

Docket No.: U1250.70000US00
U.S. Patent No. 8,510,768

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Reexamination of:	U.S. Patent No. 8,510,768	Reexam Control No.: Not yet assigned
First Named Inventor:	Jonathan S. Steuer	Reexam Filing Date: Not yet assigned
Application Serial No.:	12/283,354	Examiner: Not yet assigned
Issued:	August 13, 2013	Art Unit: Not yet assigned
Filed:	September 11, 2008	
For:	MEDIA USAGE MONITORING AND MEASUREMENT SYSTEM AND METHOD	

Certificate of Electronic Filing under 37 C.F.R. § 1.8

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being transmitted via the Office's electronic filing system in accordance with 37 C.F.R. § 1.6(a)(4).

Dated: January 16, 2024

Electronic Signature for Jennifer W. Hibbard: /Jennifer W. Hibbard /

REQUEST FOR EX PARTE REEXAMINATION OF U.S. PATENT NO. 8,510,768

Mail Stop *Ex Parte* Reexam
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Commissioner:

Pursuant to 35 U.S.C. §§ 301-037 and 37 C.F.R. § 1.510 *et seq.*, Unified Patents, LLC (“Unified” or “Requestor”) requests *ex parte* reexamination of claims 1-3, 5-7, 9-11, and 13-15 of U.S. Patent No. 8,510,768 (“‘768 patent;” Ex. A), titled “Media Usage Monitoring and Measurement System and Method,” which issued August 13, 2013, from U.S. Patent Application No. 12/283,354, filed September 11, 2008. The Patent Owner of record is Autonomous Media Research Holdings, LLC, as recorded July 27, 2023, at Reel 064426 and Frame 0829. Pursuant to 37 C.F.R. § 1.510(b)(6), Requestor certifies that the statutory estoppel provisions of 35 U.S.C. 315(e)(1) or 35 U.S.C. 325(e)(1) do not prohibit Requestor from filing this *ex parte* reexamination request.

The *ex parte* reexamination fee of \$12,600.00 is submitted herewith. The undersigned authorizes the Office to charge any deficiency to Deposit Account No. 23/2825.

Roku EX1017
IPR2024-01054
U.S. Patent No. 8,510,768

CONTENTS

I.	IDENTIFICATION OF CLAIMS FOR WHICH REEXAMINATION IS REQUESTED.....	1
II.	THE '768 PATENT	1
A.	Prosecution History	4
III.	CHALLENGED CLAIMS	7
IV.	STATEMENT POINTING OUT EACH SUBSTANTIAL NEW QUESTION OF PATENTABILITY	8
V.	DETAILED EXPLANATION UNDER 37 U.S.C. § 1.510	10
A.	Claims 1-3 and 5-6 Are Obvious Over O'Hagan in view of Apel.....	10
1.	Overview of O'Hagan.....	10
2.	Overview of Apel.....	12
3.	O'Hagan and Apel are Analogous Art.....	14
4.	A POSA Would Have Been Motivated to Modify O'Hagan in view of Apel's Teachings	15
5.	Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim	17
B.	Claims 1-3, 5-6, 9-11 and 13-14 Are Obvious Over O'Hagan in view of Apel and Cheung	35
1.	Overview of Cheung	35
2.	O'Hagan, Apel, and Cheung are Analogous Art	41
3.	A POSA Would Have Been Motivated to Modify O'Hagan in view of Apel's and Cheung's Teachings	42
4.	Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim	44
C.	Claims 7, 9-11, and 13-15 Are Obvious Over O'Hagan in view of Apel and Wang	57
1.	Overview of Wang	57
2.	O'Hagan, Apel, and Wang are Analogous Art	59
3.	A POSA Would Have Been Motivated to Modify O'Hagan in view of Apel's and Wang's Teachings	60
4.	Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim	62
D.	Claims 7 and 15 Are Obvious Over O'Hagan in view of Apel, Cheung, and Wang	67

1. Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim	67
VI. CONCLUSION	68

LIST OF EXHIBITS

Exhibit	Description
A	U.S. Patent No. 8,510,768 (“768 patent”)
B	File History of U.S. Patent Application No. 12/283,354
C	Declaration of David S. Doermann, Ph.D. (“Doermann”)
D	U.S. Patent No. 5,581,658 (“O’Hagan”)
E	U.S. Patent No. 7,284,255 (“Apel”)
F	U.S. Patent Application Publication No. 2007/0168409 (“Cheung”)
G	File History of U.S. Provisional Patent Application 60/547,931 (“Cheung-P”)
H	U.S. Patent Application Publication No. 2002/0083060 (“Wang”)

I. IDENTIFICATION OF CLAIMS FOR WHICH REEXAMINATION IS REQUESTED

In accordance with 37 C.F.R. § 1.510(b)(2), reexamination of claims 1-3, 5-7, 9-11, and 13-15 of the '768 patent is requested, in view of the following references:

U.S. Patent No. 5,581,658 (Ex. D) ("O'Hagan");

U.S. Patent No. 7,284,255 (Ex. E) ("Apel");

U.S. Patent Application Publication No. 2007/0168409 (Ex. F) ("Cheung"); and

U.S. Patent Application Publication No. 2002/0083060 (Ex. H) ("Wang").

Reexamination of claims 1-3 and 5-6 is requested in view of O'Hagan in view of Apel.

Reexamination of claims 1-3, 5-6, 9-11, and 13-14 is requested in view of O'Hagan in view of Apel, and further in view of Cheung.

Reexamination of claims 7, 9-11, and 13-15 is requested in view of O'Hagan in view of Apel and Wang.

Reexamination of claims 7 and 15 is requested in view of O'Hagan in view of Apel, Cheung, and Wang.

A complete listing of the prior art submitted with this Request is provided in the attached Form 1449. In accordance with 37 C.F.R. § 1.510(b)(3), a copy of each prior art reference is submitted herewith in Exhibits D-H. In accordance with 37 C.F.R. § 1.510(b)(4), a copy of the '768 patent is submitted herewith as Exhibit A.

In accordance with 37 C.F.R. § 1.510(b)(2), a detailed explanation of the pertinency and manner of applying the cited prior art to claims 1-3, 5-7, 9-11, and 13-15 of the '768 patent is provided below. The detailed explanation below cites the Declaration of David Doermann, Ph.D. (Ex. C) ("Doermann"), which explains the contents of the cited prior art and the '768 patent in more detail. *See* MPEP §§ 2216, 2217(II), 2258(I)(E).

II. U.S. PATENT 8,510,768

The '768 patent is directed to a media monitoring and measurement system and method. Ex. A, Title, Abstract. Media monitoring measures the consumption patterns of audience members to determine, for example, whether particular advertisements are consumed. Ex. A, 1:14-16, 1:26-32. The '768 patent purports to improve on known media monitoring systems by analyzing additional data, such as "audience data" and "sample sequence data." Ex. A, 1:60-2:15. Fig. 1 below shows a media usage monitoring and measurement system 100 that

analyzes data from various sources including monitoring devices 101a and 101b that capture audio data from a microphone 105 and media player 170b. Ex. A, 2:66-3:54. Doermann, [0032].

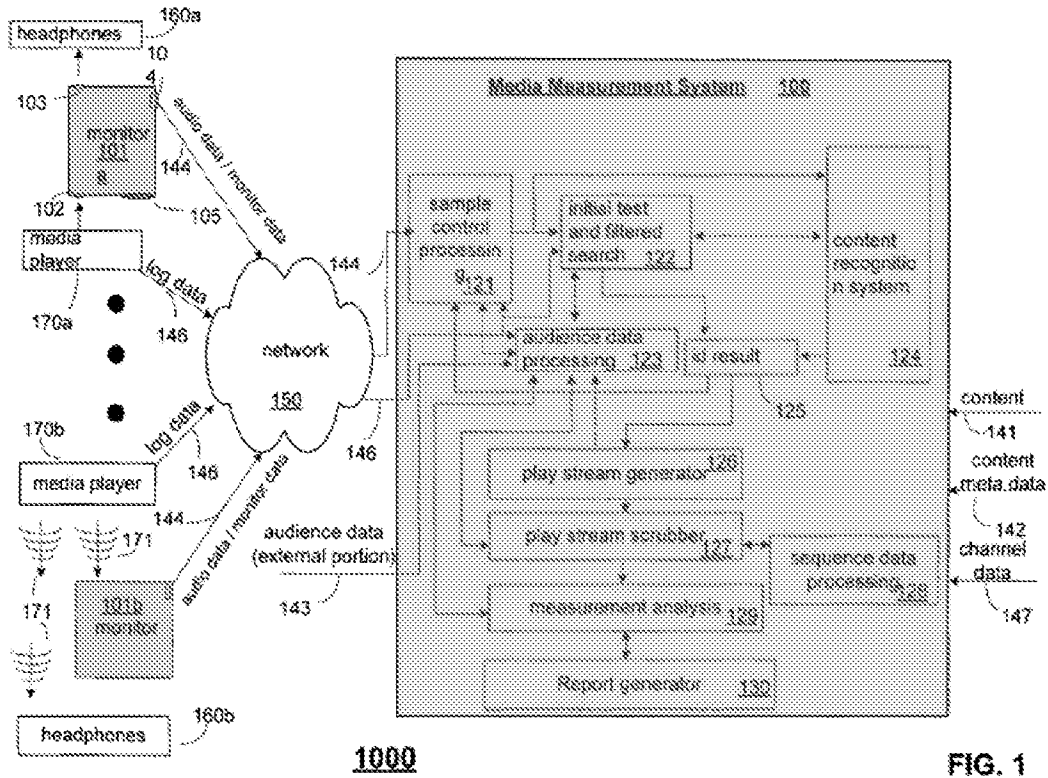


Fig. 2 of the '768 patent, below, illustrates a media measurement method:

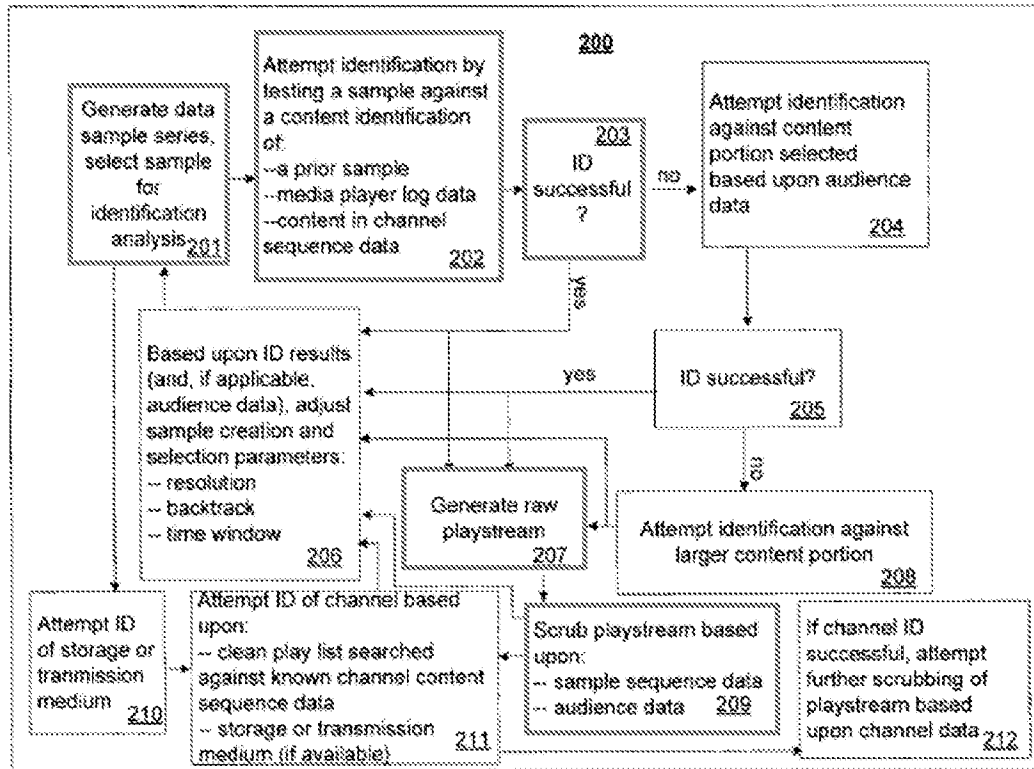


FIG. 2

The system receives audio data from a monitoring device, and a series of samples are generated for identification analysis. Ex. A, 6:62-65 (step 201). An attempt is made to identify the sample audio data. Ex. A, 7:16-24 (step 202). If identifications are successful, then a raw playstream indicating a series of samples and identification results is generated. Ex. A, 8:1-3 (step 207). The raw playstream is later scrubbed based on sample sequence data and audience data. Ex. A, 8:3-6 (step 209). Scrubbing the raw playstream results in a clean playstream that may be converted to a play list “listing the content captured and indicating the order in which it was captured. Ex. A, 8:6-8. Doermann, [0033], [0034].

The '768 patent describes performing scrubbing to correct incorrect identifications or supply missing identifications. Ex. A, 9:32-37. For example, if a raw playstream includes seven consecutive samples and identifies song51 for samples 1, 3, 4, 6, and 7, does not successfully identify content for sample 2, and identifies movie81 for sample 5, scrubbing the raw playstream will “assume that the person using the corresponding monitor listened to song51 through the time period associated with the audio data of samples 1-7, and thus, the ‘scrubbed’ version of the play stream, i.e. Clean Play Stream 450, indicates song51 for each of samples 1-7.” Ex. A, 9:38-51.

See Fig. 4 below (annotated to show the assumed errors in rows 2 and 5 of raw play stream 410 and their corrected content IDs in clean play stream 450). Doermann, [0035].

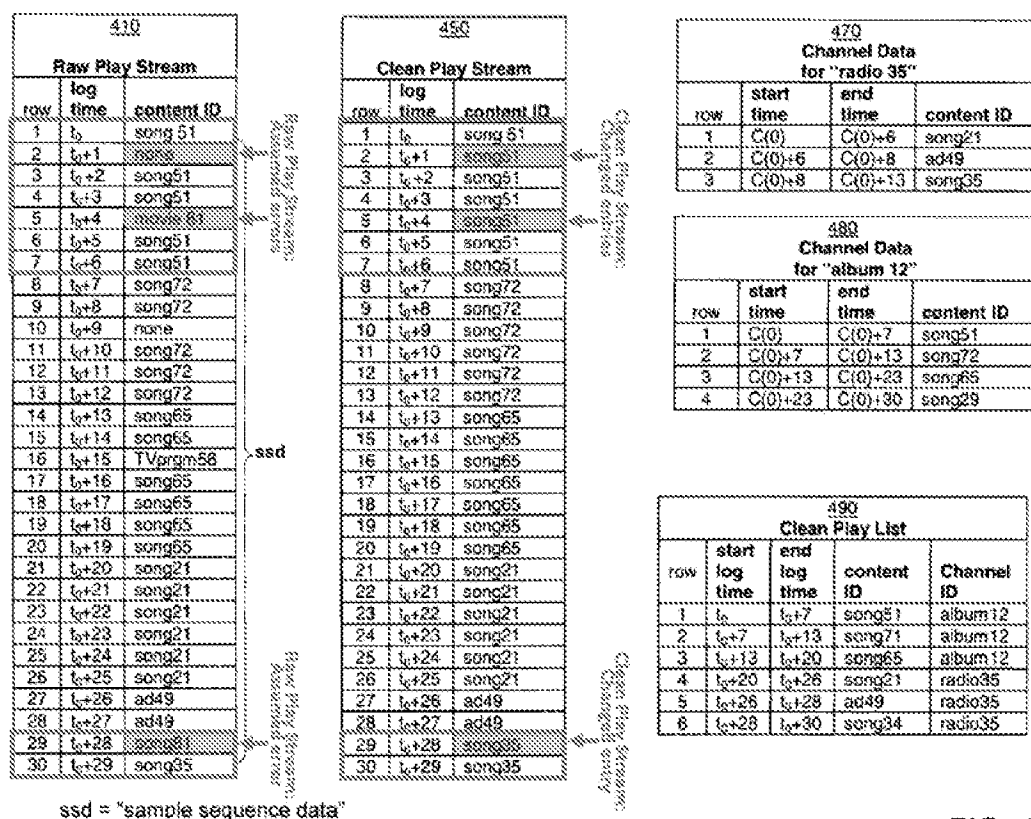


FIG. 4

"Audience data" may also be used in the scrubbing process. For example, a user's consumption habits may be used in the scrubbing process. Raw playstream 410 in Fig. 4 above includes an identification of song81 for sample 29 and song35 for sample 30. If song35 is within a user's consumption habits while song 81 is radically different from those consumption habits, then "another basis might be provided for suspecting that 'song35' rather than 'song81' is the correct content ID for sample 29 of Play Stream 410." Ex. A, 10:28-33. Audience data may be data that is collected internally or externally, and examples include age, gender, consumption patterns of other audience members, and location. Ex. A, 3:55-64, 5:11-28. Doermann, [0035].

A. Prosecution History

During prosecution, the Examiner repeatedly rejected all pending claims as obvious over combinations of prior art, necessitating significant amendments and new combinations of art at

each stage of prosecution. Ex. B, 80-85 (Rejection of 1/3/2011¹), 118-123 (Amendment of 5/3/2011), 137-146 (Rejection of 8/5/2011²), 311-316 (Amendment of 2/6/2012), 329-340 (Rejection of 10/2/2012³), and 367-373 (Amendment of 4/1/2013). Doermann, [0036].

