

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ROKU, INC.,
Petitioner

v.

ANONYMOUS MEDIA RESEARCH HOLDINGS LLC,
Patent Owner

Case IPR2024-01055
U.S. Patent No. 8,756,622

DECLARATION OF SAMUEL H. RUSS

Mail Stop "PATENT BOARD"
Patent Trial and Appeal Board
U.S. Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

Roku EX1003
IPR2024-01055
U.S. Patent No. 8,756,622

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	MY BACKGROUND AND QUALIFICATIONS	2
III.	MATERIALS CONSIDERED	4
IV.	RELEVANT LEGAL STANDARDS	7
	A. Obviousness.....	8
	B. Level of Skill in the Art.....	9
	C. Motivation to Combine and Hindsight Bias.....	10
	D. Claim Construction	12
V.	TECHNOLOGY BACKGROUND.....	12
	A. Content Identification Was Well-Known.....	13
	B. Generating Play Streams Was Well-Known	13
	C. Identifying a Channel Was Well-Known.....	16
VI.	OVERVIEW OF THE '622 PATENT	17
	A. Summary of the Disclosure	17
	B. Prosecution History	29
	C. Priority Date of the '622 Patent.....	32
VII.	OVERVIEW OF THE PRIOR ART	33
	A. Wang777	33
	1. Overview of Wang.....	37
	B. Overview of Wang060	40
	C. Overview of Herley682	42
	D. Seet	45
	1. Seet is prior art as of February 17, 2004.....	45
	2. Overview	49
	E. Overview of Pitman.....	53
	F. Overview of Schrempp.....	55
	G. Overview of Wells.....	56

VIII.	CLAIM CONSTRUCTION	57
IX.	CLAIMS 1-13 OF THE '622 PATENT ARE OBVIOUS UNDER 35 U.S.C. § 103 OVER WANG777 AND WANG060.....	58
A.	Motivation to Combine Wang777 with Wang060	58
B.	Independent Claim 1	61
1.	Preamble [1.P].....	61
2.	Element [1.1].....	65
3.	Element [1.2].....	73
4.	Element [1.3].....	75
5.	Element [1.4].....	81
C.	Independent Claims 6 and 10	82
1.	[6.P], [10.P]	82
2.	[6.1], [10.1]	83
3.	[6.2], [10.2]	83
4.	[6.3], [10.3]	83
5.	[6.4], [10.4]	83
D.	Dependent Claims 2, 7, and 11	83
E.	Dependent Claims 3, 8, and 12	85
F.	Dependent Claim 4.....	85
G.	Dependent Claims 5, 9, and 13	86
X.	CLAIMS 2-4, 7-8, AND 11-12 OF THE '622 PATENT ARE OBVIOUS UNDER 35 U.S.C. § 103 OVER WANG777, WANG060, HERLEY682, SEET, AND THE DATA IMPUTATION REFERENCES.....	90
A.	Motivation to Combine Wang777, Wang060, Herley682, and Seet ..	90
B.	Dependent Claims 2, 7, and 11	94
C.	Dependent Claims 3, 8, and 12	101
D.	Dependent Claim 4.....	102
XI.	CLAIMS 1-13 OF THE '622 PATENT ARE OBVIOUS UNDER 35 U.S.C. § 103 OVER PITMAN, SCHREMPP, AND WELLS.....	105
A.	Motivation to Combine Pitman with Schrempp and Wells	105

B.	Independent Claim 1	114
1.	Preamble [1.P].....	114
2.	Element [1.1].....	115
3.	Element [1.2].....	124
4.	Element [1.3].....	125
5.	Element [1.4].....	130
C.	Independent Claim 6 and 10.....	131
1.	[6.P], [10.P].....	131
2.	[6.1], [10.1]	132
3.	[6.2], [10.2]	132
4.	[6.3], [10.3]	132
5.	[6.4], [10.4]	132
D.	Claim 2, 7, and 11	133
E.	Dependent Claims 3, 8, and 12	140
F.	Dependent Claim 4.....	140
G.	Dependent Claims 5, 9, and 13	141
XII.	CONCLUSION.....	142

I, Samuel H. Russ, hereby declare as follows.

I. INTRODUCTION

1. I have been engaged by Sterne, Kessler, Goldstein & Fox P.L.L.C. on behalf of Roku, Inc. (“Roku”) to provide this Declaration concerning technical subject matter relevant to the *inter partes* review of U.S. Patent No. 8,756,622 (EX1001, “’622 patent”), which is entitled “Media Usage Monitoring and Measurement System and Method.”

2. I am over 18 years of age. I have personal knowledge of the facts stated in this Declaration.

3. I have reviewed, and am familiar with, the specification, claims, and prosecution history of the ’622 patent. I will cite to the specification of a U.S. Patent using the following format: Column:Line Number(s). For example, the citation to EX1001, 1:1-10, points to the ’622 patent specification at column 1, lines 1-10.

4. In forming my opinions expressed in this Declaration, I have reviewed and, in some cases, relied upon the list of materials provided in Section III. To the best of my knowledge, those exhibits listed in Section III are true and accurate copies of what they purport to be. An expert in the field would reasonably rely on them to formulate opinions such as those set forth in this Declaration.

5. I am being compensated at my standard rate by Roku for my time

spent in connection with this proceeding. My compensation is not dependent on, and in no way affects, the substance of my opinions. Nor is my compensation dependent on the outcome of this proceeding. My findings are based on my education, experience, and background in the fields discussed below.

II. MY BACKGROUND AND QUALIFICATIONS

6. My academic and professional pursuits are closely related to the subject matter of the '622 patent. Based on my education and work experience, I am well qualified to serve as a technical expert in this matter.

7. A summary of my educational and professional experience is set forth in my *curriculum vitae* ("CV"), which is submitted as Exhibit 1004. Exhibit 1004 also includes a list of my publications and the cases in which I have testified at deposition, hearing, or trial. Some of the relevant points in my CV are described below.

8. I received a Bachelor's degree in Electrical Engineering from the Georgia Institute of Technology ("Georgia Tech") in 1986 and a Ph.D. in Electrical Engineering from Georgia Tech in 1991.

9. From 2007 to the present, I have been a member of the faculty of the University of South Alabama as an Assistant and Associate Professor in the Department of Electrical and Computer Engineering. During that time, I have won awards for excellent teaching and have been actively publishing research in home

networking and digital video recording (DVR) technologies. I am active in the Institute of Electrical and Electronic Engineers (IEEE) and am a Distinguished Lecturer for the IEEE Consumer Electronics Society. As a consultant, I have conducted briefings for members of the financial community on technology trends in the cable, satellite, and IPTV sectors.

10. From 2000 to 2007, I worked for Scientific-Atlanta (now Cisco's Service Provider Video Tech. Group), where I managed a cable set-top box (STB) design group that designed four STB models, including the Explorer 4200 (non-DVR) and 8300 (DVR) models. Both models sold several million units. As design group manager, I was responsible for managing the design and prototyping activities of the group and for interfacing with other groups (especially integrated-circuit design, procurement, software developers, the factory where prototypes were built, and product managers) and for maintaining the hardware and mechanical development schedule. Since the products were produced in extremely high volumes, the projects had very high visibility in the company, and therefore carried a great deal of responsibility.

11. Also while at Scientific-Atlanta, I became a staff expert in home networking, conducting demonstrations of wireless video technology and managing a group that developed a new coaxial home networking system. The coaxial system won a Technology and Engineering Emmy® Award in 2013. I

became a staff expert in DVR reliability, and led a team that improved the software, hardware, repair, and manufacturing processes. I am a named inventor on over fifty patent applications that were filed while I was at Scientific-Atlanta, twenty-nine of which have issued as U.S. patents, as of the writing of this report. Many have also issued as European and Canadian patents.

12. From 1999 to 2000, I was a Staff Electrical Engineer and then Matrix Manager at IVI Checkmate (now Ingenico), where I managed the hardware design team that completed the design of the eN-Touch 1000 payment terminal. This terminal was in widespread use, for example, at the self-checkout at Home Depot.

13. I also served on the faculty of Mississippi State University from 1994 to 1999 as an Assistant Professor in the Department of Electrical & Computer Engineering, where I taught circuit board design and two-way interactive video classes, among other things.

14. I have also authored at least 15 journal articles and at least 31 conference papers, including a conference paper on digital video recording that won second place in a “best paper” competition at the 2011 International Conference on Consumer Electronics in Las Vegas, NV.

III. MATERIALS CONSIDERED

15. In formulating my opinions, I have relied upon my training, knowledge, and experience that are relevant to the '622 patent. Furthermore, I have

considered specifically the following documents listed below in addition to any other documents cited in this Declaration. I understand that the references are true and accurate copies of what they appear to be.

16. I have been asked to provide technical review, analysis, insights, and opinions regarding the references cited in the Petition and discussed below. My opinions are not only based on these references but also based on my over 30 years of experience as a computer architect, computer system designer, personal computer graphics designer, educator, and executive in the electronics industry.

Exhibit No.	Description
1001	U.S. Patent No. 8,756,622 to Steuer et al. (“the ’622 patent”)
1002	Prosecution File History of U.S. Patent No. 8,756,622 (“Prosecution History”)
1004	<i>Curriculum Vitae</i> of Samuel H. Russ
1005	U.S. Patent Application Publication No. 2007/0143777 to Wang (“Wang777”)
1006	U.S. Patent Application Publication No. 2002/0083060 to Wang (“Wang060”)
1007	U.S. Provisional Application No. 60/545,879 to Wang (“the ’879 application”)
1008	PCT/US05/05254 to Wang (“Wang254”)
1009	U.S. Prov. Appl. No. 60/574,836 (“the ’836 Appl.”)
1010	U.S. Appl. No. 11/139,330 (“the ’330 Appl.”)

Exhibit No.	Description
1011	U.S. Patent Application Publication No. 2004/0260682 to Herley et al. (“Herley682”)
1012	U.S. Patent Application No. 2005/0193016 to Seet et al. (“Seet”)
1013	U.S. Patent Application Publication No. 2002/0116195 Pitman et al. (“Pitman”)
1014	U.S. Patent Application Publication No. 2003/0018709 to Schrempp et al. (“Schrempp”)
1015	U.S. Patent Application Publication No. 2003/0086341 to Wells et al. (“Wells”)
1016	U.S. Provisional Application No. 60/545,681, filed February 17, 2004 (“Seet provisional”)
1017	Ingunn Myrtveit, Erik Stensrud, and Ulf H. Olsson “Analyzing Data Sets with Missing Data: An Empirical Evaluation of Imputation Methods and Likelihood-Based Methods,” in IEEE Transactions on Software Engineering, vol. 27, no. 11, pp. 999-1013, November 2001. (“Myrtveit”)
1018	Mong Li Lee, W. Hsu and Vijay Kothari, “Cleaning the spurious links in data,” in IEEE Intelligent Systems, vol. 19, no. 2, pp. 28-33, March-April 2004. (“Lee”)
1019	Kamakshi Lakshminarayan, Steven A. Harp, Tariq Samid, “Imputation of Missing Data in Industrial Databases,” Applied Intelligence 11, 259-275, (1999) (“Lakshminarayan”)
1020	Eloi Batlle, Jame Masip, Enric Gaus, “Automatic Song Identification in Noisy Broadcast Audio,” Audiovisual Institute and Department of Technology, Barcelona, Spain, (2002). (“Batlle”)
1022	U.S. Patent No. 5,581,658 to O’Hagen et al. (“O’Hagen”)
1023	Hernandez et al., “Real-world Data is Dirty: Data Cleansing and the Merge/Purge Problem,” Klumer Academic Publishers, 1998

17. This Declaration represents the opinions I have formed to date. I may consider additional documents as they become available or other documents that are necessary to form my opinions. I reserve the right to revise, supplement, or amend my opinions based on new information.

18. In light of my review of the exhibits, I recognize that the specification of the '622 patent describes systems and methods for enhancing recognition (e.g. in terms of accuracy or efficiency) of the content of a media sample by analyzing information exogenous to the sample.” EX1001, 1:66-2:3. In particular, the '622 patent purports to “analyze sequence data” to “establish channel identification.” *Id.*, 2:9-11.

19. I understand that the '622 patent purports to claim priority to U.S. Provisional Application No. 60/574,836, filed May 27, 2004. I am familiar with the technology at issue in the '622 patent and the state of the art as of May 27, 2004.

IV. RELEVANT LEGAL STANDARDS

20. When considering the '622 patent and stating my opinions, I am relying on legal principles that have been explained to me by counsel.

21. I understand that for a claim to be found patentable, the claims must be, among other requirements, novel and nonobvious from what was known at the time of the invention.

22. I understand that the information that is used to evaluate whether a

claim is novel and nonobvious is referred to as prior art.

23. I understand that in this proceeding, Petitioner has the burden of proving that each element of the challenged claims is disclosed by the prior art or would have been obvious based on the prior art.

A. Obviousness

24. I understand that a patent claim is invalid if the claims would have been obvious to a person of ordinary skill in the art (“POSA”) at the time of the invention. I understand that the obviousness inquiry should not be done in hindsight, but from the perspective of a POSA as of the time of invention of the patent claim.

25. I understand that to obtain a patent, the claims must have, as of the time of the invention, been nonobvious in view of the prior art.

26. I understand that a claim is obvious when the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to a POSA at the time the invention.

27. I understand that to prove that the prior art reference or a combination of prior art references renders a patent obvious, it is necessary to: (1) identify the particular references that, singly or in combination, make the patent obvious; (2) specifically identify which elements of the patent claim appear in each of the asserted references; and (3) explain how a POSA could have combined the prior art

references to create the claimed invention and what the combination would be.

28. I understand that to support a conclusion that a prior art reference or a combination of prior-art references renders a patent obvious, there must be some documentary evidence. Mere statements about what is basic knowledge or common sense, i.e., common knowledge as a replacement for documentary evidence, is insufficient to support a conclusion of obviousness.

29. I understand that certain objective indicia can be important evidence regarding whether a patent is obvious. Such indicia include: industry acceptance, commercial success of products covered by the patent claims; long-felt need for the invention; failed attempts by others to make the invention; copying of the invention by others in the field; unexpected results achieved by the invention as compared to the closest prior art; praise of the invention by the infringer or others in the field; taking of licenses under the patent by others; expressions of surprise by experts and those skilled in the art at the making of the invention; and the patentee proceeded contrary to the accepted wisdom of the prior art.

B. Level of Skill in the Art

30. I have been asked to consider the level of ordinary skill in the field that someone would have had at the time the claimed invention was made, which I understand is May 27, 2004. However, as I explain below in Section VI.C, it is my opinion that the earliest priority date for the '622 patent is May 26, 2005. My

opinions included herein do not change regardless of the May 27, 2004, or the May 26, 2005, dates. In deciding the level of ordinary skill, I considered the following:

- the educational level of the inventor;
- type of problems encountered in the art;
- prior art solutions to those problems;
- the rapidity with which innovations are made;
- the sophistication of the technology; and
- the educational level of active workers in the field.

31. In my opinion, a POSA at the time of the invention would have had a bachelor's degree in electrical engineering, computer engineering, computer science, or equivalent degree with at least two years of relevant industry experience, including in digital media content delivery.

32. Throughout my Declaration, even if I discuss my analysis in the present tense, I am always making my determinations based on what a POSA would have known at the time of the invention. Additionally, throughout my Declaration, even if I discuss something stating "I," I am referring to a POSA's understanding.

C. Motivation to Combine and Hindsight Bias

33. I understand that obviousness can be established by combining multiple prior art references to meet each and every claim element, but I also

understand that a proposed combination of references can be susceptible to hindsight bias.

34. I understand that references are more likely to be combinable if the nature of the problem to be solved is the same.

35. I understand that if the combination of references results in the references being unsatisfactory for their intended purposes or the combination changes the references' principle of operation, a POSA would not have a motivation to combine the references.

36. I understand that obviousness may be defeated if the prior art indicates that the invention would not have worked for its intended purpose or otherwise teaches away from the invention. I further understand that a reference teaches away when a POSA, upon reading the reference, would be discouraged from following the path set out in the reference, or would be led in a direction divergent from the path that was taken in the claim. I also understand that teaching away, e.g., discouragement, is strong evidence that the references are not combinable and/or that a POSA would not be motivated to combine them. I also understand that a disclosure of more than one alternative does not necessarily constitute a teaching away. I understand that the combination does not need to result in the most desirable embodiment.

D. Claim Construction

37. I understand that the claims of a patent are read in light of the specification of the patent as understood by POSA at the time of the invention. I understand that “the time of the invention” refers to the effective filing date of the patent in question, which I am informed by counsel is no earlier than May 27, 2004, for the purposes of this proceeding. Whether or not explicitly stated below, my analysis of the prior art is made as of the time the invention was made.

38. I understand that claim terms are given their plain and ordinary meaning as would be understood by a POSA as of the effective filing date. I understand that an inventor may provide a special meaning for a term, but that she must do so in the specification with precision and clarity.

39. In construing claims, I understand that evidence that is intrinsic to the patent and prosecution history is given more weight than extrinsic evidence. For example, if there are specific statements in the specification that define the invention, those statements are strong evidence of a definition for a term. In this Declaration, I have understood the claims in view of the general principles outlined above. Unless I indicate otherwise below, I have interpreted each claim term discussed herein to have its plain and ordinary meaning.

V. TECHNOLOGY BACKGROUND

40. The '622 patent is directed to providing “a media measurement

system and method that enhances recognition ... of the content of a media sample by analyzing information exogenous to the sample” and, specifically, analyzing “sequence data” to “establish channel identification.” EX1001, 1:66-3:3, 2:9-11.

A. Content Identification Was Well-Known

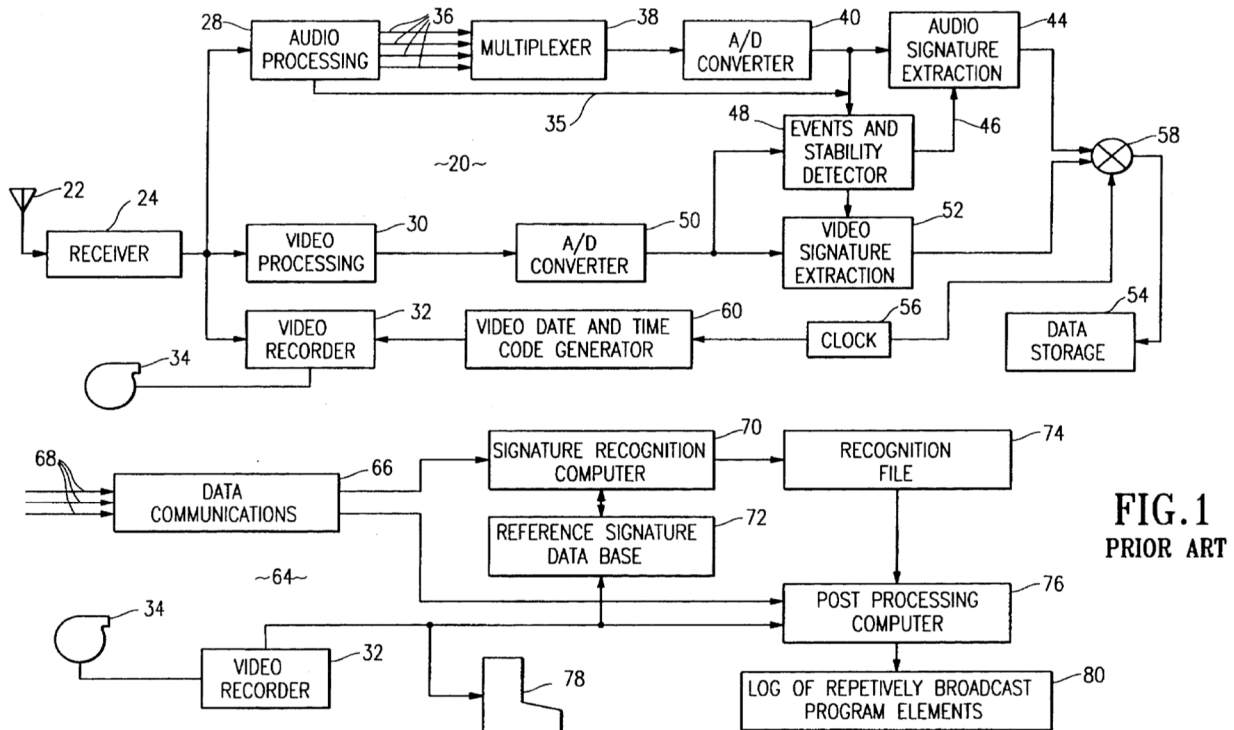
41. The '622 patent explains that “[c]ontent recognition system 124 is adapted to implement one or more known content recognition methods that identify content by extracting parameters from a media signal and applying an algorithm to those parameters to search for a content match. Those skilled in the art will recognize that many such methods and algorithms exist. One such example is described in U.S. patent application Ser. No. 09/839,476 entitled ‘System and Methods for Recognizing Sound and Music Signals in High Noise and Distortion’ by Wang et al. and published Jun. 27, 2002 with publication number US2002/0083060. Aspects of other such examples are described in the following published PCT applications by Konin-Klijke Philips Electronics N.V.: “Fingerprinting Multimedia Contents” (WO2004/044820, published May 27, 2004), “Fingerprint Extraction” (WO2004/030341 published Apr. 8, 2004), and “Improvements in and Relating to Fingerprint Searching” (WO2004/040475 published May 13, 2004).” EX1001, 5:38-55.

B. Generating Play Streams Was Well-Known

42. The '622 patent explains that a “raw play stream” functionally relates

“a sequence of samples to content identification results” EX1001, 6:14-16, 8:1-3.

But the generation of play streams in content identification systems was well known by a POSA in the art by the time of the '622 patent. For example, O'Hagan, which was issued in 1996 (nearly two decades prior to the '622 patent), describes the generation of play streams as being known to a POSA at that time. O'Hagan describes play stream generation in relation to Figure 1.



EX1021, FIG. 1

43. O'Hagan describes the automatic identification of broadcast performances, such as music and commercials. EX1021, 1:9-12. In operation, O'Hagan describes Figure 1 as including station monitoring unit 20, which

receives radio signals and process the signals to extract audio signatures representative of the received signatures. EX1021, 5:1-22. The received signals pass through audio processing circuitry 28 to output an audio signature. EX1021, 5:1-22. The extracted audio signatures are stored along with their timestamps in a storage medium 34. EX1021, 5:30-36. The extracted audio signatures are then identified by comparison to a library of reference signatures. EX1021, 5:40-51.

44. In particular, O'Hagan utilizes a signature recognition computer 70 to compare extracted signatures to a library of reference signatures stored in memory 72, "searches for instances of successive recognition of extracted signatures to identify a particular program," and then classifies and stores the signatures in recognition file 74—a play stream. EX1021, 5:52-64. Further, O'Hagan discloses that the recognition file 74 is provided to a post-processing computer 76 which, in turn, produces a file or log 80 of identified programs along with their start times and durations if successive recognized extracted signatures are detected. EX1021, 5:64-6:3. Additionally, post processing computer 76 assigns a code to non-recognized or unmatched signatures for later manual review and identification. EX1021, 6:7-10.

45. O'Hagan evidences that the generation of play streams in the content identification art was well known nearly two decades before the '622 patent – in the mid-1990s.

C. Identifying a Channel Was Well-Known

46. Wang777 explains that “[i]n the field of content identification, it may be desirable to identify not only content, but also a source (e.g., channel, stream, or station) of a broadcast transmission.” EX1005, ¶[0018]; EX1007, 1. For example, “it may be desirable to detect from a free-field audio sample of a radio broadcast which radio station a user is listening to, as well as to what song the user is listening.” EX1005, ¶[0018]; EX1007, 1.

47. In particular, the audio sample being analyzed can be “from any number of analog or digital sources, such as a stereo system, television, radio broadcast, Internet streaming broadcast, or any other suitable means of transmitting such recorded material” EX1005, ¶[0027]; EX1007, 1, 4. “Depending on the source, the sample can be in the form of acoustic waves, radio waves, a digital audio PCM stream, a compressed digital audio stream (such as Dolby Digital or MP3), or an Internet streaming broadcast.” EX1005, ¶[0027], *see also id.*, ¶[0019]; EX1007, 1-2, 4.

48. By identifying a broadcast source, more commerce possibilities become available to the content provider, such as advertisement and promotional plans. EX1005, ¶[0005]. Various techniques may be utilized to identify a broadcast source of desired content, including utilizing watermarks that are embedded by a broadcast station that identifies the station, or “performing cross-correlation

analysis between an audio sample and audio feeds captured from broadcast stations. EX1005, ¶¶[0006]-[0007]; EX1007, 2-3.

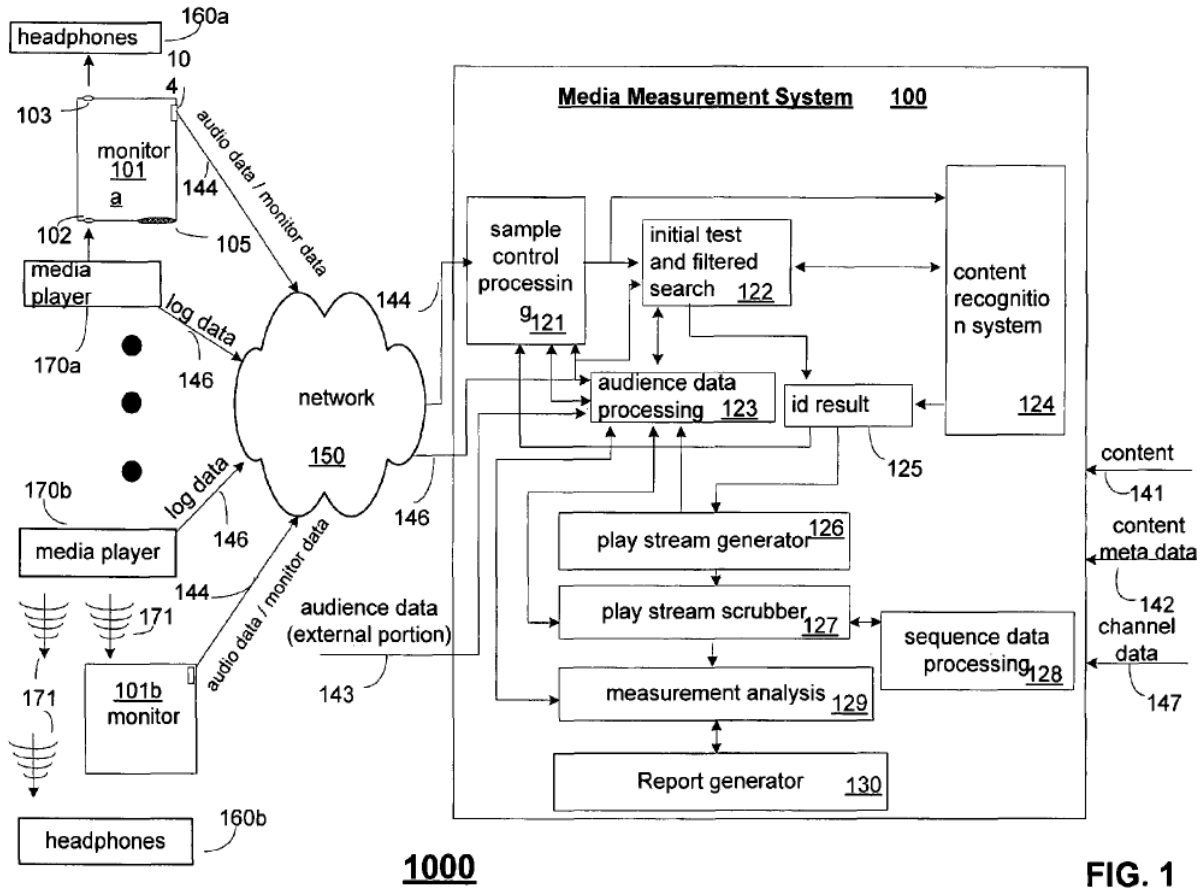
VI. OVERVIEW OF THE '622 PATENT

49. With the background discussed in Section V in mind, I now turn to the '622 patent, described below.

A. Summary of the Disclosure

50. The '622 patent purports to “provide a media measurement system and method that enhances recognition (e.g. in terms of accuracy or efficiency) of the content of a media sample by analyzing information exogenous to the sample.” EX1001, 1:60-1:64. In particular, the '622 patent purports to “analyze sequence data” to “establish channel identification.” EX1001, 2:3-5.

51. Figure 1 illustrates such a media usage monitoring and measurement system 1000 that includes a media measurement system 100 and monitoring devices 101. EX1001, 2:66-3:3, FIG. 1.



EX1001, FIG. 1.

52. Figure 2 illustrates a media measurement method 200 carried out by system 100 of Figure 1. EX1001, 6:58-62, FIG. 2. At step 201, system 100 receives audio data from a monitoring device 101 and caches a series of samples for content identification. EX1001, 6:62-65. The '622 patent explains that this audio data is uploaded to the system 100 by monitoring device 101a, which includes a microphone 105 and a media player port 102 to capture audio energy when an audience member is consuming media from media player 170a. EX1001, 3:4-43, FIG. 2.

