

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 IPR2024-01057
U.S. Patent No. 10,719,848

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 10,719,848**

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

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	IDENTIFICATION OF THE CHALLENGE (37 C.F.R. §42.104(B))	1
III.	LEVEL OF ORDINARY SKILL IN ART	2
IV.	OVERVIEW OF THE '848 PATENT	3
V.	CLAIM CONSTRUCTION	13
VI.	OVERVIEW OF THE APPLIED REFERENCES	14
	A. Overview of Herley682 (EX1005)	14
	B. Overview of Seet (EX1006)	17
	1. Seet is prior art as of February 17, 2004	17
	2. Overview	22
	C. Overview of Herley864 (EX1007)	27
	D. Overview of Wang060 (EX1008)	30
VII.	GROUND 1: CLAIMS 1-3, 5, 6, 8-11, 13, AND 14 ARE OBVIOUS OVER HERLEY682 AND SEET	32
	A. A POSA Would Have Combined Herley682 and Seet	32
	B. Independent Claim 1	35
	1. [1.P] “A media measurement method comprising:”	35
	2. [1.1] “(a) receiving over a network, at one or more computers, data for a sequence of video data samples, the video data samples comprising representations of video data captured at a media monitoring device, the plurality of video data samples being submitted over the network”	36
	3. [1.2] “(b) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers”	42
	4. [1.3] “(c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of video data samples”	43
	5. [1.4] “wherein: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective video	

	data samples in the sequence of video data samples and respective video data representations in the electronic database”	46
6.	[1.5] “a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective video data representation that is determined to be a respective likely match with a respective one of the video data samples; or (ii) an indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database”	47
7.	[1.6] “the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database”	48
8.	[1.7] “(d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data”	50
9.	[1.8] “(e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing”	55
C.	Independent Claim 9	55
1.	[9.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 one or more computers being coupled to a network to receive, over the network, data for a sequence of video data samples, the video data samples comprising representations of video data captured at a media monitoring device, the plurality of video data samples being submitted over the network, the instructions comprising instructions for:”	55

2. [9.1] “(a) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers” 57
3. [9.2] “(b) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of video data samples” 58
4. [9.3] “wherein: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective video data samples in the sequence of video data samples and respective video data representations in the electronic database” 58
5. [9.4] “a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective video data representation that is determined to be a respective likely match with a respective one of the video data samples; or (ii) an indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database” 58
6. [9.5] “the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database” 59
7. [9.6] “(c) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data” 59
8. [9.7] “(d) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing” 59
- D. Claims 2 and 10 (“wherein sample sequence data includes at least one of log times, content identifiers, and content offsets”) 60

E.	Claims 3 and 11 (“wherein: scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content identification results”).....	60
F.	Claims 5 (“wherein a clean play list is generated from the clean play stream”) and 13 (“wherein the computer program product further comprises instructions for generating a clean play list from the clean play stream”)	64
G.	Claims 6 and 14 (“wherein the plurality of video data representations include a plurality of video signatures”).....	65
H.	Claim 8 (“wherein a video data representation of the plurality of video data representations includes a signal usable for generating video energy corresponding to video energy captured at the media monitoring device”)	66
VIII.	GROUND 2: CLAIMS 1-6 AND 8-14 ARE OBVIOUS IN VIEW OF HERLEY682 AND HERLEY864.....	67
A.	A POSA Would Have Further Combined Herley864 with Herley682 and Seet	67
B.	Claims 1 and 9	69
C.	Claims 4 and 12 (“wherein: scrubbing the raw play stream further comprises using audience data as part of determining whether to change a result of the sequence of content identification results”).....	72
IX.	GROUND 3: CLAIMS 7 AND 15 ARE OBVIOUS IN VIEW OF HERLEY682 AND SEET (WITH OR WITHOUT HERLEY864) AND WANG060.....	74
A.	Claims 7 and 15 (“wherein a video data representation of the plurality of video data representations includes a plurality of signal landmarks”)	74
X.	PETITIONER IS UNAWARE OF ANY SECONDARY CONSIDERATIONS OF NON-OBVIOUSNESS	77
XI.	DISCRETIONARY DENIAL IS NOT APPROPRIATE.	77
A.	<i>Fintiv</i> does not support discretionary denial.....	77
B.	Denial under §325(d) is not appropriate.....	77
XII.	CERTIFICATION THAT THE PATENT MAY BE CONTESTED VIA <i>INTER PARTES</i> REVIEW BY THE PETITIONER AND STANDING (37 C.F.R. § 42.104(A))	78

XIII. MANDATORY NOTICES (37 C.F.R. § 42.8(A)(1)).....	78
XIV. CONCLUSION	81

EXHIBIT LIST

Exhibit No.	Description
1001	U.S. Patent No. 10,719,848 to Steuer et al. (“’848 patent”)
1002	File History of the ’848 Patent (“’848 Patent File History”)
1003	Declaration of Dr. Samuel Russ
1004	<i>Curriculum Vitae</i> of Dr. Samuel Russ
1005	U.S. Patent Application No. 2004/0260682 to Herley et al. (“Herley682”)
1006	U.S. Patent Application No. 2005/0193016 to Seet et al. (“Seet”)
1007	U.S. Patent No. 7,333,864 to Herley (“Herley864”)
1008	U.S. Patent Application No. 2002/0083060 to Wang et al. (“Wang060”)
1009	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”)
1010	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”)
1011	U.S. Patent No. 5,581,658 to O’Hagan et al. (“O’Hagan”)
1012	Maurico Hernandez and Salvatore Stolfo, “Real-world Data is Dirty: Data Clenasing and The Merge/Purge Problem, in Data Mining and Knowledge Discovery, 2, 9-37 (1998). (“Hernandez”)
1013	Kamakshi Lakshminarayan, Steven A. Harp, Tariq Samid, “Imputation of Missing Data in Industrial Databases,” Applied Intelligence 11, 259-275, (1999) (“Lakshminarayan”)

Exhibit No.	Description
1014	Eloi Batlle, Jame Masip, Enric Guaus, “Automatic Song Identification in Noisy Broadcast Audio,” Audiovisual Institute and Department of Technology, Barcelona, Spain, (2002). (“Batlle”)
1015	“Signature,” <i>Merriam-Webster Dictionary</i> , accessed at https://www.merriam-webster.com/dictionary/signature
1016	Interim Procedure for Discretionary Denials in AIA Post-Grant Proceeding with Parallel District Court Proceedings Memorandum, June 21, 2022.
1017	<i>Intentionally left blank.</i>
1018	<i>Intentionally left blank.</i>
1019	U.S. Provisional Application No. 60/545,681, filed February 17, 2004 (“Seet provisional”)
1020	U.S. Patent No. 8,510,768 to Steuer et al. (“’768 patent”)

I. INTRODUCTION

Roku, Inc. (“Roku” or “Petitioner”) petitions for *inter partes* review, seeking cancellation of claims 1-15 of U.S. Patent No 10,719,848 (“the ’848 patent”; EX1001), assigned to Anonymous Media Research Holdings LLC (“Anonymous Media Research” or “Patent Owner”).

II. IDENTIFICATION OF THE CHALLENGE (37 C.F.R. §42.104(B))

Roku requests IPR based on the grounds below.

Ground	35 U.S.C. Section	Claims	References
1	§ 103(a)	1-3, 5, 6, 8-11, 13, 14	Herley682, Seet
2	§ 103(a)	1-6, 8-14	Herley682, Seet, Herley864
3	§ 103(a)	7, 15	Herley682, Seet (with or without Herley864), Wang060

The ’848 patent was filed on October 22, 2019, and claims priority through a chain of divisionals and continuations to U.S. Appl. No. 11/139,330, filed on May 26, 2005 (now abandoned), and U.S. Prov. Appl. No. 60/574,836 filed May 27, 2004.¹

The references qualify as prior art:

¹ Petitioner reserves the right to challenge the priority date in other proceedings.

Herley682 (EX1005) was published on December 23, 2004, and filed on June 19, 2003. **Herley682** is therefore prior art under 35 U.S.C. § 102(e) (pre-AIA) regardless of the '848 patent's priority date.

Seet (EX1006) published on September 1, 2005, was filed on February 17, 2005, and claims priority to provisional application No. 60/545,681 ("the '681 application") filed on February 17, 2004. **Seet** is therefore prior art under 35 U.S.C. § 102(e) (pre-AIA) as of February 17, 2004 because it is entitled to its provisional filing date for prior art purposes (*see infra* Section VI.B.1).

Herley864 (EX1007) issued on February 19, 2008, and was filed on June 22, 2002. **Herley864** is therefore prior art under 35 U.S.C. § 102(e) (pre-AIA) regardless of the '848 patent's priority date.

Wang060 (EX1008) was published on June 27, 2002. **Wang060** is therefore prior art under 35 U.S.C. § 102(b) (pre-AIA) regardless of the '848 patent's priority date.

III. LEVEL OF ORDINARY SKILL IN ART

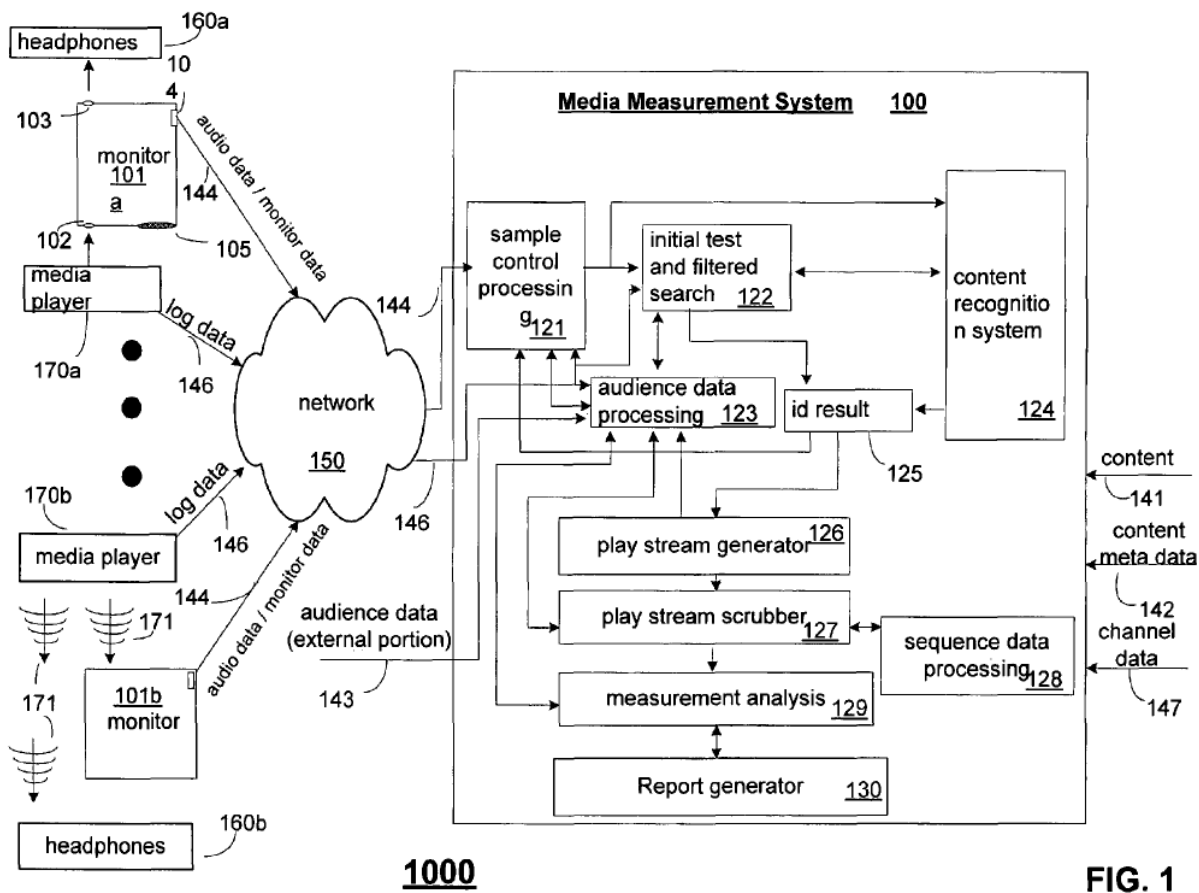
A person of ordinary skill in the art ("POSA") at the time the '848 patent was filed would have a bachelor's degree in electrical engineering, computer engineering, computer science, or equivalent degree with at least two years of relevant industry experience, including in digital media content delivery. EX1003, ¶32; *see also* EX1003, ¶¶6-15 (Dr. Russ is a POSA).

IV. OVERVIEW OF THE '848 PATENT

The '848 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, 1:66-2:3. In particular, it “analyze[s] sample sequence data” to “enhance content identification.” EX1001, 2:1-5. EX1003, ¶41.

Figure 1 illustrates such a media usage monitoring and measurement system 1000. EX1001, 3:5-9, FIG. 1.



EX1001, FIG. 1.

Figure 2 illustrates a media measurement method 200 carried out by system 100. EX1001, 7:1-5, 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, 7:5-8. The '848 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, 3:10-44, FIG. 2.

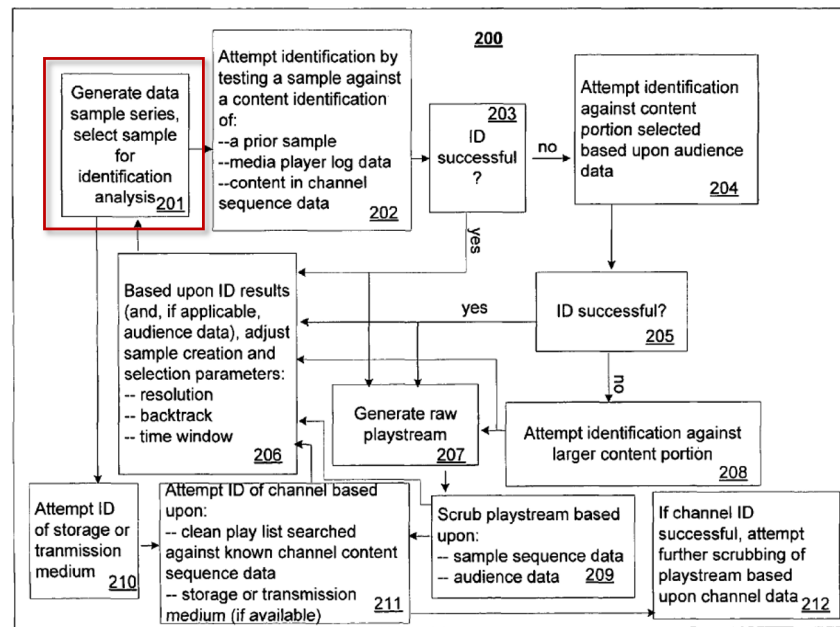


FIG. 2

EX1001, FIG. 2 (annotated).

At step 202, the initial test and filtered search module 122 attempts to identify a sample in a 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, 5:7-10, 7:26-30.

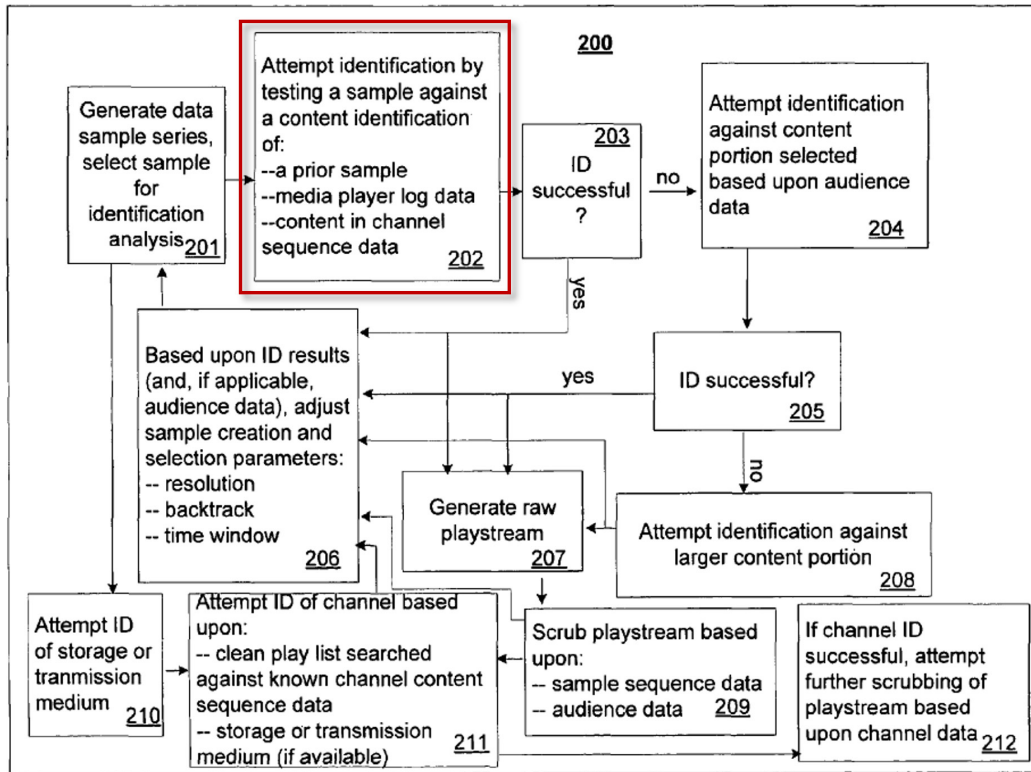


FIG. 2

EX1001, FIG. 2 (annotated).