Applicant's final Amendment of 4/1/2013 cancelled all eight pending claims and added fifteen new claims based on a telephonic interview in which the Examiners agreed to allowable subject matter. Ex. B, 368-372, 381-383. There is no discussion on the record of any particular aspect of the prior art in relation to any particular aspect of the newly added claims. As shown in the table below, Applicant made significant changes to the independent claim's scope.

Doermann, [0037].

Last Rejected Independent Claim (introduced 2/6/2012; clean version shown, amendment formatting and removed claim language omitted; see Ex. B, 312)	Allowed Independent Claim (introduced 4/1/2013; see Ex. B, 368-369) (bold emphasis added)
1. (Currently Amended) A media measurement method comprising: generating a raw play stream including a sequence of content identification results corresponding to a sequence of a plurality of audio data samples captured at a media monitoring device, wherein generating the raw play stream comprises (i) receiving the plurality of audio data samples at one or more computers; (ii) with respect to an electronic database including a plurality of audio data representations, each of the plurality of audio data representations being associated with a content identifier, using one or more computers to query the electronic database to obtain a sequence of content identification results by searching the electronic database to obtain respective likely matches between	9. (New) A media measurement method comprising: (a) receiving over a network, at one or more computers, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network; (b) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers; (c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:

¹ Including rejections based on Lamb (U.S. 5,437,050), Wang (U.S. 2002/0083060; Ex. H), and Hansen (U.S. 2002/0038456).

² Including rejections based on Deng (WO 2005/036877), Hansen, Lee (WO 2005/006768), Ellis (U.S. 5,436,653), Barton (U.S. 2002/0072982), and Herley (U.S. 2004/0001161).

³ Including rejections based on Iyoku (U.S. 2004/0055445), Deng, Hansen, Lee, Ellis, Barton, and Wang.

respective audio data samples of the plurality of audio data samples and respective audio data representations of the plurality of audio data representations, the respective audio data representations corresponding to respective content identifiers; and (iii) using the respective content identifiers to generate the sequence of content identification results in the raw play stream; and scrubbing the raw play stream by analyzing sample sequence data including the sequence of content identification results, wherein analyzing includes: determining at least certain results to be correct in view of their sequence relative to other results in the sequence of content identification results; and determining other results, if any, to be incorrect in view of their sequence relative to other results in the sequence of content identification results.	the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database; a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; (d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and (e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.
---	---

A Notice of Allowance issued on 4/24/2013, reciting the entirety of the newly added independent claim 9 as allowable subject matter without further explanation.

Although no discussions in the prosecution history reveal that any particular limitation was important or even discussed in overcoming the art of record, a comparison of the newly added independent claim with the cancelled claim now *suggests* that, in addition to the whole of the claims being allegedly non-obvious over the art of record, the allowable subject matter may be related to the functionality of the scrubbing. For example, Applicant had previously argued that the Deng prior art “does not disclose scrubbing ‘a sequence of content identification results.’” Ex. B, 315. Allowed new claim 9 claimed that the raw play stream includes a sequence of two different content identifiers or one content identifier and an indication of the absence of a likely match, and that scrubbing the raw play stream determines whether to change a result of the sequence. Additionally, allowed new claim 9 recited generating a clean play stream based on the scrubbing. Neither of these concepts were explicitly claimed in the now-cancelled claims and *may* have served as a basis for the allowance. Doermann, [0039].

III. CHALLENGED CLAIMS

Independent claim 1 is reproduced below with corresponding labels as used in this Request:

1[pre]	1. A media measurement method comprising:
1[a]	(a) receiving over a network, at one or more computers, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network;
1[b]	(b) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers;
1[c]	(c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:
1[c.i]	the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database;
1[c.ii]	a content identification result of the sequence of content identification results comprises either: (i) a content

	identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and
1[c.iii]	the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database;
1[d]	(d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and
1[e]	(e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.

Independent claim 9 is directed to a computer program product including instructions executable by one or more processors of one or more computers for performing a method that is substantially similar to the method claimed in independent claim 1.

The dependent claims recite additional features such as: “sample sequence data includes at least one of log times, content identifiers, and content offsets” (claims 2, 10); “scrubbing the raw play stream further comprises analyzing channel data ...” (claims 3, 11); “a clean play list is generated from the clean play stream” (claims 5, 13); “the plurality of audio data representations include a plurality of audio signatures” (claims 6, 14); and “an audio data representation ... includes a plurality of signal landmarks” (claims 7, 15).

IV. STATEMENT POINTING OUT EACH SUBSTANTIAL NEW QUESTION OF PATENTABILITY

To secure allowance of the claims, Applicant introduced significant amendments to the claims—replacing all pending claims in its last amendment. As discussed in relation to the ’768 patent’s prosecution history in Section II.A above, although no discussions reveal that any particular limitation was important in overcoming the art of record, the language of the allowed

claim *suggests* that, in addition to the whole of the claims being allegedly non-obvious over the art of record, the Examiner may have believed that the prior art of record did not disclose scrubbing a raw play stream that includes a sequence of two different content identifiers or one content identifier and an indication of the absence of a likely match to change a result of the sequence, thus generating a clean play stream. However, these concepts—and indeed the claims as a whole—are found in the prior art and well-known to persons of ordinary skill in the art (“POSAs”). For example, O’Hagan (Ex. D) discloses an identification system configured to produce a file recording the details of successive classified audio signatures and then later revisit and classify any audio signatures that could not be classified earlier—*i.e.*, scrubbing the file records. Apel (Ex. E) further discloses the analysis of sequence data to determine how to classify an audio sample. And Cheung (Ex. F) discloses other well-known implementation details that would have been practiced in these similar systems.

Neither O’Hagan, Apel, nor Cheung were before the Examiner during prosecution of the ’768 patent. Although Wang was used as a secondary reference in some rejections during prosecution of the ’768 patent, Wang was not considered in view of O’Hagan, Apel, or Cheung, which provide the new teachings discussed herein and were not before the Examiner. As demonstrated herein, the prior art renders obvious the claims of the ’768 patent, including the added limitations that may have served as a basis for allowance. O’Hagan, Apel, and Cheung therefore present substantial new questions of patentability of the claims of the ’768 patent.

SNQ 1: O’Hagan teaches that raw play streams having unidentified or unmatched audio data need to be scrubbed to identify the unidentified or unmatched audio data. This would have been important to a reasonable examiner.

SNQ 2: Apel teaches scrubbing including analyzing a sequence of a raw play stream and comparing the sequence to an expected sequence to identify unidentified or unmatched audio data. This would have been important to a reasonable examiner, as would the known combination with systems such as O’Hagan.

SNQ 3: Cheung teaches content identification results including (i) content identifiers associated with respective audio data representations, and (ii) indications of the absence of respective likely matches to audio data representations. This would have been important to a reasonable examiner, as would the known combination with systems such as O’Hagan and Apel.

The table below summarizes the proposed rejections based on the above substantial new questions of patentability.

References		Claims	Basis
1	O'Hagan (Ex. D) in view of Apel (Ex. E)	1-3, 5-6	§ 103
2	O'Hagan (Ex. D) in view of Apel (Ex. E) and Cheung (Ex. F)	1-3, 5-6, 9-11, 13-14	§ 103
3	O'Hagan (Ex. D) in view of Apel (Ex. E) and Wang (Ex. H)	7, 9-11, 13-15	§ 103
4	O'Hagan (Ex. D) in view of Apel (Ex. E), Cheung (Ex. F), and Wang (Ex. H)	7, 15	§ 103

Doermann, [0039]-[0195].

V. DETAILED EXPLANATION UNDER 37 U.S.C. § 1.510

A. Claims 1-3 and 5-6 Are Obvious Over O'Hagan in view of Apel

1. Overview of O'Hagan

O'Hagan (Ex. D) qualifies as prior art under at least pre-AIA 35 U.S.C. § 102(b) because it issued on 12/3/1996, more than one year prior to the '768 patent's earliest possible priority date of 5/27/2004.⁴ O'Hagan is directed to the automatic identification of broadcast performances, such as music, commercials, and the like. Ex. D, 1:9-12. Identification systems are used, for example, to measure the broadcast frequency of any particular program and to provide information about advertising audience numbers, commercial repetitions, and licensed broadcast performances. Ex. D, 1:14-24. O'Hagan describes a typical prior art identification system with respect to Fig. 1 (below), which includes a *station monitoring unit 20* for monitoring broadcasts and *signature and recognition post-processing unit 64* for identifying the broadcasts.

Doermann, [0043].

⁴ The '768 patent claims priority to U.S. provisional application no. 60/574,836, filed 5/27/2004. Requestor takes no position as to the propriety of the priority claims since the art presented herein pre-dates the earliest filing.

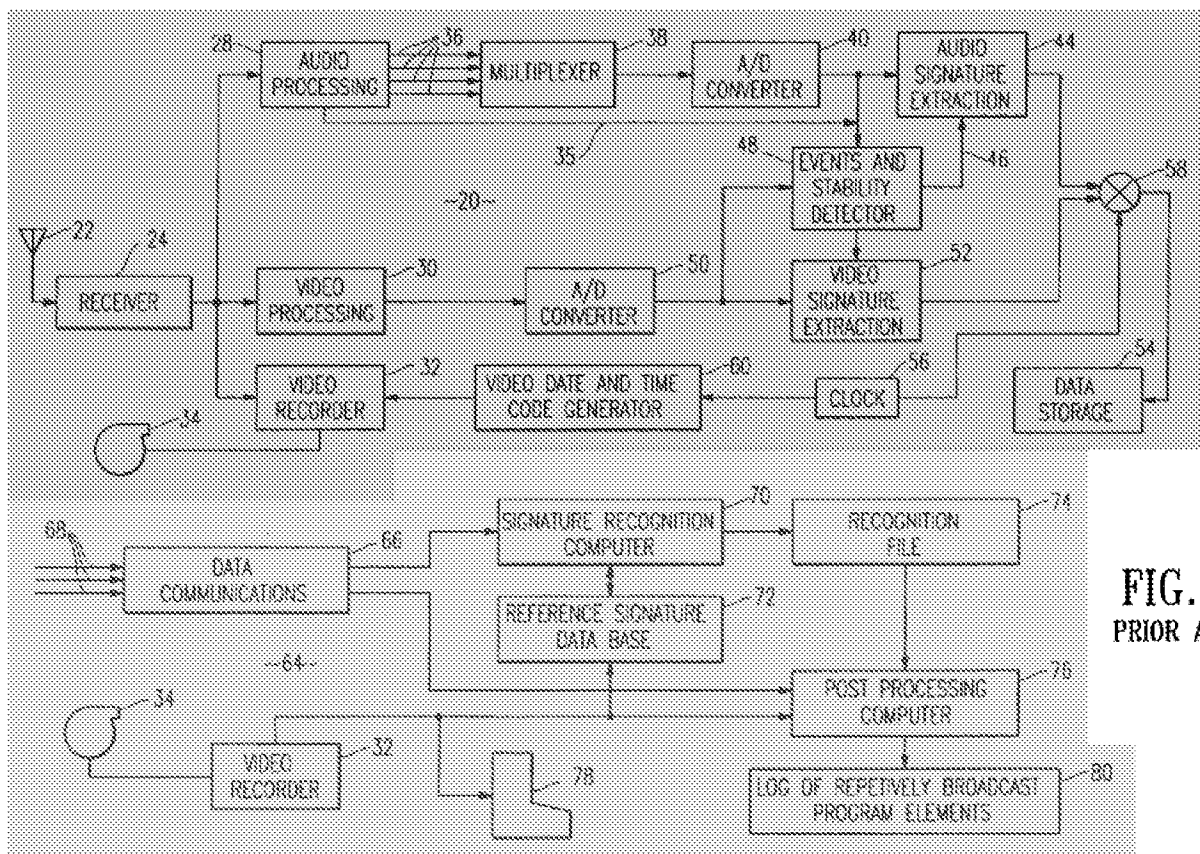


FIG. 1
PRIOR ART

Station monitoring unit 20 receives radio signals (via antenna 22 and receiver 24) and processes the signals (via at least audio processing circuitry 28 and audio signature extraction circuit 44) to extract audio signatures that are based, in part, on the energies found in different frequency bands of the received radio signals. Ex. D, 5:1-22. The extracted audio signatures are stored along with their timestamps. Ex. D, 5:30-36. Doermann, [0044].

Signature and recognition post-processing unit 64 receives signatures from station monitoring units 20 (via communication lines 68) and compares the signatures with a plurality of reference signatures to identify corresponding broadcast programs (via at least signature recognition computer 70 and reference signature database memory 72). Ex. D, 5:40-51. Signature recognition computer 70 “searches for instances of successive recognition of extracted signatures to identify a particular program” and classifies and stores the signatures in recognition file 74. Ex. D, 5:52-64. Doermann, [0045].

The recognition file 74 is further provided to post-processing computer 76, which produces a file or log 80 of identified programs along with their start times and durations. Ex. D, 5:64-6:3. Additionally, post-processing computer 76 assigns a code to non-recognized or

unmatched signatures for later manual review and identification. Ex. D, 6:7-10. Doermann, [0045].

2. Overview of Apel

Apel (Ex. E) qualifies as prior art under at least pre-AIA 35 U.S.C. § 102(e) because it is a patent granted on U.S. patent application 09/441,539, filed 11/16/1999, which is before the '768 patent's earliest possible priority date of 5/27/2004. Apel is directed to systems and methods for performing audience surveys of broadcast audio. Ex. E, Abstract. Survey systems are used, for example, to determine the popularity of programs and broadcast stations, determine advertising pricing structures, and the like. Ex. E, 1:27-32. Apel describes a system with respect to Fig. 1 (below), which includes portable monitoring units 4 and portable units and docking stations 10 for monitoring broadcasts and central computer 7 for identifying the broadcasts. Doermann, [0046].

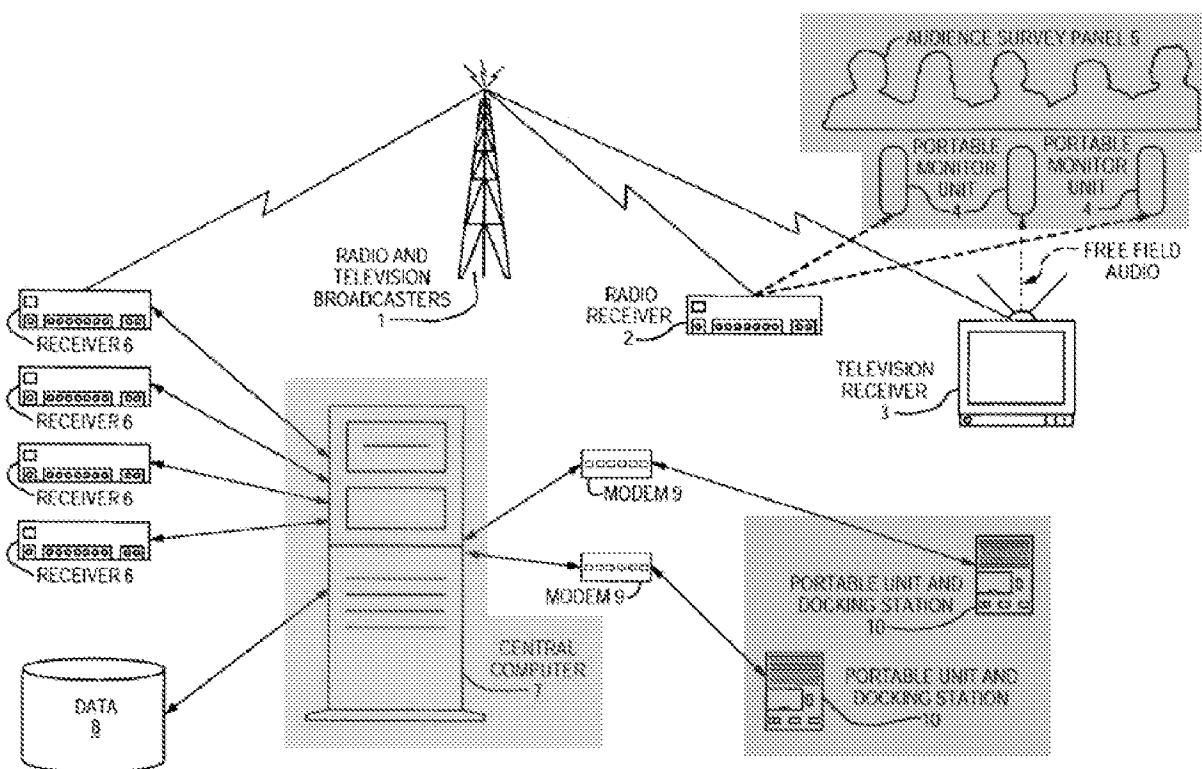


FIG. 1

Portable monitoring units 4 include body-worn units that periodically sample the audio environment and store features representing the structure of the audio. Ex. E, 6:19-25. For example, audio signals received by a microphone in a portable monitoring unit 4 are amplified, low-pass filtered, digitized, and sent to a microprocessor for audio processing and feature

extraction. Ex. E, 7:58-67. Audio processing and feature extraction includes partitioning audio into frequency bands and determining a measure of the power in each of the bands to provide audio data measurements that are used as features in the pattern recognition process. Ex. E, 8:1-29. When *portable monitoring units 4* dock into *docking stations 10*, they transfer the audio feature data, which includes timer values and is accompanied by location data, to *central computer 7* via dialup modems 9. Ex. E, 3:30-39, 5:35-41, 6:28-33, 9:12-38, 9:50-58. In a larger market, data is also transferred over a wide area network. Ex. E, Fig. 6, 7:8-28. Doermann, [0047].