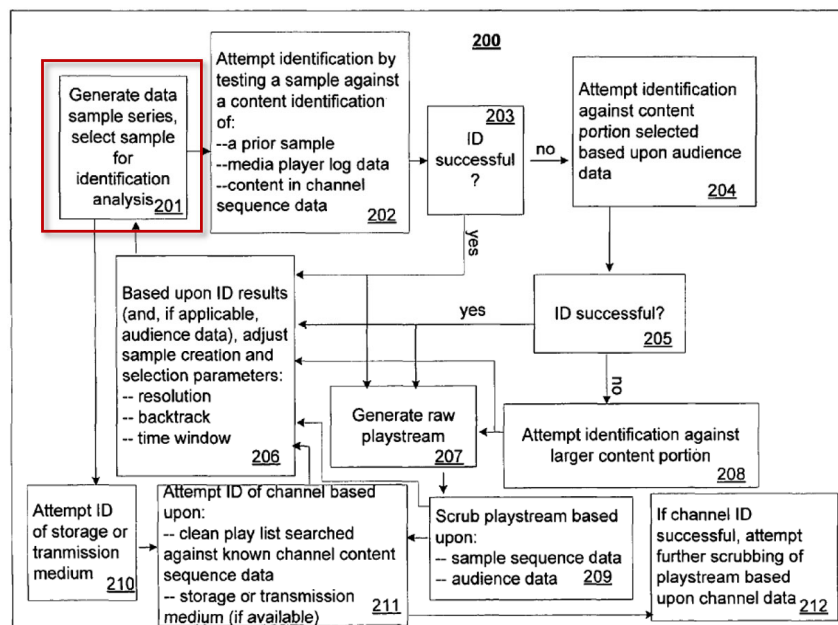


FIG. 2

EX1001, FIG. 2 (annotated).

53. At step 202, the initial test and filtered search module 122 attempts to identify a sample in the series of samples via initial test methods by determining whether the audio data in the tested sample matches data associated with the content ID corresponding to the most recently identified sample. EX1001, 4:63-5:28, 7:17-24, FIG. 2.

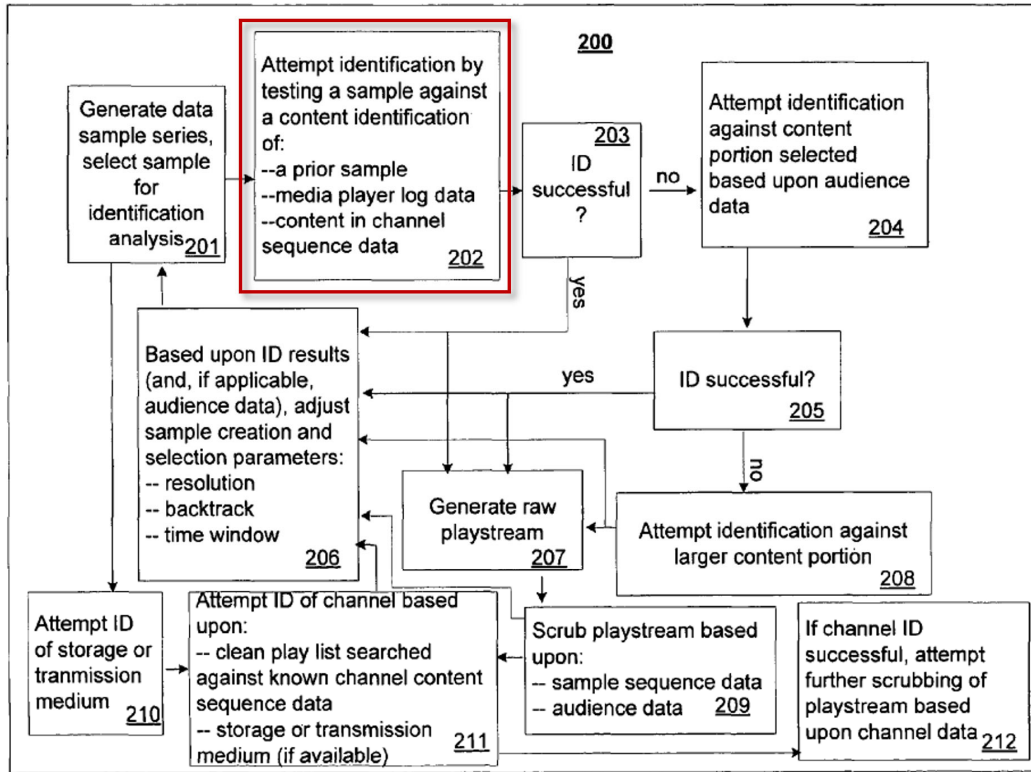


FIG. 2

EX1001, FIG. 2 (annotated).

54. If the sample is identified, the result is provided to play stream generator 126 (step 207). EX1001, 7:30-31. If not, then the initial test and filtered search module 122 may search (step 204) for a content match against targeted content selected based at least in part upon additional audience data. EX1001, 7:24-27, FIG. 2.

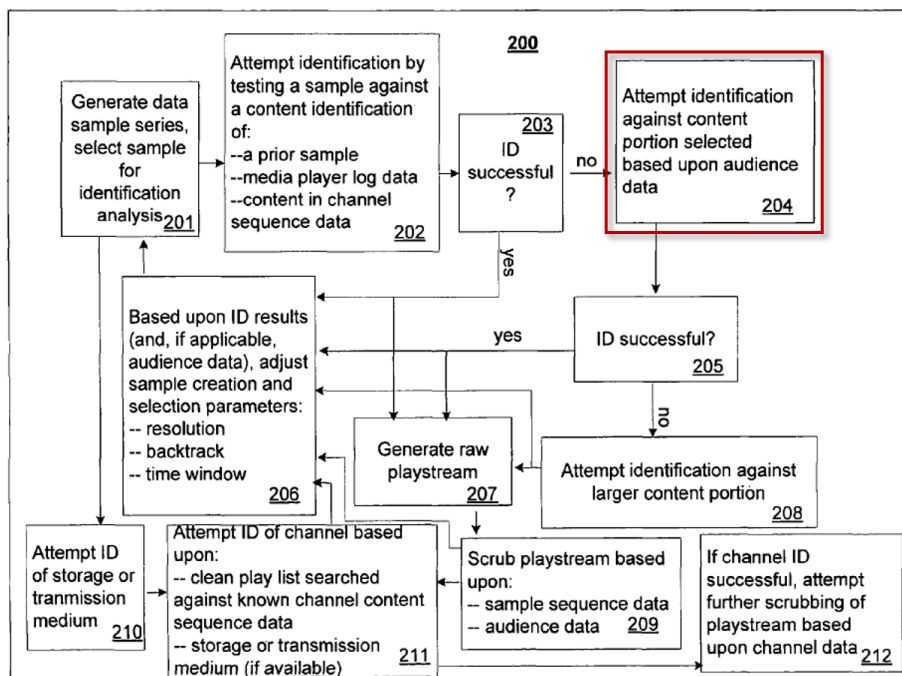


FIG. 2

EX1001, FIG. 2 (annotated).

55. If the sample is not identified at step 204, the content recognition system 124 performs a search using a larger portion of its content to identify the data sample as step 208. EX1001, 5:33-35, 7:31-35, FIG. 2.

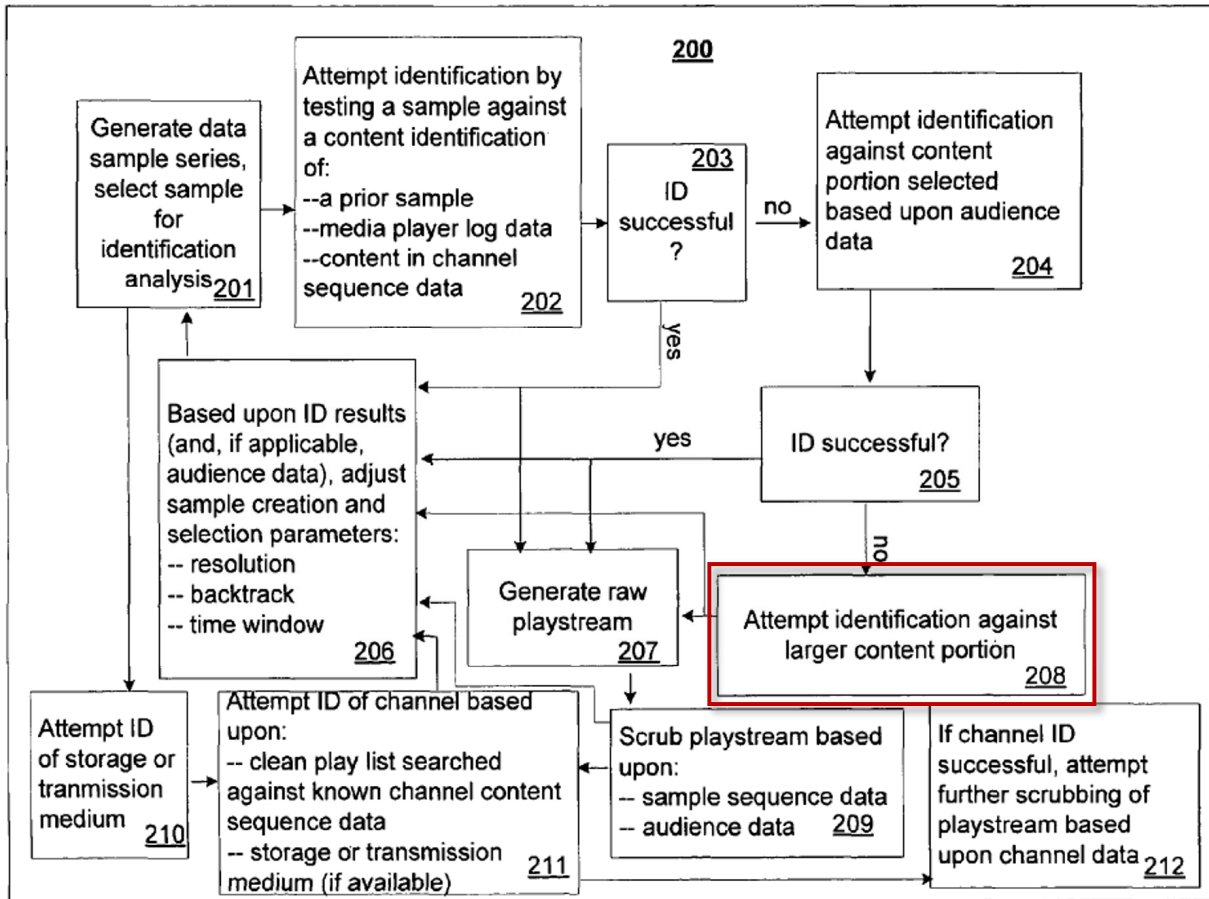


FIG. 2

EX1001, FIG. 2 (annotated).

56. The '622 patent explains, however, that the content recognition system 124 did not implement a new or novel method or technique of identifying content but instead simply “implements one or more known content recognition methods that identify content by extracting parameters from a media signal and applying an algorithm to those parameters to search for a content match.” EX1001, 5:38-62. Indeed, the '622 patent explains that:

Those skilled in the art will recognize that many such methods and

algorithms exist. One such example is described in U.S. patent application Ser. No. 09/839,476 entitled “System and Methods for Recognizing Sound and Music Signals in High Noise and Distortion” by Wang et al. and published Jun. 27, 2002 with publication number US2002/0083060. Aspects of other such examples are described in the following published PCT applications by Koninklijke Philips Electronics N.V.: “Fingerprinting Multimedia Contents” (WO2004/044820, published May 27, 2004), “Fingerprint Extraction” (WO2004/030341 published Apr. 8, 2004), and “Improvements in and Relating to Fingerprint Searching” (WO2004/040475 published May 13, 2004). Such distinct methods and algorithms may have varying degrees of efficiency, accuracy, and applicability to certain types of content (or different settings of a particular algorithm might have different efficiency depending on content type). Thus, in some embodiments, a system such as system 100 may intelligently select among multiple such algorithms (or multiple settings of a single algorithm) based on expected content characteristics.

EX1001, 5:42-5:62.

57. At step 207, a raw play stream is generated by play stream generator 126 indicating a series of samples and corresponding identification results.

EX1001, 5:31-33, 7:57-60, FIG. 2.

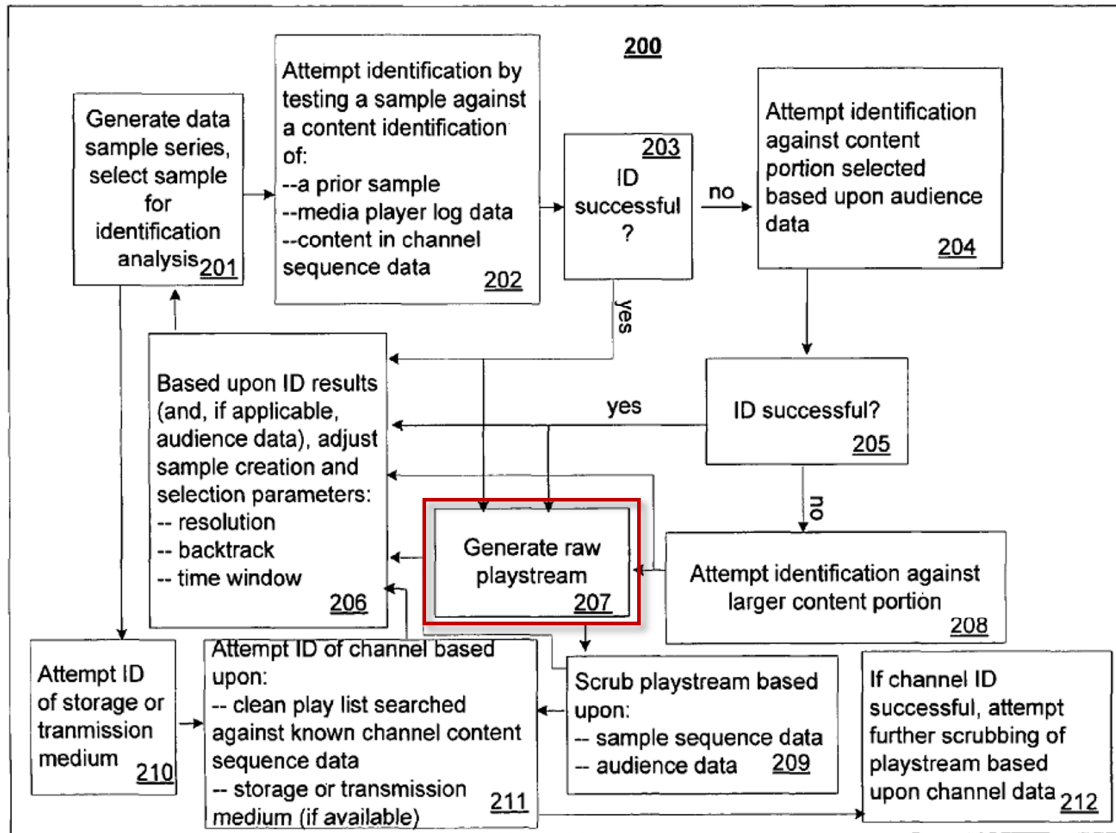


FIG. 2

EX1001, FIG. 2 (annotated).

58. Figure 4 illustrates a raw play stream 410, generated at step 207, which itemizes a sequence of analyzed samples and corresponding identification results (content ID column) where the content ID result indicates either a reference to a particular piece of content (e.g., song 51) or indicates that no identification was made (e.g., none). EX1001, 9:12-31. The '622 patent explains that the “times” in the raw play stream 410 may be relative to an arbitrary value (e.g. “t₀”) rather than representing an absolute time or indicate the actual time the sample started to

be captured. EX1001, 9:27-31. The '622 patent further explains that in the example in Figure 4, no user play-altering activities took place but that in other examples, such as that of Figure 5, sample sequence data (“ssd”) can includes content offset data. EX1001, 9:21-26.

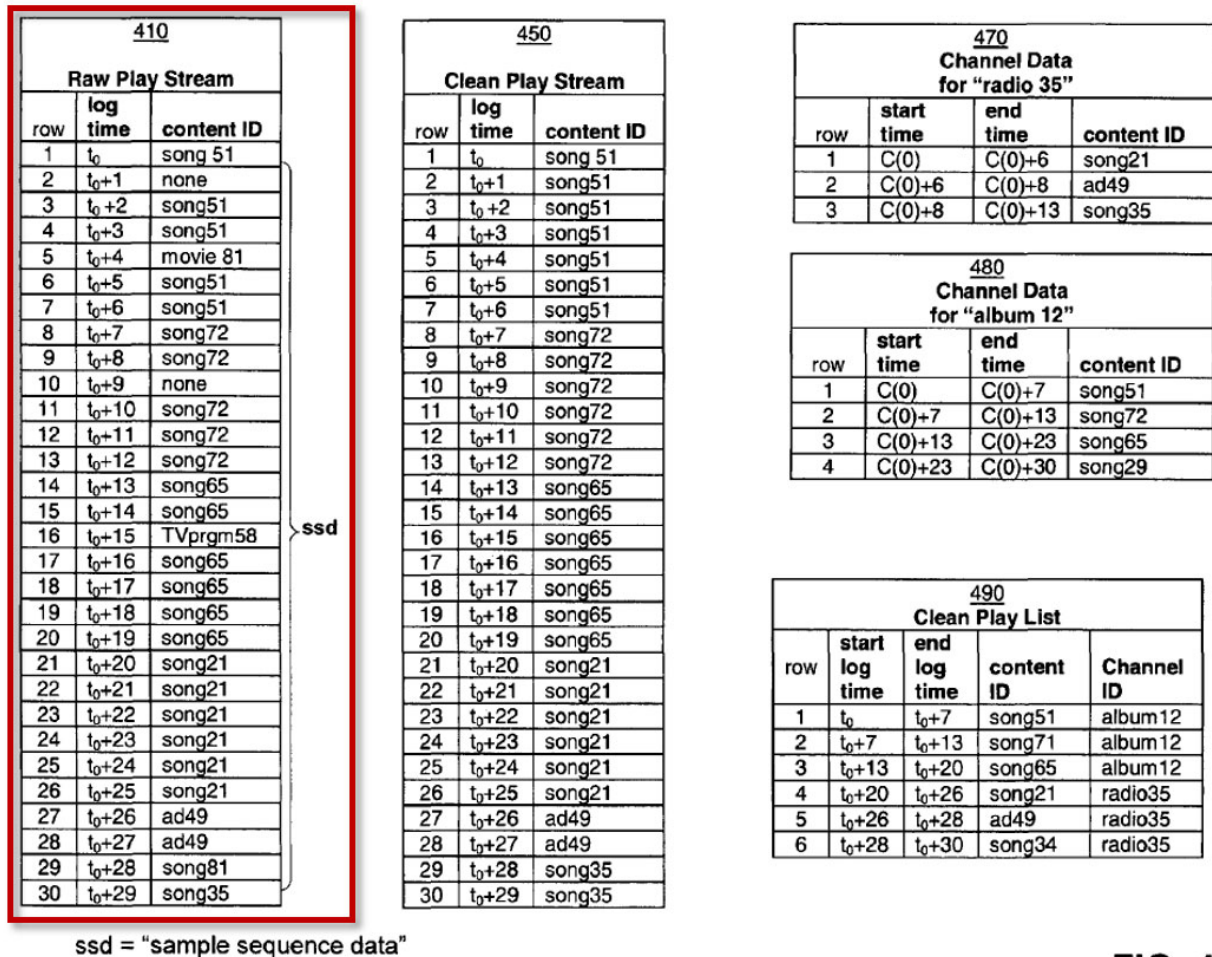


FIG. 4

EX1001, FIG. 4 (annotated).

59. At step 209, the “raw play stream” is scrubbed “based upon sample sequence data . . . and audience data” resulting in a clean play stream that may be a play list listing the content captured and indicating the order in which it was

captured. EX1001, 8:3-8. The '622 patent explains that the method then “attempts to identify the channel by searching known channel data to find an apparent content sequence match to the play list” (step 211). EX1001, 7:11-15, 8:9-11. “If step 211 successfully identifies the channel, information such as content sequence data for the identified channel may be used by step 212 to further scrub or rescrub the play stream.” EX1001, 8:17-22.

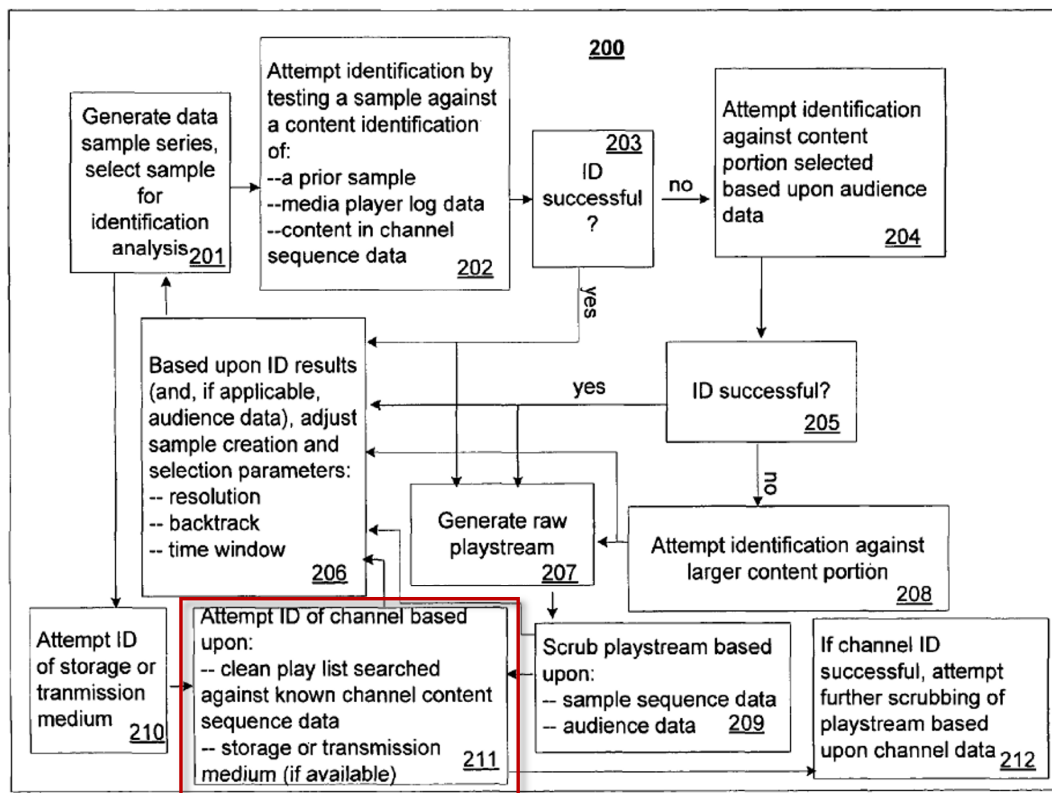


FIG. 2

EX1001, FIG. 2 (annotated).

60. Figure 3 illustrates another embodiment that includes functions similar to that of steps 207, 209, 211, and 212 of Figure 2. EX1001, 8:23-27. For

example, the '622 patent explains that process 300 includes step 303 that converts a raw play stream into “a format useful in matching to a channel (e.g. converts a play stream to a play list where a play list format is useful for channel matching, or converts a play stream or play list to a mathematical representation that is useful for channel matching).” EX1001, 8:44-49.

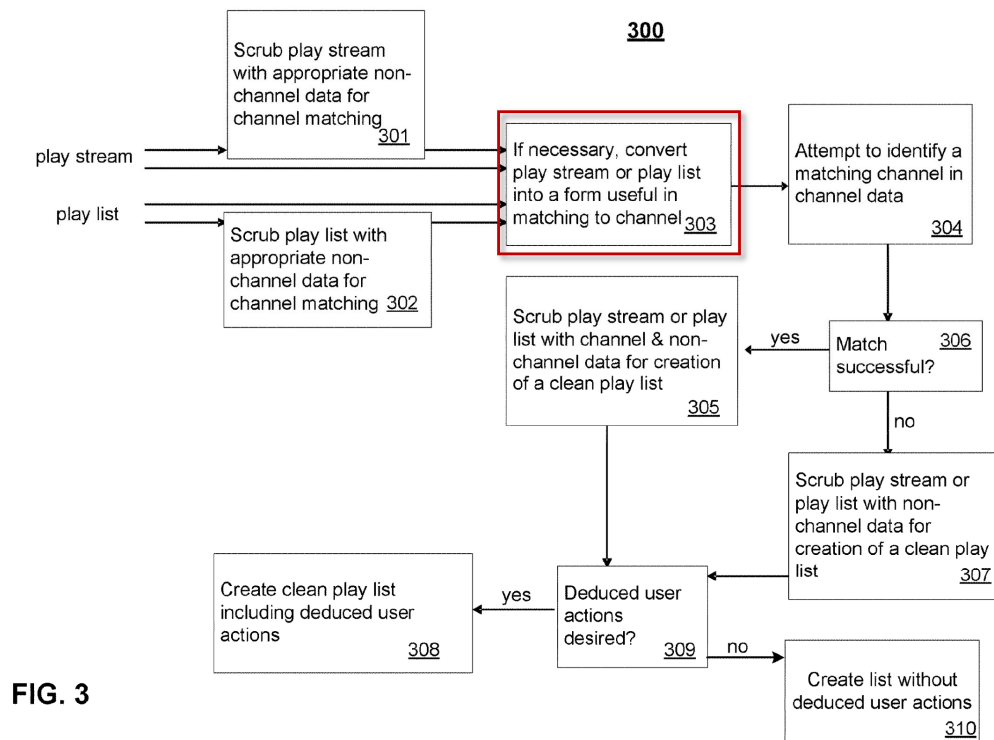


FIG. 3

EX1001, FIG. 3 (annotated).

61. Step 304 then “attempts to match the play stream or play list (or corresponding data format representing the play stream or play list) against known channel data to identify a channel (e.g., an album, samples of an album presented for marketing—e.g. on a web page—, radio broadcast, television broadcast, theater version of a movie, DVD version of a movie, etc.) associated with the elements of

the play list or play stream.” EX1001, 8:50-56. And step 308, a clean play list can be created from the raw play stream. EX1001, 9:4-11.

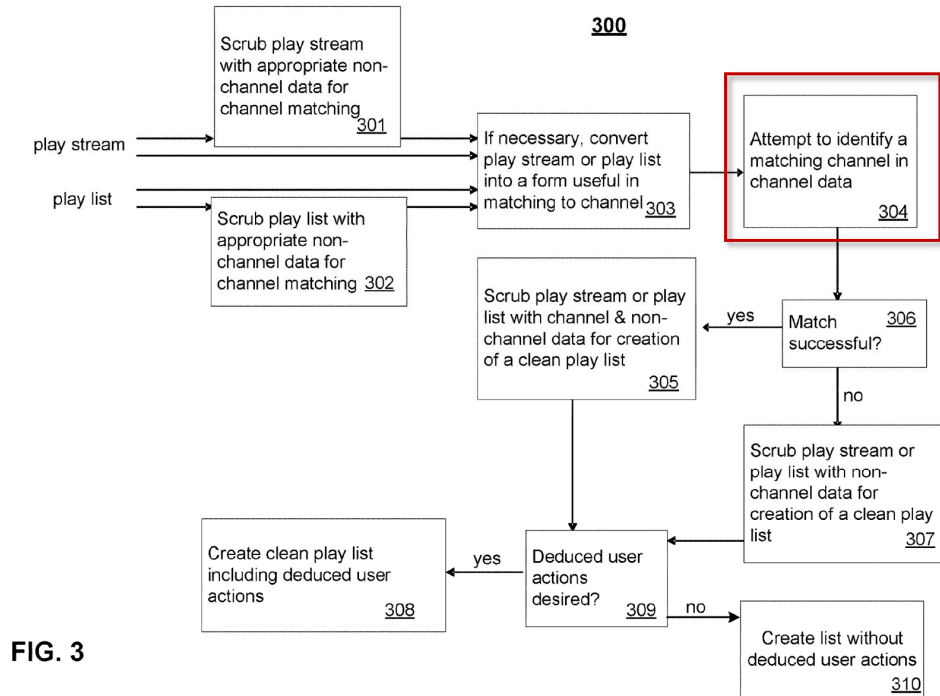
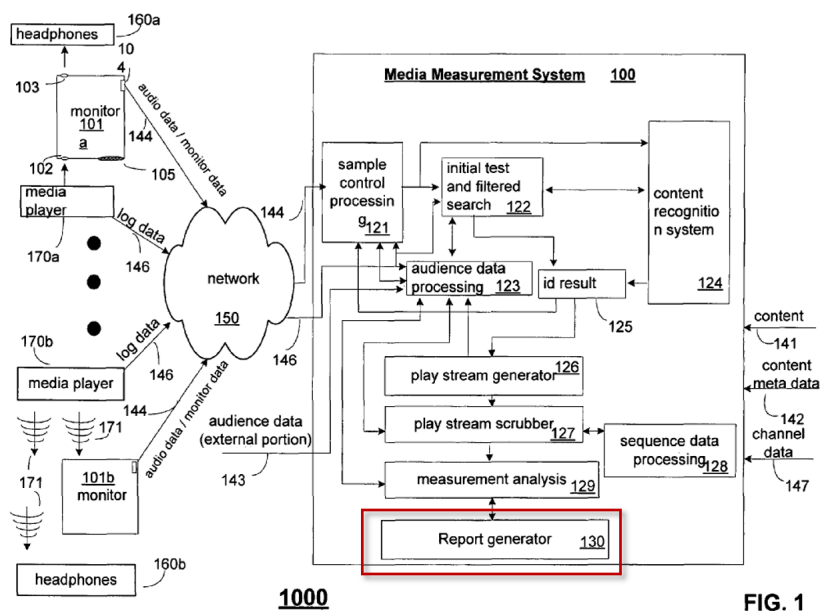


FIG. 3

EX1001, FIG. 3 (annotated).

