam at mirror M11. Alternatively the THz and visible beams may be combined colinearly using a beam splitter, for example, a pellicle beam splitter. Such a device would be placed before or after M11 and would eliminate the requirement for a hole in M11.

Figure 21 shows a variation on the imaging section 91 of Figure 20. The extended path length over which the THz beam 53 travels is purged with nitrogen to remove water vapour and hence improve the quality of the image.

Due to diffraction effects associated with the large wavelengths in the THz range, the cross-sectional size of the THz beam 53 and imaging applications is not sufficiently large that it may be treated as plain parallel. If diffraction effects are such that the radiation is paraxial so that it can be represented by a scalar field distribution, Gaussian beam mode optics and optical techniques can be used. The simplest case for system design is to assume that the fundamental mode dominates the beam profile. The use of Gaussian mode optics and design applied to conventional THz radiation and systems (generated in the Fourier transform machines, far-infrared lasers or Gunn diodes) is applicable and important to THz imaging systems.

A number of design rules or guidelines should be followed when constructing a THz imaging system to obtain a good quality image. For transmission optics such as lenses, geometric losses are kept to a minimum by ensuring that the ratio of the lens thickness to focal length and diameter to focal length is less than 0.2. If this is satisfied, then losses in the lenses will be primarily due to absorption and reflection. In this case, choice of materials is important.

A requirement which arises in pulsed systems is the need for the material to be non-dispersive so that pulse broadening does not occur. Given these requirements, high density polyethylene (DHPE), polytetrafluorethylene (PTFE), high resistivity silicon (Si), and TPX are some of the best materials and can also be machined in a lathe; any material combining low absorption and low dispersion at THz frequencies is a good candidate for fabrication of transmission optics, provided its shape can be suitably fabricated for a lens. Reflection losses in lenses tend to be highly frequency dependent at THz frequencies. Therefore care must be taken in lens design to ensure that all frequencies across the pulse bandwidth undergo the same reflection (and absorption) losses.

Ideally, reflective optics (mirrors) are used wherever possible instead of transmission optics (lenses) in order to minimise a number of losses associated with transmission optics, which include (i) frequency-dependent reflection losses and amplitude pattern distortion at dielectric (e.g. air-lens) interfaces, (ii) frequency-dependent absorption losses, (iii) diffraction effects and distortions to field distribution due to power falling on the lens's surface at an angle.

An additional property of importance in imaging (and not particular to Gaussian mode beam optics) is that if two mirrors are separated by the sum of their focal lengths, then the size of the beam waist (minimum beam diameter in plane normal to optical axis) on the optical axis after the reflection from the second mirror will be frequency-independent. This is true of the last mirror (focusing element) in a chain provided there are an even number of focusing elements in the chain. This provides a major advantage for THz imaging as the pulse is comprised of a wide range of frequency components, and it is desired to keep the object at a fixed position on the optical axis whilst images are being recorded at various (x, y) points and at all THz frequencies in the pulse. This is particularly important THz imaging as the spectral coverage (bandwidth) of THz pulses increases into the mid-infrared and even higher frequencies.

The system in Figure 21 will produce beams with 1/e diameters (for the fundamental Gaussian mode in the beam) of 1-2 mm at the sample in the THz frequency range (e.g. at 300GHz, diameter = 2mm). In the system of Figure 21, six mirrors and two lenses are used as opposed to the two mirrors of Figure 20. The direction of the beam in Figure 21 is reversed to that in Figure 20. In the imaging section, the beam is first reflected off first OAP mirror 101 onto second OAP mirror 103 and then onto third OAP mirror 105. Second 103 and third OAP mirrors each have a focal length of 250 nm. They are separated by 500 nm.

The beam is reflected from the third OAP mirror 105 onto plano-convex lens 107 which has a focal length of 10 mm and a diameter of 10 mm. Third OAP mirror

105 is separated from plano convex lens 107 by 260nm (i.e. the sum of their focal lengths). The lens 107 is made from polyethylene or high resistivity Si. The lens 107 is placed 10 mm from the motorised stage (not shown) on which sample 109 is mounted. The beam has traversed through an even number of focussing optics and mirrors (101, 103, 105 and 107) which are all spaced apart by the sum of their focal lengths. Hence, the waist of the beam at the sample is independent of the frequency. Here, the beam diameter is chosen to be 2 mm, independent of frequency in the frequency range of about 300 GHz (0.30THz).

Once the beam has passed through sample 65, the transmitted THz radiation falls onto second plano convex lens 111. Plano convex lenses 107 and 111 are identical in optical characteristics. Lens 111 focuses the THz radiation onto the fourth OAP mirror 113. Fourth OAP mirror 113 has a focal length of 250 mm and reflects the THz beam onto fifth OAP mirror 115. Fifth OAP mirror 115 also has a focal length of 250 mm and lies 500 nm away from the fourth OAP mirror 113 (i.e. the sum of the focal length of the fourth and fifth OAP mirrors).

The beam is reflected from the fifth OAP mirror 115 to the sixth OAP mirror 117. Sixth OAP mirror has a focal length of 30 mm and is located 280 mm away from the fifth OAP mirror (i.e. the sum of the focal length of the fifth and sixth OAP mirrors).

The sixth OAP mirror 117 is provided with a hole 119. The visible light beam 55 is passed through this hole to combine it with the THz beam 69 for detection.

Further improvements in spatial resolution may be achieved by inserting condenser cones (made of brass or copper, highly polished on insides, with electro-plating and/or gold/silver evaporated coating) adjacent to the sample to be imaged as shown in Figure 22. In Figure 22, condenser cones 121 and 123 are located on either side of the sample 125 between the sample 125 and plano convex lenses 127 and 129 respectively. The plano convex lenses are the same as those described with reference to

Figure 21. They have a focal length of 10 mm and are placed 10 mm away from condenser cones 121 and 123. The cones have a typical entrance aperture of 2 mm and an exit aperture of between 50 to 100 μm.

The sample 125, is typically placed within a few wavelengths of the exit aperture of condenser cone 121 e.g. about 100 μm, such that near field imaging techniques may be used to realise THz spot sizes at the sample which are less than the diffraction-limited spot size.

Another advantage of this design is that the beam waist size is frequency independent at the aperture entrance, so that all frequencies in the pulse should fit into the condenser cone.

The plano-convex lenses 127, 129 condenser cones 123, 121 and sample 125 are placed between OAP mirrors 131, 133. The mirrors have a focal length of 250 mm. THz beam 53 is reflected from OAP mirror 131 onto plano convex lens 127 which focuses beam 153 onto condenser cone 121. The beam 53 enters through the widest aperture of the condenser cone and exits through the narrowest aperture onto sample 125. Once beam 53 has passed through sample 125 it enters condenser cone 123 and exits the condenser cone 123 through its narrowest aperture onto plano convex lens 129. The beam is then reflected off OAP mirror 133 onto the detection crystal 73. The OAP mirror 133 has a hole 135. Visible light from the generator is combined with the THz beam 69 at mirror 133. It should be noted also that the optical configuration in Figure 7 can also be used with a multiplicity of other mirrors, such as the arrangement in Figure 21. It should be noted, however, that a variety of different focal lengths are possible for mirrors 133 and 131.

Also, the arrangement of condenser cones used here can easily be inserted into the system of Figure 21 using similar guidelines to beam size and mirror placement as those already elucidated.

It should be noted that simpler coupling systems such as that in Figure 23 are possible which utilise only two off-axis parabolic mirrors. These reduce the path length of the beam and therefore minimise losses due to any water vapour or other absorbing gases in the beam path. However, transmission optics are necessary in order to create frequency independent beam waists at the same and/or to realise higher spatial resolution.

In Figure 23, the THz beam 53 is reflected from OAP mirror 141 onto sample 143. The focal length of OAP mirror 141 is 30 mm and sample 143 is placed 30 mm away from OAP mirror 141. Optionally, further optical components such as lenses and condenser cones as described in Figure 22 may be added between mirror 141 and sample 143.

Once beam 53 has passed through sample 143 it is encoded with imaging information and is referred to as beam 69. Beam 69 is reflected from OAP mirror 145 onto the detection crystal. The OAP mirror is provided with a hole 147 which allows the visible beam 55 to be mixed with the THz beam 69 for detection.

Figure 24 shows a further example of an imaging system. The details of these components will not be repeated here. In Figure 24, the delay section of Figure 20 is replaced with a grating pair or an optical fibre 151 which chirps the visible pulse, extending its temporal width from 50 fs to about 20 ps. The different wavelength components in the visible pulse travel through the detection crystal 73 at different times.

Thus, when a grating spectrometer 153 is used to spatially disperse the wavelengths and a CCD camera 155 is used to record the spatial diversion, each pixel in the (for example) X-direction corresponds to a different wavelength and hence a different time. The result is that a given row of pixels in the x-direction on the CCD 155 effectively map out the temporal form of the THz beam which co-propagates through the detector crystal 73 and rotates the polarisation of the visible beam at

different times during the pulse by varying amounts. Thus, transmission through the object being imaged is plotted as a function of time along one direction in the CCD array. Hence, the rotation of the polarisation of the reference beam 55, is measured by crossed polarises 161, 163 which are arranged on either side of the detection crystal 73.

The imaged object may then be stepped in the y direction on the translation stage in the usual way to develop a 2D THz image. Alternatively, if the probe beam is focused down by a cylindrical lens to a line (say 400 μm in x by 10mm in y) on the sample, the THz transmission along the y axis of the sample can be measured by the pixels along y direction of the CCD, i.e. the y-pixels of the CCD may then be used to image the object in the y-direction without resorting to the translation stage moving in y. A full image is then completed by stepping the translation stage in x. Both of these abilities (to measure time delay along the x-axis of the CCD and y image information without mechanical movement) resulting in much quicker acquisition times if sufficient THz power is available as in this intra-cavity design to affect higher signal to noise ratios. Quicker data acquisition and potentially cheaper cost for more compact systems are the result.

The primary advantage of this system is the fast data acquisition owing to the lack of moving parts such as translation stages; using this system, both a) imaging along the y-direction of the object and b) the sampling of the time domain is very fast, limited the creation of a time delay are very fast, limited only by the speed of the CCD camera and the need to average many frames from the camera to get adequate signal to noise rations (SNR) on the images. The latter is the chief mechanism which limits the application of this technique, and hence the realisation of real-time imaging. Poor SNR results in part from the fact that the balanced photodiode detection scheme outlined in Figure 5 can no longer be used because the quarter wave plate would introduce background light onto the CCD as strong as one-half of the total probe power. Small signal detection in this scenario will be overwhelmed by photon shot noise if a CCD camera is used. To reduce the "ambient" light level on the camera, crossed polarizers are used in which the signal on the CCD falls to zero in the absence of a THz electric

field. Such as detection system is optimal for a CCD, but still does not provide as high signal-to-noise as the system in Figure 5, especially if lock-in detection is used in the latter case.

To overcome this SNR problem, regenerative amplifiers are used (not shown) to boost the optical peak power which nonlinearly generates the THz pulse, resulting in a larger THz field. Such a system suffers, however, from numerous disadvantages. Regenerative amplifiers are extremely expensive (~£100K) and tend to be large and bulky. Also, a second pump laser to drive the amplifier is required. Lastly, such systems operate at low repetition rates (50Hz-250kHz), resulting in a relative decrease in average power. The bright intracavity sources designed here would overcome all of these disadvantages. The intracavity design could therefore be a major step forward in the realisation of a THz imaging system with sufficiently quick data acquisition at sufficiently high signal to noise ratios to realise THz images at video frame rates (~38 frames/sec), so-called "THz movies".

Figure 25 shows a further example of the imaging system of Figure 24. In Figure 25, the motorised stage (of Figure 24) is redundant. Instead, the imaging area of the CCD camera 155 is matched to the imaging area of the detection crystal 73. This area is typically several mm^2. The reference beam 55 is expanded by optics 170 (e.g. telescopic or analogous optics), such that the reference beam has a larger cross sectional area than the THz beam and ideally fills all the pixels in the CCD camera 155. The distribution of the rotated polarisation of the reference beam in the x-y plane (proportional to the THz power transmitted through the sample 65 in the x-y plane) is transferred to the pixels of the CCD camera, resulting in a THz image of the object appearing on a CCD or a computer screen (not shown) attached to the output of CCD camera 155. The time delay in this system is created by an optical delay line (as described with reference to Figure 20). This is the only mechanical moving part of the system.

Figure 26 shows an imaging system for breast cancer detection.

The conventional imaging technologies which have been applied to mammography include 1) X-ray camera, X-ray CT imaging, 2) supersonic echo imaging, and 3) MRI. 1) X-ray camera and X-ray CT imaging have the general advantage of high success in the detection of diseased or abnormal tissue or body parts. However, X-ray imaging has a disadvantage of not being able to obtain high contrast and high sensitivity between different soft tissue types and abnormalities in soft tissue. This is particularly important in the breast, a major component of which is fat. Thus abnormalities in the breast, in particular breast cancer, do not have as good constrast in X-Ray mammography as is desired. X-ray camera and X-ray CT imaging also have the historic disadvantage that side-reaction due to ionization may occur in human body given sufficient exposure.

2) In Supersonic Scan, the quality of images is intensively degraded for the fatty tissues in which the velocity of sound is relatively slower (1476 m/sec) than the cases of muscle and liver (1568 m/sec and 1570 m/sec, respectively). 3) In MRI, The image is degraded by heterogeneous resonant signals due to chemical shift of fat (about 3.3 ppm from proton signal of H2O). In the case of breast images, very uniform magnetic field is required with 1 to 2 ppm in order to measure the images by suppressing the signal attributed to fat. However, it is quite difficult to obtain such uniform magnetic field in breast region, because of different magnetizations between air and organism, and because of complicated inside structure of breast.

In THz Mammography

Firstly, a breast part is sandwiched by two THz transparent plates (which made of materials which are transparent to the Terahertz pulse, including high quality of z-cut quartz or semiconductor materials such as high resistivity Si, GaAs, ZnSe and ZnTe or polymers such as polyethylene, polypropylene, PMMA and poly acrironitoryl, TPX) and is pressed to become as thin as possible. THz radiation with a plurality of frequencies in the range from 50 GHz to 84 THz (which was generated from THz-generator (description of THz-generator) is irradiated onto the surface of the pressed breast, where it does not matter whether this occurs from the top or from underneath.

THz radiation penetrating through breast is sensed by a THz detector and recorded electro-optically using free space electro-optic sampling, photoconductive sampling, or other techniques. THz signals are collected in 2D to reproduce a 2D image of the pressed breast in some appropriate level of spatial resolution which will be typically less than 1 or 2 mm. By using one or a combination of the contrast mechanisms noted on pages 11 and 12 above, a 2D image of the breast can be constructed. In addition, other incarnations of the THz mammograph allow THz images based on THz radiation reflected or scattered from the breast surface and inside the breast to be constructed

The utility of the THz mammography stems in part from the fact that soft tissues such as fat are relatively transparent to THz radiation compared to other tissue types. Moreover, fat has a markedly different spectrum in the THz range compared to other tissue types. As noted above, breast contains concentrated fat which prevents the conventional X-ray, Supersonic and MRI imaging technology from distinguishing breast cancer with high and sensitive contrast. THz radiation can easily penetrate through the fatty part. Also, it is well known that the breast cancer frequently causes lime deposits (calcium carbonate is deposited around cancerous region), and THz absorption or reflectance is changed by lime, due to specific absorption and also due to exclusion of water from that region. Also THz frequency domain imaging can easily distinguish lime areas because of existence of its own specific absorption or reflection frequency.

Figure 26 shows an imaging system for breast cancer detection. A THz imaging beam 501 is generated by irradiating a generator crystal 503 with visible light 505 in the same way as previously described. (Many different generation techniques could be used in addition to the non-linear technique. For example, surface field effect or current surge techniques.) The THz radiation is then reflected off axis parabolic mirror 507 to the output port 509. The THZ radiation 501 is then directed towards the sample. The radiation carrying the imaging information is then fed into the detection system via input port 511. The radiation carrying the imaging information 513 is then combined

with a reference beam 515 using an off-axis parabolic mirror 517. The detection mechanism uses the AC pockels effect that is described in detail in reference to Figure 19.

Figure 27 shows a further variation on the system of Figure 26. However, instead of the detection mechanism based on the AC pockels effects, a CCD camera 521 is used. The CCD camera is advantageous because the THz beam does not have to be focused to a fine spot in order to extract the imaging information. Focusing of the THz beam for detection can cause imaging information to be lost under certain circumstances. Another major advantage of the CCD camera is that a larger portion of the object (e.g. breast) can be imaged without moving the stepper motor described below, which leads to a reduction of acquisition time.

Figure 28 is a schematic overview of a breast screening apparatus that can use the generation or detection systems of Figures 26 and 28. The output port 509 and the input port 511 are shown. These ports correspond to the output and input ports of Figures 26 and 27. The imaging radiation 501 is reflected through plate 523 onto sample 525. The radiation passes through base plate 527 and is then reflected to input port 511. The generator 531 and the detector 533 are located within a movable platform 535. Scanning means 537 are provided which can move the platform 535 such that the whole of the sample can be irradiated and hence imaged.

Figure 29 shows a detail of the sample receiving area. As shown in Figure 4 the sample is sandwiched between plates 523 and 527. Plate 523 is movable in relation to plate 527. To image the sample, plate 523 is moved towards plate 527. The sample receiving area is designed for transmission measurements.

Figure 30 shows a further variation of the coupling optics shown in the previous figures. Here, polyethylene or resistivity Si plane or convex lenses 561 and 563 are used to focus the beam onto the sample 525. Details of the types of lenses have been hereinbefore described with reference to Figures 21 and 22.

Figure 31 shows a further variation on the coupling optics. Here, the off-axis parabolic mirrors are positioned so that the beam diameter of the sample is frequency independence. These optics are described in more detail with reference to Figure 21.

Figure 32 shows mammography apparatus where the reflected THz beam 601 is directed into input port 511.

Figure 33 again shows a reflection geometry arrangement for breast screening. The reflected beam is passed through a series of off-axis parabolic mirrors arranged in a configuration similar to that described in relation to Figure 21 before being directed into the input port 511.

Further, the radiation from the output port 509 is also passed through a series of off-axis parabolic mirrors in an arrangement similar to that of Figure 21.

Figure 34 shows a further variation on the system of Figure 26. Here, there are two input ports 611 and 613. The input ports can be used to collect either both transmitted radiation, both reflected radiation or reflected and transmitted information individually.

Figure 35 shows a variation on the system of Figure 34 with two input ports 611 and 613. The radiation for both input ports is fed into separate detection mechanisms, both of which comprise CCD cameras 615 and 617.

Figure 36 shows an imaging system with both transmitted THz 701 and reflected THz 703 are collected.

Figure 37 shows a system for use in THz in the detection of osteoporosis.

Bone may be weakened or rarefied by diseases such as osteoporosis or metastatic cancer. In these cases, there is a change in the density and/or composition of the bone in the effected area, which leads to the bone becoming brittle or weak. Because bone is partially transparent to THz and because both absorption and the refractive index depend on the density and/or composition. Contrast mechanisms such as those previously described on pages 11 and 12 of the specification can be used to both a THz image of the diseases portion and investigate the bone spectroscopically.

Osteoporosis also results in the thinning of the bone.

The osteoporosis apparatus shown in Figure 37 are similar to those of the breast imaging apparatus described in relation to Figures 26 to 36. The sample 701 is placed on a flat stage 703. A polyethylene or other THz transparent material plate 705 is lowered to keep the sample 701 in place. The THz is directed onto the bone and collected in any of the manners described with reference to the breast scanning apparatus (Figures 26 to 36).

Figure 38a shows a visible light image of bones. Figure 38b shows the corresponding THz image of the bone.

Figure 39 shows both visible and THz images taken of a tooth. Appropriate analysis of the time of light of THz light through the tooth allows the enamel layer on the outside of the tooth to be separated from the inner region where the enamel and dentine are both present. Time of flight information can be used to assess the thickness and quality of the enamel and dentine of the tooth. The extra option of the absorption of THz light inside the tooth can be used to identify a cavity inside the tooth. Such images can identify pulp stones and the rate of blood flow into the cavity and inspection through resin-based tooth fillings.

Figure 40 shows THz imaging penetrating through blood. THz in the range (300GHz to 3THz) can penetrate through at least 90 microns of blood, the sample used

was dried blood. However, this technique can also be used to penetrate liquid water. Data for this blood measurement is shown in Figure 40a. Either the time of the frequency domain information can be used to allow identification of increased blood flow in certain areas and establish the certain probability of the presence of a tumour.

CLAIMS:

1. A method of imaging a sample, the method comprising the steps of:

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample which is to be imaged into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies; and

c) generating an image of the area of the sample from the radiation detected in step (b) using a frequency or a selection of frequencies from the plurality of frequencies in the pulsed electro-magnetic radiation.

2 A method according to claim 1, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz

3 A method according to claim 2, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 500 GHz to 10 THz.

4. A method according to any preceding claim, wherein the input radiation is transmitted through the sample.

5. A method according to any of claims 1 to 3, wherein the input radiation is reflected from the sample.

6. A method according to any preceding claim, wherein the image is generated from a single frequency.

7. A method according to any of claims 1 to 5, wherein a single image is generated using a selected frequency range..

8. A method according to claim 7, wherein the selected frequency range is less than a third of the frequency range of the pulsed electro-magnetic radiation which irradiates the sample.

9. A method according to any preceding claim, wherein a plurality of images is generated from a corresponding plurality of different frequencies.

10. A method according to any preceding claims, wherein, in step (b), the radiation is detected in the time domain and is Fourier transformed to obtain frequency dependent data.

11. A method according to any preceding claim, wherein a fraction of the input pulsed radiation does not irradiate the area of the sample, which is to be imaged, and is detected as a reference signal.

12. A method according to any of claims 1 to 10 , wherein the method further comprises the step of detecting a reference signal when the area of the sample to be imaged is absent from the path of the pulsed electro magnetic radiation.

13. A method according to either of claims 11 or 12, wherein step (c) comprises the steps of deriving frequency dependent image data from the detected radiation and plotting the frequency dependent image data for each pixel to obtain an image of the sample.

14. A method according to claim 13, wherein step (c) further comprises the steps of calculating the Power Spectrum for the radiation detected in step (b) and calculating the Power Spectrum of the reference signal.

15. A method according to claim 14, wherein the frequency dependent image data is calculated by subtracting the Power Spectrum of the reference signal from the Power Spectrum of the radiation detected in step (b).

16. A method according to claim 14, wherein the frequency dependent image data is calculated by dividing the Power Spectrum of the radiation detected in step (b) by the Power Spectrum of the reference signal.

17. A method according to claim 13 , wherein the frequency dependent image data is calculated by deriving the frequency dependent absorption coefficient from a Fourier transform of the radiation detected in step (b) and a Fourier transform of the reference signal.

18. A method according to claim 13, wherein the frequency dependent image data is calculated by deriving the frequency dependent refractive index from a Fourier transform of the radiation detected in step (b) and a Fourier transform of the reference signal.

19. A method according to any of claims 13 to 18, wherein the frequency dependent image data is derived for two or more frequencies.

20 A method according to claim 19, wherein the frequency dependent image data for the two frequencies is added, subtracted, multiplied or divided for each pixel.

21. A method according to any preceding claim, wherein the frequency dependent image data is obtained for a narrow range of frequencies and the image data is integrated over this narrow range.

22. A method according to any of claims 13 to 21, wherein the frequency dependent image data is processed by subdividing the data into a plurality of bands, wherein each band represent a range of magnitudes of the image data for a single frequency or narrow frequency range and each band is assigned a single value or colour.

23. A method according to claim 22, wherein the bands are not of equal widths in magnitude.

24. A method according to either of claims 22 or 23, wherein the image data is derived for more than one frequency.

25. A method according to claim 24, wherein the image data for two frequencies is combined according to a predetermined rule.

26. A method according to claim 24, wherein there are two bands assigned the values '0' and '1' and the predetermined rule is a BOOLEAN algebraic rule.

27. A method according to any preceding claim, the method further comprising the step of displaying a sequence of images generated in step (c) for a plurality of different frequencies.

28. A method according to claim 27, wherein the image generated in step (c) is scanable through a continuum of frequencies.

29. A method according to claim 27, wherein the image generated in step (c) can be stepped through a plurality of discrete frequencies.

30. A method of imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies; and

c) generating an image of the area of the sample from the radiation detected in step (b) by calculating the difference between the temporal positions of a maxima and a minima of the radiation detected in step (b) for each pixel.

31 A method according to claim 30, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz.

32. A method according to claim 31, wherein the image is generated by calculating the difference between the temporal positions of the main minimum and maximum peaks.

33. A method according to any of claims 30 to 32, wherein step (c) further comprises the step of multiplying the difference between the temporal position of the maxima and minima, by a magnitude of the radiation.

34. A method according to claim 33, wherein the magnitude of the radiation is the difference in magnitude between the minimum and maxima of the detected radiation.

35. A method according to claim 33, wherein the magnitude of the radiation is the minimum or maximum of the detected radiation.

36. A method of imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies; and

c) generating an image of the area of the sample from the radiation detected in step (b) by calculating the difference between the magnitude of a maxima and a minima of the radiation detected in step (b) for each pixel.

37. A method according to claim 36, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz

38. A method for imaging a sample, according to either of claims 36 or 37, wherein the peak minima and maxima are the main minimum and maximum.

39. A method of imaging according to any preceding claim, configured to image the difference between ordinary and diseased tissue.

40. A method according to claim 39, wherein the difference between diseased tissue and non diseased tissue is determined by the different and/or absorption reflectance and/or dispersion of THz light which is irradiated to organisms on living tissue in-vivo or in-vitro.

41. A method according to any preceding claim, wherein the sample to be imaged is a human or animal.

42. A method according to any preceding claim, wherein the method is configured to distinguish differences in bone density of the imaged sample.

43. A method according to any preceding claim, wherein the method is configured to distinguish between the different types of tissue which comprise human or animal teeth.

44. A method according to any preceding claim, wherein the method is configured to image Breast cancer.

45. An apparatus for imaging a sample, the apparatus comprising:

a) means for irradiating the sample with pulsed electro-magnetic radiation having a plurality of frequencies in the range from 50 GHz to 84 THz;

b) means for subdividing an area of the sample into a two dimensional array of pixels;

c) means for detecting radiation from each pixel over a plurality of frequencies; and

d) means for generating an image of the area of the sample from the detected radiation by using a frequency or a selection of frequencies from the plurality of frequencies in the pulsed electro-magnetic radiation.

46. An apparatus for imaging a sample, the apparatus comprising:

a) means irradiating the sample with pulsed electro-magnetic radiation having a plurality of frequencies in the range form 50 GHz to 84 THz;

b) means for subdividing an area of the sample into a two dimensional array of pixels;

c) means for detecting radiation from each pixel over a plurality of frequencies; and

d) means generating an image of the area of the sample from the detected radiation by calculating the difference between the temporal positions of a maximum and a minimum of the detected radiation detected for each pixel.

47. An apparatus for imaging a sample, the apparatus comprising:

a) means for irradiating the sample with pulsed electro-magnetic radiation having a plurality of frequencies in the range from 50 GHz to 84 THz;

b) means for subdividing an area of the sample into a two dimensional array of pixels;

c) means for detecting radiation from each pixel over a plurality of frequencies; and

d) means for generating an image of the area of the sample from the detected radiation by calculating the difference between the magnitude of a maximum and a minimum of the detected radiation for each pixel.

48 An apparatus according to any of claims 45 to 47, wherein the sample to be imaged is a human or animal breast.

49. An apparatus according to claim 48, comprising means for holding the breast to be imaged.

50. An apparatus according to any of claims 45 to 49, wherein the apparatus further comprises means for detecting the time of flight of a pulse of radiation through the sample.

51. An apparatus according to any of claims 45 to 50 wherein the apparatus comprises means for collecting both transmitted and reflected radiation.

CLAIMS:

1. A method of imaging a sample, the method comprising the steps of:

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample which is to be imaged into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image of the area of the sample from the radiation detected in step (b) using a frequency or a selection of frequencies from the plurality of frequencies in the pulsed electro-magnetic radiation; and

(d) displaying a sequence of images generated in step (c) for a plurality of different frequencies.

2 A method according to claim 1, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz

3 A method according to claim 2, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 500 GHz to 10 THz.

4. A method according to any preceding claim, wherein the input radiation is transmitted through the sample.

5. A method according to any of claims 1 to 3, wherein the input radiation is reflected from the sample.

6. A method according to any preceding claim, wherein the image is generated from a single frequency.

7. A method according to any of claims 1 to 5, wherein a single image is generated using a selected frequency range.

8. A method according to claim 7, wherein the selected frequency range is less than a third of the frequency range of the pulsed electro-magnetic radiation which irradiates the sample.

9. A method according to any preceding claim, wherein a plurality of images is generated from a corresponding plurality of different frequencies.

10. A method according to any preceding claims, wherein, in step (b), the radiation is detected in the time domain and is Fourier transformed to obtain frequency dependent data.

11. A method according to any preceding claim, wherein a fraction of the input pulsed radiation does not irradiate the area of the sample, which is to be imaged, and is detected as a reference signal.

12. A method according to any of claims 1 to 10 , wherein the method further comprises the step of detecting a reference signal when the area of the sample to be imaged is absent from the path of the pulsed electro magnetic radiation.

13. A method according to either of claims 11 or 12, wherein step (c) comprises the steps of deriving frequency dependent image data from the detected radiation and

plotting the frequency dependent image data for each pixel to obtain an image of the sample.

14. A method according to claim 13, wherein step (c) further comprises the steps of calculating the Power Spectrum for the radiation detected in step (b) and calculating the Power Spectrum of the reference signal.

15. A method according to claim 14, wherein the frequency dependent image data is calculated by subtracting the Power Spectrum of the reference signal from the Power Spectrum of the radiation detected in step (b).

16. A method according to claim 14, wherein the frequency dependent image data is calculated by dividing the Power Spectrum of the radiation detected in step (b) by the Power Spectrum of the reference signal.

17. A method according to claim 13 , wherein the frequency dependent image data is calculated by deriving the frequency dependent absorption coefficient from a Fourier transform of the radiation detected in step (b) and a Fourier transform of the reference signal.

18. A method according to claim 13, wherein the frequency dependent image data is calculated by deriving the frequency dependent refractive index from a Fourier transform of the radiation detected in step (b) and a Fourier transform of the reference signal.

19. A method according to any of claims 13 to 18, wherein the frequency dependent image data is derived for two or more frequencies.

20 A method according to claim 19, wherein the frequency dependent image data for the two frequencies is added, subtracted, multiplied or divided for each pixel.

21. A method according to any preceding claim, wherein the frequency dependent image data is obtained for a narrow range of frequencies and the image data is integrated over this narrow range.

22. A method according to any of claims 13 to 21, wherein the frequency dependent image data is processed by subdividing the data into a plurality of bands, wherein each band represent a range of magnitudes of the image data for a single frequency or narrow frequency range and each band is assigned a single value or colour.

23. A method according to claim 22, wherein the bands are not of equal widths in magnitude.

24. A method according to either of claims 22 or 23, wherein the image data is derived for more than one frequency.

25. A method according to claim 24, wherein the image data for two frequencies is combined according to a predetermined rule.

26. A method according to claim 24, wherein there are two bands assigned the values '0' and '1' and the predetermined rule is a BOOLEAN algebraic rule.

27. A method according to any preceding claim, wherein the image generated in step (c) is scanable through a continuum of frequencies.

28. A method according to any of claims 1 to 26, wherein the image generated in step (c) can be stepped through a plurality of discrete frequencies.

29. A method according to any preceding claim, configured to image the difference between ordinary and diseased tissue.

30. A method according to claim 29, wherein the difference between diseased tissue and non diseased tissue is determined by the different and/or absorption reflectance and/or dispersion of THz light which is irradiated to organisms on living tissue in-vivo or in-vitro.

31. A method according to any preceding claim, wherein the sample to be imaged is a human or animal.

32. A method according to any preceding claim, wherein the method is configured to distinguish differences in bone density of the imaged sample.

33. A method according to any preceding claim, wherein the method is configured to distinguish between the different types of tissue which comprise human or animal teeth.

34. A method according to any preceding claim, wherein the method is configured to image Breast cancer.

35. An apparatus for imaging a sample, the apparatus comprising:

a) means for irradiating the sample with pulsed electro-magnetic radiation having a plurality of frequencies in the range from 50 GHz to 84 THz;

b) means for subdividing an area of the sample into a two dimensional array of pixels;

c) means for detecting radiation from each pixel over a plurality of frequencies;

d) means for generating an image of the area of the sample from the detected radiation by using a frequency or a selection of frequencies

from the plurality of frequencies in the pulsed electro-magnetic radiation; and

e) means for displaying a sequence of images generated by the means for generating an image for a plurality of different frequencies.

36. An apparatus according to claim 35, wherein the sample to be imaged is a human or animal breast.

37. An apparatus according to claim 36, comprising means for holding the breast to be imaged.

38. An apparatus according to any of claims 35 to 37, wherein the apparatus further comprises means for detecting the time of flight of a pulse of radiation through the sample.

39. An apparatus according to any of claims 35 to 38 wherein the apparatus comprises means for collecting both transmitted and reflected radiation.

Application No: GB 9913087.4 **Examiner:** Eamonn Quirk
Claims searched: 1-51 **Date of search:** 31 August 1999

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.Q): G1A (AAMA, ACDL, ACDX) H5R(RAT, RTQ)

Int Cl (Ed.6): G01N(1/47, 1/49) (A61B5/00)

Other: Online: WPI, EPODOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
X	EP 0 841 548 A1	(Lucent Technologies) Whole Document especially claim 1; page 1, lines 5-8; page 4, lines 16-20;	1-5,7-9,21 30-32,45, 46, 51.
A	EP 0 828 162 A2	(Lucent Technologies)	
A,&	US 5 710 430	(Lucent Technologies)	
A,&	US 5 623 145	(Lucent Technologies)	

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : G01J 3/00, 3/28, 11/00 — A1

(11) International Publication Number: **WO 00/50859**

(43) International Publication Date: 31 August 2000 (31.08.00)

(21) International Application Number: PCT/GB00/00632

(22) International Filing Date: 23 February 2000 (23.02.00)

(30) Priority Data:

9904166.7	23 February 1999 (23.02.99)	GB
9913087.4	4 June 1999 (04.06.99)	GB

(71) Applicant *(for all designated States except US)*: TOSHIBA RESEARCH EUROPE LIMITED [GB/GB]; 260 Cambridge Science Park, Milton Road, Cambridge, Cambridgeshire CB4 0WE (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: ARNONE, Donald, Dominic [US/GB]; Toshiba Research Europe Limited, Cambridge Research Laboratory, 260 Cambridge Science Park, Milton Road, Cambridge CB4 0WE (GB). CIESLA, Craig, Michael [GB/GB]; Toshiba Research Europe Limited, Cambridge Research Laboratory, 260 Cambridge Science Park, Milton Road, Cambridge CB4 0WE (GB).