Central computer 7 then compares packets downloaded from the *portable monitoring units 4* with sections of audio files that occurred at the same date and time. Ex. E, 9:62-65. In particular, *central computer 7* generally generates the audio files used for comparison in the same way that the *portable monitoring units 4* generates its audio feature data, except that *central computer 7* must generate audio files for the several audio sources that are available. Ex. E, 9:65-10:45. Comparisons are performed by first synchronizing the times of audio features from the *portable monitoring units 4* with audio files stored by *central computer 7*. Ex. E, 9:62-65, 10:46-51, 11:48-56. Then a pattern recognition process is performed to determine a best match with reference patterns from the several sources to ultimately determine that a specific broadcast was being listened to at a particular time. Ex. E, 3:53-57, 4:15-25, 6:31-37, 7:23-29, 11:66-13:49. If a result of a comparison meets a minimum threshold, then the corresponding source identification is recorded; otherwise, the source is recorded as unknown. Ex. E, 12:5-14, 13:43-49. In cases where matches are not initially found, *central computer 7* will “repeat all of the pattern matches using an expanded search window.” Ex. E, 11:54-61. Doermann, [0048].

To reduce the amount of broadcast data that must be transferred for analysis, if an initial identification is made for one packet, then attempts are made to match subsequent packets with broadcast source features from the same channel as the initial identification. Ex. E, 7:28-34. This is possible because, typically, “a listener will continue to listen (or stay tuned) to the same station for some amount of time.” Ex. E, 7:34-37. Doermann, [0049].

3. O'Hagan and Apel are Analogous Art

O'Hagan and Apel are analogous to the '768 patent because they are in the same field of endeavor as the '768 patent and reasonably pertinent to the alleged problems discussed therein.⁵

For example, the '768 patent is in the field of “media monitoring and measurement” and the “recognition” of media content (*see* Ex. A, Abstract, 1:60-2:15), and, similarly, O'Hagan is in the field of “monitoring, identifying, classifying and logging musical work performance broadcasts” (*see* Ex. D, Abstract, 1:9-12) and Apel is in the field of “identifying ... audio sources” and the “determination” of “which broadcast audio signals ... [are] listen[ed] to at different times of the day” (*see* Ex. E, 1:10-23). Doermann, [0050].

Moreover, O'Hagan seeks to improve on prior art systems that rely on “special codes” or “predictable cues” by providing “high recognition over all portions of [a] musical work” and, furthermore, seeks to recognize a variety of musical works, including works from radio stations, CD players, audio tracks from cinema media, and the like. Ex. D, 3:58-66, 6:22-37. O'Hagan's disclosure would therefore have been reasonably pertinent to problems alleged in the '768 patent, which include the desire to track usage “across several types of media” such as radio, CDs, and others, and to avoid relying on the use of “codes” that “are limited in that they can only track content that has been properly encoded.” *See* Ex. A, 1:14-36. Doermann, [0050].

Likewise, Apel seeks to improve on prior art systems that rely on the “injection of a coded identification signal within the audio” and therefore Apel's disclosure would have been reasonably pertinent to problems alleged in the '768 patent, which include prior art reliance on the use of “codes” that “are limited in that they can only track content that has been properly encoded.” *See* Ex. E, 1:48-2:25; Ex. A, 1:33-36. Additionally, Apel teaches that its body-worn portable monitoring units “sample the acoustic environment ... using a microphone. Ex. E, 3:12-21. Thus, Apel's disclosure would have also been reasonably pertinent to other problems alleged

⁵ A reference is analogous art to the claimed invention if: (1) the reference is from the same field of endeavor as the claimed invention (even if it addresses a different problem); or (2) the reference is reasonably pertinent to the problem faced by the inventor (even if it is not in the same field of endeavor as the claimed invention). *See In re Bigio*, 381 F.3d 1320, 1325, 72 USPQ2d 1209, 1212 (Fed. Cir. 2004) (describing the “[t]wo separate tests”).

in the '768 patent, which include false positives from determining that a person is tuned into a particular channel but actually “mute[s] a broadcast during commercial periods” (thus “not adequately track[ing] whether the audience member listened to a particular advertisement”) and having the ability to “track usage of particular media content independent of channel” and “across several types of media.” *See* Ex. A, 1:14-32. Doermann, [0050].

Accordingly, O'Hagan and Apel are each analogous to the '768 patent.

4. A POSA Would Have Been Motivated to Modify O'Hagan in view of Apel's Teachings

While O'Hagan discloses a system in which several station monitoring units 20 communicate with a signature and recognition post-processing unit 64 over “communication lines 68,” O'Hagan does not explicitly describe those communications as being over a network. Ex. D, 5:47-51. Additionally, while O'Hagan discloses that its signature and recognition post-processing unit 64 codes non-recognized or unmatched extracted signatures so that a later manual review and identification of those signatures can be performed, O'Hagan does not explicitly describe the specifics of that additional review process. Ex. D, 6:7-10. Apel teaches the use of network communications, and Apel further teaches supplementing initial identifications in an audio identification processes with matches from the same channel. Ex. E, 3:30-39, 5:35-41, 6:28-33, 7:8-37, 9:12-38. A POSA would have been motivated to modify O'Hagan's system and method to utilize network communications to broaden the range of potential customers and to match non-recognized or unmatched signatures with broadcasts from the same channel as an initial identification, as taught by Apel, for numerous reasons. Doermann, [0051].

First, networks facilitate communications between devices that are located remotely from each other. Apel teaches that its portable monitoring unit and central computers communicate via modems and telephone lines—*i.e.*, over a public switched telephone network. Ex. E, 3:30-39, 5:35-41, 6:28-33, 9:12-38. The use of telephone networks advantageously allows for the remote transfer of data at any time, such as “late at night or early in the morning when the portable monitoring unit is not in use and the phone line is available.” Ex. E, 3:33-38, 6:28-31. Moreover, Apel teaches that in an “expanded survey system that is intended to operate in multiple cities or markets,” groups of portable monitoring units and central computers are connected via “wide area network.” Ex. E, 5:46-54, 7:8-26. The use of larger scale wide area networks advantageously allows for combined access to users in geographically diverse and remote areas

and facilitates the sharing of processing loads by networked computers. A POSA would have therefore been motivated to modify O'Hagan's communication lines to utilize one or more networks to remotely access audio data from users in different locations, as taught by Apel. Doermann, [0052]. Moreover, the modification would have been, at least, the application of a known technique (*e.g.*, Apel's teaching of communication over a network) to a known system ready for improvement (*e.g.*, O'Hagan's non-limited disclosure of communication lines) to yield predictable results (*e.g.*, the communication of data from remote locations). Doermann, [0053]. *See* MPEP § 2143 (citing *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007)).

Second, O'Hagan contemplates reviewing non-recognized or unmatched extracted signatures for identification but leaves to a POSA how that additional review would be performed. Apel "assum[es] that a listener will continue to listen (or stay tuned) to the same station for some amount of time," and so "[o]nce an initial identification has been made, it is confirmed by matching subsequent packets with broadcast source features from the same channel as the previous recognition." Ex. E, 7:30-37. This shorthand matching saves processing time and energy by preemptively and automatically having a best guess in hand that only needs to be confirmed as a match, rather than requiring the performance of a new and separate manual analysis and identification based on original audio recordings. Ex. D, 5:4-9, 6:9-10. A POSA would have therefore been motivated to modify O'Hagan's system to match non-recognized or unmatched signatures with subsequent packets from the same channel of a previous recognition in order to take advantage of the known behavioral characteristic of users to listen to the same music or station for an amount of time, as taught by Apel. Doermann, [0054]. Moreover, the modification would have been, at least, the application of a known technique (*e.g.*, Apel's teaching of subsequent matching based on the assumption that a user will remain on the same channel) to a known system ready for improvement (*e.g.*, O'Hagan's non-limited disclosure of subsequent review of unmatched signatures that leaves open to the POSA how to perform the subsequent review) to yield predictable results (*e.g.*, the matching of unmatched signatures based on the channel of previous matched signatures). Doermann, [0055]. *See* MPEP § 2143 (citing *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007)).

A POSA also would have had a reasonable expectation of success modifying O'Hagan's system to communicate over a network and to scrub the raw play stream based on an expected pattern of sample sequence data. For example, Apel specifically describes using a frame relay

network, and a POSA would have recognized that a variety of other network types were also well-known and used for audio recognition. Ex. E, Fig. 6; Ex. H, 6:47-7:8 (describing connections via LAN, WAN, enterprise-wide office networks, intranets, and the Internet, and through devices including servers, routers, modems, and other devices). Additionally, Apel teaches steps to capture channel data, *e.g.*, via collection and storage of extracted feature data and their timings. Ex. E, 9:59-10:4. The storage and analysis of data was a basic computer function already practiced and described by O'Hagan (*see* Ex. D, 5:52-58, describing the comparison of extracted signatures to stored signatures in a reference database), and the storage, retrieval, and analysis of additional data would have been well within the skill of a POSA. Doermann, [0056].

Claims 1-3 and 5-6 are therefore obvious over O'Hagan in view of Apel as discussed herein. Doermann, [0057]. The examiner did not appear to have these teachings before him when allowing the claims, which seemingly were allowed for teaching a raw playstream including unmatched samples and a scrubbing correction mechanism for changing those unmatched samples.

5. Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim

a. Claim 1

i. 1[pre]: A media measurement method comprising:

To the extent that the preamble of claim 1 is limiting, O'Hagan discloses it. Doermann, [0058].

O'Hagan discloses a "broadcast program identification" system and method. Ex. D, 4:65-66. The method includes, for example, receiving broadcast radio signals (*i.e.*, media) and identifying corresponding broadcast programs for the purpose of determining the sizes of their audiences and repetitions (*i.e.*, measuring their performances). Ex. D, 4:66-5:4, 5:40-47. Accordingly, O'Hagan discloses a media measurement method. Doermann, [0059].

- ii. 1[a]: (a) receiving over a network, at one or more computers, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network;**

O'Hagan in view of Apel renders obvious this limitation. Doermann, [0060].

O'Hagan discloses receiving over a network (*e.g.*, receiving over communication lines 68), at one or more computers (*e.g.*, receiving at ~~signature and recognition post-processing unit 64~~), data for a sequence of audio data samples (*e.g.*, successively extracted and stored signatures and times of extraction for monitored audio signals from a broadcast radio antenna), the audio data samples comprising representations of audio data captured at a media monitoring device (*e.g.*, the signatures represent energy found in different frequency bands of the audio signals captured at ~~station monitoring unit 20~~), the plurality of audio data samples being submitted over the network (*e.g.*, the data is transmitted over the communication lines 68). Doermann, [0061].

For example, O'Hagan discloses that ~~station monitoring unit 20~~ includes an antenna 22 and a receiver 24 (which includes a tuner and amplifier) for receiving broadcast radio signals and producing a program signal 26 from a monitored broadcast station, *i.e.*, sampling the broadcast radio signals to produce the program signal. Ex. D, 5:1-6; *see also* Ex. D, 6:52-57, 10:38-43 and Ex. E, 3:14-21, 4:38-5:34, 7:58-8:53 (audio that is monitored and that has its information recorded is sampled; audio waveforms, their extracted features, and time-divided portions thereof are all samples). The program signal 26 is processed by audio processing circuitry 28, multiplexer 38, A/D converter 40, events detector 48, and audio signature extraction circuit 44 to: detect energies representative of audio signals within specified frequency bands of the program signal, digitize the audio signals, and "extract an audio signature from [the] digitized audio signals" based on the detected energies. Ex. D, 5:10-22; *see also* Ex. D, 7:63-8:16 (describing a spectral signature of audible musical work performances useful for identifying the musical work). The extracted audio signatures are stored with their extraction times and dates. Ex. D, 5:30-34; *see also* Ex. D, 9:30-34 (describing identification results including a time of signal reception of an original audible signal from its source; further describing an identification library storing catalog information). O'Hagan further discloses that "successive" extracted audio

signatures may be recognized and identified as belonging to a particular program, and therefore the extracted audio signatures are a sequence of audio data samples that, collectively, represent audio data of the one or more programs from which they were generated. Ex. D, 5:58-64. Doermann, [0062].

At a later time, station monitoring unit 20 submits the stored extracted signatures to signature and recognition post-processing unit 64, which receives them “by way of several communication lines 68.” Ex. D, 5:40-51. The signature and recognition post-processing unit 64 is a computer because, *e.g.*, it analyzes extracted signatures with signature recognition computer 70, compares extracted signatures to reference signatures in reference database memory 72, and stores classification results in recognition file 74. Ex. D, 5:52-61. Doermann, [0063].

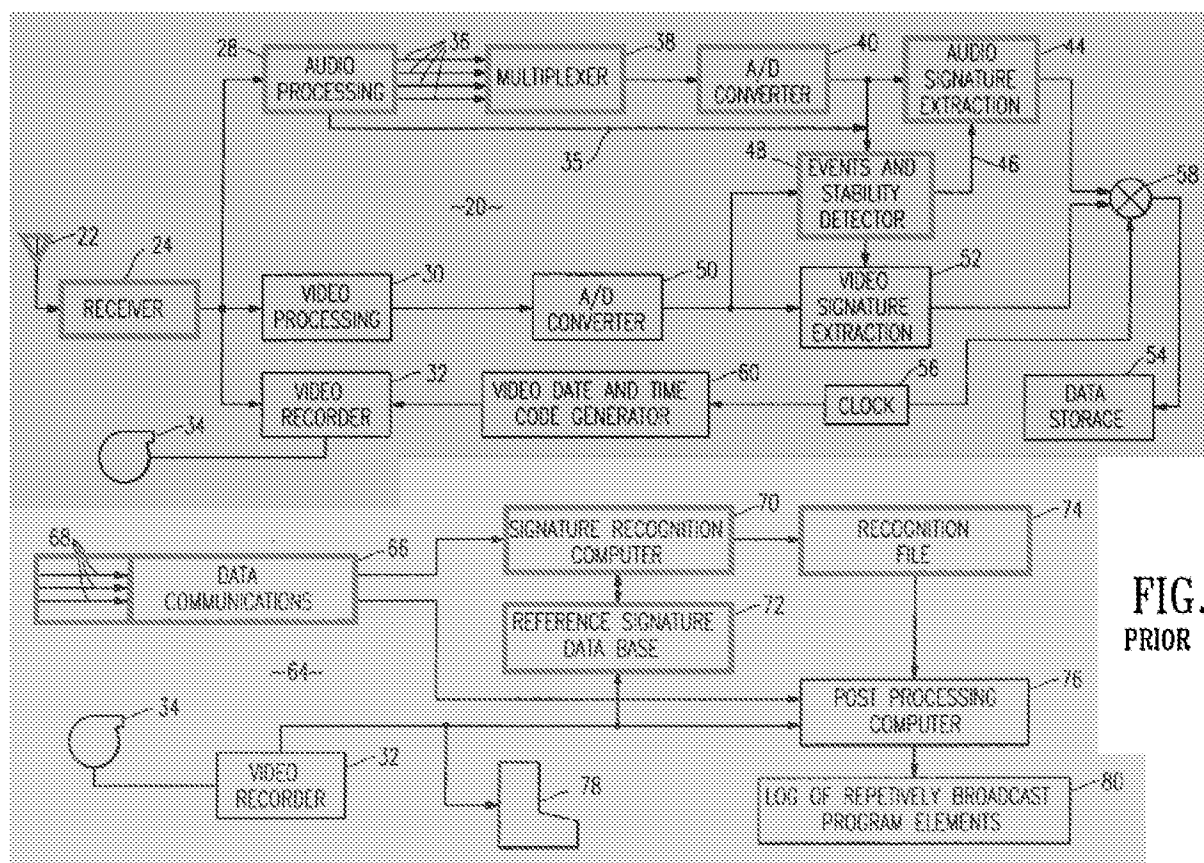


FIG. 1
PRIOR ART

Apel teaches that central computers receive audio feature samples representative of the structure of audio presented to wearers of portable monitoring units via “telephone connections,” “local network,” or “wide area data communication network.” Ex. E, 3:30-39, 5:35-41, 5:46-54, 6:28-33, 7:8-26, 9:12-38. Doermann, [0064].

To the extent Patent Owner argues that O'Hagan's communication lines 68, with which signature and recognition post-processing unit 64 communicates with a plurality of station monitoring units 20, do not explicitly include communications over a network, a POSA would have been motivated to modify O'Hagan to utilize network communications for the remote transfer and centralized processing of information, as discussed in Section V.A.4 above, in light of Apel's teachings. Doermann, [0065].

iii. 1[b]: (b) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers;

O'Hagan discloses this limitation. Doermann, [0066].

O'Hagan discloses using the one or more computers (*e.g.*, using signature and recognition post-processing unit 64, which includes signature recognition computer 70) to query an electronic database (*e.g.*, accessing a reference signature database memory 72 to compare data) of a plurality of audio data representations (*e.g.*, a library of reference signatures used for comparison) and corresponding content identifiers (*e.g.*, information used for classifying recognized signatures). Doermann, [0067].

For example, O'Hagan discloses that its signature and recognition post-processing unit 64 includes signature recognition computer 70 and reference signature database 72. Doermann, [0068].