If the sample is identified, the result is provided to play stream generator 126 (step 207). EX1001, 7:34-36. 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. *Id.*, 7:36-38, FIG. 2.

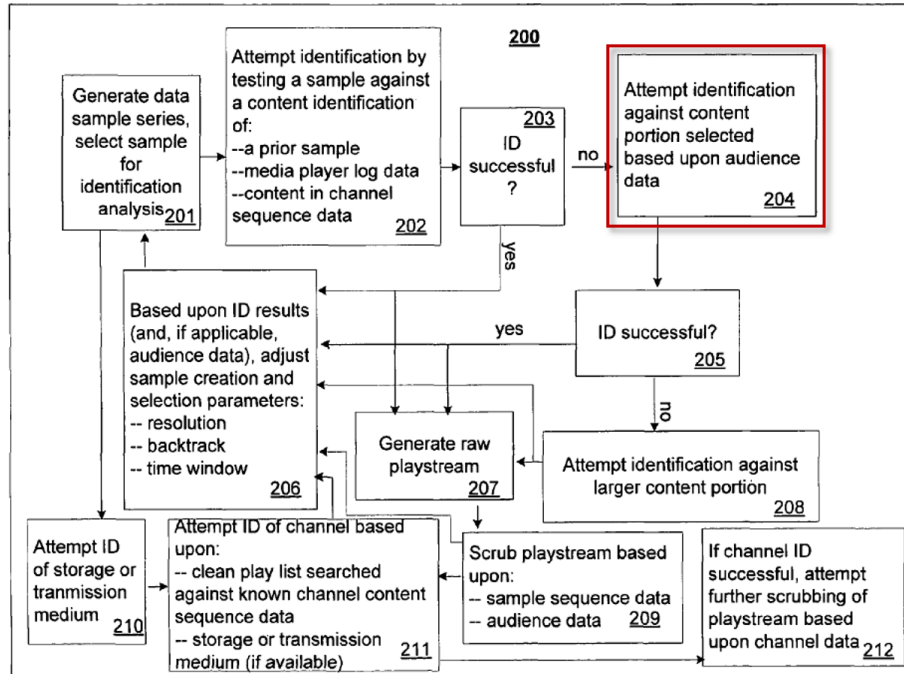


FIG. 2

EX1001, FIG. 2 (annotated).

If the sample is identified here, the result is provided to play stream generator 126 (step 207). EX1001, 7:40-41. If not, the content recognition system 124 performs a search using a larger portion of the content to identify the data sample at step 208. EX1001, 5:42-46, 7:41-42, FIG. 2.

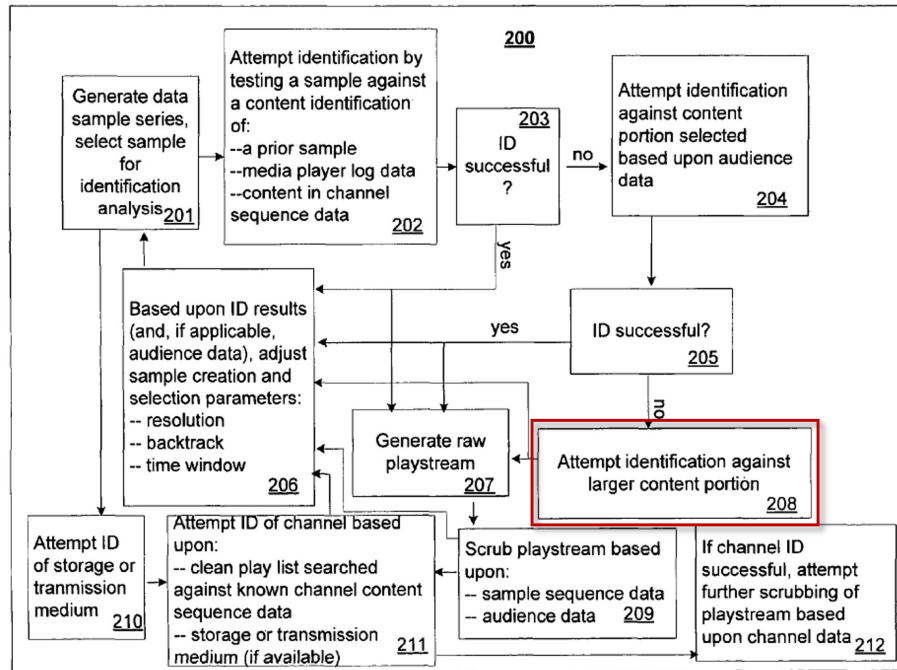


FIG. 2

EX1001, FIG. 2 (annotated).

The '848 patent admits that the content recognition system 124 did not implement a new or novel method or technique of identifying content but instead simply “implement[s] 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, 5:49-53; *see also* EX1001, 5:42-62; EX1003, ¶42.

The '848 patent explains that there are two possible results of steps 202, 204, and 208: an identification has been made or no identification has been made. EX1001, 7:42-45. Whatever the result is at step 208 (identification or no identification) that result is passed to steps 207 and 206. EX1001, 7:56-59.

At step 207, a raw play stream is generated by play stream generator 126 indicating a series of samples and corresponding identification results. EX1001, 6:24-27, 8:13-15, FIG. 2.

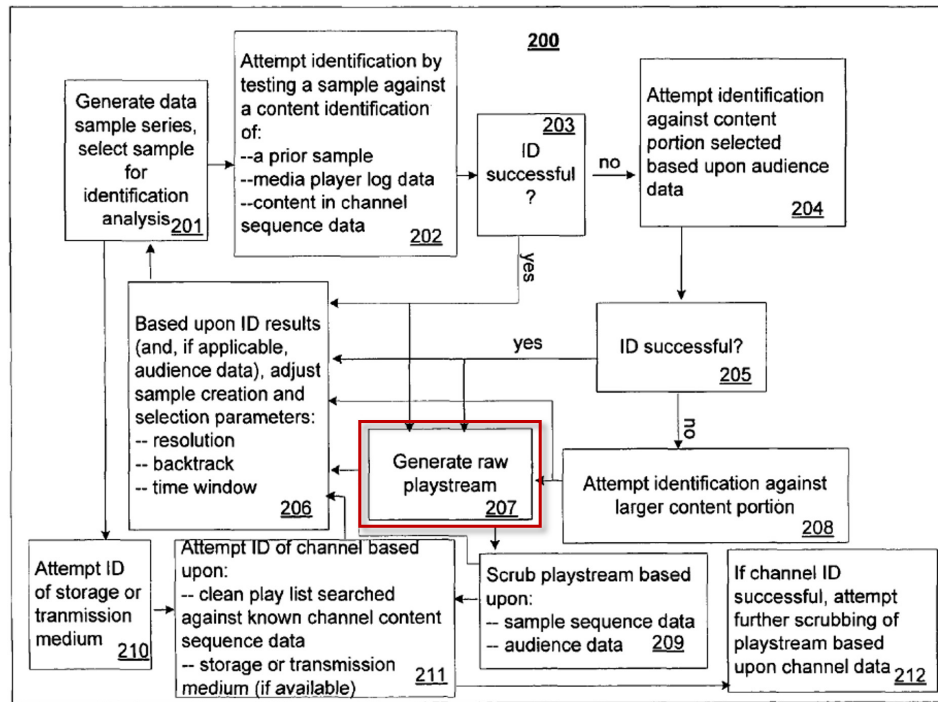


FIG. 2

EX1001, FIG. 2 (annotated).

Figure 4 illustrates a raw play stream 410, generated at step 207, which includes 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, 9:25-45. This sample sequence data (“ssd”) also includes time data (“log time” column); “the times in list 410 may be relative to an

arbitrary value (e.g., ‘ t_0 ’) rather than representing an absolute time,” or the times may “indicate the actual time the sample started to be captured.” EX1001, 9:40-44.

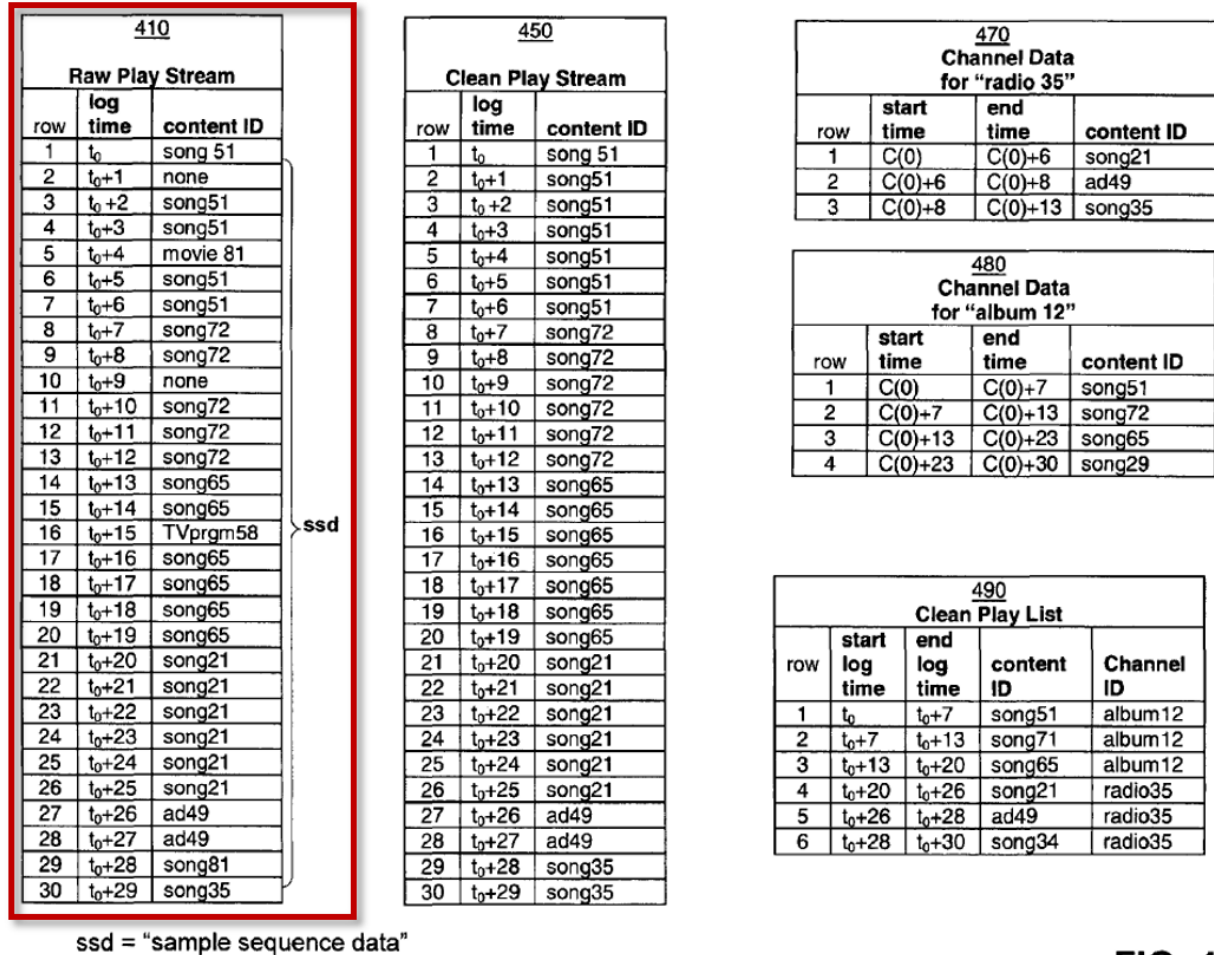


FIG. 4

EX1001, FIG. 4 (annotated).

At step 209, play stream scrubber 127 scrubs the raw play stream based upon the sample sequence data and, where appropriate, audience data, and creates a clean play stream. EX1001, 6:30-37, 8:15-18; FIG. 2.

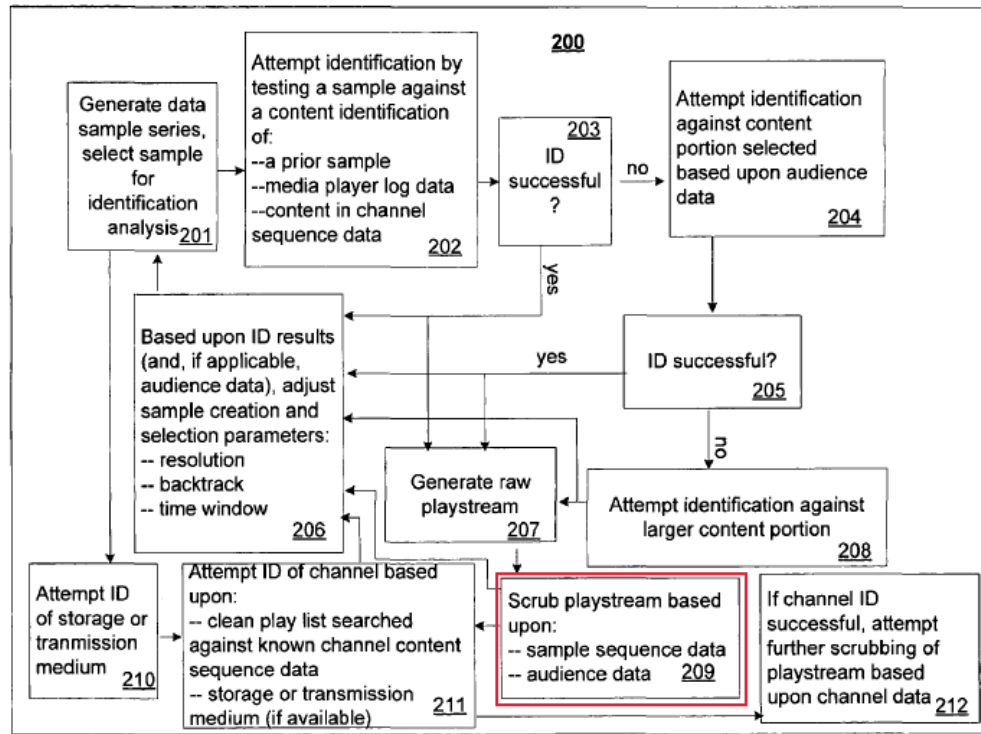


FIG. 2

EX1001, FIG. 2 (annotated).

During this scrubbing process, the raw play stream 410 is reviewed for missing or incorrect data in order to generate a clean play stream 450. EX1001, 8:44-47, 9:25-65, FIG. 4.