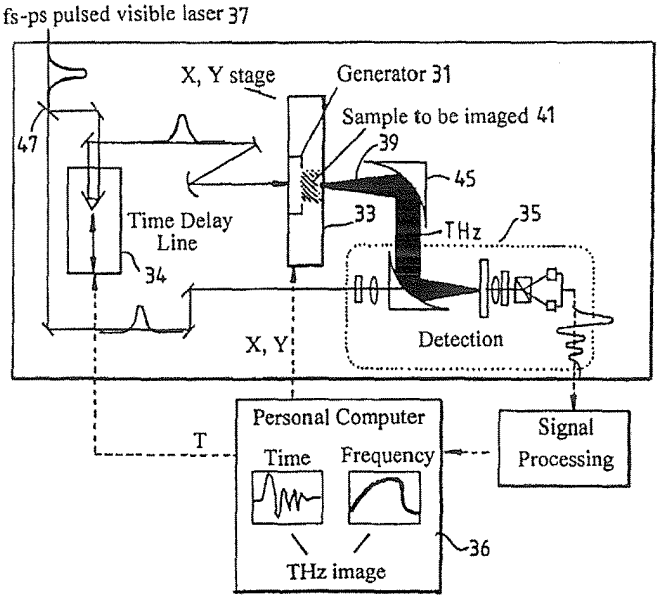
(74) Agent: GRANLEESE, Rhian, Jane; Marks & Clerk, 57–60 Lincoln's Inn Fields, London WC2A 3LS (GB).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: METHOD AND APPARATUS FOR TERAHERTZ IMAGING



(57) Abstract

In an apparatus and method for imaging a sample: a) the sample to be imaged is irradiated with pulsed electro–magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz; b) an area of the sample is subdivided into a two dimensional array of pixels, and radiation from each pixel is detected over a plurality of frequencies; and c) an image is generated from the radiation detected in step (b) preferably using a frequency or a selection of frequencies from the plurality of frequencies in the pulsed electro–magnetic radiation. The method can be used as a medical imaging technique and can be used to image cancer tumours.

IPR2023-00033
Apple EX1002 Page 448

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

METHOD AND APPARATUS FOR TERAHERTZ IMAGING

The present invention relates to the field of imaging samples with radiation in the infra-red (IR) frequency range. More specifically, the present invention relates to apparatus and methods for improving contrast in images obtained using electromagnetic radiation in the higher Gigahertz (GHz) and the Terahertz (THz) frequency ranges. However, in such imaging technology, all such radiation is colloquially referred to as THz radiation, especially that in the range from 50 GHz to 84 THz.

Recently, there has been much interest in using THz radiation to look at a wide variety of samples using a range of methods. THz radiation has been used for both imaging samples and obtaining spectra. Recently, work by Mittleman *et al*, IEEE Journal of Selected Topics in Quantum Electronics, Vol. 2, No. 3, September 1996, page 679 to 692 illustrates the use of using THz radiation to image various objects such as a flame, a leaf, a moulded piece of plastic and semiconductors.

THz radiation penetrates most dry, non metallic and non polar objects like plastics, paper, cardboard and non polar organic substances. Therefore, THz radiation can be used instead of x-rays to look inside boxes, cases etc. THz has lower energy, non-ionising photons than X-rays, hence, the health risks of using THz radiation are expected to be vastly reduced compared to those using conventional X-rays.

The use of THz imaging for medical purposes has also been suggested in the above referenced paper. However, it has previously been believed that strong-water absorption prevents the use of THz in many bio-medical research areas. Previously, a large amount of THz imaging has been analysed on the basis of obtaining contrasts between strongly water absorbing regions and non strongly water absorbing regions.

The present invention addresses the above problems and is concerned with methods for enhancing image contrasts in such THz images to allow THz to be used to not only image contrast based on water absorption but to allow a fine contrast to be shown in THz imaging over a range of different samples.

In a first aspect, the present invention provides a method of imaging a sample, the method comprising the steps of:

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image from the radiation detected in step (b) using a frequency or a selection of frequencies from the plurality of frequencies in the incident pulsed electro-magnetic radiation.

Preferably, the sample is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz; more preferably from 500 GHz to 10 THz.

The detected radiation may be analysed for a single frequency, or, it may be analysed over a selected frequency range. A selected frequency range is taken to be a frequency range which is typically, less than a third of the total frequency range of the pulsed e-m radiation used to irradiate the sample. More preferably, the selected frequency range is less than 10% of the total frequency range of the pulsed e-m radiation used to irradiate the sample.

For example, it has previously been mentioned that water is a strong absorber of THz. There are 'windows' in the water absorption spectra from 50 GHz to 500 GHz, from 30 THz to 45 THz and from 57 THz to 84 THz. If the sample is a irradiated with a range of frequencies from 50 GHz to 84 THz, it may be preferable to generate the image using one or more of the following selected frequency ranges: 50 GHz to 500 GHz, 30 THz to 45 THz and 57 THz to 84 THz. The image may be generated by integrating over the selected frequency range.

Thus, the present invention allows an image to be created from single frequency or from a selected frequency range. Also, the present invention allows a plurality of images to be derived from a plurality of frequencies or allows a single image to be derived from data from two or more distinct frequencies.

The present invention can either be used to image a sample by detecting radiation transmitted through the sample or reflected from the sample.

The image or images generated by the present invention my be displayed in a number of ways. For example, the method of the first aspect of the present invention may further comprise the step of displaying a sequence of images generated in step (c) for a plurality of different frequencies.

The image generated in step (c) may be scanable through a continuum of frequencies. Alternatively, the image generated in step (c) can be stepped through a plurality of discrete frequencies.

For many imaging contrast techniques discussed herein, a reference signal is required. Ideally, the reference signal is obtained from THz radiation which has not been passed through or reflected from the area of the sample which is to be imaged.

The reference signal may be obtained from a fraction of the electro-magnetic pulsed radiation which has not been passed through the sample, alternatively, the

reference signal may be obtained by passing the pulsed electro-magnetic radiation through a different part of the sample. For example, the present invention may be used to image tumours tissue, the area of the sample which is to be imaged would be the tumour, the reference signal could be obtained by passing the radiation through a healthy part of the tissue. The reference signal may also be measured when the sample is absent. For example, the reference signal could be measured before the sample is positioned in the path of the radiation beam or the reference signal could be measured after the sample has been imaged.

In general, the present invention will be performed using imaging apparatus which is configured to detect temporal data at each pixel. Preferably, the data is Fourier transformed to give the complex THz electric field in the frequency domain E (ω).

The image can be obtained in a number of ways from the complex THz electric field E(ω), e.g.:

(i) The power spectrum P_{sample} (ω) of the sample and the power spectrum P_{ref} (ω) of the reference signal may be calculated. The image could then be generated by plotting the difference between the two Power spectrums for a given frequency for each pixel at a selected frequency over integrated over a selected frequency range.

(ii) The power spectrum P_{sample} of the sample and the reference power spectrum P_{ref} may be divided to give the transmittance. The transmittance may then be plotted for each pixel at a selected frequency over integrated over a selected frequency range.

(iii) The frequency dependent absorption coefficient α(ω) may be calculated from the complex electric field E(ω) and plotted for each pixel at a selected frequency over integrated over a selected frequency range.

(iv) The frequency dependent refractive index $\eta(\omega)$ may also be calculated from the complex electric field and plotted for each pixel at a selected frequency over integrated over a selected frequency range.

The detected temporal electric field contains both phase and amplitude information which give a complete description of the complex dielectric constant of the medium in the beam path. The sample to be characterised is inserted into the beam and the shape of the pulses that have propagated through the sample or have been reflected from the sample are compared with the reference temporal profile acquired without the sample. The ratio of the complex electric field $E(\omega)$ and the reference signal $E_{ref}(\omega)$ is calculated to give the complex response function of the sample, $S(\omega)$. In the most simple case, the complex response function is given by:

$$S(\omega) = \frac{E(\omega)}{E_{ref}(\omega)} \propto \exp\left(\frac{i\omega d}{c}\left(\eta(\omega) - 1\right)\right)\exp\left(-\alpha(\omega)d\right) \qquad (1)$$

where d is the sample thickness, c is the velocity of light in vacuum, η is the refractive index and α is the absorption coefficient. The experimental absorption coefficient $\alpha(\omega)$ and the refractive index $\eta(\omega)$ may then be easily extracted from the magnitude $M(\omega)$ and the phase $\phi(\omega)$ of $S(\omega)$, respectively, according to

$$\alpha(\omega) = -\frac{1}{d}\ln\left(M(\omega)\right) \qquad (2)$$

$$\eta(\omega) = 1 + \left(\frac{c}{\omega d}\right)\phi(\omega) \qquad (3)$$

Additional terms may be included in equations (1) to (3) to account for reflections at dielectric interfaces of a sample, thus allowing accurate analysis of multilayered samples.

These parameters are simply related to the complex dielectric function $\varepsilon(\omega)$ of the sample

$$\varepsilon(\omega) = \left(\eta(\omega) + ik(\omega)\right)^2 = \left(\eta(\omega) + i\alpha(\omega)c / 2\omega\right)^2 \qquad (4)$$

The data derived as discussed in (i) to (iv) above, may be directly plotted either as a colour or a grey scale image where the colour or shade of grey of each pixel represents a given magnitude.

Instead of a single frequency, a selected frequency range could be chosen and the result and data of (i) to (iv) integrated over that range. The integrated data could then be plotted.

It may also be preferred to subdivide the magnitude of the data process in accordance with any of (i) to (iv) above into various bands. For example, all data below a certain value could be assigned the value 0, all data in the next magnitude range could be assigned the value 1, etc. These ranges may have equal widths in magnitude or they may have different widths. Different widths may be preferable to enhance contrast e.g. to emphasise contrast in regions of the sample where there is little variation in the sample absorption of THz.

Preferably, the present invention uses two or more frequencies. The data from say two frequencies is processed in accordance with any of (i) to (iv) above. The data is then banded as described for a single frequency above.

The data may be split into two bands, one assigned the value "0" and the other "1". The data from both frequencies can then be added together using a rule such as a Boolean algebraic expression e.g. AND, OR, NOT, NAND, XOR, etc.

Of course, the present invention also allows images to be compared from two different frequencies. This may be particularly useful to identify a substance where the absorption to THz changes over a certain frequency range.

Other methods are also possible for such THz imaging. It is known to plot the maximum or minimum of the electric field. Indeed, this has been done by Mittleman in

the earlier reference paper. However, far greater contrast is achieved by plotting the peak to peak signal, i.e. the distance between the minimum and maxima of the electric field. Therefore, in a second aspect, the present invention provides a method for imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with a pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image from the radiation detected in step (b) by calculating the difference between the magnitude of a maxima and a minima of the radiation detected in step (b) for each pixel.

Preferably, the main peak minima and maxima are chosen i.e. the highest and lowest parts of the temporal trace.

More preferably, the sample is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz;

It is also known to derive temporal information from the data. For example, the time difference between the maxima of the electric field which is passed through the sample and the field which is not passed through or reflected from the area of the sample, which is to imaged, can be plotted. However, there has been no suggestion of plotting the temporal difference between the maximum and the minimum of the electric field. This provides data on the frequency dependent absorption. Therefore, in a third aspect, the present invention provides a method for imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image from the radiation detected in step (b) by calculating the difference between the temporal positions of a maxima and a minima of the radiation detected in step (b) for each pixel.

Preferably, the main peak minima and maxima are chosen i.e. the highest and lowest parts of the temporal trace.

More preferably, the sample is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz;

To obtain even better contrast, the temporal difference between the maximum and the minimum field may also multiplied by a value of the field at that pixel. For example, the minimum, maxima or difference between the minima and maxima.

As mentioned above, the method of the present invention is suitable for medical imaging and is ideally suited to probing certain tissue abnormalities and conditions. It is also useful for imaging items inside containers, for investigating the internal structure and composition of foods, as well as authentication of documents, labels, banknotes or other printed matter and for analogous security applications. Some particular medical applications will now be explained in more detail. First, a description of some examples of the different types of tissue types which can be imaged will follow:

Cancerous tissue:

Normal tissue contains a large number of mature cells of uniform shape and size. Each cell is characterised by a nucleus of uniform size.

Benign neoplasm:

Benign neoplasm involve cellular proliferation of adult or mature cells growing slowly in an orderly manner in a capsule. These tumours do not invade surrounding tissue but may cause harm through pressure on vital structures within an enclosed structure such as the skull.

Malignant cells:

A malignant cell is one in which the basic structure and activity have become arranged in a manner that is unknown and from a cause or causes that are still poorly understood. The malignant cells lose the normal specialised function of the normal cell or may take on new characteristics and functions. A characteristic of malignant cells that can be observed through a microscope is loss of differentiation, or loss of likeness to the original cell (parent tissue) from which the tumour growth originated. This loss of differentiation is called anaplasia, and its extent is a determining factor in the extent of malignancy of the tumour. Anaplasia is one of the most reliable indicators of malignancy. It is seen only in cancers and does not appear in benign neoplasms.

Other characteristics of malignant cells that can be seen through a microscope are the presence of nuclei of various sizes, many of which contain unusually large amounts of chromatin, and the presence of mitotic figures (cells in the provision of division), which denotes rapid and disorderly division of cells. The proportion of cells actively proliferating in malignant tumours is generally greater than that of normal cells. Malignant cells have no enclosing capsule; thus they invade adjacent or surrounding tissue.

Spread of Cancer:

It has been calculated in general that a tumour mass will double in size 30 times before it is 1cm in size, when there is a chance for clinical detection by conventional means. The rate of growth of a malignant neoplasm determines its capacity to spread. Cancer may spread by direct extension, by gravitational metastasis, or by metastatic spread.

Direct extension or invasion of neighbouring tissue produces the typical local effects of ulcerating, bulky, hemorhagic masses or indurative, fibrosing lesions with tissue fixation, distortion of structure, and pitting of the skin seen in some breast cancer. Infection may accompany this spread.

Gravitional metastasis involves erosion of cancer cells in the body cavities and their dropping onto the serous membrane lining in the cavity.

Metastatic spread occurs when the cancer cells invade vascular or lymphatic channels and travel to distant parts of the body where implantation occurs. In metastatic spread, there is almost always a high degree of histologic, cytologic, and functional similarity between the primary cancer and these metastases. Consequently, the type of cell and probable site of the primary tumour can be identified from the morphology of the metastasies. In addition, the mestastes usually mimic the primary tumour in the formation of cell products and secretions.

Anaplasia should result in the modification of the transmission, absorption, and reflection properties of tissue 50GHz-40THz frequency range which can be studied by the method of the present invention. The loss of differentiation and additional ulcerating, bulky, hemorhagic masses or indurative, fibrosing lesions with tissue fixation, distortion of structure will modify properties such as the refractive index, the thickness of the tissue, etc, which can be imaged by a method of the present invention.

Taking the refractive index, for a given thickness, the probability of interaction between radiation and tissue is dependent on the density of the material. The appearance of fibrous material and additional (mutated) cells in a tumorous region suggests that the density of the tissue in this region will be modified relative to that of healthy tissue, leading to a density change. This in turn should modify the refractive index of the medium (either at a fixed frequency ω or over a range of frequencies) relative to that in normal tissue. These changes in the refractive index can be detected by either 1) simply, by shifts in the time domain of a transmitted or reflected THz pulse or 2) least squares fits to frequency domain information, which will yield the refractive index at each frequency $\eta(\omega)$. In the cancerous regions, we expect increases in the refractive index due to increases in the density of cells in that region. However, exclusion of water in the cancer will also modify the refractive index. Due to either effect, the refractive index at different pixels in the image can be used as a contrast mechanism.

Changes in thickness d between abnormal and normal tissue will lead to changes in either

a) the absorption in a given region $a(\omega)d$ or

b) the time delay of the pulse in that region $\eta(\omega)d$.

Both a) and b) can be determined using electric field and time domain spectra, respectively, forming the basis of a contrast mechanism. Because malignant cancer grows rapidly relative to the tissue which it invades, it is likely that a malignant tumour and the normal tissue adjacent to it will differ appreciably in thickness, which can be determined by the methods of the present invention using a) and/or b) above.

For a given thickness, the probability of interaction between radiation and tissue is dependent on the density of the material. The appearance of fibrous material and additional (mutated) cells in a tumourous region suggests that the density of the tissue in this region will be modified relative to that of healthy tissue, leading to a density change between normal and cancerous tissue. The linear attenuation coefficient $a(\omega)$ (usually

quoted in units of cm^{-1}) is dependent on the density of the material. Thus, if a tumour represents a larger amount of tissue in the same volume, the linear attenuation coefficient $a(\omega)$ will increases proportionally to the increase in density. For a given thickness, this change will manifest itself as increased absorption in the tumour relative to the normal tissue.

Convolved with such density-induced changes in $a(\omega)$ will be variations in $\alpha(\omega)$ due to changes in the composition of the tissue itself. Changes in chemical composition are the most obvious reason for this. For example, the presence of ulcerating, bulky, hemorhagic masses in tumours implies a change in chemical composition relative to healthy tissue. In certain types of malignant tumours such as cytologic anaplasia, there is increased or altered nucleic acid synthesis in growing tissue, which implies a change in chemical content. Moreover, simpler changes such as the exclusion or inclusion of additional water in a tumour will change its chemical composition relative to that of healthy tissue.

$\alpha(\omega)$ may also change in the THz range due to increased disorder in the system. Thus, even if there is no change in the type of molecules in tumourous vs normal tissue, the tumour is likely to be more disordered in terms of the arrangements between cells. A similar phenomenon occurs in liquids, where increasing disorder (or lack of crystallisation) leads to increased absorption. It is therefore probable that $\alpha(\omega)$ will be larger in randomly arranged tumour tissue than in more "regularly arranged" tissues.

The methods of the present invention can also be used to detect increased vascularity around tumours and cancerous regions. As noted in the text above, the increased number of blood vessels necessary to feed a tumour can be a first indication of the appearance of cancer. Therefore, detection of increased blood flow or the increased presence of blood in a certain region can be used for the detection of cancer.

In the fourth aspect, the present invention provides a method for detecting cancer, the method comprising the steps of:

a) irradiating the sample with pulsed electromagnetic radiation with a plurality of frequencies in the range from 50GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image of the area of the sample from the radiation detected in step (b).

The imaging step (c) can be generated in a number of ways. For example, it can be generated by plotting the refractive index of the sample for each pixel or it can be generated by plotting the absorption coefficient. Simpler ways to obtain the image would be to plot either the maxima or minima of the electric field or even the time of flight of the radiation through the sample, which would give crucial information about the thickness of the tumour.

The method could also comprise a step of analysing the detected radiation for the presence of lime in the sample. For example, this could be done by obtaining an image of the sample at a frequency where lime is known to strongly absorb. Or, it could be achieved by looking at the individual spectra of the image to isolate the lime fingerprint in the THz region.

The method may also comprise a step of analysing the water content in the sample. This again could be done by looking at the absorption frequencies for water. The image could be stepped through a plurality of different frequencies where water is known to strongly absorb to obtain an accurate result. The method can be used for imaging breast cancer or imaging tumours in human skin. It can also be applied to any type of tumour.

In practice, usually, a reference signal will be taken from a healthy part of the sample in order to compare with the tumourous tissue.

Further aspects of the present invention provide apparatus for effecting the methods, respectively, of the first, second and third aspects of the present invention, each respectively comprising means for carrying-out each of the steps of those methods.

The present invention will now be described by way of example and with reference to the following non-limiting embodiments in which:

Figure 1 shows a typical THz time domain data obtained from imaging a sample;

Figure 2 shows a power spectrum of a THz trace originally measured in the time domain;

Figure 3a is an image of a pork sample at visible light wavelengths and figure 3b is the same pork sample at THz frequencies;

Figure 4 shows three power spectra for different pixels of the image of figure 3b;

Figure 5 shows data for absorption coefficient for dried pork skin and dried pork kidney samples;

Figure 6 shows THz electric field data and time domain and power spectrum frequency domain for different types of chicken tissue;

Figure 7 shows power spectra obtained from the data of Figure 6 plotted on a single axis;

Figure 8 shows a visible image of a spider mounted over a hole in the metal plate;

Figure 9 shows a visible image of the spider in a box with the box open (Figure 9a) and the box closed (Figure 9b);

Figure 10 shows the THz image of the spider in the box of Figure 9b;

Figure 11 shows the THz electric field and power spectrum derived from a blood sample;

Figure 12 shows the THz electric field and power spectra derived from a water sample;

Figure 13a shows a visible image of a commercial package transistor. Figure 13b shows a THz image of Figure 13a and Figure 13c shows a cut away image of the transistor of Figure 13a;

Figure 14a shows the THz image of doped and undoped GaAs. Figure 14b shows the same sample using visible radiation;

Figure 15a shows a monochromatic image of the semiconductor of Figures 13 and 14 and Figure 15b shows a monochromatic image similar to Figure 15 but at a different frequency;

Figure 16 shows schematic apparatus in accordance with an embodiment of the present invention;

Figure 17 is a schematic diagram of a THz imaging system;

Figure 18 shows a variation of imaging system of figure 17;

Figure 19 shows a schematic of a detection section which can be used with the imaging system of Figure 18;

Figure 20 shows the imaging system of Figure 18 with the detection section of Figure 19;

Figure 21 shows a variation on the imaging section of Figure 20;

Figure 22 shows a variation on the imaging section of Figure 20;

Figure 23 shows a variation on the imaging section of Figure 20;

Figure 24 shows a further example of an imaging system; and

Figure 25 shows another example of an imaging system.

Figure 26 shows a THz imaging system configured for detecting breast cancer;

Figure 27 shows a variation on the system of Figure 26;

Figure 28 shows a schematic view of breast screening apparatus using the detection systems of Figures 26 and/or Figure 27;

Figure 29 shows in detail the sample receiving area of the breast screening apparatus of Figure 28;

Figure 30 shows an alternative configuration of the optics of the breast screening apparatus;

Figure 31 shows yet another configuration of the optics of the breast screening apparatus;

Figure 32 shows a variation on the sample receiving area of the breast screening apparatus;

Figure 33 shows a further variation on the sample receiving area and optics of the breast screening apparatus;

Figure 34 shows a further variation on the optics of the breast screening apparatus;

Figure 35 shows yet a further variation on the optics of the breast screening apparatus;

Figure 36 shows a further variation on the optics and sample receiving area of the breast screening apparatus;

Figure 37 shows apparatus for studying bone:

Figure 38 shows visible and THz images of chicken bone;

Figure 39 shows THz and visible images of teeth;

Figure 40 shows THz images of clotted blood;

Figure 41 shows a THz image in the time domain of healthy skin tissue;

Figure 42 shows frequency domain data of a sample of healthy skin;

Figure 43 shows a plot of the transmittance in the frequency domain taken from a sample of healthy skin;

Figure 44 is a plot of absorption against THz frequency for a tumourous region of skin and healthy skin;

Figure 45a shows a visible image of a skin tumour and Figure 45b shows the corresponding image taken using THz radiation;

Figure 46 shows a further THz image of a skin tumour;

Figure 47 shows a plot of absorption coefficient against frequency for two types of moisturisers and water;

Figure 48 shows a plot of the amplitude of a THz pulse in the time domain to illustrate the effect of water on the results; and

Figure 49 shows a plot of the absorption coefficient against percentage of moisturiser by volume in water.

The sample is imaged by irradiating with pulsed electro-magnetic radiation in 50GHz to 84THz. The THz radiation is either passed through the sample or reflected from the sample. Figure 1 shows a typical trace of detected electric field over time detected from the sample for a single pixel. If there is a variation in the sample composition from pixel to pixel, then this will show up in the THz trace for each pixel.

As can be seen from Figure 1, there is a vast amount of information obtained for each pixel therefore, the information needs to be processed to provide a meaningful image. There are four main parameters which can be derived from this time dependent signal. These are: the maxima of the electric field (E_1), the minima of the electric field (E_2), the temporal position of the maxima of the electric field (T_1) and the temporal position of the minima of the electric field (T_2).

IPR2023-00033
Apple EX1002 Page 467

To obtain a THz image, one or more of these parameters may be plotted directly. However, further contrast may be obtained by plotting the following functions:

1. The peak to peak height i.e. E_1 and E_2;

2. The difference in the temporal spacing between a position of the maxima E_1 and the position of the minima E_2 i.e. T_1-T_2. This parameter is particularly useful as it gives information on the frequency dependent absorption of the sample.

3. The product of the temporal peak position with one or more electric field parameters. For example:

$$(T_2-T_1) \times E_1$$

$$(T_2-T_1) \times E_2$$

$$T_2-T_1 \times (E_1+E_2)$$

Further information may also be obtained from Figure 1 by performing a frequency analysis. Figure 2 shows a power spectrum for a typical trace. The power spectrum is obtained in the conventional way from Fourier transforming the temporal data (i.e. the type of data shown in Figure 1). The magnitude of the signal is plotted logarithmically on the y-axis and the frequency is plotted on the x-axis. The data for Figure 2 was derived from a sample which had regions of fat, kidney and meat. This sample will be described later in more detail with reference to Figure 3. A reference signal is also shown which is the power spectrum of radiation which has not passed through the sample. By viewing the data in this form, the spacing between the data (along the y axis) for the three tissue types is seen to vary with frequency. Therefore, by choosing an appropriate frequency (f_{select}) an image with strong contrast may be obtained.

Also, to distinguish certain tissue types it is often advantageous to plot the image at two separate frequencies for comparison. This is particularly useful when there is

little contrast between the tissue types at various frequencies. For example, taking the data of Figure 2, it is seen that the power spectrum corresponding to kidney tissue and fat tissue has an almost identical value at 1 THz. Therefore, in an image generated at this frequency, both fat and kidney would look very similar. However, looking at the trace at around 2.5 THz, it is seen that the kidney data has a similar value to that of the meat data. Therefore, in an image taken at 2.5 THz there is very little contrast between the meat and the kidney data. However, comparing the image taken at 1 THZ with the image taken at 2.5 THz, allows the three tissue types to be equally distinguished. Two or more images at different frequencies can also be compared by a mathematical operation. For example, the data taken at 2.5 THz could be subtracted from the data taken at 1 THZ to clearly show the characteristic signal due to absorption by kidney tissue. Other mathematical operations are also possible such as multiplication, addition, and division etc.

To further enhance contrast, the data can be subdivided into a series of bands based on magnitude. In Figure 2, four bands are shown. Each of these bands can be assigned a single numerical value and this data can then be plotted. In the data of Figure 2, it can be seen by assigning a single value to band 2 and a single integer value to band 2 and a single integer value to band 3 where substantially enhanced contrasts between the fat tissue and the kidney tissue. This analysis can further be extended to just two amplitude bands. In one amplitude band, the amplitude is assigned a value of zero and in the second amplitude band the amplitude is assigned a value of one. This analysis allows data at two different frequencies, f_1 and f_2 to be compared using certain rules. Such a rule could be:

Signal (f_1)	Signal (f_2)	New Value
0	0	1
1	0	2
0	1	3
1	1	4

Also, certain Boolean algebraic expressions could be used, for example, AND, NAND, XOR, OR, NOR. For an AND comparison, the following table would be used:

Signal (f_1)	Signal (f_2)	New Value
0	0	0
1	0	0
0	1	0
1	1	1

Instead of the power spectrum, the frequency dependent absorption coefficient $a(\omega)$ or the frequency dependent refractive index $\eta(\omega)$ could be analysed in the same way as the power spectrum.

Figures 3a and 3b show images obtained from a pork sample.

The pork sample consists of a variety of different pork tissue types, meat, fat and kidney. Each tissue type is approximately 1mm thick, and the area measured using THz is approximately $13x10mm^2$ with a 500μm step size. The temporal scan at each pixel was 10ps long, giving a frequency resolution of 100GHz, and each pixel took approximately 1 minute to measure. The sample was mounted on cellulose nitrate film and a 1mm thick polythene window, before being placed behind a THz radiation source (<110>ZnTe; 1mm thick, $20x25mm^2$). Details of the apparatus used to obtain the images will be described later with reference to Figures 18 to 26.

At each pixel, the tissue type was identified using both the change in the magnitude and temporal position of the electric field maximum. All tissue types were found to transmit THz pulses, with the absorption increasing in the order: fat, meat, and kidney. The THz image shown in Figure 3(b) is a panchromatic THz image since the analysis of the absorption was not frequency specific.

Comparing the visible and THz images, the different tissue types have been successfully imaged and identified at THz frequencies. Due to the simplistic nature of this analysis there is an increased error along the boundary of each tissue region due to some scattering, which results in tissue types being incorrectly identified in these regions. Figure 4 shows power spectra for each tissue type . The data was taken at 3 different pixels. Clear differences in the absorption characteristics for each tissue type are shown.

Figure 5 shows data of the absorption coefficient for skin and kidney using the attenuation of the peak of the THz electric field. All samples were dried to minimise water absorption effects in the data. The data illustrates that the absorption coefficients for the two tissue types are different, which may be used to identify the tissues in an image; the thickness can be found from the temporal shift. Thus, the maximum tissue thickness that can be measured assuming a signal to noise ration (SNR) of 10^4:1. For this example, these values are 3.0mm for kidney and 7.5mm for skin. With the addition of an amplifier to boost visible, and hence THz, powers 10^5:1 - 10^6:1 can be achieved, which allows several cm of tissue to be probed in theory.

Figure 6 shows data from six types of chicken tissue: bone, cartilage, skin (cooked and uncooked), meat (cooked and uncooked). All tissue types were found to transmit THz by varying amounts. The meat has the strongest absorption of all the tissue types, with the electric field results multiplied by five for clarity. The frequency domain data (NH5) shows a clear difference in the nature of the absorption for each type of tissue.

Figure 7 shows the data from Figure 6 plotted on a single axis for comparison.

THz can be used to image the contents of objects that are otherwise opaque at particularly visible wavelengths, much in the same way as x-rays are currently used. However, due to the non-ionising nature of THz plus the low average powers, THz is inherently safe. For example, common packaging materials such as plastics, paper and

cardboard have been found to be transparent at THz frequencies. Figure 8 shows a visible image of a spider mounted over a hole in a metal plate. The spider is placed in a plastic box for imaging as shown in Figure 9a. At visible wavelengths it was not possible to observe the contents of the plastic box with the lid closed. Figure 9b.

The spider and the box were imaged over a 10x10 mm region with a 200μm step size. Analysing the image using the change in peak of the THz electric field, a THz image was formed of the contents of the box, Figure 10. The light areas correspond to regions of highest THz transmission. The spider is clearly resolved inside the box, with fine detail, such as a break in one of legs (top right corner) resolvable. Closer inspection of the image shown in Figure 10 also revealed some weak internal structure inside the spider.

Of considerable interest to medical applications would be the ability to measure the properties of blood using a non-invasive method on a human subject, i.e. a "contactless" in vivo method of assessing the constituents of the subject's blood. Thus, a technique that combines imaging and spectroscopic capabilities to locate and identify, respectively, different tissue types and their constituents would be of enormous medical and commercial value. Also, for imaging purposes blood is (partially) transmitting at THz frequencies. A sample of dried blood (black pudding) was taken and a thin layer spread on a piece of cellulose nitrate film. The transmission was measured at a single point on the sample, Figure 11a, with the peak electric field decreasing by 55%. In the frequency domain, Figure 11b, the absorption is found to be strongest between 1.00 THz and 1.90 THz, with a strong absorption around 0.66 THz.

One of the most important materials that must be considered for biological applications of THz is liquid water. Water is known to have a number of strong absorption bands in the infrared and far-infrared/THz regions due to the polar nature of the molecule, and ultimately it may be absorption due to liquid water in biological samples that limits the applications of THz technology. To quantify the strength of water absorption a thin (90μm) layer of liquid water was measured. This layer was

formed on the surface of a piece of cellulose nitrate film. The time-domain results for the sample and a reference (dry cellulose nitrate) are shown in Figure 12a, and the frequency domain results in Figure 12b. From the time-domain results alone the absorption is clearly stronger than in the blood sample. Considering the frequency domain results, the absorption is found to increase with increasing frequency; at frequencies above 2.5 THz the repeatability of the spectra is less reliable and therefore only the frequency range 0.6 THz to 2.5 THz should be examined. It is therefore important for biological applications that the frequency range of THz spectroscopy is extended to higher and lower frequencies where absorption windows are present. It should be noted, however, that the ability of THz to probe the skin, blood, and water thicknesses noted above suggests that it is possible to use the present invention.

In Figures 14 a to c, a commercially available, discrete transistor was taken which has the active elements encapsulated in a black casing that is opaque at visible wavelengths, Figure 13a. Fortunately, because the black casing is partially transparent to THz, the internal structure of the device can be imaged. The THz image in Figure 13b was constructed using 150μm steps over a 5x6mm^2 region centred on the transistor. Where there is only plastic, THz pulses are transmitted, but where there are metal tracks of the three leads, no THz is transmitted. To confirm the accuracy of the THz image, an identical transistor was cut open and the same lead arrangement was found, as recorded in the visible photograph of Figure 13c. Additional information on the depth of the leads below the outer surface could be obtained using reflection-based imaging. Scanning the THz beam across the front of the sample, strong reflections were recorded in regions with the metal tracks, with the temporal delay of the pulses giving the distance of the tracks below the surface.

A simple image of the distribution of the doping in a semiconductor can be formed using the change in the peak of a transmitted THz pulse, due to the approximately linear increase of THz free-carrier absorption with carrier density. Figure 14b shows a visible image of a sample consisting of two strips of undoped GaAs and two strips of n-type ($n \sim 1 \times 10^{18}$ cm^{-3}) with the strips arranged alternately side by side.

Each strip was 10x1x0.5mm. An area of approximately $8x8mm^2$ was imaged using THz, with a 200μm step size. While at visible wavelengths there was no means of identifying the GaAs, the THz image clearly shows a difference in the absorption. Figure 14a. Here the peak of the THz field is plotted for each pixel. The highest transmission (at the bottom of the image) is where there was no GaAs present. Moving across the GaAs strips there is variations in the absorption due to the different dopings. with absorption strongest in the n-type strips. Strip #3 also shows a step in the absorption at the boundary of the etch through the n-type doping.

Figures 15a and 15b show frequency dependent (monochromatic) images derived from the THz data of the semiconductor sample. Instead of plotting the peak of the field, power spectrum for each pixel is found using a fast Fourier transform and the power at a given frequency selected. Figures 15a and 15b show the images for f=2.38Thz and f=890Ghz respectively. At 2.75 THz the image is similar to Figure 14a. with the difference in the absorption between the two type of GaAs clearly resolvable. However, for the 890GHz image there is no difference in the transmission for the undoped and areas with no sample, with both appearing as the same shade of grey indicating that no absorption in the GaAs at this frequency. Thus showing that sequentially viewing the images at a plurality of different frequencies can be used to determine information about the sample.

Figure 16 shows a schematic apparatus which can be used to look at a THz image for a plurality of different frequencies. The apparatus, in its most basic form involves a module 1001 having a screen 1003 and a knob 1005. The knob 1005 is a frequency selector and controls the frequency of the image to be shown on screen 1003.

The detection system will detect a plurality of different frequencies. The image which is to be displayed will be dependent on the setting of frequency selector 1005. In this simplified arrangement, frequency selector 1005 has three settings, A, B, and C. The sample has three concentric rings. The composition of each of the rings is different so that each ring will strongly absorb (become shaded) at a particular frequency.

At frequency A, the inner ring is strongly absorbing. However, the two outer rings do not absorb. Therefore, looking at the sample at just frequency A provides virtually no information about the outer two rings. Switching to frequency B using frequency selector knob 1005 changes the image so that the middle ring is now strongly absorbing. However, the inner ring and the outermost ring do not strongly absorb at this frequency. Comparing images A and B allows information about the whole structure of the sample to be determined. This is far more useful than just looking at image A or just image B. It is also less complicated than looking at a full panchromatic image of the structure.

