

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.
Petitioner

v.

ANONYMOUS MEDIA RESEARCH HOLDINGS, LLC,
Patent Owner

Patent No. 10,719,848

DECLARATION OF TIMOTHY WAHLERS

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	BACKGROUND AND QUALIFICATIONS.....	2
III.	MATERIALS CONSIDERED AND SUMMARY OF OPINIONS.....	10
IV.	PERSON OF ORDINARY SKILL IN THE ART	13
V.	THE '848 PATENT.....	15
VI.	CLAIM CONSTRUCTION	24
VII.	OVERVIEW OF THE PRIOR ART	25
A.	Overview of Steuer750 (EX1005) and Steuer768 (EX1006)	25
B.	Overview of Neumeier (EX1007).....	27
VIII.	DISCLOSURE OF CERTAIN CLAIM LIMITATIONS IN THE '848 PATENT CLAIMS AND ITS PARENT APPLICATIONS.....	31
A.	The Priority Applications Do Not Support The Video Data Features of The Challenged Claims	32
1.	The Priority Applications Do Not Provide Support for the Claimed "Video Data Samples" and Related Video Features	32
a)	The '836 Provisional Application Does Not Support the Claimed Video Data Features.....	37
b)	The '330, '604, '870, and '811 Priority Non- Provisional Applications Do Not Provide Adequate Support For The Challenged Claims	51
c)	The Combination of The '836 Provisional Application's Disclosure with the Priority Non- Provisional Applications Does Not Support the Video Data Features of the Challenged Claims	61
2.	Additional Materials Mentioned in The Priority Non- Provisional Applications Do Not Support the Video Data Features of the Claims of the '848 Patent.....	63

3.	New Disclosures Added to The '204 Application That Issued as The '896 Patent	66
IX.	THE PRIOR ART DISCLOSES AND/OR SUGGESTS ALL RECITED FEATURES OF CLAIMS 1-15 OF THE '848 PATENT.....	70
A.	Grounds 1 and 2: Ground 1: The Combination of Steuer750 and Neumeier Discloses and/or Suggests Claims 1-15 Ground 2: The Combination of Steuer768 and Neumeier Discloses and/or Suggests Claims 1-15	70
1.	Claim 1	71
a)	[pre] A media measurement method comprising:	71
b)	(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;	75
c)	(b) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers;	92
d)	(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;	96
e)	(d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and.....	111

f)	(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.....	119
2.	Claim 2.....	122
a)	The method of claim 1 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.....	122
3.	Claim 3.....	126
a)	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.....	126
4.	Claim 4.....	128
a)	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.	128
5.	Claim 5.....	130
a)	The method of claim 1 wherein a clean play list is generated from the clean play stream.....	130
6.	Claim 6.....	135
a)	The media measurement method of claim 1 wherein the plurality of video data representations include a plurality of video signatures.	135
7.	Claim 7.....	136
a)	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.....	136
8.	Claim 8.....	138

- a) 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.....138
- 9. Claim 9 140
 - a) [pre] A computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers, the one or more computers being coupled to a network to receive, over the network, data for a sequence of 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:.....140
 - b) (a) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers; 143
 - c) (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; 144
 - d) (c) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and.....146

e)	(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.....	147
10.	Claim 10.....	147
a)	The computer program product of claim 9 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.	147
11.	Claim 11.....	148
a)	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.	148
12.	Claim 12.....	148
a)	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.	148
13.	Claim 13.....	149
a)	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.	149
14.	Claim 14.....	149
a)	The computer program product of claim 9 wherein the plurality of video data representations include a plurality of video signatures.	149
15.	Claim 15.....	150

a)	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.....	150
X.	CONCLUSION.....	151

I, Timothy Wahlers, declare as follows:

I. INTRODUCTION

1. I have been retained by Samsung Electronics Co., Ltd. (“Petitioner”) as an independent expert consultant in this proceeding before the United States Patent and Trademark Office (“PTO”) regarding U.S. Patent No. 10,719,848 (“the ’848 patent”) (EX1001). I have been asked to consider whether a person of ordinary skill in the art would have understood that the named inventors of the ’848 patent were in possession of the subject matter recited in certain claims of the ’848 patent at the time the priority applications were filed based on the disclosures of the priority applications. I have also been asked to consider whether prior art references disclose or suggest the features recited in claims 1 through 15 (“the challenged claims”) of the ’848 patent. My opinions are set forth below.

2. I am being compensated at my standard consulting rate for the time I spend on this matter. My compensation is in no way contingent on the nature of my findings, the presentation of my findings in testimony, or the outcome of this or any other proceeding. I have no other interest in this proceeding.

II. BACKGROUND AND QUALIFICATIONS

3. In this section, I have summarized my educational background, career history, publications, and other relevant qualifications. My Curriculum Vitae, which states my qualifications more fully, is filed as a separate Exhibit (EX1003).

4. I am the Owner and Principal at The Wahlers Group which offers technology, management, and expert witness consulting services to the digital video, cable, satellite, software, and consumer electronics industries. With over 35 years of industry experience, I specialize in digital video platforms, client and cloud-based platforms, content management via cloud systems, video-on-demand, and content streaming.

5. In 1981, I earned my B.Sc. in Nuclear Engineering from the University of Wisconsin, Madison. During my studies, I conducted post-graduate research in the field of fusion energy and microcomputer process control. I developed microprocessor monitoring software for the linear plasma experimental laboratory and conducted reactor core fuel analysis for the University Nuclear Reactor.

6. After completing my B.Sc., I worked in the field of nuclear engineering for 5 years, focusing on nuclear plant process computer monitoring and control. I developed software for various process computer systems and emergency response facilities. I also designed and developed PC-controlled laser disc application software for a simulated journey through a nuclear containment environment.

Additionally, I contributed to the development of an exhibit for the electric industry, which was showcased at the Chicago Science and Industry Museum. This work ignited my interest in digital media platforms, multimedia, and their use in education and entertainment.

7. In 1987, at the beginning of the digital media era, I joined OptImage Interactive Services as the Chief Engineer. OptImage was a joint venture of N.V. Philips, R.R. Donnelly, Sun Microsystems, and Microware. Over a span of 10 years, I was promoted to Vice President, Chief Technology Officer, and eventually President. My team was tasked with developing pioneering technology and tools for CD-ROM and CD-I, VideoCD, DVD, and Blu-ray digital MPEG video multimedia platform standards, which helped to advance the interactive television industry. During my tenure at OptImage, I led the development of software tools, interfaces, scripting languages, and applications that influenced the development of many of the early interactive video platforms. I have served and contributed to many industry standards bodies. I also led a project for an industry consortium that produced the world's first commercially sold MPEG A/V encoding solution.

8. In 1996, I founded Vidiom Systems Corporation, a company specializing in professional services and digital video products. As president and sole owner, I steered the company's technical direction, cultivated business relationships, and oversaw the development of our marketed products and service

contracts. Vidiom provided software development, integration, and testing services to leading consumer electronics, cable, satellite, and telecommunications companies across North America and Europe. My multifaceted role involved direct client engagement, serving as an architect, analyst, software developer, project lead, and management consultant. Our expertise centered on software for various platforms including PCs, mobile devices, set-top boxes, and TVs, as well as head-end systems for video and data delivery over broadband and IP networks. A key focus was the development of Electronic Program Guides and media-centric applications, including those synchronized with video delivery, audience measurement, and video-on-demand services.

9. From 1998 to 2001, while at Vidiom, I was contracted by CableLabs to serve as the Director of OpenCable Software for the cable industry. In this pivotal role, I led the development of the OpenCable Application Platform (OCAP) Java standard, functioning as its architect, analyst, and program lead. OCAP was designed as a standard platform for cable industry set-top boxes, enabling interoperability and supporting features such as program guides and DVR functionality. The platform's impact was substantial, achieving widespread adoption with deployments in over 20 million televisions and set-top boxes, significantly shaping the landscape of cable technology.

10. Vidiom grew to become the preeminent service provider in this space with over 110 employees and 60 subcontractors before I sold the company to the European entity, Advanced Digital Broadcast (ADB) in 2006.