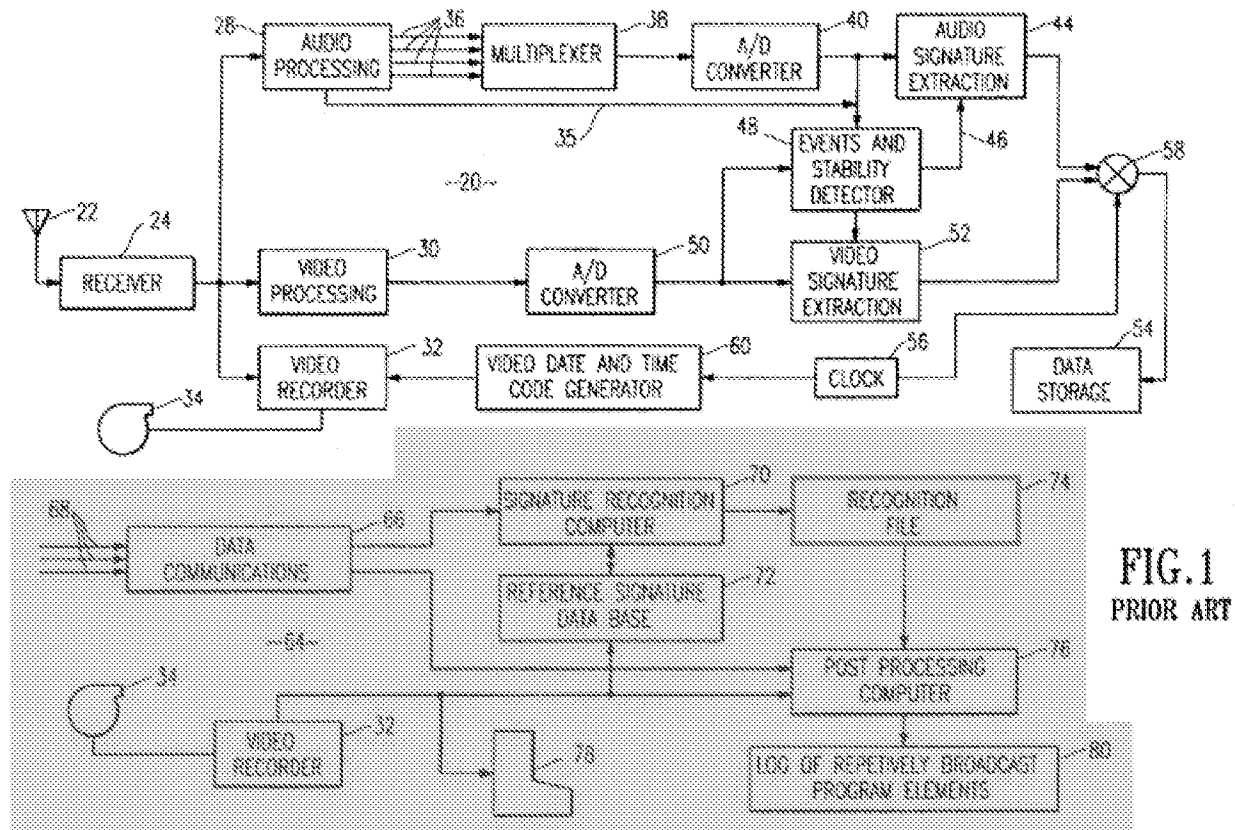
reference features with their source and time), 12:5-14, 13:43-49 (recording a source identification when a pattern recognition process determines a match); Ex. F, [0019]-[0021] (audio signatures stored in a database with cross-referenced program content identity), [0031], [0033], [0076] (returning a program-id for a database search match). Doermann, [0069].

iv. 1[c]: (c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:

O'Hagan discloses this limitation. Doermann, [0070].

O'Hagan discloses generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples (*e.g.*, generating recognition file 74 containing recognition results and classifications of the extracted signatures in sequential order such that “successive recognitions” in the recognition file 74 identify particular played programs). Doermann, [0071].

For example, O'Hagan discloses generating a recognition file 74 that stores the results of the comparisons between the extracted signatures and the reference signatures. Ex. D, 5:58-61; *see also* Ex. D, 10:28-30 (describing storage of a decision output log and identification results). *See also* the discussion of querying the database in claim 1[b] above. Doermann, [0072].



Furthermore, O'Hagan discloses that the recognition file is analyzed and processed to identify "successive" recognized extracted signatures, their start times, and durations. Ex. D, 5:58-6:3. From this disclosure, a POSA would have recognized that the recognition file is a play stream containing a sequence of identification results corresponding to audio samples from the programs that were played (*i.e.*, listened to or tuned to) and is furthermore an ordered listing of identification results, from which successive—*i.e.*, consecutive in time—recognition results are recognized as particular played programs and from which such played programs and their start times and durations are determined. Identifying audio samples and producing results relating to the order in which they were played was well-known in audio recognition systems. *See, e.g.*, Ex. E, 7:29-40 (performing identification by matching feature packets and then "matching subsequent packets"); Ex. F, Fig. 2 (showing analysis of consecutive audio frames over time), [0028] (detecting frames of an audio program in order), [0148] (identifying programs by recognitions across consecutive sets of frames) Doermann, [0073]. Furthermore, because the recognition file 74 contains the recognition and classifications from comparing the extracted

signatures to the reference signatures and, *e.g.*, has not yet been altered by scrubbing (*see* claims 1[d]-1[e] below) or some other post-processing step, it is a raw play stream. Doermann, [0073].

- v. **1[c.i]: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database;**

O'Hagan discloses this limitation. Doermann, [0074].

O'Hagan discloses that the sequence of content identification results (*e.g.*, the recognition results and classifications stored in the recognition file 74; *see* discussion of claim 1[c] above) is obtained by querying the electronic database (*e.g.*, obtained by accessing a reference signature database memory 72 to compare data); *see* discussion of claim 1[b] above) to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database (*e.g.*, to compare extracted signatures to references signatures from the database and classify the extracted signatures for identification). Doermann, [0075].

For example, as discussed with respect to claims 1[b] and 1[c] above, O'Hagan discloses that received extracted signatures are compared to respective reference signatures in the library of reference signatures stored in a reference signature database memory to generate the recognition file including the sequence of content identification results. Ex. D, 5:52-6:3. Furthermore, O'Hagan explains that its system “compares the extracted signatures from the various station monitoring units 20 with a plurality of reference signatures taken from programs to identify particular corresponding broadcast programs,” and thus the comparison is performed to determine likely matches between the extracted signatures and the reference signatures from the database. Ex. D, 5:41-47. Doermann, [0076].

- vi. 1[c.ii]: a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and**

O'Hagan discloses this limitation. Doermann, [0077].

O'Hagan discloses that a content identification result of the sequence of content identification results (*e.g.*, a particular recognition and classification result in recognition file 74) comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples (*e.g.*, a result from an extracted signature that is successfully classified); or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database (*e.g.*, a result from an extracted signature that is not classified). Doermann, [0078]-[0079].

For example, O'Hagan discloses that the recognition file 74 has stored therein the classifications of extracted signatures that were compared to the reference signatures. Ex. D, 5:58-61. Instances of "successive recognition of extracted signatures" serve as the basis for identifying particular programs. Ex. D, 5:61-64. A POSA would have therefore recognized that a recognized extracted signature would have been classified with a content identifier (*e.g.*, an identification of a program) associated in the electronic database with a respective audio data representation (*e.g.*, associated with a reference signature) that is determined to be a respective likely match with a respective one of the audio data samples (*e.g.*, determined to be a respective likely match with an extracted signature). O'Hagan thus satisfies part "(i)" of this limitation that requires "either" of two alternatives "(i)" "or" "(ii)". Doermann, [0080].

Additionally, O'Hagan discloses that analysis of the recognition file 74 having the above-described classified extracted signatures allows for "check[ing] and flag[ging] incomplete or otherwise defective broadcast identifications," for the "detection of false signal recognitions," and identification of "non-recognized or unmatched extracted signature[s]." Ex. D, 5:64-65, 6:3-10. From this disclosure, a POSA would have recognized that, for these non-recognized and

unmatched extracted signatures, the recognition file 74 must include an indication of the absence of a respective likely match in order for post-processing computer 76 to programmatically determine that any particular extracted signature is non-recognized or unmatched. That is, a “non-recognized or unmatched extracted signature” would have been classified in the recognition file in such a way that the system recognizes the classification as being non-recognized or unmatched—*i.e.*, an indication of the absence of a likely match. If, by contrast, the recognition file 74 provided no indication of the absence of a likely match, then the post-processing computer 76 would not be able to analyze the recognition file 74 to determine that an extracted signature was non-recognized and unmatched because there is no indication of that status. Such an indication would have taken any known form of indication, such as an explicit indication, examples of which were well-known in the art. *See, e.g.*, Ex. F, [0031], [0033], [0076], Figs. 2 and 7 (describing a “NO-MATCHED” indication). Even if an implicit indication were used in O’Hagan, such as an extracted signature with a blank or unfilled classification in the recognition file 74, such would have been recognized by a POSA as an indication of an absence of a match because the post-processing computer 76 is configured to understand that an extracted signature without a classification is non-recognized or unmatched based on its analysis of the recognition file 74, as disclosed by O’Hagan and discussed above. A POSA would have therefore recognized that the recognition file also includes, for non-recognized or unmatched extracted signatures, an indication of the absence of a respective likely match with a respective reference signature. O’Hagan thus further satisfies the alternative part “(ii)” of this limitation. Doermann, [0081].

- vii. 1[c.iii]: the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database;**

O’Hagan discloses this limitation. Doermann, [0082].

O’Hagan discloses that the raw play stream (*e.g.*, the recognition file) includes either: at least two different content identifiers obtained from the electronic database (*e.g.*, extracted signatures classified as programs and commercial advertisements); or at least one content identifier obtained from the electronic database and at least one indication of the absence of a

respective likely match between a respective audio data sample and an audio data representation in the electronic database (*e.g.*, extracted signatures classified as programs and commercial advertisements, along with extracted signatures classified as non-recognized and unmatched). Doermann, [0083].

For example, O'Hagan discloses that the recognition file 74, which includes classified extracted signatures, is analyzed by post-processing computer 76 to identify "programs, commercial advertisements or other repetitively broadcast programs ... together with a starting time and the program duration." Ex. D, 5:64-6:3. From this disclosure, a POSA would have therefore recognized that the recognition file includes at least two different content identifiers corresponding to two different contents, such as a first program and a second program of the plurality of programs or a program and a commercial advertisement. O'Hagan thus satisfies the first part of "at least two different content identifiers" of this limitation that requires "either" of two "or" alternatives. Doermann, [0084], [0086].

Additionally, O'Hagan discloses that not every extracted signature may have a successful match during classification, and thus the classifications will include indications of "non-recognized or unmatched extracted signature[s]." Ex. D, 6:7-10. A POSA would have therefore recognized that the recognition file includes at least one content identifier corresponding to a program or a commercial advertisement, and one indication of the absence of a likely match, for example in the case where at least one extracted signature is classified as a program or commercial advertisement and at least one extracted signature is classified as non-recognized or unmatched. O'Hagan thus satisfies the second part of "at least one content identified ... and at least one indication of the absence of a respective likely match" of this limitation that requires "either" of two "or" alternatives. Doermann, [0085], [0086].

Moreover, a POSA would have recognized that this claim language describes obvious natural variances in performing audio recognition of audio channels, such as radio stations and CDs, which include series of different audio programs from which any particular user may switch between or back and forth (in addition to switching between radio stations and CDs altogether). *See, e.g.*, Ex. D, 14:65-66 (describing a tape recording including ten different musical works), 16:20-29 (describing an audio signal including a plurality of musical works). Because the audio recognition system is expected to successfully recognize extracted signatures at least some of the time (*see* Ex. D, 5:61-64 (describing the successive recognition of extracted

signatures), 6:8-9 (describing non-recognized and unmatched extracted signatures); Ex. F, [0019], [0031], [0033], [0076], Figs. 2 and 7 (describing database searches resulting in matches and no-matches), the plurality of classifications of the programs will include two different content identifiers if more than one is successfully identified or will include one content identifier and one indication of the absence of a likely match if one is successfully identified and one is not. Doermann, [0086].

viii. 1[d]: (d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and

O'Hagan in view of Apel renders obvious this limitation. Doermann, [0087].

O'Hagan discloses scrubbing the raw play stream to determine whether to change a result of the sequence of content of identification results (*e.g.*, to determine whether to change a non-recognized or unmatched extracted signature based on a manual identification). Doermann, [0088].

For example, O'Hagan discloses that the recognition file includes the classification results of the extracted signatures, which may include classifications of the extracted signatures to "programs, commercial advertisements or other repetitively broadcast programs" or as otherwise being "non-recognized or unmatched." Ex. D, 5:58-6:10. A temporary code identifies the non-recognized or unmatched extracted signatures, which O'Hagan discloses are manually reviewed and identified. Ex. D, 6:7-10. Doermann, [0089].

Apel teaches the advantageous use of time tagged audio features to efficiently identify packets based on an initial successful identification. Apel teaches that reference features are stored with their source and time. Ex. E, 6:65-66, 10:18-32. Before an initial identification is made, Apel teaches that its recognition process "locate[s] time intervals in the stored feature blocks that are time aligned ... with the unknown feature packet" that is to be identified. Ex. E, 6:65-7:7:1. The unknown feature packet is compared to stored feature blocks that are aligned in time with the unknown feature packet. Ex. E, 6:65-7:1, 7:27-29. Those stored feature blocks being used for initial comparison include all relevant feature blocks from a desired market, which in one example may be 10 TV stations and 50 radio stations for a small region. Ex. E,

6:54-64. Based on a comparison of the unknown feature packet to the multitude of stored feature blocks from the same time interval, an initial identification is made, and afterward, “[o]nce an initial identification has been made, it is confirmed by matching subsequent packets with broadcast source features from the same channel as the previous recognition.” Ex. E, 7:28-32. Therefore, the additional comparisons are limited to stored feature blocks from the same channel—*i.e.*, channel data—but subsequent in time to the initial identification and need not include a broad search among the entire market’s many channels. In comparing and matching the additional, subsequent in time packets with packets from the single channel because a pattern of the feature packets is expected to match an expected pattern of feature packets exhibited by the channel, Apel is able to advantageously make use of packet times and broadcast times expected to occur in sequence. Ex. E, 6:65-7:1, 7:27-29. Apel explains that this method of identification is used because “a listener will continue to listen (or stay tuned) to the same station for some amount of time.” Ex. E, 7:34-37. That is, Apel teaches that packets are expected to follow a pattern of staying on the same channel, and so once an initial identification of a feature packet is made on one channel, subsequent feature packets are expected to also be on the same channel and in the same order because the user is expected to remain listening to the same channel. Doermann, [0090].

To the extent Patent Owner argues that O’Hagan does not explicitly describe a computer process of scrubbing the raw play stream in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data, a POSA would have been motivated to modify O’Hagan’s system to perform that process and scrub the raw play stream in view of a pattern of the initial successful identifications or classifications compared to a pattern of an expected pattern of sample sequence data, to take advantage of known behavioral tendencies with respect to listening patterns as discussed in Section V.A.4 above, in light of Apel’s teachings. *See also* MPEP § 2144.04(III) (providing an automatic means to replace a manual activity is obvious). Accordingly, the recognition process disclosed by O’Hagan, and as modified by Apel’s teachings, would have included:

- generating a recognition file that is a raw play stream of a sequence of content identification results (as discussed with respect to claims 1[c]-1[c.iii] above),

- which may at times include an indication of non-recognized or unmatched extracted signatures that are indications of an absence of a match for an audio data sample (as discussed with respect to claims 1[c.ii]-1[c.iii] above),
- automatically recognizing that non-recognized or unmatched extracted signatures require further analysis for identification (as discussed above; *see* Ex. D, 6:3-10), and
- utilizing existing classified extracted signatures near the non-recognized or unmatched extracted signatures as bases for determining that the non-recognized or unmatched extracted signatures in the sequence of content identification results should be classified—and thus changed from being unrecognized or unmatched to recognized or matched—in comparison to an expected pattern of signatures from a particular channel as if continuously played (as discussed above; *see* Ex. E, 7:28-37), and
- attempting to change and match the non-recognized or unmatched extracted signatures to reference signatures from the same channel as the existing classified extracted signatures (as discussed above; *see* Ex. E, 7:28-37).

In other words, the recognition and post-processing unit 64 disclosed by O'Hagan, as modified by Apel's teachings, would have recognized that one or more extracted signatures were non-recognized or unmatched—thus being part of a pattern of sample sequence data that is incomplete compared to an expected pattern of sample sequence data having complete recognitions from a channel because a user is expected to stay tune to the same channel for a period of time—analyzed existing classified extracted signatures in the vicinity of the non-recognized or unmatched extracted signatures to determine that existing classified extracted signatures are from a particular channel, and compared and matched the non-recognized or unmatched extracted signatures in the incomplete pattern of extracted signature classification results stored in the recognition file with reference signatures corresponding to the signatures in the expected pattern of signatures based on the channel of the recognized extracted signatures. Additionally, because classifications of extracted signatures are stored in recognition file 74, the recognition and identification of non-recognized or unmatched extracted signatures—*i.e.*, the result of classifying non-recognized or unmatched extracted signatures, which is disclosed as a manual process by O'Hagan but performed automatically by computer as taught by Apel—is an updated classification that would have also been stored in the same recognition file 74 that stores the classification of all extracted signatures. Ex. D, 5:58-60. Doermann, [0091].

ix. 1[e]: (e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.