62. Results of the media measurement analysis can then be obtained via report generator 130. EX1001, 6:13-16.



EX1001, FIG. 1 (annotated).

B. Prosecution History

63. I understand that the '622 patent claims priority to U.S. Application No. 60/574,836, filed May 27, 2004, and is a division of U.S. Application No. 11/139,330, filed May 26, 2005. EX1001, Cover.

64. I understand that on February 2, 2011, the Examiner issued a non-final Office Action rejecting the pending claims as obvious over Wang777 in view of U.S. Patent Publication No. 2002/0038456 to Hansen. EX1002, 0106-0110. On June 28th and 29th of 2011, the Applicants conducted interviews with the Examiner and in the June 30, 2011 Response, the Applicants amended the only independent claim to require using “an order of at least two different content identifiers in the play stream.” EX1002, 0137-0138. The Applicants argued, without substantive

discussion of the prior art, that “claim 1 is now allowable over the prior art of record.” EX1002, 0138.

65. I understand that on May 14, 2012, the Examiner issued a final Office Action rejecting the pending claims as obvious over Wang060 in view of U.S. Patent No. 8,131,389 to Hardwick and Hansen. EX1002, 0180-0185. On November 7, 2012, the Applicants conducted an interview with the Examiner and in the November 12, 2012 Response, the Applicants amended the only independent claim to require “wherein a content identifier of the play stream’s content identification results is obtained by using a corresponding data sample to search a computerized database of known content” and argued, without substantive discussion of the prior art, that claim 1 “is now allowable over the prior art of record.” EX1002, 0213-0216.

66. I understand that on April 11, 2013, the Examiner issued a non-final Office Action rejecting the pending claims as obvious over Wang060 in view of U.S. Publication No. 2004/0055445 to Iyoku and Hansen. EX1002, 0235-0241. On October 8, 2013, the Applicants conducted an interview with the Examiner and in the October 11, 2013 Response, the Applicants amended the only independent claim to require that “utilize[ing] an order of at least two content identifiers” is “in response to obtaining the content identifiers” and argued, again without any substantive discussion of the prior art, that claim 1 “is now allowable over the prior

art of record.” EX1002, 0248-0251.

67. I understand that on January 29, 2014 the Examiner issued a Notice of Allowance and Examiner’s amendment, which added claims 6-13 and amended claim 1 as follows:

3. The Application has been amended as follows:

Claim 1 is amended as follows:

A media measurement method comprising:

using one or more computers to generate a play stream of content identification results corresponding to a sequence of data samples, the data samples comprising representations of audio data captured at a media monitoring device, wherein content identifiers of the play stream’s content identification results are obtained by using corresponding data samples to search a computerized database of known content; and

in response to obtaining the content identifiers, using one or more computers to utilize [[an]]a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples, the at least two different obtained content identifiers identifying different media content items.

EX1002, 0269-0279.

68. While the Examiner cited Wang777 and Wang060 individually during examination, I understand that the Examiner did not rely on a combination of the teachings of Wang777 and Wang060, and did not cite many of the portions of Wang777 and Wang060 that are relied upon below.

C. Priority Date of the '622 Patent

69. I understand that the '622 patent claims priority to U.S. Provisional Application No. 60/574,836 (“the '836 application”), filed May 27, 2004.

70. The alleged invention here is that “in response to obtaining the content identifiers,” one or more computers “utilize a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples.” However, I note that the '836 application simply states “[a]nother exemplary method determines the source of transmission or storage of the media played by an individual based on audio characteristics of that signal (e.g., recognizes particular types of encoding or compression, or hears screen refresh on tv or computer, or recognizes compression or frequency range of FM or AM radio, or CD or DVD).” EX1009, 0003. Moreover, the '836 application never uses the word “channel.” In contrast, the '622 patent makes clear that a channel can be a non-broadcast collection of content, such as a CD, and that a sequential order of at least two different obtained content identifiers in the play stream are used to identify a channel (e.g., a CD). EX1001, 8:11-16, 9:55-59. It is thus my opinion that the claimed subject matter is not disclosed in the '836 application.

71. Thus, it is my opinion that the earliest priority date of the '622 patent is May 26, 2005, the filing date of aforementioned U.S. Pat. Appl. No. 11/139,330

VII. OVERVIEW OF THE PRIOR ART

A. Wang777

72. I understand that Wang777 (EX1005) was published on June 21, 2007, filed on October 31, 2006, is a National State Entry of PCT/US05/05254 (EX1008) filed on February 18, 2005 (published in English and designating the U.S.), and claims priority to U.S. Provisional Application No. 60/545,879 (EX1007), filed February 19, 2004.

73. I understand that a non-provisional patent application is prior art as of the earlier filing date of a provisional application if the disclosure of the provisional provides support for at least one claim in the publication, and the provisional describes the subject matter relied upon as prior art.

74. It is my opinion that the '879 provisional application provides support for at least Wang777's claim 1 and describes the subject matter relied upon as prior art, as shown below in the claim-by-claim analysis. *See infra* Sections IX-X.

Wang777's Claim 1	Wang777's '879 Provisional Appl.
1. A method for identifying a broadcast source of content comprising:	"Method and Apparatus for Identification of Broadcast Source" EX1007, Title.
recording an audio sample;	"In one embodiment of the system and methods described herein, a user has an audio sampling device containing a microphone and optional data transmission means. The user hears an audio program being broadcast from some broadcast means, such as radio or television. He then records a sample

	using the audio sampling device.” EX1007, 2.¹ “A user audio sample collected by the user is identified using a content identification means such as the one described by Wang and Smith for identifying an audio sample out of a database of audio content files (such as songs).” EX1007, 5.
recording a time at which the audio sample was recorded;	“One technique is to identify a radio station by performing a timestamped recording of a radio broadcast channel, converted into a fingerprint stream. In the field, a user collects a sample, where the sample time is also timestamped in terms of ‘realtime’ offset from a common timebase.” EX1007, 4. “While the user audio sample is collected, a user sample timestamp (UST) is taken to mark the beginning time of the audio sample based on a standard reference clock.” EX1007, 5.
identifying characteristics of the audio sample and an estimated time offset of the audio sample, the estimated time offset defining from a time difference between a start time of the audio sample and the time at which the audio sample was recorded;	“The technique invented by Wang and Smith, (which is described in International Publication Number WO 02/11123 A2, entitled System and Methods for Recognizing Sound and Music Signals in High Noise and Distortion and claiming priority to US Provisional Application No. 60/222,023 filed July 31, 2000 and US Application Serial No. 09/839,476 filed April 20, 2001, the entire contents of

¹ Emphasis added throughout unless otherwise noted.

	each of which are incorporated herein by reference (hereinafter ‘Wang and Smith’)), can be used to identify free-field audio samples of music playing from various sources, such as radio and television.” EX1007, 2. “Method 1 Use spectrogram peaks, correlate spectrogram peaks rather than direct signal. Use ‘combinatorial hash’ + peak verification” EX1007, 3. “A recognition is performed using the technique of Wang and Smith to generate an estimated time offset of the sample within the ‘original’ recording. The absolute times are calculated and compared.” EX1007, 4.
comparing the characteristics and the estimated time offset of the audio sample with characteristics and time offsets of samples taken from broadcast stations and taken at approximately the time at which the audio sample was recorded; and	“The technique invented by Wang and Smith, (which is described in International Publication Number WO 02/11123 A2, entitled System and Methods for Recognizing Sound and Music Signals in High Noise and Distortion and claiming priority to US Provisional Application No. 60/222,023 filed July 31, 2000 and US Application Serial No. 09/839,476 filed April 20, 2001, the entire contents of each of which are incorporated herein by reference (hereinafter ‘Wang and Smith’)), can be used to identify free-field audio samples of music playing from various sources, such as radio and television.” EX1007, 2. “A recognition is performed using the technique of Wang and Smith to

	generate an estimated time offset of the sample within the ‘original’ recording. The absolute times are calculated and compared. If the realtime offsets are within a certain tolerance, say 1 second, then the identification is considered to be originating from the same source, as the probability that a random performance of the same audio content (such as a hit song) is so synchronized in time is extremely low.” EX1007, 4; <i>see also</i> EX1007, 6 (describing comparison of USRTO and BSRTTO offsets). “A further step of verification may be used in which spectrogram peaks may be aligned. Because the Wang/Smith technique generates a relative time offset, it is possible to temporally align the spectrogram peak records within about 10 ms in the time axis. Then we can just count the number of matching time+frequency peaks. That is then the score.” EX1007, 5.
based on the comparison, identifying a broadcast station from which the audio sample was broadcast.	“One technique is to identify a radio station by performing a timestamped recording of a radio broadcast channel, converted into a fingerprint stream.” EX1007, 4. “This may be done by direct comparison of the fingerprint streams (from broadcast channel and from the user) and the relative time offsets. If the relative offset is near zero then it is likely that the streams are being monitored from the same source. Longer and random time delays could mean that the user is listening to an

	independent but coincident copy of the same audio program.” EX1007, 4.
--	--

1. Overview of Wang

75. Wang777 explains that “[i]n the field of content identification, it may be desirable to identify not only content, but also a source (e.g., channel, stream, or station) of a broadcast transmission. For example, it may be desirable to detect from a free-field audio sample of a radio broadcast which radio station a user is listening to, as well as to what song the user is listening.” EX1005, ¶[0018].

76. Wang777 is thus directed to “a method for identifying a broadcast source of an audio sample” and, in particular, “matching audio or media file samples to a broadcast source from which the sample was transmitted.” EX1005, ¶¶[0002]-[0003], [0017]; EX1007, 2 (stating that an audio sample is recorded and analyzed to determine which broadcast station a user is listening to).

77. Figure 1 illustrates Wang777’s “system for identifying a broadcast source of desired content.” EX1005, ¶[0020]. The system includes an audio sampling device 102 (e.g., a mobile phone, blue) used to “record an audio sample broadcast or transmitted by a broadcaster 104” (e.g., a radio or television content provider, red). EX1005, ¶[0020].

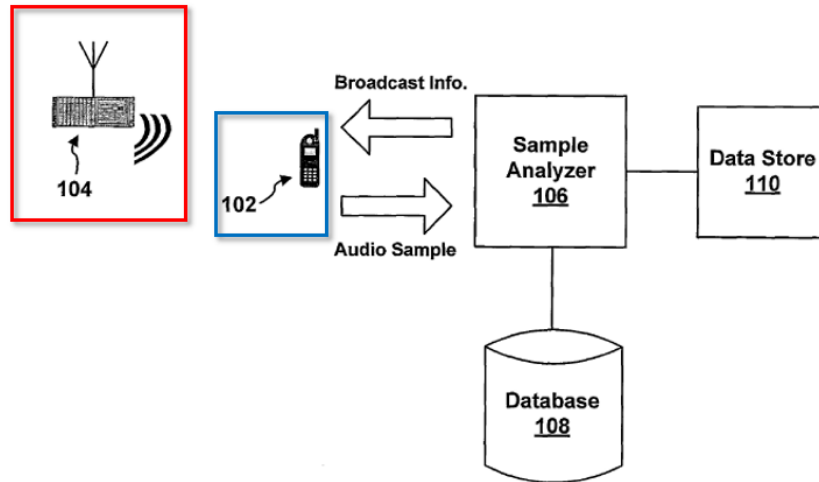


Figure 1

EX1005, FIG. 1 (annotated).

78. Figure 4 illustrates “a method of identifying a broadcast source” that includes identifying the content of a media sample (402) and ultimately identifying a broadcast source (422). EX1005, ¶¶[0016], [0047]-[0056], FIG. 4; EX1007, 2-3, FIG. 2 (stating that watermarking or cross-correlation may be used to identify a broadcast station that a user is listening to).

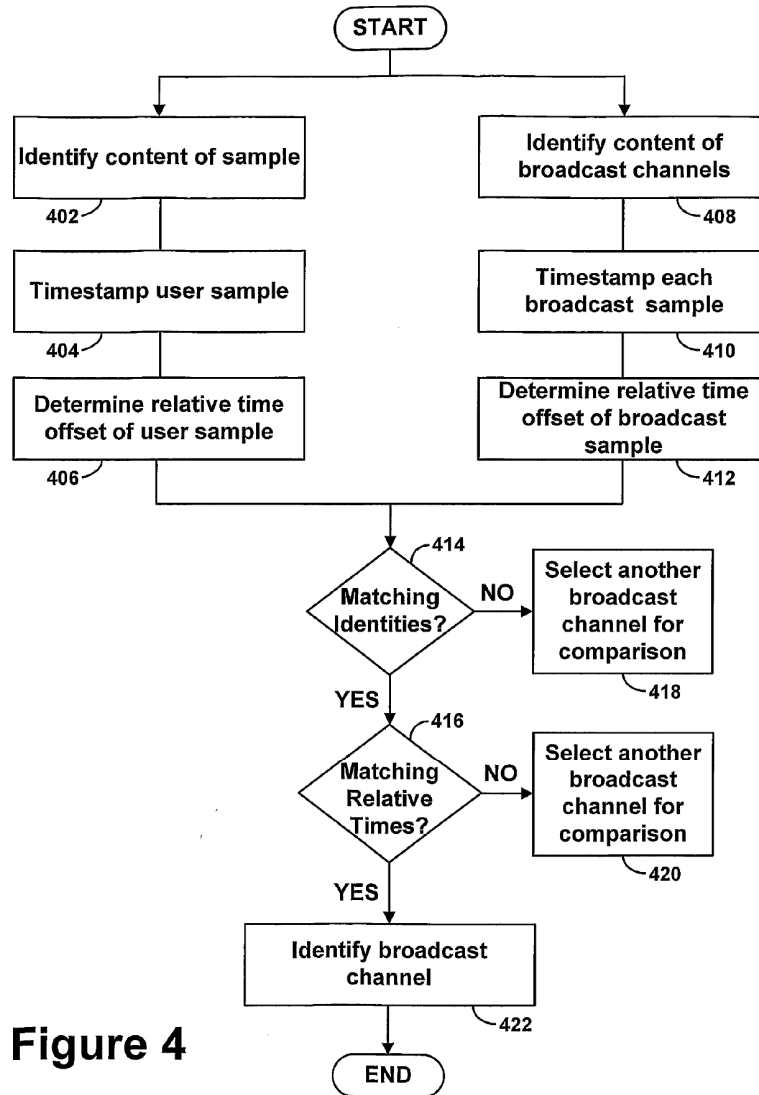


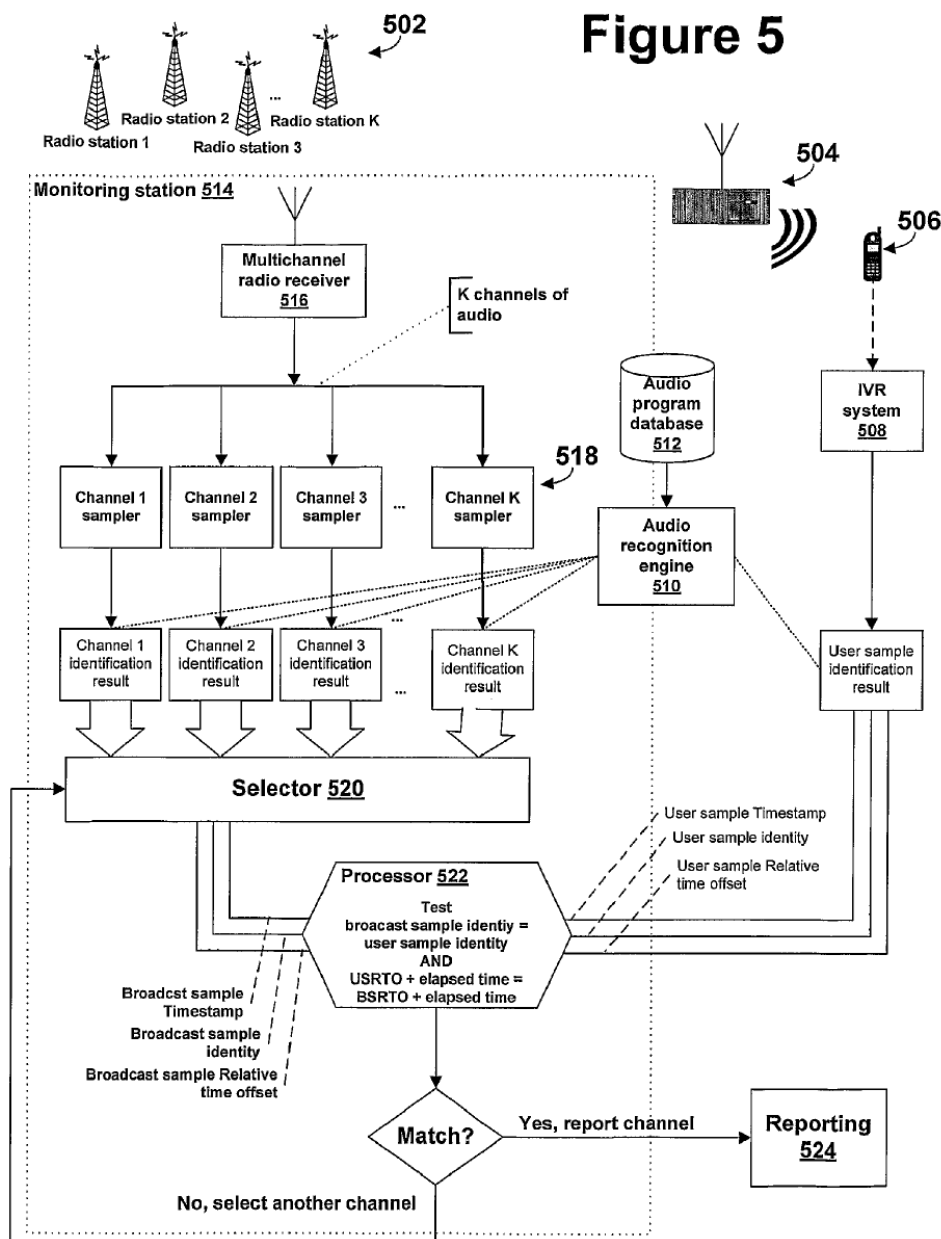
Figure 4

EX1005, FIG. 4

79. Figure 5 “illustrates one example of a system for identifying a broadcast source of an audio sample according to the method illustrated in FIG. 4.”

EX1005, ¶[0057], FIG. 5.

Figure 5



EX1005, FIG. 5.

B. Overview of Wang060

80. Wang060 is directed to a “method for recognizing an audio sample locates an audio file that most closely matches the audio sample from a database indexing a large set of original recordings.” EX1006, Abstract. Figure 1 illustrates

a method for recognizing a sound sample:

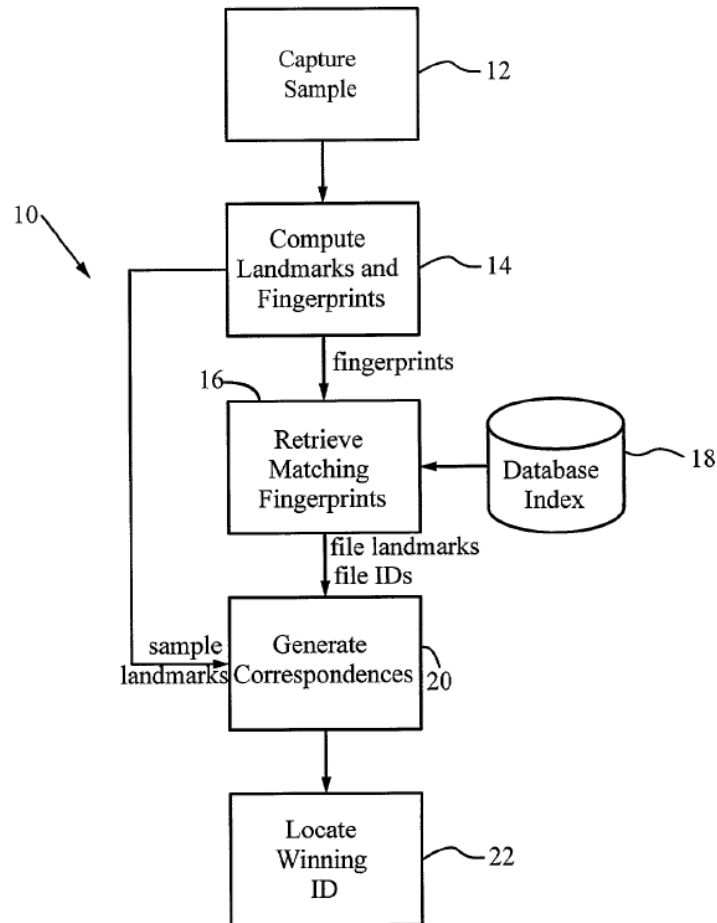


FIG. 1

EX1006, FIG. 1.

EX1006, ¶¶[0026], [0039]-[0042];

81. Wang060 explains that after the content is identified, a report for the identified sound samples, such as a log file of information about the identified sound recordings (e.g., title, artist, album cover art, lyrics, or purchase information), as well as “a list of music played from any sound stream (radio, internet streaming radio, hidden microphone, telephone call, etc.)” can be created.

EX1006, ¶[0107].

C. Overview of Herley682

82. Herley682 is directed to “a system and method for sampling one or more channels of a broadcast spectrum, such as a radio frequency spectrum, identifying and storing content information for each sampled channel, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1011, ¶[0002].

83. Figure 2 of Herley682, reproduced below, illustrates an exemplary architectural diagram of an implementation of an embodiment thereof. EX1011, ¶[0038].

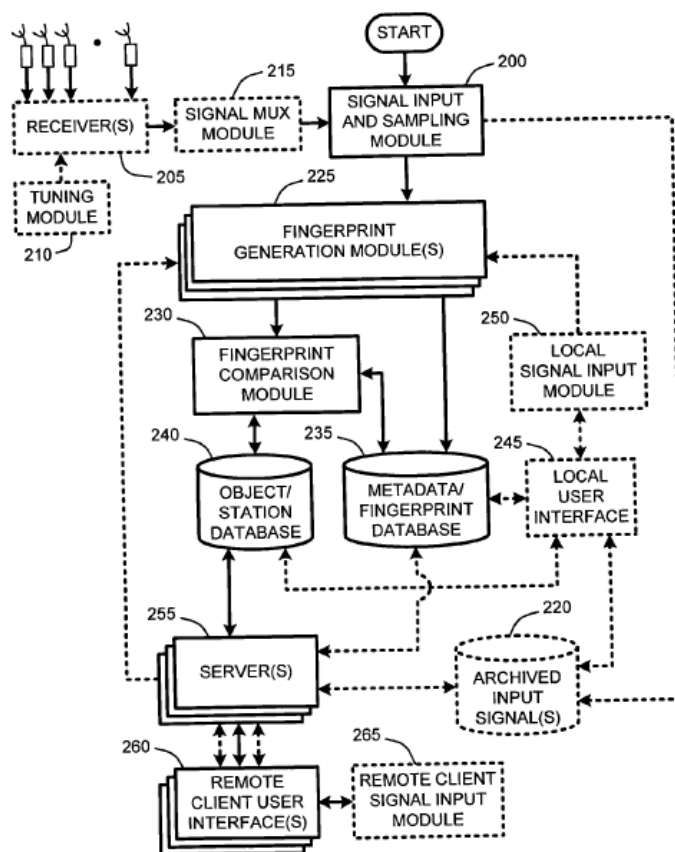


FIG. 2

EX1011, FIG. 2

84. “[S]ampling module 200 . . . receive[s] and sample[s] one or more signals, such as, for example, audio broadcast signals, video broadcast signals, etc.” EX1011, ¶[0075]. “[S]ignal samples are continuously provided in real-time to one or more fingerprint generation modules 225 for extracting fingerprints from samples of the input signals.” EX1011, ¶[0084].

85. Extracted fingerprints are provided to fingerprint comparison module 230, which searches a metadata/fingerprint database 235 to determine if there is a match to fingerprints stored in database 235. EX1011, ¶[0086]. If a match is

identified, then the sampled portion of the input signal from which the fingerprint was extracted is identified as belonging to an object of interest corresponding to the fingerprint and metadata in the metadata/fingerprint database 235. EX1011, ¶[0086]. “This information, including the metadata, is stored in an object/station database 240.” EX1011, ¶[0086].

86. The object/station database 240 includes statistical information (e.g., the time that the object was identified, the station, channel, frequency, etc., where the input signal was monitored) and any metadata for each identified object of interest. EX1011, ¶[0086]. “Metadata can include any identifying information that it . . . user desires to have associated with particular objects or object types.” EX1011, ¶[0068]. For a song, the metadata may include, for example, the song’s name, artist, album title, copyright year, length, music genre, etc. EX1011, ¶[0026], [0028], [0068].

87. “[I]f the fingerprint extracted from the sampled input signal does not match any entries in the metadata/fingerprint database 235, that fingerprint is stored to the metadata/fingerprint database [235] as an ‘unknown object’ entry.” EX1011, ¶[0088]. A copy of the sample and any statistical information available for the sample also are stored in database 235. EX1011, ¶[0088]. Metadata/fingerprint database 235 allows for unknown objects to be manually identified by a user. EX1011, ¶[0089].

D. Seet

1. Seet is prior art as of February 17, 2004

88. I understand that Seet is prior art as of the February 17, 2004 filing date of U.S. provisional patent application 60/545,681 (EX1016, “Seet provisional”) to which it claims priority.

89. I also understand that a nonprovisional patent application publication is prior art as of the earlier filing date of a provisional application if the disclosure of the provisional provides support for at least one claim in the publication, and the provisional describes the subject matter relied upon as prior art.

90. It is my opinion that the Seet provisional provides support for at least Seet’s claim 9, as shown below in the claim-by-claim analysis.

Seet Claim 9	Seet ’681 Provisional Appl. (EX1016)
9. A method for generating entries for a media content database, the method comprising:	“Media Content Database Creation and Maintenance Using Time-Delayed Auto-Correlation of Frequency Characteristics,” EX1016, 4, Title. “The invention described herein attempts to solve the issue of content acquisition by utilizing the signal itself as the data source, and discusses a method by which this database can be created in an automated fashion.” EX1016, 4, Background.
receiving a first media stream;	“Instead of using a high-quality original master recording (say a CD, or digital video tape, etc), we suggest using the same audio signal that is to

	be identified, but from a previous period. In the case of radio or television monitoring, we propose that the previous period (i.e. day or day-period) for that particular station (or genre, or group of similar stations), be fingerprinted, and inserted into a fingerprint database.” EX1016, 4. “DB of Previous Day’s Samples,” EX1016, 5, FIG. 1, EX1016, 6, FIG. 2.
receiving a second media stream;	“Instead of using a high-quality original master recording (say a CD, or digital video tape, etc), we suggest using the same audio signal that is to be identified, but from a previous period. In the case of radio or television monitoring, we propose that the previous period (i.e. day or day-period) for that particular station (or genre, or group of similar stations), be fingerprinted, and inserted into a fingerprint database.” EX1016, 5. “During the subsequent period, matches are then performed against this database of previous samples, and if matched, then the identifier (or identifiers if multiple matches are found) and a time stamp are recorded.” EX1016, 2. “Next Day’s Samples’ Results Are Matched,” EX1016, 5, FIG. 1. “Next Day’s Samples’ Results,” EX1016, 6, FIG. 2.
matching at least some of the second media stream to the first media stream to determine repeating content between	“During the subsequent period, matches are then performed against this database of previous samples, and if

the streams; and	matched, then the identifier (or identifiers if multiple matches are found) and a time stamp are recorded.” EX1016, 5. “Now the relationship has been established between the previous period’s database and the current period’s samples, it can be asserted that portions of the previous database that were matched are repeated content like advertisements, motions, songs, or the like.” EX1016, 5. “Next Day’s Samples’ Results Are Matched,” EX1016, 5, FIG. 1. “Matched content to be reviewed by human to add content description,” EX1016, 5, FIG. 1.
presenting the repeating content to a reviewer for identification.	“Now a human reviewer can manually listen to just the series of identified samples and add metadata (like add name, or artist, and title information) to the series of samples.” EX1016, 5. “Matched content to be reviewed by human to add content description,” EX1016, 5, Figure 1. “Finally, an elegant user interface could present the sets of matched samples to the user in such a way as the user can just listen to as much of the sequence of samples as he or she likes, and enter in the content meta-data before jumping to the next sequence.” EX1016, 6.