11. In 2008, I reacquired Vidiom's assets from ADB to establish EnableTV, a new venture that built upon Vidiom's mission and client base while expanding into cutting-edge technologies. EnableTV developed innovative products for the evolving digital television landscape, including: 1) an interactive television dashboard, 2) the OCAP Software Stack, 3) the VCERT testing platform, 3) an EnableTV application storefront, and 4) a next-generation content browser as a part of an IPTV program guide. These technologies positioned EnableTV at the forefront of the digital television industry, serving our established clientele while pushing boundaries in interactive and IP-based TV solutions.

12. In 2012, I sold the EnableTV assets and established The Wahlers Group, a strategic management, business development, and tactical consultancy specializing in digital video, interactive television platforms, and cloud-based content management platforms.

13. I have given many invited lectures and courses around the world on topics related to the software development lifecycle, processes, and DevOps. I have spoken at various conferences, including CES, NAB, Cable Show, and IBC. As part

of my management consultancy, I have conducted training sessions for clients in software development and delivery.

14. In 2018, I co-founded Quantum Radius and currently serve as a board member. I provide technical, intellectual property, financing, and operations management services to the company. Quantum Radius is developing distributed, ultra-low-latency video processing technology for the VR/AR ecosystem. The technology is designed for simulation, medical, government, remote piloting, and eSports markets.

15. I have served on various company Boards of Directors and Technical Advisory Committees, including MyPhonix, MystroTV, RapidView Inc., and Quantum Radius.

16. I am honored to have received an industry award recognizing my contributions to my field, the *2006 ITVT Award for Leadership in Interactive Television*, <https://itvt.com/previous-awards/>, for my groundbreaking work in the digital video and interactive television industry.

17. Through my many decades of experience in software development and establishing companies in the media technology space, Automatic Content Recognition and audience measurement have been frequently recurring areas of interest. Partners and clients alike have developed ACR solutions that employ watermarking and fingerprinting techniques. The services offered through my many

endeavors in the media technology industry have included software programming, program management, business development, partnering, and investment/merger opportunities, and those services have been provided to numerous prominent entities in the field of ACR, including:

- a. Philips (1997 – 2008): I worked for OptImage, a company mostly owned by Philips, from 1987 to 1996 before starting my own company, Vidiom Systems. In the late 1990s, Philips had a content identification group that I was familiar with. They made significant advances in content recognition technology. We looked into integrating their technology with ours, focusing on targeting the US cable industry. This group was spun off from Philips in 2008 as Civolution.
- b. Civolution; (2011 - 2013): We pursued the development of a demo of a 2nd screen application synchronized to video playing on a cable system OCAP Set-top.
- c. Enswers, Inc., which specialized in video/image fingerprinting and was later acquired by Gracenote (2012 – 2013): I evaluated Enswers ACR technology for investor, Stage 1 Ventures. The ACR technology utilized digital video frame fingerprinting within the TV. Various business relationships were considered

utilizing EnableTV core technology to enhance Enswers offering for use in the cable industry and Verizon FIOS set-top boxes.

- d. CableLabs (1997 – 2012): One of the largest customers of Vidiom and EnableTV. I served as Director of OpenCable, Software in 1998 to 2000, and was contracted to provide a team, to assist the cable industry with software standards and reference implementations, from 2000 to 2012. I led the initial phase of the OCAP software platform definition and development. In that role, most industry vendors (including those involved in ACR technology), explored technology relationships with the cable industry and implementations on the OCAP platform.
- e. OpenTV/Nagra/Kudelski (1997 – 2008): A very large and first customer of Vidiom. I built a team under contract to develop tools and applications for OpenTV. We maintained their entire documentation collection from 1997 to 2001. ACR applications, mostly involving watermarking, were investigated on multiple occasions.
- f. The Nielsen Company (1995 – 2014): Nielsen was a major contributor to cable industry projects I led or participated in. I reviewed and commented on a proposed audience measurement

data protocol specification. We explored the use of our technology to assist in Nielsen audience measurement solutions by tagging and watermarking audio and video content. We developed prototype software to generate media viewing log data for external delivery to Nielsen's platform.

- g. FourthWall Media (2010 – 2013): I worked with FourthWall on a project exploring television audience data collection on the OCAP platform.
- h. Gracenote, acquired by Nielsen in 2017 (2004): Explored joint projects and applications in media metadata technology. Using ACR to feed 2nd screen applications. The focus was on tagging and watermarking.
- i. IBM (1996-1999): Multimedia platform work. Content identification, watermarking of both analog and digital video.

III. MATERIALS CONSIDERED AND SUMMARY OF OPINIONS

18. The opinions contained in this Declaration are based on the documents I reviewed and my professional judgment, as well as my education, experience, and/or knowledge regarding technologies relating to, among other things, networking technologies.

19. In forming my opinions expressed in this Declaration, I reviewed the following materials (and any other materials that I reference in this Declaration):

EX1001	U.S. Patent No. 10,719,848
EX1003	Curriculum Vitae of Timothy Wahlers
EX1004	File History of U.S. Patent Application No. 16/660,615 (U.S. Patent No. 10,719,848)
EX1005	U.S. Publication No. 2005/0267750 to Steuer <i>et al.</i>
EX1006	U.S. Patent No. 8,510,768 to Steuer <i>et al.</i>
EX1007	U.S. Publication No. 2010/0306808 to Neumeier <i>et al.</i>
EX1008	Computer-generated Redline Between Specification of U.S. Publication No. 2005/0267750 to Steuer <i>et al.</i> (EX1005) and Specification of U.S. Patent No. 8,510,768 to Steuer <i>et al.</i> (EX1006)
EX1009	Computer-generated Redline Between Claims 1-15 of U.S. Patent No. 10,719,848 (EX1001) and Claims 1-15 of U.S. Patent No. 8,510,768 (EX1006)
EX1010	File History of U.S. Provisional Application No. 60/574,836
EX1011	File History of U.S. Patent Application No. 11/139,330
EX1012	File History of U.S. Patent Application No. 12/954,604
EX1013	File History of U.S. Patent Application No. 13/914,870

EX1014	File History of U.S. Patent Application No. 14/175,811
EX1015	File History of U.S. Patent Application No. 16/403,204 (U.S. Patent No. 10,572,896)
EX1016	U.S. Publication No. 2007/0174059 to Rhoads <i>et al.</i>
EX1017	Computer-generated Redline Between As-Filed Specifications in U.S. Patent Application No. 11/139,330 (EX1011) and U.S. Patent Application No. 12/954,604 (EX1012)
EX1018	Computer-generated Redline Between As-Filed Specifications in U.S. Patent Application No. 11/139,330 (EX1011) and U.S. Patent Application No. 13/914,870 (EX1013)
EX1019	Computer-generated Redline Between As-Filed Specifications in U.S. Patent Application No. 11/139,330 (EX1011) and U.S. Patent Application No. 14/175,811 (EX1014)
EX1020	U.S. Patent Publication No. 2002/0083060
EX1021	WO2004/044820
EX1022	WO2004/030341
EX1023	WO2004/040475
EX1028	Comparison between video-related paragraph in U.S. Provisional Application No. 60/574,836 and video- related paragraph in U.S. Patent Application No. 16/403,204
EX1030	TRECVID 2004 – An Overview (February, 2005)
EX1031	Guidelines for the TRECVID 2004 Evaluation (February, 2005)
EX1032	TRECVID-2005: Search Task (Dublin City University and NIST)

20. In support of my opinions, I have taken into account how a person of ordinary skill in the art (as defined below in Section IV) would have understood the

claims and the specification of the '848 patent at the time of the alleged invention. My opinions take into account how a person of ordinary skill in the art would have understood the '848 patent and the state of the art at the time of the alleged invention.

21. Based on my experience and expertise, it is my opinion (i) that a person of ordinary skill in the art would not have understood that the named inventors of the '848 patent were in possession of the subject matter recited in certain claims of the '848 patent at the time the priority applications were filed based on the disclosures of the priority applications and (ii) that certain references disclose and/or suggest all the features recited in the challenged claims of the '848 patent, as I discuss in detail below.