O'Hagan in view of Apel renders obvious this limitation. Doermann, [0092].

O'Hagan in view of Apel teaches scrubbing the raw play stream to determine whether to change a result of the sequence of content of identification results. *See* discussion of claim 1[d] above). Furthermore, O'Hagan discloses that the purpose of the scrubbing is for "identification" of the extracted signatures that were classified as being "non-recognized or unmatched extracted signature[s]." Ex. D, 6:7-10. Classified extracted signatures are stored in the recognition file 74 (*i.e.*, the raw play stream), and so from this disclosure a POSA would have recognized that the scrubbing, as taught by O'Hagan in view of Apel, identifies and fixes the classifications of the "non-recognized or unmatched extracted signatures" that are in the recognition file 74 (*i.e.*, the raw play stream)—the purpose of Apel's teaching of subsequent identifications of audio signatures based on initial identifications is to successfully identify those audio signatures for the first time, and O'Hagan explains that those classification results are stored in the recognition file 74. Doermann, [0093]. Thus, upon classifying previously non-recognized or unmatched extracted signatures based on the determinations made by the scrubbing, the recognition file 74 is updated to include classified extracted signatures of previously non-recognized or unmatched extracted signatures now recognized and matched—*i.e.*, a "clean play stream." Doermann, [0093].

b. Claim 2: The method of claim 1 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.

O'Hagan in view of Apel renders obvious the method of claim 1. O'Hagan in view of Apel further renders obvious the limitations of this claim. Doermann, [0094].

O'Hagan in view of Apel teaches scrubbing the raw play stream by analyzing sample sequence data, which Apel explains is performed based on corresponding packet times and broadcast times expected to occur in sequence from the same channel. *See* discussion of claim 1[d] above. As discussed, Apel teaches "matching subsequent packets with broadcast source features from the same channel as the previous recognition." Ex. E, 7:30-32. From this disclosure, a POSA would have recognized that an analysis of sample sequence data for

matching O'Hagan's non-recognized or unmatched extracted signatures therefore includes analyses of log times—*i.e.*, finding packets that are “subsequent” to an existing matched packet means finding packets that are next or later in time—and content identifiers—*i.e.*, the subsequent broadcast source features must be “from the same channel” and so the identity of the channel must be determined. Apel further explains that relevant times for sequence analysis and matching include “the time of acquisition”—*i.e.*, a log time. Ex. A, 9:30-31, 11:13-14 (a log time indicates a time of capture); Ex. E, 3:19-21. Doermann, [0095].

c. Claim 3: The method of claim 1 wherein: scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content of identification results.

O'Hagan in view of Apel renders obvious the method of claim 1. O'Hagan in view of Apel further renders obvious the limitations of this claim. Doermann, [0096].

O'Hagan in view of Apel teaches scrubbing the raw play stream by determining whether to change a result of the sequence of content of identification results. *See* discussion of claim 1[d] above. Apel further explains that the determination includes analyzing channel data: “matching subsequent packets with broadcast source features from the same channel as the previous recognition” “based on the assumption that a listener will continue to listen (or stay tuned) to the same station for some amount of time.” Ex. E, 7:30-37. That is, Apel teaches that a listener listening to a channel will, for matching purposes, be assumed to remain on the same channel, and so identifications subsequent to any successful identification are assumed to come from the same channel and matched with subsequent channel data until a subsequent match fails. Ex. E, 7:37-53, 7:10:18-32. From this disclosure, a POSA would have recognized that the scrubbing process of O'Hagan in view of Apel includes, for example, analyzing an existing classified extracted signature's channel to determine the channel of the existing classified extracted signature so that the identity of appropriate channel data may be determined for use in retrieving information to identify subsequent signatures, retrieving an appropriate channel's channel data such as reference signature and timing information, and analyzing the appropriate channel's channel data to determine if an extracted signature matches, and should be change to correspond with, the channel's reference signatures in comparison. *See* also Ex. E, 7:42-45 (comparisons are made with downloaded channel data), 7:30-34 (when matching subsequent packets, only broadcast source features—*i.e.*, channel data—from the same channel are

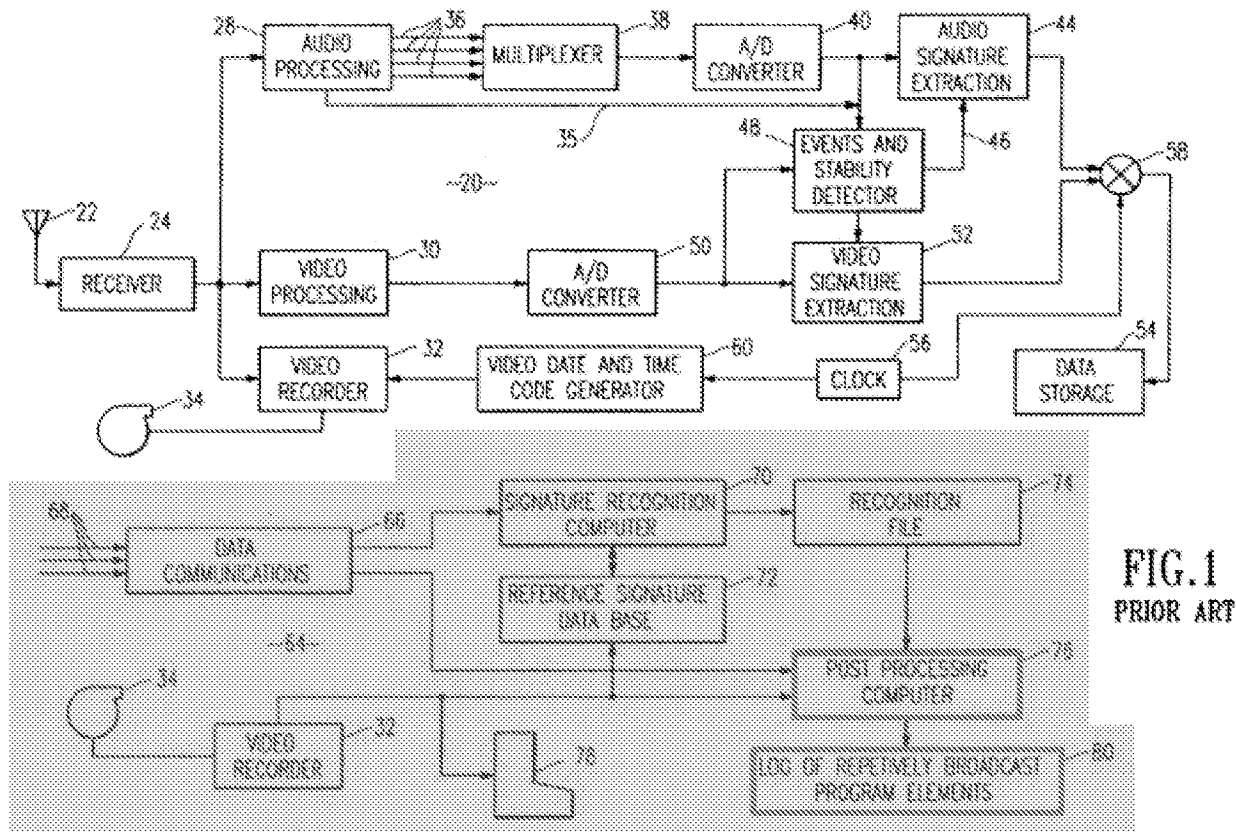
downloaded), 7:37-39, 8:13-29 (downloaded channel data are used for the pattern recognition and matching process). Doermann, [0097].

d. Claim 5: The method of claim 1 wherein a clean play list is generated from the clean play stream.

O'Hagan in view of Apel renders obvious the method of claim 1. O'Hagan in view of Apel further renders obvious the limitations of this claim. Doermann, [0098].

O'Hagan in view of Apel teaches generating the clean play stream after scrubbing as an updated recognition file 74, which included the previously non-recognized or unmatched extracted signatures—now recognized and matched as a result of the scrubbing. *See* discussion of claim 1[e] above. O'Hagan further discloses generating a clean play list (*e.g.*, generating a file or log containing records of programs with their starting times and durations) from the clean play stream (*e.g.*, from the recognition file of classified extracted signatures generated after scrubbing). Doermann, [0990].

For example, O'Hagan discloses that after extracted signatures are classified and stored in recognition file 74, “successive recognized extracted signatures are provided to post-processing computer 76, which produces a file or log 80 containing records of the identified programs, commercial advertisements or other repetitively broadcast programs from each monitored broadcast station together with a starting time and the program duration.” Ex. D, 5:58-6:3. Doermann, [0100].



Thus O'Hagan in view of Appel renders obvious generating a clean play list of identified programs, with their start times and durations, from the play stream of classified extracted signatures, which is a clean play stream generated after scrubbing. Doermann, [0101].

e. Claim 6: The media measurement method of claim 1 wherein the plurality of audio data representations include a plurality of audio signatures.

O'Hagan in view of Apel renders obvious the method of claim 1. O'Hagan further discloses the limitations of this claim. Doermann, [0102].

O'Hagan discloses that the plurality of audio data representations (*e.g.*, reference signatures stored in reference signature database memory 72) include a plurality of audio signatures (*e.g.*, "reference signatures" compared with "audio signature[s]" to identify corresponding broadcast programs). Doermann, [0103].

For example, as discussed with respect to claims 1[b] and 1[c.i] above, O'Hagan discloses that the process "compares the extracted signatures from the various station monitoring units 20 with a plurality of reference signatures taken from programs to identify particular corresponding broadcast programs." Ex. D, 5:41-47, 5:52-54. The extracted signatures include

“audio signatures,” and thus a POSA would have recognized that the reference signatures used for comparison likewise include audio signatures. Doermann, [0104].

B. Claims 1-3, 5-6, 9-11 and 13-14 Are Obvious Over O’Hagan in view of Apel and Cheung

1. Overview of Cheung

Cheung (Ex. F) qualifies as prior art under at least pre-AIA 35 U.S.C. § 102(e) because it is a publication of an application for patent having priority to U.S. provisional application no. 60/547,931, filed 2/26/2004, which is before the ’768 patent’s earliest possible priority date of 5/27/2004. *See* priority analysis in Section V.B.1.a below. Cheung is directed to a system and method for detecting and identifying broadcast audio and video signals. Ex. F, [0001]. The system and method, implemented via computer software and/or hardware, operates by first registering programming content and storing a series of “fingerprints” or “signatures” (also referred to as pattern vector data) of the content in a database, cross-referenced with the identities of the content. Ex. F, [0019]-[0021], [0219]. Doermann, [0106]. Next, unidentified monitored content is input, and the system calculates the signatures of the monitored content and attempts to match the calculated signatures with signatures of the already registered content in the database. Ex. F, [0019], [0024]-[0031]. Database searches return identifying information if successful (*e.g.*, a “program-id” and “frame-id”) or an indication of failure if not (*e.g.*, a “NO-MATCHED flag”). Ex. F, [0019], [0031], [0033], [0076], Figs. 2 and 7. If a number of matches are found within a certain time frame, then a program is deemed detected, and relevant information, such as a program title and time of detection, is logged. Ex. F, [0019], [0021], [0032]. Results are also displayed on a screen, for example. Ex. F, [0021]. Doermann, [0107].

a. *Dynamic Drinkware Analysis: Cheung is Prior Art*

On its face, Cheung is a U.S. patent application publication published 7/19/2007, based on non-provisional application 10/598,283 filed 8/23/2006, claiming priority to PCT application PCT/US05/04802 filed 2/16/2005, and claiming priority to provisional application 60/547,931 filed 2/26/2004 (Ex. G; “Cheung-P”). The 2/26/2004 filing date of Cheung-P predates the earliest possible priority date of the ’768 patent, which claims priority to provisional application

60/574,836 filed 5/27/2004.⁶ As shown below, Cheung is supported by Cheung-P and therefore predates the earliest possible priority date of the '768 patent. Doermann, [0108]-[0109].

i. Cheung-P Supports a Claim of Cheung

As shown in the table below, Cheung-P supports at least one claim of Cheung.

Cheung Claim 1	Exemplary Support in Cheung-P
1. A method executed by a digital signal processing system of	Ex. G at 9-11 6:20-8:2 (“The present invention describes a system and method whereby identification of known audio or video programming can be done without any reliance on a tandem video signal (in the audio case) or normative markers in the signal indicating a known time in the program and with unique and novel ways to calculate codes representing the characteristics of the audio program without requiring impractical computational capabilities. Benefits of this system and method are accuracy, speed, robustness to playback speed variation and the ability to perform the identification process in real time, without reliance on any embedded cue or watermark. In addition, the present invention takes advantage of the availability of low cost, high performance computing platforms in order to implement a high speed database searching methodology. Detailed Description. A. Overview The broadcast monitoring and detection system embodying the invention works in two phases: registration and detection. During the registration phase, known programming content is registered with the system by sending the program, as digital data, into the system. A series of signatures, in the case here, a pattern vector and more generally in the art a “fingerprint” or “signature”, are stored as a sequence of data records in a database, with the identity of the program content cross-referenced to them as a group. During the second phase, unidentified programming is input into the system. Such programming can include radio, television, internet broadcasts or any other source of audio or video programming, whether terrestrial broadcast, satellite, internet, cable television or any other medium of delivery, whether now known or devised in the future. While such programming is being monitored, the pattern vectors of the programming (or any other signature generating technique) are continually calculated. The calculated pattern vectors are then used to search for a match in the database. When a match is found and

⁶ Requestor takes no position as to the propriety of the priority claims since the art presented herein pre-dates the earliest filing.

	confirmed, the system uses the cross-referenced identity in the database to provide the identity of the content that is currently being played. In the preferred embodiment, the system is software running on a computer, however, it is envisioned that special purpose hardware components may replace parts or all of each module in order to increase performance and capacity of the system.”)
generating a signature associated with a [known] signal, the signature comprised of a set of numeric values of at least one element and corresponding to at least one time frame of the signal,	Ex. G at 12 9:21-23 (“During the registration phase, the results of the pattern generation module, which creates signatures for known audio or video content, are stored in the database...”) Ex. G at 13 10:18-25 (“The invention organizes the samples into a series of time frames, which consist of a predetermined number of samples. The time frames are stored in sequence. Alternatively, data structures, stored in the computer memory (which includes the hard drive if the operating system supports paging and swapping), may be used where the time frames are not physically stored in sequence, but logically may be referenced or indexed in the sequence that they were detected by means of memory addressing.”) Ex. G at 14-15 11:20-12:5 (“The PG module operating during the detection phase, (2), fetches the stored digital audio or video samples that were detected and stored by the SA Module. Once a frame of the samples is received, the PG module will compute the pattern vector of the frame and, when in detection phase, send the pattern vector to the Database Search Module in the form of a database query. During the registration phase, the PG module calculates the pattern vector in order that it be stored in the database, in correlation with the other relevant information about the known audio or video program. The calculation of the pattern vector is described further below.”) Ex. G at 60 Claim 1 (“For the pattern vector method: 1. A method of generating a signature comprised of a set of at least one member each member being a real number, for at least one time frame of an audiosignal comprised of a set of predetermined frequency bands, such audio signal being identified by an identification index and such time frame being identified by a time frame index, comprising:”)
such [known] signal being identified by an	Ex. G at 13

identification index and such time frame being identified by a time frame index, comprising:	10:18-25 (“The invention organizes the samples into a series of time frames, which consist of a predetermined number of samples. The time frames are stored in sequence. Alternatively, data structures, stored in the computer memory (which includes the hard drive if the operating system supports paging and swapping), may be used where the time frames are not physically stored in sequence, but logically may be referenced or indexed in the sequence that they were detected by means of memory addressing.”) Ex. G at 60 Claim 1 (“For the pattern vector method: 1. A method of generating a signature comprised of a set of at least one member each member being a real number, for at least one time frame of an audiosignal comprised of a set of predetermined frequency bands, such audio signal being identified by an identification index and such time frame being identified by a time frame index, comprising:”)
Calculating for at least one time frame of the signal a pre-determined number of spectral magnitude values grouped in at least one frequency band of pre-determined width;	Ex. G at 19 16:1-22 (“The PG module reads in a frame of signal, preferably consisting of 16,384 samples, with sampling rate preferably set at 8,000 samples per second. Thus, the frame length is approximately two seconds in time. More or less samples or frame widths in time may be used with varying results. Given $x = [x_1 \ x_2 \ \dots \ x_{16384}]$, the vector containing a frame of signal, where each x_i is the value of the ith audio sample, an N element pattern vector is calculated with the following steps. In the preferred embodiment, N is equal to 31. Practitioners of ordinary skill will recognize that the value of N is arbitrary, and can be increased or decreased with varying results. For example, decreasing N reduces the compute time and memory requirements, but may reduce accuracy. Increasing N may do the opposite. Also, the method presented will assume that a 31 element pattern vector is being used in the calculation in order to simplify the presentation of the invention. Practitioners of ordinary skill will recognize that the same methodology will work when N is increased or decreased, depending on whether the goal is increased accuracy or reduced computer complexity. 1. The Fourier transform of x is calculated with the number of points equal to the number of samples in the frame, in order to get the spectrum vector $X = \{X_1 \ X_2 \ \dots \ X_{16384}\}$. The spectral resolution of the transform is = $\frac{8000 \text{ samples/sec}}{16,384 \text{ samples}} = 0.488 \text{ Hz}$ Segregate the FFT spectral values into frequency bands of a specified width, where in the preferred embodiment, the width is 64 Hz.”)