91. It is my opinion that the Seet provisional also describes the subject

matter relied upon as prior art. In particular, Figure 2, reproduced below, illustrates the “scrubbing” aspects relied upon.

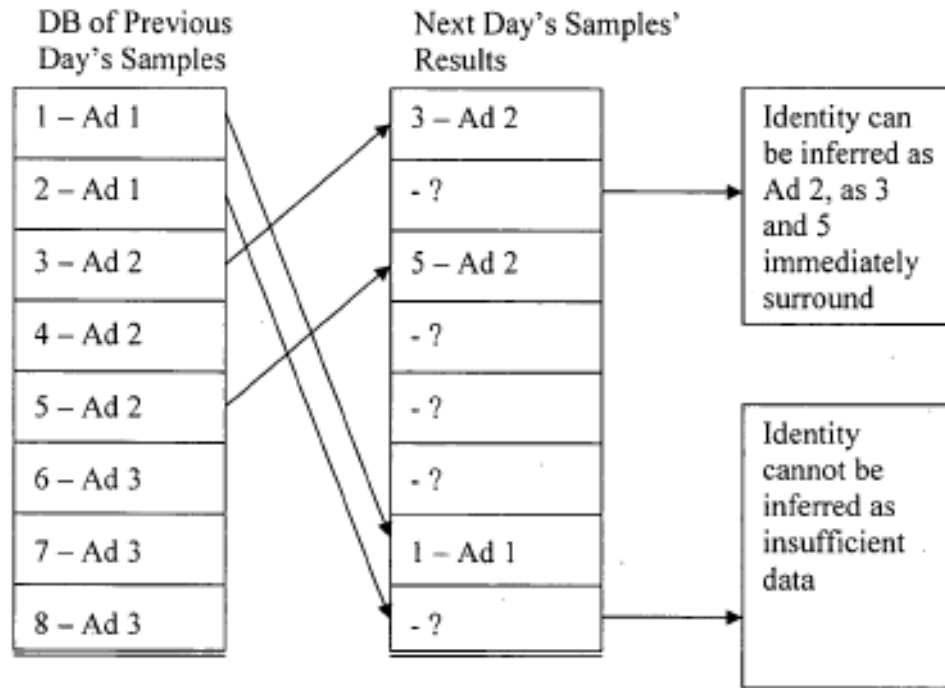


Figure 2. Auto-correlation Identity Inference Example

EX1016, FIG. 2.

92. The Seet provisional describes “[i]dentified content is left in the database for all future days until a set expiration date is reached since the last identification, and any samples from the current day that did not match are now added into the database for the next Day’s identification; the cycle repeats.”

EX1016, 5. Thus, the Seet provisional makes clear that unidentified samples are stored in the database just as the identified data are stored.

93. Continuing, the Seet provisional describes “[a]dditionally, should the

database itself contain duplicates, the post-processing algorithm should include measures to group matches into sets correspond the duplicate sample sets.”

EX1016, 6. “This requires that the underlying fingerprinting software returns all matches, not just the first match.” EX1016, 6. Further, “if some matches are not made, but surrounding samples are in the correct sequence and timing to infer the identity of the missing sample, that sample can thus be identified.” EX1016, 6.

2. Overview

94. Seet is directed to “Generation of a Media Content Database by Correlating Repeating Media Content in Media Streams.” EX1012, Title.

95. Seet’s Figure 1, reproduced below, illustrates a system for creating and maintaining a media content database. EX1012, ¶15.

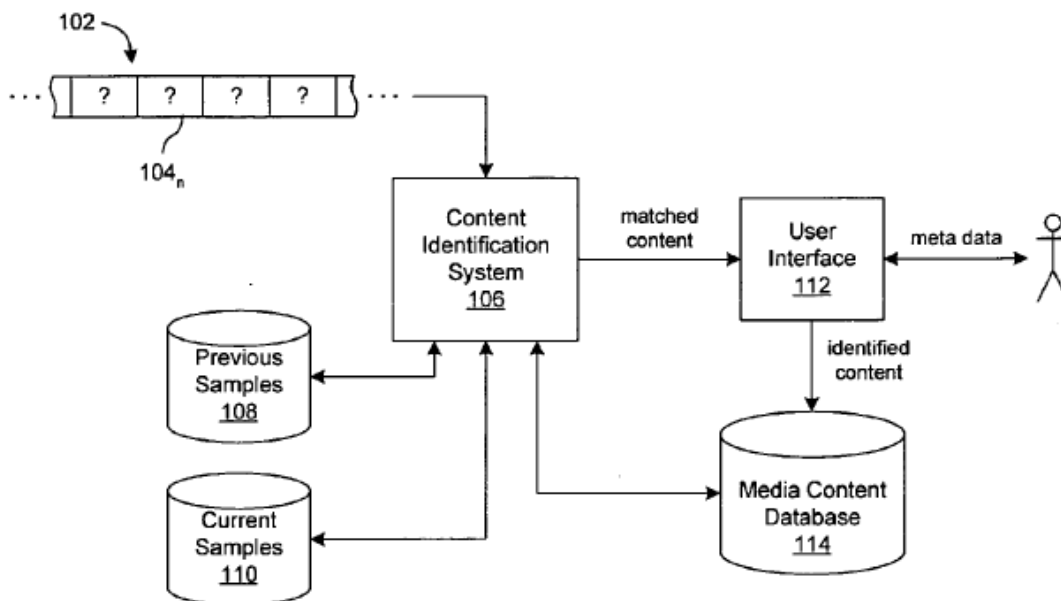


FIG. 1

EX1012, FIG. 1.

96. Content identification system 106 samples portions of media items 104_n from an incoming media data stream 102 to extract characteristic information and store that information in the previous samples database 108. EX1012, ¶19. The content identification system 106 attempts to match samples in the current samples database 110 with samples in the previous samples database 108. EX1012, ¶19. Matched samples are made accessible for review through a user interface 112, allowing a reviewer to identify meta-data for matched content. EX1012, ¶19. The identified media content is then added to the media content database 114, where it can be used by the content identification system 106 at a later time in an application that requires media content recognition. EX1012, ABSTRACT, ¶¶14, 19.

97. Figure 2, reproduced below, shows Seet's general process for creating and maintaining a media content database 114.

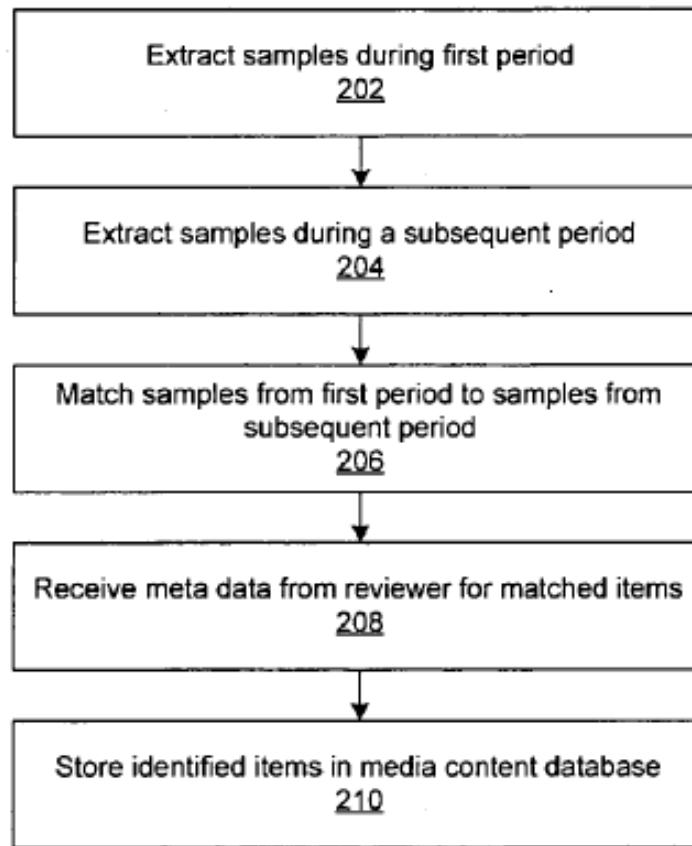


FIG. 2

EX1012, FIG. 2.

98. A set of samples is extracted from a first period of media stream 102 (step 202). EX1012, ¶23. Samples 104_n (unknown media) are subsequently extracted (step 204). EX1012, ¶23. The system then attempts to match (step 206) the second set of samples to the first set of samples. EX1012, ¶24. Matched content is then provided to a reviewer, from which meta-data are received (step 208). EX1012, ¶24. This metadata and characteristic information of associated samples

are stored (step 210) in the media content database 114. EX1012, ¶24. The characteristic information (i.e., the fingerprint) and a marker (such as a timestamp) are associated with each other. EX1012, ¶27.

99. Because the samples are typically obtained in noisy conditions, they may not match exactly (or within a predetermined tolerance) even where they are of the same content. EX1012, ¶31. Figure 4, reproduced below, illustrates a process for inferring a match for an unmatched sample. EX1012, ¶18.

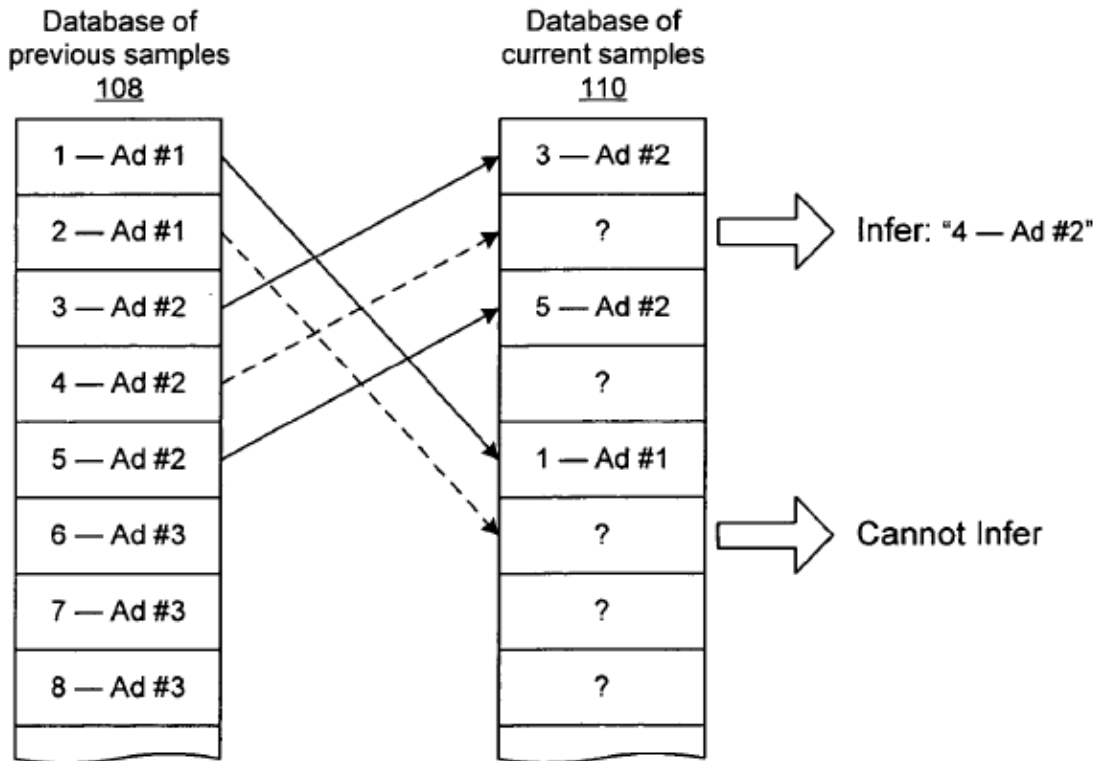


FIG. 4

EX1012. FIG. 4.

100. System 106 may infer additional matches, for example based on a

sample's temporal position in media stream 102 relative to other matched samples. EX1012, ¶32. “[I]f a match is not made for a particular sample, its surrounding samples may still be in the correct sequence and timing so that a match for the unmatched sample can be made.” EX1012, ¶32. In Figure 4, the second sample in the current samples database 110 was not matched to any sample in the previous samples database 108. EX1012, ¶32. However, because its previous sample was matched to sample 3 and its following sample was matched to sample 5, the system can infer that the second sample matches sample 4 of the previous samples database. EX1012, ¶32. The temporal consistency of the bounding samples allows one to infer that they are of the same media item 104_n (in this case, an advertisement labeled #2); therefore, the middle sample is also of the same media item 104_n. EX1012, ¶32.

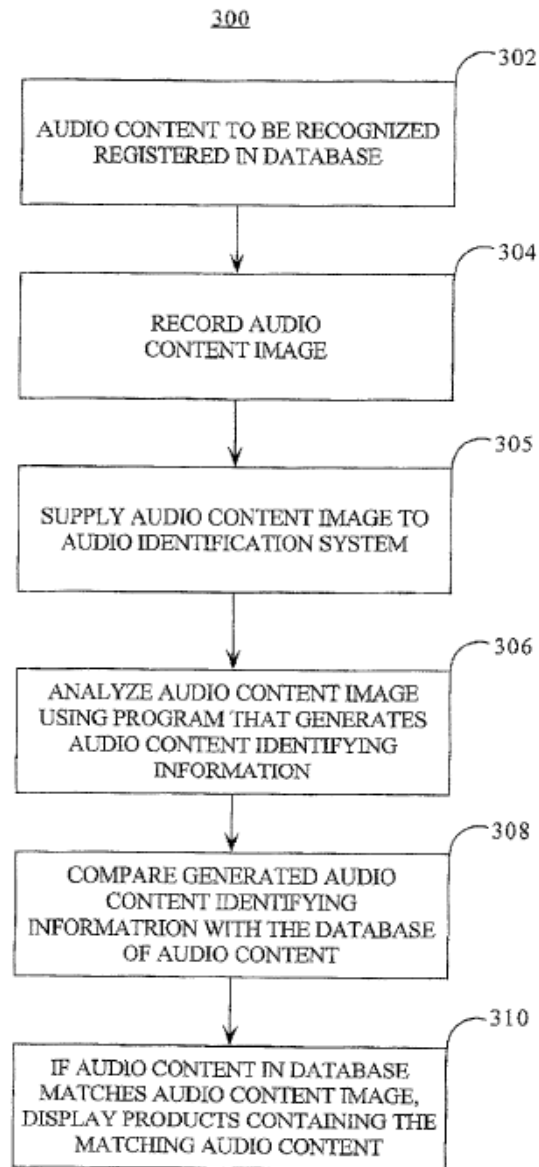
E. Overview of Pitman

101. Pitman is directed to a method of identifying audio information generated for audio content that matches audio identifying information in an audio content database, and providing information to a user if the audio identifying information matches. EX1013, Abstract.

102. Figure 3 illustrates a process 300 to automatically identify audio content that includes the steps of: recording content (304), generating a signature/fingerprint (306), comparing the signature/fingerprint to

signatures/fingerprints in a database of known content (308), and, if the content is matched, identifying the matching content (e.g., the artist and title of a song) along with providing one or more products containing (or relating to) the matching content to a user. EX1013, ¶¶[0030]-[0036].

FIG. 3



EX1013, FIG. 3.

F. Overview of Schrempp

103. Schrempp is directed to a method and system for playlist generation.

EX1014, Abstract. Figure 1 illustrates a playlist generation system 100 that includes a detection module 102 for receiving content (e.g., a radio broadcast being listened to) and a lookup module 106 that identifies the received content by comparing the audio against a database 108 of known reference signatures.

EX1014, ¶[0007], FIG. 1. If a match is made, the module 102 will keep a record of all matches made during a predetermined period of time on the database 108 (e.g., a record of song titles detected during a 24-hour period). EX1014, ¶[0008].

Schrempp further explains that a playlist server 110 may receive information (e.g., titles of identified songs) from the detection module 102 and generate a playlist of the identified content. EX1014, ¶¶[0009]-[0010].

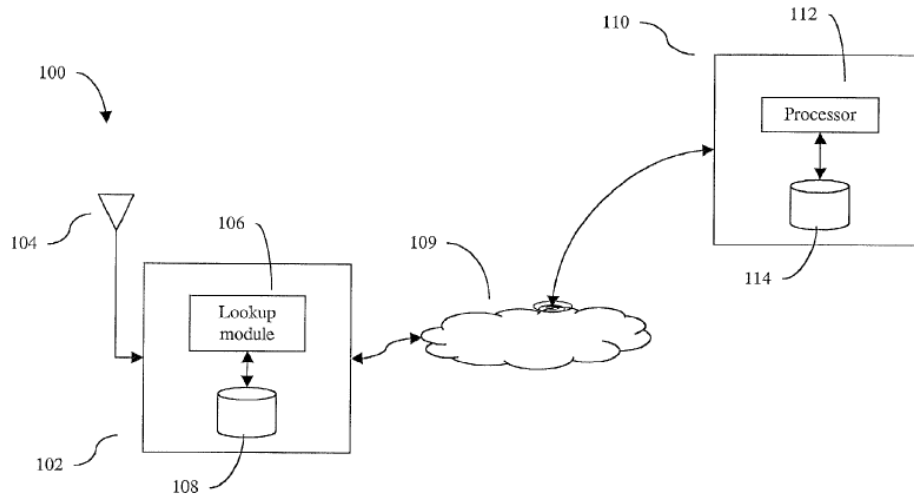


FIG. 1
Prior Art

EX1014, FIG. 1.

G. Overview of Wells

104. Wells is “directed to the identification of recordings and, more particularly, to the identification of sound recordings, such as recordings of music or spoken words.” EX1015, ¶[0003]. In particular, Wells discloses methods to identify streams of music, such as a radio broadcast, by extracting fingerprints at regular intervals and searching a database for matches. EX1015, ¶[0249]. Wells further explains that using multiple fingerprints will reduce the likelihood of mismatch errors that occur with the use of just one fingerprint. EX1015, ¶[0251].

105. In particular, Wells imposes a condition that multiple (consecutive) fingerprints must be matched to identify the stream as shown in Figure 12, which

includes the steps of: extracting a fingerprint every frame (e.g., a 3 second frame) of a stream (1602), comparing the extracted fingerprints to a database (1603), and if one or more fingerprints match, determining whether a number of consecutive candidate fingerprints (e.g., 10) match with the same song in the database (1606).

EX1015, ¶¶[0259]-[0265], FIG. 12.

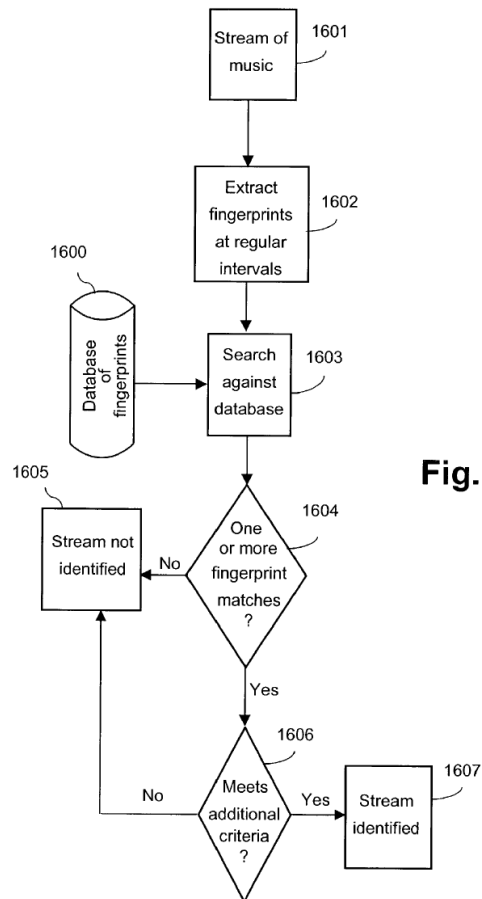


Fig. 12

EX1015, FIG. 12.

VIII. CLAIM CONSTRUCTION

106. I have been asked to interpret the claim terms recited in the '622

patent consistent with their plain and customary meaning to a person of ordinary skill in the art, as informed by the specification.

107. I reserve the right to offer claim construction opinions, however, in the event that Patent Owner proposes, or the Board adopts, particular constructions for claim terms.

IX. CLAIMS 1-13 OF THE '622 PATENT ARE OBVIOUS UNDER 35 U.S.C. § 103 OVER WANG777 AND WANG060

108. In my opinion, a POSA would have found claims 1-13 of the '622 patent obvious in light of the disclosure from Wang777 (EX1005) and Wang060 (EX1006).

A. Motivation to Combine Wang777 with Wang060

109. In my opinion, a POSA would have found it obvious to combine Wang777 and Wang060 for multiple reasons.

110. Wang777, as an initial matter, incorporates Wang060 by reference. EX1005, ¶[0022]; EX1007, 2. Wang777 also explicitly describes using the technique of Wang060 in the system for identifying a broadcast source. *See* EX1005, ¶¶[0022]-[0023], [0033], [0035], [0042], [0045], [0047], [0058], [0062]-[0065], [0074]. These explicit statements alone would have motivated a POSA to combine the teachings of Wang777 and Wang060.

111. Wang777 explains that “[o]nce the correct audio sound has been identified, the result is reported to the user or a system 324 by any suitable

method.” EX1005, ¶[0044]. Wang060 similarly teaches that a report for the identified sound samples, such as a log file of information about the identified sound recordings (e.g., title, artist, album cover art, lyrics, or purchase information) as well as “a list of music played from any sound stream (radio, internet streaming radio, hidden microphone, telephone call, etc.),” can be created. EX1006, ¶[0107].

112. To the extent that it is argued that Wang777 does not disclose a “play stream” of the content identification results of the identified sample along with its UST and USRTO in sequential time order (e.g., n , $n+1$, etc. where $[n]$ is the n th sample in time), this feature would have been obvious in view of Wang060. Wang060 teaches generating a candidate list of the identified sound samples that includes a sound_ID for each match and landmark timepoints. EX1006, ¶[0081], FIG. 9B.

113. In my opinion, it would have thus been obvious to store the content identification results of Wang777 in a play stream with their corresponding USTs and USRTOs (e.g., user sample identity $[n]$ and user sample identity $[n+1]$) in chronological order in a play stream such that each user sample can then be compared to identified broadcast samples having similar timestamps (i.e., in a play stream) in view of Wang060. Doing so would have required minimal modifications (e.g., modifying Wang777’s monitoring station 514 to report the result to a user in the form of a log file). Additionally, a POSA would recognize that adding the

teachings of Wang 060 to improve the output and reporting of the system would result in one that was more useful to end users.

114. In addition to this reason, it is my opinion that a POSA would have been motivated to combine Wang777 and Wang060 because they are in the same field and are directed to the same thing—identifying audio samples and the songs or stations from which the samples originate. *See* EX1005, Abstract; EX1006, Abstract. In particular, they are both directed to identifying audio samples from various sources and reporting that information to a user.

115. It is further my opinion that a POSA would have combined the system of Wang777 with the teachings of Wang060 by known methods, and the results of the combination would have been predictable to a POSA (e.g., identifying audio samples and reporting that information to a user). Indeed, as Wang777 states, it is desirable to identify the audio content of a broadcast transmission. EX1005, ¶[0018].

116. Further, the proposed combination of Wang777 and Wang060 would have involved nothing more than incorporating a well-known concept (i.e., Wang060's process for identifying audio samples and generating a log file) into a well-known system (i.e., Wang777's monitoring system).

B. Independent Claim 1

1. Preamble [1.P]

[1.P] <i>A media measurement method comprising:</i>

117. To any extent that the preamble is limiting, it is my opinion that Wang discloses a “*media measurement method.*”

118. For example, Wang777 discloses “a method for identifying a broadcast source of an audio sample” and, in particular, “matching audio or media file samples to a broadcast source from which the sample was transmitted.” EX1005, ¶¶[0002]-[0003], [0009], [0017]; EX1007, 2. The identification of a broadcast source provides “more commerce possibilities” (e.g., advertisements and promotional plans) to a content provider. EX1005, ¶[0005]. One method for identifying a broadcast source is illustrated in Figure 4 of Wang777. EX1005, ¶[0016]. In accordance with Figure 4, the method includes identifying the content of a media sample (402), for example, using the technique described in Wang060, and ultimately identifying a broadcast source (422) based on a comparison between timestamps and relative time offsets associated with the media sample and timestamps and relative time offsets associated with various broadcast samples for different broadcast channels. EX1005, ¶¶[0047]-[0056], FIG. 4; *compare* EX1007, 2-3, FIG. 2 *with* EX1001, 6:58-60, FIG. 2.

119. Wang777 further discloses that broadcast information is “reported to a

consumer tracking agency, or otherwise aggregated for statistical examples, for example.” EX1005, ¶[0018]; *see also* EX1005, ¶¶ [0019] (“The information may also be reported to a consumer tracking agency, or otherwise aggregated for statistical purposes, for example.”); [0020] (stating that broadcast information is reported to a “consumer tracking agency” or “statistical center”). Thus, Wang777’s “method for identifying a broadcast source of an audio sample” is a “media measurement method.” EX1005, ¶[0009]. Thus, Wang777’s “method for identifying a broadcast source of an audio sample” is the “media measurement method” described by the ’622 patent. EX1005, ¶[0009]; EX1001, 2:3-5 (stating that “embodiments analyze sequence data to enhance content identification and/or channel identification”).