IV. PERSON OF ORDINARY SKILL IN THE ART

22. I have considered the time of the alleged invention for the '848 patent to be around the mid-2019 timeframe, including up to May 3, 2019, the filing date of U.S. Patent Application No. 16/403,204 ("the '204 application"), which issued as U.S. Patent Application No. 10,572,896 ("the '896 patent"). (EX1001 at Cover.) U.S. patent application no. 16/660,615 ("the '615 application") which issued as the '848 patent, is a continuation of the '204 application, which added new matter not contained in any of the prior applications. I may refer to this time frame in this declaration as the time of the alleged invention, the relevant time, or the time of the '848 patent.

23. Based on my knowledge and experience, I understand what a person of ordinary skill in the art ("POSITA") would have known at the time of the alleged invention of the '848 patent. In my opinion, based on my review of the '848 patent, the technology, the educational level and experience of active workers in the field, the types of problems faced by workers in the field, the solutions found to those problems, the sophistication of the technology in the field, and drawing on my own experience, a person of ordinary skill in the art would have had an undergraduate degree in Electrical Engineering, Computer Engineering, Computer Science or a related field along with at least 4 years of work experience in the field of media-related software development, or in the alternative, equivalent experience, such as 6

years of work or research experience in the field of media-related software development. (*See, e.g.*, EX1001 at 1:20-62 (discussing the background of the '848 patent), 1:64-2:22 (discussing a summary of the disclosure).) More education can substitute for practical experience and *vice versa*. For example, graduate level education could supplement relevant work experience and significant work experience could supplement education experience.

24. My analysis of the '848 patent and my opinions in this declaration are from the perspective of one of ordinary skill in the art, as I have defined it above, at the time of the alleged invention for that patent.

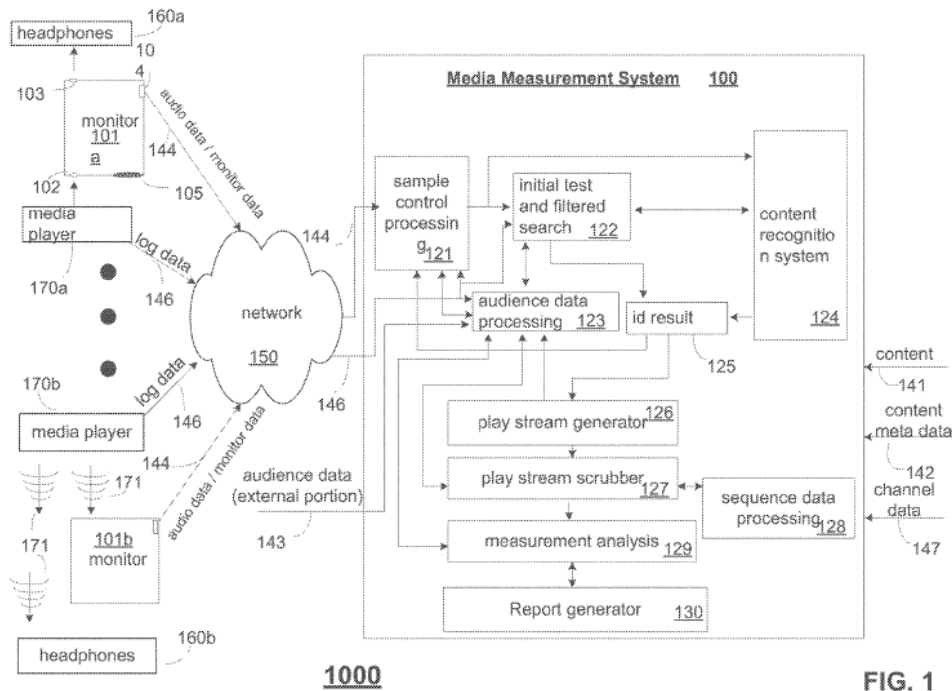
25. I note that my opinions in this declaration do not change even if the time of the alleged invention for the '848 patent is considered to be around the late 2010 timeframe, including up to the November 24, 2010 filing date of U.S. Patent Application No. 12/954,604 ("the '604 application"), which I understand is an intervening application to which the '615 application claims priority.

26. I further note that my opinions in this declaration do not change even if the time of the alleged invention for the '848 patent is considered to be around the mid-2004 timeframe, including up to the May 27, 2004 filing date of U.S. Provisional Patent Application No. 60/574,836 ("the '836 provisional application"), which I understand is the earliest application to which the '615 application claims priority.

V. THE '848 PATENT

27. The '848 patent is directed to media usage monitoring and measurement systems and methods. (EX1001 at Title, Abstract, 1:65-2:22.) The media measurement method, in particular, relates to recognizing content being consumed by a user and performing additional actions with the recognized content, such as creating a clean play list for the user, identifying a source of the content, and deducing play-altering activities a user engages in (e.g., fast-forwarding, rewinding).

28. As shown in Figure 1 below, the '848 patent discloses its media measurement methods in the context of capturing audio data from content consumed by a user whereby a monitoring device, such as monitors 101a/101b, captures the audio data. (EX1001 at 3:5-44, FIG. 1.) The audio data, or a signature calculated from the audio data, is then sent via a network to a system 100 to be recognized by comparing it against data stored in a database containing data representing known content. (EX1001 at 3:45-4:6, 4:22-57.) System 100 identifies the content and generates content identification results using the content recognition system 124. (EX1001 at 4:58-6:23.)



(EX1001 at FIG. 1.)

29. Next, the content identification results are sent to play stream generator 126 to generate a raw play stream 410, as shown in Figure 4 below. (EX1001 at 6:24-37, 8:13-34, 9:25-45.) The raw play stream 410 “itemizes a sequence of analyzed samples and relates that sequence to corresponding identification results.” (EX1001 at 9:25-45.) Once the raw play stream 410 is generated, the media measurement method includes performing additional processing and analysis of the raw play stream 410, such as scrubbing the play stream to generate a clean play stream 450, identifying a channel associated with the recognized content, or deducing play-altering activities as shown in the clean play list 590 in figure 5. (EX1001 at 6:24-37, 8:13-34, 8:42-50 (describing the process of scrubbing as

including “reviewing for missing or incorrect data”), 9:46-51 (describing the scrubbing process as including “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 identification”), 9:52-10:50, 11:8-28.)

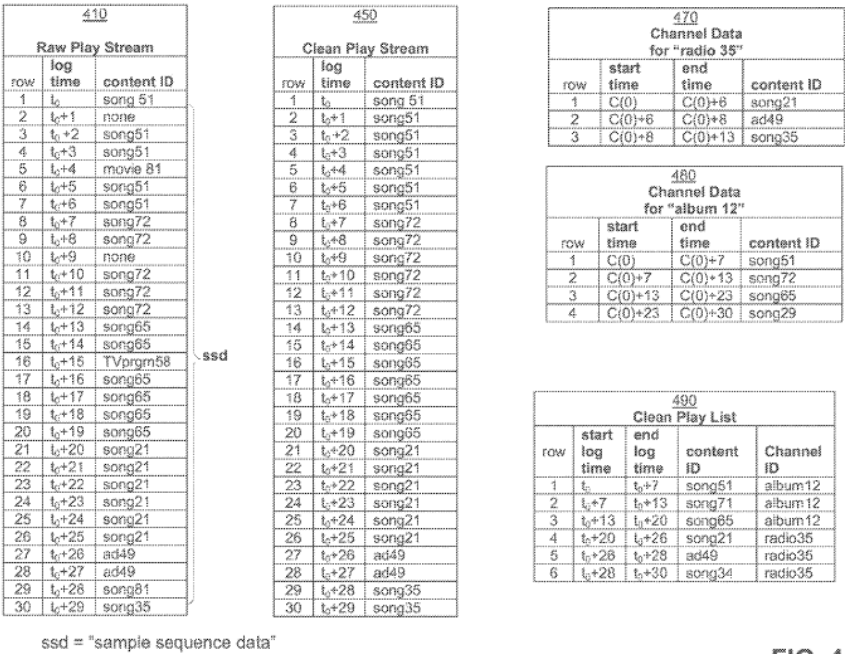


FIG. 4

(EX1001 at FIG. 4.)

