

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

APPLE INC., AMAZON WEB SERVICES, INC., and
AMAZON.COM SERVICES LLC,
Petitioner,

v.

ZENTIAN LIMITED,
Patent Owner.

IPR2023-00034¹
Patent 7,979,277 B2

Before KEVIN F. TURNER, JEFFREY S. SMITH, and
CHRISTOPHER L. OGDEN, Administrative Patent Judges.

OGDEN, Administrative Patent Judge.

JUDGMENT
Final Written Decision
Determining All Challenged Claims Unpatentable
35 U.S.C. § 318(a)

¹ IPR2023-01193 has been joined with this proceeding.

I. INTRODUCTION

In response to a Petition (Paper 1, “Pet.”) filed by Petitioner Apple Inc. (“Apple”), the Board instituted an *inter partes* review of claims 1, 4, 5, 7, 9, 10, 12, and 14–16 of U.S. Patent No. 7,979,277 B2 (Ex. 1001, “the ’277 patent”). Paper 10 (“Dec.”). Patent Owner Zentian Limited (“Zentian”) filed a Patent Owner Response (Paper 19), Apple filed a Reply to the Patent Owner Response (Paper 21 (“Pet. Reply”)), and Zentian filed a Sur-reply (Paper 27 (“PO Sur-reply”)).

We held an oral hearing on March 12, 2024, and the transcript is entered into the record (“Tr.”). Paper 31.

This is a final written decision under 35 U.S.C. § 318(a) as to whether the claims challenged in the *inter partes* review are unpatentable. For the reasons below, we conclude that Apple has shown that all the challenged claims are unpatentable.

II. BACKGROUND

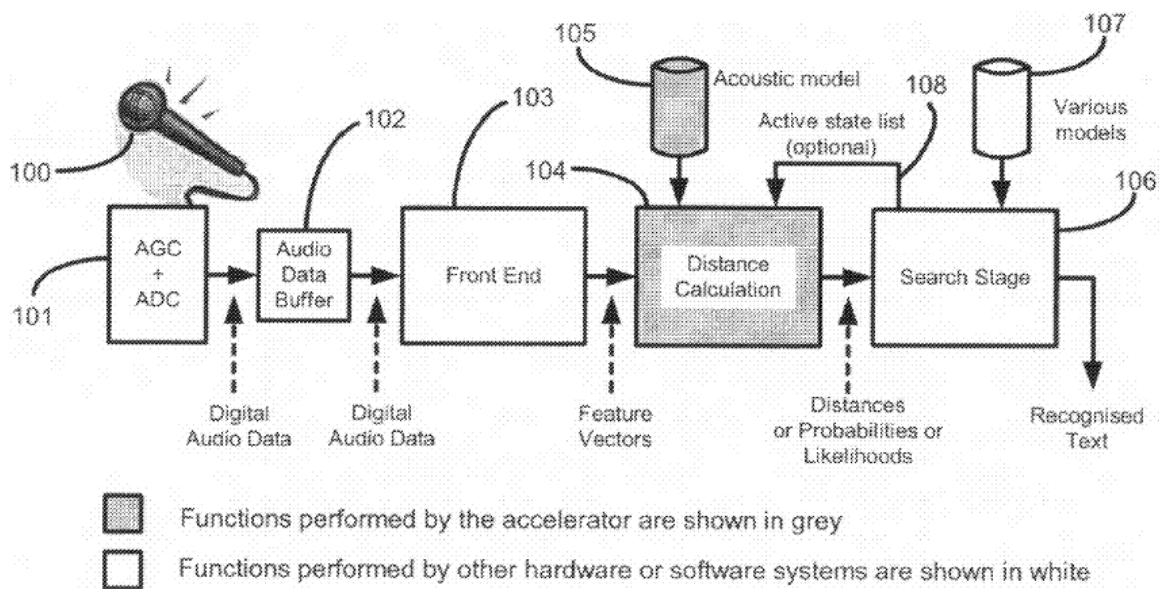
A. RELATED PROCEEDINGS

As related matters, the parties identify *Zentian Ltd. v. Apple Inc.*, No. 6:22-cv-00122-ADA (W.D. Tex. filed Feb. 2, 2022) and *Zentian Ltd. v. Amazon.com*, No. 6:22-cv-00123-ADA (W.D. Tex. filed Feb. 2, 2022), as well as co-pending *inter partes* review proceedings IPR2023-00033 and -00035 to -00037. Pet. 74; Paper 3, 1.

B. THE '277 PATENT (EX. 1001)

The '277 patent is directed to a speech recognition circuit comprising multiple parts that perform different steps of the recognition process.

Ex. 1001, code (57), 5:58–6:3, 6:17–26. Figure 16, reproduced below, illustrates one embodiment comprising a front end 103, distance calculation engine 104, and search stage 106, including data flow between these components:



Ex. 1001, 8:38–40, 23:33–34. As shown above, a speech signal from microphone 100 is converted to digital audio data and sent to front end 103, where it “is split into spectral components to generate a feature vector” capturing key features of the audio input. *Id.* at 23:34–53. Front end 103 does this by dividing the digital audio signal into time frames (e.g., 10 millisecond periods) and, for each time frame, “splitting the audio signal into spectral components” (e.g., a 39-component vector with twelve frequency bands, one energy band, plus the first and second derivative for each band). *Id.* at 12:1–10.

Front end 103 then sends these feature vectors to distance calculation engine 104, which may take place on a separate speech accelerator to reduce the computational load on the primary CPU. Ex. 1001, 3:35–50, 5:18–33, 24:6–11. Using a hidden Markov model (“HMM”) or a neural network, distance calculation engine 104 compares each feature vector with all possible states within a predetermined acoustic model stored in memory 105. *Id.* at 4:15–18, 12:12–14, 23:58–61. That comparison is in the form of a “distance” (such as a Mahalanobis distance²) to possible states within the acoustic model, which indicates “the probability or likelihood that a feature vector corresponds to a particular state” within the acoustic model. *Id.* at 4:13–15.

Finally, distance calculation engine 104 sends the calculated distances to search stage 106, which uses word or language models stored in memory 107 “to generate and output recognized text.” Ex. 1001, 23:61–67. Search stage 106 uses the “calculated distances to identify words within a lexical tree, the lexical tree comprising a model of words.” *Id.* at 5:65–67.

In one optional embodiment, search stage 106 sends active state list 108 back to distance calculation engine 104 as “an optimization that may reduce computation time and/or power consumption.” Ex. 1001, 24:25–30. Alternatively, in the preferred embodiment, “the front end 103, distance calculation engine 104 and the search stage 106 operate in a pipelined

² The Mahalanobis distance “is a distance between two N-dimensional points, scaled by the statistical variation in each component.” Ex. 1001, 12:27–29. In this context, it is “the distance in the N-dimensional space from the [f]eature [v]ector to the probability distribution for each state.” *Id.* at 12:42–44; *see also id.* at 12:51–63 (noting that the Mahalanobis distance “is extensively described in the literature,” and providing a formula).

manner.” *Id.* at 24:31–33. According to the ’277 patent, “[w]hen operating in a pipelined manner it is unlikely that the search stage 106 will be able to provide the active state list 1[08] early enough for the distance calculation engine 104 to implement the optimization of computing only the distances that will be required by the search stage 106.” *Id.* at 24:33–38. Thus, “distance calculation circuit 104 may calculate the [Mahalanobis distance] values for every state in the lexicon, per frame, whether it is subsequently required by the search stage or not,” which “allows the accelerator and software system to operate in a concurrent pipelined manner, which maximizes the system throughput.” *Id.* at 24:38–43.

C. CHALLENGED CLAIMS AND GROUNDS

Representative claim 1 is as follows:

1. A speech recognition circuit comprising:
 - [a] an audio front end for calculating a feature vector from an audio signal,
 - [b] wherein the feature vector comprises a plurality of extracted and/or derived quantities from said audio signal during a defined audio time frame;
 - [c] a calculating circuit for calculating distances indicating the similarity between a feature vector and a plurality of predetermined acoustic states of an acoustic model; and
 - [d] a search stage for using said calculated distances to identify words within a lexical tree, the lexical tree comprising a model of words;
 - [e] wherein said audio front end and said search stage are implemented using a first processor, and said calculating circuit is implemented using a second processor, and
 - [f] wherein data is pipelined from the front end to the calculating circuit to the search stage.