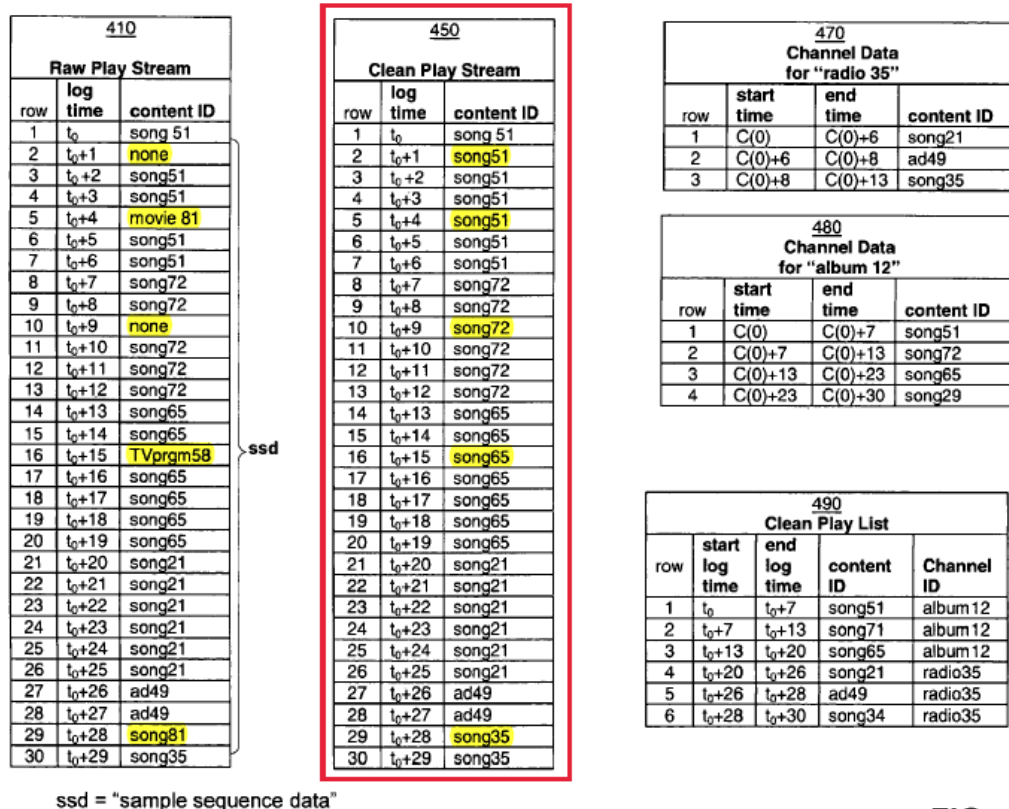


FIG. 4

EX1001, FIG. 4 (annotated).

The '848 patent "'scrub[s]' Raw Play Stream 410" "[b]y analyzing the sequence of data on Raw Play Stream" such that "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 identification." EX1001, 9:46-51. This is how the '848 patent allegedly implements "scrubbing." EX1001, 9:46-47. For example, the content ID value associated with sample 2 in the raw play stream 410 indicates that no content ID has yet been successfully associated with that sample. EX1001, 9:57-60, FIG. 4.

Given that the content IDs associated with samples 1 and 3 indicate the media content referenced as “song 51,” the ’848 patent appears to expect that sample 2 also corresponds to “song 51” and the clean play stream 450 indicates as such. EX1001, 9:52-65, FIG. 4. Similarly, the content ID value associated with sample 5 in the raw play stream 410 indicates media content referenced as “movie 81,” but given that the content IDs associate with samples 4 and 6 both indicate the media content referenced as “song 51,” the ’848 patent appears to expect that sample 5 also corresponds to “song 51” and the clean play stream 450 indicates as such. EX1001, 9:52-65, FIG. 4. The ’848 patent also describes that “channel data” or “audience data” may be used in the scrubbing process but does not describe this alleged “scrubbing” in detail. EX1001, 10:20-50.

The clean play stream 450 may then “be converted to a play list [(i.e., clean play list 490)] listing the content captured and indicating the order in which it was captured.” EX1001, 8:18-21, 9:25-29, FIG. 4.

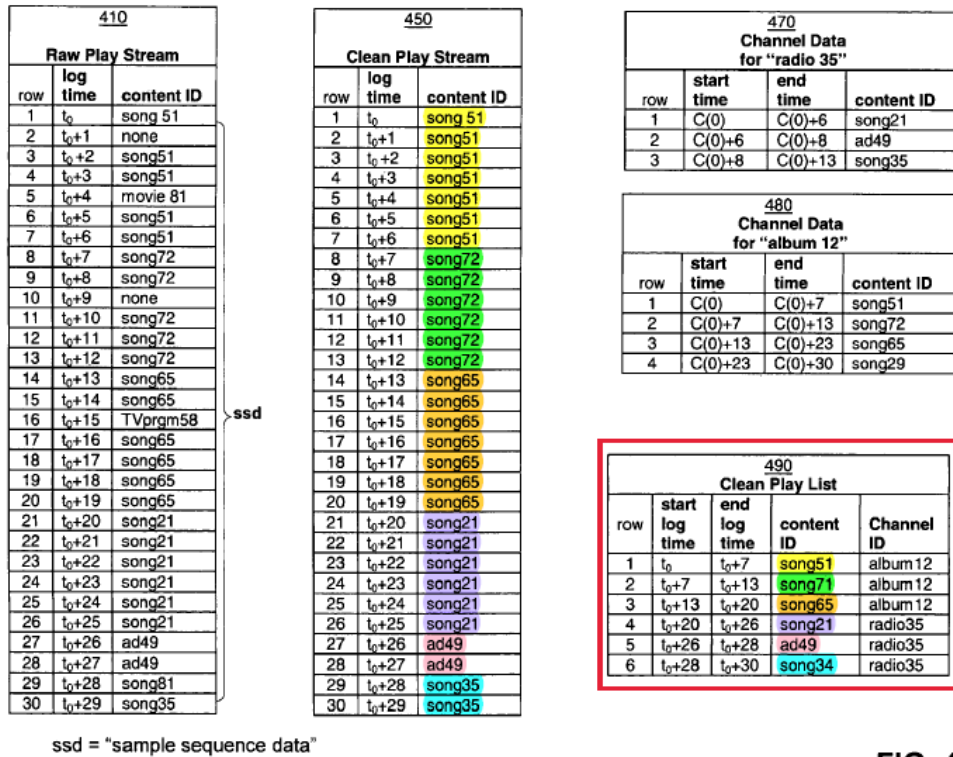


FIG. 4

EX1001, FIG. 4 (annotated)².

See also EX1003, ¶¶55-84.

V. CLAIM CONSTRUCTION

Claims must be given their ordinary and customary meaning in light of the specification—"the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention." *Phillips v. AWH Corp.*, 415 F.3d

² Based on the context of Figure 4, there appear to be typographical errors in the Clean Play List 490, where "song71" should be "song72" and "song34" should be "song35". EX1003, ¶77.

1303, 1312-1313 (Fed. Cir. 2005) (en banc); *see also* 37 C.F.R. §42.100(b); 83 Fed. Reg. 51,340, 51,358 (Oct. 11, 2018).

Solely for purposes of the present Petition, Petitioner believes that none of the claim terms require explicit constructions. Accordingly, all claim terms have been given their plain and ordinary meaning, in the context of the '848 patent specification, under the *Phillips* standard. EX1003, ¶¶89-90.

VI. OVERVIEW OF THE APPLIED REFERENCES

A. Overview of Herley682 (EX1005)

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.” EX1005, ¶2.

Figure 2 of Herley682, reproduced below, illustrates an exemplary architectural diagram of an implementation of an embodiment thereof. EX1005, ¶38.

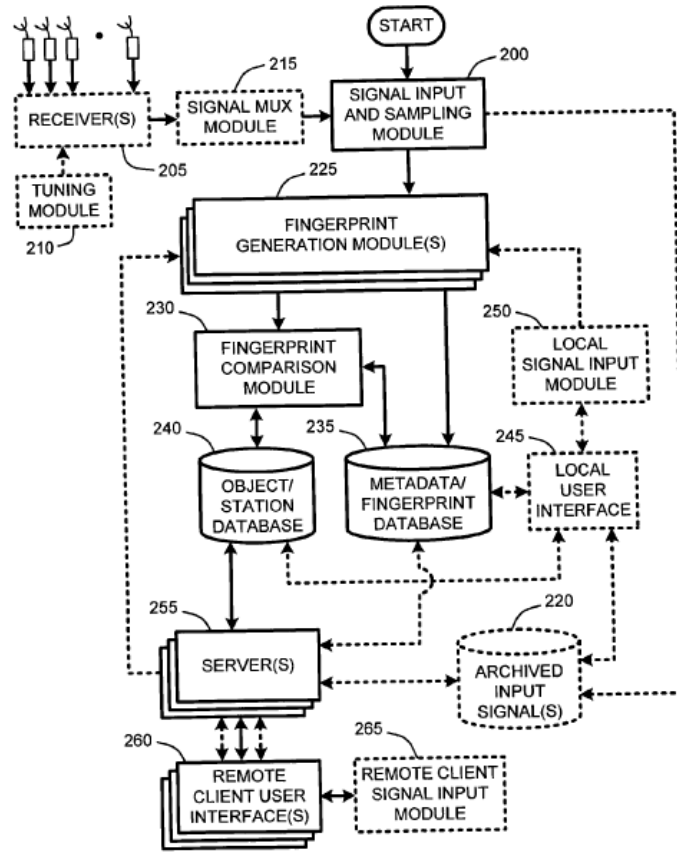


FIG. 2

EX1005, FIG. 2.

“Sampling module 200 receives and samples one or more signals, such as, for example, audio broadcast signals, video broadcast signals, etc.” EX1005, ¶75.

“[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.” EX1005, ¶84. A “separate instance of the fingerprint extraction module

225 is provided for each input signal.” EX1005, ¶84.

Extracted fingerprints are provided to fingerprint comparison module 230. EX1005, ¶86. Fingerprint comparison module 230 searches a metadata/fingerprint database 235 to determine if there is a match between the extracted fingerprint and pre-computed fingerprints stored in database 235. EX1005, ¶86. “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.” EX1005, ¶86. “This information, including the metadata, is stored in an object/station database 240.” EX1005, ¶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. EX1005, ¶86. “Metadata can include any identifying information that it the user desires to have associated with particular objects or object types.” EX1005, ¶68. For a song, the metadata may include, for example, the song’s name, artist, album title, copyright year, length, music genre, etc. EX1005, ¶¶26, 28, 68. For commercials, the metadata may include, for example, the advertisement title, name of the advertised product, etc. EX1005, ¶68.

“[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.” EX1005, ¶88.

A copy of the sample and any statistical information available for the sample also are stored in database 235. EX1005, ¶88. Metadata/fingerprint database 235 allows for unknown objects to be manually identified by a user. EX1005, ¶89.

“[W]henver any metadata in the metadata/fingerprint database 235 is edited or modified, any corresponding entry in the object/station database 240 is automatically updated to reflect the changes in the metadata. In the fashion, the two databases are kept synchronized with respect to metadata content.” EX1005, ¶89. “Alternately such [unknown objects’] metadata can be added by automatically querying a more authoritative and up-to-date fingerprint database.” EX1005, ¶17.

Herley682 expressly explains that its “interactive signal analyzer is not restricted to music or songs in fingerprinting and identifying objects in an audio stream,” and that “any type of signal or object of interest may be monitored and processed by the interactive signal analyzer,” such as, for example, “in a video stream, a given image frame or images sequence can be fingerprinted.” EX1005, ¶101. *See also* EX1003, ¶¶91-99.

B. Overview of Seet (EX1006)

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

Seet is prior art as of the February 17, 2004 filing date of U.S. provisional

patent application 60/545,681 (EX1019, “Seet provisional”) to which it claims priority. 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 § 112 support for at least one claim in the publication, and the provisional describes the subject matter relied upon as prior art. *Dynamic Drinkware LLC v. Nat’l Graphics, Inc.*, 800 F.3d 1375, 1381-82 (Fed. Cir. 2015).

The Seet provisional provides support for at least Seet’s claim 9, as shown below in the claim-by-claim analysis. EX1003, ¶¶107-109.

Seet Claim 9	Seet ’681 Provisional Appl. (EX1019)
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,” EX1019, 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.” EX1019, 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

Seet Claim 9	Seet '681 Provisional Appl. (EX1019)
	fingerprinted, and inserted into a fingerprint database.” EX1019, 4. “DB of Previous Day’s Samples,” EX1019, 5, FIG. 1, EX1019, 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.” EX1019, 4. “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.” EX1019, 5. “Next Day’s Samples’ Results Are Matched,” EX1019, 5, FIG. 1. “Next Day’s Samples’ Results,” EX1019, 6, FIG. 2.
matching at least some portions of the second media stream to the first media stream to determine repeating content between the streams; and	“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.” EX1019, 5.

Seet Claim 9	Seet '681 Provisional Appl. (EX1019)
	“Now that a relationship has been established between the previous period’s database and the current period’s samples, it can be asserted that the portions of the previous period database that were matched are repeated content like advertisements, promotions, songs, or the like.” EX1019, 5. “Next Day’s Samples’ Results Are Matched,” EX1019, 5, FIG. 1. “Matched content to be reviewed by human to add content description,” EX1019, 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 to add metadata (like ad name, or artist, and title information) to the series of samples.” EX1019, 5. “Matched content to be reviewed by human to add content description,” EX1019, 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.” EX1019, 6.

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.

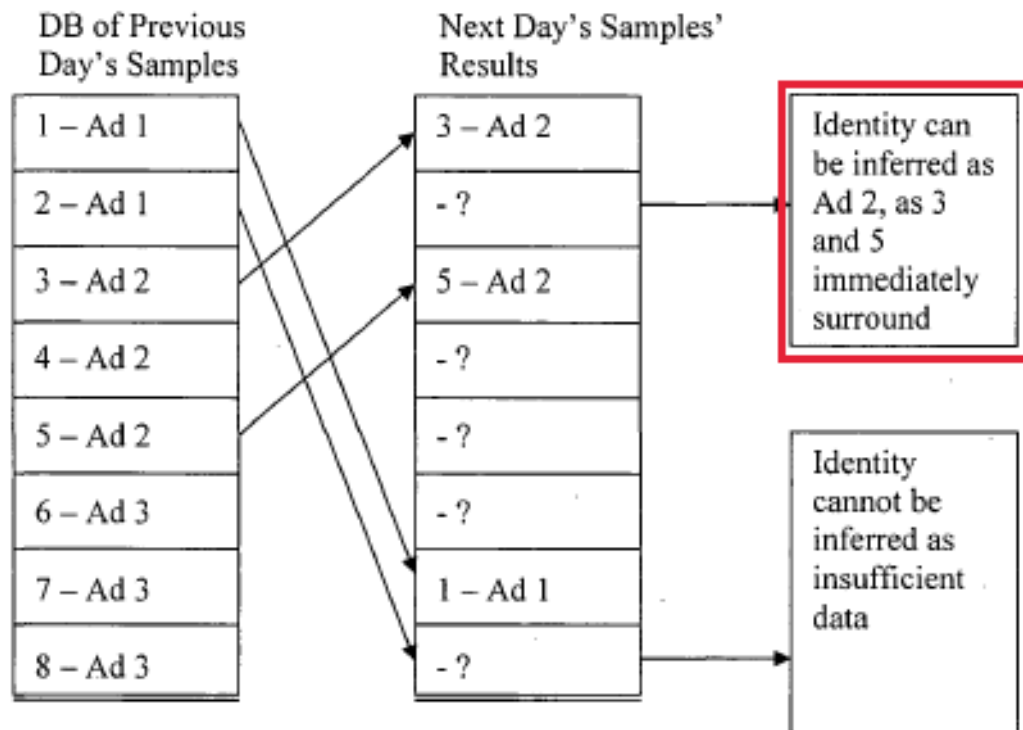


Figure 2. Auto-correlation Identity Inference Example

EX1019, FIG. 2 (annotated).

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.” EX1019, 5. Thus, the

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

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 that correspond to the duplicate sample sets.” EX1019, 6. “This requires that the underlying fingerprinting software returns all matches, not just the first match.” EX1019, 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.” EX1019, 6. .
See also EX1003, ¶¶110-116.

2. Overview

Seet is directed to “Generation of a Media Content Database by Correlating Repeating Media Content in Media Streams.” EX1006, Title.

Seet’s Figure 1, reproduced below, illustrates a system for creating and maintaining a media content database. EX1006, ¶15.

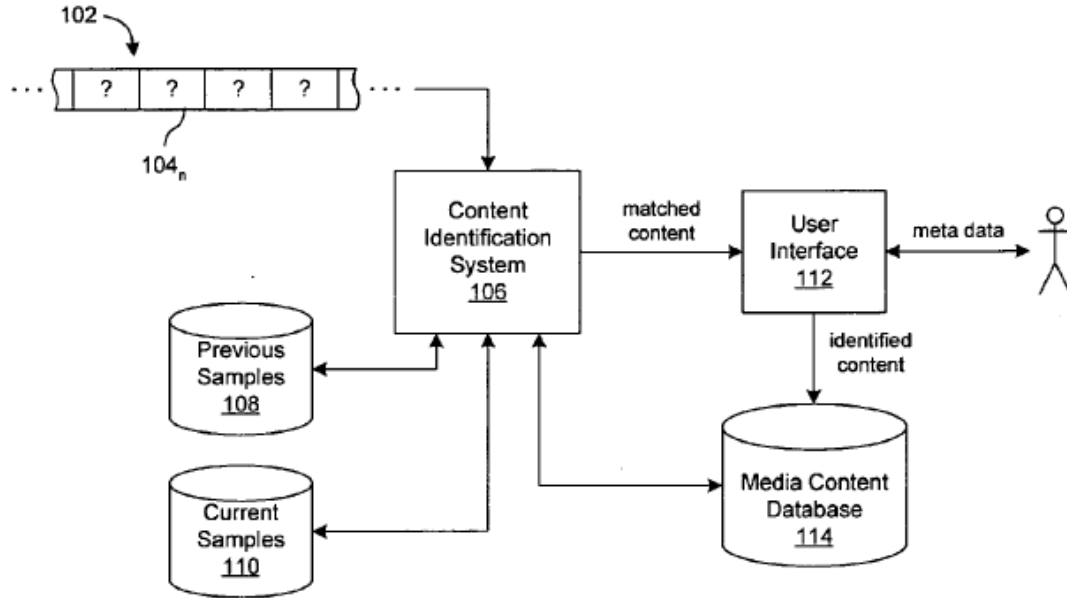


FIG. 1

EX1006, FIG. 1.