510 Clean Playstream			
row	log time	content ID	content offset
1	t_0	TVprgm 27	(0)
2	t_0+1	TVprgm 27	(0)+1
3	t_0+2	none	n/a
4	t_0+3	none	n/a
5	t_0+4	none	n/a
6	t_0+5	TVprgm 27	(0)+2
7	t_0+6	TVprgm 27	(0)+3
8	t_0+7	TVprgm27	(0)+9
9	t_0+8	TVprgm27	(0)+10
10	t_0+9	TVprgm27	(0)+9
11	t_0+10	TVprgm27	(0)+10
12	t_0+11	none	n/a
13	t_0+12	TVprgm27	(0)+29
14	t_0+13	TVprgm27	(0)+30
15	t_0+14	ad23	(0)
16	t_0+15	none	n/a
17	t_0+16	ad25	(0)+2
18	t_0+17	TVprgm28	(0)
19	t_0+18	TVprgm28	(0)+1
20	t_0+19	TVprgm28	(0)+2

560 Clean Play List				
row	start log time	end log time	content ID	content offset
1	t_0	t_0+2	TVprgm27	(0)
2	t_0+2	t_0+5	none	n/a
3	t_0+5	t_0+7	TVprgm27	(0)+2
4	t_0+7	t_0+9	TVprgm27	(0)+9
5	t_0+9	t_0+11	TVprgm27	(0)+9
6	t_0+11	t_0+12	none	n/a
7	t_0+12	t_0+14	TVprgm27	(0)+29
8	t_0+14	t_0+15	ad23	(0)
9	t_0+15	t_0+16	none	n/a
10	t_0+16	t_0+17	ad25	(0)+2
11	t_0+17	t_0+20	TVprgm28	(0)

570 Channel Data for TV C4			
row	start time	end time	content ID
1	C(0)	C(0)+31	TVprgm27
2	C(0)+31	C(0)+34	ad23
3	C(0)+34	C(0)+37	ad24
4	C(0)+37	C(0)+40	ad25
5	C(0)+40	C(0)+50	TVprgm28

580 Clean Play List with Deduced Actions				
row	log time	end log time	content ID and <deduced action>	content offset
1	t_0	t_0+2	TVprgm27	(0)
2	t_0+2	t_0+5	<pause for 3 units>	n/a
3	t_0+5	t_0+7	TVprgm27	(0)+2
4	t_0+7	t_0+7	<skip frwrd 5 units>	n/a
5	t_0+7	t_0+9	TVprgm27	(0)+9
6	t_0+9	t_0+9	<skip bckwrd 2 units>	n/a
7	t_0+9	t_0+11	TVprgm27	(0)+9
8	t_0+11	t_0+12	<fast frwrd 18 units>	n/a
9	t_0+12	t_0+14	TVprgm27	(0)+29
10	t_0+14	t_0+15	ad23	(0)
11	t_0+15	t_0+16	<fast frwrd 7 units>	n/a
12	t_0+16	t_0+17	ad25	(0)+2
13	t_0+17	t_0+20	TVprgm28	(0)

FIG. 5

(EX1001 at FIG. 5.)

30. According to the '848 patent, a raw play stream “relat[es] a sequence of samples to content identification results,” and is then “processed by play stream scrubber 127 which analyzes the sequence of sample identification results utilizing sequence data” in order to “generate[] a clean play stream and clean play list.” (EX1001 at 6:24-35.) More specifically, as I noted earlier, the '848 patent explains that raw play stream 410 may be “scrubbed” by analyzing “sample sequence data” in order to “locate and correct any identifications that are likely to be incorrect or to supply missing identifications.” (EX1001 at 9:46-51; *see also id.*, 8:29-34, 8:47-50, 10:20-22.) The output of the process of scrubbing raw play stream 410 is clean play

stream 450, which reflects changes made as a result of the scrubbing process.

(EX1001 at 9:52-65, 10:20-50; *see also id.*, 6:24-35, 8:13-21.)

410					
Raw Play Stream					
row	log time	content ID			
1	t_0	song 51			
2	t_0+1	none			
3	t_0+2	song51			
4	t_0+3	song51			
5	t_0+4	movie 81			
6	t_0+5	song51			
7	t_0+6	song51			
8	t_0+7	song72			
9	t_0+8	song72			
10	t_0+9	none			
11	t_0+10	song72			
12	t_0+11	song72			
13	t_0+12	song72			
14	t_0+13	song65			
15	t_0+14	song65			
16	t_0+15	TVprgm58			
17	t_0+16	song65			
18	t_0+17	song65			
19	t_0+18	song65			
20	t_0+19	song65			
21	t_0+20	song21			
22	t_0+21	song21			
23	t_0+22	song21			
24	t_0+23	song21			
25	t_0+24	song21			
26	t_0+25	song21			
27	t_0+26	ad49			
28	t_0+27	ad49			
29	t_0+28	song81			
30	t_0+29	song35			

450					
Clean Play Stream					
row	log time	content ID			
1	t_0	song 51			
2	t_0+1	song51			
3	t_0+2	song51			
4	t_0+3	song51			
5	t_0+4	song51			
6	t_0+5	song51			
7	t_0+6	song51			
8	t_0+7	song72			
9	t_0+8	song72			
10	t_0+9	song72			
11	t_0+10	song72			
12	t_0+11	song72			
13	t_0+12	song72			
14	t_0+13	song65			
15	t_0+14	song65			
16	t_0+15	song65			
17	t_0+16	song65			
18	t_0+17	song65			
19	t_0+18	song65			
20	t_0+19	song65			
21	t_0+20	song21			
22	t_0+21	song21			
23	t_0+22	song21			
24	t_0+23	song21			
25	t_0+24	song21			
26	t_0+25	song21			
27	t_0+26	ad49			
28	t_0+27	ad49			
29	t_0+28	song35			
30	t_0+29	song35			

ssd = "sample sequence data"

(EX1001 at FIG. 4.)

31. Throughout the '848 patent, the disclosures describe only embodiments related to recognizing, analyzing, or processing samples of audio data and do not describe any features that specifically rely on using video data samples (to identify content). (*See, e.g.*, EX1001 at 1:37-39, 2:17-22, 3:13-24, 4:23-33, 6:13-17, 7:14-

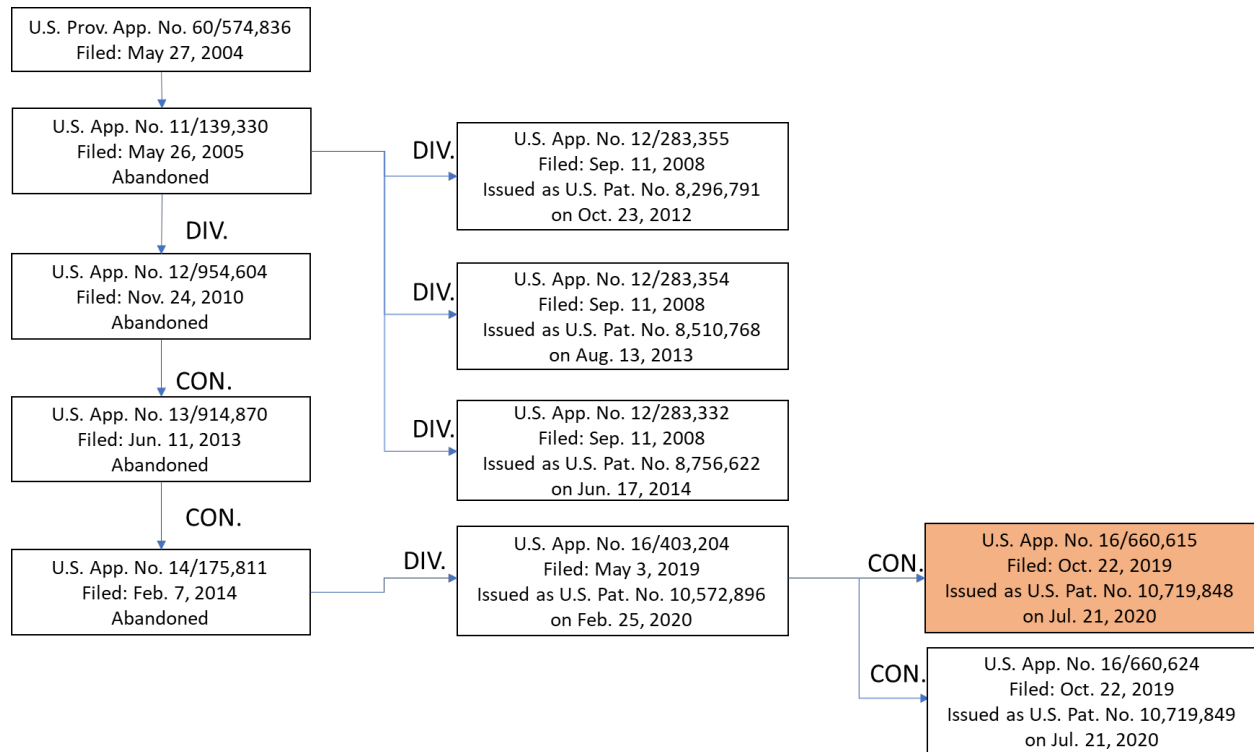
21, 11:47-52.) Even where the '848 patent discusses video-related content (e.g., television, TV, movie, etc.), it does so only in the context of recognizing, analyzing, or processing the audio data portion of the captured video (with audio) content. (*See, e.g.,* EX1001 at 1:56-62, 4:23-33, 7:14-21, 11:47-52.)