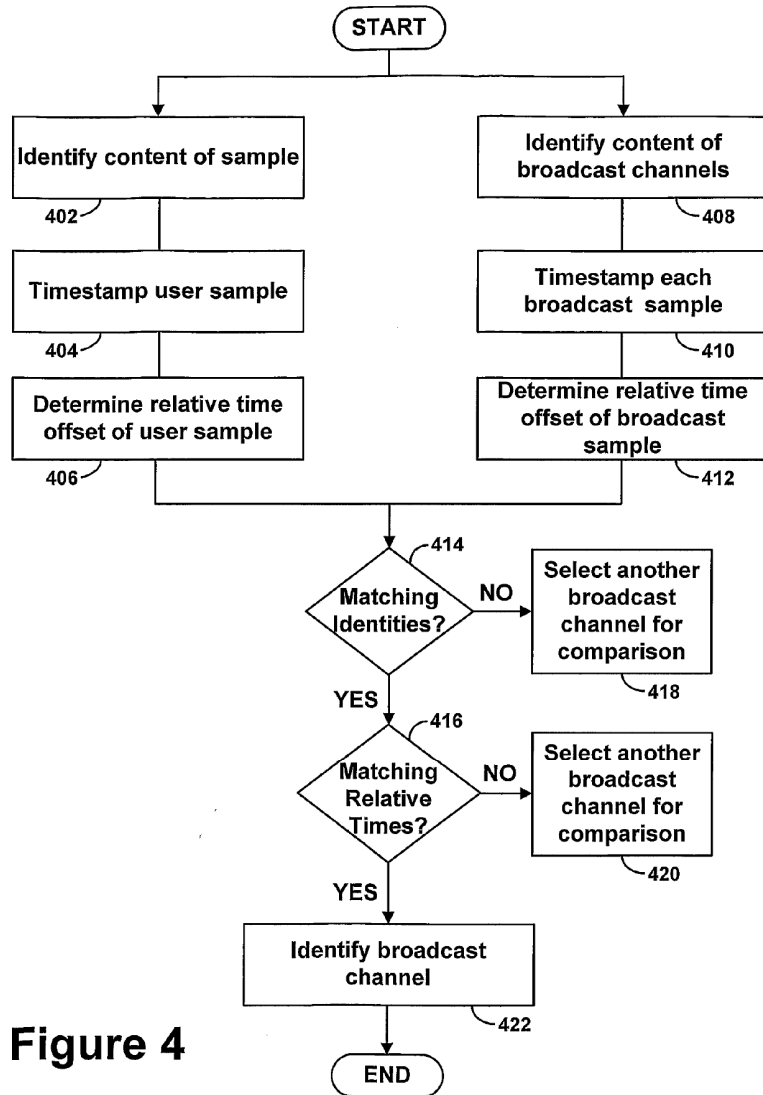
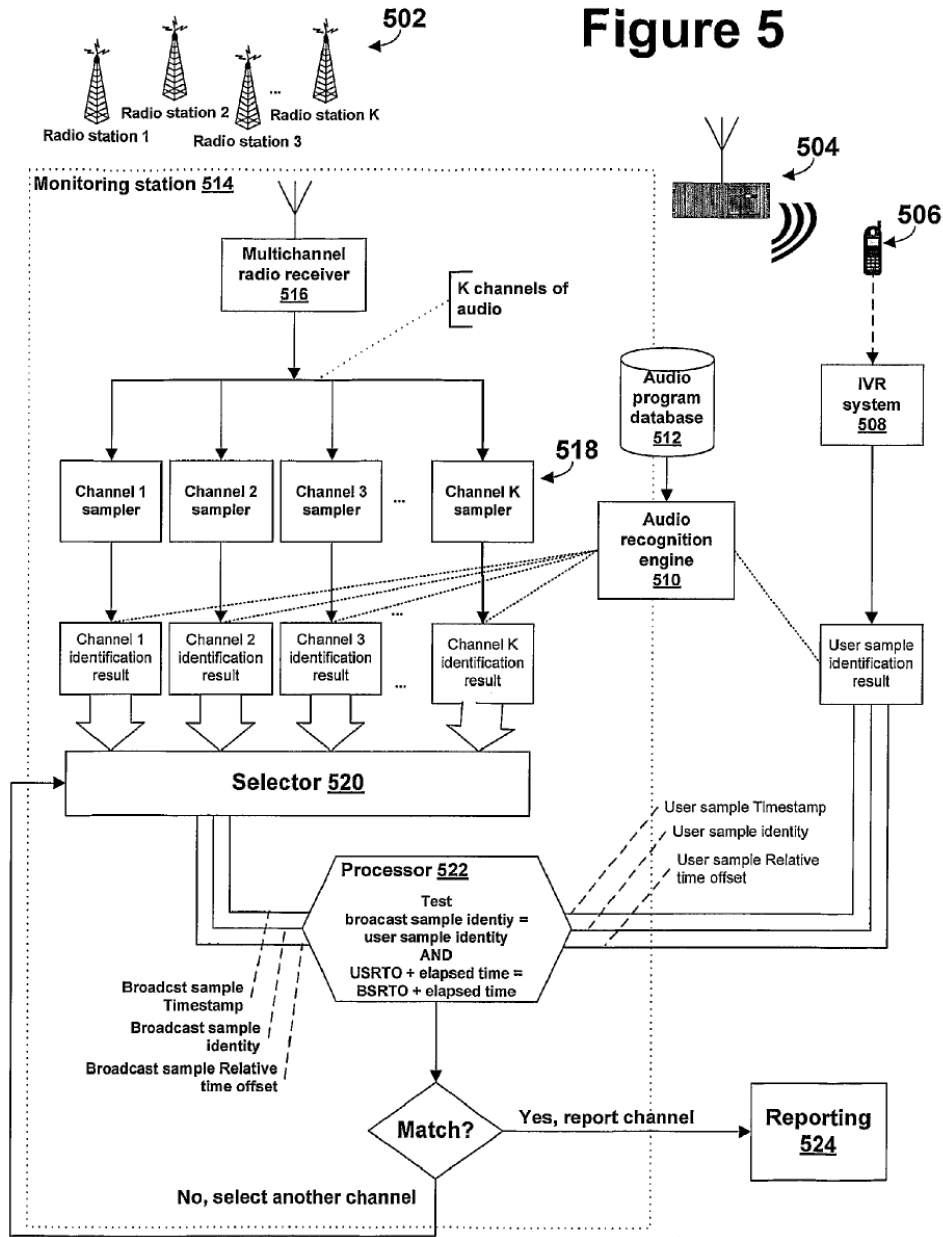


Figure 4

EX1005, FIG. 4

Figure 5



EX1005, FIG. 5.

2. Element [1.1]

[1.1] *using one or more computers to generate a play stream of content identification results corresponding to a sequence of data samples, the data samples comprising representations of audio data captured at a media monitoring device,*

120. In my opinion, Wang777 discloses and/or renders obvious this element in view of Wang060.

121. First, Wang777 teaches first capturing a sequence of data samples (e.g., audio samples) comprising representations of audio data (e.g., broadcast radio) captured at a media monitoring device (e.g., audio sampling device 102, 506, blue). EX1005, ¶¶[0009], [0019]-[0020], [0022], [0047], [0057], FIGs. 1-2, 4-5; EX1007, 2-3 (stating that audio samples may be recorded by an audio sampling device). Wang777 explains that the audio sampling device “may be a mobile cellular telephone, a PDA, or any device with processing means” and that the audio data samples may be from any source such as a radio station, television, internet streaming, etc. (e.g., broadcaster 104, 502/504, red). EX1005, ¶¶[0018]-[0020], [0057], FIGs. 1-2, 4-5; EX1007, 2-3 (stating that samples are recorded by an audio sampling device, such as a mobile phone).

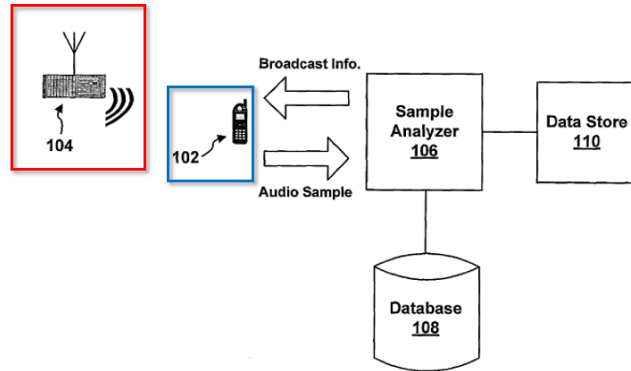


Figure 1

EX1005, FIG. 1 (annotated).

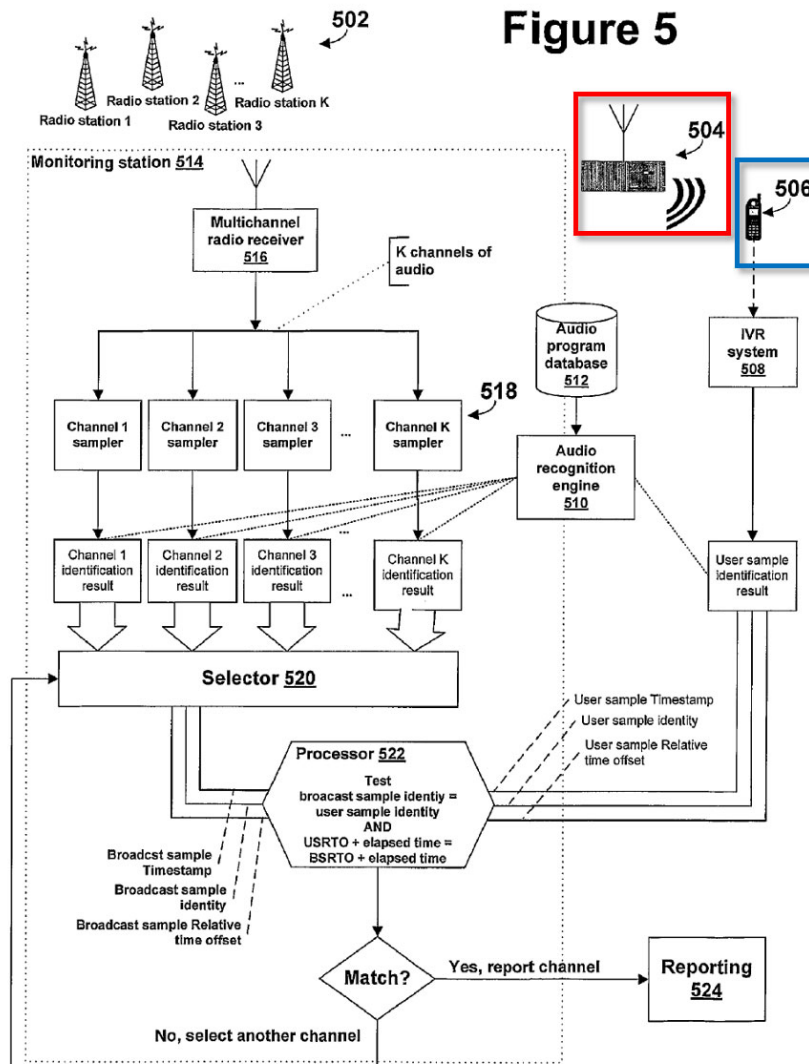


Figure 5

EX1005, FIG. 5 (annotated).

122. Second, Wang⁷⁷⁷ teaches generating content identification results (e.g., title, artist(s), and record label) corresponding to the sequence of audio data samples. EX1005, ¶¶[0004], [0072]; EX1007, 1, 5-6 (stating that a decision is made as to whether a sample originated from a given broadcast source and reporting that channel to a user), FIG. 2. The audio data samples are captured by audio sampling device 102, 506 over time (e.g., time [n], time [n+1], etc). EX1005, ¶¶[0004], [0072]; EX1007, 1, 5-6 (stating the samples are taken periodically), FIG. 2. Multiple audio data samples are captured over time to “verify that the user is actually listening to a given broadcast channel” rather than listening via a CD player. EX1005, ¶[0072].

123. As shown in Figure 4, for example, each audio sample collected by the audio sampling device 102/506 is identified (402). EX1005, ¶[0047]; EX1007, 5-6 (stating that a broadcast sample identity is identified for each sample), FIG. 2. A user sample timestamp (UST) is taken to mark the beginning time of each audio sample based on a standard reference clock (404). EX1005, ¶¶[0047], [0022]; EX1007, 5-6, FIG. 2. A user sample relative time offset (USRTO) is also determined (406). EX1005, ¶¶[0047], [0022], [0024]; EX1007, 5-6, FIG. 2. Wang⁷⁷⁷ explains that the USRTO is a time offset between a beginning of the identified content file from the database and a beginning of the audio sample being

analyzed, e.g., a user may record a ten second sample of a song that was 67 seconds into a song. EX1005, ¶[0047]; EX1007, 5.

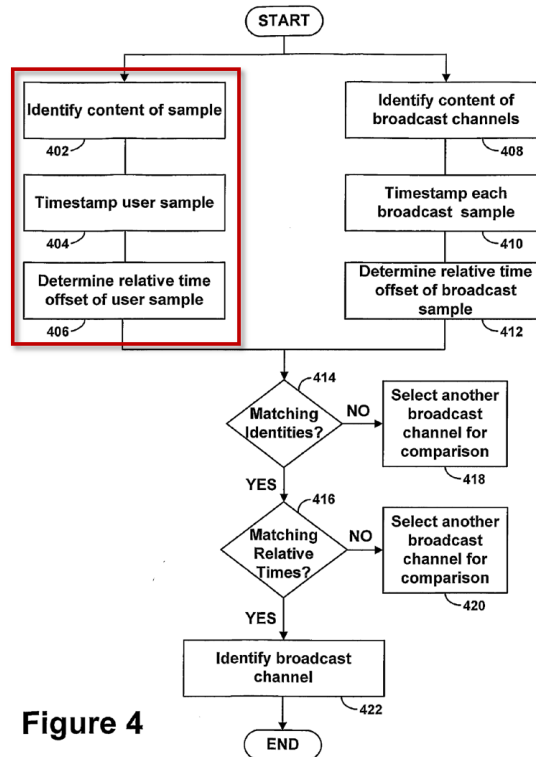


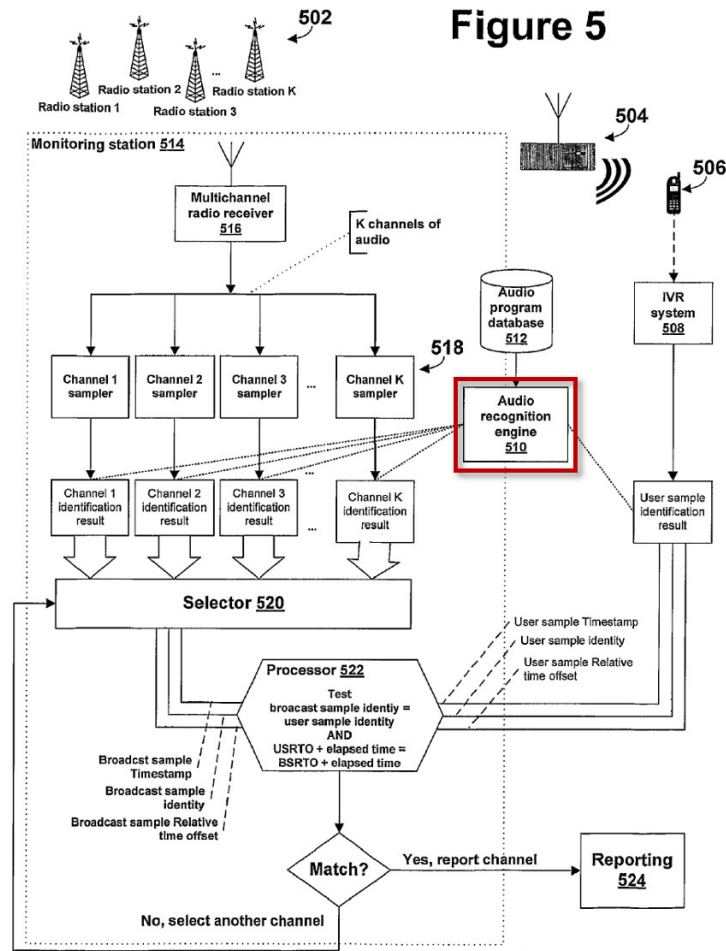
Figure 4

EX1005, FIG. 2 (annotated).

124. Figure 5 of Wang777 illustrates a system for identifying a broadcast source of an audio sample according to the method illustrated in Figure 4. EX1005, ¶[0057]; EX1007, 5-6, FIG. 2. With respect to Figure 5, Wang777 explains that the audio recognition engine 510 will identify each audio sample by performing a lookup within an audio program database 512 using the technique described within Wang060, which is incorporated by reference into Wang777 and described in detail relative to the embodiment shown in Wang777's Figure 2. EX1005, ¶[0058],

see also id., ¶[0020] (describing that a sample analyzer uses an audio sample to identity information pertaining to a broadcast by accessing a database containing the sample and broadcast information), [0028]-[0033], [0037], [0039]-[0040], [0059]-[0065]; EX1007, 1, 5-6, FIG. 2. Specifically, an audio recognition engine 510 identifies an audio sample by comparing fingerprints of the audio sample to fingerprints of media files stored in a database. EX1005, ¶¶[0058]-[0059]; EX1007, 1, 5-6, FIG. 2.

125. Wang777 thus discloses the exact same technique for generating content identification results corresponding to a sequence of data samples as the '622 patent. *See* EX1001, 5:42-62 (admitting that content recognition techniques were known and citing Wang060 as a known technique); *supra* Section VI.A.



EX1005, FIG. 5 (annotated).

126. After matching the fingerprints, a relative time difference between the fingerprints of the audio sample and the fingerprints of the matched media file are determined. EX1005, ¶¶[0061]-[0062]; EX1007, 1, 5-6, FIG. 2. The identity of the audio sample and its relative time offset are returned by audio recognition engine 510. EX1005, ¶[0065]; EX1007, 1, 5-6, FIG. 2.

127. Wang777 further discloses generating a play stream of the content identification results. As explained above, Wang777 identifies each sample along

with its UST and USRTO. EX1005, ¶[0047]; EX1007, 1, 5-6, FIG. 2. Wang⁷⁷⁷ further explains that samples (e.g., “user sample identity[n]” and “user sample identity[n+1]” where “[n] is the nth sample in time”) captured over a period of time (e.g., longer than a typical audio program) can be compared to identified broadcast samples. EX1005, ¶[0072]; EX1007, 7.

128. This is performed to track the continuity or transition between songs and “verify that the user is actually listening to a given broadcast channel” rather than listening via, for example, a CD player. EX1005, ¶[0072]. For instance, suppose an identity of a first audio sample (e.g., user sample identity[n]) identifies a first song title, and an identity of a second audio sample (e.g., user sample identity[n+1]) identifies a second song title that is different than the first song title. Such sample identities are compared to consecutive broadcast sample identities (e.g., broadcast sample identity[n] and broadcast sample identity[n+1]). If the broadcast sample identities[n] and [n+1] match the user sample identities[n] and [n+1], respectively, then a determination is made that the user was listening to the broadcast source from which broadcast sample identities[n] and [n+1] are based with greater confidence. EX1005, ¶[0072]. If there is a mismatch between broadcast sample identity[n+1] and user sample identity[n+1], then a determination is made that the user has changed channels. EX1005, ¶[0073].

129. It is my opinion that a POSA would have understood from Wang⁷⁷⁷

that to compare multiple content identification results captured over a period of time to identified broadcast samples, the content identification results would have been stored with their corresponding USTs and USRTOs such that each user sample can be compared to identified broadcast samples having similar timestamps. EX1005, ¶[0073]. And it would have been obvious in my opinion to store the identified samples in a known format such as a play stream for later comparison to the known broadcast samples. *See* EX1001, 10:34-56 (broadly describing “play stream”).

130. To the extent that it is argued that Wang777 does not disclose a “play stream” of the content identification results of the identified sample along with its UST and USRTO in sequential time order (e.g., n , $n+1$, etc. where $[n]$ is the n th sample in time), this feature would have been obvious in view of Wang060. Wang060 teaches generating a candidate list of the identified sound samples that includes a sound_ID for each match and landmark timepoints. EX1006, ¶[0081], FIG. 9B.

131. In my opinion, it would have been obvious to store the content identification results of Wang777 in a play stream with their corresponding USTs and USRTOs (e.g., user sample identity $[n]$ and user sample identity $[n+1]$) in chronological order in a play stream such that each user sample can then be compared to identified broadcast samples having similar timestamps (i.e., in a play

stream) in view of Wang060.

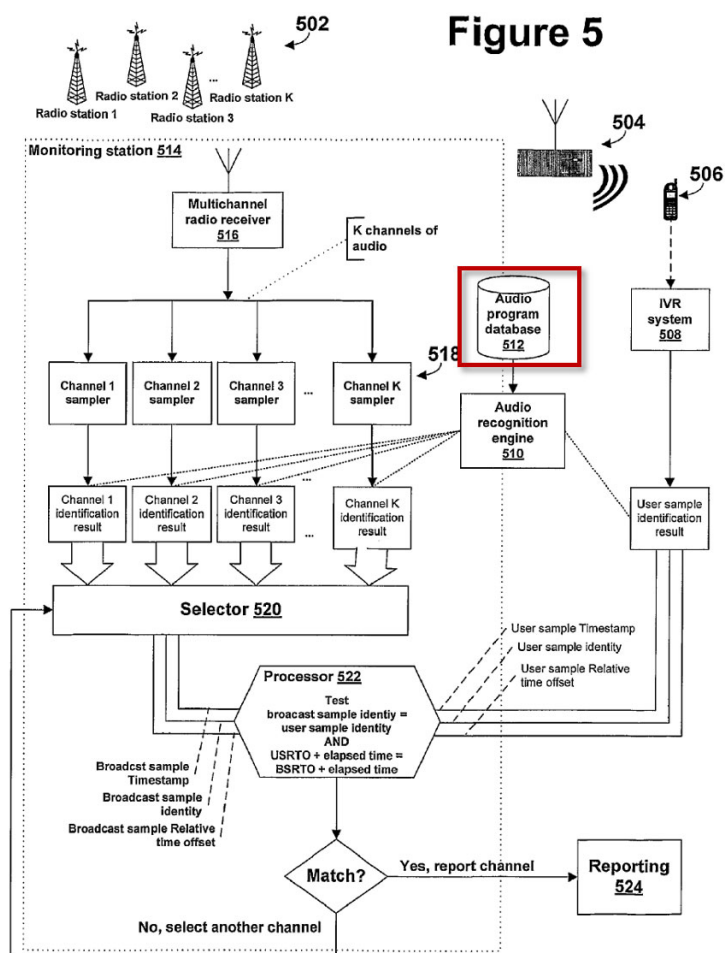
132. Lastly, Wang777 teaches generating the play stream of content identification results using one or more computers. Wang777 states, for example, that “the apparatus and methods described herein may be implemented in hardware, software, or a combination, such as a general purpose or dedicated processor running a software application through volatile or non-volatile memory. EX1005, ¶[0078], *see also id.*, ¶¶[0011], [0068]-[0070] (describing that a “processor” performs broadcast station/channel identification); EX1007, 1.

3. Element [1.2]

<i>[1.2] wherein content identifiers of the play stream's content identification results are obtained by using corresponding data samples to search a computerized database of known content; and</i>
--

133. In my opinion, Wang777 discloses and/or renders obvious this element in view of Wang060.

134. As I described above, Wang777’s audio recognition engine 510 will identify each audio sample by performing a lookup within an audio program database 512 using the technique described within Wang060 and generate a play stream of the content identification results. *See supra* Section IX.B.2; EX1006, ¶[0058]; EX1007, 1, 5-6, FIG. 2. This again is similar to the disclosure of the ’622 patent. EX1001, 5:42-62 (describing that Wang060 may be utilized to identify an audio sample).



EX1005, FIG. 5 (annotated).

4. Element [1.3]

[1.3] in response to obtaining the content identifiers, using one or more computers to utilize a sequential order of at least two different obtained content identifiers in the play stream to identify a channel² corresponding to the data samples,

135. In my opinion, Wang777 discloses and/or renders obvious this element in view of Wang060.

136. As I explained above for element [1.2], Wang777 obtains content identifiers (e.g., title, artist(s), and record label) corresponding to a sequence of audio data samples (e.g., the audio samples captured by audio sampling device 102, 506 over time) at steps 402-406. EX1005, ¶¶[0004], [0072]; EX1007, 1, 5-6, FIG. 2. Specifically, at step 402, content identifiers for audio samples are obtained using the techniques described in Wang060. EX1005, ¶[0047]. In addition to obtaining the content identifiers, a user sample offset (USO) indicating the beginning time of the collected audio sample is obtained based on a standard

² I note that the '622 patent describes the claimed “channel” broadly—including “an album, samples of an album presented for marketing—e.g. on a web page—, radio broadcast, television broadcast, theater version of a movie, DVD version of a movie, etc.” EX1001, 8:50-56.

reference clock at 404, and a user sample relative offset (USRTO) between the beginning of the identified content file from the database and the beginning of the audio sample being analyzed is obtained at 406. EX1005, ¶[0047].

137. After the content identifiers have been obtained (i.e., in response to obtaining the content identifiers), Wang⁷⁷⁷ identifies a broadcast source/channel corresponding to the audio data sample at steps 414-422. EX1005, ¶[0051]-[0056], FIG. 4; EX1007, 1, 5-6, FIG. 2. Specifically, the content identification of each audio data sample is compared to identified broadcast audio samples at 414. EX1005, ¶[0051]. If a match is found, the “relative times” are compared at 416. EX1005, ¶[0051]. If matched “relative times” are found, the broadcast source/channel is identified and reported back to the user at 422. EX1005, ¶[0051]. The comparisons of the user and broadcast samples are performed according to the following equations:

$$\text{“(User sample identity)=(Broadcast sample identity)”}$$

Equation (1)

$$\text{USRTO}+(\text{ref. time}-\text{UST})=\text{BSRTO}+(\text{ref. time}-\text{BST})+\text{delay}$$

Equation (2)

where the ref. time is a common reference clock time, and (ref. time-UST) and (ref. time-BST) take into account the possibility for different sampling times by the user audio sampling device and the monitoring station (e.g.,

(ref. time-BST)=elapsed time since last broadcast sample and now).” EX1005,

¶[0052]; EX1007, 1, 5-6, FIG. 2.

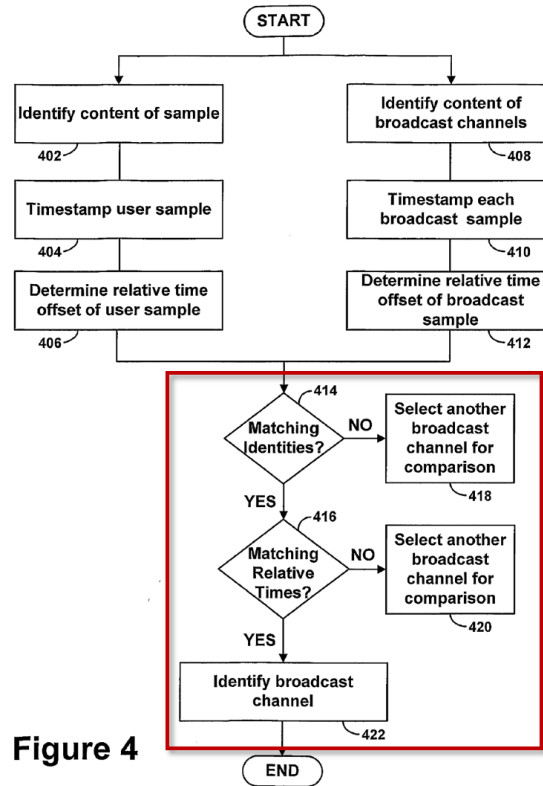


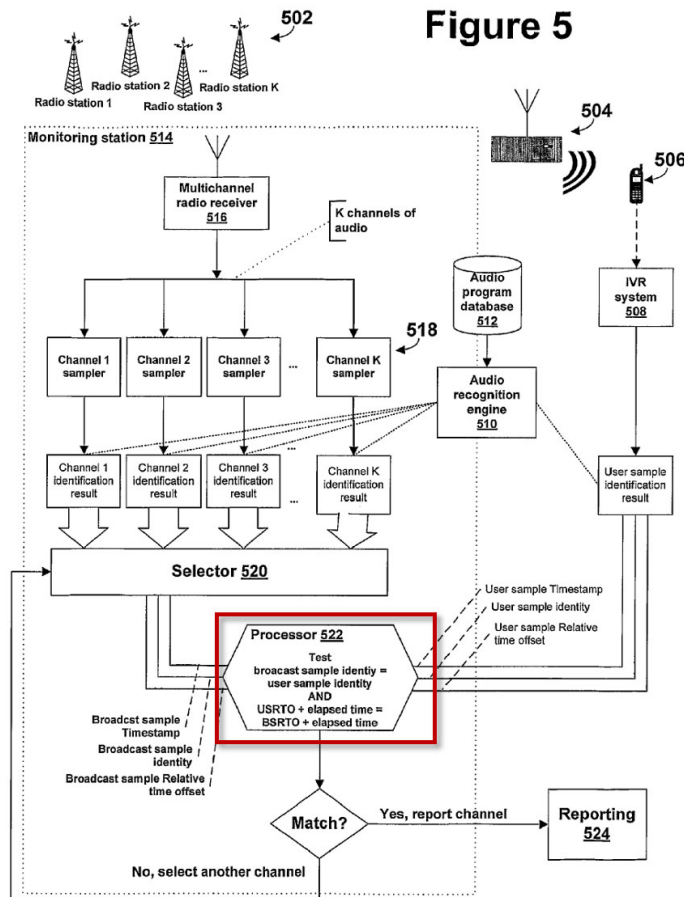
Figure 4

EX1005, FIG. 4 (annotated).

138. Wang777 Figure 5 illustrates processor 522 implementing steps 414-422 of Figure 4 to determine a broadcast source/channel using the content identification result, UST, and USRTO. EX1005, ¶¶[0068]-[0070], FIG. 5; EX1007, 5-6, FIG. 2. Specifically, upon receiving an inquiry from a user to determine broadcast information for a given audio sample, processor 522 first selects a given broadcast channel, from a plurality of broadcast channels, to determine if a broadcast sample identity of a broadcast sample recorded at or near

the user sample time matches the user audio sample identity. EX1005, ¶[0068].

Once an identity match is found, processor 522 determines if the user sample relative time matches the broadcast sample relative time for the selected broadcast channel. EX1005, ¶[0069]. If the relative time matches, then processor 522 considers the audio sample and the broadcast sample to be a match. EX1005, ¶[0069]. After finding the match, processor 522 reports the identified broadcast channel, for example, to the user. EX1005, ¶[0070].



EX1005, FIG. 5 (annotated).

139. While the description of Figures 4-5 focus on identifying the

source/channel for a single audio data sample, Wang⁷⁷⁷ further discloses identifying a source/channel corresponding to the audio data samples using a sequential order of at least two different obtained content identifiers in the play stream. Wang⁷⁷⁷, in particular, explains that to verify which channel a user is actually listening to, content identifiers obtained over a period of time (e.g., longer than a typical audio program, such as over a transition between audio programs on the same channel) are used. EX1005, ¶[0072]; EX1007, 7. Wang⁷⁷⁷ explains that if the correct channel is identified, the content identifiers (e.g., user sample identity[n] and user sample identity[n+1]) will match even across a transition between audio programs (e.g., a transition from one song to another). EX1005, ¶[0072]; EX1007, 7 (stating that program transitions are tracked because the “continuity of identity over a program transition is a strong indicator that the correct broadcast channel is being tracked”).