When the frequency selector knob 1005 is turned to C, the frequency is selected so that the outer ring strongly absorbs and the inner two rings do not. The situation is similar to Figure A in that no real information can be established about differences between the two inner rings.

Figure 17 shows a basic THz imaging system. The system can be simplified into three main sections, a generator 31, an imaging section 33 and a detection section 35. THz radiation is generated in the generating section 31 by using a THz emitter which is supplied by a visible pulsed laser 37.

A THz beam 39 is emitted from generation section 31 and is directed onto sample 41 of the imaging section 33. The THz beam 39 is then directed via further optics 45 into the detection section 35. The system of Figure 17 is an example of a near field imaging system, where the sample to be imaged is placed immediately behind the THz source.

The detection section reads the information carried in the detected THz signal via a visible light signal and AC Pockels effect. The visible light is ideally obtained from laser 37 via beam splitter 47. A time delay is added to the THz pulse via time delay line 34. The system (e.g. the control of the sample 41 movement, the time delay 34 and the detected signal processing) is controlled by computer 36.

Details of the AC pockels effect will be described with reference to Figure 19.

Figure 18 shows another imaging system. Here, for simplicity, details of the detection part of the system will be omitted. These will be described with reference to Figure 19.

The THz generation section is indicated by the components within box 51. The imaging system requires both a visible light pulse and a THz pulse to be emitted from the generation section 51. Therefore, the output coupler 23 should not 100% reflective to visible radiation to allow some visible radiation to be emitted from the THz generation section 51.

The emitted THz beam 53 and visible 55 from the generation system are incident on beam splitter 57. This beam splitter 57 allows transmission of the THz beam 53 but reflects the visible light beam 55 onto mirror 59 which reflects the beam 55 into optical delay line 61. The delayed beam 55 is then inputted into the THz detection unit 63.

The THz beam 53 is directed into the imaging section 52 and onto sample 65 via THz imaging optics 67. The sample 65 is located on a motorised X-Y translation stage (not shown) so that the whole sample 65 can be imaged. (The x-y plane is orthogonal to the beam axis). The THz radiation 69 carrying the imaging information from the sample is reflected into the THz detection system 63 via THz imaging optics 71.

The presence of visible radiation 55 as well as THz radiation 69 allows for imaging and electro-optic detection to be performed inside a single nitrogen-purged unit.

The sample 65 is mounted on a X-Y motorised translation stage (not shown) which is controlled by a PC computer (not shown). Each section (pixel) of the object may then be imaged. To improve the spatial resolution of the technique, off-axis

parabolic mirrors, condenser cones, and lenses may be used to focus the beam to a diffraction limit spot. By mounting the sample in the near field of a condenser cone, the diffraction limit may be overcome and spatial resolution of about 50 μm may be achieved. The imaging system can function with or without such objects depending on the nature of the object to be imaged and the nature of the detection circuit.

Figure 19 shows the detection system in detail. The THz beam 69 carrying the imaging information and a visible light beam 55 are inputted into the THz detection system. The visible light beam 55 acts as a reference beam which is incident on the detection crystal 73. The reference beam 55 is linearly polarised and the polarisation is orientated such that it has components along both the ordinary and extraordinary axis of the detection crystal 73. Each of the axes has distinct refractive indices n_o and n_e along the ordinary and extraordinary axis of the crystal 73 respectively. In the absence of a second (THz) radiation beam 69, the linearly polarised reference beam 55 passes through the detection crystal 73 with negligible change in its polarisation.

The applicant wishes to clarify that the angle Θ through which the polarisation is rotated by is negligible. However, the linearly polarised beam can become slightly elliptical. This effect is compensated for by a variable retardation waveplate, e.g. a quarter waveplate 81.

The emitted beam 77 is converted into a circularly polarised beam 83 using quarter wave plate 81. This is then split into two linearly polarised beams by a Wollaston Prism 79 (or equivalent device for separating orthogonal polarisation components) which directs the two orthogonal components of the polarised beam onto a balanced photodiode 85. The balanced photodiode signal is adjusted using wave plate 81 such that the difference in outputs between the two diodes is zero.

However, if the detector 73 also detects a secondary beam 69 (in this case a beam with a frequency in the THz range) as well as a reference beam, the angle Θ through which the polarisation is rotated by is not negligible. This is because the THz

electric field modifies the refractive index of the visible (fundamental) radiation along one of the axes n_e, n_o. This results in the visible field after the detector 73 being elliptical and hence the polarisation components separated by the prism 79 are not equal. The difference in the voltage between the output diodes gives a detection voltage.

The reference beam 55 and the THz beam 69 should stay in phase as they pass through the crystal 73. Otherwise the polarisation rotation Θ is obscured. Therefore, the detection crystal 73 has phase matching means to produce a clear signal.

All of the items shown in Figure 20 sit on an optical bread board of dimensions 36 inches by 36 inches. The only external units required are a power supply for the diode laser and a cooling unit for the generation section 51.

The imaging section 91, has a motorised stage which is movable in the x-y plane, i.e. along two orthogonal axis which are perpendicular to the incident beam of THz radiation.

The imaging section 91 has two mirrors M11 and M12. Mirror M12 directs the THz beam 53 onto the sample 65. Mirror M11 is positioned to reflect the THz radiation transmitted through sample 65 onto the detection crystal 73. Mirrors M11 and M12 are off axis parabolic (OAP) mirrors. Such mirrors are configured so that the phase difference between the incident and reflected beams is the same at all points on the mirror. The parameters resulting in an off axis parabolic surface are characterised by the focal length of the mirror.

An optical delay section 93 is also shown. The visible light beam emitted from the generating section is reflected by beamsplitter 57 into delay section 93. The delay section 93 has a corner cube mirror M9 which is moveable along the beam axis. The beam is directed onto corner cube mirror M9 via mirror M8. The beam is reflected off corner cube mirror M9 onto mirror M10. Corner cube mirror M9 is oscillated back and forth along the beam direction with an oscillation frequency of several 10s of Hz. This

increases or decreases the path length of the visible beam 55 as required. A Clark ODL-150 system may be used to drive the mirror, this is capable of delays of 150ps. The emitted beam is then combined with the emitted THz beam at mirror M11. Alternatively the THz and visible beams may be combined colinearly using a beam splitter, for example, a pellicle beam splitter. Such a device would be placed before or after M11 and would eliminate the requirement for a hole in M11.

Figure 21 shows a variation on the imaging section 91 of Figure 20. The extended path length over which the THz beam 53 travels is purged with nitrogen to remove water vapour and hence improve the quality of the image.

Due to diffraction effects associated with the large wavelengths in the THz range, the cross-sectional size of the THz beam 53 and imaging applications is not sufficiently large that it may be treated as plain parallel. If diffraction effects are such that the radiation is paraxial so that it can be represented by a scalar field distribution. Gaussian beam mode optics and optical techniques can be used. The simplest case for system design is to assume that the fundamental mode dominates the beam profile. The use of Gaussian mode optics and design applied to conventional THz radiation and systems (generated in the Fourier transform machines, far-infrared lasers or Gunn diodes) is applicable and important to THz imaging systems.

A number of design rules or guidelines should be followed when constructing a THz imaging system to obtain a good quality image. For transmission optics such as lenses, geometric losses are kept to a minimum by ensuring that the ratio of the lens thickness to focal length and diameter to focal length is less than 0.2. If this is satisfied, then losses in the lenses will be primarily due to absorption and reflection. In this case, choice of materials is important.

A requirement which arises in pulsed systems is the need for the material to be non-dispersive so that pulse broadening does not occur. Given these requirements, high density polyethylene (DHPE), polytetrafluorethylene (PTFE), high resistivity silicon

(Si), and TPX are some of the best materials and can also be machined in a lathe; any material combining low absorption and low dispersion at THz frequencies is a good candidate for fabrication of transmission optics, provided its shape can be suitably fabricated for a lens. Reflection losses in lenses tend to be highly frequency dependent at THz frequencies. Therefore care must be taken in lens design to ensure that all frequencies across the pulse bandwidth undergo the same reflection (and absorption) losses.

Ideally, reflective optics (mirrors) are used wherever possible instead of transmission optics (lenses) in order to minimise a number of losses associated with transmission optics, which include (i) frequency-dependent reflection losses and amplitude pattern distortion at dielectric (e.g. air-lens) interfaces, (ii) frequency-dependent absorption losses, (iii) diffraction effects and distortions to field distribution due to power falling on the lens's surface at an angle.

An additional property of importance in imaging (and not particular to Gaussian mode beam optics) is that if two mirrors are separated by the sum of their focal lengths, then the size of the beam waist (minimum beam diameter in plane normal to optical axis) on the optical axis after the reflection from the second mirror will be frequency-independent. This is true of the last mirror (focusing element) in a chain provided there are an even number of focusing elements in the chain. This provides a major advantage for THz imaging as the pulse is comprised of a wide range of frequency components, and it is desired to keep the object at a fixed position on the optical axis whilst images are being recorded at various (x, y) points and at all THz frequencies in the pulse. This is particularly important THz imaging as the spectral coverage (bandwidth) of THz pulses increases into the mid-infrared and even higher frequencies.

The system in Figure 21 will produce beams with 1/e diameters (for the fundamental Gaussian mode in the beam) of 1-2 mm at the sample in the THz frequency range (e.g. at 300GHz, diameter = 2mm). In the system of Figure 21, six mirrors and two lenses are used as opposed to the two mirrors of Figure 20. The direction of the

beam in Figure 21 is reversed to that in Figure 20. In the imaging section, the beam is first reflected off first OAP mirror 101 onto second OAP mirror 103 and then onto third OAP mirror 105. Second 103 and third OAP mirrors each have a focal length of 250 nm. They are separated by 500 nm.

The beam is reflected from the third OAP mirror 105 onto plano-convex lens 107 which has a focal length of 10 mm and a diameter of 10 mm. Third OAP mirror 105 is separated from plano convex lens 107 by 260nm (i.e. the sum of their focal lengths). The lens 107 is made from polyethylene or high resistivity Si. The lens 107 is placed 10 mm from the motorised stage (not shown) on which sample 109 is mounted. The beam has traversed through an even number of focusing optics and mirrors (101, 103, 105 and 107) which are all spaced apart by the sum of their focal lengths. Hence, the waist of the beam at the sample is independent of the frequency. Here, the beam diameter is chosen to be 2 mm, independent of frequency in the frequency range of about 300 GHz (0.30THz).

Once the beam has passed through sample 65, the transmitted THz radiation falls onto second plano convex lens 111. Plano convex lenses 107 and 111 are identical in optical characteristics. Lens 111 focuses the THz radiation onto the fourth OAP mirror 113. Fourth OAP mirror 113 has a focal length of 250 mm and reflects the THz beam onto fifth OAP mirror 115. Fifth OAP mirror 115 also has a focal length of 250 mm and lies 500 nm away from the fourth OAP mirror 113 (i.e. the sum of the focal length of the fourth and fifth OAP mirrors).

The beam is reflected from the fifth OAP mirror 115 to the sixth OAP mirror 117. Sixth OAP mirror has a focal length of 30 mm and is located 280 mm away from the fifth OAP mirror (i.e. the sum of the focal length of the fifth and sixth OAP mirrors).

The sixth OAP mirror 117 is provided with a hole 119. The visible light beam 55 is passed through this hole to combine it with the THz beam 69 for detection.

Further improvements in spatial resolution may be achieved by inserting condenser cones (made of brass or copper, highly polished on insides, with electro-plating and/or gold/silver evaporated coating) adjacent to the sample to be imaged as shown in Figure 22. In Figure 22, condenser cones 121 and 123 are located on either side of the sample 125 between the sample 125 and plano convex lenses 127 and 129 respectively. The plano convex lenses are the same as those described with reference to Figure 21. They have a focal length of 10 mm and are placed 10 mm away from condenser cones 121 and 123. The cones have a typical entrance aperture of 2 mm and an exit aperture of between 50 to 100 μm.

The sample 125, is typically placed within a few wavelengths of the exit aperture of condenser cone 121 e.g. about 100 μm, such that near field imaging techniques may be used to realise THz spot sizes at the sample which are less than the diffraction-limited spot size.

Another advantage of this design is that the beam waist size is frequency independent at the aperture entrance, so that all frequencies in the pulse should fit into the condenser cone.

The plano-convex lenses 127, 129 condenser cones 123, 121 and sample 125 are placed between OAP mirrors 131, 133. The mirrors have a focal length of 250 mm. THz beam 53 is reflected from OAP mirror 131 onto plano convex lens 127 which focuses beam 153 onto condenser cone 121. The beam 53 enters through the widest aperture of the condenser cone and exits through the narrowest aperture onto sample 125. Once beam 53 has passed through sample 125 it enters condenser cone 123 and exits the condenser cone 123 through its narrowest aperture onto plano convex lens 129. The beam is then reflected off OAP mirror 133 onto the detection crystal 73. The OAP mirror 133 has a hole 135. Visible light from the generator is combined with the THz beam 69 at mirror 133. It should be noted also that the optical configuration in Figure 7 can also be used with a multiplicity of other mirrors, such as the arrangement in Figure

21. It should be noted, however, that a variety of different focal lengths are possible for mirrors 133 and 131.

Also, the arrangement of condenser cones used here can easily be inserted into the system of Figure 21 using similar guidelines to beam size and mirror placement as those already elucidated.

It should be noted that simpler coupling systems such as that in Figure 23 are possible which utilise only two off-axis parabolic mirrors. These reduce the path length of the beam and therefore minimise losses due to any water vapour or other absorbing gases in the beam path. However, transmission optics are necessary in order to create frequency independent beam waists at the same and/or to realise higher spatial resolution.

In Figure 23, the THz beam 53 is reflected from OAP mirror 141 onto sample 143. The focal length of OAP mirror 141 is 30 mm and sample 143 is placed 30 mm away from OAP mirror 141. Optionally, further optical components such as lenses and condenser cones as described in Figure 22 may be added between mirror 141 and sample 143.

Once beam 53 has passed through sample 143 it is encoded with imaging information and is referred to as beam 69. Beam 69 is reflected from OAP mirror 145 onto the detection crystal. The OAP mirror is provided with a hole 147 which allows the visible beam 55 to be mixed with the THz beam 69 for detection.

Figure 24 shows a further example of an imaging system. The details of these components will not be repeated here. In Figure 24, the delay section of Figure 20 is replaced with a grating pair or an optical fibre 151 which chirps the visible pulse, extending its temporal width from 50 fs to about 20 ps. The different wavelength components in the visible pulse travel through the detection crystal 73 at different times.

Thus, when a grating spectrometer 153 is used to spatially disperse the wavelengths and a CCD camera 155 is used to record the spatial diversion. each pixel in the (for example) X-direction corresponds to a different wavelength and hence a different time. The result is that a given row of pixels in the x-direction on the CCD 155 effectively map out the temporal form of the THz beam which co-propagates through the detector crystal 73 and rotates the polarisation of the visible beam at different times during the pulse by varying amounts. Thus, transmission through the object being imaged is plotted as a function of time along one direction in the CCD array. Hence. the rotation of the polarisation of the reference beam 55. is measured by crossed polarises 161, 163 which are arranged on either side of the detection crystal 73.

The imaged object may then be stepped in the y direction on the translation stage in the usual way to develop a 2D THz image. Alternatively. if the probe beam is focused down by a cylindrical lens to a line (say 400 μm in x by 10mm in y) on the sample, the THz transmission along the y axis of the sample can be measured by the pixels along y direction of the CCD, i.e. the y-pixels of the CCD may then be used to image the object in the y-direction without resorting to the translation stage moving in y. A full image is then completed by stepping the translation stage in x. Both of these abilities (to measure time delay along the x-axis of the CCD and y image information without mechanical movement) resulting in much quicker acquisition times if sufficient THz power is available as in this intra-cavity design to affect higher signal to noise ratios. Quicker data acquisition and potentially cheaper cost for more compact systems are the result.

The primary advantage of this system is the fast data acquisition owing to the lack of moving parts such as translation stages; using this system. both a) imaging along the y-direction of the object and b) the sampling of the time domain is very fast, limited the creation of a time delay are very fast, limited only by the speed of the CCD camera and the need to average many frames from the camera to get adequate signal to noise rations (SNR) on the images. The latter is the chief mechanism which limits the application of this technique, and hence the realisation of real-time imaging. Poor SNR

results in part from the fact that the balanced photodiode detection scheme outlined in Figure 5 can no longer be used because the quarter wave plate would introduce background light onto the CCD as strong as one-half of the total probe power. Small signal detection in this scenario will be overwhelmed by photon shot noise if a CCD camera is used. To reduce the "ambient" light level on the camera, crossed polarizers are used in which the signal on the CCD falls to zero in the absence of a THz electric field. Such as detection system is optimal for a CCD, but still does not provide as high signal-to-noise as the system in Figure 5, especially if lock-in detection is used in the latter case.

To overcome this SNR problem, regenerative amplifiers are used (not shown) to boost the optical peak power which non-linearly generates the THz pulse, resulting in a larger THz field. Such a system suffers, however, from numerous disadvantages. Regenerative amplifiers are extremely expensive (~£100K) and tend to be large and bulky. Also, a second pump laser to drive the amplifier is required. Lastly, such systems operate at low repetition rates (50Hz-250kHz), resulting in a relative decrease in average power. The bright intracavity sources designed here would overcome all of these disadvantages. The intracavity design could therefore be a major step forward in the realisation of a THz imaging system with sufficiently quick data acquisition at sufficiently high signal to noise ratios to realise THz images at video frame rates (~38 frames/sec), so-called "THz movies".

Figure 25 shows a further example of the imaging system of Figure 24. In Figure 25, the motorised stage (of Figure 24) is redundant. Instead, the imaging area of the CCD camera 155 is matched to the imaging area of the detection crystal 73. This area is typically several mm^2. The reference beam 55 is expanded by optics 170 (e.g. telescopic or analogous optics), such that the reference beam has a larger cross sectional area than the THz beam and ideally fills all the pixels in the CCD camera 155. The distribution of the rotated polarisation of the reference beam in the x-y plane (proportional to the THz power transmitted through the sample 65 in the x-y plane) is transferred to the pixels of the CCD camera, resulting in a THz image of the object

appearing on a CCD or a computer screen (not shown) attached to the output of CCD camera 155. The time delay in this system is created by an optical delay line (as described with reference to Figure 20). This is the only mechanical moving part of the system.

Figure 26 shows an imaging system for breast cancer detection.

The conventional imaging technologies which have been applied to mammography include 1) X-ray camera, X-ray CT imaging, 2) supersonic echo imaging, and 3) MRI.

1) X-ray camera and X-ray CT imaging have the general advantage of high success in the detection of diseased or abnormal tissue or body parts. However, X-ray imaging has a disadvantage of not being capable of obtaining high contrast and high sensitivity between different soft tissue types and abnormalities in soft tissue. This is particularly important in the breast, a major component of which is fat. Thus abnormalities in the breast, in particular due to breast cancer, do not have as good contrast in X-Ray mammography as is desired. X-ray camera and X-ray CT imaging also have the historic disadvantage that side-reaction due to ionization may occur in human body given sufficient exposure.

2) In Supersonic Scan, the quality of images is intensively degraded for the fatty tissues in which the velocity of sound is relatively slower (1476 m/sec) than the cases of muscle and liver (1568 m/sec and 1570 m/sec, respectively). 3) In MRI, The image is degraded by heterogeneous resonant signals due to chemical shift of fat (about 3.3 ppm from proton signal of H^2O). In the case of breast images, a very uniform magnetic field is required for example 1 to 2 ppm in order to measure the images by suppressing the signal attributed to fat. However, it is quite difficult to obtain such uniform magnetic field in breast region, because of different magnetizations between air and the organism, and also because of complicated internal breast structure.

In Mammography using a Terahertz radiation

A breast is sandwiched by two THz transparent plates (which are made of materials which are transparent to the Terahertz pulse, including high quality z-cut quartz or semiconductor materials such as high resistivity Si, GaAs, ZnSe and ZnTe or polymers such as polyethylene, polypropylene, PMMA and poly acrironitoryl, TPX) and is pressed to become as thin as possible. THz radiation with a plurality of frequencies in the range from 50 GHz to 84 THz is irradiated onto the surface of the pressed breast, the breast can be irradiated from the top, underneath or from the sides.

THz radiation which penetrates through breast is sensed by a THz detector and recorded electro-optically using free space electro-optic sampling, photoconductive sampling, or other techniques.

THz signals are collected in 2D to reproduce a 2D image of the pressed breast in some appropriate level of spatial resolution which will be typically less than 1 or 2 mm. The 2D data can be collected by scanning the incident THz beam across the breast.

By using one or a combination of the contrast mechanisms noted on pages 11 and 12 above, a 2D image of the breast can be constructed. In addition, other incarnations of the THz mammograph allow THz images based on THz radiation reflected or scattered from the breast surface and inside the breast to be constructed

The utility of the THz mammography stems in part from the fact that soft tissues such as fat are relatively transparent to THz radiation compared to other tissue types. Moreover, fat has a markedly different spectrum in the THz range compared to other tissue types. As noted above, breast contains concentrated fat which prevents the conventional X-ray, Supersonic and MRI imaging technology from distinguishing breast cancer with high sensitive contrast. THz radiation can easily penetrate through the fatty part. Also, it is well known that the breast cancer frequently causes lime deposits (calcium carbonate is deposited around cancerous region), and THz absorption or reflectance is changed by lime, due to specific absorption and also due to exclusion of water from that region.

Also THz frequency domain imaging can easily distinguish lime areas because of existence of the specific absorption or reflection characteristics of time.

Figure 26 shows an imaging system for breast cancer detection. A THz imaging beam 501 is generated by irradiating a generator crystal 503 with visible light 505 in the same way as previously described. (Many different generation techniques could be used instead of using an optically non linear crystal (as previously described). For example, surface field effect or current surge techniques.) The THz radiation is then reflected off axis parabolic mirror 507 to the output port 509. The THz radiation 501 is then directed towards the sample. The radiation carrying the imaging information is then fed into the detection system via input port 511. The radiation carrying the imaging information 513 is then combined with a reference beam 515 using an off-axis parabolic mirror 517. The detection mechanism uses the AC pockels effect which is described in detail in reference to Figure 19.

Figure 27 shows a further variation on the system of Figure 26. However, instead of the detection mechanism based on the AC pockels effects, a CCD camera 521 is used. The CCD camera is advantageous because the THz beam does not have to be focused to a fine spot in order to extract the imaging information. Focusing of the THz beam for detection can cause imaging information to be lost under certain circumstances. Another major advantage of the CCD camera is that a larger portion of the object (e.g. breast) can be imaged without moving the stepper motor described below, which leads to a reduction of acquisition time.

Figure 28 is a schematic overview of a breast screening apparatus that can use the generation or detection systems of Figures 26 and 27. The output port 509 and the input port 511 are shown. These ports correspond to the output and input ports of Figures 26 and 27. The imaging radiation 501 is reflected through plate 523 onto sample 525. The radiation passes through base plate 527 and is then reflected to input port 511. The generator 531 and the detector 533 are located within a movable platform

535. Scanning means 537 are provided which can move the platform 535 such that the whole of the sample can be irradiated and hence imaged.

Figure 29 shows a detail of the sample receiving area. As shown in Figure 26 the sample is sandwiched between plates 523 and 527. Plate 523 is movable in relation to plate 527. To image the sample, plate 523 is moved towards plate 527. The sample receiving area is designed for transmission measurements.

Figure 30 shows a further variation of the coupling optics shown in the previous figures. Here, polyethylene or resistivity Si plane or convex lenses 561 and 563 are used to focus the beam onto the sample 525. Details of the types of lenses have been hereinbefore described with reference to Figures 21 and 22.

Figure 31 shows a further variation on the coupling optics. Here, the off-axis parabolic mirrors are positioned so that the beam diameter of the sample is frequency independence. These optics are described in more detail with reference to Figure 21.

Figure 32 shows mammography apparatus where the reflected THz beam 601 is directed into input port 511.

Figure 33 again shows a reflection geometry arrangement for breast screening. The reflected beam is passed through a series of off-axis parabolic mirrors arranged in a configuration similar to that described in relation to Figure 21 before being directed into the input port 511.

Further, the radiation from the output port 509 is also passed through a series of off-axis parabolic mirrors in an arrangement similar to that of Figure 21.

Figure 34 shows a further variation on the system of Figure 26. Here, there are two input ports 611 and 613. The input ports can be used to collect either both

transmitted radiation, both reflected radiation or reflected and transmitted information individually.

Figure 35 shows a variation on the system of Figure 34 with two input ports 611 and 613. The radiation for both input ports is fed into separate detection mechanisms, both of which comprise CCD cameras 615 and 617.

Figure 36 shows an imaging system with both transmitted THz 701 and reflected THz 703 are collected.

Figure 37 shows a system for use in THz in the detection of osteoporosis.

Bone may be weakened or rarefied by diseases such as osteoporosis or metastatic cancer. In these cases, there is a change in the density and/or composition of the bone in the effected area, which leads to the bone becoming brittle or weak. Because bone is partially transparent to THz and because both absorption and the refractive index depend on the density and/or composition. Contrast mechanisms such as those previously described on pages 11 and 12 of the specification can be used to both a THz image of the diseases portion and investigate the bone spectroscopically.

Osteoporosis also results in the thinning of the bone.

The osteoporosis apparatus shown in Figure 37 are similar to those of the breast imaging apparatus described in relation to Figures 26 to 36. The sample 701 is placed on a flat stage 703. A polyethylene or other THz transparent material plate 705 is lowered to keep the sample 701 in place. The THz is directed onto the bone and collected in any of the manners described with reference to the breast scanning apparatus (Figures 26 to 36).

Figure 38a shows a visible light image of bones. Figure 38b shows the corresponding THz image of the bone.

Figure 39 shows both visible and THz images taken of a tooth. Appropriate analysis of the time of light of THz light through the tooth allows the enamel layer on the outside of the tooth to be separated from the inner region where the enamel and dentine are both present. Time of flight information can be used to assess the thickness and quality of the enamel and dentine of the tooth. The extra option of the absorption of THz light inside the tooth can be used to identify a cavity inside the tooth. Such images can identify pulp stones and the rate of blood flow into the cavity and inspection through resin-based tooth fillings.

Figure 40 shows THz imaging penetrating through blood. THz in the range (300GHz to 3 THz) can penetrate through at least 90 microns of blood, the sample used was dried blood. However, this technique can also be used to penetrate liquid water. Data for this blood measurement is shown in Figure 40a. Either the time of the frequency domain information can be used to allow identification of increased blood flow in certain areas and establish the certain probability of the presence of a tumour.

THz imaging can also be used to detect the presence of skin tumours. Figure 41 shows a part of the detected THz electric field in the time domain. The sample used to produce these traces is healthy human skin. The three traces are taken from different points on the skin. It can be seen that there is fairly good reproducibility between the traces.

Figure 42 shows a plot of the power spectrum (obtained by a Fourier transform of the time data) against the THz frequency. The upper trace is a reference trace, obtained by measuring the signal in the absence of the sample. This is taken in the absence of the sample. The three lower parts show the data through healthy skin. As explained for Figure 41, the three traces correspond to different parts of the sample. There is good repeatability between measurements.

Figure 43 shows a plot of the transmittance the signal through the skin divided by the reference signal for the data of Figure 42.

Figure 44 shows a plot of the absorption (i.e. the reciprocal of the transmission) for both healthy and tumourous skin for frequencies in the THz range. The four upper traces correspond to the signal measured in the tumourous region. The lower trace corresponds to the healthy skin. There is a difference between the tumourous traces. However, it can be clearly seen that the tumour can be easily distinguished from the healthy skin.

Figure 45a shows a visible image of a tumour. Figure 45b shows the corresponding THz image of the same sample. To obtain the THz image, the sample has been subdivided into a plurality of pixels. The magnitude of the minimum of the E-field (i.e. the directly measured magnitude of the THz signal) has been plotted at each pixel. It can be seen that the tumourous region (in the centre of the image) has a far higher absorption than that of the healthy skin or the region of the sample. This image can be used to determine the lateral extent of the tumour.

Figure 46 shows a further image of the tumourous sample. This uses the time of flight of the THz pulse through the sample in order to determine the density and thickness of the tumour into the skin. In this particular figure, the time and position of the e-field minima has been plotted for each pixel.

Combining the straight absorption image of Figure 45 with the time of flight image of Figure 46 allows a 3D picture of the tumour to be built up.

As previously mentioned, THz frequencies can be used to determine the composition of liquid. Figure 47 shows the absorption coefficient measured at different THz frequencies for a well known moisturiser, Oil of Olay (registered trade mark), a non-branded moisturiser which will be referred to as "M" and water. Water absorbs radiation in the THz region. Therefore, the water trace has the highest absorption

coefficient, M has a medium composition and Olay has the lowest absorption. Figure 48 shows a plot of THz fields (arbitrary units - which is closely related to the absorption and transmission of the sample) in the time domain. Looking at the maxima of the trace at 2.5picoseconds, pure Olay is the largest trace. Pure water has the lowest field peak because water is more strongly absorbed than pure Olay. The traces in between are for samples with the following contents: 80% Olay, 20% water; 60% Olay, 40% water; 40% Olay and 60% water; and 20% Olay and 80% water. The pulse size of the mixture decreases as more water is added to the Olay.

Figure 49 shows a plot of the average absorption coefficient against the percentage of Olay by volume taken from Figure 48. Comparing this with the average absorption coefficient of M allows the percentage of Olay within the Aloe to be calculated. The finding was that M is equivalent to a mixture of 55% Olay and 43% water.

CLAIMS:

1. A method of imaging a sample, the method comprising the steps of:

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample which is to be imaged into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies; and

c) generating an image of the area of the sample from the radiation detected in step (b) using a frequency or a selection of frequencies from the plurality of frequencies in the pulsed electro-magnetic radiation.

2 A method according to claim 1, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz

3 A method according to claim 2, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 500 GHz to 10 THz.

4. A method according to any preceding claim, wherein the input radiation is transmitted through the sample.

5. A method according to any of claims 1 to 3, wherein the input radiation is reflected from the sample.

6. A method according to any preceding claim, wherein the image is generated from a single frequency.

7. A method according to any of claims 1 to 5, wherein a single image is generated using a selected frequency range..

8. A method according to claim 7, wherein the selected frequency range is less than a third of the frequency range of the pulsed electro-magnetic radiation which irradiates the sample.

9. A method according to any preceding claim, wherein a plurality of images is generated from a corresponding plurality of different frequencies.

10. A method according to any preceding claims, wherein, in step (b), the radiation is detected in the time domain and is Fourier transformed to obtain frequency dependent data.

11. A method according to any preceding claim. wherein a fraction of the input pulsed radiation does not irradiate the area of the sample. which is to be imaged, and is detected as a reference signal.

12. A method according to any of claims 1 to 10 . wherein the method further comprises the step of detecting a reference signal when the area of the sample to be imaged is absent from the path of the pulsed electro magnetic radiation.

13. A method according to either of claims 11 or 12, wherein step (c) comprises the steps of deriving frequency dependent image data from the detected radiation and plotting the frequency dependent image data for each pixel to obtain an image of the sample.

14. A method according to claim 13, wherein step (c) further comprises the steps of calculating the Power Spectrum for the radiation detected in step (b) and calculating the Power Spectrum of the reference signal.

15. A method according to claim 14, wherein the frequency dependent image data is calculated by subtracting the Power Spectrum of the reference signal from the Power Spectrum of the radiation detected in step (b).

16. A method according to claim 14, wherein the frequency dependent image data is calculated by dividing the Power Spectrum of the radiation detected in step (b) by the Power Spectrum of the reference signal.

17. A method according to claim 13 , wherein the frequency dependent image data is calculated by deriving the frequency dependent absorption coefficient from a Fourier transform of the radiation detected in step (b) and a Fourier transform of the reference signal.

18. A method according to claim 13, wherein the frequency dependent image data is calculated by deriving the frequency dependent refractive index from a Fourier transform of the radiation detected in step (b) and a Fourier transform of the reference signal.

19. A method according to any of claims 13 to 18, wherein the frequency dependent image data is derived for two or more frequencies.

20 A method according to claim 19, wherein the frequency dependent image data for the two frequencies is added, subtracted, multiplied or divided for each pixel.

21. A method according to any preceding claim, wherein the frequency dependent image data is obtained for a narrow range of frequencies and the image data is integrated over this narrow range.

22. A method according to any of claims 13 to 21, wherein the frequency dependent image data is processed by subdividing the data into a plurality of bands, wherein each

band represent a range of magnitudes of the image data for a single frequency or narrow frequency range and each band is assigned a single value or colour.

23. A method according to claim 22, wherein the bands are not of equal widths in magnitude.

24. A method according to either of claims 22 or 23, wherein the image data is derived for more than one frequency.

25. A method according to claim 24, wherein the image data for two frequencies is combined according to a predetermined rule.

26. A method according to claim 24, wherein there are two bands assigned the values '0' and '1' and the predetermined rule is a BOOLEAN algebraic rule.

27. A method according to any preceding claim, the method further comprising the step of displaying a sequence of images generated in step (c) for a plurality of different frequencies.

28. A method according to claim 27, wherein the image generated in step (c) is scanable through a continuum of frequencies.

29. A method according to claim 27, wherein the image generated in step (c) can be stepped through a plurality of discrete frequencies.

30. A method of imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels. and detecting radiation from each pixel over a plurality of frequencies; and

c) generating an image of the area of the sample from the radiation detected in step (b) by calculating the difference between the temporal positions of a maxima and a minima of the radiation detected in step (b) for each pixel.

31 A method according to claim 30, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz.

32. A method according to claim 31, wherein the image is generated by calculating the difference between the temporal positions of the main minimum and maximum peaks.

33. A method according to any of claims 30 to 32, wherein step (c) further comprises the step of multiplying the difference between the temporal position of the maxima and minima, by a magnitude of the radiation.

34. A method according to claim 33, wherein the magnitude of the radiation is the difference in magnitude between the minimum and maxima of the detected radiation.

35. A method according to claim 33, wherein the magnitude of the radiation is the minimum or maximum of the detected radiation.

36. A method of imaging a sample, the method comprising the steps of :

a) irradiating the sample to be imaged with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 50 GHz to 84 THz;

b) subdividing an area of the sample into a two dimensional array of pixels. and detecting radiation from each pixel over a plurality of frequencies; and

c) generating an image of the area of the sample from the radiation detected in step (b) by calculating the difference between the magnitude of a maxima and a minima of the radiation detected in step (b) for each pixel.

37. A method according to claim 36, wherein the sample to be imaged is irradiated with pulsed electro-magnetic radiation with a plurality of frequencies in the range from 100 GHz to 20 THz

38. A method for imaging a sample, according to either of claims 36 or 37, wherein the peak minima and maxima are the main minimum and maximum.

39. A method of imaging according to any preceding claim, configured to image the difference between ordinary and diseased tissue.

40. A method according to claim 39, wherein the difference between diseased tissue and non diseased tissue is determined by the different and/or absorption reflectance and/or dispersion of THz light which is irradiated to organisms on living tissue in-vivo or in-vitro.

41. A method according to any preceding claim, wherein the sample to be imaged is a human or animal.

42. A method according to any preceding claim, wherein the method is configured to distinguish differences in bone density of the imaged sample.

43. A method according to any preceding claim, wherein the method is configured to distinguish between the different types of tissue which comprise human or animal teeth.