Ex. 1001, 38:12–27 (Apple’s reference numbers added).

Apple argues four grounds for inter partes review, as summarized in the following table:

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
1, 5, 7, 12, 14–16	103(a) ³	Jiang, ⁴ Baumgartner, ⁵ Brown ⁶
4	103(a)	Jiang, Baumgartner, Brown, Kazeroonian ⁷
9, 10	103(a)	Jiang, Baumgartner, Brown, Vensko ⁸
1, 5, 7, 12, 14–16	103(a)	Jiang, Baumgartner, Brown, Smyth ⁹
4	103(a)	Jiang, Baumgartner, Brown, Smyth, Kazeroonian
9, 10	103(a)	Jiang, Baumgartner, Brown, Smyth, Vensko

Pet. 4.

³ 35 U.S.C. § 103(a) (2006), *amended by* Leahy–Smith America Invents Act, Pub. L. No. 112-29 § 103, sec. (n)(1), 125 Stat. 284, 287, 293 (2011) (effective Mar. 16, 2013). This version of § 103 applies because the ’277 patent issued from an international application filed on September 14, 2005, which is before the effective date of the AIA amendments. *See* Ex. 1001, code (22).

⁴ Jiang, US 6,374,219 B1 (issued Apr. 16, 2002) (Ex. 1004).

⁵ Baumgartner et al., US 2002/0049582 A1 (published Apr. 25, 2002) (Ex. 1007).

⁶ Brown et al., US 5,699,456 (issued Dec. 16, 1997) (Ex. 1036).

⁷ Kazeroonian et al., WO 99/41684 (published Aug. 19, 1999) (Ex. 1037).

⁸ Vensko et al., US 4,567,606 (issued Jan. 28, 1986) (Ex. 1038).

⁹ Smyth et al., US 5,819,222 (issued Oct. 6, 1998) (Ex. 1005).

D. EXPERT TESTIMONY

Apple submits a declaration by Christopher Schmandt. Ex. 1003.
Zentian likewise submits a declaration by Dr. David Anderson. Ex. 2020.

III. GROUNDS OF THE PETITION

For the reasons below, we determine that Apple has shown, by a preponderance of the evidence, that claims 1, 4, 5, 7, 9, 10, 12, and 14–16 of the '277 patent are unpatentable as obvious.

A. LEVEL OF ORDINARY SKILL IN THE ART

The level of ordinary skill in the pertinent art at the time of the invention is a factor in how we construe patent claims. *See Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–13 (Fed. Cir. 2005) (en banc). It is also one of the factors we consider when determining whether a patent claim is obvious over the prior art. *See Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

To assess the level of ordinary skill, we construct a hypothetical “person of ordinary skill in the art,” from whose vantage point we assess obviousness and claim interpretation. *See In re Rouffet*, 149 F.3d 1350, 1357 (Fed. Cir. 1998). This legal construct “presumes that all prior art references in the field of the invention are available to this hypothetical skilled artisan.” *Id.* (citing *In re Carlson*, 983 F.2d 1032, 1038 (Fed. Cir. 1993)).

Relying on Mr. Schmandt’s testimony, Apple argues that a person of ordinary skill in the art would have had “a master’s degree in computer engineering, computer science, electrical engineering, or a related field, with . . . two years of experience in the field of speech recognition, or a

bachelor's degree in the same fields with . . . four years of experience in the field of speech recognition.” Pet. 3 (citing Ex. 1003 ¶¶ 28–30).

Zentian does not contest Apple's proposed definition of the level of ordinary skill. PO Resp. 5. Because Apple's articulation of that level of skill is uncontested, supported by testimonial evidence, and appears consistent with the types of problems and solutions in the '277 patent, we adopt it for this decision. *See, e.g.*, Ex. 1001, 1:4–8:10.

B. CLAIM CONSTRUCTION

In an inter partes review, we construe a patent claim “using the same claim construction standard that would be used to construe the claim in a civil action under 35 U.S.C. 282(b).” 37 C.F.R. § 42.100(b). This generally includes “construing the claim in accordance with the ordinary and customary meaning of such claim as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent.” *Id.* The ordinary and customary meaning of a claim term “is its meaning to the ordinary artisan after reading the entire patent,” and “as of the effective filing date of the patent application.” *Phillips*, 415 F.3d at 1313, 1321. There are only two circumstances in which a construction departs from the ordinary and customary meaning: “1) when a patentee sets out a definition and acts as [their] own lexicographer, or 2) when the patentee disavows the full scope of a claim term either in the specification or during prosecution.” *Thorner v. Sony Comput. Entm't Am. LLC*, 669 F.3d 1362, 1365 (Fed. Cir. 2012). Any such special meaning of a term “must be sufficiently clear in the specification that any departure from common usage would be so understood

by a person of experience in the field of the invention.” *Multiform Desiccants Inc. v. Medzam Ltd.*, 133 F.3d 1473, 1477 (Fed. Cir. 1998).

To construe the claim terms, “we look principally to the intrinsic evidence of record, examining the claim language itself, the written description, and the prosecution history, if in evidence.” *DePuy Spine, Inc. v. Medtronic Sofamor Danek, Inc.*, 469 F.3d 1005, 1014 (Fed. Cir. 2006).

Apple argues that all terms should be given their plain and ordinary meaning, except that the limitation “calculating means for calculating a distance indicating the similarity between a feature vector and a predetermined acoustic state of an acoustic model” in claim 14 (Ex. 1001, 39:6–8) is a means-plus-function limitation and thus invokes 35 U.S.C. § 112 para. 6 (2006). Pet. 5–6.

Zentian does not contest these arguments or propose any explicit constructions. *See* PO Resp.

We do not find it necessary to explicitly construe any terms for this decision. To the extent that we need to interpret claim language, we address that language below in the context of the prior art. *See Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (“[W]e need only construe terms ‘that are in controversy, and only to the extent necessary to resolve the controversy’” (quoting *Vivid Techs., Inc. v. Am. Sci & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999))).

C. GROUNDS BASED ON JIANG, BAUMGARTNER, BROWN, AND
OPTIONALLY SMYTH (CLAIMS 1, 5, 7, 12, AND 14–16)

Apple’s first ground challenges claims 1, 5, 7, 12, and 14–16 as obvious over Jiang, Baumgartner, and Brown, and its fourth ground challenges the same claims based on the same references with the addition

of Smyth. Pet. 4. Because of the common issues and evidence, we address these grounds together.

A claim is unpatentable under § 103 for obviousness “if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains.” 35 U.S.C. § 103; *see also KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007).

We base our obviousness inquiry on factual considerations including (1) the scope and content of the prior art, (2) any differences between the claimed subject matter and the prior art, (3) the level of skill in the art, and (4) any objective indicia of obviousness or non-obviousness that may be in evidence. *See Graham*, 383 U.S. at 17–18.

Considering these factors,¹⁰ we determine that Apple has shown, by a preponderance of the evidence, that claims 1, 5, 7, 12, and 14–16 are unpatentable under § 103(a) as obvious over Jiang in view of Baumgartner and Brown, with or without Smyth.