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. EX1006, ¶19. The content identification system 106 attempts to match samples in the current samples database 110 with samples in the previous samples database 108. EX1006, ¶19. Matched samples are made accessible for review through a user interface 112, allowing a reviewer to identify meta-data for matched content. EX1006, ¶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. EX1006, ABSTRACT, ¶¶14,

19.

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

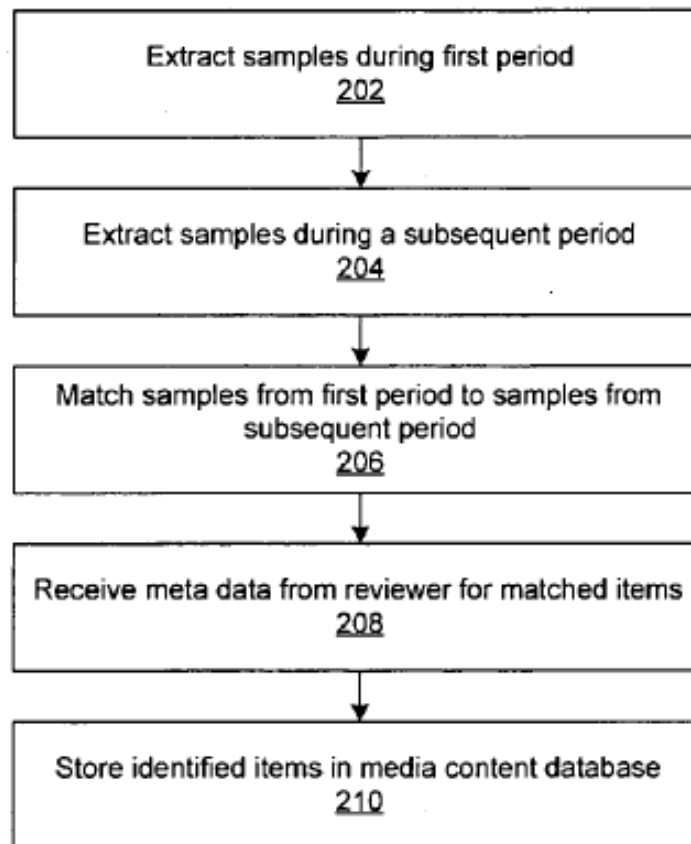


FIG. 2

EX1006, FIG. 2.

A set of samples is extracted from a first period of media stream 102 (step 202). EX1006, ¶23. Samples 104_n (unknown media) are subsequently extracted (step 204). EX1006, ¶23. The system then attempts to match (step 206) the second

set of samples to the first set of samples. EX1006, ¶24. Matched content is then provided to a reviewer, from which meta-data are received (step 208). EX1006, ¶24. This metadata and characteristic information of associated samples are stored (step 210) in the media content database 114. EX1006, ¶24. The characteristic information (i.e, the fingerprint) and a marker (such as a timestamp) are associated with each other. EX1006, ¶27.

Because the samples typically obtained in noisy conditions, they may not match exactly (or within a predetermined tolerance) even where they are of the same content. EX1006, ¶31. Figure 4, reproduced below, illustrates a process for inferring a match for an unmatched sample. EX1006, ¶18.

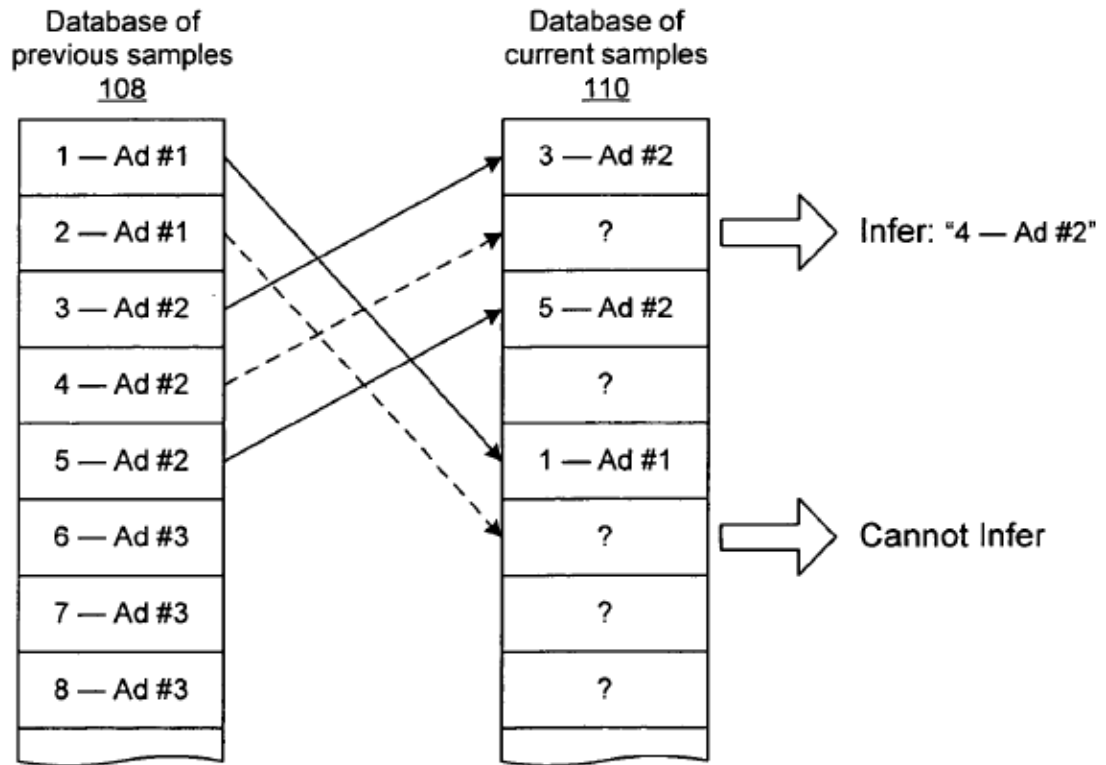


FIG. 4

EX1006, FIG. 4.

System 106 may infer additional matches, for example based on a sample's temporal position in media stream 102 relative to other matched samples. EX1006, ¶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.” EX1006, ¶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. EX1006, ¶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.

EX1006, ¶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.

EX1006, ¶32.

Seet recognizes that although embodiments are described with respect to audio content, its “media content database system could be used with any other type of media content,” including a video file. EX1006, ¶50; EX1006, ¶26 (describing that the media stream may be a television broadcast). *See also* EX1003, ¶¶100-106.

C. Overview of Herley864 (EX1007)

Herley864 is directed to a “computationally efficient” way to “identif[y] repeating objects or music in an audio stream, and in particular, to a system and method for jointly segmenting and identifying repeating objects from one or more broadcast audio streams.” EX1007, 1:16-19, 55-60.

Figure 2 of Herley864, reproduced below, illustrates an exemplary architectural diagram of an implementation of an embodiment thereof. EX1007, 4:15-17.

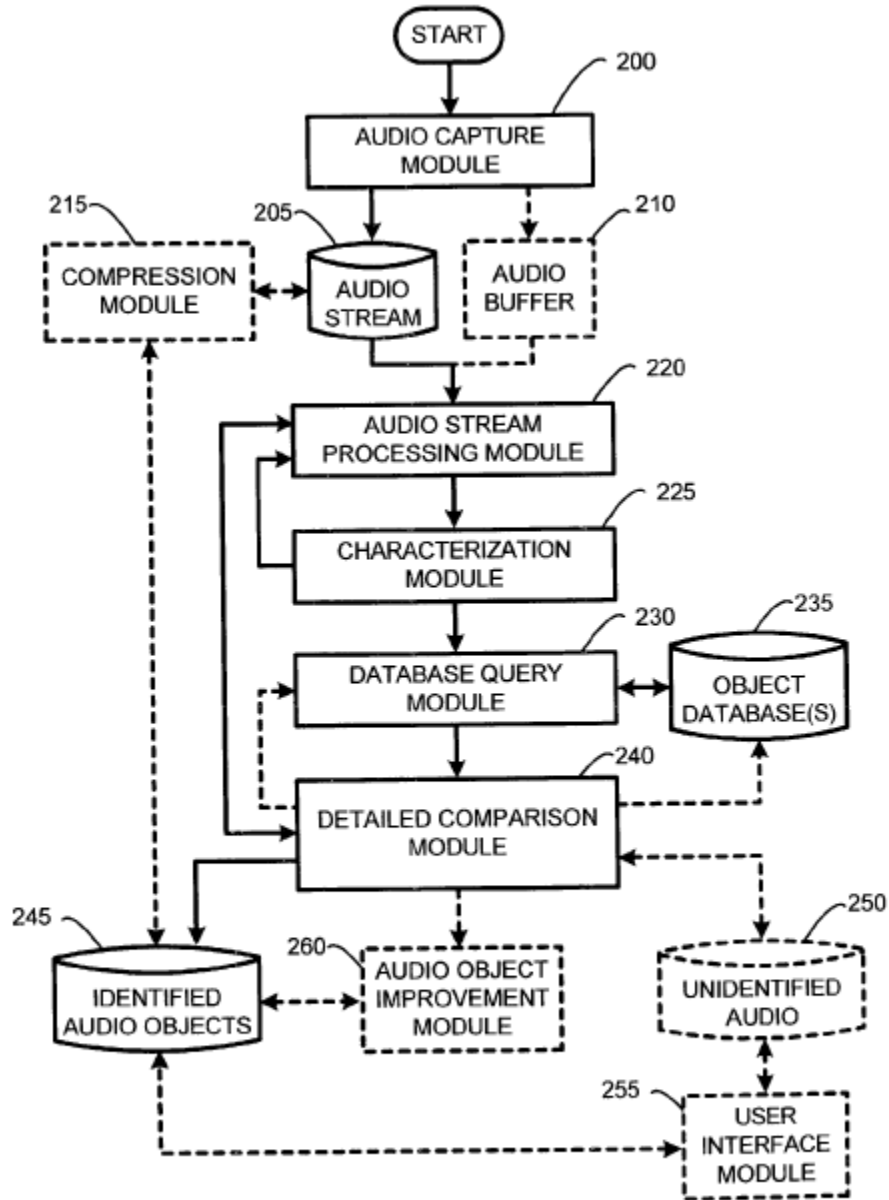


FIG. 2

EX1007, FIG. 2.

Audio capture module 200 captures one or more audio streams containing audio information, using well-known audio stream capture techniques. EX1007, 14:59-15:2. “[A]udio stream processing module 220 then begins examining the

audio stream, either stored or buffered, 205 or 210, respectively.” EX1007, 15:11-

13. In particular, audio stream processing module 220 provides segments or windows of the audio stream to a characterization module 225 for computing parametric information of the audio object. EX1007, 15:14-16. For example, if audio object represents a song, “the parametric information includes elements such as, for example, beats per minute (BPM), energy occupancy of various frequency bands, zero-crossings, etc., as estimated or calculated over each current window.” EX1007, 8:49-54.

“[D]atabase query module 230 ... queries the object database 235 in order to identify potential matches to the portion of the audio stream being examined.” EX1007, 15:29-32. “[T]he object database contains parametric information for characterizing each known audio object.” EX1007, 8:39-40. Potential matches are identified, and the detailed comparison module 240 “perform[s] a detailed comparison between the portion of the audio stream being analyzed and one or more of the potentially matching object.” EX1007, 16:4-10. Once identified, detailed comparison module 240 “uses the duration information [of the known audio object] to identify the nominal extents [(i.e., the endpoints)] of the identified audio object.” EX1007, 16:24-26.

If any portion of the audio stream cannot be identified, the “unidentified audio portions 250 are saved ... for later manual browsing, classification, or

identification via a user interface module 255.” EX1007, 16:37-41. The “parametric and detailed information for characterizing the manually identified audio object is then made available to the object database for use in future automatic identifications of repeat instances of the manually identified audio object.” EX1007, 10:12-16. “[T]he extents of any such unknown or unidentified audio can be determined by identifying audio objects existing on either side of the unidentified audio within the stream.” EX1007, 13:26-28

Herley864 recognizes that it “is well known” that “the audio portion of an audio/video broadcast coincides with the video portion of the broadcast” and can be used “to identify repeating video objects within the audio/video stream.” EX1007, 19:37-20:2; *see also* EX1007, claims 17, 40, 60, 69 (claiming identifying video objects). *See also* EX1003, ¶¶117-133.

D. Overview of Wang060 (EX1008)

Wang060’s embodiments describe 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.” EX1008, Abstract. Wang060 expressly recognizes that its “invention can be applied to any type of media samples and media files, including ... video.” EX1008, ¶37. Figure 1 illustrates a method for recognizing a sound sample:

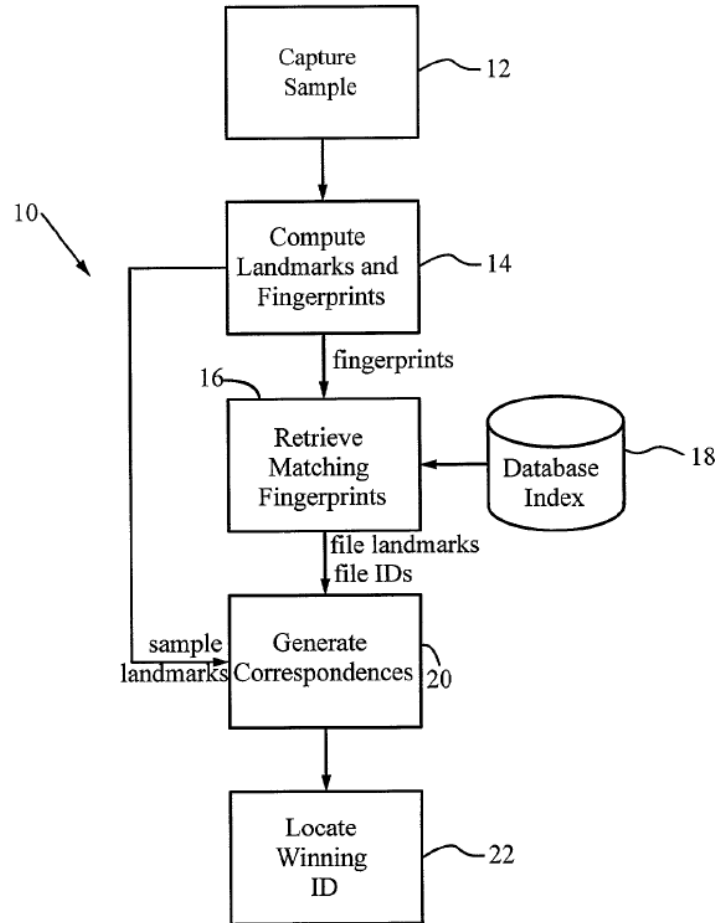


FIG. 1

EX1008, FIG. 1.

EX1008, ¶¶26, 39-42.

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.

EX1008, ¶107. *See also* EX1003, ¶¶134-143.

VII. GROUND 1: CLAIMS 1-3, 5, 6, 8-11, 13, AND 14 ARE OBVIOUS OVER HERLEY682 AND SEET

Claims 1-3, 5, 6, 8-11, 13, and 14 would have been obvious over Herley682 and Seet. EX1003, ¶¶144-247.

A. A POSA Would Have Combined Herley682 and Seet

A POSA would have found it obvious to combine Herley682 and Seet for multiple reasons. EX1003, ¶¶145-150.

Both Herley682 and Seet relate to methods and systems of identifying objects of interest (e.g., songs) within signals (e.g., a radio broadcast). Both also expressly recognize the methods and systems thereof can alternatively be used to identify video content. They both 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. Herley682 and Seet both also describe how to handle the situation when there is not a match. EX1003, ¶145.

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.” EX1005, ¶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.” EX1005, ¶10. Herley682 teaches that these monitored signals may be, for example, audio signals, video signals, or combined audio/video signals. EX1005, ¶11; EX1003, ¶¶145-147. In Herley682, unmatched objects are stored as “unknown objects” and can be manually identified by a user. EX1005, ¶¶16-17, 71.