44. A method according to any preceding claim, wherein the method is configured to image Breast cancer.

45. A method for detecting cancer in a sample, the method comprising the steps of:

a) irradiating the sample with pulsed electromagnetic radiation with a plurality of frequencies in the range from 50GHz to 84THz;

b) subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) generating an image of the area of the sample from the radiation detected in step (b).

46. A method according to claim 45, wherein step (c) comprises the step of plotting the refractive index of the sample for each pixel to generate the image.

47. A method according to claim 45, wherein step (c) comprises the step of plotting the absorption coefficient of the sample to generate the image.

48. A method according to claim 45, wherein step (c) comprises the step of plotting the maxima or minima of the electric field for each pixel.

49. A method according to claim 45, wherein step (c) comprises the step of plotting the time of flight of the THz pulse through the sample for each pixel.

50. A method according to any of claims 45 to 49, the method further comprising the step of analysing the detected radiation for the presence of lime in the sample.

51. A method according to any of claims 45 to 50, wherein the method further comprises the step of analysing the water content in the sample.

52. A method according to any of claims 45 to 51, wherein the sample is a human or animal breast.

53. A method according to any of claims 45 to 51, wherein the sample comprises human or animal skin.

54. A method according to any of claims 45 to 53, wherein the said area of the sample comprises an area where there is a suspected tumour and a healthy area of the sample.

55. An apparatus for imaging a sample, the apparatus comprising:

a) means for irradiating the sample with pulsed electro-magnetic radiation having a plurality of frequencies in the range from 50 GHz to 84 THz;

b) means for subdividing an area of the sample into a two dimensional array of pixels;

c) means for detecting radiation from each pixel over a plurality of frequencies; and

d) means for generating an image of the area of the sample from the detected radiation by using a frequency or a selection of frequencies from the plurality of frequencies in the pulsed electro-magnetic radiation.

56. An apparatus for imaging a sample, the apparatus comprising:

a) means irradiating the sample with pulsed electro-magnetic radiation having a plurality of frequencies in the range form 50 GHz to 84 THz;

b) means for subdividing an area of the sample into a two dimensional array of pixels;

c) means for detecting radiation from each pixel over a plurality of frequencies; and

d) means generating an image of the area of the sample from the detected radiation by calculating the difference between the temporal positions of a maximum and a minimum of the detected radiation detected for each pixel.

57. An apparatus for imaging a sample, the apparatus comprising:

a) means for irradiating the sample with pulsed electro-magnetic radiation having a plurality of frequencies in the range from 50 GHz to 84 THz;

b) means for subdividing an area of the sample into a two dimensional array of pixels;

c) means for detecting radiation from each pixel over a plurality of frequencies; and

d) means for generating an image of the area of the sample from the detected radiation by calculating the difference between the magnitude of a maximum and a minimum of the detected radiation for each pixel.

58. An apparatus according to any of claims 55 to 57, wherein the sample to be imaged is a human or animal breast.

59. An apparatus according to claim 58, comprising means for holding the breast to be imaged.

60. An apparatus according to any of claims 55 to 59, wherein the apparatus further comprises means for detecting the time of flight of a pulse of radiation through the sample.

61. An apparatus according to any of claims 55 to 60 wherein the apparatus comprises means for collecting both transmitted and reflected radiation.

62. An apparatus for detecting cancer in a sample, the apparatus comprising:

a) means of irradiating the sample with pulsed electromagnetic radiation with a plurality of frequencies in the range from 50GHz to 84THz;

b) means of subdividing an area of the sample into a two dimensional array of pixels, and detecting radiation from each pixel over a plurality of frequencies;

c) means of generating an image of the area of the sample from the radiation detected in step (b).

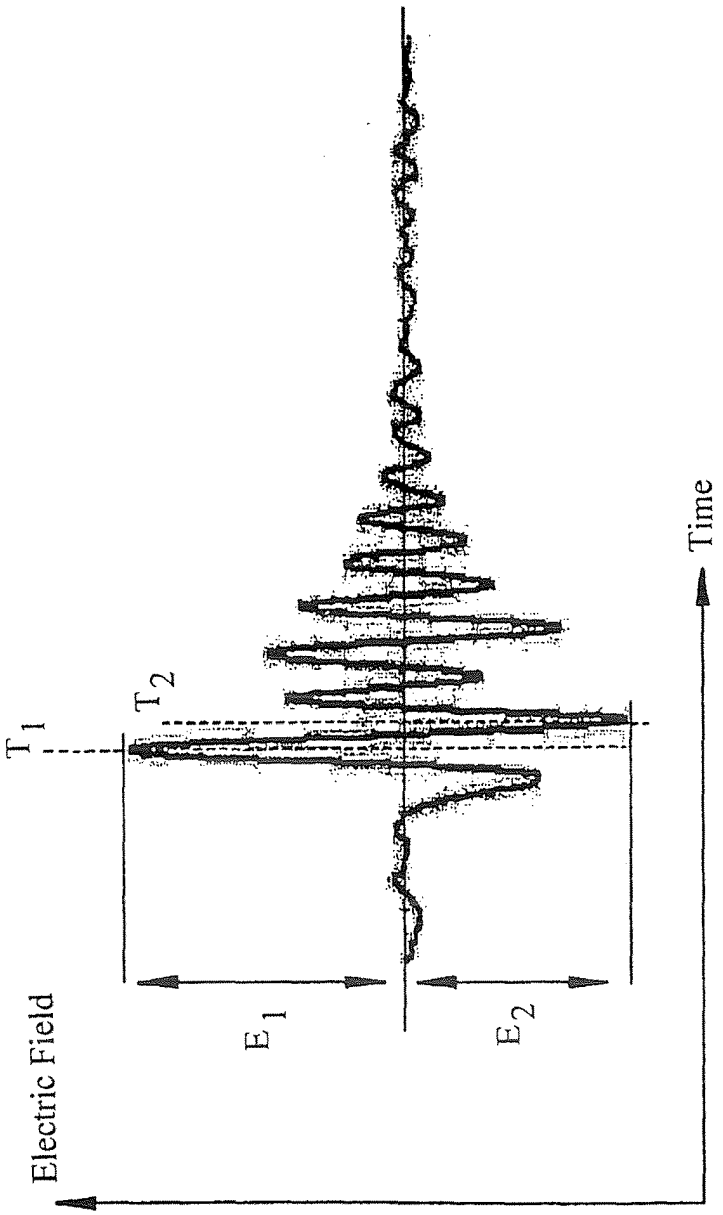


Fig. 1

IPR2023-00033
Apple EX1002 Page 504

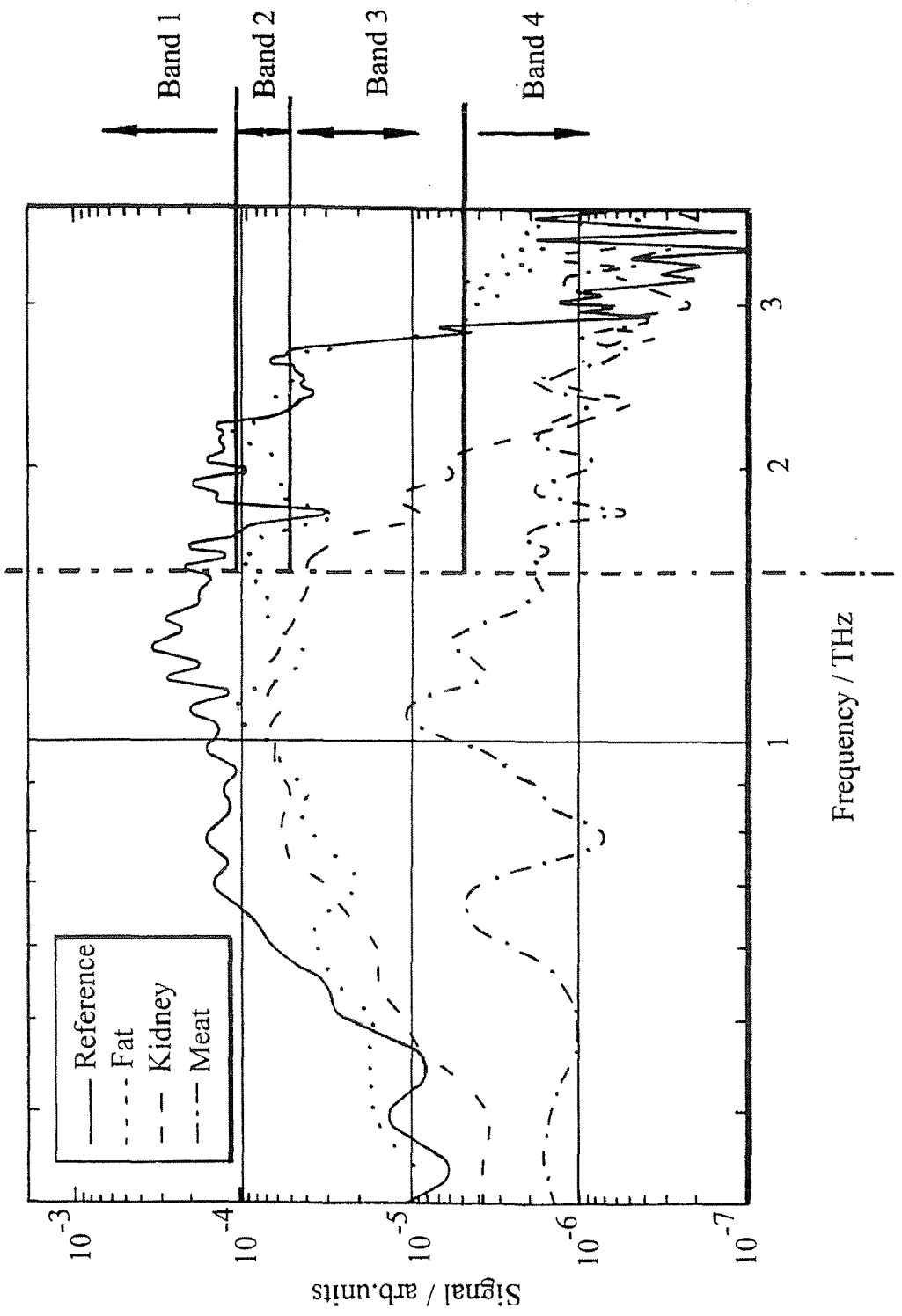


Fig. 2

IPR2023-00033
Apple EX1002 Page 505

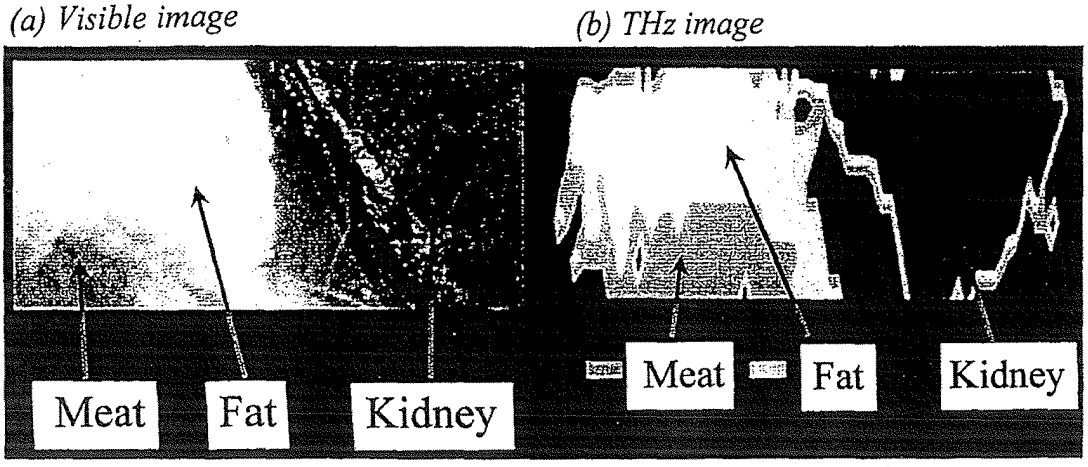


Fig. 3

IPR2023-00033
Apple EX1002 Page 506

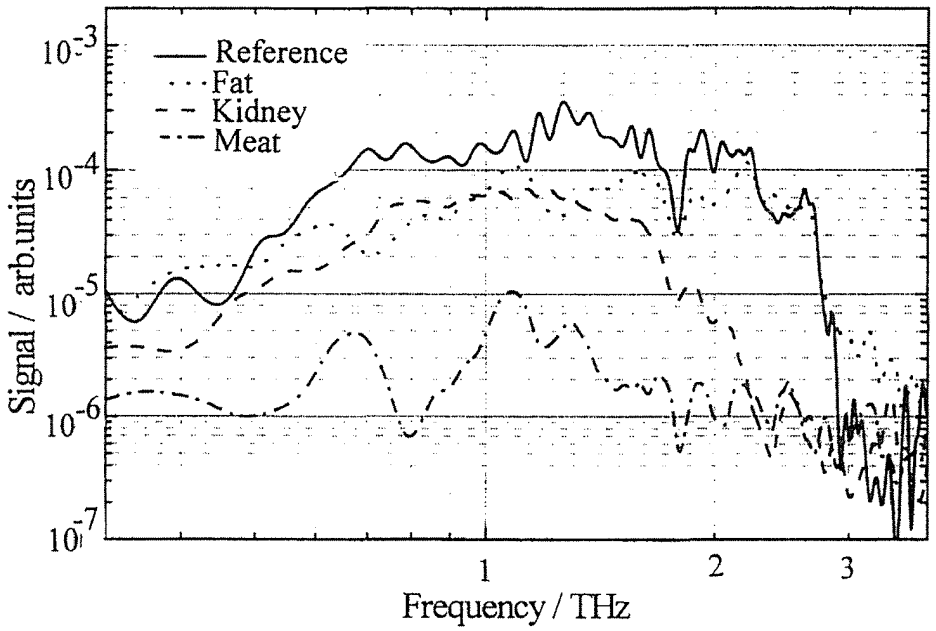


Fig. 4

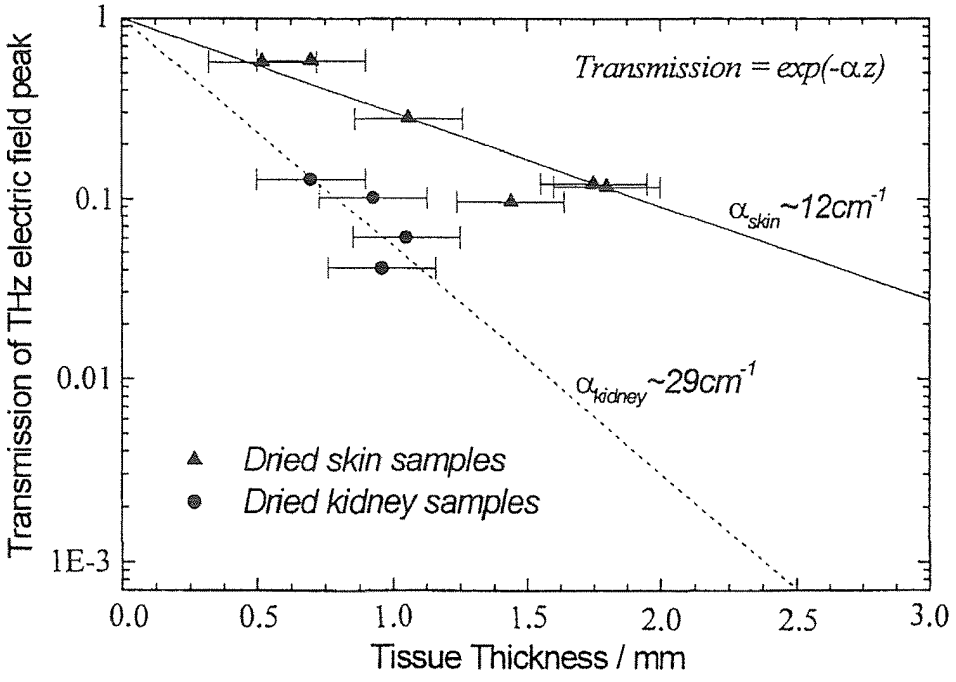


Fig. 5

IPR2023-00033
Apple EX1002 Page 507

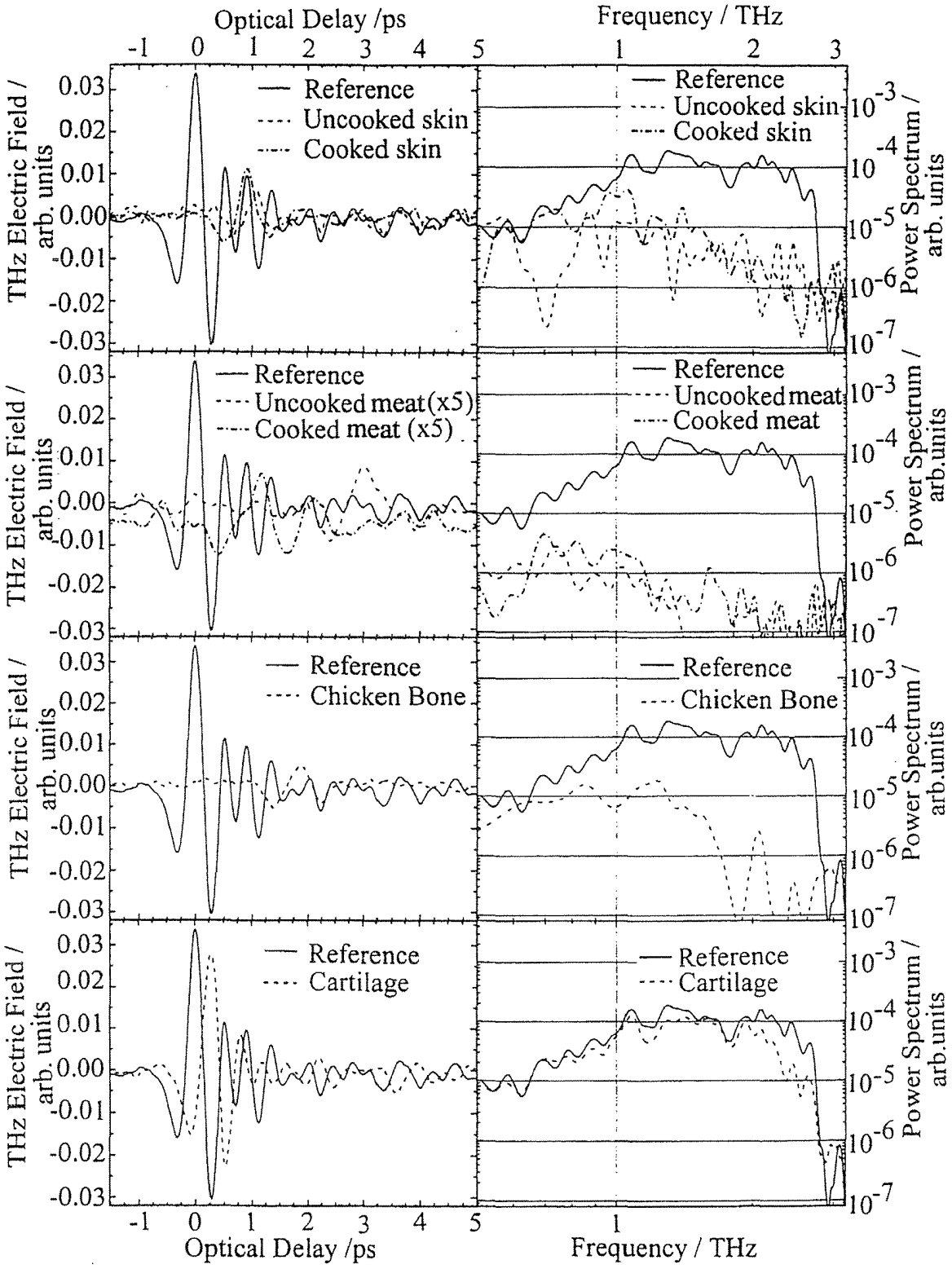


Fig. 6

IPR2023-00033
Apple EX1002 Page 508

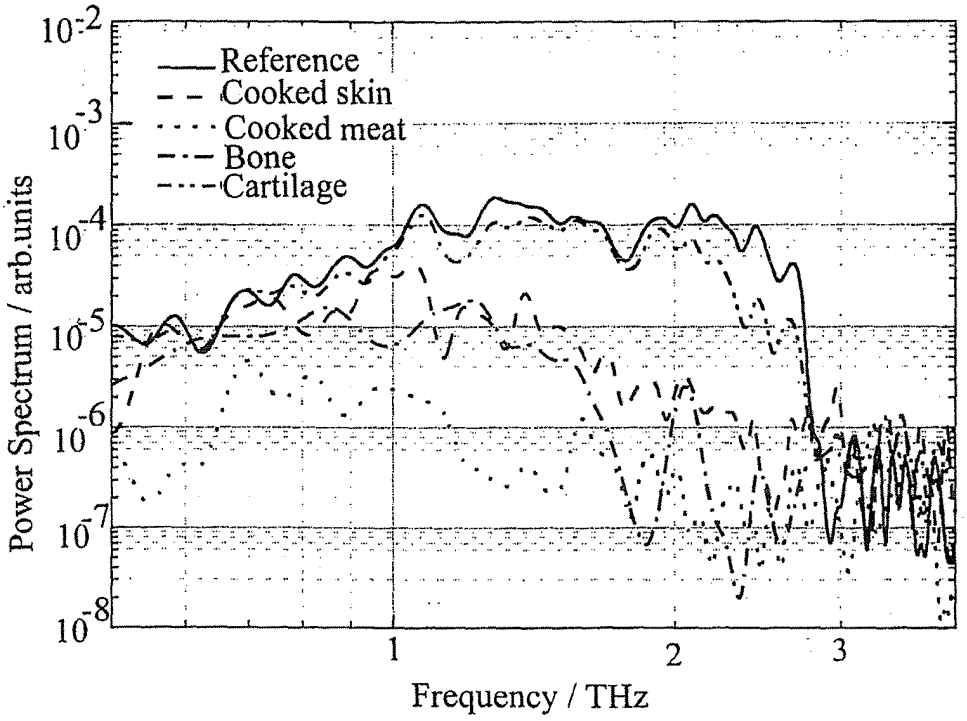


Fig. 7

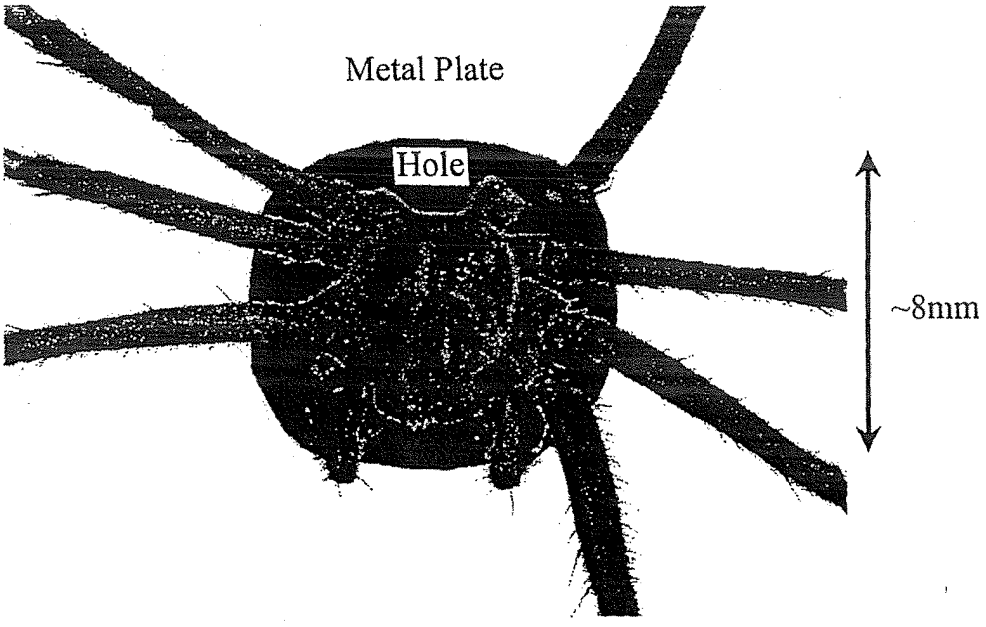


Fig. 8

IPR2023-00033
Apple EX1002 Page 509

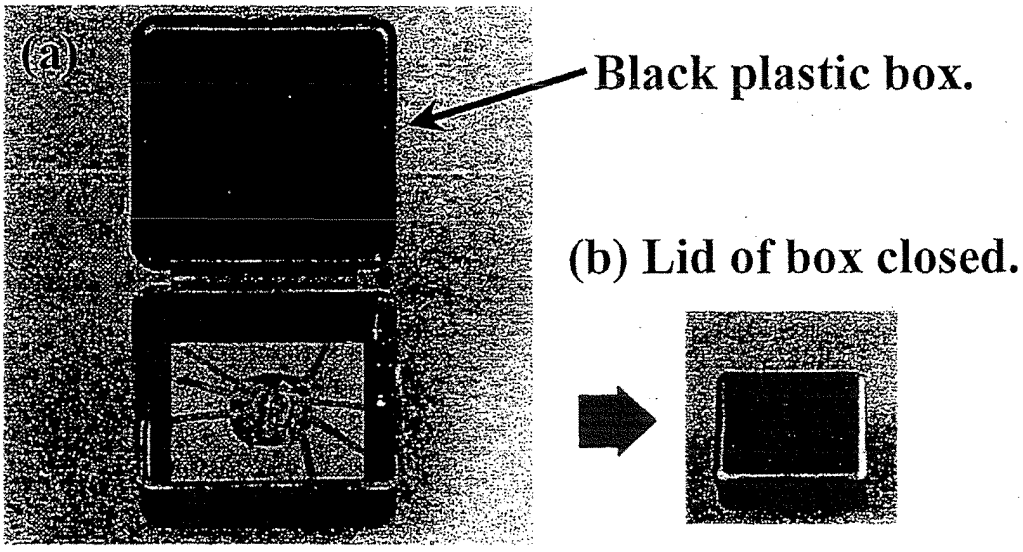


Fig. 9

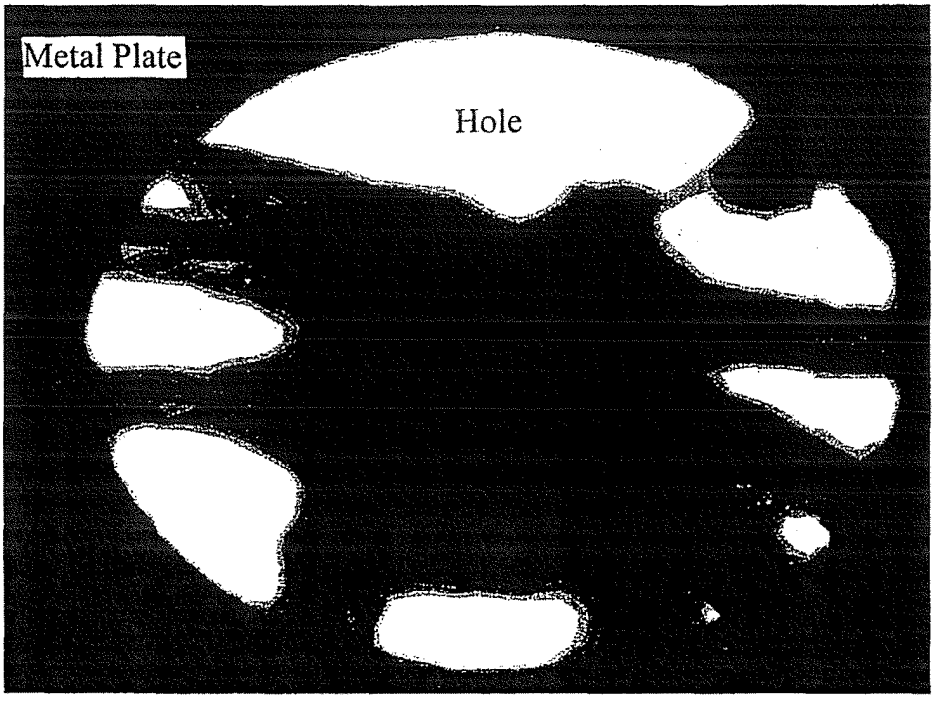


Fig. 10

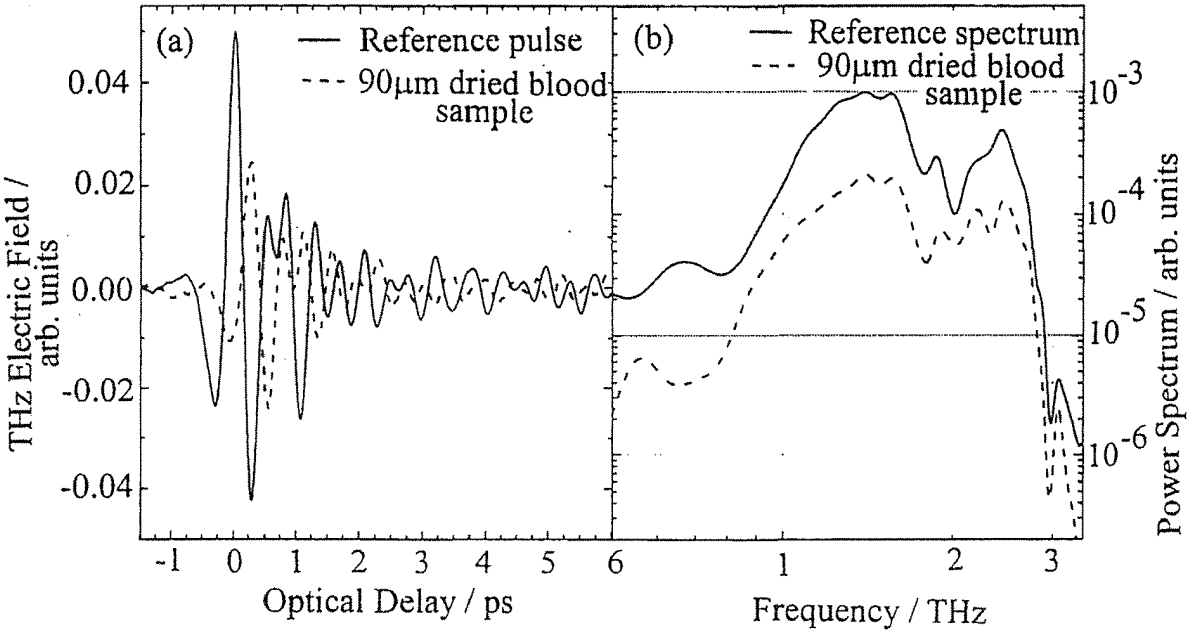


Fig. 11

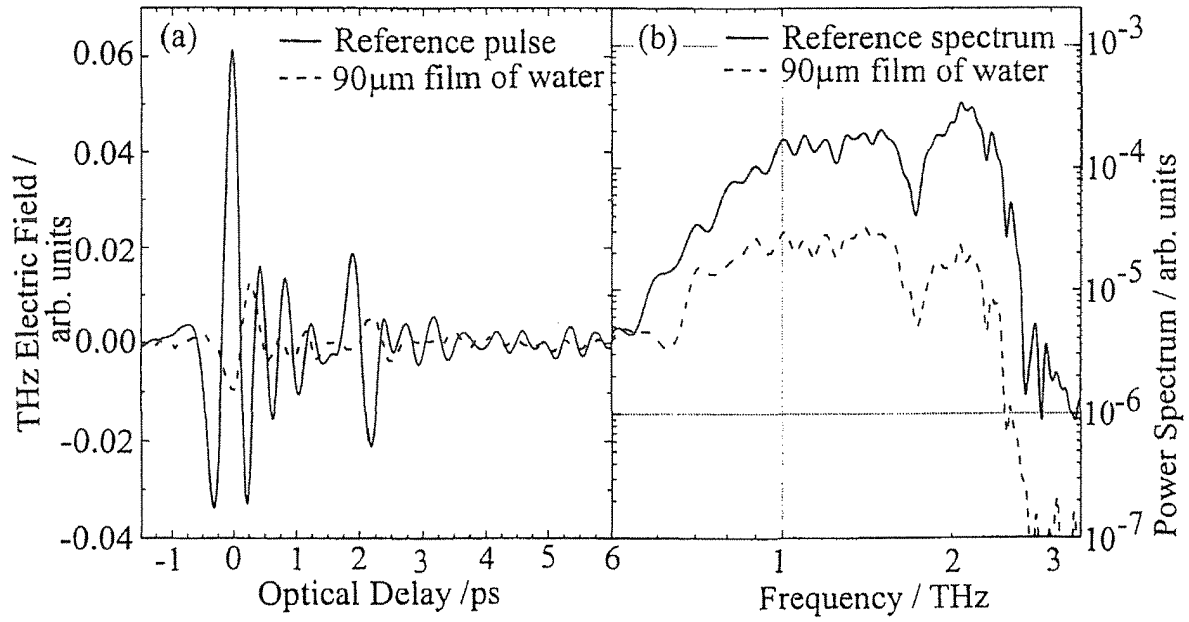


Fig. 12

IPR2023-00033
Apple EX1002 Page 511

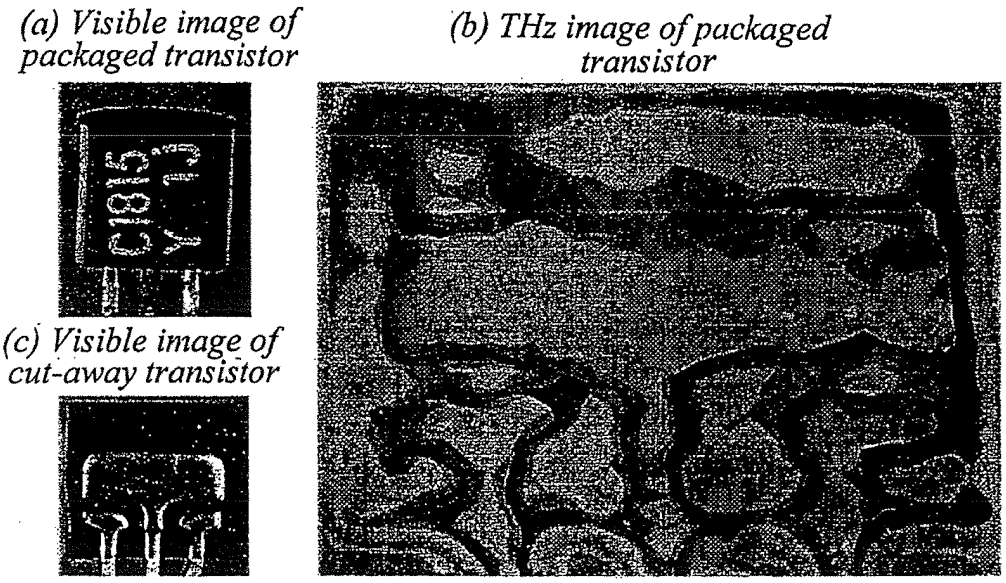


Fig. 13

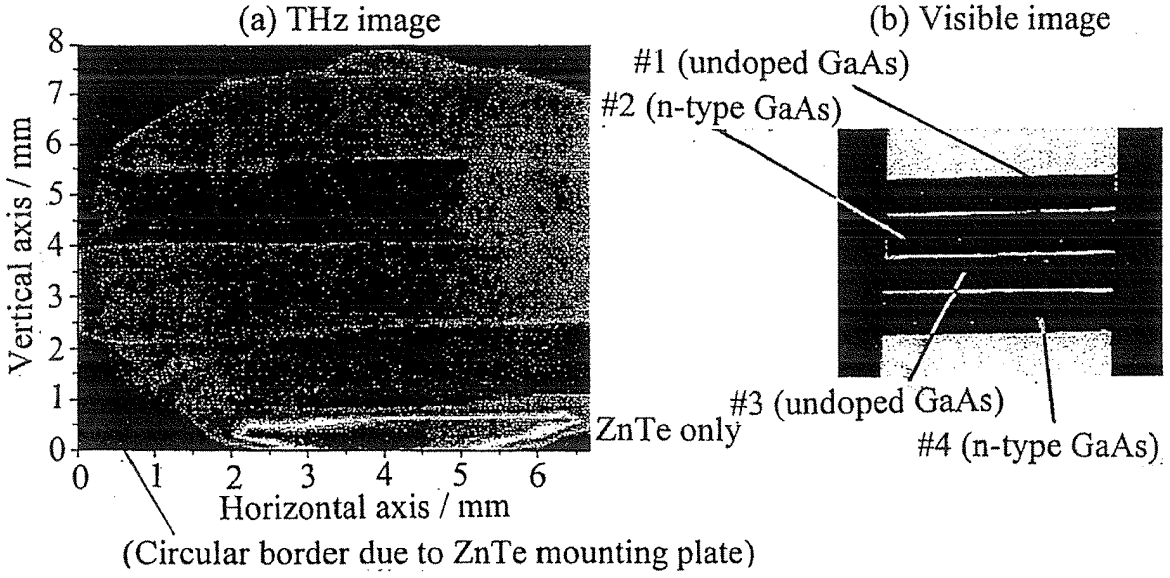


Fig. 14

IPR2023-00033
Apple EX1002 Page 512

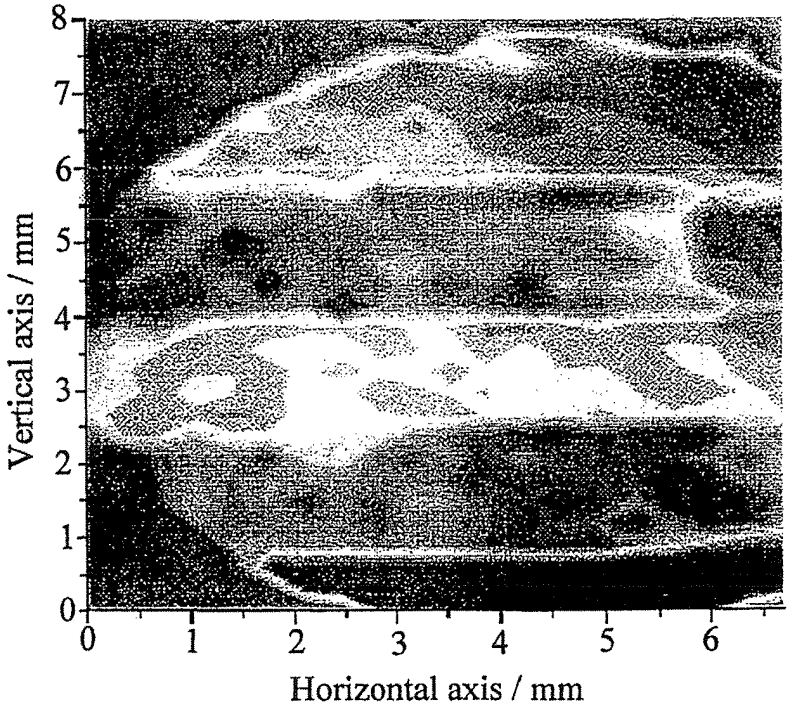


Fig. 15(a)