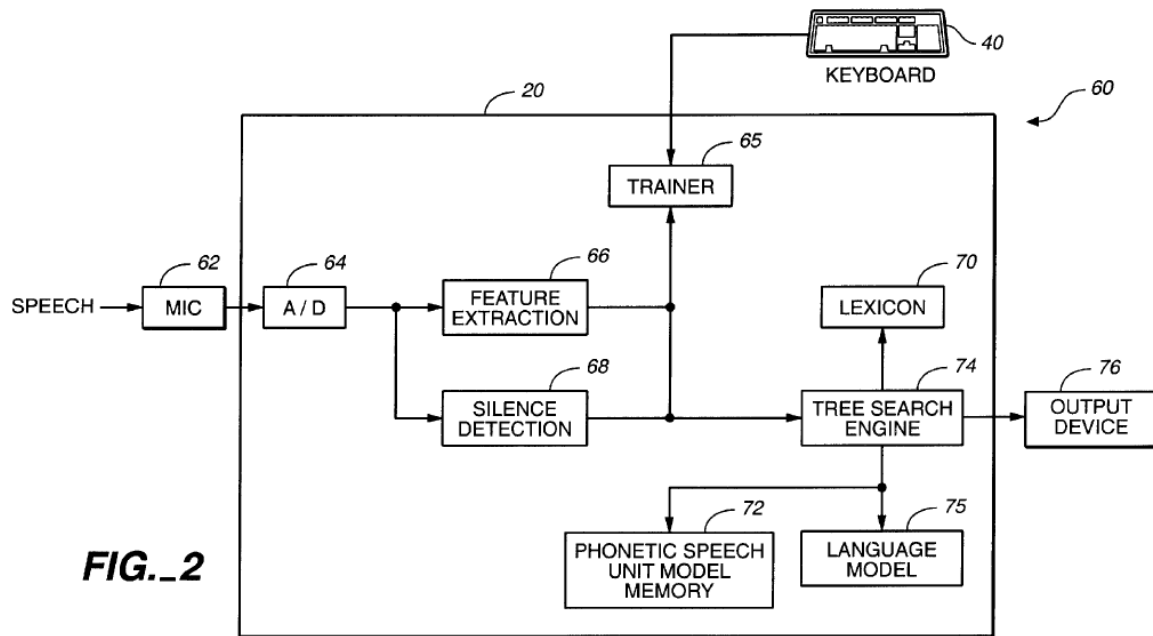
1. Overview of Prior Art

(a) Jiang

Jiang is directed to “computer speech recognition performed by conducting a prefix tree search of a silence bracketed lexicon.” Ex. 1004, 1:15–18. The system recognizes possible words from an input data stream by comparing them to a prefix tree including a plurality of phoneme

¹⁰ Neither party presents evidence of objective indicia of obvious or non-obviousness, so no such evidence factors into our decision.

branches connected at nodes. *Id.* at 4:8–12. Figure 2, reproduced below, is a block diagram of the system:



As shown above, Jiang’s system 60 has components including microphone 62, analog-to-digital (A/D) converter 64, training module 65, feature extraction module 66, silence detection module 68, lexicon storage module 70, phonetic speech unit storage module 72, tree search engine 74, and output device 76. Ex. 1004, 6:20–25.

A user speaks into microphone 62, which converts the audible speech into an analog electrical signal, and then A/D converter 64 converts the analog signals into a sequence of digital signals and then sends them to feature extraction module 66. Ex. 1004, 6:46–52. Feature extraction module 66 divides the digital signal into frames, each lasting about 10 milliseconds. *Id.* at 6:62–65.

Feature extraction module 66 also encodes each frame into a feature vector reflecting the spectral characteristics for a plurality of frequency bands. Ex. 1004, 6:65–7:1. “In the case of discrete and semi-continuous

hidden Markov modeling, feature extraction module 66 also preferably encodes the feature vectors into one or more codewords using vector quantization techniques and a codebook derived from training data.” *Id.* at 7:1–5.

Upon receiving the codewords or feature vectors from feature extraction module 66 and a boundary detection signal provided by silence detection module 68, tree search engine 74 accesses information stored in the phonetic speech unit model memory 72. Ex. 1004, 7:5–7, 7:30–34. Based upon the HMMs stored in memory 72, tree search engine 74 determines a most likely phoneme represented by the codeword or feature vector received from feature extraction module 66, and hence representative of the utterance received by the user of the system. *Id.* at 7:11–15, 7:37–42. Search engine 74 also references lexicon 70 (containing a prefix tree for all words in the system’s vocabulary) as well as a language model stored in module 75 to “identify[] a most likely word or word sequence represented by the input data,” which search engine 74 then sends to output device 76. *Id.* at 7:48–8:4.

While traversing the lexical tree in lexicon 70, search engine 74 may use a “pruning technique” such that, if a score for a particular node of the tree is sufficiently low, “it is assumed that the branches leading to the [low-scoring] node . . . will not likely be part of the completed most likely phoneme sequence Thus, that branch is dropped (or pruned) from [the tree] and is no longer considered in further processing.” Ex. 1004, 8:52–64.

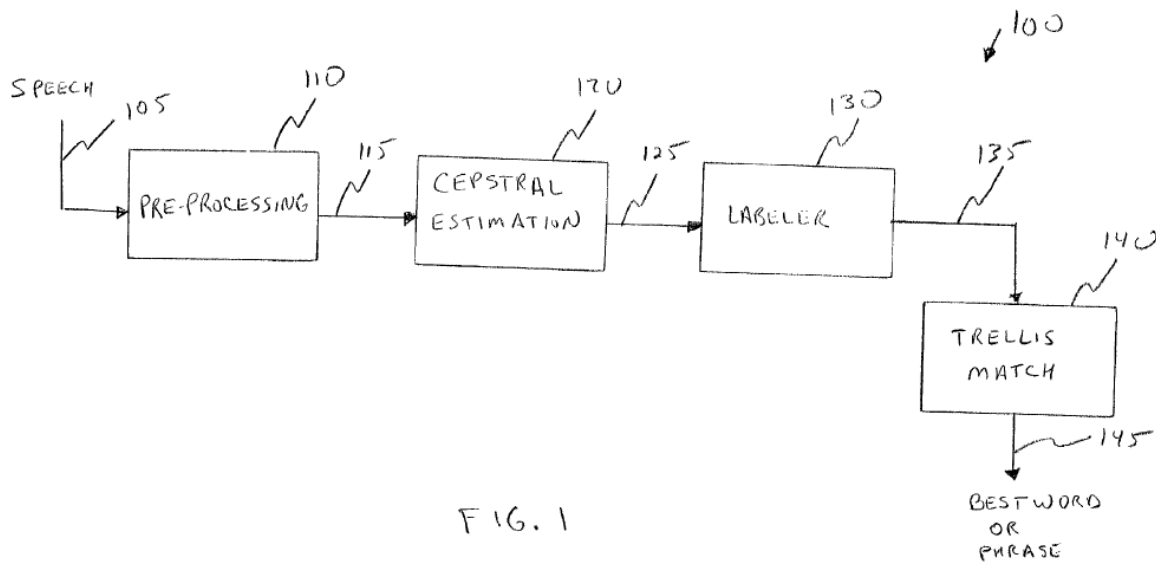
According to Jiang, feature extraction module 66 may take place on a different processor than tree search engine 74: “In a preferred embodiment, feature extraction module 66 is a conventional array processor which

performs spectral analysis on the digital signals.” Ex. 1004, 6:53–55.

“Further, tree search engine 74 is preferably implemented in [a CPU] (which may include one or more processors) or may be performed by a dedicated speech recognition processor employed by [a] personal computer.” *Id.* at 6:39–42.

(b) Baumgartner

Baumgartner is directed to a speech recognition system that includes “Speech Label Accelerators (SLAs),” distinct processors that “speed the labeling phase of the speech recognition computation” and “reduce the computational load of a processor during speech recognition.” Ex. 1007 ¶¶ 2, 6, 31. Figure 1, reproduced below, is a block diagram of a speech recognition system that includes an SLA (130):

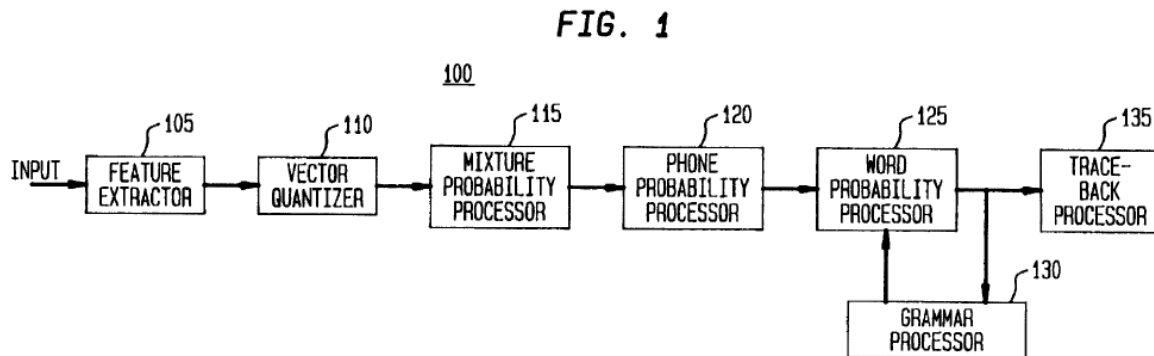


As shown in the figure above, “[s]peech recognition system 100 operates on a speech signal 105 to produce a best word or phrase 145.” Ex. 1007 ¶ 36. In doing so, pre-processing block 110 and cepstral estimation block 120 generate “a series of observation vectors” 125 from speech signal 105. *Id.*