140. Specifically, to provide further confidence that a correct match has been made (e.g., when both the user and the broadcast sources change synchronously or the user listens to more than one song via the same broadcast source), the following equations are used:

$$\text{“User sample identity}[n]=\text{Broadcast sample identity}[n]$$

Equation (3)

$$\text{User sample identity}[n+1]=\text{Broadcast sample identity}[n+1]$$

Equation (4)

$\text{User sample identity}[n] \neq \text{User sample identity}[n+1]$

Equation (5)

where $[n]$ is the n th sample in time.”

EX1005, ¶[0072]; EX1007, 7.

141. In accordance with Equation 3, if the user sample identity $[n]$ equals broadcast sample identity $[n]$, then a determination is made that a user is listening to a broadcast source from which broadcast sample identity $[n]$ is based. In accordance with Equations 4 and 5, if the next user sample identity (e.g., user sample identity $[n+1]$) equals the next broadcast sample identity (e.g., broadcast sample identity $[n+1]$), but user sample identity $[n]$ does not equal user sample identity $[n+1]$, then the continuity or transition between different songs has been tracked and a determination is made that both the user and the broadcast sources changed synchronously or that the user listened to more than one song via the same broadcast source. EX1005, ¶[0072].

142. As such, in response to obtaining the user sample identities (content identifiers), it is my opinion that Wang⁷⁷⁷ teaches using one or more computers to utilize user sample identity $[n]$ and user sample identity $[n+1]$ (a sequential order of at least two different obtained content identifiers) in the play stream, *see supra* Section IX.A, to identify a broadcast source (channel) corresponding to the data

samples.

5. Element [1.4]

[1.4] the at least two different obtained content identifiers identifying different media content items.

143. It is my opinion that Wang777 discloses this element.

144. As explained above for element [1.3], Wang777 teaches that the content identifiers, user sample identity[n] and user sample identity[n+1], used to determine the source/channel can be from different songs (i.e., different media content items). EX1005, ¶[0072]; EX1007, 7. This is exemplified in Equation 5:

$$\text{User sample identity}[n] \neq \text{User sample identity}[n+1]$$

Equation (5)

where[n] is the nth sample in time.

EX1005, ¶[0072]; EX1007, 7.

145. In accordance with Equation 5, if a first sample identity (e.g., user sample identity[n]) does not match a second sample identity (e.g., user sample identity[n+1]), then a transition between two different songs is tracked. EX1005, ¶[0072]; EX1007, 7. Accordingly, Wang777 discloses two different obtained content identifiers that identify different media content items.

C. Independent Claims 6 and 10

1. [6.P], [10.P]

[6.P] *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 instructions comprising instructions for:*

[10.P] *A computer system comprising one or more computers configured by computer code stored in the one or more computers, the computer code comprising instruction executable by one or more processors of the one or more computers, the instructions comprising instructions for:*

146. It is my opinion that Wang777 discloses these preambles to the extent they are limiting for the reasons set forth above for element [1.P].

147. Wang777 states that “the present application is capable of being distributed in the form of a computer-readable medium of instructions in a variety of forms, and that the present application applies equally regardless of the particular type of signal bearing media used to actually carry out the distribution. Examples of such computer-accessible devices include computer memory (RAM or ROM), floppy disks, and CD-ROMs, as well as transmission-type media such as digital and analog communication links.” EX1005, ¶[0077].

148. Wang777 further states that “the apparatus and methods described herein may be implemented in hardware, software, or a combination, such as a general purpose or dedicated processor running a software application through

volatile or non-volatile memory. EX1005, ¶[0078], *see also id.*, ¶¶[0011], [0068]-[0070]; EX1007, 1.

2. [6.1], [10.1]

149. It is my opinion that Wang777 discloses and/or renders obvious this element in view of Wang060 for at least the reasons set forth above for element [1.1].

3. [6.2], [10.2]

150. It is my opinion that Wang777 discloses and/or renders obvious this element in view of Wang060 for at least the reasons set forth above element [1.2].

4. [6.3], [10.3]

151. It is my opinion that Wang777 discloses and/or renders obvious this element in view of Wang060 for at least the reasons set forth above element [1.3].

5. [6.4], [10.4]

152. It is my opinion that Wang777 discloses this element for at least the reasons set forth above for element [1.4].

D. Dependent Claims 2, 7, and 11

[2], [7], [11] <i>converting a raw play stream to a format usable for searching channel data to identify a channel to associate with content of the raw play stream.</i>

153. It is my opinion that Wang777 alone or in combination with Wang060 discloses this element.

154. As explained above, Wang777-Wang060's content identification results are stored with their corresponding USTs and USRTOs (e.g., user sample identity[n] and user sample identity[n+1]) in chronological order in a play stream. It is my opinion that a POSA would have understood that this play stream is a "raw play stream." EX1001, 10:34-40 (describing "raw play stream" broadly).

155. Wang060 further teaches generating a log file of information about the identified sound recordings (e.g., title, artist, album cover art, lyrics, or purchase information) as well as "a list of music played from any sound stream (radio, internet streaming radio, hidden microphone, telephone call, etc.)" based on the determined identities of the sound content. EX1006, ¶[0107]. Wang060 states that the list may also be made "only when the identity of recognized sound changes" "to avoid redundancy." EX1006, ¶[0107].

156. In my opinion, a POSA would have thus understood that in Wang777-Wang060, the content identification results stored with their corresponding USTs and USRTOs (raw play stream) are converted to a list of music played (a play list), which would also be a format usable for searching channel data to identify a channel to associated with content of the raw play stream because it includes the information necessary to determine the source/channel (e.g., songs identified over a specific period of time). EX1001, 8:44-49 (describing how data is converted to a format useful in matching to a channel (e.g. by converting a play stream to a play

list format)), 10:34-40 (describing “play list” broadly); *see also* EX1014, ¶¶[0004]-[0005] (describing known play lists), ¶[0006] (describing that a playlist is “the documentation of the performance of one or more works at a particular time over a particular media”).

E. Dependent Claims 3, 8, and 12

[3], [8], [12] *wherein the raw play stream is converted to a play list.*

157. It is my opinion that Wang777 alone or in combination with Wang060 discloses this element.

158. As explained above, Wang777-Wang060 teaches converting the content identification results stored with their corresponding USTs and USRTOs (raw play stream) to a list of music played (a play list). *See supra* Section XI.D.

F. Dependent Claim 4

[4] *wherein the raw play stream is converted to a mathematical representation of the raw play stream.*

159. It is my opinion that Wang777 alone or in combination with Wang060 discloses this element.

160. Wang060 explains that when the log is created, the reported results can include an “absolute score of the sound file or its score in comparison to the next highest scored file,” which corresponds to the identified content. EX1006, ¶[0105]; *see also* ¶¶ [0040] (describing that file identifier with the highest score is

identified as the winning file identifier), [0089]-[0090] (describing that library sound files not having a minimum threshold score are not utilized for identification), [0097]-[0099] (describing how sound files are ranked based on a probability score), [0101] (describing that the search for a target file is terminated when a score exceeds a predetermined threshold value). In my opinion, A POSA would have thus understood the scores in the log generated from the content identification results (raw play stream) are a mathematical representation converted from the raw play stream. *See* EX1001, 8:44-49 (describing the mathematical representation without detail).

G. Dependent Claims 5, 9, and 13

[5], [9], [13] wherein using one or more computers to utilize a sequential / utilizing an order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples comprises searching channel information that includes a plurality of channel types including at least a broadcast media channel type and a non-broadcast playable media channel type.³

161. It is my opinion that Wang777 alone or in combination with Wang060 discloses this element.

³ I note that the '622 patent does not describe the term “non-broadcast playable media channel type” in any detail. *See generally* EX1001.

162. As I explained above, Wang777 teaches “using one or more computers to utilize a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples.” *See supra* Section IX.B.4.

163. Wang777 also explains that “[i]n the field of content identification, it may be desirable to identify not only content, but also a source (e.g., channel, stream, or station) of a broadcast transmission.” EX1005, ¶[0018]; EX1007, 1. In particular, the audio sample being analyzed can be “from any number of analog or digital sources, such as a stereo system, television, radio broadcast, Internet streaming broadcast, or any other suitable means of transmitting such recorded material” EX1005, ¶[0027]; EX1007, 1, 4 (stating that audio samples may be from varying sources, such as radio, television, or internet streaming). “Depending on the source, the sample can be in the form of acoustic waves, radio waves, a digital audio PCM stream, a compressed digital audio stream (such as Dolby Digital or MP3), or an Internet streaming broadcast.” EX1005, ¶[0027], *see also id.*, ¶[0019]. EX1007, 1-2, 4.

164. It is my opinion that a POSA would have thus understood that Wang777’s identification of a broadcast source for audio “from any number of analog or digital sources, such as a stereo system, television, radio broadcast, Internet streaming broadcast, or any other suitable means of transmitting such

recorded material” corresponds to searching channel information that includes a plurality of channel types including at least a broadcast media channel type (e.g., radio broadcast over radio waves) and a non-broadcast playable media channel type (e.g., a CD player or a stereo system playing recorded material in MP3 format). EX1005, ¶[0072] (which suggests the inference of a user taking a recording from a CD player (or some other non-broadcast source) if consecutive user sample identities do not match any broadcast sample identifiers).

165. To the extent that it may be argued that Wang777 does not disclose this element, Wang060 does. Wang060 explains that the system provides “a method for recognizing an exogenous media sample given a database containing a large number of known media files” and a “a method for generating a database index that allows efficient searching using the recognition method. EX1006, ¶[0037]. “The sound database index contains an index set for each file or recording in the selected collection or library of recordings, which may include speech, music, advertisements, sonar signatures, or other sounds.” EX1006, ¶[0047]. “To index the sound database, each recording in the collection is subjected to a landmarking and fingerprinting analysis that generates an index set for each audio file.” EX1006, ¶[0049].

166. Thus, after the sound database is generated, a “microphone or other source of sound is continually sampled into a buffer to obtain a record of the

previous N seconds of sound.” EX1006, ¶[0107]. The contents of the sound buffer are periodically analyzed to determine the identity of the sound content from the database. EX1006, ¶[0107]. A report or log file of information about the identified sound recordings (e.g., title, artist, album cover art, lyrics, or purchase information) as well as “a list of music played from any sound stream (radio, internet streaming radio, hidden microphone, telephone call, etc.)” is created based on the determined identities of the sound content. EX1006, ¶[0107]. Wang060 also explains that the audio sample can be “from any number of analog or digital sources, such as a stereo system, television, compact disc player, radio broadcast, answering machine, telephone, mobile telephone, Internet streaming broadcast, FTP, computer file as an email attachment.” EX1006, ¶[0079].

167. It is my opinion that it would have thus been obvious to a POSA in view of Wang060 for the audio program database 512 of Wang777 to store not only the identified content from radio stations 502 but to also store content identified from other non-broadcast playable media channel type sources (e.g., a compact disc player) to enable proper identification of a source of the audio. Thus, for example, if a user is listening to an album played from a compact disc player via a stereo system, audio sample device 506 may sample that audio and Wang777 can properly identify the album on the compact disc and report back to the user “an advertisement of a promotion, prize notification, discount offers” as well as

“biographical information, information about concerts in the vicinity, and other information of interest to fans can be provided” (i.e. searching channel information that includes a plurality of channel types including at least a broadcast media channel type and a non-broadcast playable media channel type). EX1005, ¶¶[0019], [0044]. Otherwise, Wang777 would not be able to identify the source unless a radio station broadcasts the album.

X. CLAIMS 2-4, 7-8, AND 11-12 OF THE '622 PATENT ARE OBVIOUS UNDER 35 U.S.C. § 103 OVER WANG777, WANG060, HERLEY682, SEET, AND THE DATA IMPUTATION REFERENCES

168. In my opinion, a POSA would have found claims 2-4, 7-8, and 11-12 of the '622 obvious in light of the disclosures from Wang777 (EX1005), Wang060 (EX1006), Herley682 (EX1011), Seet (EX1012), and the data imputation references (EX1017, EX1018, EX1019, EX1020, and EX1023).

A. Motivation to Combine Wang777, Wang060, Herley682, and Seet

169. It is my opinion that a POSA would have found it obvious to combine Wang777, Wang060, Herley682, and Seet for multiple reasons.

170. As an initial matter, I note that each reference relates to methods and systems of identifying objects of interest (e.g., songs) within signals (e.g., a radio broadcast). EX1005, ¶¶[0002]-[0003], [0009], [0017]; EX1006, Abstract, ¶¶[0026], [0039]-[0042]; EX1011, ¶[0002]; EX1012, ¶ [0019]. They all do so by extracting fingerprints from samples and comparing these to fingerprints of other,

e.g., known, objects in order to determine if there is a match. EX1005, ¶¶ [0021] and [0061] (stating that “fingerprints” of recordings taken at similar times can be compared”); EX1006 ¶¶[0026], [0039]-[0042] (illustrating that fingerprints of an audio sample are compared to fingerprints of audio samples stored in a database); EX1011, ¶[0086] (stating that fingerprints extracted from a sample are compared to fingerprints stored in a database); EX1012, ¶¶ [0021] (stating that a fingerprint of audio of a radio broadcast is compared to fingerprints of a media content database). As described below, Herley682 and Seet both also describe how to handle the situation when there is not a match.

171. In particular, Herley682 discloses a “system and method for sampling one or more channels of a broadcast spectrum, such as a radio frequency spectrum, identifying and storing content information for each sampled channel, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1011, ¶2. Herley682’s “interactive signal analyzer” among other things, “monitors one or more signals, identifies objects of interest within such signals, [and] stores statistical and metadata information describing such objects to an object database.” EX1011, ¶10. Herley682 teaches that these monitored signals may be, for example, audio signals, video signals, or combined audio/video signals. EX1011, ¶11. In Herley682, unmatched objects are stored as “unknown objects” and can be manually identified by a user. EX1011, ¶¶16-17,

71. It is my opinion that POSA would have understood that content in Wang777-Wang060 may not be matched in view of Herley682.

172. Seet teaches using system 106 to “infer additional matches” for unmatched samples “based on their position in a media stream.” EX1012, ¶32; EX1016, 6. In particular, an unmatched sample may be identified using “surrounding samples [that are] in the correct sequence and timing so that a match for the unmatched sample can be made.” EX1012, ¶32; EX1016, 6. The example in Seet’s Figure 4 identifies unmatched samples in the same way as later disclosed in Figure 4 of the ’622 patent. Namely, to identify the unmatched second sample, Seet looks to the previous sample and the following sample, and “[t]he temporal consistency of the bounding samples allows one to infer that they are of the same media item ... (in this case, an advertisement labeled #2); therefore, the middle sample is also of the same media item.” EX1012, ¶32; EX1016, 6. Using data imputation techniques, such as that disclosed in Seet, to correct missing and/or incorrect data in a data set, such as the unidentified sample/fingerprint in Wang777-Wang066, was well-known to a skilled artisan. *See, e.g., infra* Section X.B. (describing multiple data imputation techniques with reference to EX1017, EX1018, EX1019, EX1020, and EX1023).

173. While Herley682 provides a way to manually correct an unmatched sample/fingerprint in Wang777-Wang066, in my opinion a POSA would have been

motivated to use Seet's teachings of inferring matches to identify unmatched content in Wang777-Wang060 at least because in my opinion a POSA would have been motivated to automate that which is done manually because it would be faster. *See* EX1011, ¶80.

174. It is my opinion that a POSA would have further understood that using Seet's method to automatically correct the unmatched content in Wang777-Wang060 would have increased data reliability. Further, Seet provides express motivation for using its methods to improve media identification systems, such as that of unmatched content in Wang777-Wang060. *See* EX1012, ¶¶11-13. In particular, Seet's method provides "[s]ignificant operational and competitive advantages" and "improve[d] accuracy of matches," as well as "largely automat[ing]" "creating and maintaining the media content database." EX1012, ¶13.

175. Further still, Wang777, Wang060, Herley682, and Seet are in the same field of endeavor. *Compare* EX1005, ¶2, with EX1006, ¶2, EX1011, ¶2 EX1012, ¶3. In my opinion, a POSA would have been motivated to combine these references, and would have had a reasonable expectation of success in doing so, because they both relate to methods of identifying objects of interest in a broadcast signal. In my opinion, applying the teachings of Seet to Wang777-Wang060 would have simply been the combination of prior-art elements according to known

methods to yield predictable results. That is, a POSA, having implemented Wang777-Wang060, would recognize that there may be unmatched samples (e.g. as taught in Herley682) and see the clear advantages of using a system, such as that in Seet, to address unmatched samples easily and in an automated manner.

176. Similarly, it is further my opinion that the combination would have merely been the application of known techniques (Seet's method of inferring additional matches for unmatched samples) to a known system/method ready for improvement (Wang777-Wang060's system/method) to yield the predictable result of identifying objects of interest in a broadcast signal.

177. In my opinion, it would have been well within the level of ordinary skill to implement Seet's teachings in Wang777-Wang060 because they are similar media content identification methods implemented in conventional database systems implemented in conventional hardware/software. EX1005, ¶50; EX1006, ¶51.

178. Additional motivations for including specific features of Herley682 and Seet are discussed with respect to particular claim elements herein.

B. Dependent Claims 2, 7, and 11

[2], [7], [11] <i>converting a raw play stream to a format usable for searching channel data to identify a channel to associate with content of the raw play stream.</i>

179. It is my opinion that Wang777-Wang060 in view of Herley682 and

Seet render obvious this element.

180. The '622 specification, in one example, states that “[p]lay stream generator 126 generates a raw play stream relating a sequence of samples to content identification results (for an example of a portion of a raw play stream, see FIG. 3[] and accompanying text).” EX1001, 6:14-17, FIG. 4; *see also id.*, 8:1-3. Figure 4 illustrates an example of such a “series of samples and corresponding identification results” as a table including columns for “row,” “log time,” and “contentID” but is clear that a “raw play stream” “may vary considerably from that illustrated in FIG. 4” EX1001, 10:34-40, FIG. 4.

181. The '622 specification further states that a raw play stream may be prepared for further processing by “scrubbing (i.e., reviewing for missing or incorrect data).” EX1001, 8:32-35. However, the '622 specification does not claim “scrubbing” nor does the specification describe how this “scrubbing” is actually performed.

182. I note that the '622 specification simply discusses the “scrubbing” process with reference to Figure 4, stating that:

Sample sequence data may be used to “scrub” Raw Play Stream 410.

By analyzing the sequence of data on Raw Play Stream, it is possible to locate and correct any identifications that are likely to be incorrect or to supply missing identifications where the system has not

otherwise been able to provide an indication.

EX1001, 9:32-37. Figure 4 illustrates an example of missing identifications as “none” entries in 410, which are replaced to create Clean Play Stream 450, and ultimately identify a channel. EX1001, 9:38-55.

183. Thus to the extent it is argued that the claimed “converting” corresponds to reviewing data in a “raw play stream” for missing or incorrect data and correcting this data, it is my opinion that this would have been obvious in view of the Wang070-Wang060-Herley682-Seet combination.

184. As I noted above, Wang777-Wang060 identifies samples (e.g., “user sample identity[n]” and “user sample identity[n+1]” where “[n] is the nth sample in time”) along with corresponding USTs and USRTOS, which would have been stored in a known format such as a play stream for later comparison to the identified broadcast samples, also referred to as a “candidate list.” *See supra* Section IX.B.2.

185. Herley682 teaches that certain samples/fingerprints may not be identified and thus discloses a method for handling these unmatched samples/fingerprints, which correspond to the ’622 patent’s “none” entries in the raw play stream. EX1001, 9:42-44, FIG. 4; *see also* EX1006, ¶84. In Herley682, where “fingerprints generated from samples do not match any fingerprints in the database, those trace fingerprints are added as fingerprints to the fingerprint

database as ‘unknown objects.’” EX1011, ¶¶17-18, 71, 89, 164-165. Thus, in my opinion, a POSA would have understood that samples in Wang777-Wang060 may not be identified and it would have been obvious to store an indication that it was not identified together with the identified samples as taught by Herley682 (i.e., a “raw play stream”).

186. Herley682 teaches that the unidentified sample/fingerprint may be manually changed via a user interface. EX1011, ¶¶17-18, 71, 89, 164-165. Seet further teaches inferring matches as an improved way to identify Herley682’s unmatched samples/fingerprints, which, in my opinion, would have motivated a POSA to automate that which is done manually in Herley682 because it would be faster. EX1012, ¶¶26-32; EX1011, ¶80. Further, it is my opinion that it was well-known in the art to use other surrounding data in order to clean (scrub) a data set, as explicitly suggested by Seet. EX1012, ¶32; EX1016, 6.

187. For example, the publications to Lee (EX1018) and Myrtveit (EX1017) describe examples of known data imputation techniques. Lee notes that data quality problems can arise from a variety of sources. EX1018 28. “When mining these data warehouses for decision making information, making logical and well-informed decisions becomes problematic if the data’s quality is doubtful.” EX1018, 28. Lee teaches that “clean data is crucial” for many database implementations, including data mining. Lee notes “[d]ata cleaning refers to

processes for detecting and removing errors and inconsistencies from data.”

EX1018, 28. “Research in this area as included outlier detection, noise handling for classification, and duplicate elimination.” EX1018, 28. “Given the ‘garbage in, garbage out’ principle, clean data is crucial for database integration, data warehousing, and data mining, thus leading to data cleaning’s strong association with these research areas.” EX1018, 28.

188. Lee further explains that “[d]ata cleaning refers to processes for detecting and removing errors and inconsistencies from data.” EX1018, 28. Lee notes a type of erroneous data is known as ‘spurious links’ and Lee identifies a “context-based solution” to identify data that needs correcting. EX1018, 28.

189. Lee describes automated software-based methods for identifying and correcting data. EX1018, 30. “We implemented our algorithms in Java and ran the experiments on a Pentium 4 1.6-GHz system with 256 Mbytes of RAM running Windows XP Professional. We used IntelliClean4 on the three data sets to find record sets with high similarity in the spurious attribute, and we used DM2-CBA7 to generate association rules to determine the context attributes.” EX1018, 30.

190. Myrtveit discusses automated software-based imputation methods for data sets containing missing data. EX1017, 999. As described in Myrtveit, “imputation-based techniques” are one type of “missing data techniques (MDTs)” to handle missing data. EX1017, 1001. One such imputation technique is “similar

response pattern imputation (SRPI)” in which the most similar entry without missing data is determined, and those values are copied into the missing data. EX1017, 1002. In SRPI, “the idea is to identify the most similar project without missing [data] and to copy the values of this project to fill the [data] holes in the project with the missing [data].” EX1017, 1002. Myrtviet describes another technique known as “[h]ot deck imputation (HDI)” which “replaces missing values by drawing from an *estimated* distribution for each missing value.” EX1017, 1002. Myrtviet teaches “[i]t is common [in hot deck imputation] to use the *sample* distribution [...] as the distribution (or deck) to draw from.” EX1017, 1002, 1003. Continuing, Myrtviet teaches “SRPI belongs to the hot deck class,” and “picks the nearest neighbor rather than drawing from the sample at random.” EX1017, 1003. Myrtviet articulates clearly that “[i]mputation-based methods replace missing values by suitable estimates.” EX1017, 1001. In my opinion, a POSA would have immediately understood the desirability and advantages of replacing missing data values with suitable data estimates in order to enhance the integrity and reliability of the data, given the known problem of “dirty data” and known desirability for “cleaned” or “scrubbed” data.

191. Looking back in time, the article to Hernandez, entitled “Real-world Data is Dirty: Data Cleansing and the Merge/Purge Problem” and published in 1998 describes the known problem of dirty data. Hernandez notes that “[e]rrors

due to data entry mistakes, faulty sensor readings or more malicious activities, provide scores of erroneous data sets that propagate errors in each successive generation of data.” EX1023, p. 10. Hernandez describes a basic data cleansing solution utilizing a sorted neighbor method (considering neighbor data values to resolve data errors). EX1023, pp. 11-18.

192. The Lakshminarayan article, entitled “Imputation of Missing Data in Industrial Databases” and published in 1999, further evidences the well-known problem of dirty data and well-known data scrubbing techniques. The Lakshminarayan article discusses well-known approaches for data imputation in incomplete data structures. Lakshminarayan notes “[i]ncomplete data is an unavoidable problem in dealing with most of the real world data sources, and various methods for dealing with such data have been developed.” EX1019, p. 261. In the article, Lakshminarayan concludes that the mining of databases is often hampered by missing and erroneous data entries, and notes reasons for completing incomplete databases and summarizes well-known statistical approaches for missing data imputation. EX1019, p. 274. Broadly, Lakshminarayan provides a brief overview of existing methods for dealing with missing data, as roughly grouped following known taxonomy as 1) ignoring and discarding incomplete records and attributes, 2) parameter estimation in the presence of missing data, and 3) imputation procedures. EX1019, p. 261-262.

193. Finally, I note the 2002 article by Batlle entitled “Automatic Song Identification in Noisy Broadcast Audio” as exemplifying that it was well known at the time of the ‘622 patent that a problem with broadcast audio data is that the data is dirty. Batlle notes “[o]ne of the main problems with broadcast audio is that the received audio suffers several transformations before reaching the listener . . . and, therefore, the original and the broadcast songs are very different from the signal point of view.” EX1020, Abstract. This “dirty” data can impede proper automatic song identification. Although the Batlle reference pursues a different solution to the problem of dirty data, it evidences that data cleansing was, at the time of the ‘622 patent, known as a necessary component in audio content identification systems, especially when analyzing broadcast audio.

194. It is my opinion that a POSA would have been motivated to identify and correct any unmatched samples/fingerprints in Wang070-Wang060 in order to obtain a complete and accurate data set to identify a broadcast source (channel) corresponding to the data samples. Adding the ability to identify and correct unmatched samples would be desirable, increasing the accuracy and usability of the results, and easily implemented, for reasons explained above.

C. Dependent Claims 3, 8, and 12

[3], [8], [12] <i>wherein the raw play stream is converted to a play list.</i>
--

195. It is my opinion that Wang777-Wang060 in combination with

Herley682 and Seet discloses this element.

196. As I noted above, in the Wang777-Wang060-Herley682-Seet combination, a raw play stream of identified samples (e.g., “user sample identity[n]” and “user sample identity[n+1]” where “[n] is the nth sample in time”) and unidentified samples is created and then reviewed/corrected. *See supra* Section VIII.B. And Wang060 teaches that once the samples have been identified, “a log file can be collected, or a display can be shown on a device indicating information about the music, such as title, artist, album cover art, lyrics, or purchase information.” EX1006, ¶107; EX1007, 1. Thus, in my opinion, a POSA would have understood that in the Wang777-Wang060-Herley682-Seet combination, the raw play stream of identified and unidentified samples is ultimately converted to a play list in the form of the log file. EX1001, 10:34-40, 10:46-50 (describing play list broadly).

D. Dependent Claim 4

[4] <i>wherein the raw play stream is converted to a mathematical representation of the raw play stream.</i>
--

197. It is my opinion that Wang777-Wang060 in combination with Herley682 and Seet discloses this element.

198. As I noted above, in the Wang777-Wang060-Herley682-Seet combination, a raw play stream of identified samples (e.g., “user sample

identity[n]” and “user sample identity[n+1]” where “[n] is the nth sample in time”) and unidentified samples is created. *See supra* Section X.B. For example, Wang060’s candidate list includes the identified samples (e.g., sound_IDs, landmarks). EX1006, ¶81; EX1007, 1. Wang060 explains that there are a variety of techniques for computing landmarks, including Power Norm, which is a special case of the more general Spectral Lp Norm in which p=2. EX1006, ¶[0052]; EX1007, 1. “The general Spectral Lp Norm is calculated at each time along the sound signal by calculating a short-time spectrum, for example via a Hanning-windowed Fast Fourier Transform (FFT).” EX1006, ¶[0052]; EX1007, 1.

$$\begin{array}{c} \text{Candidate List} \\ \left[\begin{array}{c} (\text{landmark}_1, \text{landmark}^*_1, \text{sound_ID}_1) \\ \vdots \\ (\text{landmark}_k, \text{landmark}^*_j, \text{sound_ID}_j) \\ \vdots \\ (\text{landmark}_m, \text{landmark}^*_r, \text{sound_ID}_n) \end{array} \right] \end{array}$$

FIG. 9B

EX1006, FIG. 9B.

199. After the candidate list is compiled, it is processed by segmenting according to sound_ID using any one of a large number of algorithms. EX1006, ¶82. “The result of this process is a list of candidate sound_IDs, each of which has

a scatter list of pairs of sample and file landmark timepoints” as shown in FIG. 9C.

EX1006, ¶83; EX1007, 1.