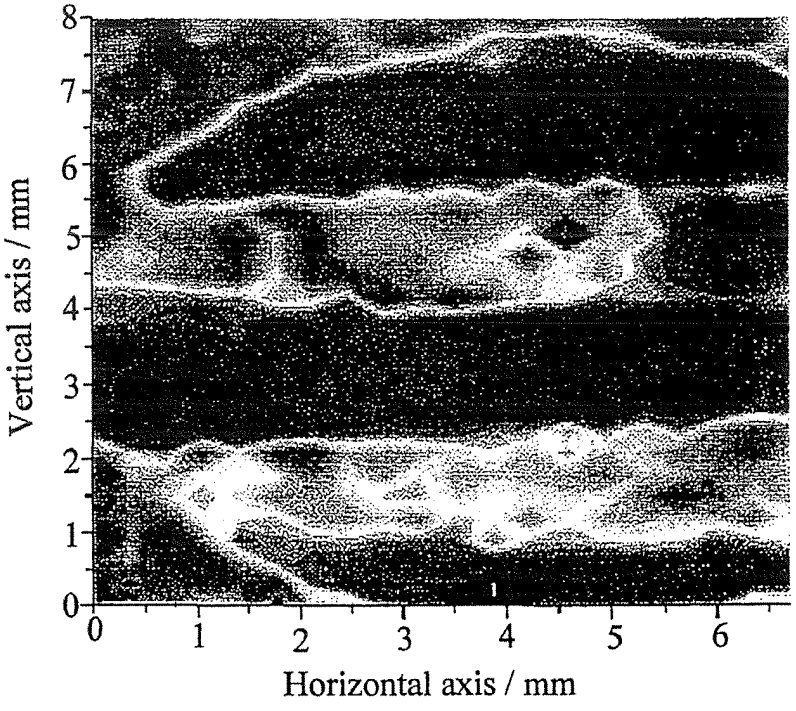


Fig. 15(b)