Seet teaches using system 106 to “infer additional matches” for unmatched samples “based on their position in a media stream.” EX1006, ¶32; EX1019, 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 “unknown objects” data in Herley682, was well-known to a skilled artisan. EX1003, ¶¶46-53, 148; *see, e.g.*, EX1009, EX1010, EX1012, EX1013, EX1014.

A POSA would have been motivated to use Seet’s teachings of inferring matches to identify Herley682’s unmatched objects at least because a POSA would have been motivated to automate that which is done manually because it would be faster, a stated goal of Herley682. EX1005, ¶80; EX1003, ¶¶147, 210, 222, 262. *See also Application of Venner*, 262 F.2d 91, 95 (C.C.P.A. 1958) (“[I]t is well settled that it is not ‘invention’ to broadly provide a mechanical or automatic means to replace manual activity which has accomplished the same result.”). A POSA would have further understood that using Seet’s method to automatically correct the unknown objects in Herley682 would increase data reliability. EX1003,

¶149. Further, Seet provides express motivation for using its methods to improve media identification systems, such as that of Herley682. *See* EX1006, ¶¶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.” EX1006, ¶13.

Further still, Herley682 and Seet are in the same field of endeavor. *Compare* EX1005, ¶2, *with* EX1006, ¶3, EX1019, 5. 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. EX1003, ¶150. Applying the teachings of Seet to the system and method taught in Herley682 would have simply been the combination of prior-art elements according to known methods to yield predictable results. EX1003, ¶150. Similarly, 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 (Herley682’s system/method, which includes identifying such unmatched samples) to yield the predictable result of identifying objects of interest in a broadcast signal. EX1003, ¶150.

It would have been well within the level of ordinary skill to implement Seet’s teachings in the system of Herley682 because they are similar media content

identification methods implemented in conventional database systems

implemented in conventional hardware/software. EX1003, ¶150; EX1005, ¶50; EX1006, ¶51.

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

B. Independent Claim 1

1. [1.P] “A media measurement method comprising:”

To the extent it is limiting, Herley682 teaches the preamble. EX1003, ¶¶151-155.

For example, 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.” EX1005, ¶2. In particular, Herley682 discloses an “interactive signal analyzer” which, 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.” EX1005, ¶10. Herley682 teaches that these monitored signals may be, for example, audio signals, video signals, or combined audio/video signals. EX1005, ¶11; *see* EX1005, ¶6 (other signal types, including video signals, may be analyzed).

Herley682 also describes performing analysis of the content of at least radio or television signals, i.e., “media.” *See, e.g.*, EX1005, ¶23; EX1003, ¶155.

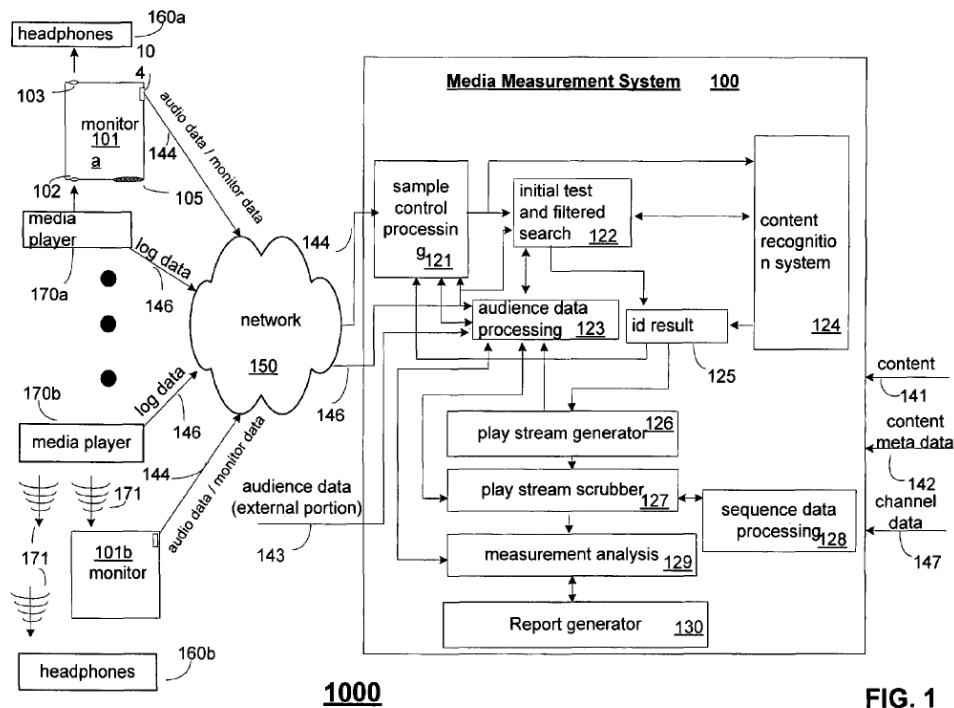
2. [1.1] **“(a) receiving over a network, at one or more computers, data for a sequence of video data samples, the video data samples comprising representations of video data captured at a media monitoring device, the plurality of video data samples being submitted over the network”**

Herley682 teaches element [1.1]. EX1003, ¶¶156-169.

As an initial matter, the ’848 patent explains, with reference to Figure 1, that audio data³ is captured by monitors 101 and uploaded to media measurement system 100, e.g., via network 150. EX1001, 3:45-50. The uploaded audio data may be, for example, (1) “a raw signal corresponding to audio energy received at monitor 101”; (2) “raw data” (e.g., data that “include[s] parameters useable to calculate audio signatures (sometimes referred to as audio fingerprints) and landmarks or other values useful in the content recognition process”); or (3) data that has undergone certain processing prior to upload (e.g., “the audio data may comprise pre-calculated signatures, landmarks, or other data useful for content recognition”). EX1001, 3:53-62; EX1003, ¶157.

³ As in much of the prior art, the ’848 patent’s exemplary embodiments are described with respect to audio data, and an alternative embodiment directed to video signals is mentioned as an alternative. *See* EX1001, 14:62-15:10.

The '848 patent's sample control processing module 121 divides the audio data into samples (or receives pre-divided audio data) and determines which data into samples (or receives pre-divided audio data) and determines which samples should be submitted for content recognition testing. EX1001, 4:62-5:5. Then, "[c]ontent recognition system 124 . . . implement[s] 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, 5:48-52; EX1003, ¶¶42, 158. The provided examples include various fingerprint searching techniques. *See* EX1001, 5:54-66.



EX1001, FIG. 1.

Herley682 teaches the same. Figure 2 of Herley682, reproduced below, illustrates an exemplary architectural diagram of an embodiment thereof. EX1005,

¶38.

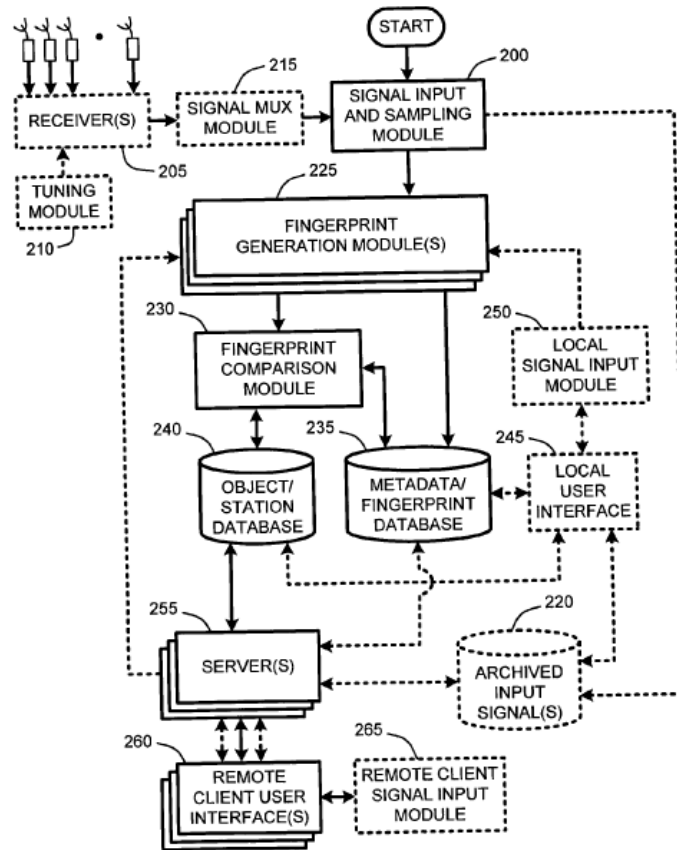


FIG. 2

EX1005, FIG. 2.

Figure 5 of Herley682, reproduced below, illustrates an exemplary flow diagram of an embodiment thereof. EX1005, ¶41.

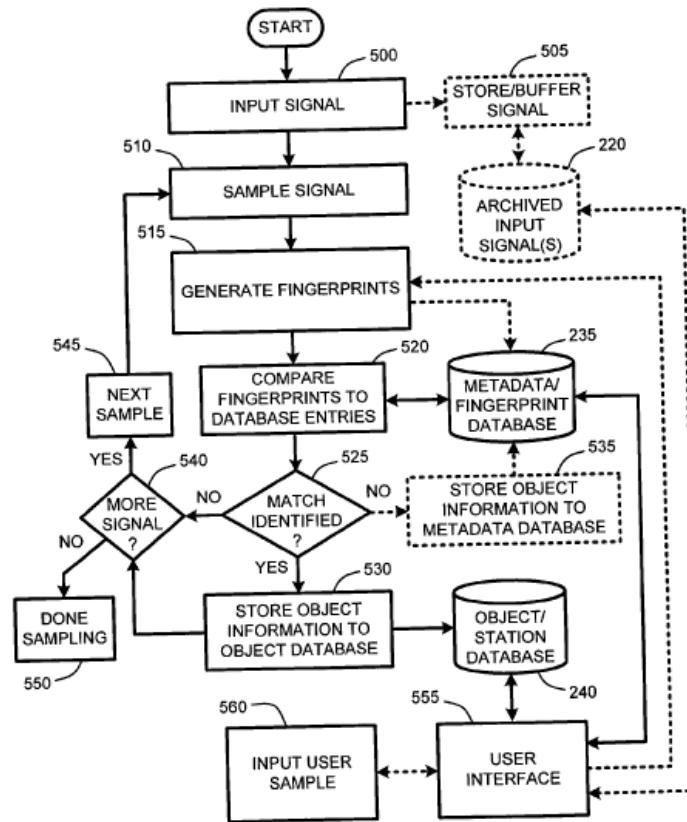


FIG. 5

EX1005, FIG. 5.

Referring to Figures 2 and 5, receiver 205 (*media monitoring device*⁴) captures a radio or television broadcast (*video data*). EX1005, ¶¶75, 160 (input signal 500). Sampling module 200 receives and samples these broadcast signals, and sends the samples (*data for a sequence of video data samples*) to fingerprint generation module(s) 225, which in turn extracts fingerprints from the samples.

⁴ Claim language is presented in italics.

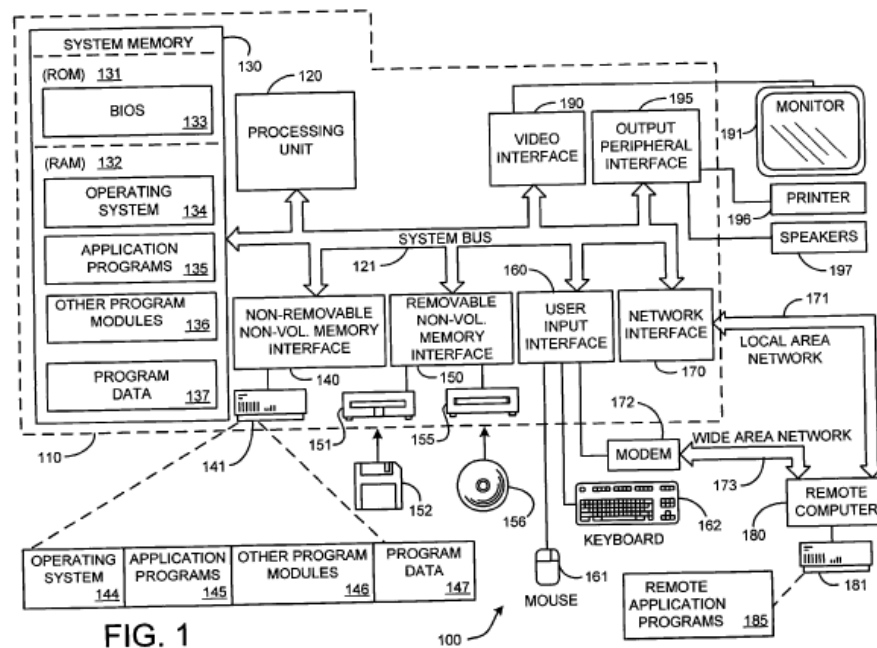
EX1005, ¶¶24, 75, 84, 161 (sample signals 510; generate fingerprints 515).

Herley682 describes that such “fingerprints” are “used to uniquely identify the signal segments from which they are derived.” EX1005, ¶13. In other words,

Herley682’s fingerprints are *representations of captured video data* as claimed.

EX1003, ¶168. “Once the fingerprint extraction module 225 has extracted a fingerprint from the sampled input signal, that fingerprint is provided to a fingerprint comparison module 230.” EX1005, ¶86.

Figure 1 of Herley682, reproduced below, illustrates an example of a suitable computing system environment 100 on which “an exemplary system for automatically identifying and storing content information for sampled signals” may be implemented. EX1005, ¶¶37, 49.



EX1005, FIG. 1.

Herley682 further describes that the “invention is operational with numerous other general purpose or special purpose computing system environments or configurations,” such as “personal computers, server computers, handheld, laptop or mobile computer or communications devices such as cell phones and PDAs, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, main frame computers, distributed computing environments that include any of the above systems or devices, and the like.” EX1005, ¶50.

A POSA would have thus understood that signal input and sample module 200, fingerprint generation module(s) 225, and fingerprint comparison module 230 are implemented as *one or more computers*. EX1003, ¶166, 169. Herley682 expressly teaches that “[t]he invention may be described in the general context of computer-executable instructions, such as program modules, being executed by a computer.” EX1005, ¶51. Herley682 further teaches that “[t]he invention may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network.” EX1005, ¶51; *see also id.* ¶¶59-60 (discussing Herley682’s operation in a networked environment). In at least this instance, a POSA would have further understood that Herley682’s broadcast signals, samples (*data for a sequence of*

audio data samples), and fingerprints (*representations of video data captured*) are provided respectively to signal input and sample module 200, fingerprint generation module(s) 225, and fingerprint comparison module 230 (*one or more computers*) over a network. EX1003, ¶169.

3. [1.2] “(b) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers”

Herley682 teaches element [1.2]. EX1003, ¶¶170-174.

In the '848 patent, the “content recognition system 124” which, as described above, is “implement[ed as] 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, 5:48-52; EX1003, ¶42. The provided examples include various fingerprint searching techniques. *See* EX1001, 5:54-66.

Herley682 teaches one such fingerprint searching technique. EX1003, ¶¶171-173. After receiving the trace fingerprints (*representations of video data captured*) from fingerprint generation module 225, Herley682's fingerprint comparison module 230 (*one or more computers*) searches (*queries*) metadata/fingerprint database 235 (*an electronic database*) having pre-computed fingerprints of known objects (*a plurality of video data representations*) and metadata describing objects represented by the pre-computed fingerprints.

EX1005, ¶¶86, 161, FIG. 5 (step 520). Herley682 describes that that the “name and other meta data” are determined, e.g., for a song, the metadata can further include artist, album title, copyright year, length, music genre or subgenre. EX1005, ¶¶26, 28, 68. This metadata corresponds to the claimed *corresponding content identifiers*, which is consistent with the ’848 patent’s description of “content ID” as “a reference corresponding to a particular piece of content,” e.g., “song51.” EX1001, 9:32-35, FIG. 4; EX1003, ¶174.

4. [1.3] “(c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of video data samples”

The Herley682-Seet combination teaches and/or renders obvious element [1.3]. EX1003, ¶¶175-181.

In the ’848 patent, “[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^[5] and accompanying text).” EX1001, 6:24-27, FIG. 4; *see also id.*, 8:13-15 (“step 207 generates a raw play stream indicating a series of samples and corresponding identification results”); EX1003, ¶¶43-45, 81. Figure 4 illustrates an example of such a “series of samples and

⁵ Presumably this is a typographical error which should refer to Figure 4, which shows “raw play stream 410.” EX1003, ¶77.

corresponding identification results” as a table including columns for “row,” “log time,” and “contentID.” EX1001, FIG. 4; *see also* EX1001, 10:51-11:7 (describing that the invention is not limited to any specific organization, division, or content of data).