32. The challenged claims, on the other hand, include claim limitations directed to recognizing, analyzing, or processing “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.” (EX1001 at claim 1).¹ However, as I have discussed above and further discuss below, the '848 patent and its priority applications only describe embodiments involving recognizing, analyzing, or processing audio data samples. (*See* my discussions and explanations below in Section VIII; *see also generally* EX1001 at 2:61-14:16 (Detailed Description section discussing aspects related to audio data).)

33. Priority Applications of the '848 patent: I understand that the '848 patent issued from U.S. Patent Application No. 16/660,615 (“the '615 application”) (EX1004), which was filed on October 22, 2019. (EX1001 at Cover.) I also

¹ I sometimes add emphases to text of a document that I quote (as I do here) in this declaration. All such emphasis is added unless I state or explain otherwise.

understand that the '615 application is a continuation of U.S. patent application no. 16/403,204 ("the '204 application"). I understand that the '204 application is a divisional of U.S. patent application no. 14/175,811 ("the '811 application"). I also understand that the '811 application is a continuation of U.S. patent application no. 13/914,870 ("the '870 application"). I understand that the '870 application is a continuation of U.S. patent application no. 12/954,604 ("the '604 application"). I also understand that the '604 application is a divisional of U.S. patent application no. 11/139,330 ("the '330 application"). I understand that the '330 application is a non-provisional application claiming priority to U.S. provisional application no. 60/574,836 ("the '836 provisional application"). The chart in Demonstrative A below helps visualize the purported priority chain for the '848 patent.



Demonstrative A²

² The chart in Demonstrative A also shows additional applications that I understand are related to the '848 patent. For example, the chart shows another application (U.S. patent application no. 16/664,624) that I understand claims priority to the '204 application, which issued as the U.S. Patent No. 10,572,896—I understand that this application issued as U.S. Patent No. 10,719,849. The chart also shows three divisional applications (U.S. patent application nos. 12/283,355, 12/283,354, and 12/283,332) that I understand claim priority to the '330 application—I understand that these divisional applications issued as U.S. Patent Nos. 8,296,791, 8,510,768, and 8,756,622, respectively.

34. Based on my review, the priority applications of the '848 patent do not contain disclosures adequately describing recognizing, analyzing, or processing video data samples of the video portion of the video content, in the ways set forth in the challenged claims. (*See* my discussion below in Section VIII.) Thus, as I discussed above, I have considered the time of the alleged invention for the '848 patent to be around the mid-2019 timeframe, including up to May 3, 2019, the filing date of the '204 application, which added new matter not contained in any of the prior applications. The '615 application, which issued as the '848 patent, is a continuation of the '204 application. (*See* my discussion above in Section IV and my discussion below in Section VIII.)

VI. CLAIM CONSTRUCTION

35. I have given the terms in the challenged claims their plain and ordinary meanings, as would have been understood by a person of ordinary skill in the art at the time of the alleged invention, taking into consideration the language of the claims, the specification, and the prosecution history of the '848 patent. I have applied these understandings of the claim terms in my analysis and in forming my opinions where appropriate in this Declaration.

VII. OVERVIEW OF THE PRIOR ART³

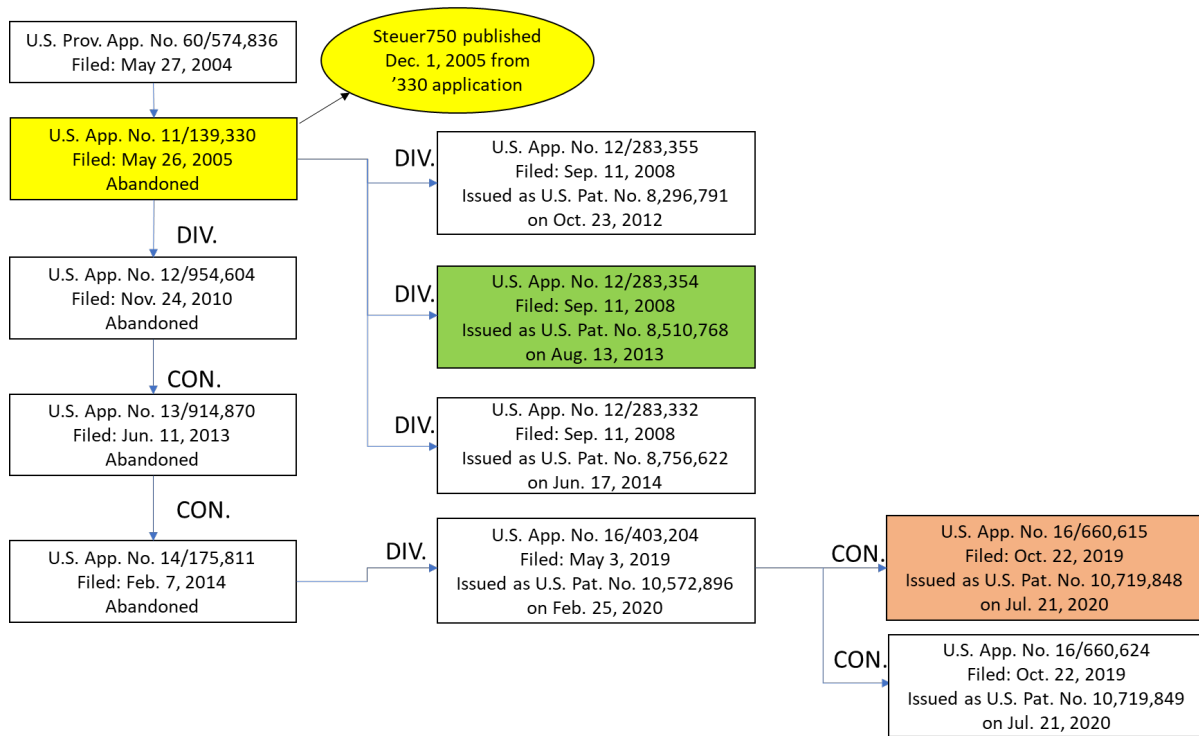
A. Overview of Steuer750 (EX1005) and Steuer768 (EX1006)

36. I understand that Exhibit 1005 is U.S. Publication No. 2005/0267750 to Steuer *et al.*, which I refer to as “Steuer750.” (EX1005 at Cover.) I understand that Steuer750 is a publication of the ’330 application, to which the ’848 patent claims priority. (*See* my discussion above in Section V.)

37. I understand that Exhibit 1006 is U.S. Patent No. 8,510,768 to Steuer *et al.*, which I refer to as “Steuer768.” (EX1006 at Cover.) I understand that Steuer768 is a patent that is part of the same family as the ’848 patent family and Steuer750, as it claims priority to the ’330 application. (*See* the demonstrative above in Section V.)

38. Demonstrative B below, which is an edited version of Demonstrative A in Section V above, shows a chart of the ’848 patent’s priority applications and related applications where the Steuer750 publication and its related ’330 application are highlighted in yellow and the Steuer768 patent is highlighted in green:

³ These summaries of the prior art in Section VII are applicable to my analysis below in Section IX. Accordingly, I incorporate these summaries, which reflect how a person of ordinary skill in the art would have understood the references, into my analysis below where appropriate (*see, e.g.*, Section IX below).