IPR2023-00033
Apple EX1002 Page 513

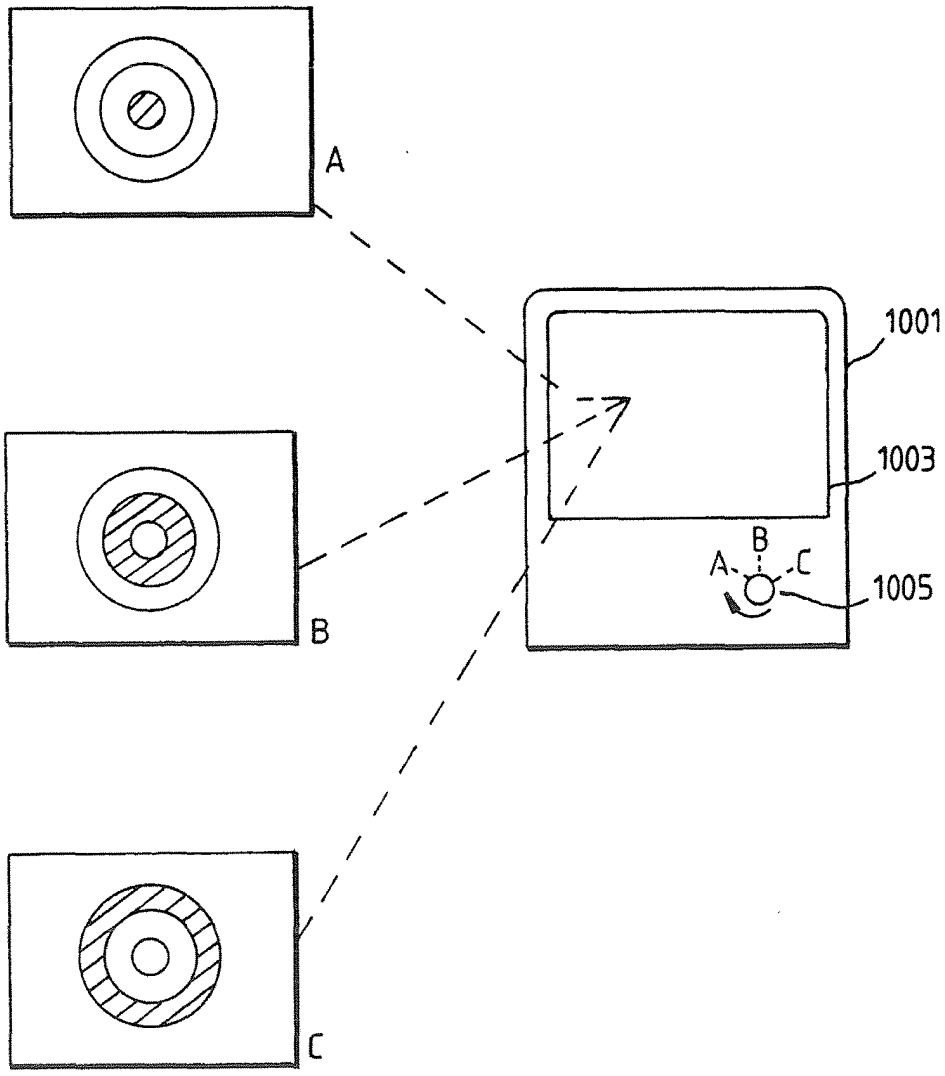


Fig. 16

IPR2023-00033
Apple EX1002 Page 514

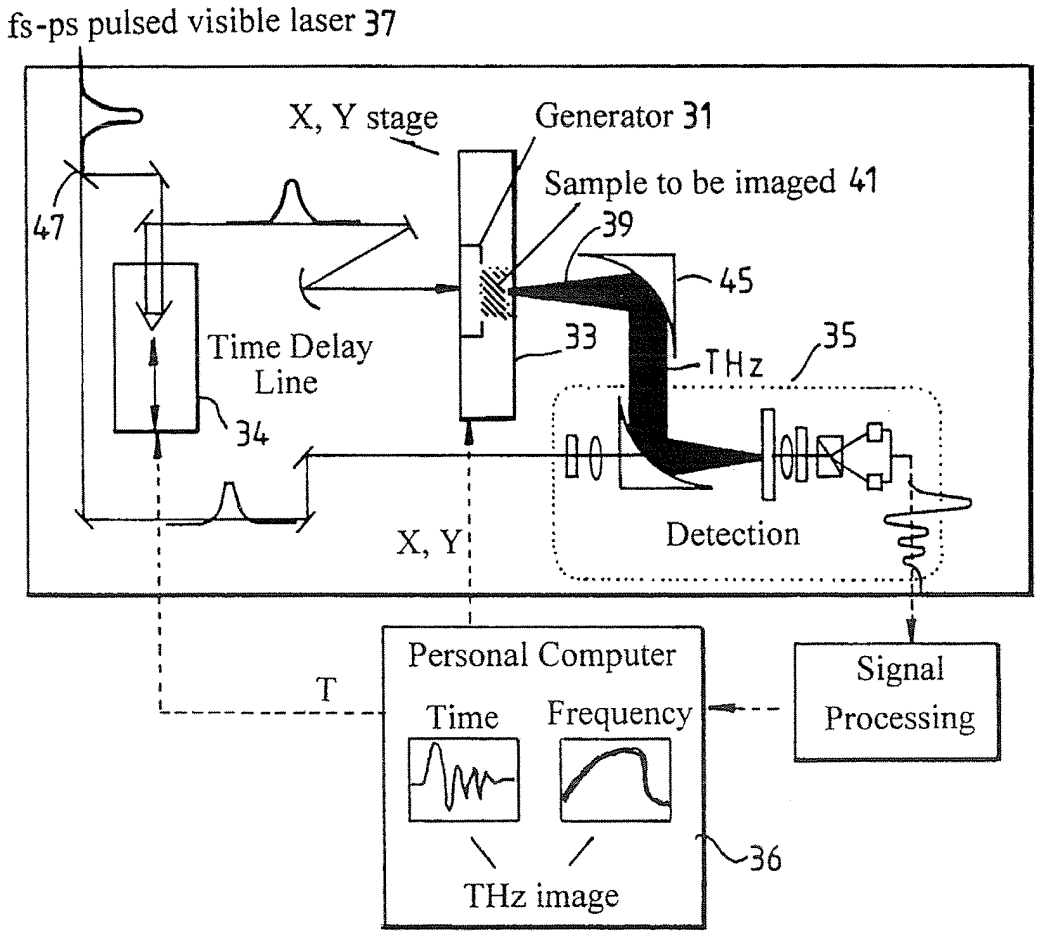


Fig. 17

IPR2023-00033
Apple EX1002 Page 515

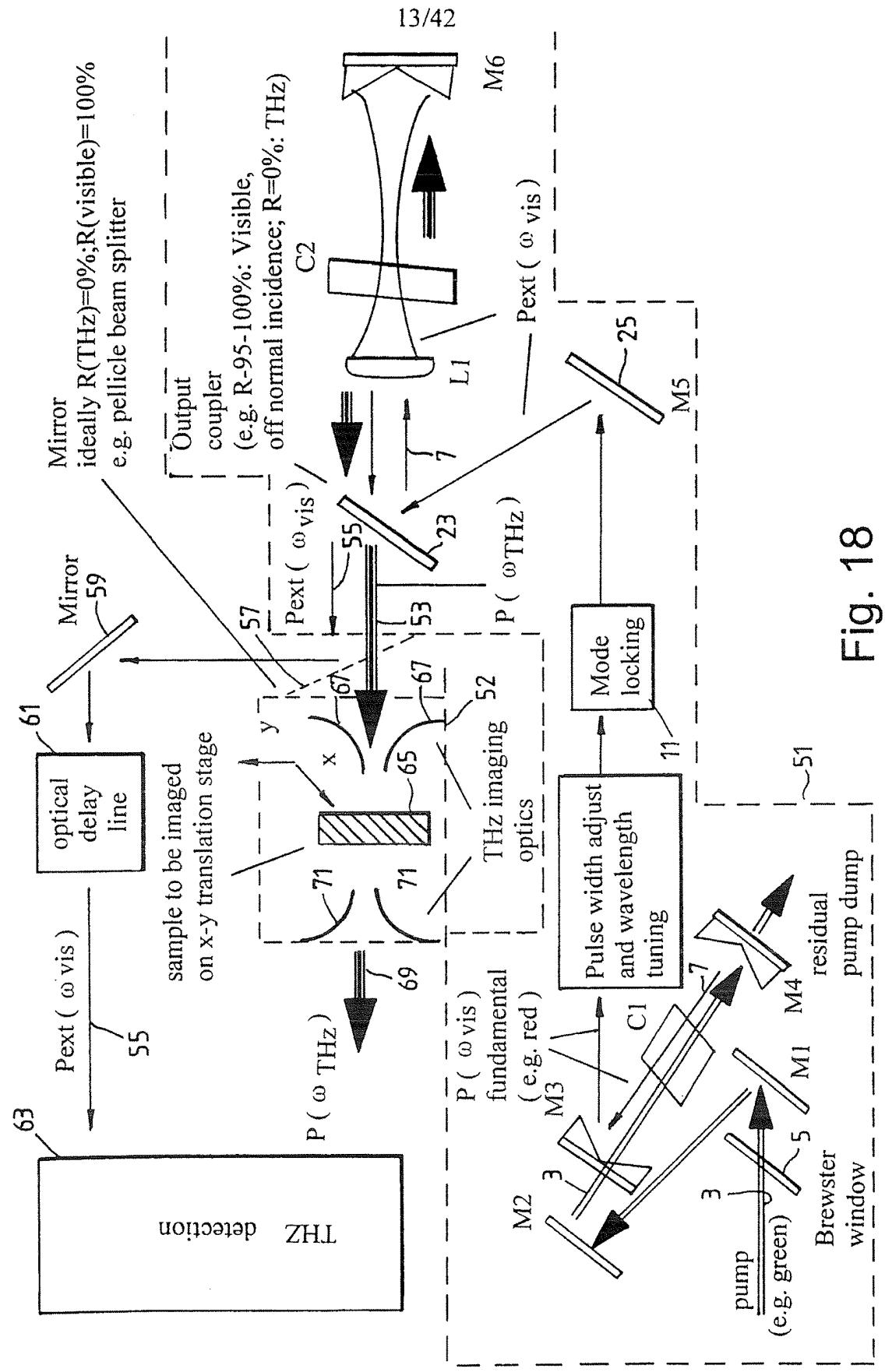


Fig. 18

IPR2023-00033
Apple EX1002 Page 516

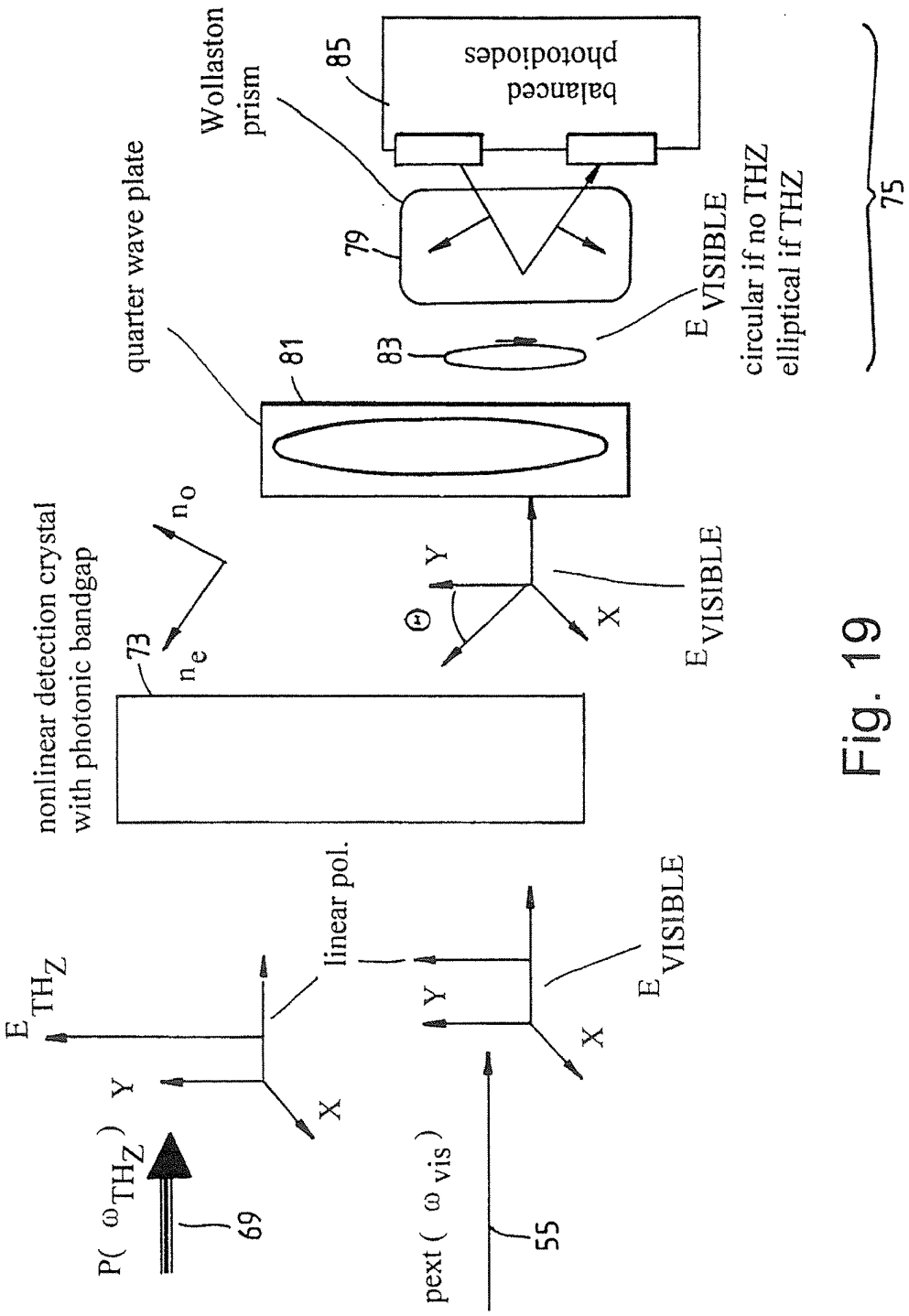


Fig. 19

IPR2023-00033
Apple EX1002 Page 517

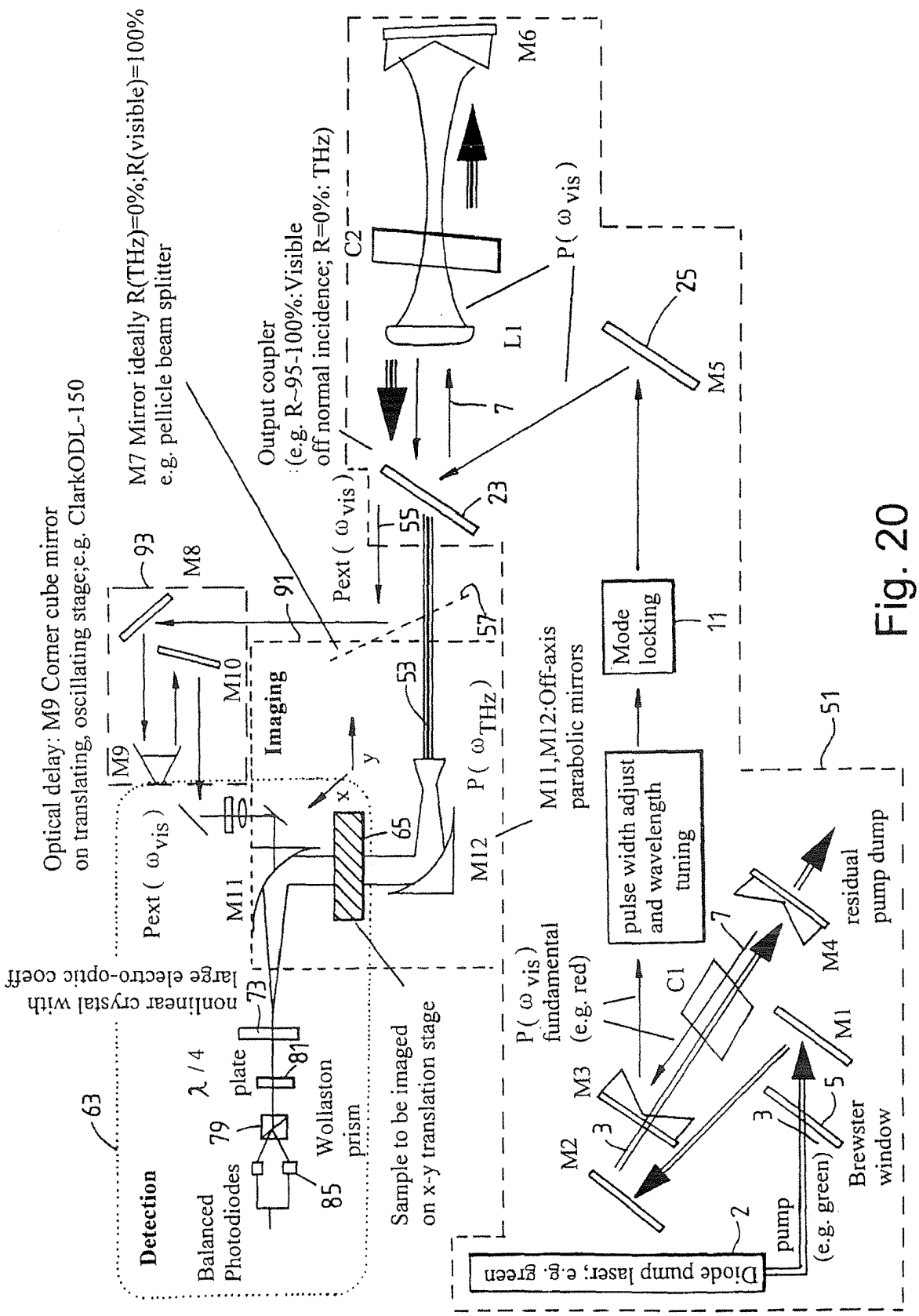


Fig. 20

IPR2023-00033
Apple EX1002 Page 518

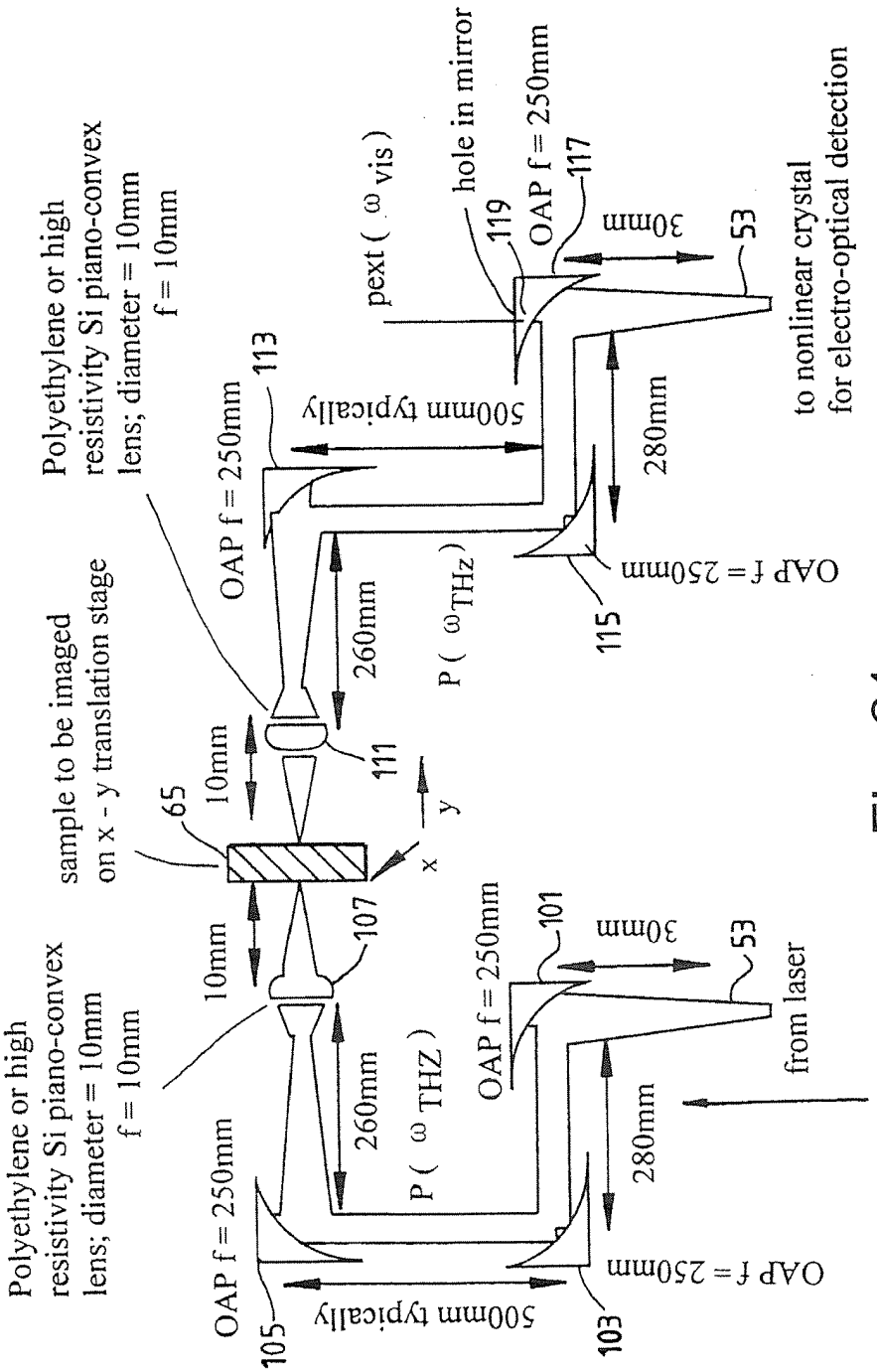


Fig. 21

IPR2023-00033
Apple EX1002 Page 519

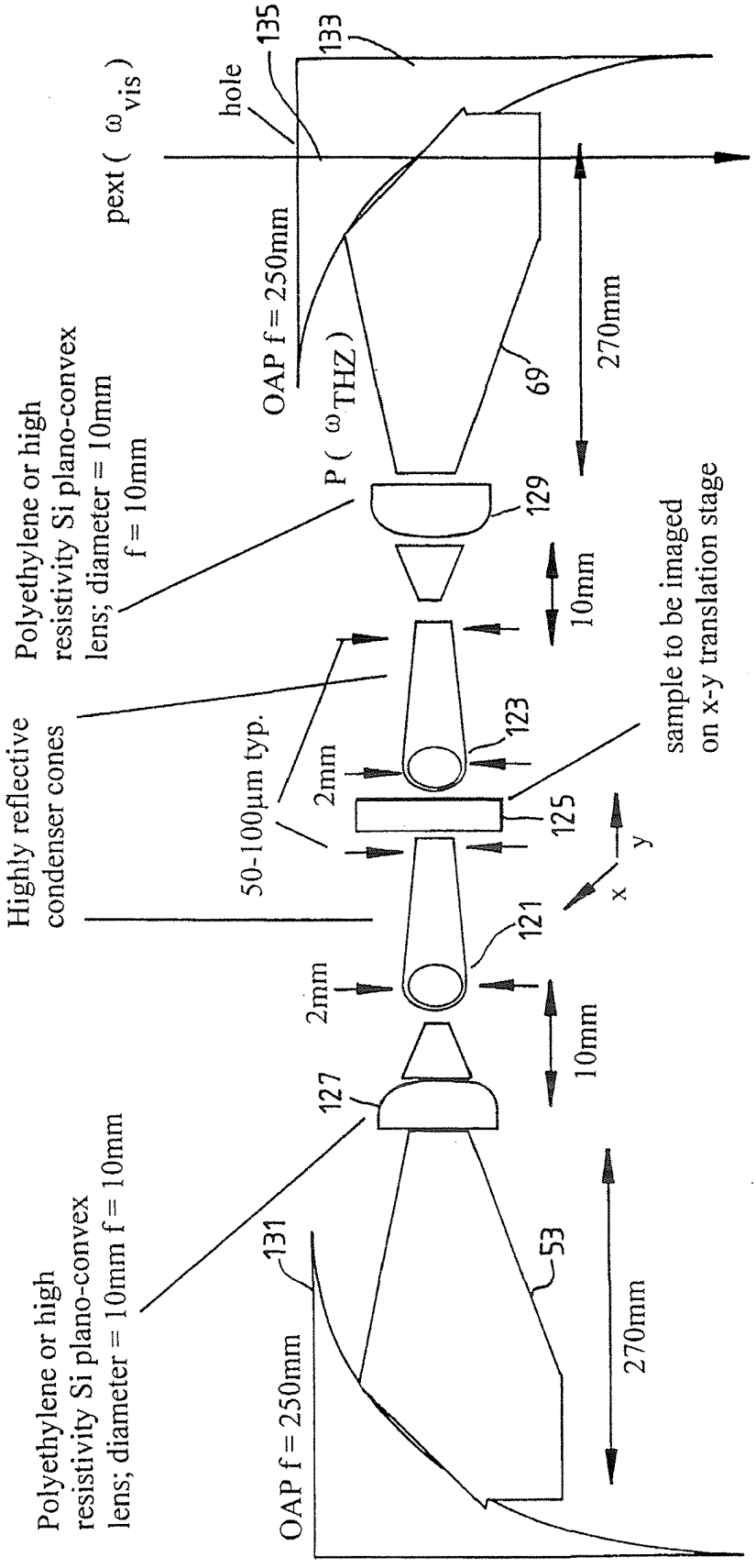


Fig. 22

IPR2023-00033
Apple EX1002 Page 520

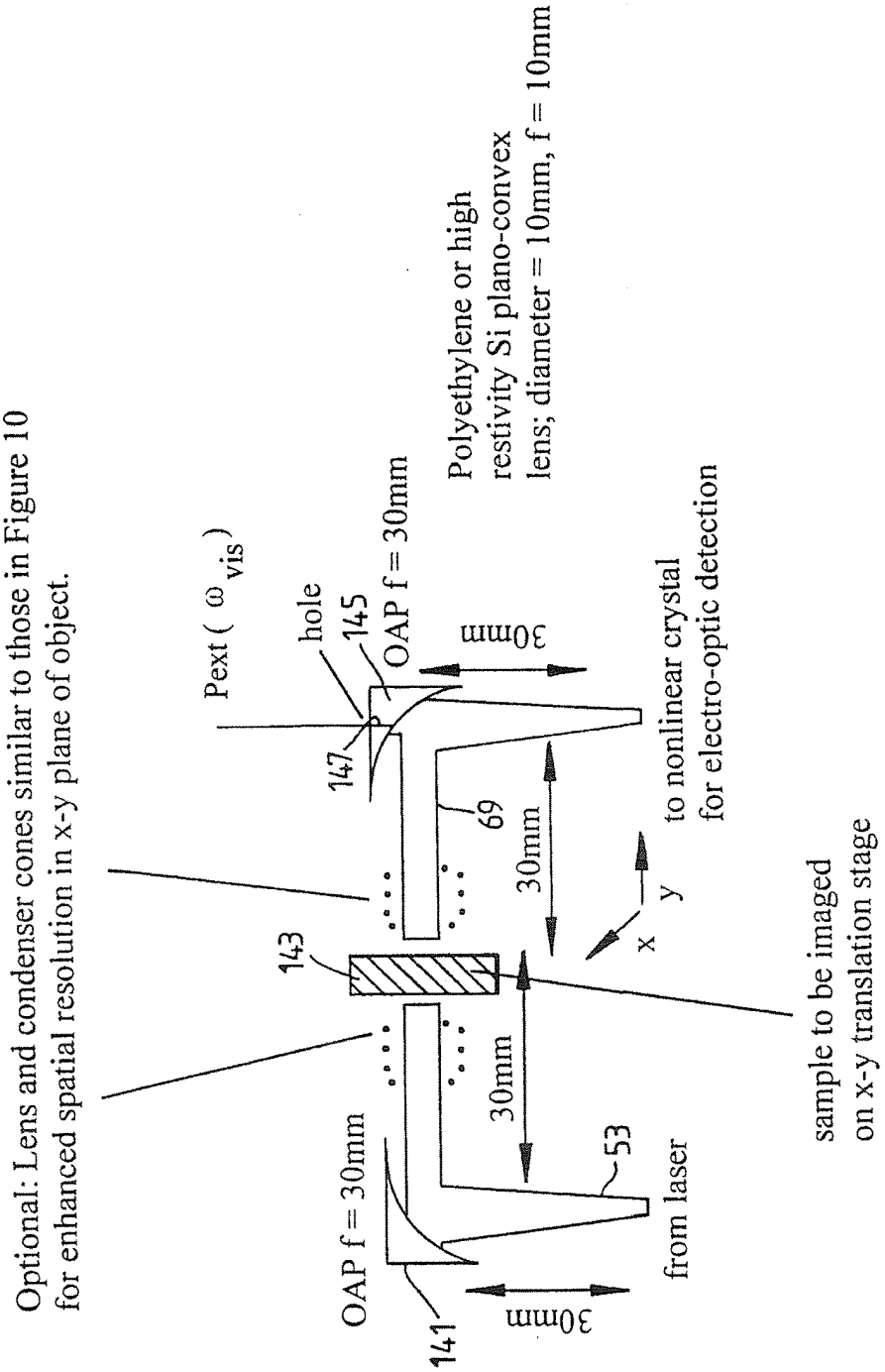


Fig. 23

IPR2023-00033
Apple EX1002 Page 521

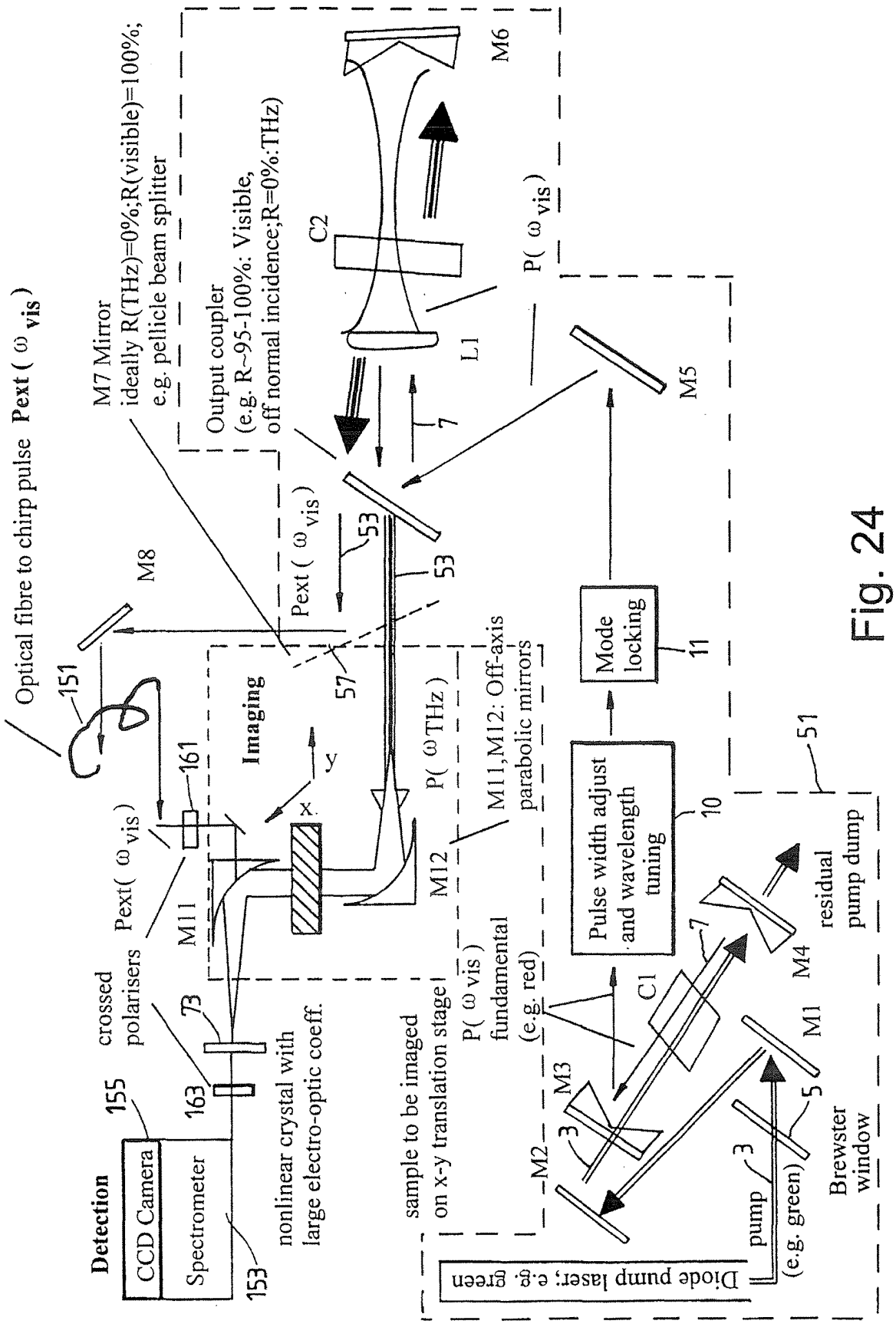


Fig. 24

IPR2023-00033
Apple EX1002 Page 522

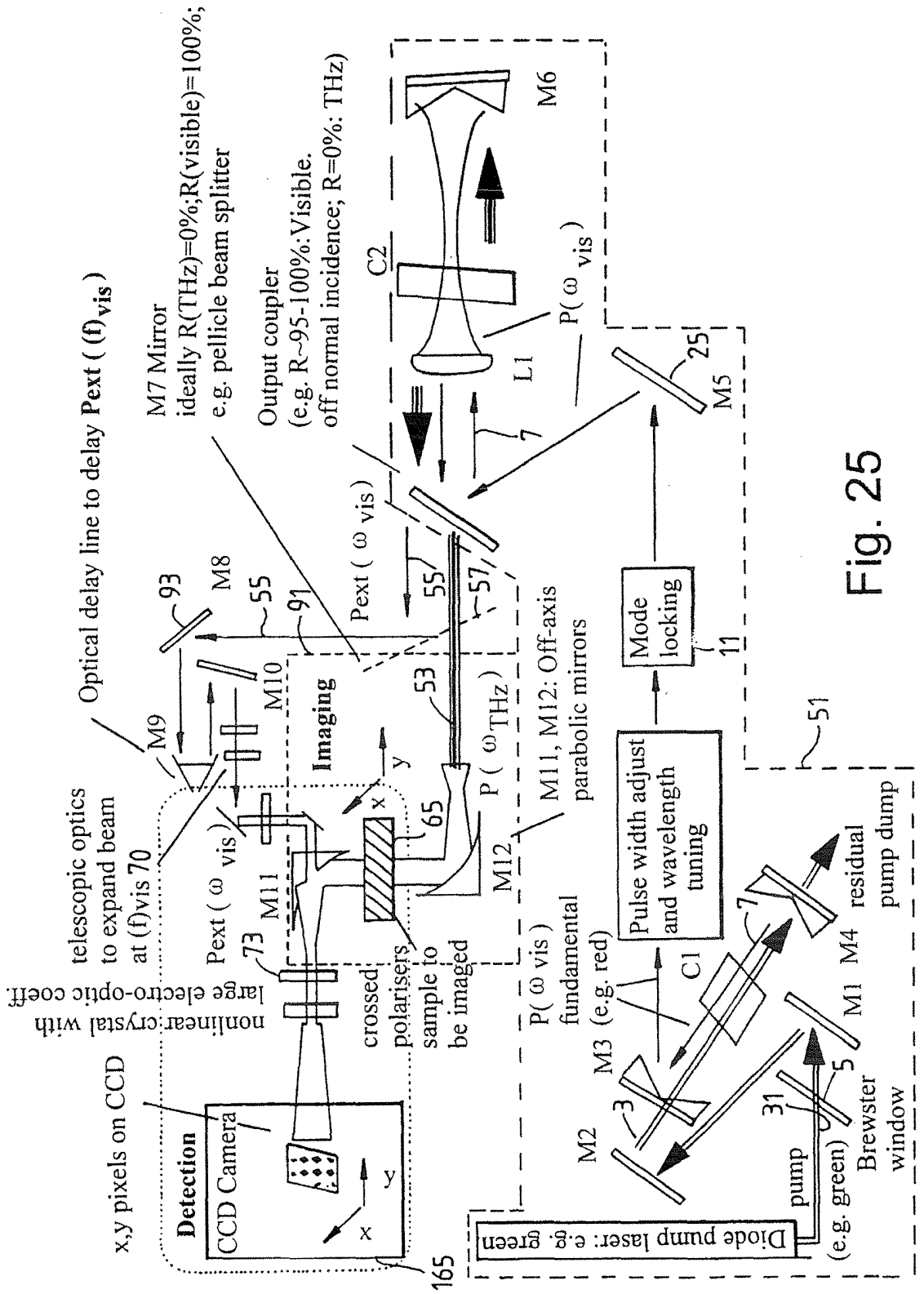


Fig. 25

IPR2023-00033
Apple EX1002 Page 523

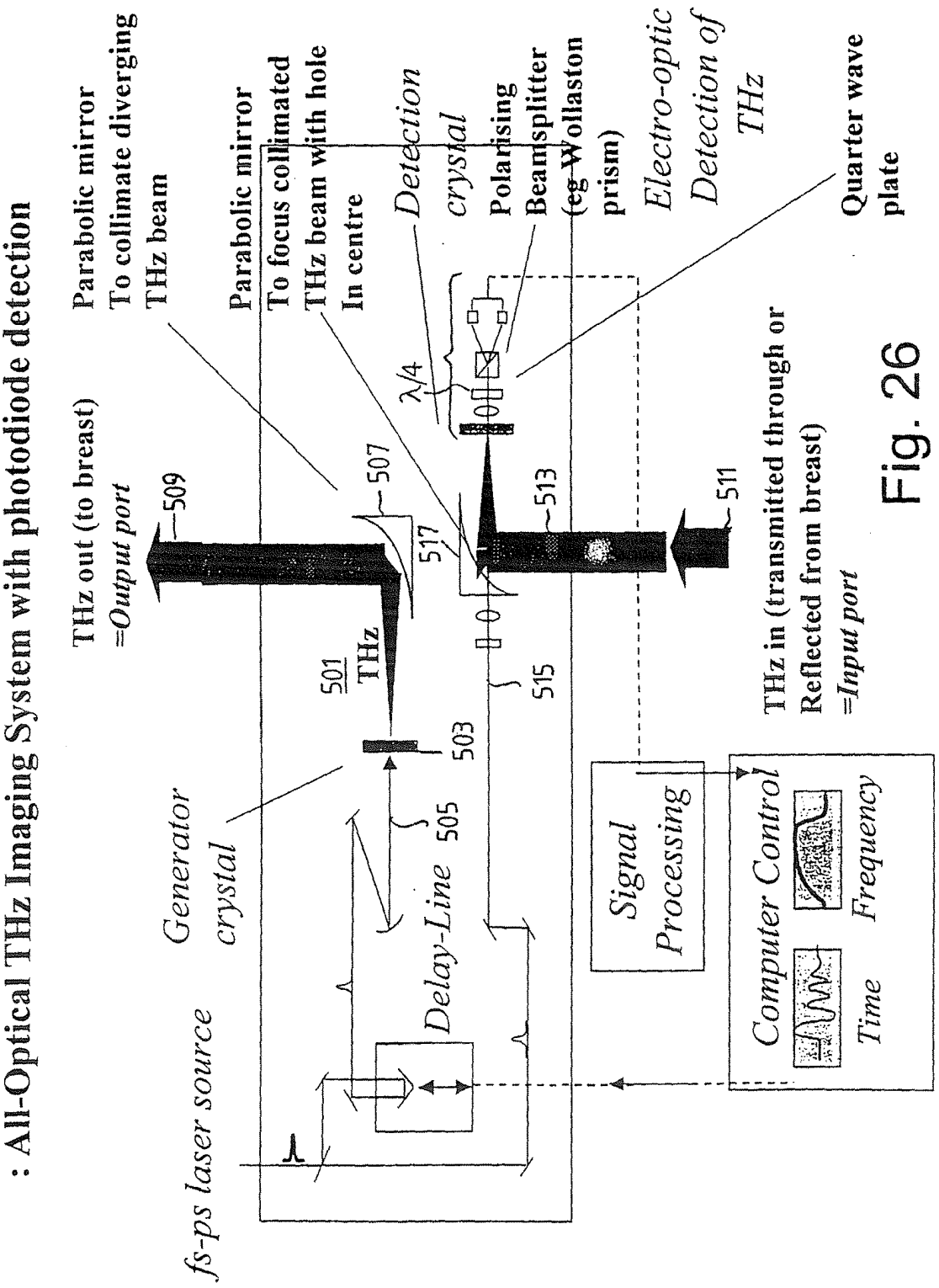


Fig. 26

IPR2023-00033
Apple EX1002 Page 524

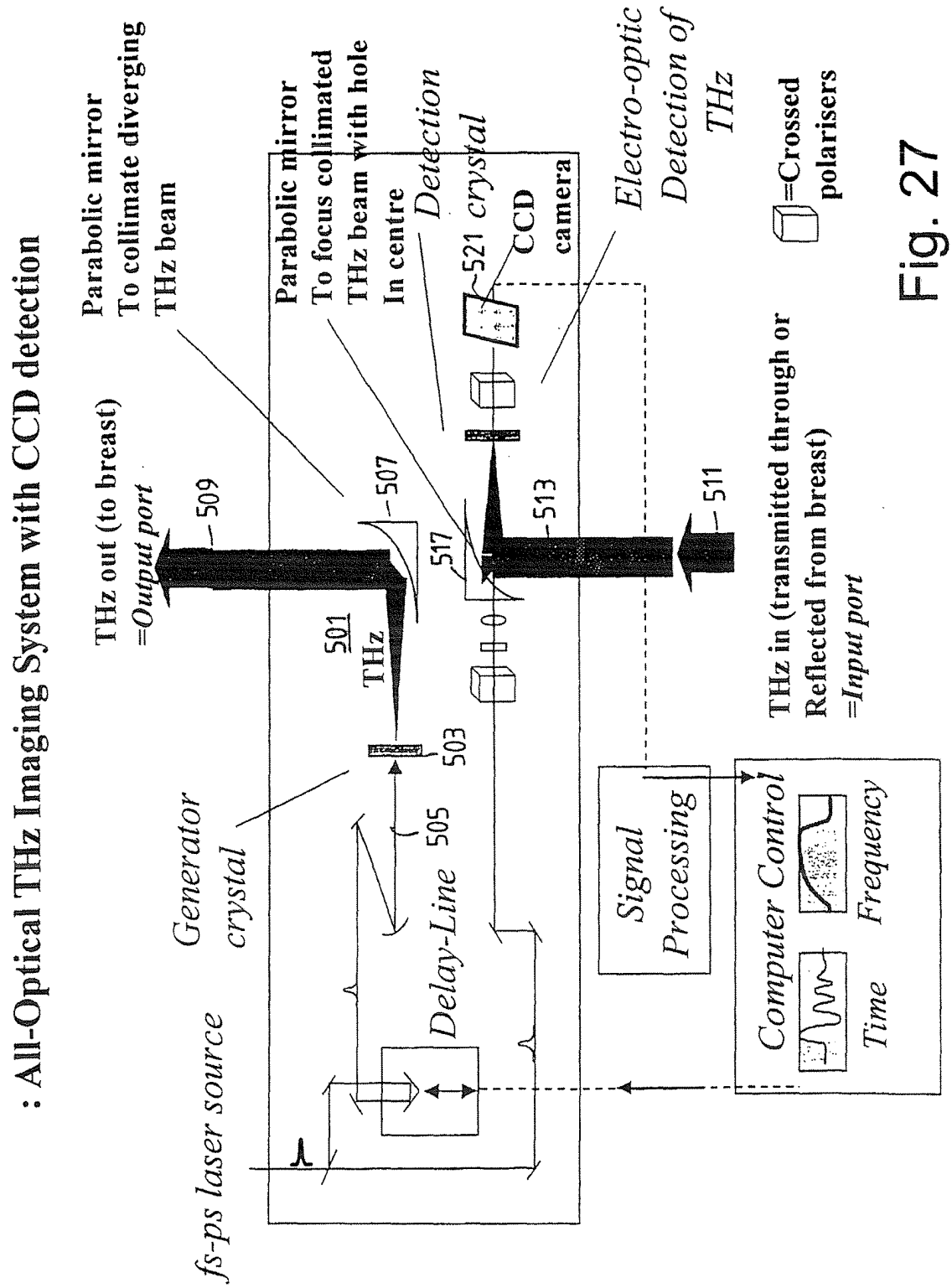


Fig. 27

IPR2023-00033
Apple EX1002 Page 525

: Sample port for THz breast screening
(for use with all-optical systems in Figures 2 and 3)

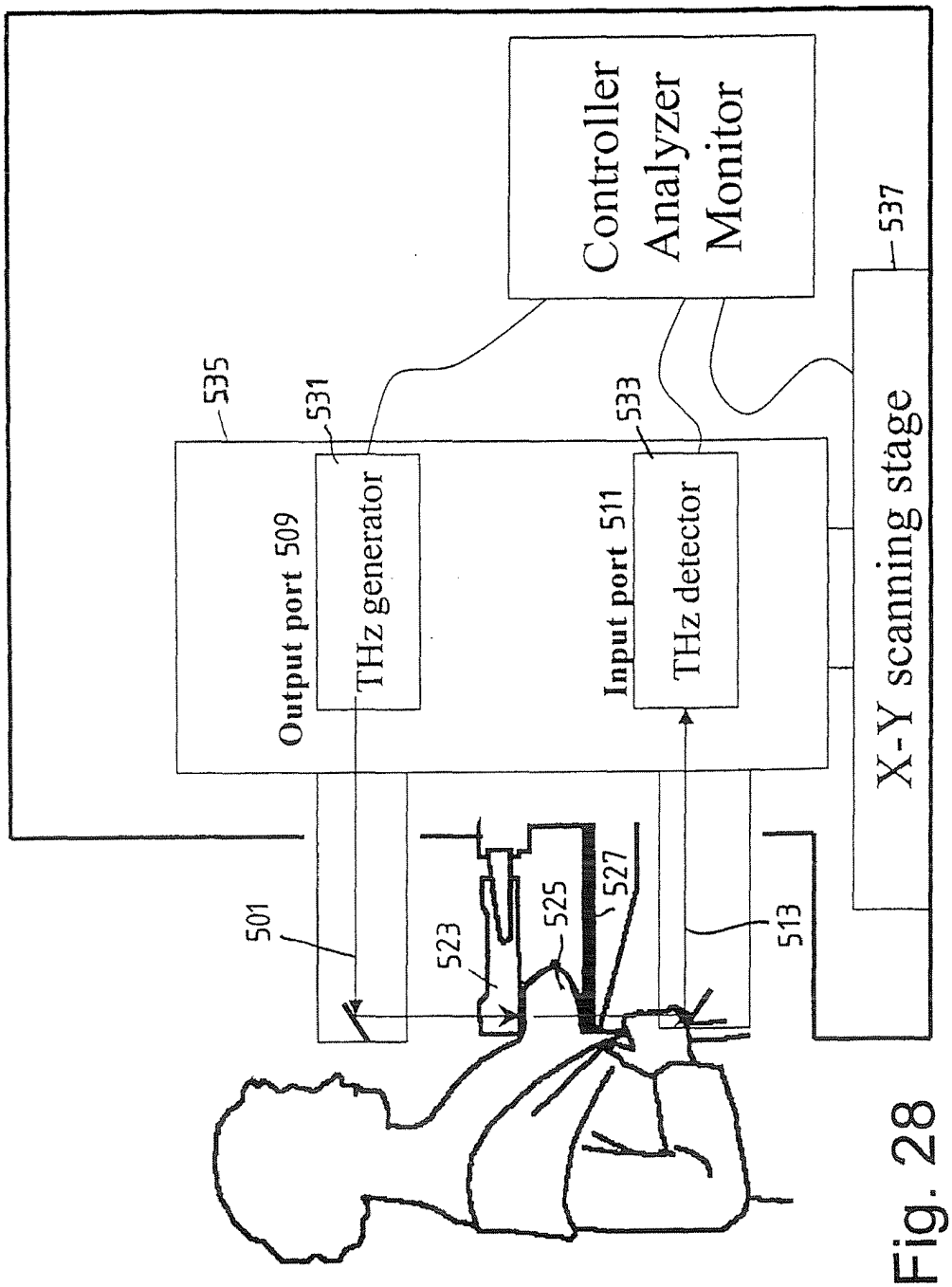


Fig. 28

IPR2023-00033
Apple EX1002 Page 526

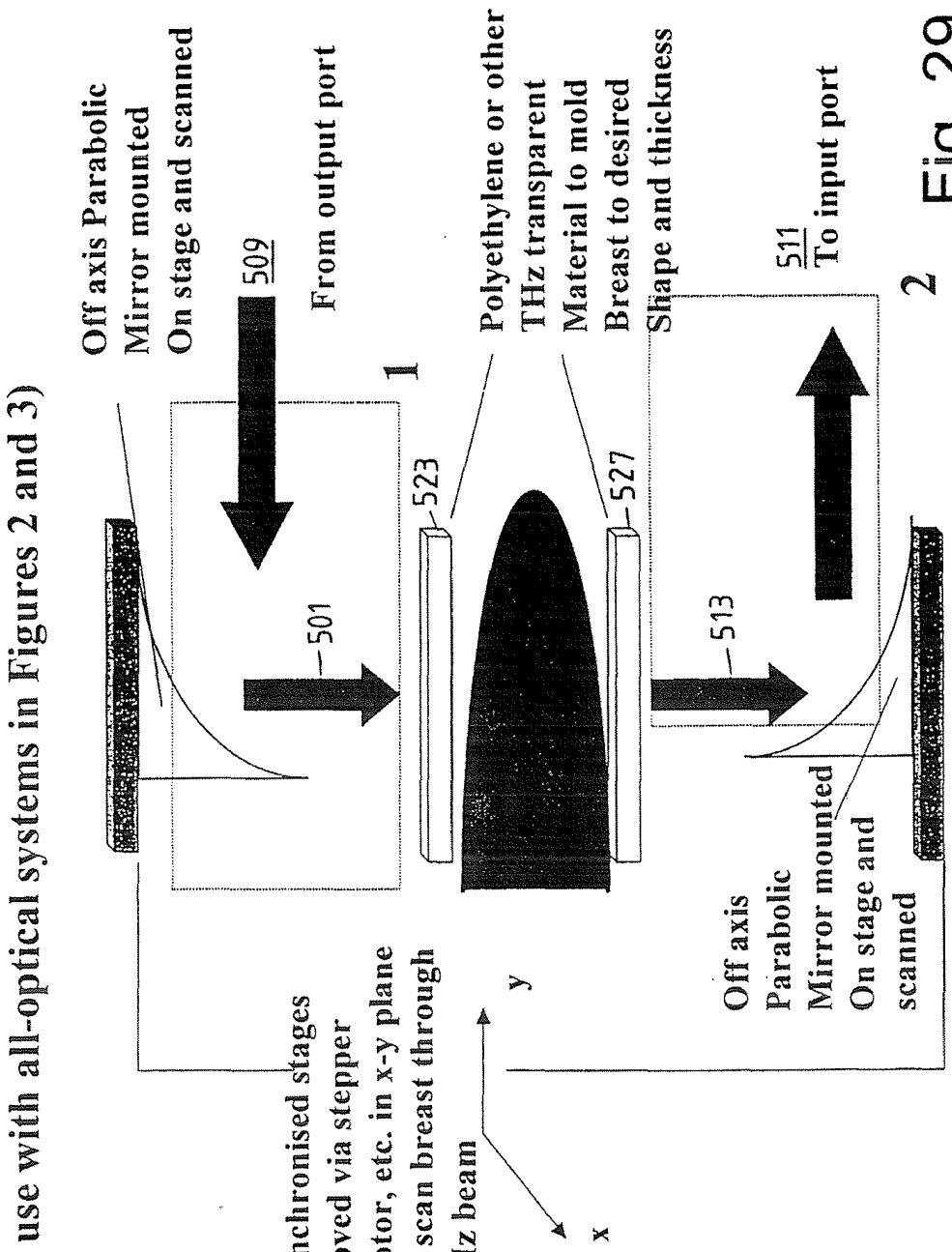


Fig. 29

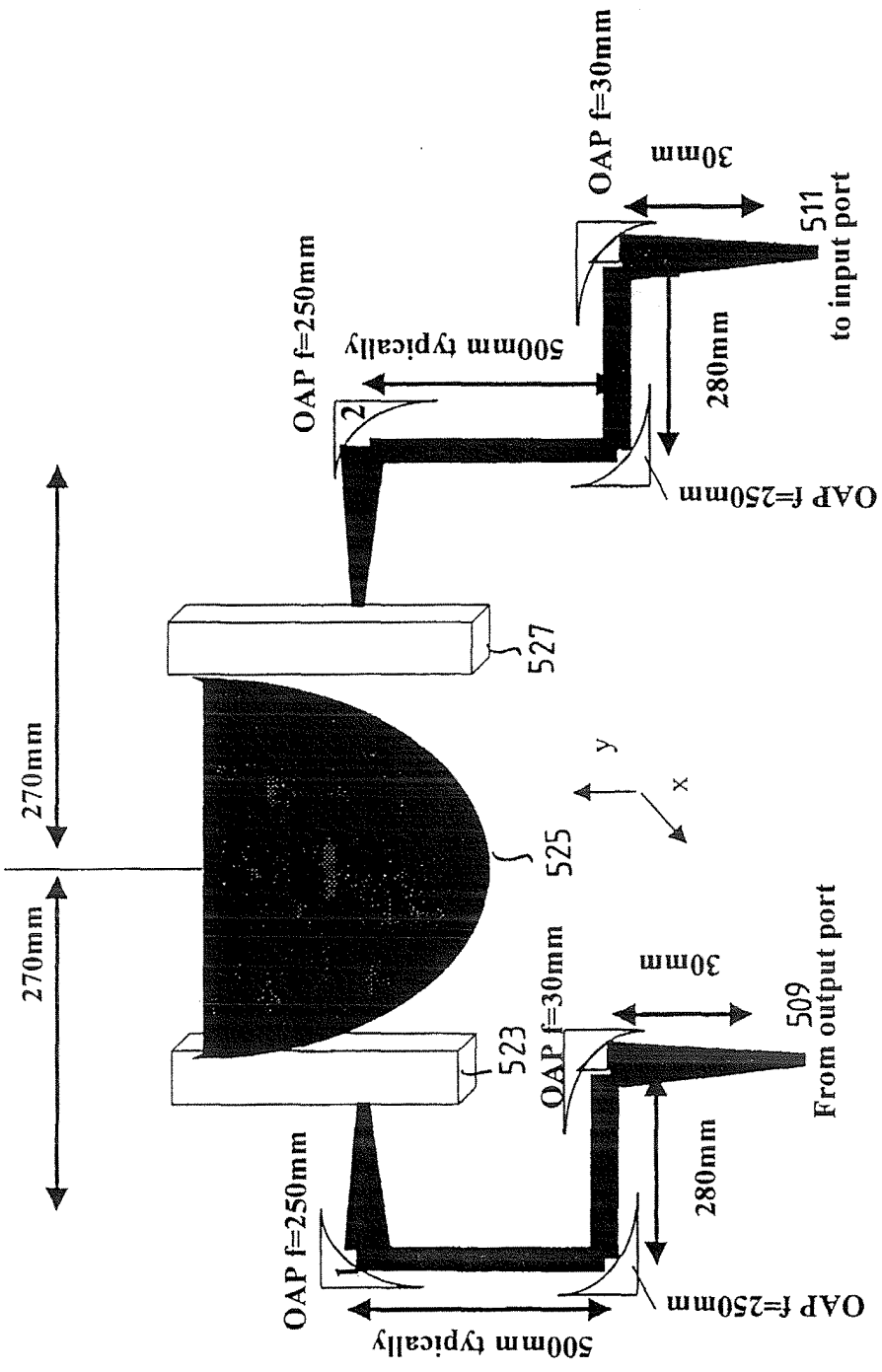
IPR2023-00033
Apple EX1002 Page 527

: Possible alternative configuration of coupling optics used in Figure 5 (regions denoted 1. And 2. In dotted boxes).

All mirrors are off-axis parabolics (OAPs).
OAP's designated by 1. And 2. Are on synchronised X-Y stepper motor stages

Fig. 30

IPR2023-00033
Apple EX1002 Page 528



: Possible alternative configuration of coupling optics used in Figure 5 (regions denoted 1. And 2. In dotted boxes).

All mirrors are off-axis parabolics (OAPs).
OAP's designated by 1. And 2. Are on synchronised X-Y stepper motor stages