Herley682 similarly discloses that if a matching fingerprint is identified by the fingerprint comparison module 230, then the sampled portion of the input signal is identified as belonging to an object of interest, e.g., a song, corresponding to the known fingerprint and metadata in metadata/fingerprint database 235. EX1005, ¶¶86, 162, FIG. 5 (step 525). This information, including the metadata, is stored in the object database 240. EX1005, ¶¶86, 162, FIG. 5 (step 530). The time that the identified object appeared in the input signal 500 (e.g., the time played) may also be included. EX1005, ¶162. A POSA would have understood that trace fingerprints are generated (step 515) and compared (step 520) for a plurality of samples taken from input signal 500 (step 510). EX1003, ¶178; EX1005, ¶161 (discussing sample size and period), ¶13 (““trace fingerprints’ ... are generated several times per second on an incoming signal”). For each identified match (step 525), the metadata associated with the matching fingerprint is stored (step 530) in object database 240. EX1005, ¶¶86, 162.

In other words, Herley682 stores, in the object database 240, a list of data relating sampled portions, metadata (*content identification results*), and times.

EX1003, ¶179. A POSA would have understood that the collection of data stored in Herley682's object database 240 teaches a *raw play stream*. EX1003, ¶179. Seet similarly teaches storing samples and corresponding identifiers in the database of current samples 110. EX1006, ¶30, FIG. 3; EX1019, 5, FIG. 1. This is consistent with the '848 patent's description that a "raw play stream relat[es] a sequence of samples to content identification results." EX1001, 6:24-26; *see also* EX1001, 8:13-15 ("[S]tep 207 generates a raw play stream indicating a series of samples and corresponding identification results."), 9:25-32, FIG. 4 (raw play stream 410); EX1003, ¶¶43-45, 81, 179.

To the extent Patent Owner argues that the claimed *sequence* must be stored in chronological order, it would have been obvious to store the data in sequential order as the data in the stream is processed or, at a minimum to sort the data in Herley682's object database 240 based on the time data to put it in sequential, chronological order. EX1005, ¶¶81, 83, 121; EX1003, ¶180. Seet provides an express teaching of maintaining the temporal integrity, i.e., chronological order, of the sample data, confirming that a POSA would have been motivated to do so. EX1003, ¶181. In Seet, the list of match identifications is maintained in sequential order to allow for using information from temporally adjacent samples to identify unmatched samples. EX1003, ¶181; EX1006, ¶¶30, 32, FIGs. 3-4; EX1019, 5-6, FIGs. 1, 2.

Additional details of how the Herley682-Seet combination renders obvious claim 1's *raw play stream* are discussed with regard to elements [1.4], [1.5], and [1.6] which further define the raw play stream.

5. [1.4] **“wherein: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective video data samples in the sequence of video data samples and respective video data representations in the electronic database”**

The Herley682-Seet combination teaches element [1.4]. EX1003, ¶¶182-185.

As discussed for elements [1.2] and [1.3], Herley682's fingerprint comparison module 230 searches (*quer[ies]*) metadata/fingerprint database 235 (*the electronic database*) to compare the trace fingerprints (*representations of video data captured* representing the *respective video data samples*) to pre-computed fingerprints of known objects (*plurality of video data representations*) in order to determine if there is a match (*to attempt to determine respective likely matches*). EX1005, ¶¶86, 161, FIG. 5 (step 520). If a matching fingerprint is identified, the list of data relating sampled portions, metadata (*content identification results*), and times are stored in the object database 240. EX1005, ¶¶86, 162, FIG. 5 (steps 525, 530). As discussed for element [1.3], this searching, matching, storing process is completed for a series of samples from input signal 500, resulting in a *sequence of content identification results* stored, in sequential

order as expressly taught by Seet, in object database 240. Thus, the claimed

sequence ... is obtained by the querying process just described. EX1005, ¶¶86,

160, 161; EX1006, ¶32, FIG. 4; EX1019, 6, FIG. 2; EX1003, ¶185.

6. [1.5] **“a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective video data representation that is determined to be a respective likely match with a respective one of the video data samples; or (ii) an indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database”**

Herley682 teaches element [1.5]. EX1003, ¶¶186-189.

As discussed for elements [1.2]-[1.4], for a particular trace fingerprint (*representation of video data captured* representing the *respective video data sample*), the metadata stored in the object database 240 when a matching fingerprint is identified corresponds to the claimed *content identification result*. EX1003, ¶187. This metadata comes from (is *associated with*) the pre-computed fingerprint of a known object (*respective video data representation*) found to be a likely match in metadata/fingerprint database 235 (*electronic database*) by the fingerprint comparison module 230. This metadata also is a *content identifier*, as recited in part (i) of element [1.5], because it describes “a reference corresponding to a particular piece of content,” such as a song name, as described in the ’848 patent. EX1005, ¶¶26, 28, 68; EX1001, 9:32-35, FIG. 4; EX1003, ¶188.

This is sufficient to meet part (i) of element [1.5], which recites parts (i) and (ii) in the alternative.

Herley682 also teaches part (ii) of element [1.5]. Specifically, Herley682 notes that if a matching fingerprint is *not* identified in the metadata/fingerprint database 235, then it is added to the metadata/fingerprint database 235 and the object database 240 as an “unknown object.” EX1005, ¶164, FIG. 5 (step 535). The inclusion of the “unknown object” in object database 240 corresponds to *an indication of the absence of a respective likely match*, as recited in part (ii) or element [1.5]. EX1003, ¶189. This is consistent with the ’848 patent in which a sample having “no content ... successfully associated with that sample,” has a content ID value of “none.” EX1001, 9:57-61, FIG. 4 (sample 2).

7. **[1.6] “the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database”**

The Herley682-Seet combination teaches element [1.6]. EX1003, ¶¶190-194.

As discussed above, the data in Herley682’s object database 240 corresponds to the claimed *raw play stream*. Describing its object database (also referred to as a “station database”) in more detail with reference to monitoring FM

radio broadcast streams, Herley682 notes that “over time, the station database, D_s, will be populated with more and more information about the content that each monitored station plays.” EX1005, ¶¶120, 122. “[S]tations will appear in the station database D_s with one or more entries, depending upon the number of objects played and identified.” EX1005, ¶123.

A POSA would have understood that radio stations typically play multiple songs. EX1003, ¶193; EX1005, ¶123 (“it was observed that some stations play a limited collection of songs that are repeated fairly often, while others play larger collections.”). A POSA would also have understood that these teachings indicate that Herley682’s object database 240 would include multiple entries therein, e.g., one for each song played on the station. EX1003, ¶193. A POSA would have understood or at a minimum found it obvious that, at least most of the time, these multiple entries would be comprised of either *different content identifiers*, identifying the different songs played on the station, or they would include *at least one content identifier* and an *indication of the absence of a respective likely match*, e.g., an “unknown object” as discussed above for element [1.5] part (ii)). EX1003, ¶193.

To the extent Patent Owner argues this would not have been obvious based on Herley682 alone, this feature would have been obvious in view of Seet. Seet expressly teaches a database of samples that includes multiple entries, both

identified samples and unmatched samples. EX1006, ¶32, FIG. 4; EX1019, 6, FIG.

2. Seet's database 110, thus, teaches both *different content identifiers*, e.g., "Ad #2" and "Ad #1", and *at least one content identifier and an indication of the absence of a respective likely match*, e.g., "Ad #2" and "?" (which corresponds also to Herley862's "unknown object" as discussed above for element [1.5] part (ii)). EX1006, ¶32, FIG. 4; EX1019, 6, FIG. 2; EX1003, ¶194. A POSA would have combined Herley682 and Seet for the reasons discussed above. Section VII.A.

8. **[1.7] "(d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data"**

The Herley682-Seet combination teaches element [1.7]. EX1003, ¶¶195-212.

Dependent claim 2 recites that "sample sequence data includes at least one of log times, content identifiers, and content offsets." Thus, the *sample sequence data* of claim 1 may include any of these elements. As discussed for elements [1.3] and [1.5], the collection of data stored in Herley682's object database 240 teaches the *raw play stream* and this data includes at least sampled portions, metadata (*content identifiers*), and the time the sample was taken from the signal, e.g., the time played (*log times*). EX1003, ¶203. Herley682, thus, teaches *sample sequence*

data in object database 240.

As to the claimed *scrubbing*, the '848 patent explains that a raw play stream is prepared for further processing by “scrubbing (i.e. reviewing for missing or incorrect data).” EX1001, 8:44-47. However, the '848 patent does not describe in detail how this “scrubbing” is actually performed. The '848 patent 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, 9:46-51. A POSA would have thus understood the claimed *scrubbing*, in light of the '848 patent's specification, to merely require reviewing the data in the raw play stream for missing or incorrect data and correcting this data. The '848 patent does not provide any technical details about *how* the scrubbing is achieved (e.g., whether the scrubbing may be performed manually by a user or whether any reviewing for missing or incorrect data is performed without user intervention). EX1003, ¶¶82, 196-199.

As to the claimed *expected pattern of sample sequence data*, the '848 patent also does not expressly discuss the meaning of this language. However, in relation to samples 1-7 of Figure 4's raw data stream 410, the '848 patent describes:

“Given this pattern of sample sequence data, it is reasonable to assume that the person using the corresponding monitor listened to song51 through the time period associated with the audio-data of samples 1-7, and thus the ‘scrubbed’ version of the play stream, i.e. Clean Play Stream 450, indicates song51 for each of samples 1-7.” EX1001, 9:60-65; *see also* EX1001, 10:51-11:7 (describing that the invention is not limited to any specific organization, division, or content of data).

Consistent with these teachings of the ’848 patent, a POSA would have understood that an indication that a video sample was not identified or is missing (e.g., “none” entries in Figure 4 of the ’848 patent) is *a pattern of the sample sequence data* that is not within an *expected pattern of sample sequence data* as recited in the claim. EX1003, ¶¶83, 200-201. Thus, as best understood in light of the specification of the ’848 patent, a POSA would have understood that *scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data* includes modifying data in *the raw play stream* that indicates a missing content identification result based on reasonable assumptions from data points (e.g., *content identification results*) near the missing data in the *raw play stream*. EX1003, ¶202.

Herley682 teaches a method for handling unmatched trace fingerprints,

which correspond to the '848 patent's "none" entries in the raw play stream.

EX1003, ¶¶204-205; EX1001, 9:57-60, FIG. 4. In Herley682, "where trace 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.'" EX1005, ¶71. These unknown objects are also added to the object database 240. EX1005, ¶¶17, 18, 164.

A POSA would have been motivated to fill in the metadata (*content identifiers*) information for these unmatched entries in Herley682, because having a complete data set would be more useful for Herley682's subsequent statistical analysis. EX1003, ¶¶206-208. Indeed, Herley682 teaches that these "unknown objects" results can be changed. EX1005, ¶¶17, 71, 89. For example, Herley682 teaches that users can "manually identify such unknown objects" and change the information in the metadata/fingerprint database 235 or the object database 240, via user interface 245, 555 provided for at least that purpose. EX1005, ¶¶17, 18, 71, 89, 165. "Any metadata updated or edited by the user in one database will be automatically updated in the other database," so the two databases are kept synchronized with respect to the metadata content. EX1005, ¶¶89, 165.

And, Herley682 teaches that a user may query the object/station database 240 with any number of different requests. *See* EX1005, ¶¶125-152. If the object/station database 240, however, contains incomplete or inaccurate data the

query may return inaccurate results. A POSA would have thus been motivated to identify and correct missing data in order to obtain a complete and accurate data set in the object/station database 240—a teaching that is explicitly taught by Seet. EX1003, ¶¶209-210.

A POSA would have looked to Seet’s teachings of inferring matches as an improved way to identify Herley682’s unmatched objects, at least because a POSA would have been motivated to automate that which is done manually because it would be faster, a stated goal of Herley682. EX1005, ¶80; EX1003, ¶215. Further, 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. EX1006, ¶32; EX1019, 6; *see* EX1009, EX1010, EX1012, EX1013, EX1014; EX1003, ¶¶46-53.

In Seet, “[a] matching inference may be made for some samples based on their position in a media stream 102 relative to other matched samples.” EX1006, ¶32; EX1019, 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.” EX1006, ¶32; EX1019, 6. The example in Seet’s Figure 4 identifies unmatched samples in the same way as disclosed in Figure 4 of the ’848 patent. Namely, to identify the unmatched second sample (*determine whether to change a result of the sequence of content identification results*), 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.” EX1006, ¶32; EX1019, 6. Seet, thus, uses *an expected pattern of sample sequence data*, to identify and correct missing data in the raw play stream. EX1003, ¶211.

9. [1.8] “(e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing”

The Herley682-Seet combination teaches element [1.8]. EX1003, ¶¶213-216.

As discussed for element [1.7], Herley682 in view of Seet teaches identifying and updating the information for unidentified content in the object database 240. EX1005, ¶¶17, 18, 71, 89, 165; EX1003, ¶215. The *updated* collection of data in object database 240 corresponds to the *clean play stream*. EX1003, ¶215. This is consistent with the discussion in the ’848 patent, in which the “none” entries in the raw play stream 410 are updated to include a content identifier in clean play stream 450. *See* EX1001, FIG. 4 (e.g., sample 2, sample 10).

C. Independent Claim 9

- 1. [9.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 one or more computers being coupled to a**

network to receive, over the network, data for a sequence of video data samples, the video data samples comprising representations of video data captured at a media monitoring device, the plurality of video data samples being submitted over the network, the instructions comprising instructions for:”

To the extent it is limiting, Herley682 teaches the preamble. EX1003, ¶¶233-236.

The preamble recites 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:*”

Herley682 expressly teaches that its “invention may be described in the general context of computer-executable instructions, such as program modules, being executed by a computer.” EX1005, ¶51. Herley682’s Figure 1 illustrates “a general system diagram depicting a general-purpose computing device constituting an exemplary system for automatically identifying and storing content information for sampled signals, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1005, ¶37. Herley682’s Figure 2 “illustrates an exemplary architectural diagram showing exemplary program modules for automatically identifying and storing content information for sampled signals, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1005, ¶38.

A POSA would have understood that Herley682's system and method, when implemented on Figure 1's computing device as program modules such as those shown in Figure 2, teaches a *computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers* as recited in element [9.P]. EX1003, ¶¶234-235.

Claim 9's preamble further recites "*the one or more computers being coupled to a network to receive, over the network, data for a sequence of video data samples, the video data samples comprising representations of video data captured at a media monitoring device, the plurality of video data samples being submitted over the network.*" This portion of the preamble is essentially the same as element [1.1] and is taught by Herley682 for the reasons discussed above.

Section VII.B.2; EX1003, ¶236.

The rest of claim 9 recites the specific *instructions* included in the *computer program product*. Each of these instructions corresponds to a step in the method recited in claim 1, and is taught by Herley682 for the same reasons as discussed above. See Section VII.B.

2. [9.1] **“(a) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers”**

Element [9.1] corresponds to element [1.2]. Herley682 teaches this element

for the reasons set forth above for element [1.2]. *See* Section VII.B.3; EX1003, ¶237.

3. **[9.2] “(b) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of video data samples”**

Element [9.2] corresponds to element [1.3]. The Herley682-Seet combination teaches and/or renders obvious this element for the reasons set forth above for element [1.3]. *See* Section VII.B.4; EX1003, ¶238.

4. **[9.3] “wherein: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective video data samples in the sequence of video data samples and respective video data representations in the electronic database”**

Element [9.3] corresponds to element [1.4]. The Herley682-Seet combination teaches this element for the reasons set forth above for element [1.4]. *See* Section VII.B.5; EX1003, ¶239.

5. **[9.4] “a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective video data representation that is determined to be a respective likely match with a respective one of the video data samples; or (ii) an indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database”**

Element [9.4] corresponds to element [1.5]. Herley682 teaches this element for the reasons set forth above for element [1.5]. *See* Section VII.B.6; EX1003,

¶240.

6. **[9.5] “the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database”**

Element [9.5] corresponds to element [1.6]. The Herley682-Seet combination teaches this element for the reasons set forth above for element [1.6].

See Section VII.B.7; EX1003, ¶241.

7. **[9.6] “(c) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data”**

Element [9.6] corresponds to element [1.7]. The Herley682-Seet combination teaches this element for the reasons set forth above for element [1.7].

See Section VII.B.8; EX1003, ¶242.