¶¶ 37–38. Labeling block 130 (an SLA) takes observation vectors 125 and uses them to generate probabilities 135 for various “phones,” which are “small, indivisible acoustic elements of speech.” *Id.* ¶ 38. Then trellis match block 140 “performs a search and selects a best word or phrase 145.” *Id.*

(c) Brown

Brown describes “[a] method of recognizing speech input [that] selectively creates and maintains grammar representations of the speech input in essentially real time.” Ex. 1036, code (57). Figure 1, reproduced below, is a block diagram of a speech recognition system implementing this method:



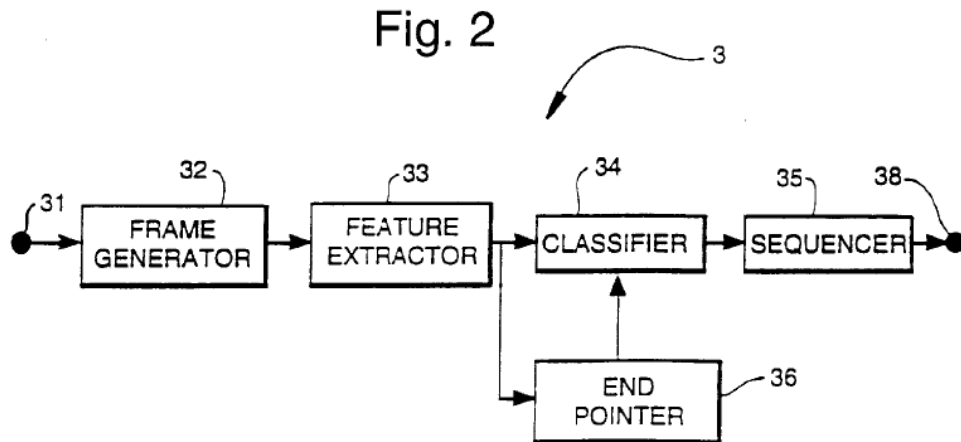
As shown in the figure above, system 100 “includes a number of functional blocks” including feature extractor 105, vector quantizer 110, and probability processors 115, 120, and 125. Ex. 1036, 3:34–40. In particular, phone probability processor 120 generates “phone scores” which “indicate[] the probability that a respective phoneme was uttered during the speech signal frame period in question.” *Id.* at 4:10–12. Word probability processor 125 contains a word model and uses the phone scores to generate word probability scores. *Id.* at 4:13–29.

Brown teaches that “[t]he functionality of each block of the recognition system may be performed by a respective different processor,”

and that “[t]he stages are pipelined and their operations are synchronized with the rate at which frames of the incoming speech signal are generated.” Ex. 1036, 3:40–48.

(d) Smyth

Smyth is directed to “[a] speech recognition system [that] recognizes connected speech using a plurality of vocabulary nodes.” Ex. 1005, code (57). Figure 1, reproduced below, is a block diagram of a recognition processor used in this system:



As shown in the above figure, recognition processor 3 includes frame generator 32, which takes digital speech input 31 and produces a series of frames, and feature extractor 33, which processes the frames to produce a set of feature vectors. Ex. 1005, 5:1–7. Classifier 34 receives the feature vectors and “operat[es] on each with a plurality of model states, to generate recognition results.” *Id.* at 5:7–9. End pointer 36 determines the end of speech, and sequencer 35 “determine[s] the predetermined utterance to which the sequence of classifier output indicates the greatest similarity,” which sequencer 35 outputs to port 38. *Id.* at 5:9–18.

Frame generator 32 and feature extractor 33 are “provided by a single suitably programmed digital signal processor (DSP) device.” Ex. 1005, 5:49–51. Classifier 34 may be a separate processor with its own state memory. *See id.* at 5:55–57.

2. *Independent Claim 1*

(a) Preamble and Limitations 1(a) and 1(b)

The preamble of claim 1 recites “[a] speech recognition circuit.” Ex. 1001, 38:12. Limitations 1(a) and 1(b) recite “an audio front end for calculating a feature vector from an audio signal, wherein the feature vector comprises a plurality of extracted and/or derived quantities from said audio signal during a defined audio time frame.” *Id.* at 38:13–16.

“To the extent the preamble is limiting,” Apple contends that Jiang teaches a speech recognition system and contends that Jiang’s feature extraction module 66 is an audio front end that calculates feature vectors as recited in limitations 1(a) and 1(b). Pet. 10–15 (citing Ex. 1004, 1:15–19, 1:58–67, 2:7–10, 4:30–32, 4:51–7:10; Ex. 1003 ¶¶ 111–123).

Zentian does not specifically contest Apple’s arguments about the preamble of claim 1 or limitations 1(a) or 1(b). *See* PO Resp.

Because the evidence of record suggests that Jiang teaches a speech recognition circuit (*see, e.g.*, Ex. 1004, Fig. 2) and thus discloses the preamble, we need not decide whether the preamble is limiting. Also, we find Apple’s arguments persuasive that Jiang’s feature extraction module 66 corresponds to the “audio front end” recited in limitation 1(a) and that the resulting feature vectors comprise “a plurality of extracted and/or derived quantities from said audio signal during a defined audio time frame” as

recited in limitation 1(b). *See* Ex. 1003 ¶¶ 111–123. Thus, we determine that Jiang teaches limitations 1(a) and 1(b).

(b) Limitation 1(c)

Limitation 1(c) recites “a calculating circuit for calculating distances indicating the similarity between a feature vector and a plurality of predetermined acoustic states of an acoustic model.” Ex. 1001, 38:17–19.

Apple asserts that Jiang discloses this limitation alone or in combination with Baumgartner. Pet. 15. In its fourth ground, Apple alternatively relies on Smyth for teaching the recited distance calculation. Pet. 59–60. For the reasons below, we find both of Apple’s arguments persuasive.

First, Apple argues that Jiang discloses a calculating circuit and a plurality of predetermined acoustic states of an HMM acoustic model. Pet. 19–20 (citing Ex. 1004, 1:15–2:15, 2:53–56, 6:39–42, 7:30–8:51, 11:25–47; Ex. 1003 ¶¶ 129–136). Apple also argues that Jiang calculates distances indicating the similarity between a feature vector and acoustic states. Pet. 20–25. In particular, Apple contends that Jiang’s feature extraction module 66 encodes each frame of spoken speech into a feature vector, and also encodes feature vectors into codewords, which Apple contends are “vector quantized feature vectors.” Pet. 20–21 (citing Ex. 1004, 6:62–7:10, 7:1–10, 7:38–43; Ex. 1003 ¶¶ 137–139). According to Apple, Jiang calculates each distance in the form of a “probability score,” which “is a probability or likelihood representative of a similarity or distance measurement.” Pet. 21–23 (citing Ex. 1004, 7:30–47, 8:29–33; Ex. 1003 ¶¶ 140–147).

Apple also contends that a person of ordinary skill in the art would have understood that, to the extent Jiang’s distance calculation involves codewords, the distance would “indicat[e] the similarity between a *feature vector* and HMM states.” Pet. 23 (citing Ex. 1003 ¶ 148). Also, Apple contends that “because the codeword is simply a vector quantized feature vector,” a person of ordinary skill in the art would have understood the comparison between Jiang’s codewords and phonemes in the HMM “to likewise be determining the likelihood the considered phoneme is indicated by one or more *feature vectors*.” Pet. 23 (citing Ex. 1003 ¶¶ 148–149).