	Ex. G at 60 Claim 1 (“Transforming the at least one time frame of signal into the frequency domain, such that there is a finite number of frequency magnitude values stored in a computer memory;”)
Calculating for each frequency band a numeric value that is equal to a pre-determined function of one or more of the calculated spectral magnitude values grouped within the frequency band;	Ex. G at 20-21 17:8-8:8 (“2. Compute the centroid (or center.-of-gravity COG) of each band: $p_k = \frac{\sum_{m=1}^{131} m \times X_{131k+m}}{\sum_{m=1}^{131} X_{131k+m}}$ In the preferred embodiment, only Band 2 to 32 is used because Band 1 is the lowest band including zero Hz, which is normally not useful in FM radio transmission; and Band 32 covers the band up to 1,800Hz, which is typically sufficient bandwidth to encode a fingerprint of the audio. Of course, higher bands or lower bands can be used if required. The inclusion of higher or lower bands to account for signal characteristics can be determined empirically. The first step, where the FFT coefficients are collected in order to calculate the centroid in step 2 is different in the case of video. In the video case, the FFT coefficients have to be selected from locations either in the complex plane or on the 2-dimensional spatial frequency plane as described on page 23 of Poynton, incorporated herein by reference. These locations are analogous to the frequency bands on the audio case. In a manner analogous to using predetermine frequency bands in audio, predetermined regions on the vertical/horizontal plane in the frequency domain can be defined and the FFT coefficient values in each regions used to calculate an element corresponding to that region. Once this selection is made, the centroid can be calculated in an equivalent manner. It is advantageous to ignore the frequency region encompassing the frame rate, sync rate, subcarrier, or line rate. The end result is essentially equivalent to the case of audio: that each time frame of video will have a pattern vector associated with it that is stored in a database.”) Ex. G at 60 Claim 1 (“Calculating a single value for each frequency band that is a linear combination of the frequency magnitude values that reside within the frequency band produced by said transformation;”)
Storing the calculated numeric values in a computer database with a reference to	Ex. G at 4 1:10-16 (“The system initially registers a known program by digitally sampling the program and separating the digital sample stream into a large set of short segments in time. These segments are then

their corresponding time frame index and their corresponding identification index.	processed to extract particular feature sets that are characteristic of the segment. The invention processes each set of features to produce a numerical code that represents the feature set for a particular segment of the known program. These codes and the registration data identifying the program populate a database as part of the system.”) Ex. G at 10 7:9-14 (“During the registration phase, known programming content is registered with the system by sending the program, as digital data, into the system. A series of signatures, .in the case here, a pattern vector and more generally in the art a “fingerprint” or “signature”, are stored as a sequence of data records in a database, with the identity of the program content cross-referenced to them as a group.”) Ex. G at 11 8:6-9 (“During the registration phase, the CPU fetches the audio or video data from the sound card, calculates the pattern vector data, and then, along with timing data and the identity of the program, these results are stored in a database, as further described below.”) Ex. G at 12 9:21-23 (“During the registration phase, the results of the pattern generation module, which creates signatures for known audio or video content, are stored in the database”) Ex. G at 15 12:2-4 (“During the registration phase, the PG module calculates the pattern vector in order that it be stored in the database, in correlation with the other relevant information about the known audio or video program.”) Ex. G at 22 19:17-18 (“In the preferred embodiment, each element is then converted to the unsigned 16-bit integer format for convenient storage and further processing.”) Ex. G at 60 Claim 1 (“Storing said single value in a computer memory with a reference to the time frame index and the identification index.”)
--	--

ii. Cheung-P Supports Every Relied On Disclosure of Cheung

Every paragraph, claim, and figure of Cheung relied upon in this Request is directly supported in Cheung-P, most if not all supported verbatim, at the locations indicated below.

Citation to Cheung	Exemplary Support in Cheung-P	
	Page of Ex. G	Location on Page

[0001]	4	1:4-25
[0018]	9-10	6:20-7:4
[0019]	10-11	7:8-8:2
[0020]	11	8:4-24
[0021]	12	9:1-17
[0024]	13	10:3-25
[0025]	14	11:1-8
[0026]	14	11:9-17
[0027]	14-15	11:20-12:8
[0028]	15	12:7-21
[0029]	15-16	12:22-13:9
[0030]	16	13:11-18
[0031]	16-17	13:20-14:5
[0032]	17-18	14:7-15:7
[0033]	18	15:9-13
[0076]	34	31:8-11
[0219]	58	55:11-25
Claim 35	67	Claim 25
Fig. 2	69	Fig. 2
Fig. 7	73	Fig. 7

2. O'Hagan, Apel, and Cheung are Analogous Art

O'Hagan and Apel are analogous to the '768 patent, as discussed in Section V.A.3 above. Cheung is also analogous to the '768 patent because it is in the same field of endeavor as the '768 patent and reasonably pertinent to the alleged problems discussed therein. Doermann, [0110].

For example, the '768 patent is in the field of “media monitoring and measurement” and the “recognition” of media content (*see* Ex. A, Abstract, 1:60-2:15), and, similarly, Cheung is in the field of the “automatic detection and identification of broadcast programming” (*see* Ex. F, Abstract, [0001]. Doermann, [0110].

Moreover, Cheung seeks to provide a system that works with “any readily available source of content” and operates “without any reliance on ... normative markers in the signal” and “without reliance on any embedded cue or watermark.” Ex. F, [0001], [0018]. Cheung's disclosure would therefore have been reasonably pertinent to problems alleged in the '768 patent, which include the desire to track usage “across several types of media” such as radio, CDs, and others, and to avoid relying on the use of “codes” that “are limited in that they can only track content that has been properly encoded.” *See* Ex. A, 1:14-36. Doermann, [0110].

Accordingly, O'Hagan, Apel, and Cheung are each analogous to the '768 patent.

3. A POSA Would Have Been Motivated to Modify O’Hagan in view of Apel’s and Cheung’s Teachings

O’Hagan in view of Apel renders obvious claims 1-3 and 5-6, as discussed in Section V.A above. For the same reasons, O’Hagan in view of Apel renders obvious the method described in claims 9-11 and 13-14, which is substantially similar to the method described in claims 1-3 and 5-6 but performed by executing instructions embodied on a computer program product. To the extent Patent Owner may argue that O’Hagan and Apel do not explicitly describe specific implementation examples or details in claims 1-2 and 9-10, Cheung teaches that such implementations would have been well-known and obvious to a POSA. Doermann, [0111].

Regarding claim elements 1[b]-1[c.iii] and 2, O’Hagan discloses querying an electronic database and comparing extracted signatures to reference signatures to classify extracted signatures. To the extent argued that O’Hagan does not explicitly describe classifying extracted signatures to corresponding “content identifiers,” Cheung teaches that comparisons of signatures in a database return, if successful, a “cross-referenced identity in the database to provide the identity of the content that is currently being played.” Ex. F, [0019], [0031], [0033]. [0076]. That is, Cheung teaches implementing a searchable database with signatures (*i.e.*, content to be matched) and identity information corresponding to the signatures (*i.e.*, content identifiers that describe the content to be matched). A POSA would have been motivated to implement O’Hagan’s database-based classifications with content identifiers—to the extent argued that the classification of content can be performed without any content identifiers—to advantageously return descriptive identifying information about matches. Indeed, one purpose of O’Hagan’s system is to “identify particular corresponding broadcast programs.” Ex. D, 5:46-47; *see also* Ex. D, 1:9-12, 1:25-29, 3:30-35. Doermann, [0112]. Moreover, the modification would have been, at least, the application of a known technique (*e.g.*, Cheung’s teaching of cross-referenced identities in a database) to a known system ready for improvement (*e.g.*, O’Hagan’s database used for classifying extracted signatures to identify the signatures) to yield predictable results (*e.g.*, the production of information that identifies the extracted signatures upon finding a match). Doermann, [0113]. *See* MPEP § 2143 (citing *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007)).

Regarding claim elements 1[c.ii]-1[c.iii], O’Hagan discloses classifying extracted signatures so that they either identify particular programs or be “non-recognized or unmatched.”

Ex. D, 5:58-6:10. To the extent it is argued that O'Hagan does not explicitly describe the manner of classifying non-recognized or unmatched extracted signatures such as with "an indication of the absence of a respective likely match," Cheung teaches that comparisons of signatures in a database return, if not successful, a "NO MATCHED" flag. Ex. F, [0031], [0033], [0076], Figs. 2 and 7. That is, Cheung teaches that when a matching signature is not found in a database, that absence of a match is indicated by a NO MATCHED indication and not indicated, hypothetically, by the absence of any indication. A POSA would have been motivated to implement O'Hagan's database-based classifications with an indication of the absence of a match, such as with a "NO MATCHED" indication, to positively identify extracted signatures that are not matched to reference signatures in the database. In this way, the system is assured that the extracted signature has indeed been analyzed and that a determination, whether matched or unmatched, has actually been made. Doermann, [0114]. Moreover, the modification would have been, at least, the application of a known technique (*e.g.*, Cheung's teaching of identifying unmatched signatures with a "NO MATCHED" indication) to a known system ready for improvement (*e.g.*, O'Hagan's system of classifying extracted signatures that leaves to a POSA how to describe the classifications) to yield predictable results (*e.g.*, indicating the absence of a match with a NO MATCHED result). Doermann, [0115]. *See* MPEP § 2143 (citing *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007)).

Regarding independent claim 9 and claims depending therefrom, O'Hagan discloses implementing its system and method with computers. *See, e.g.*, Ex. D, 4:65-6:12 (describing various processing "circuitry," "units," and "computers," an A/D converter, processing digital signals, using a database memory, etc.). To the extent argued that O'Hagan does not explicitly describe its systems (*e.g.*, station monitoring unit 20 and signature and post-processing unit 64) implemented as one or more computers executing instructions stored on a non-transitory computer readable medium, Cheung teaches implementation via a computer with a processing unit running "software ... implemented in ROM (read-only memory) form or stored on any kind of computer readable media, including CD-ROM, ... or transmitted as digital data files stored in a computer's memory." Ex. F, [0019]-[0021], [0219], claim 35. A POSA would have been motivated to implement O'Hagan's station monitor units and post-processing units as computers with processing units executing software stored on computer readable media to take advantage of commodity hardware and readily transportable software. Indeed, O'Hagan contemplates the use

of “hardware as simple as a Macintosh computer” to avoid “resorting to supercomputers or expensive data storage facilities.” Ex. D, 4:5-13, 9:35-43. Cheung’s teachings of general-purpose computers executing instructions on computer readable media facilitates O’Hagan’s desired implementation. Moreover, the implementation of O’Hagan’s system with computers would have been obvious because known work in a field of endeavor (*e.g.*, Wang’s use of processors executing software stored on computer readable media) may prompt variations of it for use in the same field (*e.g.*, using the same for O’Hagan’s recognition system) based on design incentives or other market forces if the variations are predictable to one of ordinary skill in the art (*e.g.*, making use of commercially available hardware and distributable software taught by O’Hagan and Wang). Doermann, [0116]. *See* MPEP § 2143 (citing *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007)).

A POSA also would have had a reasonable expectation of success modifying O’Hagan’s system to utilize particular content identifiers, “NO MATCHED” result indicators, and executable instructions stored on computer readable media, as discussed above. The modifications amount merely to particular implementation details for O’Hagan’s existing system and method, and Cheung teaches specific implementation examples, *e.g.*, classifying content with “program-ids” or as “NO MATCHED,” and storing and reading software from a CD-ROM, that were obvious and well-known to POSAs. Ex. F, [0033], [0076], [0219]. Doermann, [0117].

Claims 1-3, 5-6, 9-11, and 13-14 are therefore obvious over O’Hagan in view of Apel and Cheung as discussed herein. Doermann, [0118]. The examiner did not appear to have these teachings before him when allowing the claims, which seemingly were allowed for teaching a raw playstream including unmatched samples and a scrubbing correction mechanism for changing those unmatched samples.

4. Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim

a. Claim 1

i. 1[pre]: A media measurement method comprising:

To the extent that the preamble of claim 1 is limiting, O’Hagan discloses it. *See* the discussion of claim 1[pre] in Section V.A above. Doermann, [0119].

- ii. 1[a]: (a) receiving over a network, at one or more computers, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network;**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[a] in Section V.A above. Doermann, [0120].

- iii. 1[b]: (b) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers;**

O'Hagan discloses this limitation. For example, O'Hagan discloses using the one or more computers (*e.g.*, using *signature and recognition post-processing unit 64*, which includes signature recognition computer 70) to query an electronic database (*e.g.*, accessing a reference signature database memory 72 to compare data) of a plurality of audio data representations (*e.g.*, a library of reference signatures used for comparison) and corresponding content identifiers (*e.g.*, information used for classifying recognized signatures). *See* the discussion of claim 1[b] in Section V.A above. Doermann, [0121].

Cheung further teaches querying an electronic database (*e.g.*, searching a database) of a plurality of audio data representations (*e.g.*, signatures stored in the database) and corresponding content identifiers (*e.g.*, program identities cross-referenced to the signatures). Doermann, [0122].

For example, Cheung discloses “search[ing] for a match” for calculated signatures in a “database,” and “[w]hen a match is found and confirmed, the system uses the cross-referenced identity in the database to provide the identity of the content that is currently being played.” Ex. F, [0019], [0021]. For example, a “program-id” and “frame-id” are returned for successful database searches. Ex. F, [0031], [0033], [0076]. The signatures and cross-referenced identities are added and stored in the database during a registration phase prior to querying the database. Ex. F, [0020]-[0021]. Doermann, [0123].

To the extent Patent Owner argues that O'Hagan does not explicitly describe reference signature database memory 72 including content identifiers, a POSA would have been motivated to modify O'Hagan's system to include content identifiers in reference signature database memory 72, such as cross-referenced program identities like program-ids, to return useful

identifying information in response to querying reference signature database memory 72 as discussed in Section V.B.3 above, in light of Cheung's teachings. A purpose of querying reference signature database memory 72 is to search for potential matches to an extracted signature, and so the inclusion of content identifiers in reference signature database memory 72 to return classifying information would have been obvious to a POSA. Doermann, [0124].

iv. 1[c]: (c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:

O'Hagan discloses this limitation. For example, O'Hagan discloses generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples (*e.g.*, generating recognition file 74 containing recognition results and classifications of the extracted signatures in sequential order such that "successive recognitions" in the recognition file 74 identify particular programs). *See* the discussion of claim 1[c] in Section V.A above. Doermann, [0125].

O'Hagan in view of Cheung further renders obvious the recognition file 74 including content identification results because Cheung teaches returning content identifiers in response to a database query, as discussed with respect to claim 1[b] above. The recognition file 74 includes the classifications of extracted signatures determined by comparing the extracted signatures to reference signatures stored in reference database memory 72. *See* the discussion of claim 1[c] in Section V.A above. Therefore, if a database query returns content identification results as taught by Cheung, then the recognition file 74 containing the classifications of extracted signatures based on database queries will include content identification results. Doermann, [0126].

v. 1[c.i]: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database;

O'Hagan discloses this limitation. For example, O'Hagan discloses that the sequence of content identification results (*e.g.*, the recognition results and classifications stored in the recognition file 74; *see* discussion of claim 1[c] in Section V.A above) is obtained by querying the electronic database (*e.g.*, obtained by accessing a reference signature database memory 72 to

compare data); *see* discussion of claim 1[b] in Section V.A above) to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database (*e.g.*, to compare extracted signatures to references signatures from the database and classify the extracted signatures for identification). *See* the discussion of claim 1[c.i] in Section V.A above. Doermann, [0127].

O'Hagan in view of Cheung further renders obvious that the results of querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database are obtained content identification results because Cheung teaches returning content identifiers in response to a database query, as discussed with respect to claim 1[b] above. Like O'Hagan, Cheung's querying of an electronic database is an attempt to determine respective likely matches between respective audio data samples in a sequence of audio data samples and respective audio data representations in an electronic database. *See* claim 1[b] above. Additionally, Cheung describes:

2. Pattern Vector Generation [(PG)] Module

The PG module operating during the detection phase, (2), fetches the stored digital audio or video samples that were detected and stored by the SA Module. ***Once a frame of the samples is received, the PG module will compute the pattern vector of the frame and, when in detection phase, send the pattern vector to the Database Search Module in the form of a database query. During the registration phase, the PG module calculates the pattern vector in order that it be stored in the database, in correlation with the other relevant information about the known audio or video program.*** The calculation of the pattern vector is described further below.

[...]

3. Database Search (DBS) Module

Upon the reception of a query generated by the [Pattern Vector Generation (PG)] module, this module, [Database Search (DBS) Module] (3), will search the database containing the sequence of pattern vectors of known programming. If a match is found, then the module returns a set of registration numbers otherwise referred to herein as program-ids and frame-ids, referred to also as frame numbers, corresponding to the identities of a set of audio or video programs and the time frame numbers within these programs where the match occurred. If the search of the database

fails to find a match, the DBS Module will issue a NO-MATCHED flag. It is contemplated that aspects of the invention for the DBS Module are applicable to any kind of data set containing signal signatures, even signatures derived using techniques distinct from those used in the Pattern Vector [Generation] module.