8. **[9.7] “(d) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing”**

Element [9.7] corresponds to element [1.8]. The Herley682-Seet combination teaches this element for the reasons set forth above for element [1.8].

See Section VII.B.9; EX1003, ¶243.

D. Claims 2 and 10 (“wherein sample sequence data includes at least one of log times, content identifiers, and content offsets”)

The Herley682-Seet combination teaches this element. EX1003, ¶¶217-218, 244.

As explained above for elements [1.3], [1.5], and [1.7], the data stored in Herley682’s object database 240 includes at least sampled portions, metadata (*content identifiers*), and the time the sample was taken from the signal, e.g., the time played (*log times*). EX1003, ¶¶217-218. Thus, Herley682 teaches the sample sequence data as recited in claims 2 and 10. EX1003, ¶¶217-218, 244.

E. Claims 3 and 11 (“wherein: scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content identification results”)

Herley682 teaches and/or renders obvious this element. EX1003, ¶¶219-224, 245.

As to using “channel data” in scrubbing, the ’848 patent describes that “[i]nformation about the sequence of content in a play stream can be used to generate a play list and identify a channel, and the identified channel data can also potentially be used to help further ‘scrub’ a play stream.” EX1001, 9:66-10:2. The ’848 patent continues, with reference to Figure 4:

With respect to further “scrubbing” the Raw Play Stream 410, in this example, channel data may be used to correct the apparent misidentification of sample 29. Note that the sample sequence data alone in this example does not provide a clear basis for correcting

sample 29. Sample 28 is identified as “ad49”, sample 29 as “song81,” and sample 30 as “song35.” However, given the sequence “song21, ad49, and song35” is associated with a known sequence of the radio broadcast referenced as “radio 35,” there is a basis in channel data 470 for believing that the proper content ID for sample 29 is “song35” rather than “song81.”

EX1001, 10:20-30.

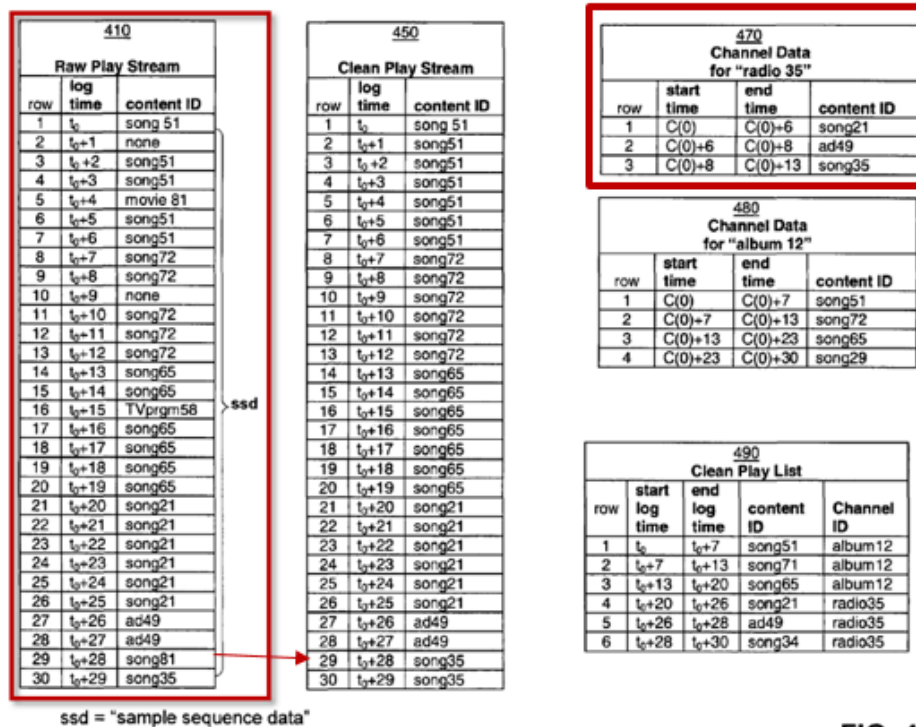


FIG. 4

EX1001, FIG. 4 (annotated).

The '848 patent, thus, describes using known information about the identification of songs played on a particular channel to provide information helping to identify missing or incorrect data in the raw play stream. In the

embodiments of the '848 patent, the specific channel also may be identified as part of the process, but claims 3 and 11 are broad enough to include the case where the identity of the channel is already known. EX1003, ¶¶219-220.

Herley682 similarly teaches using known information about the songs played on a particular channel to help identify content. EX1003, ¶221. “For example, when listening to a station that plays pop hits from the 1980’s repetitively it makes sense to search the fingerprints of songs previously identified on that station, before searching the larger database.” EX1005, ¶80. Herley682 specifically teaches that “when listening to a particular channel, the fingerprints of previously identified objects from that channel are searched first in the database, thereby *speeding the search*.”⁶ EX1005, ¶80. One way in which Herley682 describes identifying its “unknown objects” (*scrubbing*) is by “automatically querying a more authoritative and up-to-date fingerprint database.” EX1005, ¶71. Matches for all samples cannot be inferred using Seet’s inference method (e.g., sample “1 – Ad #1” in Figure 4). EX1006, ¶32 (matching for “the sixth sample ... is not inferred because it is not bounded by matched samples that correspond to the same media item”); EX1019, 6, FIG. 2 (identity cannot be inferred as insufficient data). In such cases, a POSA would have turned next to querying Herley682’s “more

⁶ Emphasis added throughout unless otherwise noted.

authoritative ... database” to determine a match in order to compile as complete a data set as possible. EX1003, ¶¶222-223. In order to “speed[] the search” when querying this additional database for identifying “unknown objects,” it would have been obvious to likewise use Herley682’s teaching of first searching previously identified objects from the channel (*analyzing channel data*) when querying this additional database. EX1003, ¶223. This would have been an obvious, natural extension of Herley682’s teaching of using the channel information when identifying the objects in the first instance, because it is simply an additional query in a different database. EX1003, ¶223.

Alternatively, Seet’s inference method itself relies on *channel data* because it takes into account previous samples, e.g., from the channel, when it determines whether it can infer matches. EX1006, ¶¶27 (previous samples may be from a particular station), ¶32 (relying on previous samples to infer matches in current samples database); EX1019, 4, 6. In the example of Seet’s Figure 4, a match is inferred when the sequence in the current samples (database 110) is consistent with the sequence in the previous samples (database 108) from that channel. A POSA would have, thus, understood that when combining Seet’s inference method with Herley682’s teaching of identifying such unmatched samples, that the combined teachings result in *analyzing channel data as part of determining whether to change a result of the sequence of content identification results*. EX1003, ¶224.

F. Claims 5 (“wherein a clean play list is generated from the clean play stream”) and 13 (“wherein the computer program product further comprises instructions for generating a clean play list from the clean play stream”)

Herley682 renders obvious this element. EX1003, ¶¶225-228, 246.

The ’848 patent describes that after scrubbing, “[t]he resulting clean play stream may be converted to a play list *listing the content captured and indicating the order in which it was captured.*” EX1001, 8:18-21, FIG. 4; *id.*, 9:66-67 (“Information about the sequence of content in a play stream can be used to generate a play list.”); 11:47-48 (“[Clean] Play List 560 contains a row for each internally continuous media consumption event.”). In other words, and consistent with the ’848 patent’s specification, a POSA would have understood a “clean play list” to encompass a list of the content identified in the video data. EX1003, ¶225.

Herley682 teaches allowing users to query the object/station database 240 to “view information describing monitored input streams.” EX1005, ¶94; EX1005, ¶¶176-186 (describing exemplary queries). In one example, “a predefined database query string” is provided, such as “a dropdown menu listing call signs of each monitored radio station.” EX1005, ¶94. “User selection of a particular radio call sign may then automatically call up a display of statistical information regarding that radio station, and the objects of interest, songs, commercials, etc., identified for that radio station.” EX1005, ¶94. Herley682 also expressly teaches that the “identification information” of individual objects in the signal (the data in

object/station database 240) can be “used for any number of purposes” including “*generation of play lists* or the like for cataloging a media stream type signal.” EX1005, ¶7.

Based on these disclosures, it would have been obvious to generate a clean play list of songs, as claimed, from the identification information in the updated collection of data in object database 240 (*clean play stream*, see element [1.8]). EX1003, ¶226. Herley682 recognizes that the available “queries are not intended to be limited to the user interface and query examples provided,” and also expressly suggests creating a “play list[]” from the “identification information.” EX1005, ¶¶7, 186. Regarding claim 13, a POSA would have understood a predefined database query that results in a play list to teach *instructions for generating a clean play list*. EX1003, ¶¶227-228.

G. Claims 6 and 14 (“wherein the plurality of video data representations include a plurality of video signatures”)

Herley682 teaches this element. EX1003, ¶¶229-230, 247.

As described above for element [1.2], Herley682’s “pre-computed fingerprints” of known objects stored in metadata/fingerprint database 235 correspond to the recited *plurality of video data representations*. EX1005, ¶¶86, 161, FIG. 5 (step 520). Herley682 describes:

One common method for automatically identifying such objects involves analyzing an input signal, or a predefined portion or segment

of such a signal, to produce a *set of parameters or features that are derived from the signal*. These *parameters or features are then stored as “fingerprints” that uniquely identify such objects*. These fingerprints may then be used to identify subsequent occurrences of objects in a similar signal.

EX1005, ¶5. A POSA would have understood Herley682’s “pre-computed fingerprints” include *a plurality of video signatures* because they uniquely identify video objects. EX1003, ¶229. This is consistent with the plain and ordinary meaning of a “signature.” EX1015 (“6: something ... that serves to set apart or identify”; “7: FINGERPRINT”). This also is consistent with the ’848 patent’s description that “audio signatures” are “sometimes referred to as audio fingerprints.” EX1001, 3:53-57; EX1003, ¶229.

H. Claim 8 (“wherein a video data representation of the plurality of video data representations includes a signal usable for generating video energy corresponding to video energy captured at the media monitoring device”)

Herley682 teaches this element. EX1003, ¶¶231-232.

A POSA would have understood claim 8 to encompass at least a video data representation that includes the ability to play back the captured media. EX1003, ¶232.

As explained above for element [1.2], the *plurality of video data representations* correspond to Herley682’s pre-computed fingerprints, as stored in metadata/fingerprint database 235. EX1005, ¶¶86, 161. Herley682 further teaches,

for example, that unidentified samples are “stored to the metadata/fingerprint database 235 as an ‘unknown object’ entry, *along with either a copy of the sample, or a pointer to a location in the input signal where the sample was taken.*” EX1005, ¶164. Henceforth, the metadata/fingerprint database 235 would include a copy of the unidentified sample. A POSA would have understood that this copy of the sample is a copy of a video signal having the ability to play back the captured media. EX1003, ¶232; *see* EX1005, ¶26 (Herley682 describing that “audio clips or samples, such as a particular song, commercial, speaker, etc.” are provided to the fingerprint engine), ¶93 (“user entry of signal samples is not to be limited to audio clips”), ¶121 (describing that storing a copy of a sample allows the user to listen to the sample). In other words, the stored sample is *a signal usable for generating video energy corresponding to video energy captured at the media monitoring device.* EX1003, ¶232.

VIII. GROUND 2: CLAIMS 1-6 AND 8-14 ARE OBVIOUS IN VIEW OF HERLEY682 AND HERLEY864

Claims 1-6 and 8-14 would have been obvious over Herley682, Seet, and Herley864. EX1003, ¶¶248-264.

A. A POSA Would Have Further Combined Herley864 with Herley682 and Seet

A POSA would have been motivated to combine Herley682 and Seet for the reasons discussed above. Section VII.A; EX1003, ¶248. A POSA would have

found it obvious to further look to Herley864 for multiple reasons. EX1003, ¶248.

Like Herley 682, Herley864 discloses a “system and method for jointly segmenting and identifying repeating objects from one or more broadcast audio streams.” EX1007, 1:16-19. Herley864’s “repeating object extractor” “automatically segments and identifies audio objects such as songs, station identifiers, jingles, advertisements, emergency broadcast signals, etc., in an audio stream.” EX1007, Abstract. Herley864 also describes performing analysis of the content of at least radio or television signals, i.e., “media.” *See, e.g.*, EX1007, 2:13-26; EX1003, ¶¶249-250.

Herley682 and Herley864 are in the same field of endeavor, and even share an inventor. *Compare* EX1005, ¶2, *with* EX1007, 1:16-19; *compare* EX1005, (75), *with* EX1007, (75). 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. EX1003, ¶251. Applying the teachings of Herley864 to the system and method taught in Herley682 would have simply been the combination of prior-art elements according to known methods to yield predictable results. EX1003, ¶251. Similarly, the combination would have merely been the application of known techniques (Herley864’s system/method) to a known system/method ready for improvement (Herley682’s system/method) to yield the predictable result of

identifying objects of interest in a broadcast signal. EX1003, ¶251.

A POSA further would have had a reasonable expectation of success in combining these references because they both teach implementing their inventions in similar operating environments. EX1003, ¶252. In particular, both teach performing the methods thereof in the same “general-purpose computing device constituting an exemplary system for automatically identifying” either “sampled signals” (in Herley682) or “audio objects in an audio stream” (in Herley864). EX1003, ¶252; EX1005, ¶37; EX1007, 4:11-14; *compare* EX1005, FIG. 1, *with* EX1007, FIG. 1. It would have been well within the level of ordinary skill to implement Herley864’s teachings in the system of Herley682 because these exemplary suitable operating environments are described in the same way. EX1003, ¶252; EX1005, ¶¶48-61; EX1007, 4:34-7:15.

B. Claims 1 and 9

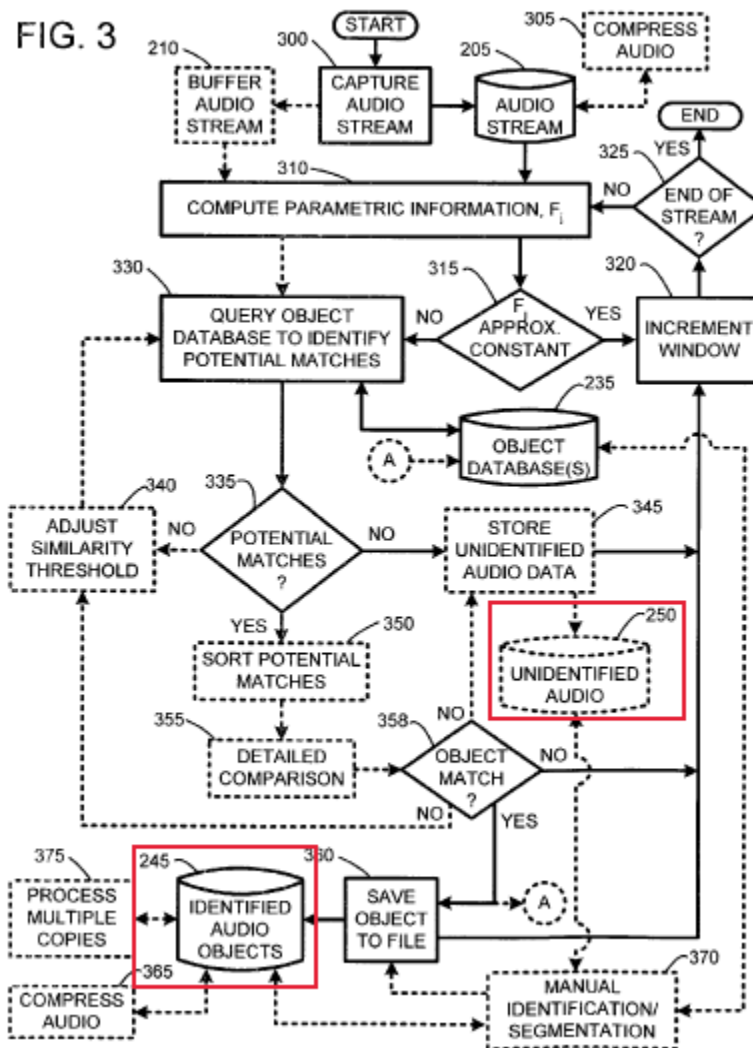
Claims 1 and 9 would have been obvious over the Herley682-Seet combination for all the reasons discussed above. *See* Sections VII.B, VII.C; EX1003, ¶254. Elements [1.8] and [9.7] further would have been obvious in view of Herley682, Seet, and Herley864 for the reasons discussed below.

The Herley682-Seet combination teaches element [1.8], [9.7] as discussed above. *See* Sections VII.B.9, VII.C.8; EX1003, ¶255. To the extent Patent Owner argues that *generating a clean play stream* as recited in element [1.8], [9.7]

requires generation of a new data set, rather than simply updating data in

Herley682's object database, this element would have nonetheless been obvious in view of Herley682, Seet, and Herley864. EX1003, ¶255.