Alternatively, Apple argues that Smyth, in a similar speech recognition system to Jiang, teaches a direct comparison between feature vectors and acoustic states within an acoustic model. Pet. 65–67 (citing Ex. 1005, 1:34–37, 5:7–9, 5:55–57, 5:59, 5:63–6:4, Figs. 2, 3; Ex. 1003 ¶¶ 285–288). Apple also contends that a person of ordinary skill in the art “would have been motivated to replace Jiang’s probability scores for each phoneme branch . . . with Smyth’s probability calculation that an input feature vector corresponds to a given state.” Pet. 68 (citing Ex. 1004, 8:29–33; Ex. 1005, 5:63–67; Ex. 1003 ¶¶ 289–297). According to Apple, this is because “[p]erforming distance calculations to determine the most likely phonemes would have entailed use of a known technique, i.e., calculating distances by comparing a feature vector to a state of an acoustic model, such as an HMM, to improve similar devices, i.e., Jiang’s speech recognition system,” and “would have obtained predictable results.” *Id.* at 68–69 (citing Ex. 1012, 21, 28–29; Ex. 1003 ¶¶ 295–296).

Zentian does not specifically contest Apple’s arguments. *See* PO Resp.

For the reasons Apple has provided in its Petition, we determine that Apple has shown that Jiang, in light of either Baumgartner or Smyth, teaches limitation 1(c).

(c) Limitation 1(d)

Limitation 1(d) recites “a search stage for using said calculated distances to identify words within a lexical tree, the lexical tree comprising a model of words.” Ex. 1001, 38:20–22.

Apple argues that the recited search stage corresponds to the part of Jiang’s tree search engine 74 that uses a lexical tree comprising a model of words (lexicon 70). Pet. 28–33 (citing Ex. 1004, 2:47–67, 3:1–21, 6:67–7:4, 7:30–54, 7:65–8:43, 9:37–42, 11:25–47, 7:30–8:51, Figs. 3–6; Ex. 1003 ¶¶ 162–171).

Zentian does not specifically contest Apple’s arguments for limitation 1(d). *See* PO Resp.

We find Apple’s arguments persuasive that Jiang’s tree search engine 74 corresponds to the recited “search stage” and that it uses calculated distances to identify words within a lexical tree as recited in limitation 1(d). *See* Ex. 1003 ¶¶ 162–171.

(d) Limitation 1(e)

Limitation 1(e) recites “wherein said audio front end and said search stage are implemented using a first processor, and said calculating circuit is implemented using a second processor.” Ex. 1001, 38:23–25.

Apple concedes that Jiang “does not expressly teach the hardware architecture of [limitation] 1(e),” but according to Apple, Baumgartner “teaches a similar front end and search stage as Jiang’s provided on a first

processor and a similar calculating circuit as Jiang’s provided on a second processor.” Pet. 34. Apple contends that a person of ordinary skill in the art would have modified Jiang “such that distance calculation was performed on a processor distinct from the processor performing both feature extraction . . . and word identification (i.e., the processor for the tree search engine).” Pet. 42 (citing Ex. 1003 ¶¶ 192–195). According to Apple, this is because Jiang already teaches practicing its speech-recognition system using multiple dedicated processors, and a person of ordinary skill in the art would have looked to Baumgartner’s SLAs as a way to speed up the labeling stage of a speech recognizer. Pet. 41–43 (citing Ex. 1007 ¶¶ 3–7, 200–205; Ex. 1003 ¶¶ 192–195, 201).

Apple also contends that there “would have [been] a reasonable expectation of success in implementing Jiang’s distance calculation on a separate processor without undue experimentation at least because Baumgartner and the art cited by Mr. Schmandt expressly illustrated its feasibility.” Pet. 43 (citing Ex. 1003 ¶ 202). Also, according to Apple, “using separate processors, such as Baumgartner’s SLA, or co-processors, was a well-known technique to offload computation tasks prior to the ’277 Patent.” *Id.* (citing Ex. 1003 ¶ 202).

Zentian does not specifically contest Apple’s specific arguments for limitation 1(e). *See* PO Resp.

For the reasons Apple has provided in its Petition, we determine that Apple has shown that Jiang, in light of Baumgartner, teaches limitation 1(e).

(e) Limitation 1(f)

Limitation 1(f) recites “wherein data is pipelined from the front end to the calculating circuit to the search stage.” Ex. 1001, 38:26–27. Apple relies on Brown as teaching this limitation. Pet. 44–47.

(1) *Apple’s Proposed Combination*

According to Apple, Brown teaches that processing stages in a speech-recognition system may be pipelined such that “as the frames of data are generated, the data is processed from one stage to another so that the operations are synchronized with the rate of frame generation.” Pet. 44–45 (citing Ex. 1036, 3:32–48; Ex. 1003 ¶¶ 206–209). Apple contends that a person of ordinary skill in the art would have had reason to modify the combination of Jiang and Baumgartner by pipelining data sent from the CPU to the SLA and then back to the CPU because “[p]ipelining data offered numerous well-known advantages including greater throughput of the processor, improved performance, decreased cost, and simplified programming.” Pet. 45 (citing Ex. 1003 ¶ 210). Apple also argues that a person of ordinary skill in the art “would have realized pipelining keeps all processors operating efficiently in a multi-processor system, such that wasteful idle-processing time is minimized.” Pet. 46 (citing Ex. 1003 ¶¶ 212–213). And according to Apple, “Jiang already discusses the importance of time with respect to pruning.” Pet. 46 (citing Ex. 1004, 10:65–11:25).

Apple also argues that a person of ordinary skill in the art would have had a reasonable expectation of success in adding Brown’s pipelining to the Jiang–Baumgartner combination because “[p]ipelining was well-known in

the art including in speech recognition systems,” and the three references “teach speech recognition systems using similar hardware, such as various processors to perform similar speech recognition functions,” such that the combination “would have been a trivial modification.” Pet. 46–47 (citing Ex. 1004, 4:60–65, 6:32–36, 6:39–42; Ex. 1007 ¶ 33, Fig. 19; Ex. 1036, 3:11–15; Ex. 1003 ¶ 214).

(2) *Whether Apple’s Proposed Combination Requires the Use of Jiang’s Pruning Technique*

In its Response, Zentian argues that Apple’s proposed combination to modify Jiang to add pipelining, as Brown teaches, requires the use of Jiang’s pruning technique. *See* PO Resp. 6–7. According to Zentian, Mr. Schmandt admitted during cross-examination that Jiang’s pruning technique is part of the asserted prior art combination. *Id.* at 6 (citing Ex. 2016, 15:4–9 (“Q: And so with respect to your theory as to Jiang for the ’277 patent, you also maintain Jiang’s pruning, right? A: I don’t remove it from Jiang, yes. Jiang has pruning, and it’s part of the combination.”))). Zentian also argues that Apple’s proposed combination performs real-time speech recognition, and Jiang teaches that pruning greatly reduces computation time for speech recognition in which there is a large vocabulary. *Id.* at 6–7 (citing Ex. 1004, 2:19–28; Ex. 2016, 6:24–7:2; Ex. 2020 ¶ 40). Thus, according to Zentian, “Jiang expressly teaches the use of pruning to address the problem of excessive processing delays.” *Id.* at 7 (citing Ex. 2020 ¶ 40).

In its Reply, Apple argues that its proposed combination does not rely on Jiang’s pruning embodiment, which Jiang teaches is optional. Pet. Reply 3–13, 20–21. According to Apple, Jiang only discloses that “a pruning

technique *can be used*,” not that it is required. Pet. Reply 3 (quoting Ex. 1004, 8:52). Apple argues that Zentian’s sole basis for arguing that Jiang’s pruning embodiment is part of Apple’s proposed combination is Mr. Schmandt’s statement during cross examination that Jiang’s pruning is “part of the combination.” *Id.* at 3–4 (quoting PO Resp. 6). But according to Apple, the “combination” Mr. Schmandt was referring to was in reference to his “theory” that improved pruning would have been part of the motivation to combine Jiang and Baumgartner. *See id.* at 4; Tr. 17:16–23. Apple argues that Mr. Schmandt also “explains how pipelining achieves Jiang’s desired goals *without* pruning,” because there is a tradeoff: although “pruning saves on computational resources, pruning also decreases accuracy.” Pet. Reply 4–5 (quoting Ex. 1003 ¶ 194) (citing Ex. 1003 ¶¶ 201, 213).