Demonstrative B

39. Thus, like the '848 patent, Steuer750 and Steuer768 are directed to measurement systems and methods for media usage monitoring involving recognizing content being consumed by a user and performing additional actions with the recognized content, such as creating a clean play list for the user, identifying a source of the content, and deducing play-altering activities a user engages in (e.g., fast-forwarding, rewinding). (EX1005 at Title, Abstract, [0006]; EX1006 at Title, Abstract, 1:60-2:15.) In fact, other than (1) the addition to the '204 application (the '615 application, which issued as the '848 patent, being a continuation of the '204 application) of the single paragraph purported to disclose video features (*see* my discussion and explanation below in Section VIII), and (2) differences in the

respective priority claims in the “Cross Reference to Related Applications” sections, it is my opinion that the specifications of the ’848 patent, Steuer750, and Steuer768 are substantively identical. (EX1001 at 1:6-16, 14:17-29; EX1005 at [0001]; EX1006 at 1:7-10; *see also* EX1008.)

40. In addition, based on my review, claims 1-15 of Steuer768 are nearly identical to claims 1-15 of the ’848 patent, except the word “audio” in claims 1-15 of Steuer768 is substituted with “video” in claims 1-15 of the ’848 patent. (*Compare* EX1001 at claims 1-15 *with* EX1006 at claims 1-15; *see also* EX1009 (showing a comparison of the claims).)

B. Overview of Neumeier (EX1007)

41. Like Steuer750 and Steuer768, Neumeier is directed to systems and methods for recognizing content being consumed by a user and performing additional analysis and processing in order to extract and report information on user actions and metrics based on the recognized content. (EX1007 at Abstract.) For example, Neumeier discloses identifying a channel the content was presented on and detecting viewer actions. (EX1007 at [0045]-[0047], [0073].) Neumeier additionally discloses that its methods of recognizing content can be performed using a sequence of samples of either audio or video data samples. (EX1007 at [0002], [0011]-[0014], [0025].)

42. Neumeier explains that its method of identifying media content involves “extracting audio or image data associated with [content] and then finding similar audio or image data in a content database.” (EX1007 at [0014].) As illustrated by Figure 10, television system 52 extracts patches of pixel data from each “frame,” the patches are then reduced into pixel samples that are a function of the data in the patches, and the “pixel samples for a succession of ‘frames’ (e.g., 10) are accumulated and then sent to the server.” (EX1007 at [0054], [0062], [0064]; *see also* [0068].) Channel recognition module 62 at the server receives “regular update[s] with pixel cue information,” which comprise “samples of pixel values composed of RGB combinations.” (EX1007 at [0054], [0058], [0068]-[0069].) Channel recognition module 62 then “searches the database 66 for those pixel cue points stored in the database which are nearest to the pixel cue points received” in order to identify a match for the content based on “a probability indicating the likelihood that it matches the video segment being displayed on the television screen.” (EX1007 at [0070]-[0072].) (*See also* EX1007 at [0095]-[0100].)

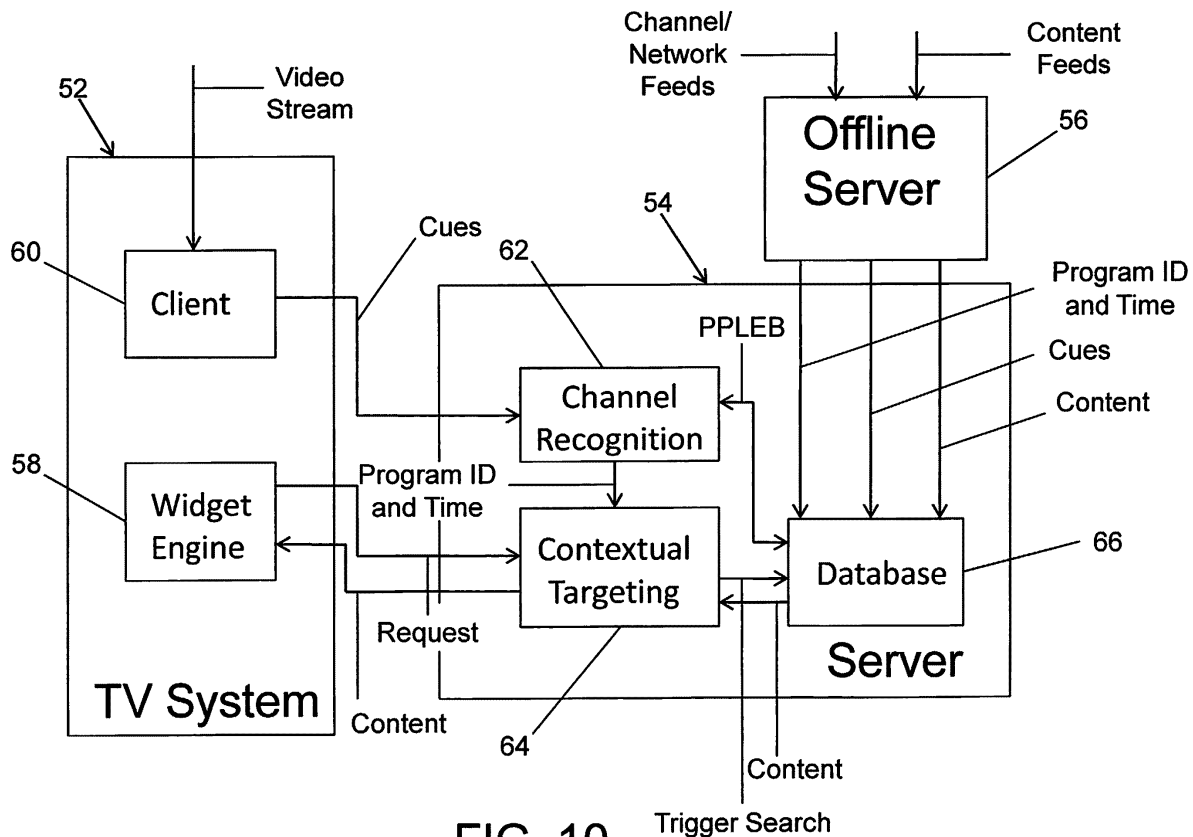


FIG. 10

(EX1007 at FIG. 10.)

43. Neumeier further discloses making determinations based on the “playing time” or “offset time” of the content consumed by the user, which “refer to a time which is offset from a fixed point in time, such as the starting time of a particular television program or commercial” and are “respectively associated with the aforementioned pixel cue points, thereby indicating the sequence in time when those pixel cue points occur.” (EX1007 at [0012], [0078]; *see also* EX1007 at [0014], [0024], [0032], [0070], [0072], [0076], [0087].) The probabilistic matching