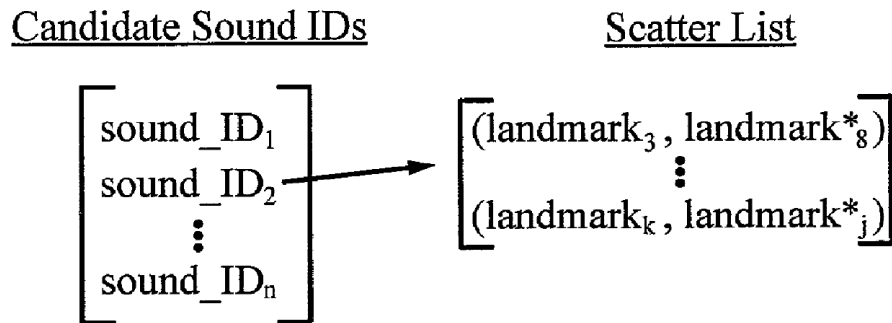


FIG. 9C

EX1006, FIG. 9C.

200. The scatter list for each candidate sound_ID is then analyzed to determine whether a given sound_ID has a linear correspondence in the form of: $\text{landmark}^*_n = m * \text{landmark}_n + \text{offset}$, where m is the slope, landmark is the timepoint within the sample, landmark^*_n is the corresponding timepoint within the sound recording indexed by sound_ID, and offset is the time offset into the sound recording corresponding to the beginning of the sound sample. EX1006, ¶¶82-85. “Landmark pairs that can be fit with the above equation for particular values of m and offset are said to be linearly related,” which “identifies the correct sound file with high probability.” EX1006, ¶85; EX1007, 1. In other words, Wang060 clearly discloses that a stream is converted to a mathematical representation of the same stream.

201. It is my opinion that a POSA would have thus understood that the scatter list is a mathematical representation of the candidate list (raw play stream). *See* EX1001, 8:44-49 (describing the mathematical representation without detail).

XI. CLAIMS 1-13 OF THE '622 PATENT ARE OBVIOUS UNDER 35 U.S.C. § 103 OVER PITMAN, SCHREMPP, AND WELLS

202. In my opinion, a POSA would have found claims 1-13 of the '622 obvious in light of the disclosure from Pitman (EX1013), Schrempp (EX1014), and Wells (EX1015).

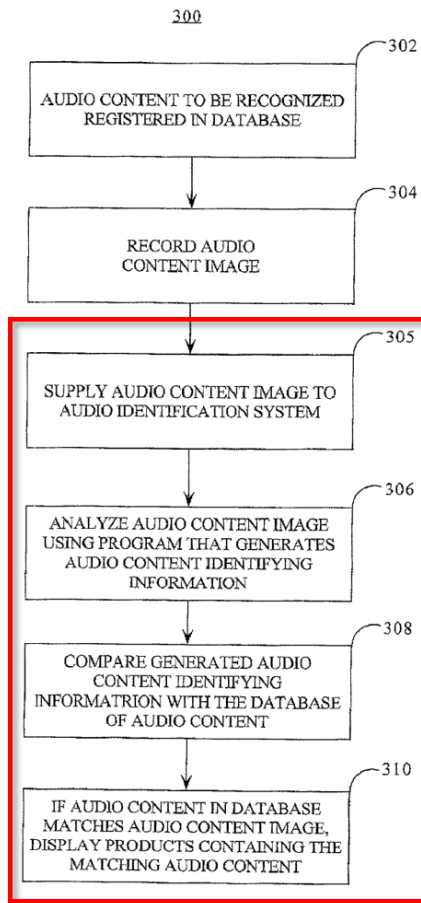
A. Motivation to Combine Pitman with Schrempp and Wells

203. It is my opinion that a POSA would have found it obvious to combine Pitman with Schrempp and Wells for multiple reasons.

204. Pitman is directed to a method for “determining whether audio identifying information generated for an audio content image matches audio identifying information in an audio content database” EX1013, Abstract. To generate audio content database 218, audio content from products available for purchase (e.g., songs from albums or concert videos) is registered along with identifying information (302). EX1013, ¶[0030], FIGs. 2A, 3. The identifying information for a particular piece of content may be a searchable, unique signature (e.g., a set of keys). EX1013, ¶[0030]. The identifying information may alternatively be a fingerprint, a watermark or an identification number. EX1013, ¶[0031], FIG. 3.

205. To identify audio content, a user records an image of a segment of the audio content (304) and supplies the image to an audio identification system (305). EX1013, ¶[0032]. A content identifier module 318 of the audio identification system analyzes the image and generates a unique feature signature (e.g., a set of keys) for the image (306). EX1013, ¶[0033], FIGs. 2A, 3. The set of keys for the image are compared to the keys stored in the database 216 (308), and content identifier module 218 “determines whether or not the audio content image matches any of the audio content in the audio content database.” *Id.* If a match between the audio content image and audio content in the database is found, the audio identification system displays an identification of the matching audio content (e.g., the artist and title) (310). EX1013, ¶[0035], FIG. 3. The audio identification system also displays a list including products related to the identified song (e.g., a cassette single of the song, an album containing the song, a music video for the song, a concert video including a live performance of the song, a movie in which the song appears, and/or a movie soundtrack including the song.). EX1013, ¶[0035]. Using the audio identification system, the user can select the product(s) in the list and purchase the product(s). EX1013, ¶[0036].

FIG. 3



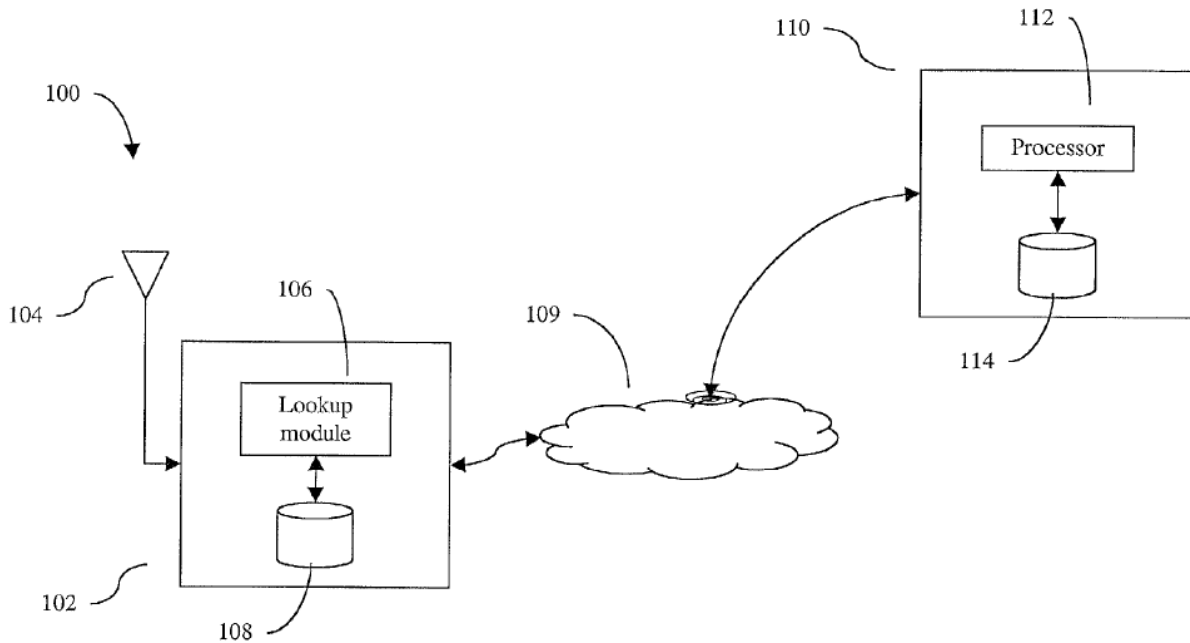
EX1013, FIG. 3 (annotated).

206. As such, Pitman describes identifying a song and displaying the content identification results to a user, but is silent on the format of that information. In my opinion, a POSA would have known that this information would have been presented in any number of formats and would have looked to prior art, such as Schrempp, which discloses as much.

207. Like Pitman, Schrempp is directed to capturing audio and identifying the audio (e.g., song titles) by matching the audio to a database. EX1014,

¶¶[0007]-[0008]. Schrempp also explains in its Background section that the generation of playlists of streamed and identified media content is well-known. EX1014, ¶[0004]. Such playlists document digital media, audio works, or video works performed/transmitted at a particular time over a particular medium (e.g., broadcast or internet stream). EX1014, ¶[0005].

208. An example system for generating a playlist is described with reference to Figure 1 of Schrempp. In particular, Schrempp discloses a playlist generation system 100 that includes a detection module 102 comprising an antenna 104 for receiving broadcast signals, which are provided for analysis to a lookup module 106. EX1014, ¶¶[0006]-[0007], FIG. 1. Lookup module 106 identifies content of the broadcast signal by comparing it to a database 108 of known reference signatures. EX1014, ¶[0007], FIG. 1. If a match is made, detection module 102 keeps a record of the match. EX1014, ¶[0008], FIG. 1. Detection module 102 may keep a record of all matches made during a predetermined period of time (e.g., a record of song titles detected during a 24-hour period). EX1014, ¶[0008], FIG. 1. System 100 also includes a playlist server 110, which receive the titles of the identified songs from detection module 102 and generates a playlist based on the titles. EX1014, ¶[0009], FIG. 1.



EX1014, FIG. 1.

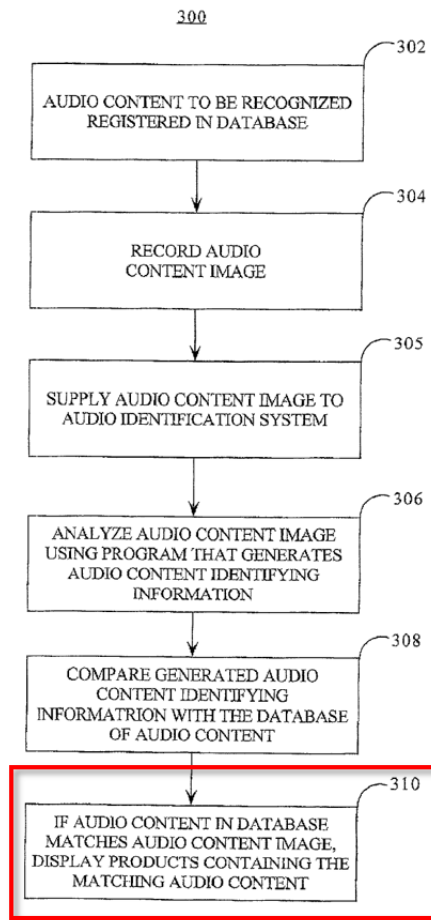
209. Based on the teachings of Schrempp, it is my opinion that it would have been obvious to a POSA to arrange the identified songs in Pitman into a play stream of content identification results corresponding to a sequence of data samples. It is also my opinion that modifying Pitman to incorporate Schrempp's teaching would have required minor modifications (e.g., modifying Pitman's software). Particularly, as Pitman already describes the generation of a list including products for purchase that are related to an identified song, EX1013, ¶¶[0035]-[0036], the modification of Pitman to generate a list of identified songs would be minimal.

210. In my opinion, a POSA would have been further motivated to

combine Pitman and Schrempp because both are in the same field of content identification. Additionally, a POSA would have combined the teaching of Pitman and Schrempp by known methods, and the result of the combination would have been predictable to a POSA (e.g., generating a play stream of the identified content). The proposed combination would have involved nothing more than incorporating a well-known concept (e.g., play streams) into a well-known system (e.g., the audio identification system of Pitman).

211. Pitman also explains “[i]f the listener enjoys the song” they may desire “to purchase a recording of the song or a product containing the song (e. g., an album or video).” EX1013, ¶[0003]. Pitman thus not only can identify a song, the system can also identify one or more products relating to the identified audio content. EX1013, ¶¶[0030]-[0031], [0035]. As described above, such products include “a cassette single of the song, an album containing the song, a music video for the song, a concert video including a live performance of the song, a movie in which the song appears, and a movie soundtrack including the song.” EX1013, *see also id.*, ¶[0071] (“the present invention provides systems and methods in which audio content is automatically identified and then one or more products containing (or relating to) the identified audio content are presented for purchase”).

FIG. 3



EX1013, FIG. 3 (annotated).

212. It is my opinion that a POSA would have thus understood that Pitman teaches identifying a channel (e.g., an album containing the song) corresponding to the data samples. Pitman, however, does not explicitly disclose using a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples. However, it is my opinion that this would have been obvious in view of Wells.

213. Wells, like Pitman, discloses techniques for identifying streams of

music. EX1015, ¶[0259], FIG. 12. For instance, with reference to Figure 12 of Wells, a stream of music is received (1601), and fingerprints are extracted for every frame of the stream (where each frame is typically between 1-5 seconds) (1602). EX1015, ¶[0261], FIG. 12. Each extracted fingerprint is matched (1603) against a database 1600 including fingerprints of “M unique songs” (referred to as the “original” song set). EX1015, ¶¶[0260], [0262], FIG. 12. If one or more fingerprints match (1604), additional criteria to identify the stream is applied (1606). EX1015, ¶[0264], FIG. 12. In particular, Wells teaches that the stream is identified only if P (e.g., 10) “consecutive candidate fingerprints match with the same song in the DB”, where greater the value of P, the less likely the possibility of a mismatch. EX1015, ¶[0264].

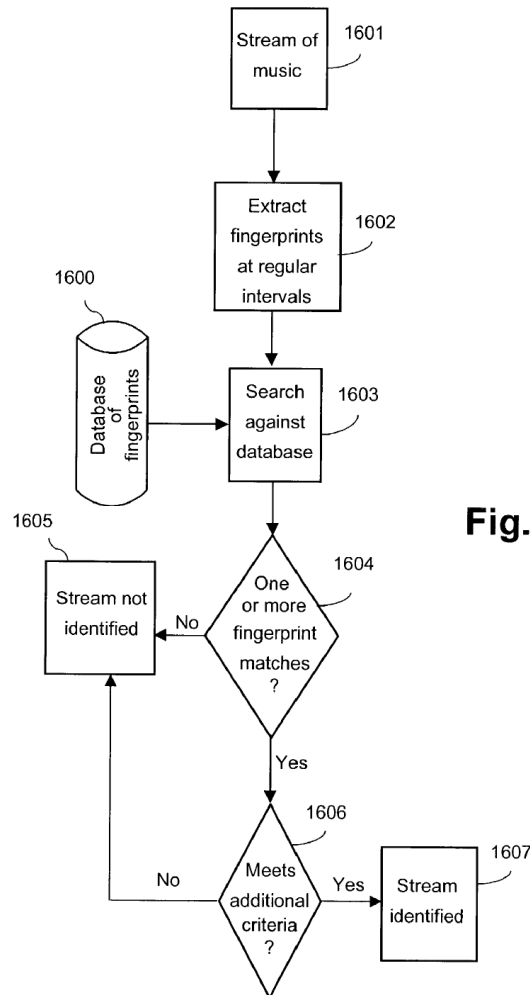


Fig. 12

EX1015, FIG. 12.

214. That is, Wells teaches the use of multiple consecutive content identifiers to identify a match.

215. Based on the teachings of Wells, it is my opinion that a POSA would have been motivated to modify Pitman-Schrempp (as combined above) to incorporate Wells' teaching of identifying a song when multiple sequential song segments correspond to the same song and also identifying an album or movie when multiple sequential songs match the album or movie. Doing so would have

reduced the possibility of an incorrect song identification (as described in paragraph [0264] of Wells), and by extension reduce the possibility of an incorrect album (or movie in which the songs appears) identification (as described in paragraph [0035] of Pitman). It is my opinion that modifying Pitman-Schrempp to incorporate Wells' teaching would have required minor modifications (e.g., modifying Pitman's software).

216. It is also my opinion that a POSA would have been further motivated to combine Pitman/Schrempp and Wells because all are in the same field of content identification. Additionally, it is my opinion that a POSA would have combined the teaching of Pitman-Schrempp and Wells by known methods, and the result of the combination would have been predictable to a POSA (e.g., identifying songs and a corresponding album or movie in which the songs appear when multiple songs correspond to the same album or movie). EX1015, ¶[0264]; EX1013, ¶[0035]. The proposed combination would have involved nothing more than incorporating a well-known concept (e.g., content identification) into a well-known system (e.g., the audio identification system of Pitman).

B. Independent Claim 1

1. Preamble [1.P]

[1.P] <i>A media measurement method comprising:</i>

217. To any extent that the preamble is limiting, it is my opinion that

Pitman discloses a “media measurement method.”

218. For example, Pitman is directed to a method for “determining whether audio identifying information generated for an audio content image matches audio identifying information in an audio content database.” EX1013, Abstract. In particular, a “method for allowing a listener to automatically identify audio content so that a product containing the audio content can be purchased.” EX1013, ¶[0006], *see also id.*, ¶¶[0003], [0071], claims 1-15.

219. Thus, Pitman’s method for automatically identifying audio content is a “media measurement method” just like that described by the ’622 patent. EX1001, 2:3-5.

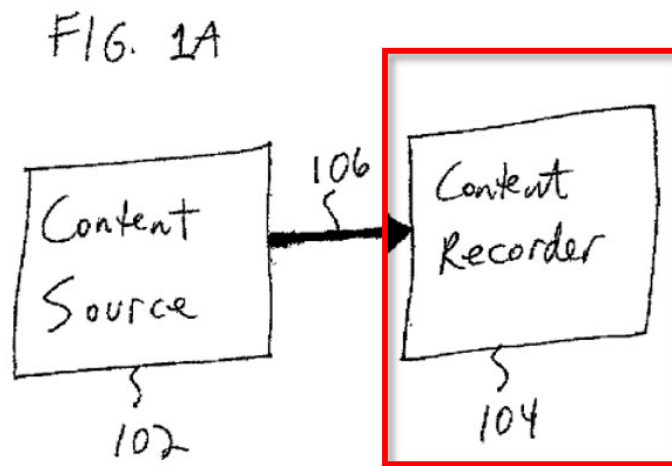
2. Element [1.1]

[1.1] <i>using one or more computers to generate a play stream of content identification results corresponding to a sequence of data samples, the data samples comprising representations of audio data captured at a media monitoring device,</i>

220. It is my opinion that Pitman in combination with Schrempp renders obvious this element.

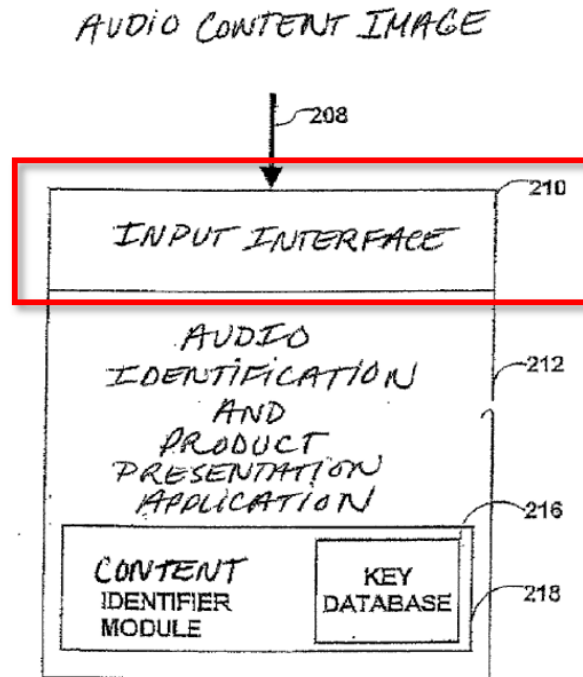
221. Pitman’s system initially captures data samples comprising representations of audio data captured at a media measurement device. Figure 1A, for example, illustrates a content source 102 and a content recorder 104. Content source 102 is “any source that supplies audio content” such as a radio station).

EX1013, ¶¶[0021], [0023]. Content recorder 104 is any device, such as a media measurement device, that can receive and record an image of the audio content from content source 102. EX1013, ¶[0021].



EX1013, FIG. 1A.

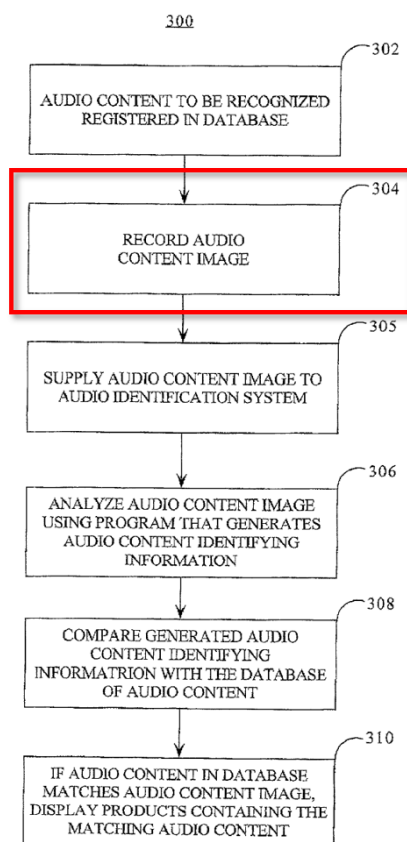
222. Figure 2A illustrates a block diagram of the system, which includes an input interface 210 for receiving content. EX1013, ¶[0025]. The content includes an analog signal that is modulated by the audio content, a digital datastream, or a computer file. EX1013, ¶[0025].



EX1013, FIG. 2A (annotated).

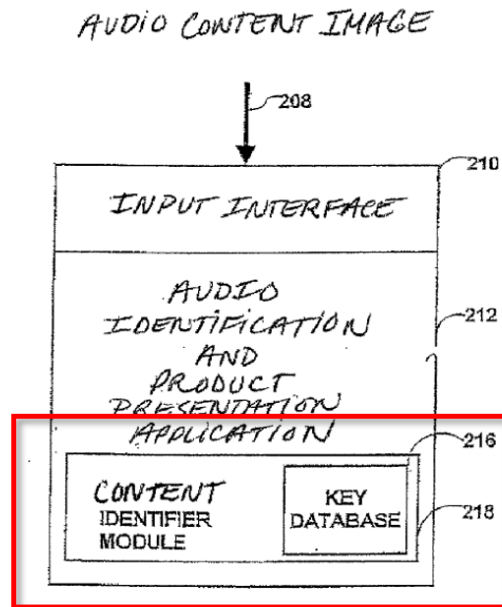
223. Pitman explains that a user wanting to identify audio content may record “an image of at least a portion of the audio content” (304). EX1013, ¶¶ [0032], [0037], FIGs. 3-4A. That audio content may correspond to a segment of a song. EX1013, ¶¶ [0026]-[0027].

FIG. 3



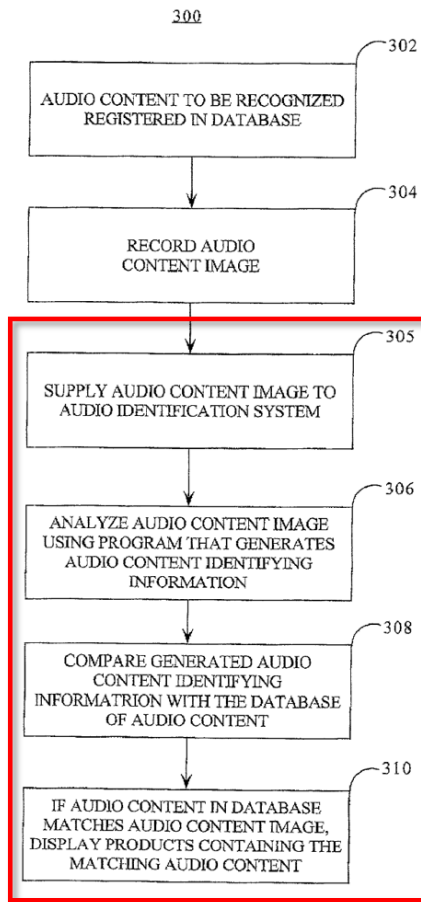
EX1013, FIG. 3 (annotated).

¶[0033], FIGs. 2A, 3. The keys stored in key database 216 correspond to audio content from products available for purchase (e.g., songs from albums or concert videos). EX1013, ¶[0030]. In the database, there is a set of keys for each song and a section of a song will have a corresponding subset of keys. EX1013, ¶[0026]. If a match is found, the audio identification system displays an identification of the matching audio content (e.g., the artist and title) (310). EX1013, ¶[0035], FIG. 3. The audio identification system also displays a list including products related to the identified song (e.g., a cassette single of the song, an album containing the song, a music video for the song, a concert video including a live performance of the song, a movie in which the song appears, and/or a movie soundtrack including the song.). EX1013, ¶[0035]. Using the audio identification system, the user can select the product(s) in the list and purchase the product(s). EX1013, ¶[0036].



EX1013, FIG. 2A (annotated).

FIG. 3



EX1013, FIG. 2 (annotated).

226. Pitman illustrates its process for identifying audio in Figures 4A-6, 8. EX1013, ¶¶[0037]-[0059], [0063]-[0070], FIGs. 4A-6, 8. That is, Pitman discloses using one or more computer to generate content identification results (song titles) corresponding to a sequence of data samples (audio samples) by searching a database of known content.

227. While Pitman describes identifying a song and displaying the content

identification results to a user, Pitman is silent on the format of that information.

However, Schrempp explains in its Background section that the generation of

playlists of streamed and identified media content is well known. EX1014,

¶[0004]. Such playlists document digital media, audio works, or video works

performed/transmitted at a particular time over a particular medium (e.g., broadcast

or internet stream). EX1014, ¶[0005]. In particular, Schrempp discloses a playlist

generation system 100 that includes a detection module 102 comprising an antenna

104 for receiving broadcast signals, which are provided for analysis to a lookup

module 106. EX1014, ¶[0007], FIG. 1. Lookup module 106 identifies content of

the broadcast signal by comparing it to a database 108 of known reference

signatures. EX1014, ¶[0007], FIG. 1. If a match is made, detection module 102

keeps a record of the match. EX1014, ¶[0008], FIG. 1. Detection module 102 may

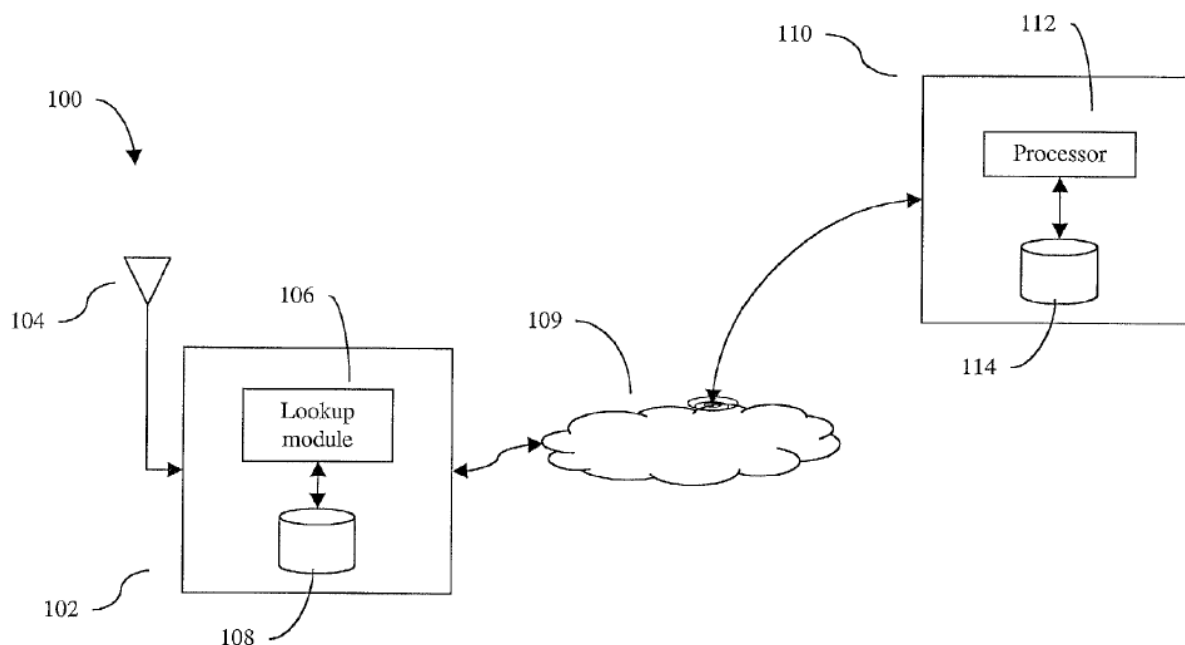
keep a record of all matches made during a predetermined period of time (e.g., a

record of song titles detected during a 24-hour period). EX1014, ¶[0008], FIG. 1.

System 100 also includes a playlist server 110, which receive the titles of the

identified songs from detection module 102 and generates a playlist based on the

titles. EX1014, ¶[0009], FIG. 1.



EX1014, FIG. 1.