In its Sur-reply, Zentian argues that Mr. Schmandt’s testimony clearly relies on Jiang’s pruning embodiment, and thus it is part of Apple’s proposed prior art combination, despite any unsupported attorney argument to the contrary. PO Sur-reply 11–13 (citing Ex. 1030 ¶¶ 194, 213; Ex. 2016, 15:4–9). Zentian also contends that even if Jiang presents pruning as an optional embodiment, Apple cannot withdraw pruning from its combination, once having presented that as part of the combination in the Petition. *Id.* at 14 (citing *Bausch & Lomb, Inc. v. Barnes-Hind/Hydrocurve, Inc.*, 796 F.2d 443, 448 (Fed. Cir. 1986)). And according to Zentian, pruning is not optional in Jiang because pruning is necessary to achieve real-time speech-recognition performance. *Id.* at 14–17 (citing Ex. 1003 ¶¶ 207, 209; Ex. 1004, 2:19–28; Ex. 2016, 6:1–8, 6:24–5:2; Ex. 2020 ¶ 40; Ex. 2021, 41:2–6, 38:11–16).

We find Apple’s arguments persuasive that pruning is not required as part of its proposed combination of Jiang, Baumgartner, and Brown. Nothing

in Jiang suggests that pruning is mandatory—it is presented as option, possibly a preferred option, but not a requirement. *See* Ex. 1004, 8:52 (“[A] pruning technique *can* be used.” (emphasis added)). Although Apple relied on the possibility of pruning in Jiang as part of the motivation to combine Jiang’s system with Baumgartner’s parallel-processing hardware, Apple never included pruning as an element needed to meet any limitation of claim 1. *See* Pet. 41, 46; Ex. 1003 ¶¶ 194, 201, 213.

Although Jiang teaches the desirability of using pruning, at least “for a large vocabulary” (Ex. 1004, 2:19), there is nothing in claim 1 limiting how small the acoustic model can be. Also, whether or not a person of ordinary skill in the art could have achieved real-time speech-recognition system without pruning is not pertinent to achieving the claimed invention because claim 1 does not require real-time speech recognition. Apple has persuasively shown that a person of ordinary skill in the art would have had reason to implement Jiang’s system on parallel-processing hardware, as Baumgartner teaches, using pipelining as Brown teaches to meet all the limitations of claim 1, even without Jiang’s pruning embodiment, and that this combination would have been within the ordinary skill in the art.

(3) *Whether Jiang’s Pruning Technique Relies on Feedback*

Zentian notes that, during cross-examination, Mr. Schmandt testified that Jiang’s “calculating circuit” calculates distance scores for all nodes in the search tree prior to the search stage, and then using those scores during the second stage to calculate probabilities. PO Resp. 16–18 (citing Ex. 2016, 64:20–25, 65:1–9, 65:12–16). Zentian disputes that Jiang calculates distance scores for all HMM states “regardless of whether those states were

associated with pruned branches and nodes,” because Jiang never teaches that those distance “calculations are unaffected by its pruning determinations,” and Mr. Schmandt’s testimony that Jiang’s distance scores are unaffected by pruning is untimely. PO Resp. 18–19 (citing Ex. 1003 ¶ 143; Ex. 1004, 8:21–37; Ex. 2016, 15:4–9, 19:8–14, 21:12–22:18, Ex. 2020 ¶ 60).

Zentian also points to language in Jiang that allegedly implies feedback-based pruning, including a statement that during the search stage, “when a branch ‘is dropped (or pruned) from [the lexical tree],’ that branch and the nodes along that branch are ‘*no longer considered in further processing,*’” and that a score is calculated for “each node connected to a phoneme branch *then under consideration* by the speech recognition system.” PO Resp. 19–21 (quoting Ex. 1004, 8:62–64, and then quoting 8:16–20) (citing Ex. 1004, 8:29–37; Ex. 2016, 64:20–25; Ex. 2020 ¶¶ 61–64). According to Zentian, because Jiang allegedly does not calculate scores for every node, a person of ordinary skill in the art would have understood that Jiang is using feedback to decide which nodes to perform calculations on. *Id.* at 21–24 (citing Ex. 1003 ¶ 194; Ex. 1004, 2:5–7; Ex. 1012, 91; Ex. 1050, 3, 11; Ex. 2016, 18:25–19:14, 64:20–65:16; Ex. 2018, 40–42, 69; Ex. 2020 ¶¶ 65–71).

In its reply, Apple argues that Jiang’s pruning technique occurs in the search stage and does not rely on feedback. Pet. Reply 3 (citing Ex. 1004, 8:52–64, 10:65–11:15; Ex. 1064, 21:5–13, 22:6–9; Ex. 2016, 8:11–15). According to Apple, it was known in the art that lexical trees can be pruned at the search stage without feedback, and such that “[p]runing reduces the search space *during the search stage* and can be advantageously performed

regardless of whether the search results are fed back to the classifier of the distance calculation stage.” *Id.* at 2 (citing Ex. 1004, 11:8–12; Ex. 1067, 6:1–9, 8:40–7, Fig. 5; Ex. 1063, 13–14).

Also, Apple argues that “Jiang includes no discussion whatsoever of feeding pruning results from the search stage and to the distance calculation stage.” Pet. Reply 5–6 (citing Ex. 8:52–64, 10:65–11:15; Ex. 2016, 18:16–19:14). Apple contends that Jiang explicitly discloses determining distance calculations for an output feature vector for “*each HMM state*,” rather than just a feedback-determined subset of states, which means that there would have been no role for feedback in Jiang’s disclosure. *Id.* at 6 (citing Ex. 1004, 2:5–7, 7:43–46, 7:37–42; Ex. 1003 ¶ 194).

Apple notes that Jiang never refers to an “active” HMM state, which might arguably imply feedback. Pet. Reply 6–7. As to Zentian’s argument (PO Resp. 20–21) that Jiang’s use of the phrase “a phoneme branch then under consideration by the speech recognition system” (Ex. 1004, 8:18–20) implies the use of feedback, Apple responds that the next paragraph in Jiang makes it clear that Jiang’s system computes a score “*for each phoneme branch considered in [the] tree*,” so Jiang teaches calculating scores for every phoneme in the search tree without using feedback. Pet. Reply 7–8 (quoting Ex. 1004, 8:29–33) (citing Ex. 1004, 1:19–2:46, 7:5–7).

Apple also disputes Zentian’s argument based on the implications of Jiang’s teaching “that once a branch is pruned at the search stage, ‘that branch is dropped (or pruned) from [the search tree] and is no longer considered in further processing.’” Pet. Reply 8 (quoting Ex. 1004, 8:62–64). According to Apple, this language refers to consideration for further processing during the search stage, and does not suggest or imply that

distance calculations were not performed on the pruned branches. *Id.* at 9–11 (citing Ex. 1004, 2:25–46, 8:52–64, 11:8–15, Figs. 4, 6; Ex. 1064, 45:16–20).

We find Apple’s arguments persuasive that a person of ordinary skill in the art would have understood Jiang’s pruning technique not to rely on feedback, particularly because Jiang makes no mention of feedback, and Jiang as a whole appears consistent with Apple’s theory that Jiang does not teach feedback-based pruning. *See, e.g.*, Ex. 1004, 2:5–7, 8:52–64, 10:65–11:15. We credit Mr. Schmandt’s testimony and supporting evidence that “[e]ven though Jiang still describes calculating all scores initially (for each state for example), Jiang explains that pruning thereafter saves on computational resources.” Ex. 1003 ¶ 194 (citing Ex. 1004, 10:65–11:25). We find this testimony more persuasive than Mr. Anderson’s contrary testimony which does not credibly find support in the text of Jiang. *See* Ex. 2020 ¶¶ 57–71.

We also disagree that Apple’s arguments that Jiang’s distance calculations do not rely on pruning are untimely. PO Resp. 19. The Petition argues that Jiang teaches calculating distances, “namely a *score* for each phoneme branch of the prefix tree.” Pet. 21 (citing Ex. 1004, 8:29–33). Also, Mr. Schmandt clearly opined, in his initial declaration submitted with the Petition, that “each phoneme branch is scored” in Jiang’s search tree. Ex. 1003 ¶ 149. Thus, Zentian had clear notice in the Petition of Apple’s argument that Jiang performs distance calculations for the entire search tree. Apple’s later argument that this process does not necessitate feedback is a natural consequence of its argument in the Petition, and is not untimely.