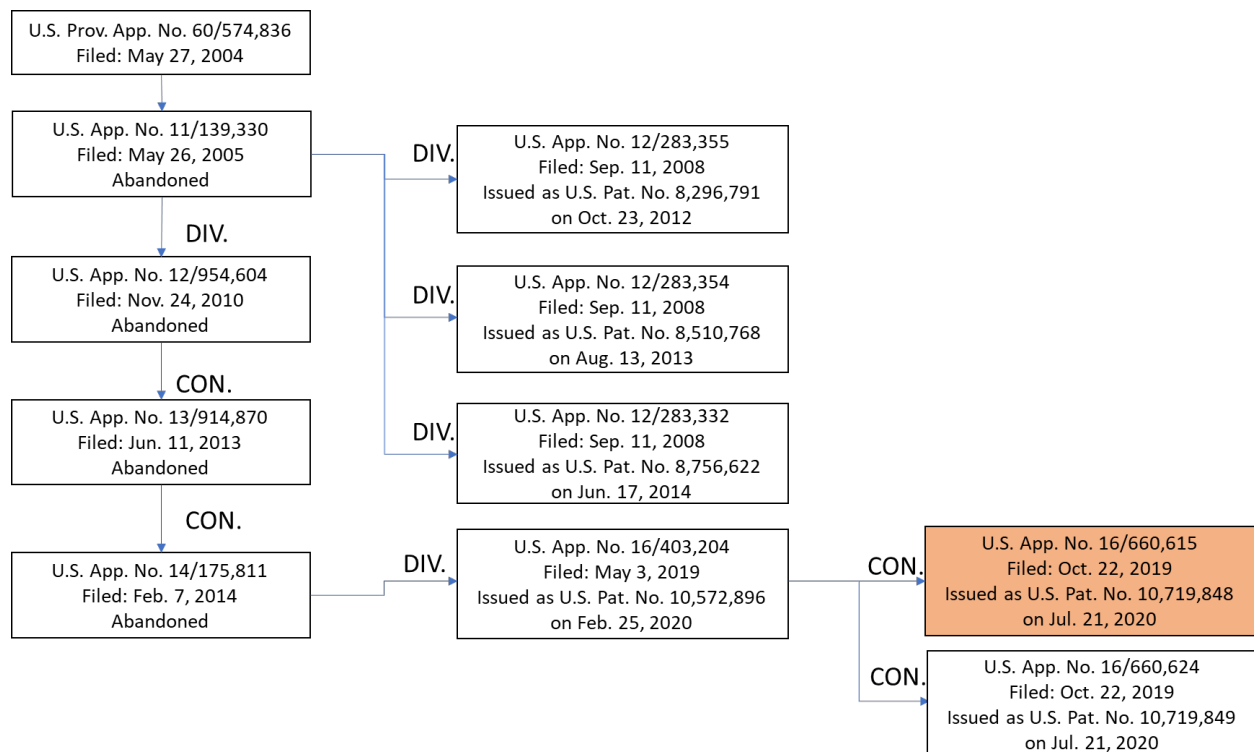
algorithm disclosed by Neumeier considers this offset time information and is able to respond to user activity, such as “viewer pausing, changing channels, etc.” (EX1007 at [0073].)

44. In addition, Neumeier discloses that using the probabilistic approach of the “path pursuit algorithm,” “an exact matching of pixel cue points at all times is not needed.” (EX1007 at [0073].) Instead, Neumeier discloses that even when certain individual frames of a video segment are not recognized, the segment may still be successfully identified by using “aggregated evidence” in order to “handle intermittent pixel cue points in a sequence that deviate from the other pixel data points in that sequence.” (EX1007 at [0073].)

VIII. DISCLOSURE OF CERTAIN CLAIM LIMITATIONS IN THE '848 PATENT CLAIMS AND ITS PARENT APPLICATIONS

45. In my opinion, a person of skill in the art would not have understood that the named inventors of the '848 patent possessed an invention with all of the features recited in the challenged claims (namely, claims 1-15) at the time of the alleged invention.

46. I understand that the '848 patent claims priority to several patents and/or patent applications, as shown in the alleged priority chain below in Demonstrative A (which I also discussed above in Section V):



Demonstrative A

47. However, as I discuss below, it is my opinion that the priority applications, namely U.S. Non-Provisional Application Nos. 11/139,330 (“the ’330 application”), 12/954,604 (“the ’604 application”), 13/914,870 (“the ’870 application”), 14/175,811 (“the ’811 application”), and 16/403,204 (“the ’204 application”) (collectively, “the priority non-provisional applications”) and U.S. Provisional Application No. 60/574,836 (“the ’836 provisional application”), do not convey to a person of ordinary skill in the art that the named inventors were in possession of all the features of the challenged claims at the time these priority applications were filed.

A. The Priority Applications Do Not Support The Video Data Features of The Challenged Claims

48. At least for the reasons I discuss below, it is my opinion that none of the priority applications support all the claim features of the challenged claims.

1. The Priority Applications Do Not Provide Support for the Claimed “Video Data Samples” and Related Video Features

49. Independent claim 1 of the ’848 patent recites:

A media measurement method comprising:

(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;

(b) using the one or more computers to query an electronic database of *a plurality of video data representations* and corresponding content identifiers;

(c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of *video data samples*;

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*;

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

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;*

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

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

(EX1001 at 15:13-59.)

50. Thus, independent claim 1 recites features relating to, amongst other things, matching captured video data samples comprising representations of video data with video data representations in an electronic database to obtain content identification results. Specifically, the claim first recites 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*. The claim recites generating a raw play stream, the raw play stream including a sequence of content identification results corresponding to the sequence of video data samples. The sequence of content identification results is obtained by querying an electronic database (of *a plurality of video data representations* and corresponding content identifiers) 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*.

51. The content identification result of the sequence of content identification results includes 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*. 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 claim then recites scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data. Finally, a clean play stream is generated from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.

52. Independent claim 9 recites similar features. (EX1001 at 16:20-67.)

53. As I discuss below, it is my opinion that the priority applications (the '836 provisional application and the priority non-provisional applications) do not provide written description support for the above video data-related features.

a) The '836 Provisional Application Does Not Support the Claimed Video Data Features

54. As shown above in Demonstrative A, the '848 patent claims priority through each of the priority non-provisional applications to U.S. Provisional Application 60/574,836 (“the '836 provisional application”). However, the '836 provisional application does not convey to a person of ordinary skill in the art that the named inventors had possession of the video data features at the time the '836 provisional application was filed (May 27, 2004). This is at least because it does not describe any specific features or embodiments that rely on video data samples comprising representations of video data (generally referred to as “video data samples” below), let alone any disclosure corresponding to the claimed subject matter.

55. The '836 provisional application discloses “measurement systems” that purport to involve recording and reporting on individuals’ media usage habits. According to the disclosures of the '836 provisional application, this is accomplished by using monitoring or detection devices to collect and recognize media broadcast by a playback device. (EX1010 at 2-3.) But, as explained further below, the disclosures of the '836 provisional application fail to adequately describe embodiments that rely on the use of video data, let alone in the manner claimed (as

I discussed above in Section VIII.A.1), and therefore the '836 provisional application does not support the claims of the '848 patent.

56. The '836 provisional application does not disclose any features relating to collecting and recognizing the visual aspects of video content. Instead, throughout the '836 provisional application, its disclosures focus on identifying media content by collecting and recognizing *audio* data associated with the content. For example, the '836 provisional application consistently discusses collecting specifically *audio* media content and using “*audio* fingerprinting” technology in order to recognize the content:

[A]n exemplary measurement system employs *audio fingerprinting and recognition technology* in order to permit the recognition of media content itself. Such an *audio fingerprinting and recognition technology* may, for example, include one or more previously developed technologies of the kind commercialized by Philips Research, Shazam Entertainment Ltd., and MusiKube LLC, among others, in the context of *song identification*.

Id. at 2 (emphasis added).

[A]n exemplary method creates a record of individual media usage over time by means of an open-ended measurement system that employs portable and stationary detection devices to *generate digital fingerprints of the audio signal portion of the media content* itself . . .

Id. at 3 (emphasis added).

[A]n exemplary method includes ***analyzing digital fingerprints of audio energy collected from portable and stationary audio detection devices . . .***

Id. at 3 (emphasis added).

[P]articular meta-data may be associated with an ***audio fingerprint . . .***

Id. at 4 (emphasis added).

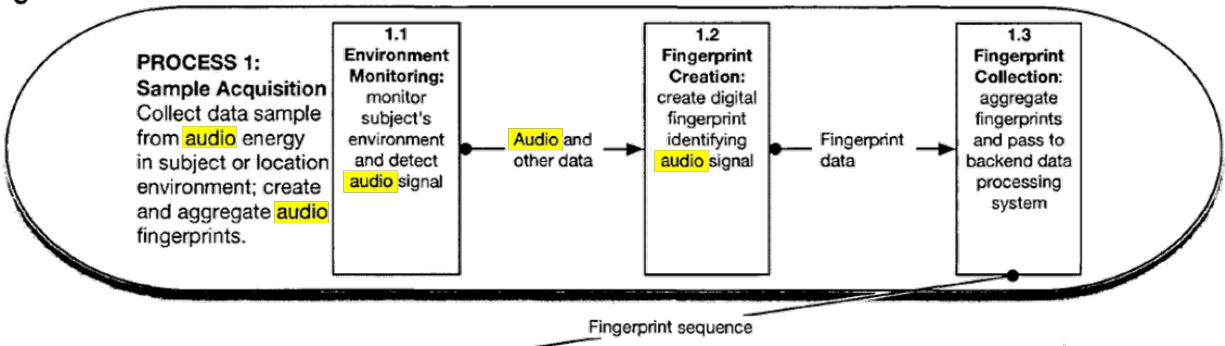
An electronic device is either impermanently affixed to the individual whose media usage is to be monitored, or placed in a stationary location to be monitored. ***This device samples audio energy*** either from the air by means of a microphone or directly from one or more sources of media content to be monitored by means of a direct electrical connection to such device or devices.