Fig. 31

IPR2023-00033
Apple EX1002 Page 529

Fig. 32

IPR2023-00033
Apple EX1002 Page 530

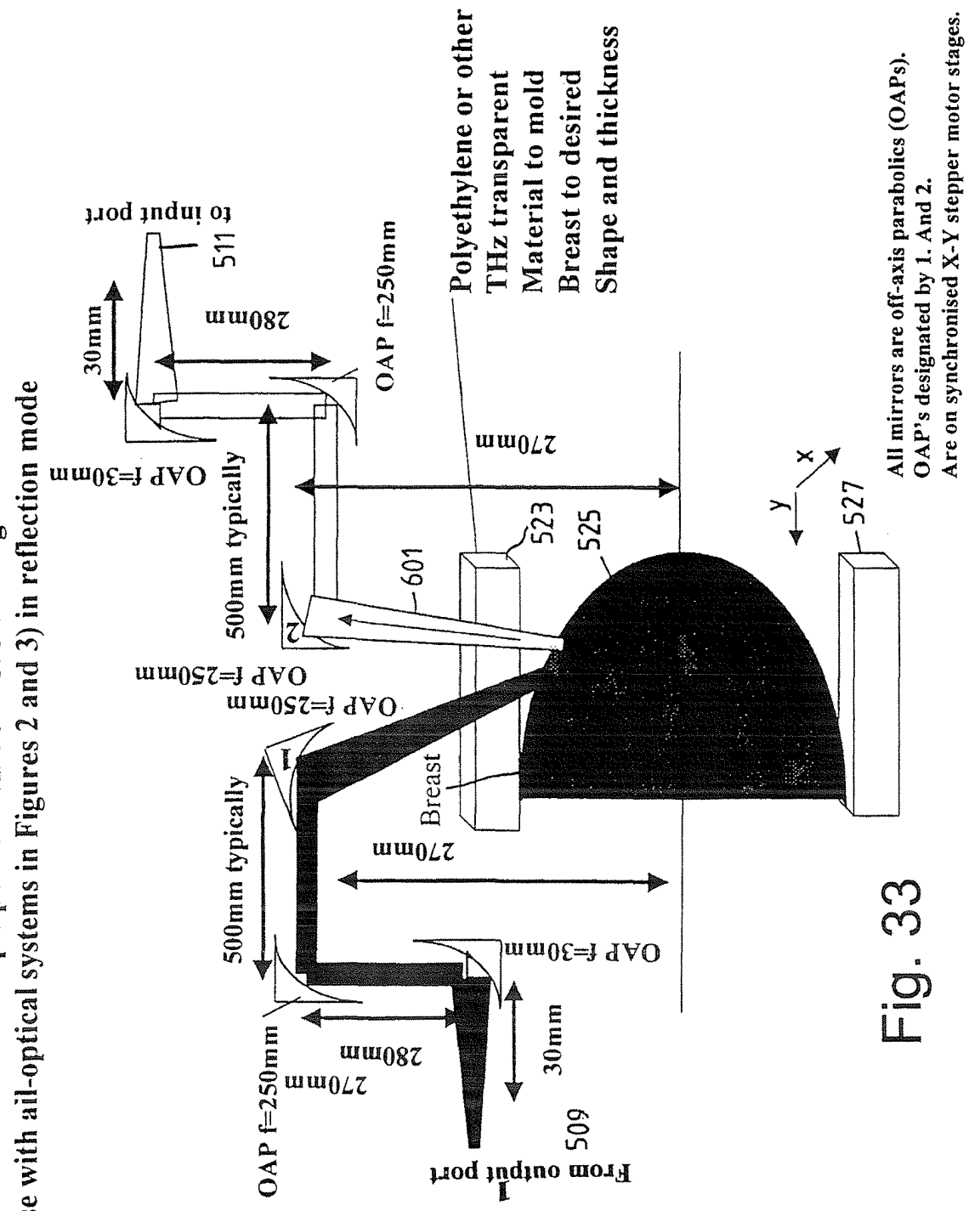


Fig. 33

IPR2023-00033
Apple EX1002 Page 531

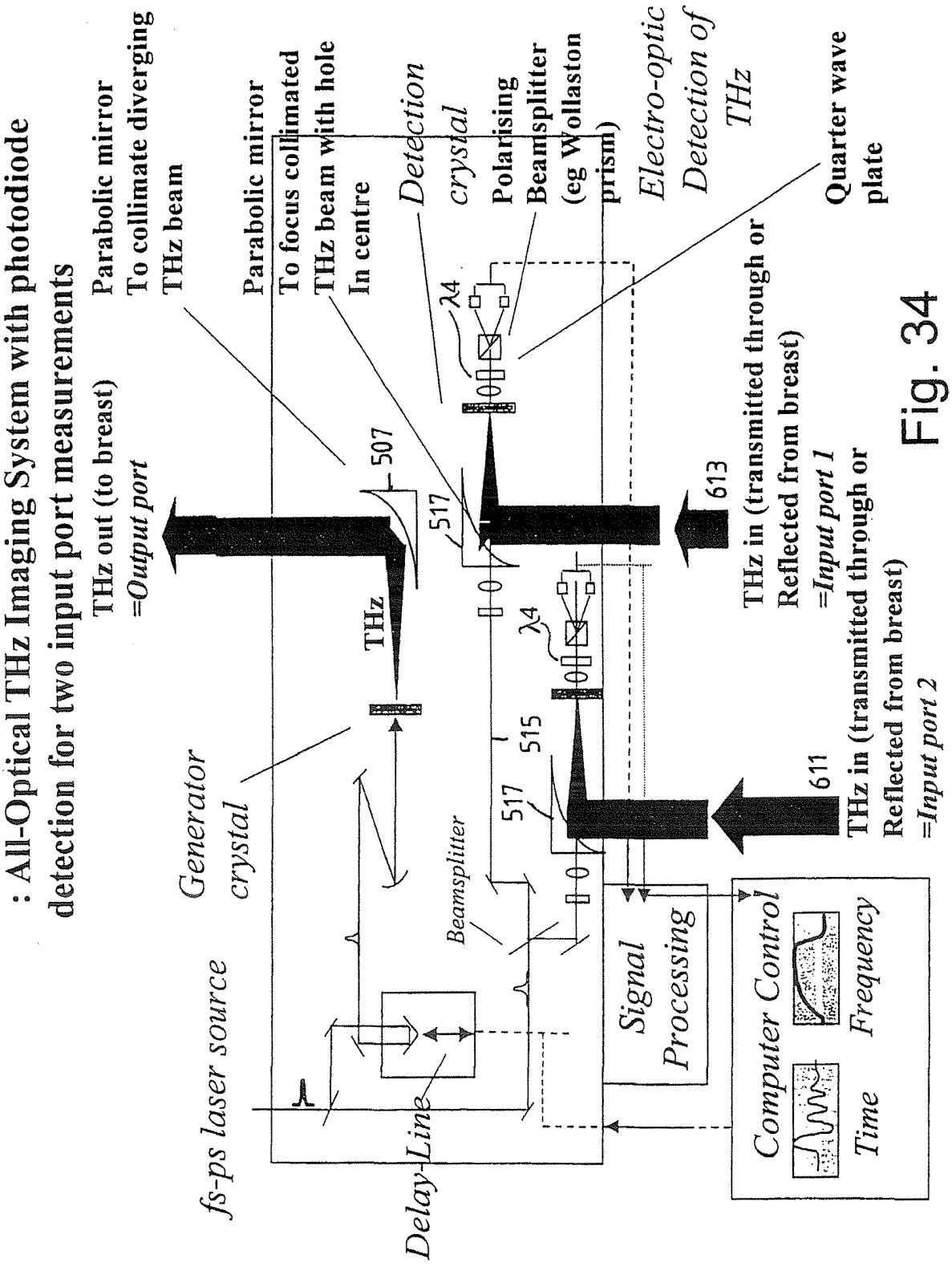


Fig. 34

IPR2023-00033
Apple EX1002 Page 532

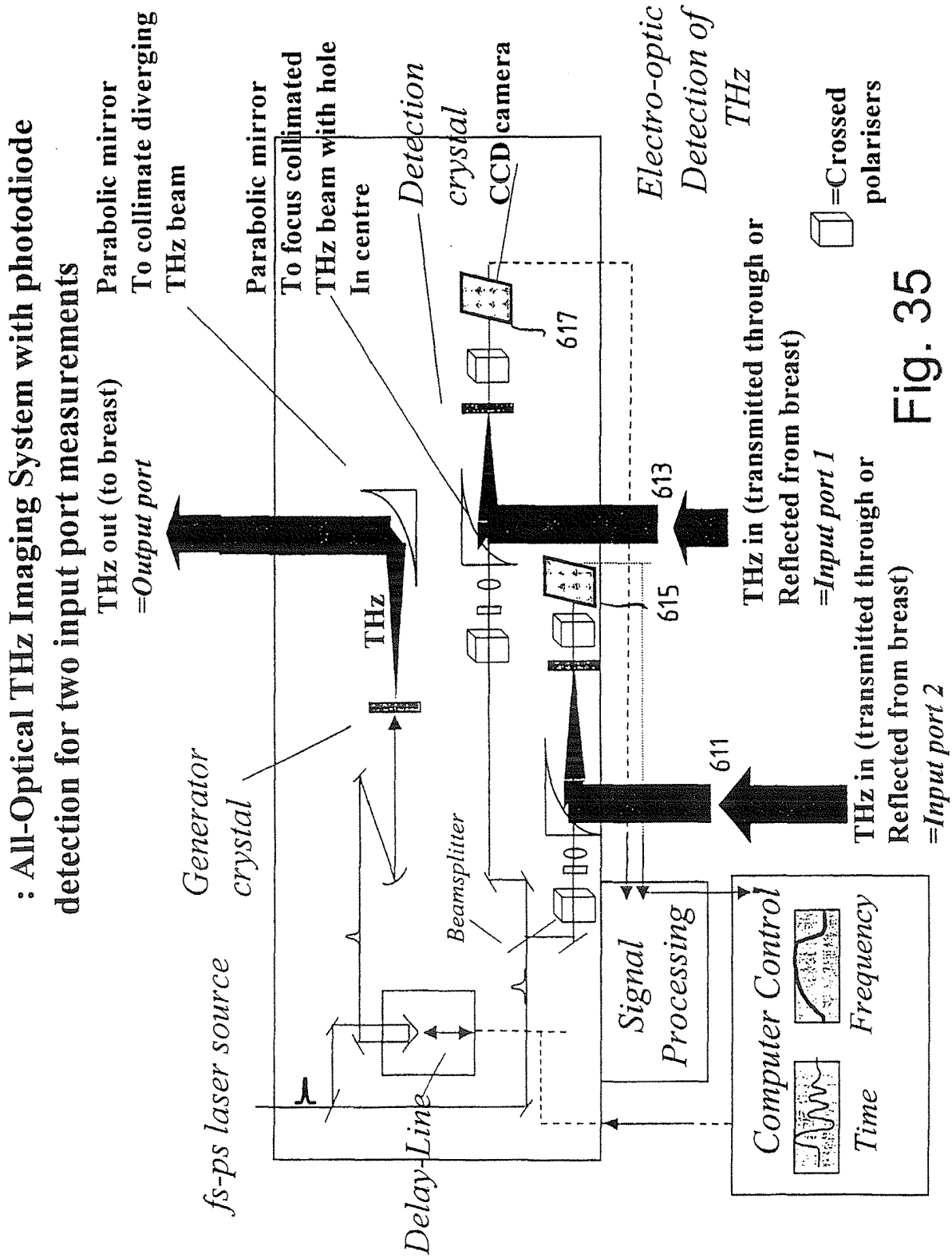


Fig. 35

IPR2023-00033
Apple EX1002 Page 533

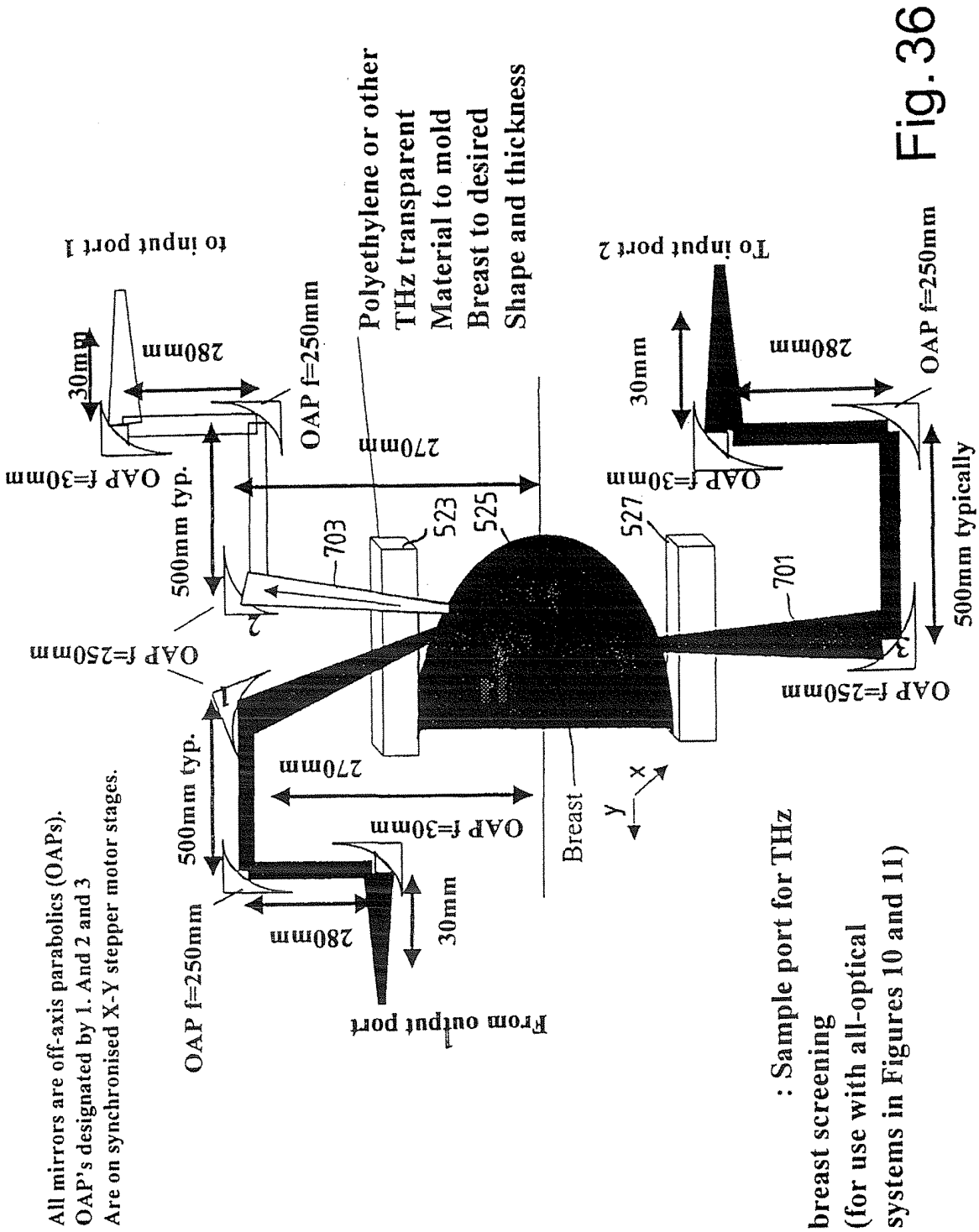


Fig. 36

IPR2023-00033
Apple EX1002 Page 534

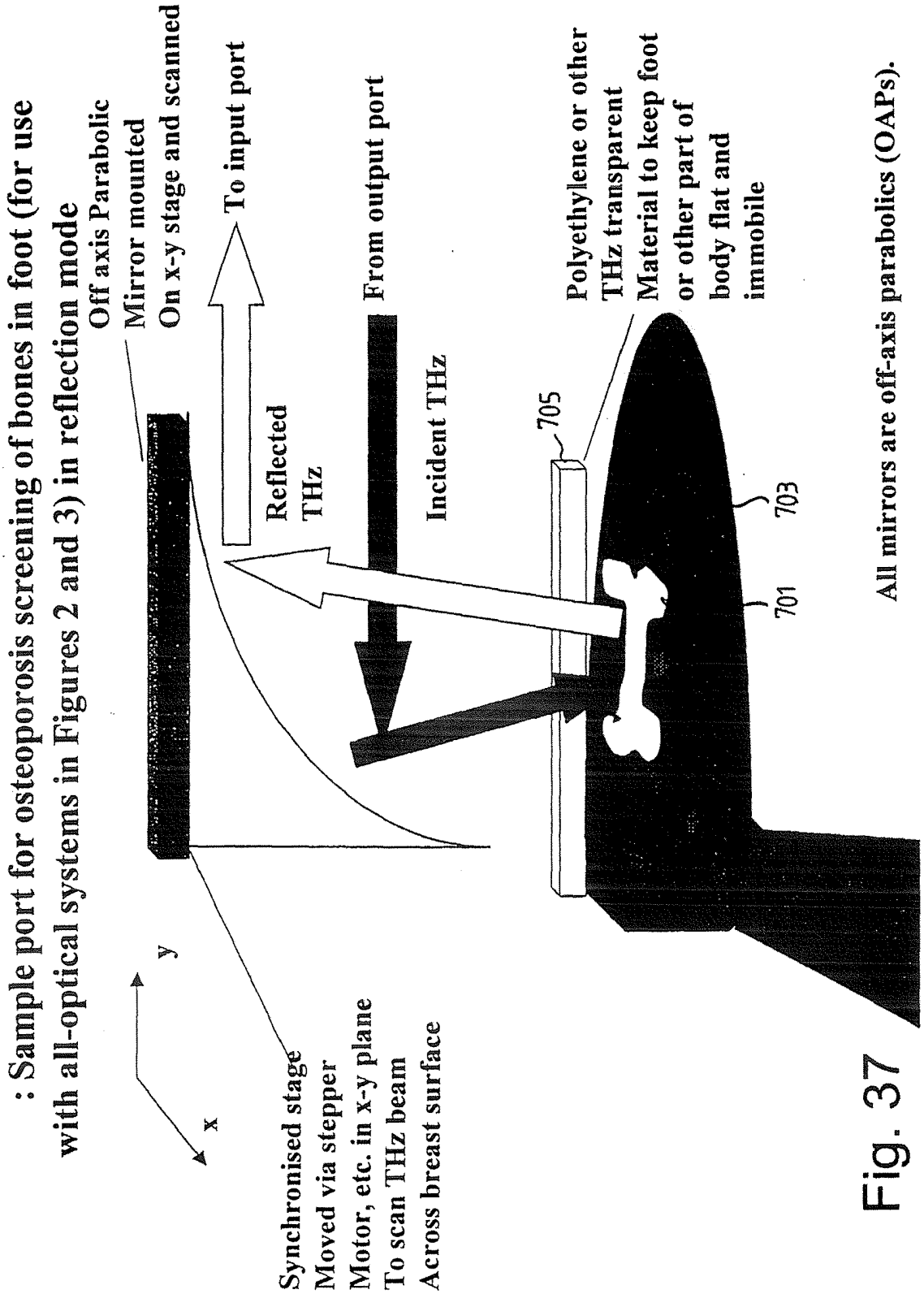


Fig. 37

IPR2023-00033
Apple EX1002 Page 535

: *THz transmission through animal bone*

Example of 2 bones - visible image

(a)

3D THz image through 2 bone sample

(b)

Fig. 38

IPR2023-00033
Apple EX1002 Page 536

: *Medical Imaging using Terahertz light in a human tooth*

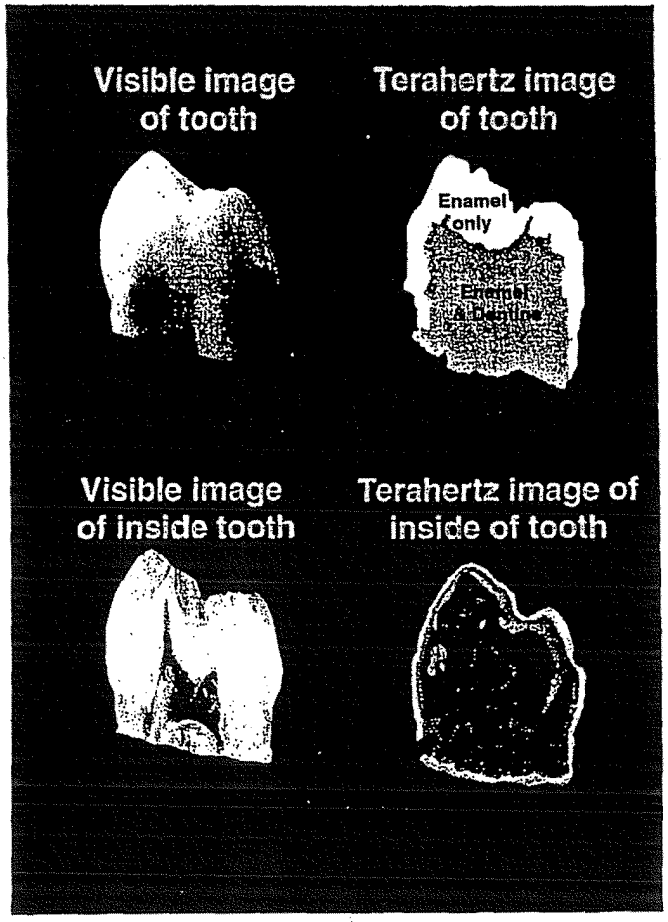


Fig. 39

IPR2023-00033
Apple EX1002 Page 537

: *THz imaging of different types of animal tissue: clotted blood*

THz time domain pulse measured through thin layer (90μm) of clotted blood

THz spectrum measured through thin layer (90μm) of clotted blood

Fig. 40

IPR2023-00033
Apple EX1002 Page 538

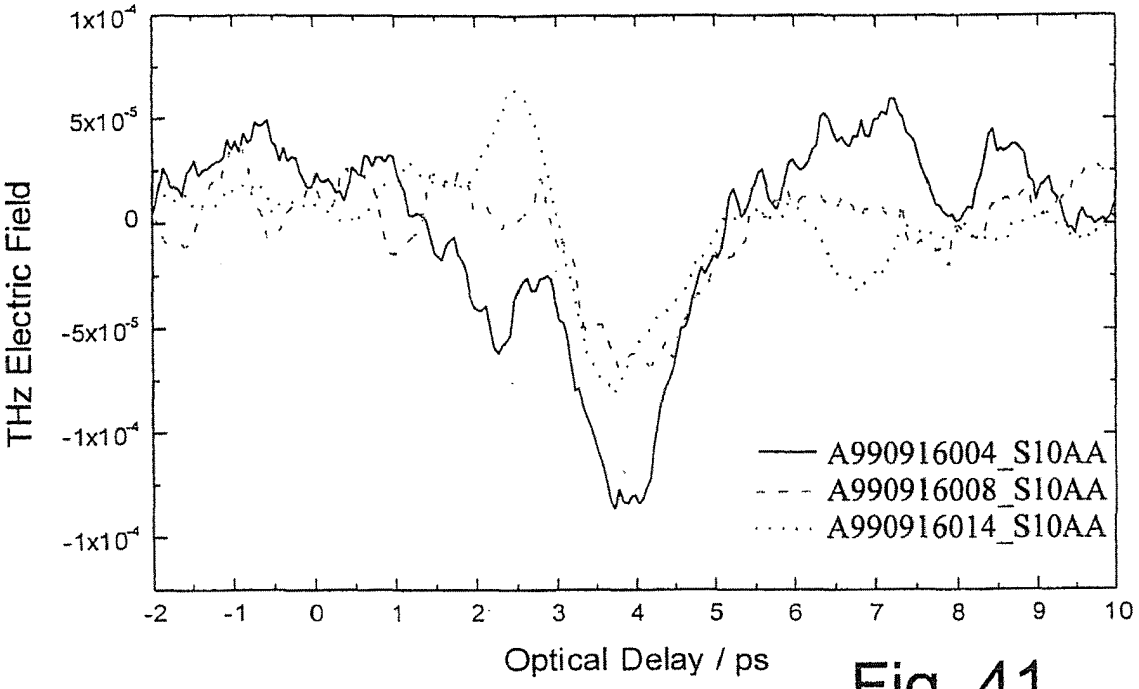


Fig. 41

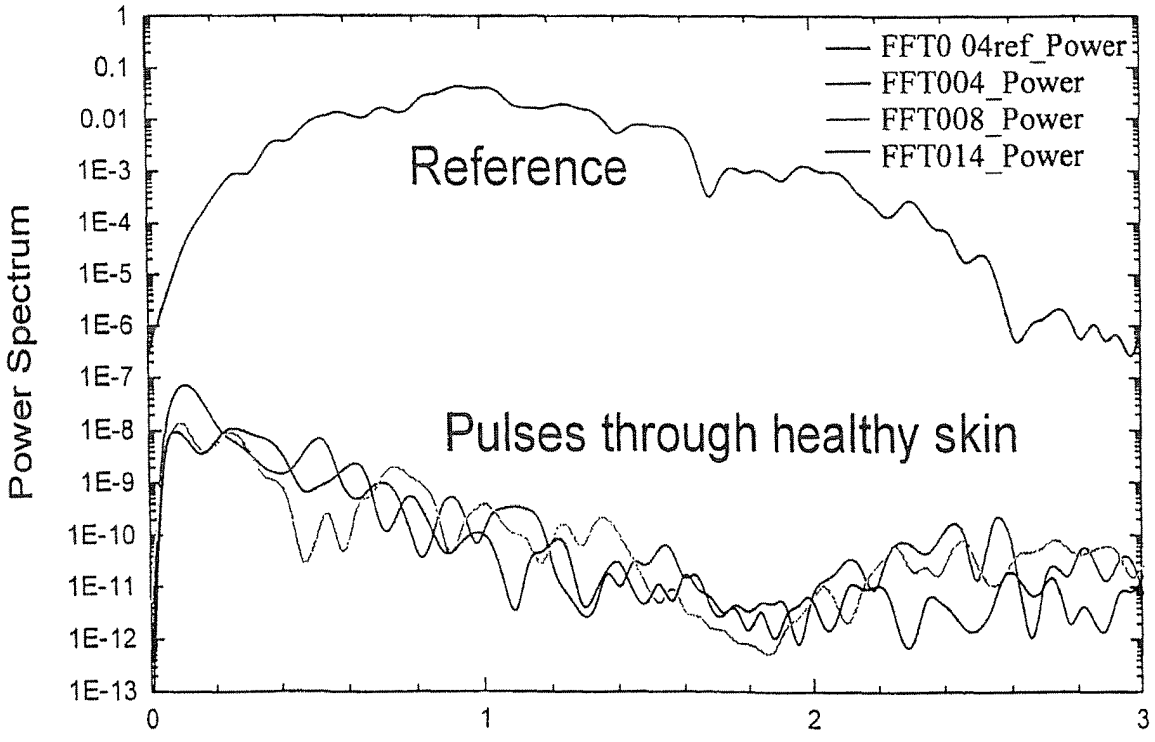


Fig. 42

IPR2023-00033
Apple EX1002 Page 539

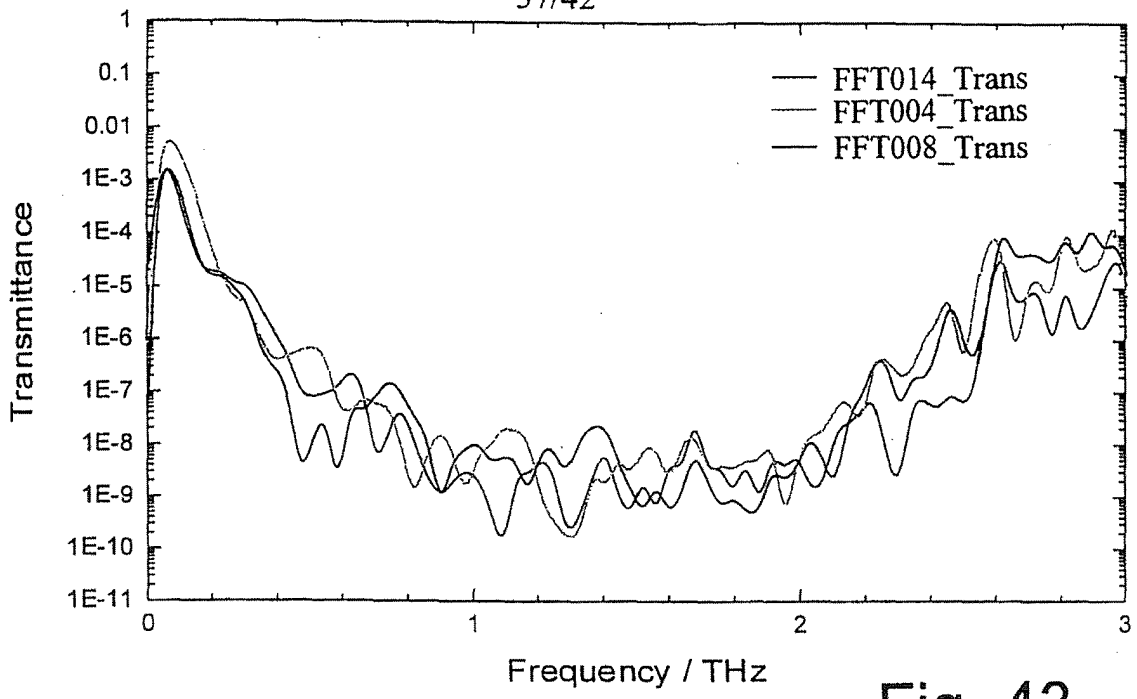


Fig. 43

Transmittance = (signal through skin)/(reference signal)

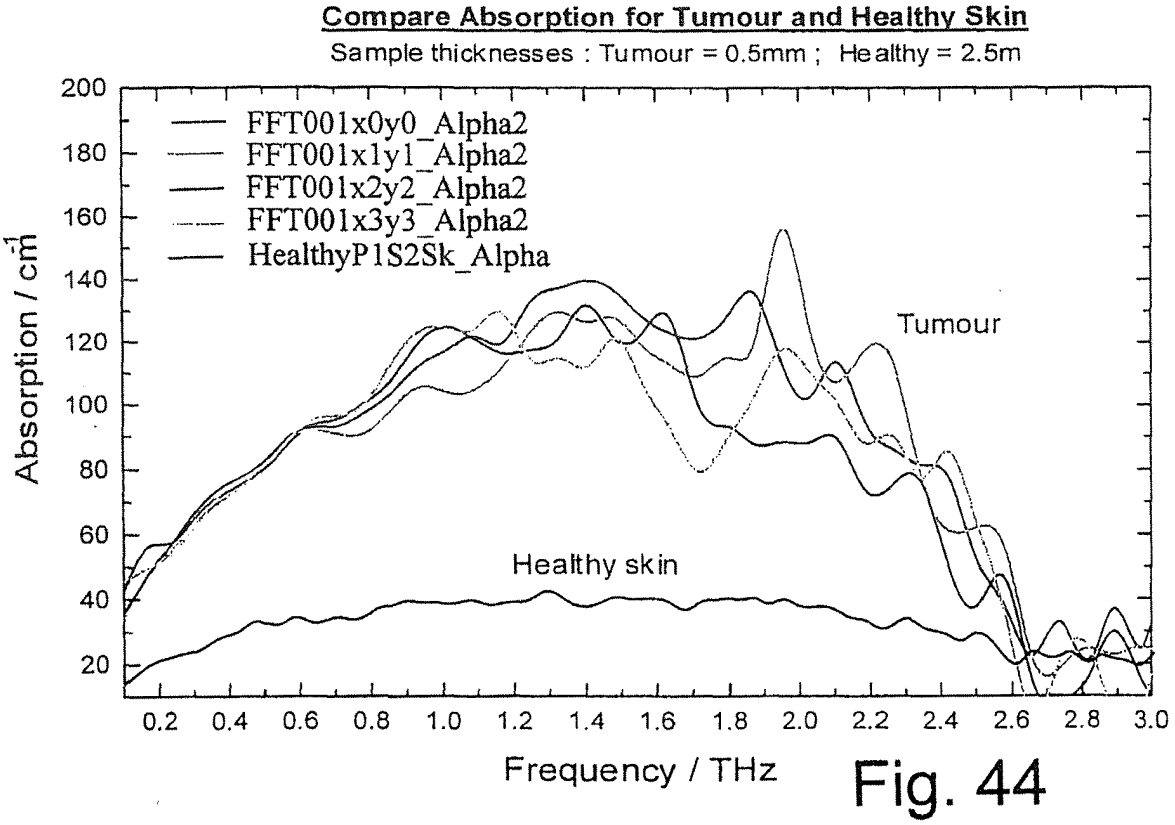


Fig. 44

IPR2023-00033
Apple EX1002 Page 540

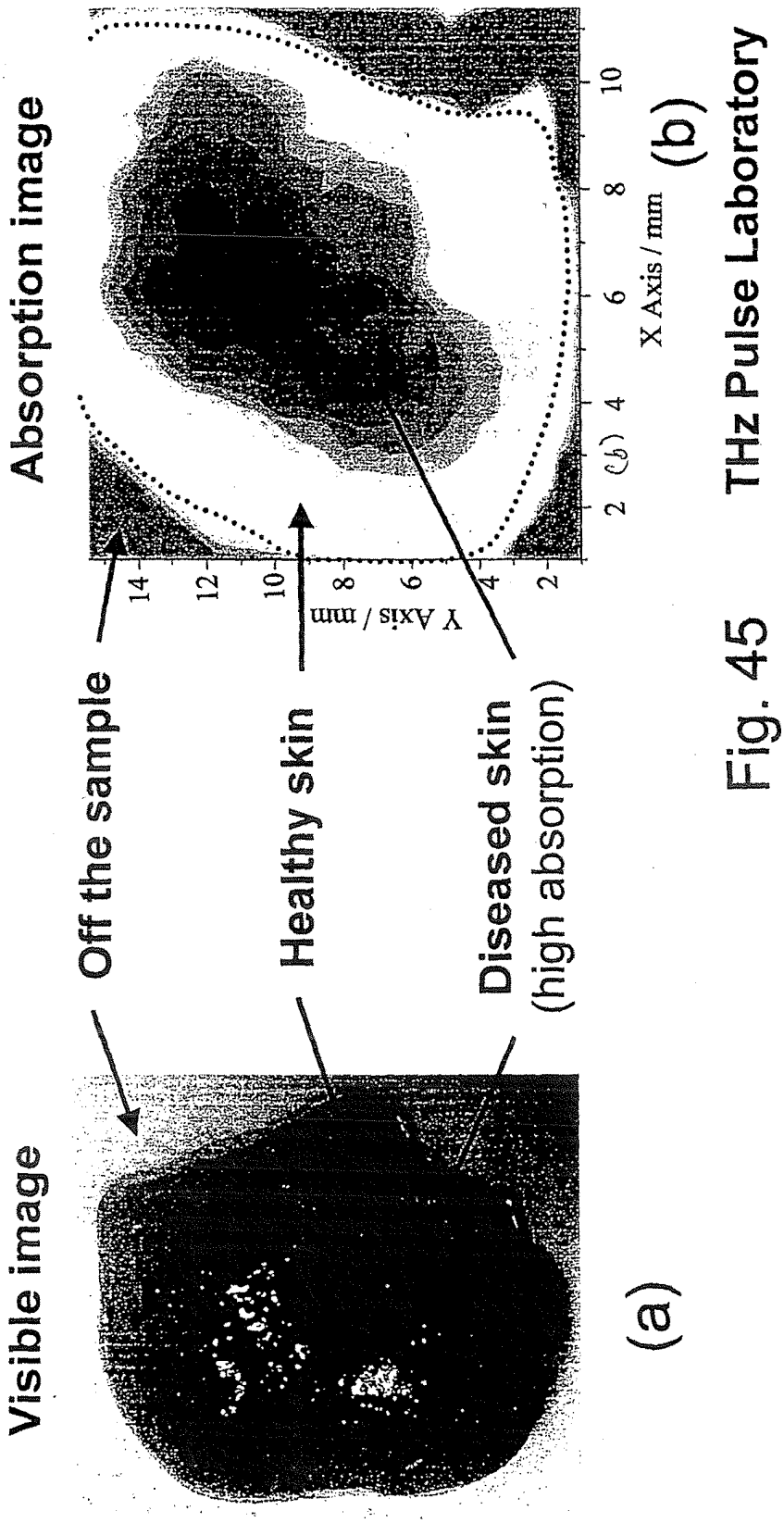


Fig. 45 THz Pulse Laboratory

IPR2023-00033
Apple EX1002 Page 541

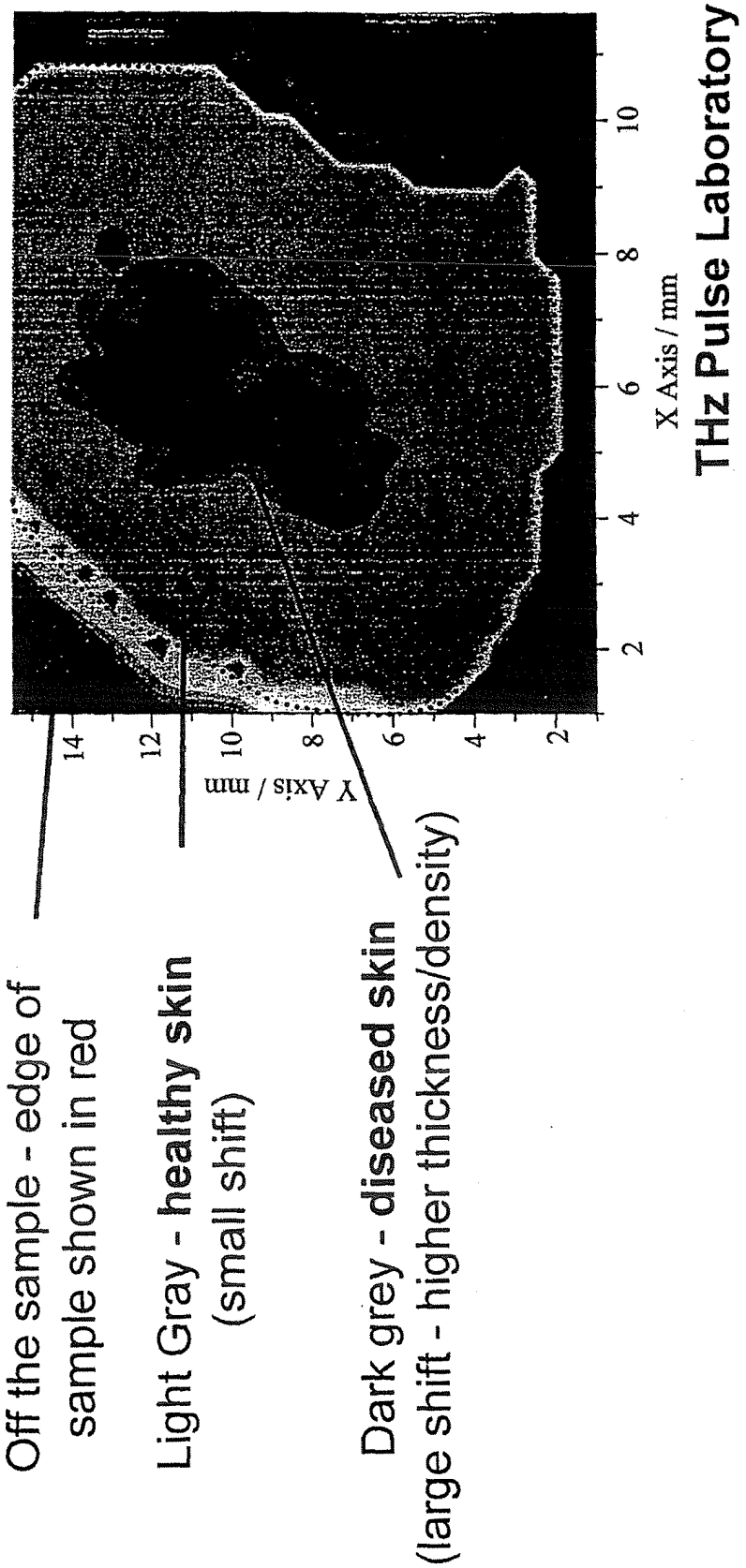


Fig. 46

IPR2023-00033
Apple EX1002 Page 542

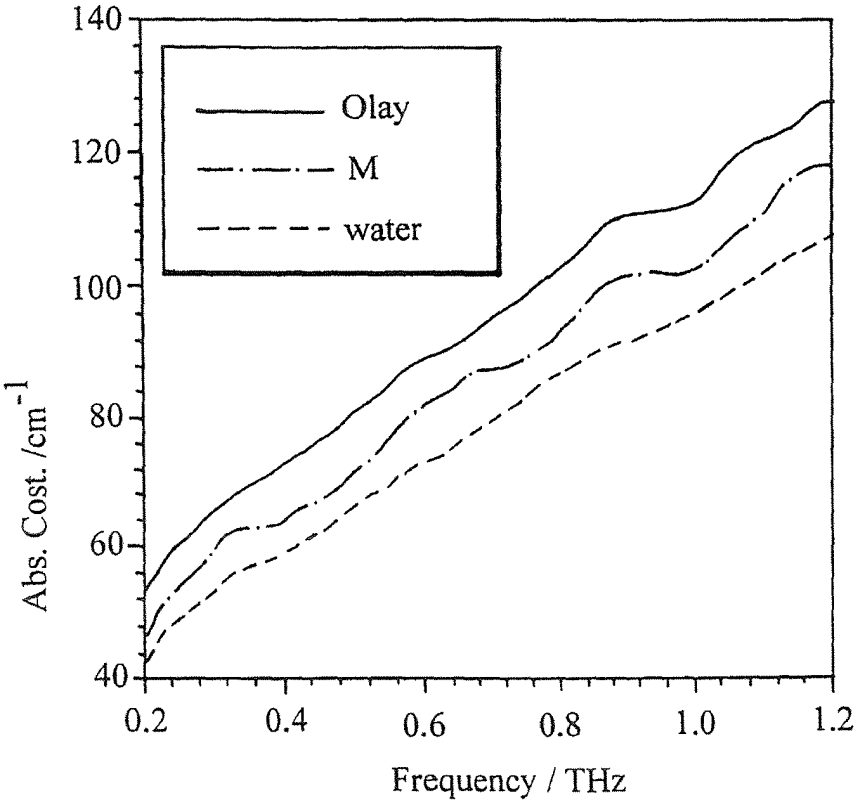


Fig. 47

IPR2023-00033
Apple EX1002 Page 543

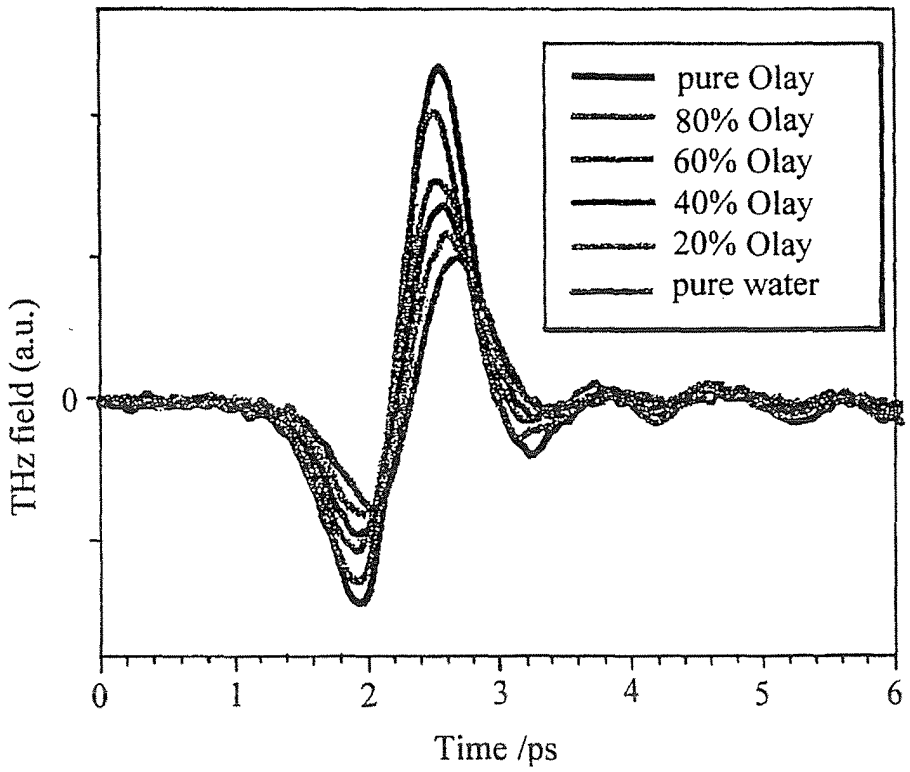


Fig. 48

IPR2023-00033
Apple EX1002 Page 544

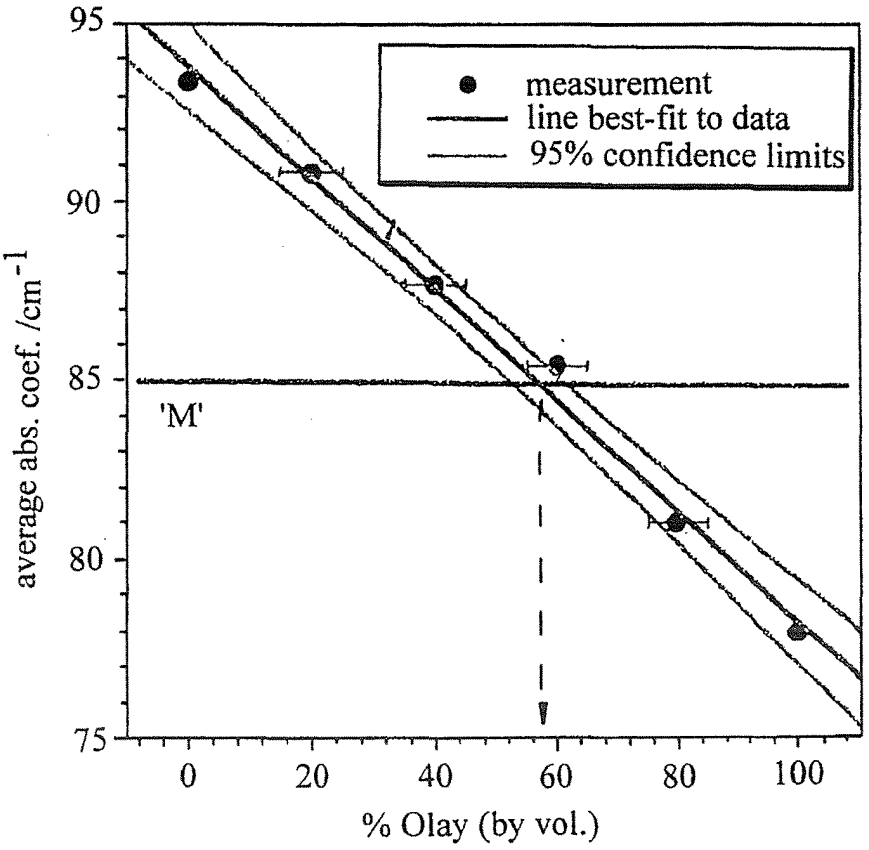


Fig. 49

IPR2023-00033
Apple EX1002 Page 545

INTERNATIONAL SEARCH REPORT

Int. Ional Application No
PCT/GB 00/00632

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G01J3/00 G01J3/28 G01J11/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G01J