(4) *Apple's Alternative Reply Arguments*

Based on the premise that Apple's proposed combination must include Jiang's pruning embodiment and that it is feedback-based pruning, Zentian contends that Apple has failed to show that Jiang is combinable with Brown's teaching of pipelining. PO Resp. 8–16. In particular, Zentian argues that pipelining is incompatible with Jiang's pruning in which, according to Zentian, feedback from a search stage of current frame N is passed to the calculation stage of the next frame N+1, because pipelining requires the search stage at frame N and the calculation stage at frame N+1 to happen at the same time. *Id.* (citing Ex. 1001, Fig. 18; Ex. 1003 ¶¶ 143, 194, 207, 209; Ex. 1004, 2:5–7, 8:16–37, 8:44–51, 8:62–64, Figs. 5, 6; Ex. 1012, 91; Ex. 1050, 3, 11; Ex. 2016, 5:15–20, 6:1–8, 6:24–5:2, 12:3–13, 12:19–13:1, 15:4–9, 18:11–15, 18:25–19:14, 21:3–22:18, 64:20–65:16; Ex. 2018, 40–42, 69; Ex. 2020 ¶¶ 41–56, 60–71).

In response, Apple presents three alternative arguments based on the assumption that Apple's proposed combination requires pruning and Jiang teaches feedback-based pruning: (1) although Apple's proposed combination relies on Jiang, it does not rely on Jiang's alleged teaching of feedback-based pruning (Pet. Reply 1–2); (2) Jiang's alleged feedback-based pruning would have been compatible with Brown (*id.* at 11–19); and (3) a person of ordinary skill in the art would have had reason to remove feedback-based pruning from Jiang's system in favor of Brown's pipelining (*id.* at 21–27). *See* PO Sur-reply 1–11, 28–31.

Because we determine that Apple's proposed combination does not include Jiang's pruning embodiment, and that embodiment is not feedback-based pruning, we need not address Apple's alternative arguments or

whether they are timely raised. *See* Pet. Reply 27–29 (arguing that Apple’s arguments are timely); PO Sur-reply 26–28 (arguing that Apple’s arguments are untimely).

(f) Conclusion as to Claim 1

For all the above reasons, we find Apple’s arguments persuasive that the combination of Jiang, Baumgartner, and Brown, with or without Smyth, teach all the limitations of claim 1 and that a person of ordinary skill in the art would have had reason to combine the references as Apple proposes, with a reasonable expectation of success. Thus, by a preponderance of the evidence, we determine that claim 1 is unpatentable under § 103(a) as obvious over the combination of Jiang, Baumgartner, and Brown, with or without the further teachings of Smyth.

3. *Claim 5*

Claim 5 depends from claim 1 and further recites “wherein the said calculating circuit is configured to autonomously calculate distances for every acoustic state defined by the acoustic model.” Ex. 1001, 38:41–44.

Apple relies on its arguments for claim 1 and further argues that Jiang’s calculating circuit, as modified by Baumgartner and Brown, autonomously calculates distances for every state of the acoustic model. Pet. 47–48 (citing Ex. 1004, 1:21–25, 2:5–6, 7:30–42; Ex. 1003 ¶¶ 215–217).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claim 5. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claim 5 (Pet. 47–48), Apple has shown by a preponderance of the evidence

that claim 5 is unpatentable as obvious over Jiang in view of Baumgartner and Brown, with or without the further teachings of Smyth.

4. *Claim 7*

Claim 7 depends from claim 1 and further recites “wherein the feature vector comprises a plurality of spectral components of an audio signal for a predetermined time frame.” Ex. 1001, 38:48–50.

Apple relies on its arguments for claim 1 and further argues that Jiang’s feature extraction module encodes feature vectors reflecting spectral characteristics, and the feature vector for each frame is for a predetermined time frame, as recited. Pet. 48 (citing Ex. 1004, 6:62–7:1, 7:7–10; Ex. 1003 ¶¶ 218–220).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claim 7. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claim 7 (Pet. 48), Apple has shown by a preponderance of the evidence that claim 7 is unpatentable as obvious over Jiang in view of Baumgartner and Brown, with or without the further teachings of Smyth.

5. *Claim 12*

Claim 12 depends from claim 1 and further recites “wherein the audio front end is configured to input a digital audio signal.” Ex. 1001, 38:64–65.

Apple relies on its arguments for claim 1 and further argues that Jiang’s feature extraction module, the recited “audio front end,” is provided with a digital signal. Pet. 48 (citing Ex. 1004, 6:46–53; Ex. 1003 ¶¶ 221–222).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claim 12. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claim 12 (Pet. 48), Apple has shown by a preponderance of the evidence that claim 12 is unpatentable as obvious over Jiang in view of Baumgartner and Brown, with or without the further teachings of Smyth.

6. *Claim 14*

Independent claim 14 is as follows:

14. A speech recognition circuit, comprising:
- [a] an audio front end for calculating a feature vector from an audio signal,
 - [b] wherein the feature vector comprises a plurality of extracted and/or derived quantities from said audio signal during a defined audio time frame;
 - [c] calculating means for calculating a distance indicating the similarity between a feature vector and a predetermined acoustic state of an acoustic model; and
 - [d] a search stage for using said calculated distances to identify words within a lexical tree, the lexical tree comprising a model of words;
 - [e] wherein said audio front end, said calculating means, and said search stage are connected to each other to enable pipelined data flow.

Ex. 1001, 39:1–15 (Apple’s reference numbers added).

For the preamble and limitations 14(a)–(e), Apple relies on its arguments for the preamble of claim 1 and limitations 1(a)–(d) and 1(f), respectively, and further argues that limitation 14(c) is a means-plus-function limitation, and that Jiang and Baumgartner teach that the “calculating means” includes a dedicated processor for performing the recited function.

Pet. 49 (citing Ex. 1004, 6:39–42, 8:16–24; Ex. 1003 ¶¶ 226–228). Apple also argues that, for limitation 14(e), Baumgartner teaches that the recited components are interconnected with a bus. Pet. 50 (citing Ex. 1007 ¶ 156; Ex. 1003 ¶¶ 230–231).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claim 14. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claim 14 (Pet. 49–50), Apple has shown by a preponderance of the evidence that claim 14 is unpatentable as obvious over Jiang in view of Baumgartner and Brown, with or without the further teachings of Smyth.

7. *Claim 15*

Independent claim 15 is as follows:

15. A speech recognition method, comprising:
- [a] calculating a feature vector from an audio signal using an audio front end,
 - [b] wherein the feature vector comprises a plurality of extracted and/or derived quantities from said audio signal during a defined audio time frame;
 - [c] calculating a distance indicating the similarity between a feature vector and a predetermined acoustic state of an acoustic model using a calculating circuit; and
 - [d] using a search stage to identify words within a lexical tree using said calculated distances, the lexical tree comprising a model of words;
 - [e] wherein data is pipelined from the front end, to the calculating circuit, and to the search stage.

Ex. 1001, 39:16–40:5 (Apple’s reference numbers added).

Apple relies on its arguments for claim 1 and further argues that Jiang discloses a speech recognition method. Pet. 50 (citing Ex. 1003 ¶ 232).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claim 15. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claim 15 (Pet. 50), Apple has shown by a preponderance of the evidence that claim 15 is unpatentable as obvious over Jiang in view of Baumgartner and Brown, with or without the further teachings of Smyth.

8. *Claim 16*

Independent claim 16 is as follows:

16. A non-transitory storage medium storing processor implementable code for controlling at least one processor to implement a speech recognition method, the code comprising:

- [a] code for controlling the processor to calculate a feature vector from an audio signal,
- [b] wherein the feature vector comprises a plurality of extracted and/or derived quantities from said audio signal during a defined audio time frame;
- [c] code for controlling the processor to calculate a distance indicating the similarity between a feature vector and a predetermined acoustic state of an acoustic model; and
- [d] code for controlling the processor to identify words within a lexical tree using said calculated distances, the lexical tree comprising a model of words,
- [e] wherein data is pipelined by the processor pursuant to the code from the feature calculation, to the distance calculation, and to the word identification.