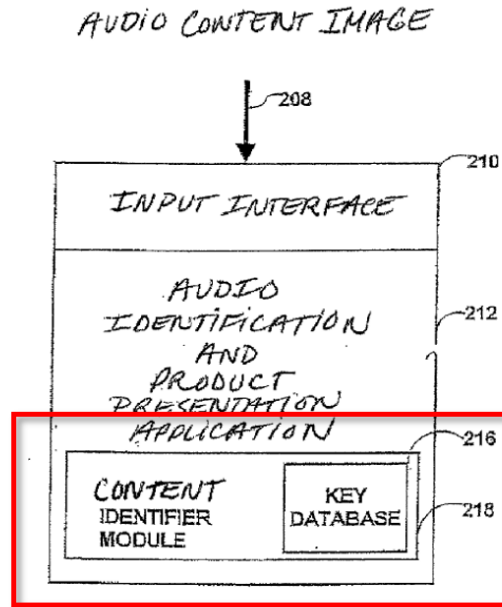
228. It is my opinion that it would have been obvious to a POSA to arrange the identified songs in Pitman into a play stream of content identification results corresponding to a sequence of data samples in light of Schrempp. *See supra* Section XI.A. Particularly, as Pitman already describes the generation of a list including products for purchase that are related to an identified song, EX1013, ¶¶[0035]-[0036], the modification of Pitman to generate a list of identified songs would be minimal.

3. Element [1.2]

[1.2] wherein content identifiers of the play stream's content identification results are obtained by using corresponding data samples to search a computerized database of known content; and

229. In my opinion, Pitman in combination with Schrempp renders obvious this element.

230. As I explained above, Pitman's processes can identify a song (or segments of a song) by analyzing each segment to generate a set of keys (a signature, 306). EX1013, ¶¶[0030]-[0036]. Particularly, in Pitman, a content identifier module 318 of its audio identification system analyzes an image and generates a unique feature signature (e.g., a set of keys) for the image (306). EX1013, ¶[0033], FIGs. 2A, 3. The set of keys for the image are compared to the keys stored in the database 216 (308), and content identifier module 218 "determines whether or not the audio content image matches any of the audio content in the audio content database." EX1013, ¶[0033], FIGs. 2A, 3. If a match between the audio content image and audio content in the database is found, the audio identification system displays an identification of the matching audio content (e.g., the artist and title) to the user (310). EX1013, ¶[0035], FIG. 3.



EX1013, FIG. 2A (annotated).

4. Element [1.3]

[1.3] *in response to obtaining the content identifiers, using one or more computers to utilize a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples,*

231. In my opinion, Pitman in combination with Schrempp and Wells renders obvious this element.

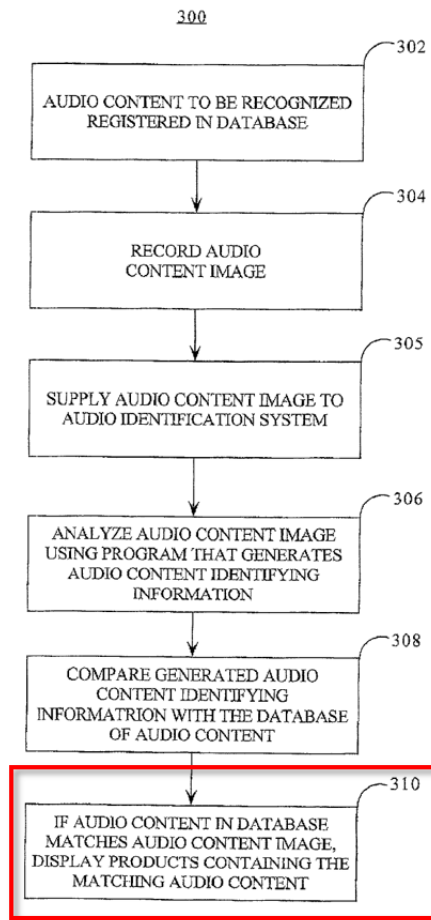
232. As explained above, Pitman's processes can identify a song (or segments of a song) by analyzing each segment to generate a set of keys (a signature, 306). EX1013, ¶¶[0030]-[0036]. Particularly, in Pitman, a content identifier module 318 of its audio identification system analyzes an image and generates a unique feature signature (e.g., a set of keys) for the image (306).

EX1013, ¶[0033], FIGs. 2A, 3. The set of keys for the image are compared to the keys stored in the database 216 (308), and content identifier module 218

“determines whether or not the audio content image matches any of the audio content in the audio content database.” EX1013, ¶[0033], FIGs. 2A, 3. If a match between the audio content image and audio content in the database is found, the audio identification system displays an identification of the matching audio content (e.g., the artist and title) to the user (310). EX1013, ¶[0035], FIG. 3.

233. Pitman also explains “[i]f the listener enjoys the song” they may desire “to purchase a recording of the song or a product containing the song (e. g., an album or video).” EX1013, ¶[0003]. Pitman thus not only can identify a song, the system can also identify one or more products relating to the identified audio content. EX1013, ¶¶[0030]-[0031], [0035]. “For example, if the audio content image is a portion of a song, the list displayed to the listener could include products such as a cassette single of the song, an album containing the song, a music video for the song, a concert video including a live performance of the song, a movie in which the song appears, and a movie soundtrack including the song.” EX1013, ¶¶[0030]-[0031], [0035], *see also id.*, ¶[0071] (“[T]he present invention provides systems and methods in which audio content is automatically identified and then one or more products containing (or relating to) the identified audio content are presented for purchase . . .”).

FIG. 3



EX1013, FIG. 3 (annotated).

234. It is my opinion that a POSA would have thus understood that Pitman teaches identifying a channel (e.g., an album containing the song) corresponding to the data samples. EX1013, ¶[0035]. However, while Pitman does not explicitly disclose using a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples, Wells does.

235. Wells, like Pitman, discloses (with reference to Figure 12 of Wells) a

stream of music is received (1601), and fingerprints are extracted for every frame of the stream (where each frame is typically between 1-5 seconds) (1602).

EX1015, ¶[0261], FIG. 12. Each extracted fingerprint is matched (1603) against a database 1600 including fingerprints of “M unique songs” (referred to as the “original” song set). EX1015, ¶¶[0260], [0262], FIG. 12. If one or more fingerprints match (1604), additional criteria to identify the stream is applied (1606). EX1015, ¶[0264], FIG. 12. In particular, Wells teaches that the stream is identified only if P (e.g., 10) “consecutive candidate fingerprints match with the same song in the DB”, where greater the value of P, the less likely the possibility of a mismatch. EX1015, ¶[0264]. That is, Wells teaches the use of multiple consecutive content identifiers to identify a match.

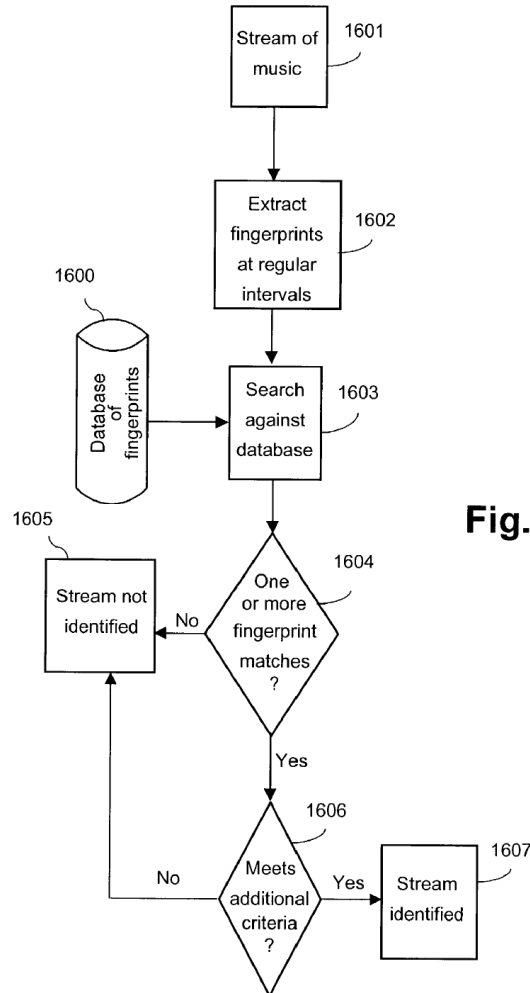


Fig. 12

EX1015, FIG. 12.

236. In addition to identifying segments of songs, Pitman “provides a means for identifying a song.” EX1013, ¶26. And Pitman teaches that not just a single song but all songs on an album or in a movie may be registered in the audio content database for presentation to a user for sale. EX1013, ¶¶30-31. As noted above, Schrempp teaches keeping a record of all matches made during a predetermined period of time (e.g., a record of song titles detected during a 24-hour period). EX1014, ¶¶[0006]-[0008], FIG. 1. Moreover, Wells teaches

identifying a source by looking at content matches over time. EX1015, ¶¶[0259]-[0264].

237. In my opinion, it would have thus been obvious to a POSA to use the identification of two (or more) sequential songs from the stored record of song titles to determine whether those songs correspond to a particular album or movie in the Pitman-Schrempp-Wells combination. Otherwise, if a song appears on multiple albums and in multiple movies, the system would not be able to identify the album that the user is actually listening to but instead could only provide every possible album or movie to the user for purchase; potentially frustrating the user and leading them to not purchase a product.

5. Element [1.4]

[1.4] <i>the at least two different obtained content identifiers identifying different media content items.</i>
--

238. It is my opinion that Pitman in combination with Schrempp and Wells renders obvious this element.

239. As noted above, it would have been obvious to identify the songs and corresponding album or movie in which the songs appear when sequential songs correspond to the same album or movie in Pitman, in view of Wells. Doing so would have reduced the possibility of an incorrect song identification. *See supra* Section XI.B.4.

C. Independent Claim 6 and 10

1. [6.P], [10.P]

[6.P] *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 instructions comprising instructions for:*

[10.P] *A computer system comprising one or more computers configured by computer code stored in the one or more computers, the computer code comprising instruction executable by one or more processors of the one or more computers, the instructions comprising instructions for:*

240. It is my opinion that Pitman discloses this preamble to the extent it is limiting, for the reasons set forth above for element [1P].

241. In particular, Pitman states that the “present invention can be realized in hardware, software, or a combination of hardware and software. Any kind of computer system—or other apparatus adapted for carrying out the methods described herein—is suited.” EX1013, ¶[0074].

242. Moreover, Pitman discloses that the “present invention can also be embedded in a computer program product, which includes all the features enabling the implementation of the methods described herein, and which—when loaded in a computer system—is able to carry out these methods. Computer program means or computer program in the present context mean any expression, in any language, code or notation, of a set of instructions intended to cause a system having an

information processing capability to perform a particular function either directly or after either or both of the following a) conversion to another language, code or, notation; and b) reproduction in a different material form.” EX1013, ¶[0075].

2. [6.1], [10.1]

243. It is my opinion that Pitman discloses and/or renders obvious this element in view of Schrempp for at least the reasons set forth above for reasons set forth above for element [1.1].

3. [6.2], [10.2]

244. It is my opinion that Pitman discloses and/or renders obvious this element in view of Schrempp for at least the reasons set forth above for reasons set forth above for element [1.2].

4. [6.3], [10.3]

245. It is my opinion that Pitman discloses and/or renders obvious this element in view of Schrempp and Wells for at least the reasons set forth above for reasons set forth above for element [1.3].

5. [6.4], [10.4]

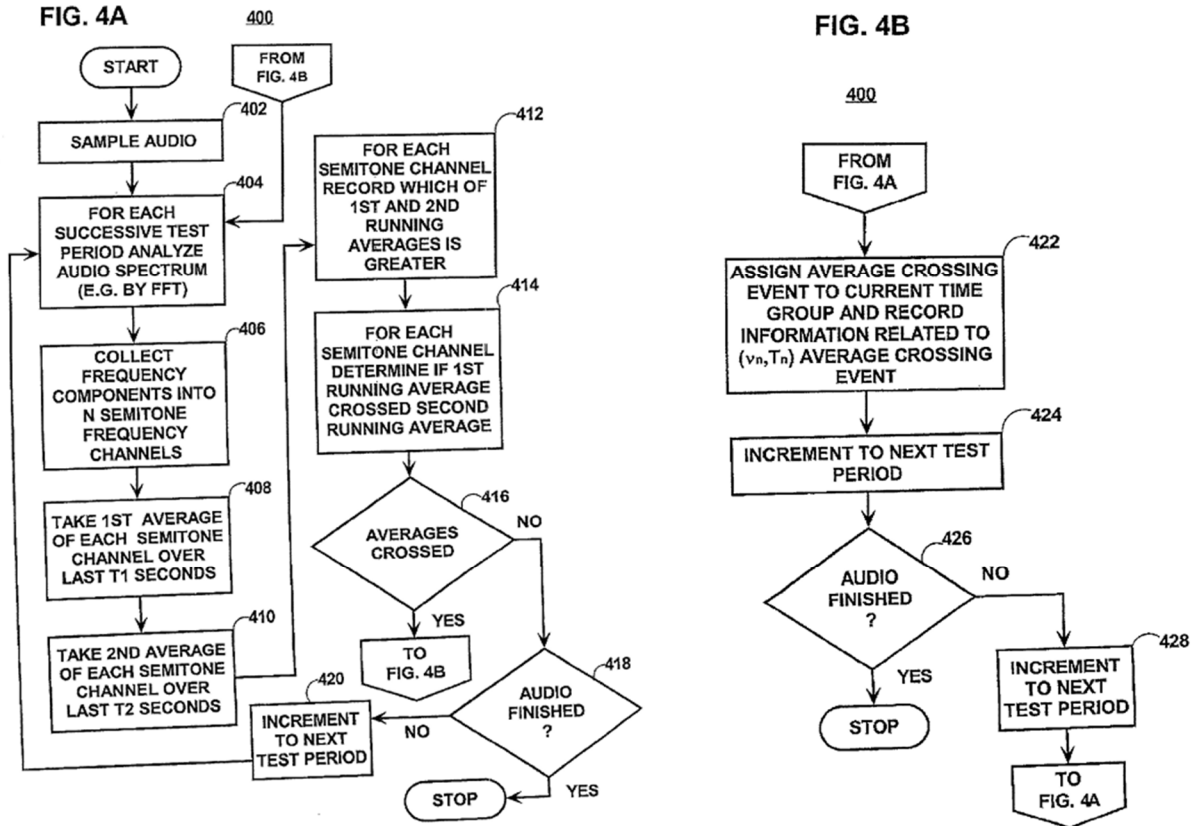
246. It is my opinion that Pitman discloses and/or renders obvious this element in view of Schrempp and Wells for at least the reasons set forth above for reasons set forth above for element [1.4].

D. Claim 2, 7, and 11

[2], [7], [11] *converting a raw play stream to a format usable for searching channel data to identify a channel to associate with content of the raw play stream.*

247. It is my opinion that Pitman discloses this element.

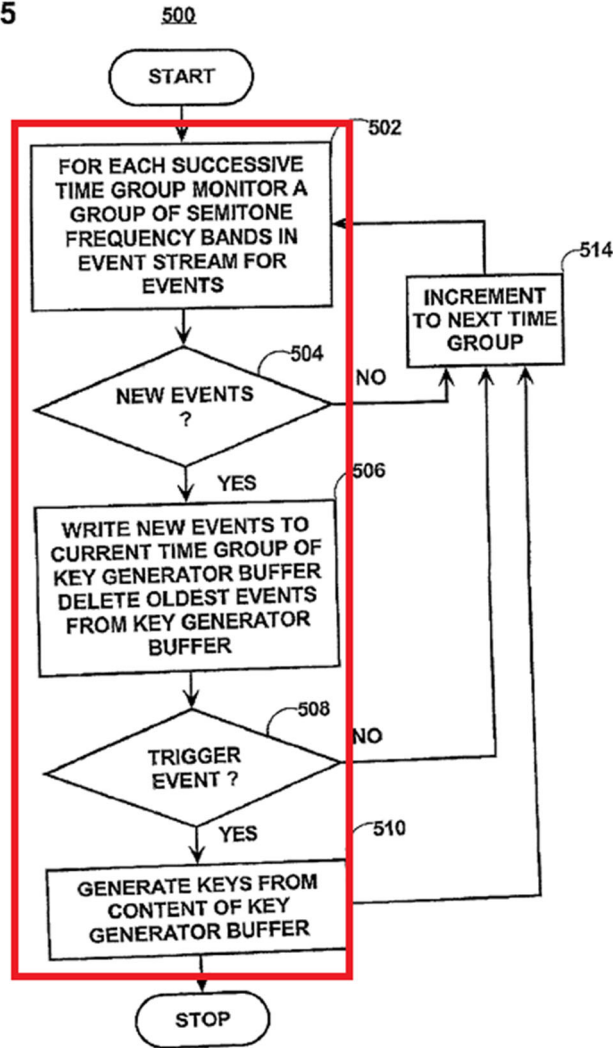
248. Pitman discloses a multistep process for identifying the sampled content. For each successive test period (e.g., ranging from 1/43 of a second to 1/10.75 of a second), the spectrum of the sampled signal is computed, for example, using a fast Fourier transform (FFT) algorithm and then analyzed to generate “event information.” (404-428). EX1013, ¶¶ [0038]-[0049], FIGs. 4A-4B. The event information includes the “time group for the event” and the “test period for the event.” EX1013, ¶[0048]. The events are then grouped into “time groups,” which are ¼ to 2 seconds and the event information is then stored in memory in the form of an event stream. EX1013, ¶¶[0048], [0050].



EX1013, FIGs. 4A-4B.

249. Next, Pitman's process analyses each time group of events spanning $\frac{1}{4}$ to 2 seconds and, depending on the analysis, generates keys for that time group of events. EX1013, ¶¶[0051]-[0054], FIG. 5.

FIG. 5

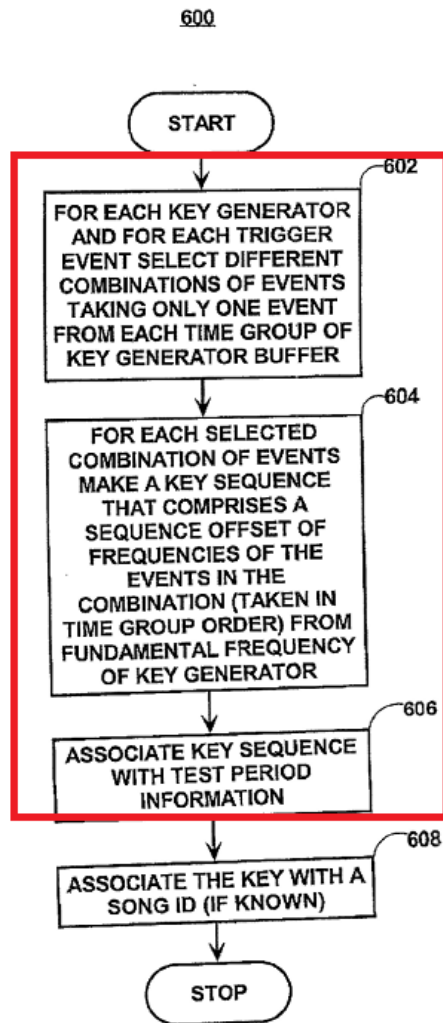


EX1013, FIG. 5 (annotated).

250. Figure 6 of Pitman illustrates the process of generating the keys in more detail, which corresponds to step 510 in Figure 5. EX1013, ¶¶[0055]-[0058]. In particular, for each key generator and trigger event, different combinations of events from the key generator buffer are selected (602), where each combination includes only one event from each time group. EX1013, ¶[0055], FIG. 6. For each selected combination of events, a key sequence is composed that includes a

sequence of numerical values of frequency offsets (relative to the main key generator frequency) (604). EX1013, ¶[0057], FIG. 6. Test period information (e.g., a sequence number for the test period of the trigger event) is associated with the key sequence (606). EX1013, ¶[0057], FIG. 6.

FIG. 6



EX1013, FIG. 6 (annotated).

251. While Figure 6 illustrates a step of associating the key with a song (608), that step is only performed when generating the database of known content.

EX1013, ¶[0058].

252. Figure 8 of Pitman illustrates “a flow diagram of a song identification process that uses the keys generated in the processes of FIGS. 5 and 6” using a song database. EX1013, ¶[0063]. Pitman explains that the song database is a keytable and that each includes a key sequence in the first column, a time group value in the second column, and a song ID (e.g., title) in the third column:

key sequence	time group value	song ID (e.g., title)

EX1013, ¶¶[0027], [0058], [0063].

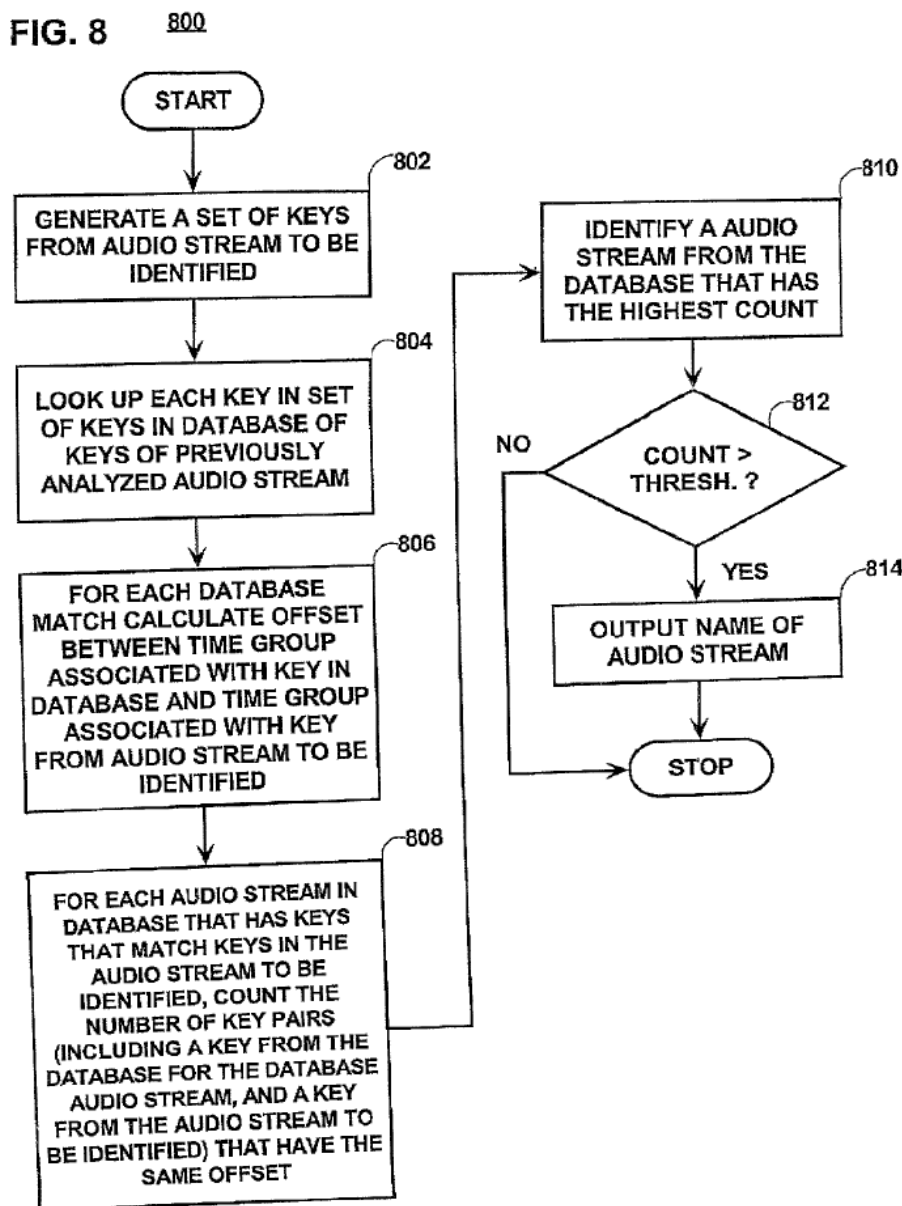
253. Pitman explains that after the keys are generated (802), each key in the set of keys is looked up in a song database (i.e., key database 218) including keys for a plurality of different songs (804). EX1013, ¶[0064], FIGs. 2A, 8. For each key in the database that matched key(s) in the song database, an offset is calculated by “taking the difference between a test period associated with the key being looked up and a test period associated with each matching key in the song database” (806). EX1013, ¶[0064], FIG. 8. The offsets are collected into offset time groups (e.g., equal to 2 to 10 test periods). EX1013, ¶[0065], FIG. 8. For each offset time group and each song that has keys that match keys in the song to be identified, the number of matching keys having the same time group offset is

counted (410). EX1013, ¶[0066], FIG. 8. For instance, Table 1 of Pitman visualizes how the number of matching keys are tracked. As shown in Table 1, the first column indicates different song titles, the second column indicates a value of the offset time group in which matching keys were found, and the third column includes the counts of the number of key matches corresponding to a particular song within a same offset time group. EX1013, ¶[0066].

TABLE 1		
SONG TITLE	OFFSET VALUE (UNITS OF TIME GROUP INTERVAL)	COUNT OF KEY SEQUENCE MATCHES FOR THIS SONG AND WITH THIS OFFSET VALUE
Title1	3	1
Title1	4	1
Title2	2	2
Title2	3	107
Title3	5	1
Title2	8	1

EX1013, Table 1.

254. If the song to be identified is in the database, then one particular time group offset value will accumulate a high count. EX1013, ¶[0066]. In the example shown in Table 1, Title2 has a count of 107 for an offset time group of 3. EX1013, ¶[0066]. The song having the highest count of matching keys within the same offset is identified (810), and the name of the identified song is output (814). EX1013, ¶¶[0068]-[0069], FIG. 8.



EX1013, FIG. 8

255. It is my opinion that a POSA would have thus understood that the data set of unprocessed audio samples in Pitman is converted to a format (e.g., keys) usable for searching channel data to identify a channel (e.g., album) to associate with content of the raw play stream. EX1013, ¶¶[0030], [0054], [0064] (stating that

an unprocessed audio content are represented (or converted) to a set of keys),
[0035] (stating that an album may be identified after identifying audio content).

E. Dependent Claims 3, 8, and 12

[3], [8], [12] *wherein the raw play stream is converted to a play list.*

256. It is my opinion that Pitman discloses this element.

257. As explained above, the data set of unprocessed audio samples in Pitman is converted to a set of keys, where each key includes a key sequence and associated test period. EX1013, ¶[0058]. The key sequence is used to identify audio. EX1013, ¶[0064]. Table 1 of Pitman represents a list of matching song titles that were determined by matching the key sequences. EX1013, ¶¶[0066]-[0067]. A POSA would understand that Table 1 represents a play lists, as it indicates a series of song titles that matched. *See supra* Section IX.D; *see also* EX1013, ¶¶[0026], [0030], [0033], [0057]-[0058], [0064].

F. Dependent Claim 4

[4] *The method of claim 2 wherein the raw play stream is converted to a mathematical representation of the raw play stream*

258. It is my opinion that Pitman discloses this element.

259. As explained above, the data set of unprocessed audio samples in Pitman is converted to a key sequence, which includes a sequence of numerical values of frequency offsets (relative to the main key generator frequency) for the

sequence of events from each combination.” *See supra* Section XI.D; *see also* EX1013, ¶[0057]. A POSA would have thus understood that Pitman converts the raw play stream to a mathematical representation of the raw play stream (i.e., a set of numerical values of frequency offsets).

G. Dependent Claims 5, 9, and 13

[5], [9], [13] *wherein using one or more computers to utilize a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples comprises searching channel information that includes a plurality of channel types including at least a broadcast media channel type and a non-broadcast playable media channel type.*

260. It is my opinion that Pitman discloses and/or renders obvious this element.

261. Pitman explains that the “content source 102 can be any source that supplies audio content, either alone or as part of a multimedia presentation.” EX1013, ¶[0020]. “For example, in one embodiment, the content source 102 is an FM radio station, the content recorder 104 is a stereo having an FM receiver and a cassette deck, the content player 112 is a portable cassette player, and the AIC 120 is a computer-powered kiosk at a record store.” EX1013, ¶[0023]. “In another embodiment, a recorded video playback device (e. g., VCR or DVD player) is the content source 102, a personal computer is both the content recorder 104 and

content player 112, and a remote web server is the AIC 120.” EX1013, ¶[0024].

Accordingly, in Pitman, audio content to be identified may originate from both broadcast media channel types (e.g., an FM radio station) or non-broadcast media channel types (e.g., a cassette deck, a VCR, a CD player, or a DVD player). Pitman further discloses that a non-broadcast media channel type (e.g., an album) is identified based on the identification of audio content. EX1013, ¶[0035].

262. In my opinion, a POSA would have thus understood that Pitman teaches searching channel information that includes a plurality of channel types including at least a broadcast media channel type (e.g., FM radio station) and a non-broadcast playable media channel type (e.g., DVD player).

XII. CONCLUSION

263. In signing this declaration, I recognize that the declaration will be filed as evidence in an *inter partes* review before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to cross-examination in the case and that cross-examination will take place within the United States. If cross-examination is required, I will appear for cross-examination within the United States during the time allotted.

264. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements were made with the

knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed on this 15th day of July 2024, in Mobile, AL.



Samuel H. Russ, Ph.D.