Ex. F, [0027], [0031] (emphasis added). Thus, it would have been obvious to similarly obtain the sequence of content identification results by querying the database, as taught by O'Hagan in view of Cheung. Doermann, [0128].

- vi. 1[c.ii]: a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and**

O'Hagan discloses this limitation. For example, O'Hagan discloses that a content identification result of the sequence of content identification results (*e.g.*, a particular recognition and classification result in recognition file 74) comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples (*e.g.*, a result from an extracted signature that is successfully classified); or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database (*e.g.*, a result from an extracted signature that is not classified). *See* the discussion of claim 1[c.ii] in Section V.A above. Doermann, [0129].

O'Hagan in view of Cheung further renders obvious the content identification results including a content identifier because Cheung teaches returning content identifiers in response to a database query match, as discussed with respect to claim 1[b] above. Doermann, [0130].

Cheung further teaches an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database (*e.g.*, returning a "NO MATCHED" indication in response to a database query that does not find a match). Doermann, [0131].

For example, Cheung discloses "search[ing] for a match" for calculated signatures in a "database," and "[w]hen a match is found and confirmed, the system uses the cross-referenced

identity in the database to provide the identity of the content that is currently being played.” Ex. F, [0019], [0021]. For example, a “program-id” and “frame-id” are returned for successful database searches. Ex. F, [0031], [0033], [0076]. If a match is not found, then a “NO MATCHED” indication is returned instead. Ex. F, [0019], [0031], [0033], [0076], Figs. 2 and 7. Doermann, [0132].

To the extent Patent Owner argues that O’Hagan does not explicitly describe an indication of the absence of a respective likely match, a POSA would have been motivated to modify O’Hagan’s system to include the indication, such as a “NO MATCHED” indication, to return an actual result that positively identifies the absence of a match and that provides positive assurance that an attempt to find match in the database was made, as discussed in Section V.B.3 above, in light of Cheung’s teachings. Thus, if an extracted signature is classified as non-recognized or unmatched in response to querying the electronic database, as disclosed by O’Hagan, it would have been obvious in view of Cheung to classify those signatures in the recognition file 74, which stores the classifications of extracted signatures, as “NO MATCHED” to indicate that a database query was performed and that the result of the database query was that the extracted signature was non-recognized or unmatched. Doermann, [0133].

- vii. 1[c.iii]: the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database;**

O’Hagan discloses this limitation. For example, O’Hagan discloses that the raw play stream (*e.g.*, the recognition file) includes either: at least two different content identifiers obtained from the electronic database (*e.g.*, extracted signatures classified as programs and commercial advertisements); or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database (*e.g.*, extracted signatures classified as programs and commercial advertisements, along with extracted signatures classified as non-recognized and unmatched). *See* the discussion of claim 1[c.iii] in Section V.A above. Doermann, [0134].

O'Hagan in view of Cheung further renders obvious the content identifiers obtained from the electronic database and the indication of the absence of a respective likely match because Cheung teaches returning content identifiers and "NO MATCH" indications in response to database queries that are successful or unsuccessful, respectively, as discussed with respect to claims 1[b] and 1[c.ii] above. Therefore, where O'Hagan discloses storing classifications of extracted signatures in recognition file 74 including two different content identifiers or one content identifier and one absence of a likely match as a result of querying the electronic database, it would have been obvious in view of Cheung to classify those signatures with content identifiers like program-ids or with "NO MATCH" indications. Doermann, [0135].

- viii. 1[d]: (d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[d] in Section V.A above. Doermann, [0136].

- ix. 1[e]: (e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[e] in Section V.A above. Doermann, [0137].

- b. Claim 2: The method of claim 1 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.**

O'Hagan in view of Apel and Cheung renders obvious the method of claim 1, including the raw play stream including a sequence of content identification results such as content identifiers (*e.g.*, program-ids) and indications of the absence of matches (*e.g.*, NO-MATCHED indications), as discussed above with respect to claims 1[c]-1[d]. O'Hagan in view of Apel further renders obvious the limitations of this claim, wherein sample sequence data includes log times and content identifiers. *See* the discussion of claim 2 in Section V.A above. And, as

discussed above, O'Hagan in view of Cheung renders obvious content identifiers such as program-ids. Doermann, [0138].

- c. Claim 3: The method of claim 1 wherein: scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content of identification results.**

O'Hagan in view of Apel and Cheung renders obvious the method of claim 1. O'Hagan in view of Apel further renders obvious the limitations of this claim. *See* the discussion of claim 3 in Section V.A above. Doermann, [0139].

- d. Claim 5: The method of claim 1 wherein a clean play list is generated from the clean play stream.**

O'Hagan in view of Apel and Cheung renders obvious the method of claim 1. O'Hagan in view of Apel further renders obvious the limitations of this claim. *See* the discussion of claim 5 in Section V.A above. Doermann, [0140].

- e. Claim 6: The media measurement method of claim 1 wherein the plurality of audio data representations include a plurality of audio signatures.**

O'Hagan in view of Apel and Cheung renders obvious the method of claim 1. O'Hagan further discloses the limitations of this claim. *See* the discussion of claim 6 in Section V.A above. Doermann, [0141].

- f. Claim 9**

- i. 9[pre]: A computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers, the one or more computers being coupled to a network to receive, over the network, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network, the instructions comprising instructions for:**

To the extent that the preamble of claim 9 is limiting, O'Hagan in view of Apel and Cheung renders it obvious. Doermann, [0142].

O'Hagan in view of Apel renders obvious one or more computers being coupled to a network to receive, over the network, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network. *See* the discussion of claim 1[a] in Section V.A above. Doermann, [0143].

O'Hagan further discloses that its system including one or more computers and processors for performing the described methods. For example, O'Hagan discloses that ~~station monitoring unit 20~~ includes audio processing circuitry 28, an A/D converter 40, and audio signature extraction circuit and events detector 48 that process digital signals. Ex. D, 4:65-5:22; *see also* 5:22-39 (describing digital video processing). Furthermore, O'Hagan discloses that ~~signature and recognition post-processing unit 64~~ includes a signature recognition computer 70, database memory 72, and post-processing computer 76, and discusses generating files, logs, flags, and codes. Ex. D, 5:40-6:12. *See* Fig. 1 below. A POSA would have recognized that ~~station monitoring unit 20~~ and ~~signature and recognition post-processing unit 64~~ each include one or more computers and one or more processors to process the above-described signals and files. Doermann, [0144].

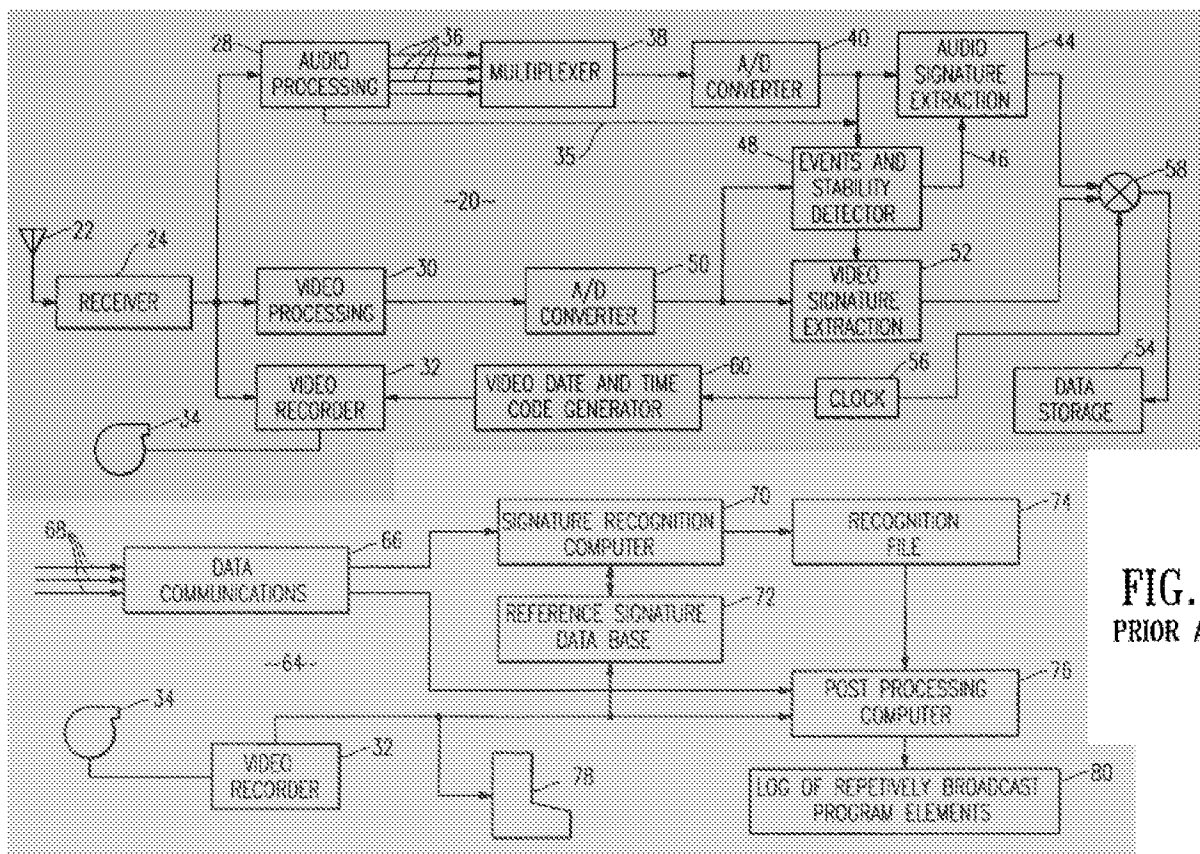


FIG. 1
PRIOR ART

Cheung teaches implementation of a system and method for broadcast audio and video detection and identification with via a “computer containing a central processing unit” or “CPU.” Ex. F, [0020]. The computer is configured to register audio signatures (*i.e.*, create database records of audio signatures and corresponding timing data and identifying information) and also detect and identify new audio data (*i.e.*, receive audio data, calculate its audio signature, and query the database to find a match). Ex. F, [0019]-[0021]. Cheung further teaches that the computer runs “software” or “programs” to perform the described methods and that the software and programs are stored on computer readable media, such as a hard drive or CD-ROM. [0019], [0021], [0219], claim 35. Doermann, [0145].

To the extent Patent Owner argues that O’Hagan does not explicitly describe a computer readable medium including instructions that, when executed by a computer, perform the method, a POSA would have been motivated to modify O’Hagan’s system to include a computer processor executing instructions stored on a computer readable medium to take advantage of commodity computer hardware and readily transportable software, as discussed in Section V.B.3, in light of Cheung’s teachings. Doermann, [0146].

- ii. 9[a]: (a) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers;**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[b] in Section V.A above. Doermann, [0147].

O'Hagan in view of Apel and Cheung further renders obvious this limitation. *See* the discussion of claim 1[b] above. Doermann, [0148].

- iii. 9[b]: (b) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c] in Section V.A above. Doermann, [0149].

O'Hagan in view of Apel and Cheung further renders obvious this limitation. *See* the discussion of claim 1[c] above. Doermann, [0150].

- iv. 9[b.i]: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database;**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c.i] in Section V.A above. Doermann, [0151].

O'Hagan in view of Apel and Cheung further renders obvious this limitation. *See* the discussion of claim 1[c.i] above. Doermann, [0152].

- v. 9[b.ii]: a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c.ii] in Section V.A above. Doermann, [0153].

O'Hagan in view of Apel and Cheung further renders obvious this limitation. *See* the discussion of claim 1[c.ii] above. Doermann, [0154].

- vi. 9[b.iii]: the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database;**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c.iii] in Section V.A above. Doermann, [0155].

O'Hagan in view of Apel and Cheung further renders obvious this limitation. *See* the discussion of claim 1[c.iii] above. Doermann, [0156].

- vii. 9[c]: (c) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[d] in Section V.A above. Doermann, [0157].

- viii. 9[d]: (d) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[e] in Section V.A above. Doermann, [0158].

- g. Claim 10: The computer program product of claim 9 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.**

O'Hagan in view of Apel and Cheung renders obvious the computer program product of claim 9. O'Hagan in view of Apel further renders obvious the limitations of this claim. *See* the discussion of claim 2 in Section V.A above. Doermann, [0159].

- h. Claim 11: The computer program product of claim 9 wherein scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content of identification results.**

O'Hagan in view of Apel and Cheung renders obvious the computer program product of claim 9. O'Hagan in view of Apel further renders obvious the limitations of this claim. *See* the discussion of claim 3 in Section V.A above. Doermann, [0160].

- i. Claim 13: The computer program product of claim 9 wherein the computer program product further comprises instructions for generating a clean play list from the clean play stream.**

O'Hagan in view of Apel and Cheung renders obvious the computer program product of claim 9. O'Hagan in view of Apel and Cheung further renders obvious the limitations of this claim. *See* the discussion of claim 5 in Section V.A above and the discussion of 9[pre] above. Doermann, [0161].

- j. Claim 14: The computer program product of claim 9 wherein the plurality of audio data representations include a plurality of audio signatures.**

O'Hagan in view of Apel and Cheung renders obvious the computer program product of claim 9. O'Hagan further discloses the limitations of this claim. *See* the discussion of claim 6 in Section V.A above. Doermann, [0162].

C. Claims 7, 9-11, and 13-15 Are Obvious Over O'Hagan in view of Apel and Wang

1. Overview of Wang

Wang (Ex. H) qualifies as prior art under at least pre-AIA 35 U.S.C. § 102(b) because it was published on 6/27/2002, more than one year prior to the '768 patent's earliest possible priority date of 5/27/2004. Wang is directed to a system and method for matching an audio sample to an audio sample in a database even when audio is highly distorted or contains a high level of noise. Ex. H, Abstract, [0002]. Wang discloses that matching is enhanced by comparing "fingerprints" at particular locations of the audio. Ex. H, [0018]. The locations used for analysis are "reproducibly computable locations ... called 'landmarks,'" where "each fingerprint represents a number of features of the media sample at each location." Ex. H, [0020]. Fig. 1, below, illustrates a method for recognizing an audio sample. Doermann, [0164].

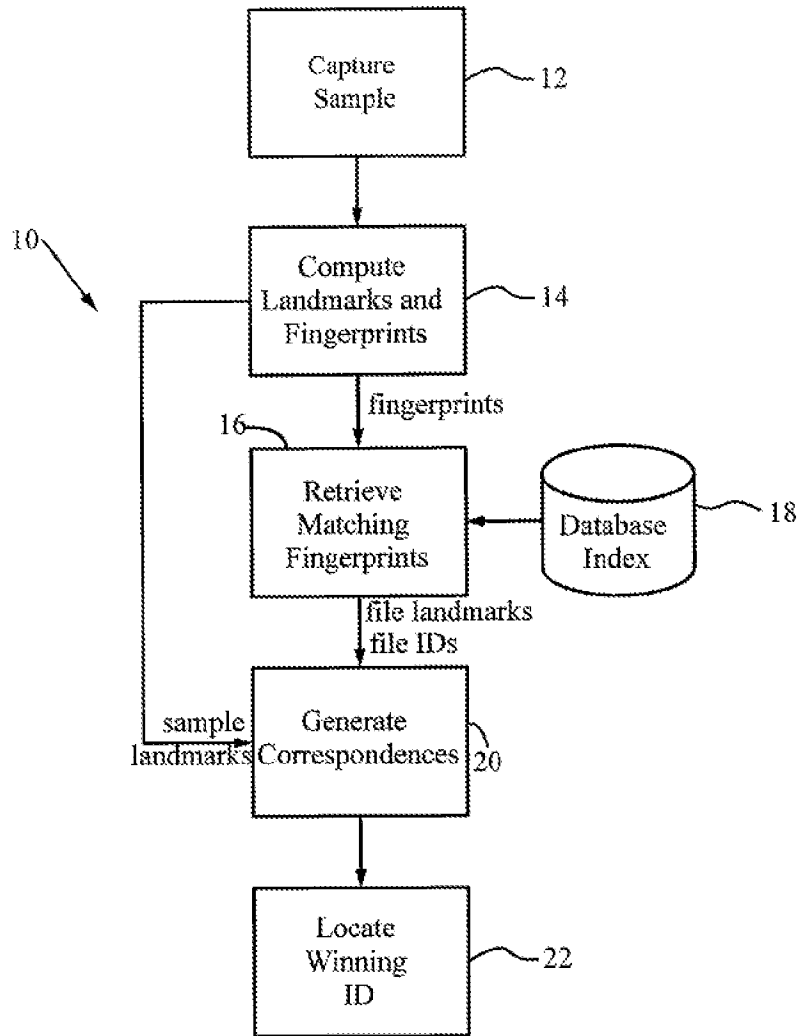


FIG. 1

An audio sample is captured, and its landmarks and fingerprints are computed. Ex. H, [0039] (steps 12 and 14).

A preferred landmarking algorithm is able to mark the same timepoints within a sound recording despite the presence of noise and other linear and nonlinear distortion.... Landmarking results in a list of timepoints {landmark_k} within the sound recording at which fingerprints are subsequently calculated.

Ex. H, [0050].