Herley864 teaches an alternative way of handling the unknown objects. As seen in Herley864's Figure 3, reproduced below, "unidentified audio portions 250 are saved to a computer readable medium for later manual ... identification via a user interface module" and the "identified audio objects 245 are stored to a [different] computer readable medium." EX1007, 16:19-45, FIG. 3; EX1003, ¶256-257.



EX1007, FIG. 3 (annotated).

As shown in Herley864's Figure 3, once an unidentified object is identified, it is stored to the identified objects database. EX1007, FIG. 3; EX1003, ¶258. It would have been obvious to implement Herley864's separation of identified and unidentified objects in Herley682's system. EX1003, ¶258.

This would have been a matter of combining prior art elements according to

known methods to yield predictable results and/or a simple substitution of one known element for another to obtain predictable results. In this combination, the “unidentified audio portions 250” correspond to the *raw play stream*, and “identified audio objects 245” correspond to the *clean play stream*. EX1003, ¶259.

C. Claims 4 and 12 (“wherein: scrubbing the raw play stream further comprises using audience data as part of determining whether to change a result of the sequence of content identification results”)

The Herley682-Seet combination in view of Herley864 teaches this element. EX1003, ¶¶260-263.

As to using “audience data” in scrubbing, the ’848 patent describes that “[a]udience data may also be used in the ‘scrubbing’ process.” EX1001, 10:44-45. The ’848 patent continues, with reference to Figure 4:

For example, if “song81” is radically different from the user’s known consumption habits and “song35” is within the user’s known consumption habits, another basis might be provided for suspecting that “song35” rather than “song81” is the correct content ID for sample 29 of Play Stream 410.

EX1001, 10:44-50. The ’848 patent provides examples of “audience data” including: “age, gender, and other demographic information about the audience member,” “media consumption patterns of the audience member,” or “location of the audience member.” EX1001, 5:24-36. A POSA would have understood that the genre of music a user typically listens to is one example of “media consumption

patterns of the audience member” because genre was commonly used to identify and/or classify user consumption habits. EX1003, ¶¶260-262. This also would apply to the genre of video programming. EX1003, ¶263. Thus, a user’s preferred or typical genre is an example of “audience data.” EX1003, ¶263. As explained above for claims 3 and 11, one way in which Herley682 describes identifying its “unknown objects” (*scrubbing*) is by “automatically querying a more authoritative and up-to-date fingerprint database.” EX1005, ¶71. Herley682 also teaches that “when listening to a particular channel, the fingerprints of previously identified objects from that channel are searched first in the database, thereby *speeding the search*.” EX1005, ¶80. As a specific example, Herley682 describes that “when listening to a station that plays pop hits from the 1980’s repetitively it makes sense to search the fingerprints of songs previously identified on that station, before searching the larger database.” EX1005, ¶80. Herley864 similarly teaches limiting searches to results of a particular genre (as opposed to results from a particular station) as a way to reduce search and comparison times. EX1007, 9:10-20, 12:44-52. With this teaching in mind, and in order to “speed[] the search” when querying the additional database for identifying “unknown objects” in Herley682, it would have been obvious to likewise use Herley682’s teaching of first searching previously identified objects from the channel, as modified by Herley864 to first search objects of a user’s preferred or typical genre (*using audience data*), when

querying this additional database. EX1003, ¶263. This would have been the use of known technique to improve similar devices (methods, or products) in the same way. EX1003, ¶263.

D. Dependent claims 2, 3, 5, 6, 8, 10, 11, 13, and 14.

The Herley682-Seet combination teaches the limitations of dependent claims 2, 3, 5, 6, 8, 10, 11, 13, and 14 for the reasons discussed above. *See* Sections VII.D-VII.H; EX1003, ¶264. Therefore, because independent claims would have been obvious in view of Herley682, Seet, and Herley864, dependent claims 2, 3, 5, 6, 8, 10, 11, 13, and 14 likewise would have been obvious over Herley682, Seet, and Herley864. EX1003, ¶264.

IX. GROUND 3: CLAIMS 7 AND 15 ARE OBVIOUS IN VIEW OF HERLEY682 AND SEET (WITH OR WITHOUT HERLEY864) AND WANG060.

Claims 7 and 15 would have been obvious over Herley682 and Seet (with or without Herley864) and Wang060. EX1003, ¶¶265-271.

A. Claims 7 and 15 (“wherein a video data representation of the plurality of video data representations includes a plurality of signal landmarks”)

Wang060 teaches this element. EX1003, ¶¶84, 265-271.

Like Herley682, Wang060 teaches a method for recognizing a media sample, such as video, by matching the media sample to one in a database index of known media files. EX1008, ¶18. Wang060’s database contains “fingerprints representing features at particular locations of the indexed media files.” EX1008,

¶18. The “particular locations” are preferably “reproducibly computable locations ... called ‘landmarks.’” EX1008, ¶20. Wang060 explicitly recognizes that its “method is particularly useful for recognizing audio samples,” but also notes that its method “can be applied to any type of media samples ... including ... video.” EX1008, ¶¶21, 37.

A POSA would have further understood that *signal landmarks* were a known method by which video data could be analyzed. EX1003, ¶268; EX1008, ¶¶37, 51. This is consistent with the ’848 patent’s description that “the uploaded audio data may include parameters useable to calculate audio signatures (sometimes referred to as audio fingerprints) and landmarks or other values useful in the content recognition process” without further elaboration. EX1001, 3:53-57.

Such *landmarks* are expressly taught in Wang060. “Landmarks occur at specific timepoints of the sound and have values in time units offset from the beginning of the file, while fingerprints characterize the sound at or near a particular landmark.” EX1008, ¶49, FIG. 4. Wang060 teaches a method for constructing a searchable sound database index, in which “landmarks are first computed, and then fingerprints are computed at or near the landmarks.” EX1008, ¶48, FIG. 3; *see also* EX1008, ¶49 (“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.”), ¶74 (“An index set for a given sound

recording is a list of pairs of values (fingerprint, landmark)”). Wang060 also teaches using the index to perform sound recognition. EX1008, ¶79.

Based on Wang060’s disclosure, a POSA would have been motivated to implement Herley682/Seet’s and Herley682/Seet/Herley864’s fingerprint-based content identification method also using landmarks, as taught in Wang060. EX1003, ¶270. This would have been the use of known technique to improve similar devices (methods, or products) in the same way. EX1003, ¶270. A POSA would have had a reasonable expectation of success in making this modification to Herley682/Seet and Herley682/Seet/Herley864. EX1003, ¶270. Each of Herley682/Seet, Herley682/Seet/Herley864, and Wang060 are fingerprint-based methods of identifying content in a media, e.g., video, stream. Wang060 further describes that “specific technical processes used to implement the landmarking schemes of the invention are known in the art.” EX1008, ¶51. It would have been within the skill of a POSA to implement these known landmarking schemes into the Herley682/Seet and Herley682/Seet/Herley864 content identification methods. EX1003, ¶270.

Accordingly, based on claim 7’s dependence from claim 1, and based on claim 15’s dependence from claim 9, and the explanations regarding these independent claims above, claims 7 and 15 would have been obvious in view of either (1) Herley682, Seet, and Wang060; or (2) and Herley682, Seet, Herley864,

and Wang060.

X. PETITIONER IS UNAWARE OF ANY SECONDARY CONSIDERATIONS OF NON-OBVIOUSNESS

Petitioner is unaware of any alleged evidence of secondary considerations, or that Patent Owner has made any non-conclusory assertions of any secondary considerations of non-obviousness. EX1003, ¶¶272-273.

XI. DISCRETIONARY DENIAL IS NOT APPROPRIATE.

A. *Fintiv* does not support discretionary denial.

Fintiv does not support discretionary denial. *Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11, at 6 (P.T.A.B. Mar. 20, 2020) (precedential) (“*Fintiv*”). If the Board institutes trial, Roku will not pursue, in a parallel district court proceeding, the same grounds as in the petition or any grounds that could have reasonably been raised in the IPR. *See Sotera Wireless, Inc. v. Masimo Corp.*, IPR2020-01019, Paper 12, at 18-19 (P.T.A.B. Dec. 1, 2020) (precedential as to § II.A). Under the Interim Procedure, “the PTAB will not discretionarily deny institution in view of parallel district court litigation where a petitioner presents a [*Sotera*] stipulation,” as Roku presents here. EX1016, 3, 7.

B. Denial under §325(d) is not appropriate.

Denial under §325(d) and *Advanced Bionics* is not appropriate. Under *Advanced Bionics* part one, Herley682, Seet, and Herley864 were not before the Examiner during prosecution. *See* EX1001, (56); EX1002. Wang060 was cited but

not applied by the Examiner during prosecution. - *See* EX1002. Thus, the Examiner during original prosecution did not consider the combinations or arguments presented in the Petition and *Advanced Bionics* part one is not met.

Even if part one were met (which it is not), *Advanced Bionics* part two requires a consideration of whether the Office erred. Here, the Examiner clearly erred by not considering or applying Herley682, Seet, or Herley864 in rejecting the claims. *See* Sections VII-IX; EX1003, ¶¶144-271.

XII. CERTIFICATION THAT THE PATENT MAY BE CONTESTED VIA *INTER PARTES* REVIEW BY THE PETITIONER AND STANDING (37 C.F.R. § 42.104(A))

Roku certifies that the '848 patent is available for IPR and Roku is not barred or estopped from requesting IPR of any '848 patent claim.

XIII. MANDATORY NOTICES (37 C.F.R. § 42.8(A)(1))

Real Party-in-Interest (37 C.F.R. § 42.8(b)(1)): The real party-in-interest is Roku, Inc,

Related Matters (37 C.F.R. § 42.8(b)(2)): The '848 patent is the subject of the following pending civil actions:

- *Anonymous Media Research Holdings LLC v. Samsung Electronics Co., Ltd. et al.*, Case No. 2:23-cv-00439 (E.D. Tex.), filed September 22, 2023; and
- *Anonymous Media Research Holdings LLC v. Roku, Inc.*, Case No.

1:23-cv-01143 (W.D. Tex.), filed September 22, 2023.

Petitioner is concurrently filing Petitions for *Inter Partes* Review of the following U.S. Patents, which are also the subject of the above-listed civil actions:

- 8,296,791 (IPR2024-01053);
- 8,510,768 (IPR2024-01054); and
- 10,572,896 (IPR2024-01056).

Lead and Back-up Counsel (37 C.F.R. § 42.8(b)(3)) for Roku are

Lead Counsel	Back-Up Counsel
Lestin L. Kenton (Reg. No. 72,314) STERNE, KESSLER, GOLDSTEIN & FOX PLLC 1101 K Street, NW, 10th Floor Washington, DC 20005 202-772-8594 (telephone) 202-371-2540 (facsimile) lkenton-PTAB@sternekessler.com	Christopher R. O'Brien (Reg. No. 63,208) Jennifer Meyer Chagnon (Reg. No. 55,440) Jessica J. Harrison (Reg. No. 72,780) Syed S. Ahmed (Reg. No. 64,587) STERNE, KESSLER, GOLDSTEIN & FOX PLLC 1101 K Street, NW, 10th Floor Washington, DC 20005 202-371-2600 (telephone) 202-371-2540 (facsimile) cobrien-PTAB@sternekessler.com jchagnon-PTAB@sternekessler.com jharrison-PTAB@sternekessler.com syedahmed-PTAB@sternekessler.com

Service Information (37 C.F.R. § 42.8(b)(4)): Please direct all correspondence regarding this Petition to counsel at the above addresses and PTAB@sternekessler.com. Roku consents to service by email at the addresses above.

Procedural Statements: This Petition is filed in accordance with 37 C.F.R.

§ 42.106(a). Concurrently filed herewith are a Power of Attorney and Exhibit List under 37 C.F.R. §§ 42.10(b) and 42.63(e), respectively. The required fee is paid through Deposit Acct. No. 19-0036 (Customer ID No. 45324). The Office is authorized to charge any fee deficiency, or credit any overpayment, to Deposit Acct. No. 19-0036 (Customer ID No. 45324).

XIV. CONCLUSION

The challenged claims are unpatentable as obvious and IPR is warranted.

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX PLLC

/Lestin Kenton/

Lestin L. Kenton
Registration No. 72,314
Lead Attorney for Petitioner

Date: July 15, 2024

1101 K Street, NW, 10th Floor
Washington, DC 20005
(202) 371-2600

CERTIFICATE OF WORD COUNT (37 C.F.R. § 42.24(d))

1. This Petition for *Inter Partes* Review complies with the type-volume limitation of 14,000 words, comprising 13,495 words, excluding the parts exempted by 37 C.F.R. § 42.24(a)(1).

2. This Petition for *Inter Partes* Review complies with the general format requirements of 37 C.F.R. § 42.6(a) and has been prepared using Microsoft® Word 2016 in 14-point Times New Roman font.

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX PLLC

/Lestin Kenton/

Lestin L. Kenton
Registration No. 72,314
Lead Attorney for Petitioner

Date: July 15, 2024

1101 K Street, NW, 10th Floor
Washington, DC 20005
(202) 371-2600

CERTIFICATE OF SERVICE (37 C.F.R. §§ 42.6(e), 42.105(a))

I certify that the above-captioned **PETITION FOR *INTER PARTES***
REVIEW FOR U.S. PATENT NO. 10,719,848 and associated Exhibits 1001-
1016 and 1019-1020 were served in their entireties upon the Patent Owner on July
15, 2024, via FedEx® Express at the following addresses:

MASCHOFF BRENNAN GILMORE ISRAELSEN & MAURIEL LLP
1389 Center Drive, Suite 300
Park City, UT 84098
UNITED STATES
Patent Owner's Correspondence Address of Record for U.S. Patent No.
10,719,848

Respectfully submitted,

STERNE, KESSLER, GOLDSTEIN & FOX PLLC

/Lestin Kenton/

Lestin L. Kenton
Registration No. 72,314
Lead Attorney for Petitioner

Date: July 15, 2024

1101 K Street, NW, 10th Floor
Washington, DC 20005
(202) 371-2600

APPENDIX A: CHALLENGED CLAIMS

Limitation	Claim
1.P	A media measurement method comprising:
1.1	(a) receiving over a network, at one or more computers, data for a sequence of video data samples, the video data samples comprising representations of video data captured at a media monitoring device, the plurality of video data samples being submitted over the network;
1.2	(b) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers;
1.3	(c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of video data samples;
1.4	wherein: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective video data samples in the sequence of video data samples and respective video data representations in the electronic database;
1.5	a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective video data representation that is determined to be a respective likely match with a respective one of the video data samples; or (ii) an indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database; and
1.6	the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database;

1.7	(d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and
1.8	(e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.
2	The method of claim 1 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.
3	The method of claim 1 wherein: scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content of identification results.
4	The method of claim 1 wherein: scrubbing the raw play stream further comprises using audience data as part of determining whether to change a result of the sequence of content of identification results.
5	The method of claim 1 wherein a clean play list is generated from the clean play stream.
6	The media measurement method of claim 1 wherein the plurality of video data representations include a plurality of video signatures.
7	The media measurement method of claim 1 wherein a video data representation of the plurality of video data representations includes a plurality of signal landmarks.
8	The media measurement method of claim 1 wherein a video data representation of the plurality of video data representations includes a signal usable for generating video energy corresponding to video energy captured at the media monitoring device.
9.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 one or more computers being coupled to a network to receive, over the network, data for a sequence of video data samples, the video data samples comprising representations of video data captured at a media monitoring device, the plurality of video data samples being submitted over the network, the instructions comprising instructions for:
9.1	(a) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers
9.2	(b) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of video data samples;
9.3	wherein: the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective video data samples in the sequence of video data samples and respective video data representations in the electronic database;
9.4	a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective video data representation that is determined to be a respective likely match with a respective one of the video data samples; or (ii) an indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database; and
9.5	the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective video data sample and a video data representation in the electronic database;
9.6	(c) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a

	result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and
9.7	(d) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.
10	The computer program product of claim 9 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.
11	The computer program product of claim 9 wherein scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content of identification results.
12	The computer program product of claim 9 wherein scrubbing the raw play stream further comprises using audience data as part of determining whether to change a result of the sequence of content of identification results.
13	The computer program product of claim 9 wherein the computer program product further comprises instructions for generating a clean play list from the clean play stream.
14	The computer program product of claim 9 wherein the plurality of video data representations include a plurality of video signatures.
15	The computer program product of claim 9 wherein a video data representation of the plurality of video data representations includes a plurality of signal landmarks.