Ex. 1001, 40:6–22 (Apple’s reference numbers added).

Apple relies on its arguments for claim 1 and further argues, for the preamble, that Jiang discloses that the claimed invention may be implemented on a computer which has RAM or ROM memory to store instructions. Pet. 51 (citing Ex. 1004, 4:54–56, 5:5–29, 6:39–42; Ex. 1003 ¶¶ 238–239). For limitations 16(a) and 16(e), Apple argues that a person of ordinary skill in the art would have understood that Jiang teaches that there is code for performing the recited functions. Pet. 51–52 (reciting Ex. 1003 ¶¶ 240–241).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claim 16. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claim 16 (Pet. 51–52), Apple has shown by a preponderance of the evidence that claim 16 is unpatentable as obvious over Jiang in view of Baumgartner and Brown, with or without the further teachings of Smyth.

D. GROUNDS BASED ON JIANG, BAUMGARTNER, BROWN, KAZEROONIAN, AND OPTIONALLY SMYTH (CLAIM 4)

Apple’s second ground challenges claim 4 as obvious over Jiang, Baumgartner, and Brown, and Kazeroonian, and its fifth ground challenges the same claims based on the same references with the addition of Smyth. Pet. 4. Because of the common issues and evidence, we address these grounds together.

Claim 4 depends from claim 1 and further recites “wherein the first processor supports multi-threaded operation, and runs the search stage and front ends as separate threads.” Ex. 1001, 38:37–40.

Kazeroonian describes “[a]n automated real-time system for processing and distribution of audio-video data.” Ex. 1037, code (57);

Pet. 52 (citing Ex. 1037, 15:33–16:1, 19:30–20:5, 21:9–11, Fig. 2).

According to Apple, Kazeroonian supports multi-threaded operation in which tasks and processing steps are divided among separate processing threads to improve performance. Pet. 53 (citing Ex. 1037, 15:33–16:1, 19:26–20:5, 21:9–11; Ex. 1003 ¶¶ 246–249). Citing testimony by Mr. Schmandt, Apple contends that Kazeroonian’s teachings would have been applicable to systems processing only audio data, as opposed to audio-video data. *Id.* (citing Ex. 1003 ¶ 249). And Apple argues that a person of ordinary skill in the art would have been motivated to run the search stage and the front end of the Jiang–Baumgartner–Brown system using separate threads to avoid data loss, and this would have improved speed and performance. Pet. 53–54 (citing Ex. 1037, 19:30–20:5, 20:10–12; Ex. 1003 ¶¶ 250–252); *see also* Pet. 72 (relying on the same arguments for Petitioner’s fifth ground including Smyth).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claim 4. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claim 4 (Pet. 52–54), Apple has shown by a preponderance of the evidence that claim 4 is unpatentable as obvious over Jiang in view of Baumgartner, Brown, and Kazeroonian, with or without the further teachings of Smyth.

E. GROUNDS BASED ON JIANG, BAUMGARTNER, BROWN, VENSKO,
AND OPTIONALLY SMYTH (CLAIMS 9 AND 10)

Apple’s third ground challenges claims 9 and 10 as obvious over Jiang, Baumgartner, and Brown, and Vensko, and its sixth ground challenges the same claims based on the same references with the addition of Smyth.

Pet. 4. Because of the common issues and evidence, we address these grounds together.

Claim 9 depends from claim 1 and further recites “wherein the speech accelerator has an interrupt signal to inform the front end that the accelerator is ready to receive a next feature vector from the front end.” Ex. 1001, 38:54–57.

Vensko describes “[a] continuous speech recognition system [that] includes a plurality of processors doing template comparisons of speech data.” Ex. 1038, code (57). According to Apple, Vensko’s system generates interrupts to inform processors that processing results are ready for transfer. Pet. 55–56 (citing Ex. 1038, 6:22–29, 6:48–51, 10:10–19; Ex. 1003 ¶¶ 253–259). Relying on testimony by Mr. Schmandt, Apple contends that a person of ordinary skill in the art would have applied Vensko’s teachings to the Jiang–Baumgartner–Brown combination so that the calculating circuit (assumed to be the “speech accelerator”) uses an interrupt signal to inform the front end that the calculating circuit is ready to receive a next feature vector. *See* Pet. 56–57 (citing Ex. 1003 ¶¶ 253–263); *see also* Pet. 72 (relying on the same arguments for Petitioner’s sixth ground including Smyth).

Claim 10 depends from claim 1 and further recites “wherein the accelerator signals to the search stage when the distances for a new frame are available in a result memory.” Ex. 1001, 38:58–60.

Apple argues that, by the same interrupt functionality discussed above for claim 9, the Jiang–Baumgartner–Brown–Vensko system would have used an interrupt to signal when distances for a new frame available in memory for processing. Pet. 58 (citing Ex. 1004, 7:31–34; Ex. 1003 ¶ 264); *see also*

Pet. 72 (relying on the same arguments for Petitioner's sixth ground including Smyth).

Apart from its general arguments that we have discussed above, Zentian makes no particular arguments for claims 9 or 10. For the reasons given above for claim 1 and the additional reasons articulated in the Petition for claims 9 and 10 (Pet. 54–58), Apple has shown by a preponderance of the evidence that claims 9 and 10 are unpatentable as obvious over Jiang in view of Baumgartner, Brown, and Vensko, with or without the further teachings of Smyth.

IV. CONCLUSION¹¹

For the reasons above, Apple has shown by a preponderance of the evidence that claims 1, 4, 5, 7, 9, 10, 12, and 14–16 of the '277 patent are unpatentable under § 103(a).

¹¹ Should Zentian wish to pursue amendment of claims in a reissue or reexamination proceeding after this decision, we draw attention to the April 2019 *Notice Regarding Options for Amendments by Patent Owner Through Reissue or Reexamination During a Pending AIA Trial Proceeding*. See 84 Fed. Reg. 16,654 (Apr. 22, 2019). If Zentian chooses to file a reissue application or a request for reexamination of the challenged patent, we remind Zentian of its continuing obligation to notify the Board of any such related matters in updated mandatory notices. See 37 C.F.R. § 42.8(a)(3), (b)(2).

In summary:

Claim(s)	35 U.S.C. §	Reference(s)/Basis	Claims Shown Unpatentable	Claims Not Shown Unpatentable
1, 5, 7, 12, 14– 16	103(a)	Jiang, Baumgartner, Brown	1, 5, 7, 12, 14–16	
4	103(a)	Jiang, Baumgartner, Brown, Kazeroonian	4	
9, 10	103(a)	Jiang, Baumgartner, Brown, Vensko	9, 10	
1, 5, 7, 12, 14– 16	103(a)	Jiang, Baumgartner, Brown, Smyth	1, 5, 7, 12, 14–16	
4		Jiang, Baumgartner, Brown, Smyth, Kazeroonian	4	
9, 10		Jiang, Baumgartner, Brown, Smyth, Vensko	9, 10	
Overall Outcome			1, 4, 5, 7, 9, 10, 12, 14–16	

V. ORDER

In consideration of the foregoing, it is

ORDERED that claims 1, 4, 5, 7, 9, 10, 12, and 14–16 of the '277 patent are unpatentable;

FURTHER ORDERED that parties to this proceeding seeking judicial review of our decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

IPR2023-00034
Patent 7,979,277 B2

For PETITIONER:

Jennifer C. Bailey
Adam P. Seitz
ERISE IP, P.A.
jennifer.bailey@eriseip.com
adam.seitz@eriseip.com

J. David Hadden
Dargaye Churne
Saina Shamilov
FENWICK & WEST LLP
dhadden@fenwick.com
dchurnet@fenwick.com
sshamilov@fenwick.com

For PATENT OWNER:

Peter C. Knops
Kayvan B. Noroozi
NOROOZI PC
peter@noroozipc.com
kayvan@noroozipc.com

Katherine Rhoades
BARTLIT BECK LLP
katherine.rhoades@bartlitbeck.com