The fingerprint is generally a value or set of values that summarizes a set of features in the recording at or near the timepoint. In a currently preferred embodiment, each fingerprint is a single numerical value that is a hashed function of multiple features. Possible types of fingerprints include spectral slice fingerprints,

multi-slice fingerprints, LPC coefficients, and cepstral coefficients.... Fingerprints can be computed by any type of digital signal processing or frequency analysis of the signal.

Ex. H, [0055]. Doermann, [0165].

The computed fingerprints are used to retrieve sets of matching fingerprints in a database, which have associated landmarks and identifiers. Ex. H, [0040] (step 16). “Correspondences” between the audio sample fingerprints and the matching fingerprints from the database are then generated. Ex. H, [0040] (step 20). That is, the system seeks to determine whether sets of matching fingerprints occur at the same corresponding timepoint landmarks such that a “time evolution of fingerprints” match. Ex. H, [0018], [0040], [0084], [0085]. A winning file identifier is identified as the set of correspondences having the most linearly related correspondences. Ex. H, [0040] (step 22). Doermann, [0166].

Wang further discloses implementing its system and method with software instructions stored on a computer-readable medium, executed by one or more computers. Ex. H, [0023], [0042], [0110]. Doermann, [0167].

2. O’Hagan, Apel, and Wang are Analogous Art

O’Hagan and Apel are analogous to the ’768 patent, as discussed in Section V.A.3 above. Wang is also analogous to the ’768 patent because it is in the same field of endeavor as the ’768 patent or reasonably pertinent to the alleged problems discussed therein. Doermann, [0168].

For example, the ’768 patent is in the field of “media monitoring and measurement” and the “recognition” of media content (*see* Ex. A, Abstract, 1:60-2:15), and, similarly, Wang is in the field of “content-based information retrieval” and the “recognition of an audio signal, including sound or music” (*see* Ex. H, Abstract, [0002]). Doermann, [0168].

Moreover, Wang seeks to provide a system that addresses “the need for automatic recognition of music or other audio signals generated from a variety of sources” that “performs well under conditions of high noise and distortion” and that improves on systems that “rely on embedding inaudible codes” or that would require searching a database of “millions of sound[] recordings” that is “economically unfeasible.” Ex. H, [0003]-[0004], [0009]-[0010]. Wang’s disclosure would therefore have been reasonably pertinent to problems alleged in the ’768 patent, which include the desire to track usage “across several types of media” such as radio, CDs, and others, to avoid relying on the use of “codes” that “are limited in that they can only track content

that has been properly encoded,” to avoid “computationally intensive” database searches, and to be less susceptible to “high-noise environments.” *See* Ex. A, 1:14-49. Doermann, [0168].

Accordingly, O’Hagan, Apel, and Wang are each analogous to the ’768 patent.

3. A POSA Would Have Been Motivated to Modify O’Hagan in view of Apel’s and Wang’s Teachings

O’Hagan in view of Apel renders obvious claims 1-3 and 5-6, as discussed in Section V.A above. For the same reasons, O’Hagan in view of Apel renders obvious the method described in claims 9-11 and 13-14, which is substantially similar to the method described in claims 1-3 and 5-6 but performed by executing instructions embodied on a computer program product. To the extent Patent Owner may argue that O’Hagan and Apel do not explicitly describe specific implementation examples or details in claims 7, 9, and 15, Cheung teaches that such implementations would have been well-known and obvious to a POSA. Doermann, [0169].

Regarding claims 7 and 15, O’Hagan discloses querying an electronic database and comparing extracted signatures to reference signatures to classify extracted signatures. To the extent argued that O’Hagan does not explicitly describe audio data representations that include “signal landmarks,” Wang teaches that the use of landmarks in addition to signatures or fingerprints significantly improves recognition results: “A key insight into the matching process is that the time evolution in matching sounds must follow a linear correspondence.... Thus, the correct landmark pairs ... must have a linear correspondence.... While it is possible for two distinct signals to contain a number of identical fingerprints, it is very unlikely that these fingerprints have the same relative time evolutions. The requirement for linear correspondences ... provides significantly better recognition than techniques that simply count the total number of features in common or measure the similarity between features. In fact, because of this aspect of the invention, sounds can be recognized even if fewer than 1% of the original recording’s fingerprints appear in the exogenous sound sample, i.e., if the sound sample is very short or if it is significantly distorted.” Ex. H, [0084]-[0085]. A POSA would have been motivated to modify O’Hagan’s signatures to have associated landmarks and to use those landmarks for classifying the signatures, as taught by Wang, to advantageously provide additional information relating to the signatures that improves matching accuracy. Moreover, the modification would have been, at least, the use of a known technique (*e.g.*, Wang’s teaching of landmarks) to improve similar devices in the same way (*e.g.*, to improve matching accuracy in recognition systems that use

signatures or fingerprints, like O'Hagan's and Wang's). Doermann, [0170]. *See* MPEP § 2143 (citing *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007)).

Regarding independent claim 9 and claims depending therefrom, O'Hagan discloses implementing its system and method with computers. *See, e.g.*, Ex. D, 4:65-6:12 (describing various processing "circuitry," "units," and "computers," an A/D converter, processing digital signals, using a database memory, etc.). To the extent argued that O'Hagan does not explicitly describe its systems (*e.g.*, station monitoring unit 20 and signature and post-processing unit 64) implemented as one or more computers executing instructions stored on a non-transitory computer readable medium, Wang teaches implementation via one or more processors of one or more computers executing code stored in memory, a CD-ROM, and the like. Ex. H, [0023], [0042], [0110]. A POSA would have been motivated to implement O'Hagan's station monitor units and post-processing units as computers with processing units executing software stored on computer readable media to take advantage of commodity hardware and readily transportable software. Indeed, O'Hagan contemplates the use of "hardware as simple as a Macintosh computer" to avoid "resorting to supercomputers or expensive data storage facilities." Ex. D, 4:5-13, 9:35-43. Wang's teachings of general-purpose computers executing instructions on various forms of computer readable media facilitates O'Hagan's desired implementation. Moreover, the implementation of O'Hagan's system with computers would have been obvious because known work in a field of endeavor (*e.g.*, Wang's use of processors executing software stored on computer readable media) may prompt variations of it for use in the same field (*e.g.*, using the same for O'Hagan's recognition system) based on design incentives or other market forces if the variations are predictable to one of ordinary skill in the art (*e.g.*, making use of commercially available hardware and distributable software taught by O'Hagan and Wang). Doermann, [0171]. *See* MPEP § 2143 (citing *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 415-421 (2007)).

A POSA also would have had a reasonable expectation of success modifying O'Hagan's system to utilize landmarks and executable instructions stored on computer readable media, as discussed above. The modifications amount merely to additional computer calculations and determinations or particular implementation details for O'Hagan's existing system and method, and Wang teaches specific landmark and implementation examples, *e.g.*, landmarks calculated at "local maxima of spectral Lp norms of the audio sample" and executing instructions stored on

CD-ROM, that were obvious and well-known to POSAs. Ex. H, [0021], [0110]. Doermann, [0172].

Claims 7, 9-11, and 13-15 are therefore obvious over O'Hagan in view of Apel and Wang as discussed herein. Doermann, [0173]. The examiner did not appear to have these teachings before him when allowing the claims, which seemingly were allowed for teaching a raw playstream including unmatched samples and a scrubbing correction mechanism for changing those unmatched samples.

4. Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim

a. Claim 7: The media measurement method of claim 1 wherein an audio data representation of the plurality of audio data representations includes a plurality of signal landmarks.

O'Hagan in view of Apel renders obvious the method of claim 1. O'Hagan in view of Wang further renders obvious the limitations of this claim. Doermann, [0174].

O'Hagan discloses the plurality of audio data representations including audio signatures, as discussed with respect to claims 1[b], 1[c.i], and 6 in Section V.A above. Doermann, [0175].

Wang teaches an audio data representation of a plurality of audio data representations including a plurality of signal landmarks (*e.g.*, “landmarks”). Ex. H, [0020], [0039], [0040], [0048]-[0051]. For example, Wang teaches that to “index the sound database, each recording in the collection is subjected to a landmarking and fingerprinting analysis.... Landmarks occur at specific timepoints of the sound and have values in time units offset from the beginning of the file, while fingerprints characterize the sound at or near a particular landmark.” Ex. H, [0049]. “A good landmarking scheme marks about 5-10 landmarks per second of sound recording.” Ex. G, [0050]. Doermann, [0176].

To the extent Patent Owner argues that O'Hagan does not explicitly describe an audio data representation including a plurality of signal landmarks, a POSA would have been motivated to modify O'Hagan's signatures to have associated landmarks and to use those landmarks for classifying the signatures to improve matching accuracy as discussed in Section V.C.3 above, in light of Wang's teachings. Doermann, [0177].

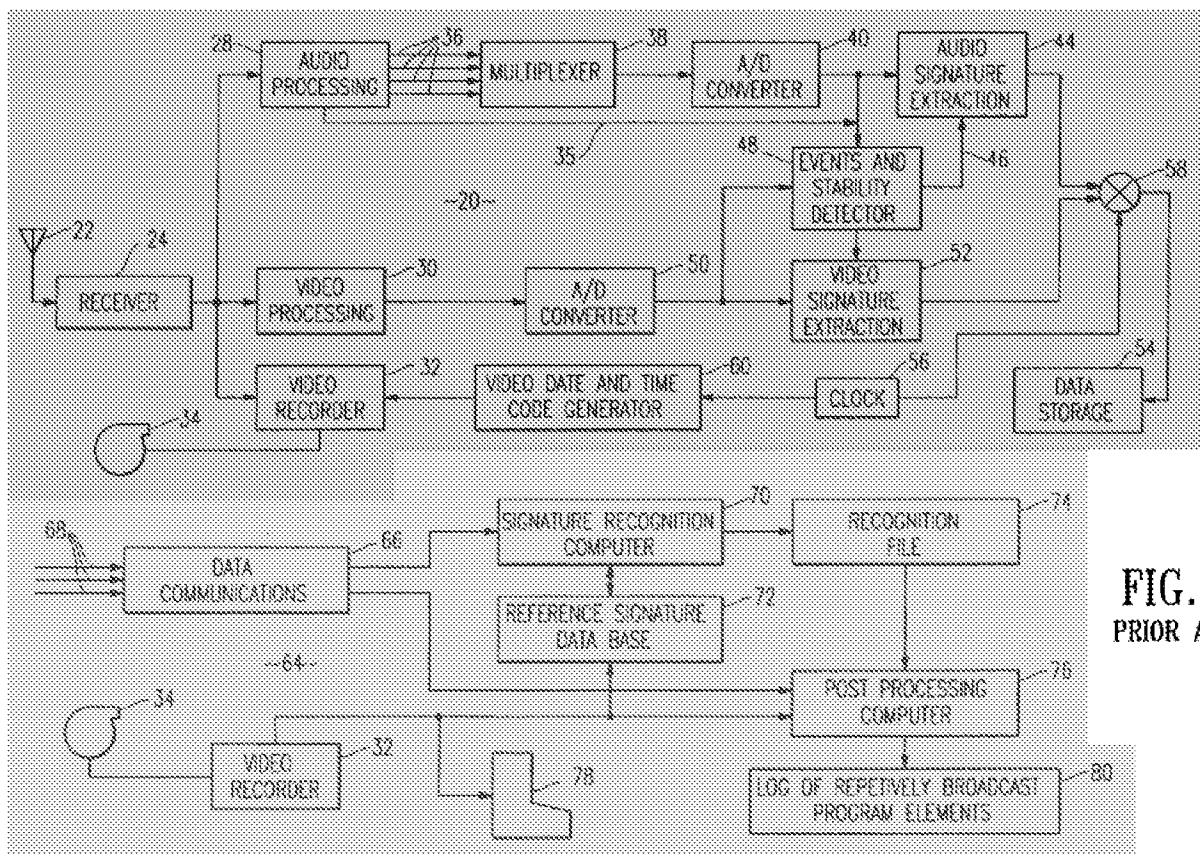
a. Claim 9

- i. 9[pre]: A computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers, the one or more computers being coupled to a network to receive, over the network, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network, the instructions comprising instructions for:**

To the extent that the preamble of claim 9 is limiting, O'Hagan in view of Apel and Wang renders it obvious. Doermann, [0178].

O'Hagan in view of Apel renders obvious one or more computers being coupled to a network to receive, over the network, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network. *See* the discussion of claim 1[a] in Section V.A above. Doermann, [0179].

O'Hagan further discloses that its system including one or more computers and processors for performing the described methods. For example, O'Hagan discloses that *station monitoring unit 20* includes audio processing circuitry 28, an A/D converter 40, and audio signature extraction circuit and events detector 48 that process digital signals. Ex. D, 4:65-5:22; *see also* 5:22-39 (describing digital video processing). Furthermore, O'Hagan discloses that *signature and recognition post-processing unit 64* includes a signature recognition computer 70, database memory 72, and post-processing computer 76, and discusses generating files, logs, flags, and codes. Ex. D, 5:40-6:12. *See* Fig. 1 below. A POSA would have recognized that *station monitoring unit 20* and *signature and recognition post-processing unit 64* each include one or more computers and one or more processors to process the above-described signals and files. Doermann, [0180].



Wang teaches implementation of a system and method for matching audio samples via one or more processors of one or more computers executing code stored in memory, a CD-ROM, and the like. Ex. H, [0023], [0042], [0110]. Systems executing the code are configured to generate landmarks and fingerprints of audio samples, search a database for matching audio samples, and identify best matching audio samples. Ex. H, [0042]. Doermann, [0181].

To the extent Patent Owner argues that O'Hagan does not explicitly describe a computer readable medium including instructions that, when executed by a computer, perform the method, a POSA would have been motivated to modify O'Hagan's system to include a computer processor executing instructions stored on a computer readable medium to take advantage of commodity computer hardware and readily transportable software, as discussed in Section V.C.3, in light of Wang's teachings. Doermann, [0182].

ii. 9[a]: (a) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers;

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[b] in Section V.A above. Doermann, [0183].

- iii. 9[b]: (b) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c] in Section V.A above. Doermann, [0184].

- iv. 9[b.i]: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database;**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c.i] in Section V.A above. Doermann, [0185].

- v. 9[b.ii]: a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c.ii] in Section V.A above. Doermann, [0186].

- vi. 9[b.iii]: the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database;**

O'Hagan discloses this limitation. *See* the discussion of claim 1[c.iii] in Section V.A above. Doermann, [0187].

- vii. 9[c]: (c) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[d] in Section V.A above. Doermann, [0188].

- viii. 9[d]: (d) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.**

O'Hagan in view of Apel renders obvious this limitation. *See* the discussion of claim 1[e] in Section V.A above. Doermann, [0189].

- b. Claim 11: The computer program product of claim 9 wherein scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content of identification results.**

O'Hagan in view of Apel and Wang renders obvious the computer program product of claim 9. O'Hagan in view of Apel further renders obvious the limitations of this claim. *See* the discussion of claim 3 in Section V.A above. Doermann, [0190].

- c. Claim 13: The computer program product of claim 9 wherein the computer program product further comprises instructions for generating a clean play list from the clean play stream.**

O'Hagan in view of Apel and Wang renders obvious the computer program product of claim 9. O'Hagan in view of Apel and Wang further renders obvious this claim. *See* the discussion of claim 5 in Section V.A above and the discussion of 9[pre] above. Doermann, [0191].

- a. Claim 14: The computer program product of claim 9 wherein the plurality of audio data representations include a plurality of audio signatures.**

O'Hagan in view of Apel and Wang renders obvious the computer program product of claim 9. O'Hagan further discloses the limitations of this claim. *See* the discussion of claim 6 in Section V.A above. Doermann, [0192].

- b. Claim 15: The computer program product of claim 9 wherein an audio data representation of the plurality of audio data representations includes a plurality of signal landmarks.**

O'Hagan in view of Apel and Wang renders obvious the computer program product of claim 9. O'Hagan in view of Wang further renders obvious the limitations of this claim. *See* the discussion of claim 7 above. Doermann, [0193].

D. Claims 7 and 15 Are Obvious Over O'Hagan in view of Apel, Cheung, and Wang

1. Detailed Explanation of the Pertinency and Manner of Applying the Prior Art to Each Claim

- a. Claim 7: The media measurement method of claim 1 wherein an audio data representation of the plurality of audio data representations includes a plurality of signal landmarks.**

Claim 7 is obvious over O'Hagan in view of Apel and Wang, as discussed in Section V.C above. For the same reasons, it would have been obvious to a POSA to modify O'Hagan in view of Apel and Cheung further in view of Wang's teachings and modify O'Hagan's signatures to have associated landmarks and to use those landmarks for classifying the signatures, further rendering obvious claim 7 over O'Hagan in view of Apel, Cheung, and Wang. Doermann, [0194]-[0195].

- b. Claim 15: The computer program product of claim 9 wherein an audio data representation of the plurality of audio data representations includes a plurality of signal landmarks.**

O'Hagan in view of Apel, Cheung, and Wang renders obvious this claim. *See* the discussion of claim 7 above. Doermann, [0194]-[0195].

VI. CONCLUSION

For the foregoing reasons, *ex parte* reexamination of claims 1-3, 5-7, 9-11, and 13-15 of U.S. Patent No. 8,510,768 is respectfully requested.

Dated: January 16, 2024

Respectfully submitted,

By /Scott A. McKeown/
Scott A. McKeown, Reg. No. 42,866
Victor Cheung, Reg. No. 66,229
WOLF GREENFIELD & SACKS, P.C.
601 Massachusetts Avenue NW
Washington, DC 20001