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 841 548 A (LUCENT TECHNOLOGIES INC) 13 May 1998 (1998-05-13) claims 1-7	1-5,7,8, 10-21, 30-38, 55-57
X	EP 0 828 143 A (LUCENT TECHNOLOGIES INC) 11 March 1998 (1998-03-11) abstract	1-3
X	EP 0 727 671 A (AT & T CORP) 21 August 1996 (1996-08-21) claims 1-4	1-3
A	EP 0 864 857 A (LUCENT TECHNOLOGIES INC) 16 September 1998 (1998-09-16) claims 1-4	49

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 May 2000	26/05/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	De Buyzer, H

1

Form PCT/ISA/210 (second sheet) (July 1992)

IPR2023-00033
Apple EX1002 Page 546

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/GB 00/00632

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	WU Q ET AL: "TWO-DIMENSIONAL ELECTRO-OPTIC IMAGING OF THZ BEAMS" APPLIED PHYSICS LETTERS,US,AMERICAN INSTITUTE OF PHYSICS. NEW YORK, vol. 69, no. 8, 19 August 1996 (1996-08-19), pages 1026-1028, XP000626128 ISSN: 0003-6951	1
A	PATENT ABSTRACTS OF JAPAN vol. 13, no. 182 (C-591), 27 April 1989 (1989-04-27) & JP 01 011530 A (HITACHI), 17 January 1989 (1989-01-17) abstract	45
A	GB 2 254 417 A (JOUZA BIJAN) 7 October 1992 (1992-10-07) abstract	45,62

1

IPR2023-00033
Apple EX1002 Page 547

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter ...nal Application No
PCT/GB 00/00632

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0841548 A	13-05-1998	US 5939721 A	17-08-1999
		JP 10153547 A	09-06-1998
EP 0828143 A	11-03-1998	US 5789750 A	04-08-1998
		JP 10104171 A	24-04-1998
EP 0727671 A	21-08-1996	US 5623145 A	22-04-1997
		JP 8320254 A	03-12-1996
		US 5710430 A	20-01-1998
EP 0864857 A	16-09-1998	JP 11108845 A	23-04-1999
JP 01011530 A	17-01-1989	NONE	
GB 2254417 A	07-10-1992	NONE	

IPR2023-00033
Apple EX1002 Page 548

Electronic Patent Application Fee Transmittal

Application Number:	10503463
Filing Date:	24-May-2005
Title of Invention:	Speech recognition circuit using parallel processors
First Named Inventor/Applicant Name:	Mark Catchpole
Filer:	Mark J. Thronson/Rupal Patel
Attorney Docket Number:	M0025.0311/P311

Filed as Large Entity

U.S. National Stage under 35 USC 371 Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
Submission- Information Disclosure Stmt	1806	1	180	180
		Total in USD ($)		**180**

Electronic Acknowledgement Receipt

EFS ID:	4727815
Application Number:	10503463
International Application Number:	
Confirmation Number:	8875
Title of Invention:	Speech recognition circuit using parallel processors
First Named Inventor/Applicant Name:	Mark Catchpole
Customer Number:	24998
Filer:	Mark J. Thronson/Rupal Patel
Filer Authorized By:	Mark J. Thronson
Attorney Docket Number:	M0025.0311/P311
Receipt Date:	03-FEB-2009
Filing Date:	24-MAY-2005
Time Stamp:	16:59:56
Application Type:	U.S. National Stage under 35 USC 371

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$180
RAM confirmation Number	2301
Deposit Account	
Authorized User	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)

IPR2023-00033
Apple EX1002 Page 551

1		ids1.pdf	15858938 00c723a8f192bb2f0e7279ff48463ac775aec8aa	yes	233

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Information Disclosure Statement Letter	1	2
Information Disclosure Statement (IDS) Filed (SB/08)	3	4
NPL Documents	5	8
Foreign Reference	9	233

Warnings:

Information:

2		ids2.pdf	21071316 e8e0231f195ff3dc6106ecf5274e8296dead05ff	yes	183

Multipart Description/PDF files in .zip description		
Document Description	**Start**	**End**
Foreign Reference	1	101
NPL Documents	102	183

Warnings:

Information:

3	Fee Worksheet (PTO-06)	fee-info.pdf	29960 c52f27a9105800ad8d925b0901be41abfaf93374	no	2

Warnings:

Information:

Total Files Size (in bytes):	36960214

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

24998 7590 05/11/2009

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403

EXAMINER	
ABEBE, DANIEL DEMELASH	
ART UNIT	PAPER NUMBER
2626	

DATE MAILED: 05/11/2009

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/503,463	05/24/2005	Mark Catchpole	M0025.0311/P311	8875

TITLE OF INVENTION: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$755	$300	$0	$1055	08/11/2009

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.

If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:

A. If the status is the same, pay the TOTAL FEE(S) DUE shown above.

B. If the status above is to be removed, check box 5b on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and twice the amount of the ISSUE FEE shown above, or

If the SMALL ENTITY is shown as NO:

A. Pay TOTAL FEE(S) DUE shown above, or

B. If applicant claimed SMALL ENTITY status before, or is now claiming SMALL ENTITY status, check box 5a on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and 1/2 the ISSUE FEE shown above.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

PTOL-85 (Rev. 08/07) Approved for use through 08/31/2010.

IPR2023-00033
Apple EX1002 Page 554

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

24998 7590 05/11/2009

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/503,463	05/24/2005	Mark Catchpole	M0025.0311/P311	8875

TITLE OF INVENTION: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$755	$300	$0	$1055	08/11/2009

EXAMINER	ART UNIT	CLASS-SUBCLASS
ABEBE, DANIEL DEMELASH	2626	704-242000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list

(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,

(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 _____

2 _____

3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. The following fee(s) are submitted:

❑ Issue Fee

❑ Publication Fee (No small entity discount permitted)

❑ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**

❑ A check is enclosed.

❑ Payment by credit card. Form PTO-2038 is attached.

❑ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)

❑ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.

❑ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

IPR2023-00033
Apple EX1002 Page 555

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/503,463	05/24/2005	Mark Catchpole	M0025.0311/P311	8875

24998 7590 05/11/2009

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403

EXAMINER	
ABEBE, DANIEL DEMELASH	
ART UNIT	PAPER NUMBER
2626	

DATE MAILED: 05/11/2009

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment to date is 823 day(s). If the issue fee is paid on the date that is three months after the mailing date of this notice and the patent issues on the Tuesday before the date that is 28 weeks (six and a half months) after the mailing date of this notice, the Patent Term Adjustment will be 823 day(s).

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

PTOL-85 (Rev. 08/07) Approved for use through 08/31/2010.

IPR2023-00033
Apple EX1002 Page 556

Notice of Allowability	Application No.	Applicant(s)	
	10/503,463	CATCHPOLE, MARK	
	Examiner	Art Unit	
	Daniel D. Abebe	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *1/29/2009*.

2. ☒ The allowed claim(s) is/are *1-44,59 and 68-91*.

3. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☐ All b) ☐ Some* c) ☐ None of the:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
3. ☒ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
7. ☐ Examiner's Amendment/Comment
8. ☒ Examiner's Statement of Reasons for Allowance
9. ☐ Other _____.

IPR2023-00033
Apple EX1002 Page 557

Allowable Subject Matter

Claims 1-44, 59 and 68-91 are allowed.

The following is an examiner's statement of reasons for allowance: the application is allowed as the result of the amendment filed on 1/29/2009 cancelling all the previously rejected claims.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Daniel D. Abebe whose telephone number is 571-272-7615. The examiner can normally be reached on monday-friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on 571-272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Daniel D Abebe/
Primary Examiner, Art Unit 2626

Used in Lieu of PTO/SB/08A/B
(Based on PTO 10-07 version)

Substitute for form 1449/PTO				*Complete if Known*	
INFORMATION DISCLOSURE STATEMENT BY APPLICANT				**Application Number**	10/503,463-Conf. #8875
				Filing Date	May 24, 2005
				First Named Inventor	Mark Catchpole
				Art Unit	2626
(Use as many sheets as necessary)				Examiner Name	D. D. Abebe
Sheet	1	of	2	Attorney Docket Number	M0025.0311/P311

U.S. PATENT DOCUMENTS					
Examiner Initials*	Cite No.[1]	Document Number Number-Kind Code[2] (*if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
	AA	5,881,312	03-09-1999	Dulong	
	AB	7,174,037 B2	02-06-2007	Arnone et al.	
	AC	5,293,213	03-08-1994	Klein et al.	
	AD	2003/0178584 A1	09-25-2003	Arnone et al.	
	AE	2008/0255839 A1	10-16-2008	Larri et al.	

FOREIGN PATENT DOCUMENTS						
Examiner Initials*	Cite No.[1]	Foreign Patent Document Country Code[3]-Number[4]-Kind Code[5] *(if known)*	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages Or Relevant Figures Appear	T[6]
	BA	WO 03/067572 A2	08-14-2003			
	BB	GB 2 331 392 A	05-19-1999			
	BC	GB 2 112 194 A	07-13-1983			
	BD	GB 2 380 920 A	04-16-2003			
	BE	GB 2 347 835 A	09-13-2000			
	BF	WO 00/50859	08-31-2000			

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. * CITE NO.: Those application(s) which are marked with an single asterisk (*) next to the Cite No. are not supplied (under 37 CFR 1.98(a)(2)(iii)) because that application was filed after June 30, 2003 or is available in the IFW. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	CA	S. Glinski et al., "Spoken Language Recognition on a DSP Array Processor", IEEE Transactions on Parallel and Distributed Systems, July 5, 1994, No. 7, New York, USA, pp. 697-703.	
	CB	S. Chatterjee et al., "Connected Speech Recognition on a Multiple Processor Pipeline", ICASSP 89, May 23, 1989, Glasgow, UK, pp. 774-777.	
	CC	S-H. Chung et al., "A Parallel Phoneme Recognition Algorithm Based on Continuous Hidden Markov Model", Proceedings of 13th International Parallel Processing Symposium and 10th Symposium on Parallel and Distributed Processing, 1999, Los Alamitos, CA, USA, pp. 453-457.	
	CD	S-H. Chung et al., "A Parallel Computational Model for Integrated Speech and Natural Language Understanding", IEEE Transactions on Computers, New York, USA, Vol. 42, No. 10, October 1, 1993, pp. 1171-1183.	

Examiner Signature	/Daniel Abebe/ (05/01/2009)	Date Considered	

IPR2023-00033
ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /D.A./
Apple EX1002 Page 560

Used in Lieu of PTO/SB/08A/B
(Based on PTO 10-07 version)

Substitute for form 1449/PTO				Complete if Known	
INFORMATION DISCLOSURE STATEMENT BY APPLICANT				Application Number	10/503,463-Conf. #8875
				Filing Date	May 24, 2005
				First Named Inventor	Mark Catchpole
				Art Unit	2626
(Use as many sheets as necessary)				Examiner Name	D. D. Abebe
Sheet	2	of	2	Attorney Docket Number	M0025.0311/P311

	CE	N. Deshmukh et al., "Hierarchical Search for Large-Vocabulary Conversational Speech Recognition: Working Toward a Solution for the Decoding Problem", IEEE Signal Processing Magazine, Vol. 16, No. 5, September 1999, pp. 84-107.	
	CF	R.M. Woodward et al., " Tissue Classification Using Terahertz Pulsed Imaging", Proceedings of the SPIE - The International Society for Optical Engineering SPIE-INT. SOC. OPT. ENG. USA, Vol. 5318, No.1, 2004, pp. 23-33. Abstract also attached.	
	CG	V.P. Wallace et al., "Biomedical Applications of THz Imaging", Microwave Symposium Digest, 2004, IEEE MTT-S International Fort Worth, TX, USA, June 6-11, 2004, Piscataway, NJ, USA, IEEE, Vol. 3, 6 June 2004 (2004 06-06), pp. 1579-1581	
	CH	V.P. Wallace et al., "Biomedical Applications of Terahertz Pulse Imaging", Second Joint EMBS-BMES Conference 2002. Conference Proceedings. 24th Annual International Conference of the Engineering in Medicinal and Biology Society. Annual Fall Meeting of the Biomedical Engineering Society, Houston, TX, Oct. 23-26, 2002, Annual Vol. 1 of 3. Conf. 24, 23 October 2002 (2002-10-23), pp. 2333-2334.	
	CI	A.J. Fitzgerald et al., "Terahertz Imaging of Breast Cancer, a Feasibility Study", Conference Digest of the 2004 Joint 29th International Conference on Infrared and Millimeter Waves and 12th International Conference on Terahertz Electronics IEEE Piscataway, NJ, USA, 2004, pp. 823-824.	
	CJ	B.E. Cole et al., "Terahertz Imaging and Spectroscopy of Human Skin, In-vivo", Proceedings of SPIE, Vol. 4276, pp. 1-10.	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1]Applicant's unique citation designation number (optional). [2]Applicant is to place a check mark here if English language Translation is attached.

Examiner Signature	/Daniel Abebe/ (05/01/2009)	Date Considered	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /D.A./

IPR2023-00033
Apple EX1002 Page 561

Search Notes

Application/Control No.	Applicant(s)/Patent under Reexamination
10/503,463	CATCHPOLE, MARK
Examiner	**Art Unit**
Daniel D. Abebe	2626

SEARCHED			
Class	Subclass	Date	Examiner
704	235, 242		
704	243,258	4/30/2009	DA

INTERFERENCE SEARCHED			
Class	Subclass	Date	Examiner
704	**235**	4/30/2009	**DA**

SEARCH NOTES (INCLUDING SEARCH STRATEGY)		
	DATE	EXMR
WEST	4/30/2009	DA

Issue Classification	Application/Control No.	Applicant(s)/Patent under Reexamination	
	10/503,463	CATCHPOLE, MARK	
	Examiner	Art Unit	
	Daniel D. Abebe	2626	

ISSUE CLASSIFICATION

ORIGINAL						CROSS REFERENCE(S)									
CLASS				SUBCLASS	CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)									
704				235	704	242	243	258							
INTERNATIONAL CLASSIFICATION															
G	1	0	L	15/00											
				/											
				/											
				/											
				/											

(Assistant Examiner) (Date)	/Daniel D Abebe/ 4/30/09	Total Claims Allowed: 69	
(Legal Instruments Examiner) (Date)	(Primary Examiner) (Date)	O.G. Print Claim(s): 1	O.G. Print Fig.: 2

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	1	31	31		61	69	91		121		151		181
2	2	32	32		62		92		122		152		182
3	3	33	33		63		93		123		153		183
4	4	34	34		64		94		124		154		184
5	5	35	35		65		95		125		155		185
6	6	36	36		66		96		126		156		186
7	7	37	37		67		97		127		157		187
8	8	38	38	46	68		98		128		158		188
9	9	39	39	47	69		99		129		159		189
10	10	40	40	48	70		100		130		160		190
11	11	41	41	49	71		101		131		161		191
12	12	42	42	50	72		102		132		162		192
13	13	43	43	51	73		103		133		163		193
14	14	44	44	52	74		104		134		164		194
15	15		45	53	75		105		135		165		195
16	16		46	54	76		106		136		166		196
17	17		47	55	77		107		137		167		197
18	18		48	56	78		108		138		168		198
19	19		49	57	79		109		139		169		199
20	20		50	58	80		110		140		170		200
21	21		51	59	81		111		141		171		201
22	22		52	60	82		112		142		172		202
23	23		53	61	83		113		143		173		203
24	24		54	62	84		114		144		174		204
25	25		55	63	85		115		145		175		205
26	26		56	64	86		116		146		176		206
27	27		57	65	87		117		147		177		207
28	28		58	66	88		118		148		178		208
29	29	45	59	67	89		119		149		179		209
30	30		60	68	90		120		150		180		210

IPR2023-00033
Apple EX1002 Page 563

WEST Search History for Application 10503463

Creation Date: 2009050113:48

Query	DB	Op.	Plur.	Thes.	Date
((lexical near2 tree) and ((lexical near3 process$5) and 704/$5)) AND @pd>20081021	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
((lexical near2 tree) and ((lexical near3 process$5) and 704/$5))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel) and (((lexical near2 tree) and ((lexical near3 process$5) and 704/$5)))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 lexcial near4 proces$4) and (lexical tree)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 lexical near4 proces$4) and (lexical tree)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 lexical near4 proces$4)	PGPB, USPT, USOC, EPAB, JPAB,	ADJ	YES		10-22-2008

IPR2023-00033
Apple EX1002 Page 564

	DWPI, TDBD				
(parallel near5 parameter near4 proces$4) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near5 parameter near4 proces$4) and (704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(pluralit near4 proces$4) and (704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(pluralit$4 near4 proces$4) and lexical and (704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
tree structure and ((pluralit$4 near4 proces$4) and lexical and (704/$5))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel adj2 model) and lexical and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near2 match$6) and lexical and 704/$5	PGPB, USPT,	ADJ	YES		10-22-2008

	USOC, EPAB, JPAB, DWPI, TDBD				
7120582.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(7120582.pn.) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near4 tree near4 search$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(pluralit$5 near4 tree near4 search$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near4 tree near4 search$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near4 tree near4 process$4) and 704/$5	PGPB, USPT, USOC, EPAB, JPAB, DWPI,	ADJ	YES		10-22-2008

	TDBD				
((parallel near4 tree near4 search$4) and 704/$5) or ((parallel near4 tree near4 process$4) and 704/$5)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
lexical adj3 processors!	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel adj3 processors!) and hmm and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel adj3 processors!) and hmm	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(k near lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(k near3 lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(parallel near2 pars$4) and (tree structure) and lexical	PGPB, USPT, USOC,	ADJ	YES		10-22-2008

	EPAB, JPAB, DWPI, TDBD				
7035802.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
7035802.pn. or 7120582.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
5995930.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn.) or 7120582.pn.	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
lexic$4 and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008

tree$4 and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(process$5) and ((pluralit$4 near4 proces$4) and lexical and (704/$5))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
(process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
((process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)) AND @pd>20081022	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
((process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
lexical and (((process$5) and (5995930.pn. or 5832428.pn. or 20010011218.pn. or 6047283.pn. or 7120582.pn.)))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-22-2008
20010011218.pn. or 6047283.pn. or 7120582.pn.	PGPB, USPT, USOC, EPAB,	ADJ	YES		10-22-2008

	JPAB, DWPI, TDBD				
(20010011218.pn. or 7120582.pn.) AND @pd>20081023	PGPB, USPT	ADJ	YES		10-23-2008
(20010011218.pn. or 7120582.pn.) and score	PGPB, USPT	ADJ	YES		10-23-2008
(history) and ((20010011218.pn. or 7120582.pn.) and score)	PGPB, USPT	ADJ	YES		10-23-2008
(20010011218.pn. or 7120582.pn.) and (stor$6)	PGPB, USPT	ADJ	YES		10-23-2008
(20010011218.pn. or 7120582.pn.) and (propagat$6)	PGPB, USPT	ADJ	YES		10-23-2008
prun$4 and ((20010011218.pn. or 7120582.pn.) and (stor$6))	PGPB, USPT	ADJ	YES		10-23-2008
(multiple or pluralit$4 or more than one) near4 (tree) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008
(704/$5) and ((multiple or pluralit$4 or more than one) near4 (tree) and lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008
((multiple or pluralit$4 or more than one) near4 (tree) near5 (process$4)) and lexical	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008
(704/$5) and (((multiple or pluralit$4 or more than one) near4 (tree) near5 (process$4)) and lexical)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-23-2008

((20010011218.pn. or 7120582.pn.) and (stor$6)) AND @pd>20081023	PGPB, USPT	ADJ	YES		10-24-2008
(7120582.pn. or 20010011218.pn.) and (control$6)	PGPB, USPT	ADJ	YES		10-24-2008
(prun$4 and ((20010011218.pn. or 7120582.pn.) and (stor$6))) AND @pd>20081023	PGPB, USPT	ADJ	YES		10-24-2008
704/$5	USPT	ADJ	YES		10-24-2008
(7120582.pn. or 20010011218.pn.)	PGPB, USPT	ADJ	YES		10-24-2008
lexical tree near4 parallel	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		10-24-2008
(704/$5) and lexical tree	USPT	ADJ	YES		10-24-2008
704/$5	USPT	ADJ	YES		10-24-2008
((7120582.pn. or 20010011218.pn.)) AND @pd>20081024	PGPB, USPT	ADJ	YES		10-25-2008
((7120582.pn. or 20010011218.pn.))	PGPB, USPT	ADJ	YES		10-25-2008
control$5 and (((7120582.pn. or 20010011218.pn.)))	PGPB, USPT	ADJ	YES		10-25-2008
(lexical near6 parallel)	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		05-01-2009
(704/$5) and ((lexical near6 parallel))	PGPB, USPT, USOC, EPAB, JPAB, DWPI, TDBD	ADJ	YES		05-01-2009

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

24998 7590 05/11/2009

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403

OIPE IAP42 AUG 03 2009 PATENT & TRADEMARK OFFICE

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/503,463	05/24/2005	Mark Catchpole	M0025.0311/P311	8875

TITLE OF INVENTION: SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$755	$300	$0	$1055	08/11/2009

EXAMINER	ART UNIT	CLASS-SUBCLASS
ABEBE, DANIEL DEMELASH	2626	704-242000

08/04/2009 LNGUYEN2 00000034 10503463
01 FC:2501 755.00 OP
02 FC:1504 300.00 OP
01 FC:1504 15.00 OP

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

Dickstein Shapiro LLP
X
X

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE: Zentian Limited

(B) RESIDENCE: (CITY and STATE OR COUNTRY): Cambridge, United Kingdom

Please check the appropriate assignee category or categories (will not be printed on the patent) : ☐ Individual ☒ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
☒ Issue Fee
☒ Publication Fee (No small entity discount permitted)
☒ Advance Order - # of Copies 5

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
☐ A check is enclosed.
☒ Payment by credit card. Form PTO-2038 is attached.
☒ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number 04-1073 (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
☒ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.
☐ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature Stephen A. Soffen Date August 3, 2009

Typed or printed name Stephen A. Soffen Registration No. 31,063

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

IPR2023-00033
EX1002 Page 572
Apple EX1002 Page 572

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/503,463	09/08/2009	7587319	M0025.0311/P311	8875

24998 7590 08/19/2009

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 823 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

Mark Catchpole, Cambridgeshire, GBN, UNITED KINGDOM;

Doc Code: PET.PTA.RCAL
Document Description: Request for Recalculation in view of Wyeth

PTO/SB/131 (01-10)
Approved for use through 02/28/2011. OMB 0651-0020
U.S. Patent and Trademark Office; U. S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

REQUEST FOR RECALCULATION OF PATENT TERM ADJUSTMENT IN VIEW OF *WYETH**

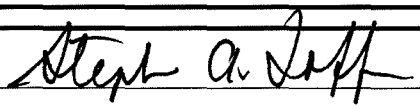
Attorney Docket Number: M0025.0311	Patent Number: 7,587,319
Filing Date (or 371(b) or (f) Date): February 4, 2003	Issue Date: September 8, 2009
First Named Inventor: Catchpole	
Title: SPEECH RECOGNITION CIRCUIT	

PATENTEE HEREBY REQUESTS RECALCULATION OF THE PATENT TERM ADJUSTMENT (PTA) UNDER 35 USC 154(b) INDICATED ON THE ABOVE-IDENTIFIED PATENT. THE PATENTEE'S SOLE BASIS FOR REQUESTING THE RECALCULATION IS THE USPTO'S PRE-*WYETH* INTERPRETATION OF 35 U.S.C. 154(b)(2)(A).

Note: This form is only for requesting a recalculation of PTA for patents issued before March 2, 2010, if the sole basis for requesting the recalculation is the USPTO's pre-*Wyeth* interpretation of 35 U.S.C. 154(b)(2)(A). See Instruction Sheet on page 2 for more information.

Patentees are reminded that to preserve the right to review in the United States District Court for the District of Columbia of the USPTO's patent term adjustment determination, a patentee must ensure that he or she also takes the steps required under 35 U.S.C. 154(b)(3) and (b)(4) and 37 CFR 1.705 in a timely manner.

**Wyeth v. Kappos*, No. 2009-1120 (Fed. Cir., Jan. 7, 2010).

Signature 	Date March 5, 2010
Name (Print/Typed) Stephen A. Soffen	Registration Number 31,063

Note: *Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required in accordance with 37 CFR 1.33 and 11.18. Please see 37 CFR 1.4(d) for the form of the signature. If necessary, submit multiple forms for more than one signature, see below**.

☑ *Total of 1 forms are submitted.

The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Electronic Acknowledgement Receipt

EFS ID:	7155700
Application Number:	10503463
International Application Number:	
Confirmation Number:	8875
Title of Invention:	SPEECH RECOGNITION CIRCUIT USING PARALLEL PROCESSORS
First Named Inventor/Applicant Name:	Mark Catchpole
Customer Number:	24998
Filer:	Stephen A. Soffen/Teresa Joel
Filer Authorized By:	Stephen A. Soffen
Attorney Docket Number:	M0025.0311/P311
Receipt Date:	05-MAR-2010
Filing Date:	24-MAY-2005
Time Stamp:	20:35:32
Application Type:	U.S. National Stage under 35 USC 371

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Request for PTA recalculation in view of Wyeth	311.PDF	69999 5bcb8b2f987341a70626519fb75686e7f1f9ea42	no	1

Warnings:	
Information:	

Total Files Size (in bytes):	69999

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

DICKSTEIN SHAPIRO LLP
1825 EYE STREET NW
Washington, DC 20006-5403

Mail Date: 04/21/2010

Applicant	: Mark Catchpole	: DECISION ON REQUEST FOR
Patent Number	: 7587319	: RECALCULATION of PATENT
Issue Date	: 09/08/2009	: TERM ADJUSTMENT IN VIEW
Appliction No	: 10/503,463	: OF WYETH AND NOTICE OF INTENT TO
Filed	: 05/24/2005	: ISSUE CERTIFICATE OF CORRECTION
		:

The Request for Recalculation is **GRANTED** to the extent indicated.

The patent term adjustment has been determined to be **1137** days. The USPTO will *sua sponte* issue a certificate of correction reflecting the amount of PTA days determined by the recalculation.

Prior to the issuance of the certificate of correction, the USPTO will afford patentee an opportunity to be heard and request reconsideration. Accordingly, patentee has **one month or thirty (30) days**, whichever is longer, to file a request for reconsideration of this patent term adjustment calculation. See 35 U.S.C. 154(b)(3)(B)(ii) and 37 CFR 1.322(a)(4). No extensions of time will be granted under 37 CFR 1.136.

Patentee should use document code PET.OP if electronically filing a request for reconsideration of this patent term adjustment calculation. The patentee must also include the information required by 37 CFR 1.705(b)(2) and the fee required by 37 CFR 1.18(e). If patentee does not file a timely request for reconsideration of this patent term adjustment calculation including the information required by 37 CFR 1.705(b)(2) and the fee required by 37 CFR 1.18(e), the USPTO will issue a certificate of correction reflecting the PTA determination noted above.

Patentee should be aware that in order to preserve the right to review in the United States District Court for the District of Columbia of the USPTO patent term adjustment determination, patentee must ensure that he or she also take the steps required under 35 U.S.C. 154(b)(4)(A) in a timely manner. Nothing in the request for recalculation should be construed as providing an alternative time frame for commencing a civil action under 35 U.S.C. 154(b)(4)(A).

Any questions concerning this decision should be directed to the Office of Patent Legal Administration at 571-272-7702.

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 7,587,319 B2
APPLICATION NO. : 10/503463
DATED : September 8, 2009
INVENTOR(S) : Mark Catchpole

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

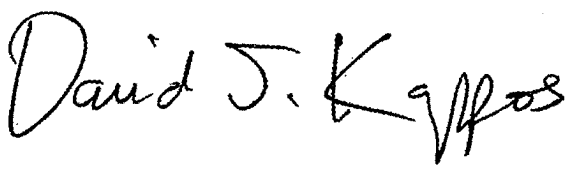
On the Title Page:

The first or sole Notice should read --

Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 1137 days.

Signed and Sealed this

Twenty-first Day of September, 2010



David J. Kappos
Director of the United States Patent and Trademark Office

AO 120 (Rev. 08/10)

TO:	REPORT ON THE FILING OR DETERMINATION OF AN ACTION REGARDING A PATENT OR TRADEMARK
Mail Stop 8 Director of the U.S. Patent and Trademark Office P.O. Box 1450 Alexandria, VA 22313-1450	

In Compliance with 35 U.S.C. § 290 and/or 15 U.S.C. § 1116 you are hereby advised that a court action has been filed in the U.S. District Court for the Western District of Texas on the following

☐ Trademarks or ☑ Patents. (☐ the patent action involves 35 U.S.C. § 292.):

DOCKET NO.	DATE FILED	U.S. DISTRICT COURT
6:22-cv-00123	2/2/2022	for the Western District of Texas

PLAINTIFF	DEFENDANT
Zentian Ltd.	Amazon.com, Inc.

	PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1	7,979,277	7/12/2011	Zentian Ltd.
2	7,587,319	9/8/2009	Zentian Ltd.
3	10,971,140	4/6/2021	Zentian Ltd.
4	10,062,377	8/28/2018	Zentian Ltd.
5	10,839,789	11/17/2020	Zentian Ltd.

In the above---entitled case, the following patent(s)/ trademark(s) have been included:

DATE INCLUDED	INCLUDED BY ☐ Amendment ☐ Answer ☐ Cross Bill ☐ Other Pleading	
PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1		
2		
3		
4		
5		

In the above---entitled case, the following decision has been rendered or judgement issued:

DECISION/JUDGEMENT

CLERK	(BY) DEPUTY CLERK	DATE

Copy 1---Upon initiation of action, mail this copy to Director Copy 3---Upon termination of action, mail this copy to Director
Copy 2---Upon filing document adding patent(s), mail this copy to Director Copy 4---Case file copy

AO 120 (Rev. 08/10)

TO:	REPORT ON THE FILING OR DETERMINATION OF AN ACTION REGARDING A PATENT OR TRADEMARK
Mail Stop 8 Director of the U.S. Patent and Trademark Office P.O. Box 1450 Alexandria, VA 22313-1450	

In Compliance with 35 U.S.C. § 290 and/or 15 U.S.C. § 1116 you are hereby advised that a court action has been filed in the U.S. District Court for the Western District of Texas on the following

☐ Trademarks or ☑ Patents. (☐ the patent action involves 35 U.S.C. § 292.):

DOCKET NO.	DATE FILED	U.S. DISTRICT COURT
6:22-cv-00122	2/2/2022	for the Western District of Texas

PLAINTIFF	DEFENDANT
Zentian Ltd.	Apple, Inc.

	PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1	7,979,277	7/12/2011	Zentian Ltd.
2	7,587,319	9/8/2009	Zentian Ltd.
3	10,971,140	4/6/2021	Zentian Ltd.
4	10,062,377	8/28/2018	Zentian Ltd.
5	10,839,789	11/17/2020	Zentian Ltd.

In the above---entitled case, the following patent(s)/ trademark(s) have been included:

DATE INCLUDED	INCLUDED BY ☐ Amendment ☐ Answer ☐ Cross Bill ☐ Other Pleading	
PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1		
2		
3		
4		
5		

In the above---entitled case, the following decision has been rendered or judgement issued:

DECISION/JUDGEMENT

CLERK	(BY) DEPUTY CLERK	DATE

Copy 1---Upon initiation of action, mail this copy to Director Copy 3---Upon termination of action, mail this copy to Director
Copy 2---Upon filing document adding patent(s), mail this copy to Director Copy 4---Case file copy