Id. at 5 (emphasis added).

Every few seconds, ***the sampler takes an audio snapshot*** of all detectable media activity within its range, ***extracts features of the audio snapshot, and creates a digital fingerprint*** identifying the signals present.

Id. at 5 (emphasis added).

Figure 3: Process Flow



Id. at 9 (FIG. 3) (annotated). Similarly, the '836 provisional application asserts that once the audio fingerprints have been generated, the media content is recognized, or matched, using those audio fingerprints:

[A]n exemplary method . . . *generate[s] digital fingerprints of the audio signal portion* of the media content itself and identifies media played in the environment by means of *matching such fingerprints* against a database of known media.

Id. at 3 (emphasis added).

[A] selected sample of *audio fingerprints* from each source *may be matched* . . .

Id. at 4 (emphasis added).

[A]dditional *audio fingerprints* from a source may be *selected for matching* . . .

Id. at 4 (emphasis added). The '836 provisional application also discloses methods “to enhance the speed and accuracy of matches of *audio fingerprints*.” *Id.* at 3 (emphasis added).

[M]ultiple *audio fingerprints* of differing depth, size, and complexity may be created during the monitoring process in order to enhance the speed and reduce the data processing power required *to match audio fingerprints . . .*

Id. at 4 (emphasis added).

[A]udio fingerprints representing multiple time windows may be created during the monitoring process in order to enhance the speed and reduce the data processing power required *to match audio fingerprints* . . .

Id. at 4 (emphasis added). Other purported methods of enhancing matching include “collaborative filtering technologies,” “comparison to known and prearranged sequences of media,” and using “time sequence data” to “correlat[e] successive fingerprint data samples.” *Id.* at 3. But in my opinion, a person of ordinary skill in the art would have understood that all of the various disclosed methods, techniques, and/or technologies rely on using audio fingerprints and the ’836 provisional application provides no equivalent or corresponding detailed disclosures that instead rely on using video data samples from the visual portion of monitored media content, identifying content based on analysis of such video data samples, and processing subsequent data processing steps based on the content identification (e.g., deducing play-altering activities by the user, identifying a channel, or scrubbing a playlist).

57. Even where the ’836 provisional application mentions video content, a person of ordinary skill in the art would have understood that those features are limited to “content including at least some audio elements.” (EX1010 at 2.) Indeed, throughout the ’836 provisional application, it focuses on audio data or only the audio portion of video content:

[A]n exemplary measurement system may enable the monitoring of individuals' media usage across multiple sources of *audio and/or audiovisual content* . . .

Id. at 2 (emphasis added).

[A] measurement system is designed to measure individual usage of *audio and audiovisual media* over long periods of time . . .

Id. at 3 (emphasis added).

Another exemplary method determines the source of transmission or storage of the media played by an individual based on *audio characteristics of that signal* (e.g. recognizes particular types of encoding or compression, or *hears screen refresh on [TV] or computer*, or recognizes *compression or frequency range of* FM or AM radio, or CD or *DVD*).

Id. at 3 (emphasis added).

A system is disclosed for measuring individuals' usage of *audio and audiovisual media* in order to understand how media content is used across a range of devices and transmission methods. The system uses portable and stationary devices to *detect audio signals* in the local environment and/or from the media devices to be measured. The system matches *digital fingerprints of the audio signal present in audio and audiovisual media* to the corresponding fingerprints of known media content stored in a database.

Id. at 6. Thus, in my opinion, the monitored media potentially including visual elements does not change my understanding of each of the exemplary methods described in the '836 provisional application because only the audio portion of video content is analyzed and processed.

58. The '836 provisional application does include a brief, vague, and conclusory portion that mentions video data, but it fails to provide adequate written description support for the claimed video data features of the claims of the '848 patent. (EX1010 at 4.) This generic disclosure states:

Another exemplary method creates a record of individual media usage over time by means of an open-ended measurement system that employs portable and stationary detection devices to generate digital fingerprints of the video signal portion of the media content itself and identifies media played in the environment by means of matching such fingerprints against a database of known media. Such a video-based system may be built upon direct electrical connections and optical sensors (rather than microphones).

These various exemplary methods may involve video rather than audio.

(EX1010 at 4.) In my opinion, this disclosure does not convey to a person of ordinary skill in the art that the named inventors had possession of the video data features at the time the '836 provisional application was filed, for the reasons explained below.

59. First, where the '836 provisional application states that “[t]hese various exemplary methods may involve video rather than audio,” it is unclear what methods the phrase “[t]hese various exemplary methods” refers to. For example, one disclosed exemplary method relates to “detect[ing] the simultaneous playback of more than one media source in a measured environment,” and illustrates such a scenario with an example where content is recognized when a “TV and radio are on at the same time,” or “if music is the background of an advertisement.” (EX1010 at

3.) It is unclear how these examples, which refer to content that only includes an audio portion—content played by radio and music content—would have instead “involve[d] video rather than audio.”

60. Further, the conclusory disclosure related to video appears only around halfway into '836 provisional application, at the end of a section with the heading “Detailed Specification,” and after which are additional disclosures. (EX1010 at 4.) It is therefore further unclear whether the phrase “[t]hese various exemplary methods” refers to the purported methods described prior to this disclosure, in the “Detailed Specification” section, to the later disclosed methods, or something else. Moreover, regardless of the examples to which the conclusory disclosure is directed, the '836 provisional application provides no details whatsoever to a person of ordinary skill in the art as to what changes would be made in order for the “various exemplary methods” to “involve video rather than audio.” (EX1010 at 4.)

61. I have been asked to assume an argument by Patent Owner that there are no substantive differences between audio and video in the context of substituting video data samples for audio data samples in the context of the '836 provisional application. But in my opinion, the '836 provisional application does not support such a conclusion. To the contrary, a person of ordinary skill in the art would have recognized the different challenges and obstacles faced with the collection and recognition of audio versus video (i.e., using audio data samples for content

recognition as opposed to using video data samples for content recognition, and for performing the subsequent claim steps like 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).

62. For example, a research competition funded by government organizations at the time, National Institute of Standards and Technology (NIST) and the Advanced Research and Development Activity (ARDA), demonstrates the early stages and challenges of video-data-based content recognition during the 2004-2005 timeframe. (EX1030 at 1-2 (Introduction); EX1031 at 1.) The 2004 research competition, Text REtrieval Conference (TREC) Video Retrieval Evaluation (VID), collectively TRECVID 2004 included participants from worldwide universities and research organizations. (EX1030 at 2 (Table 1).) TRECVID 2004 involved various tasks involving video content including shot boundary determination, story segmentation, feature-extraction, and search. (EX1030 at 1; *see also id.* at 3 (Section 2.1 describing the video content), EX1031 at 1 (Section 2), 5 (Section 3.2 describing the video test data).) The Guidelines for the TRECVID 2004 Evaluation (EX1031) set forth ten features for recognition including “Boat/ship,” “Madeline Albright,” “Bill Clinton,” “Train,” “Beach,” “Basket scored,” “Airplane takeoff,” “People walking/running,” “Physical Violence,” and “Road.” (EX1031 at 2-3 (Section 2.3); EX1030 at 10 (“The features to be detected were defined (briefly) as follows . . .”).)

As shown in Table 2 of the TRECVID 2004 overview (EX1030), the search accuracy for identifying this content was less than 10% for eight out of ten features, and about 30% or less for the remaining two features. (EX1030 at 10 (Table 2, last column showing accuracy rate for the various features).) Indeed, another TRECVID presentation from 2005 shows that the average search accuracy between 2003-2005 for identifying content was less than 10%. (EX1032, 9.) The TRECVID documents confirm my understanding that in the 2004-2005 timeframe, a person of ordinary skill in the art would have understood that substituting audio data samples with video data samples for content recognition would not have been straightforward.

63. The '836 provisional application is silent on such issues and how the disclosed audio-based embodiments would have been implemented in the context of video in light of any unique challenges posed by video data.

64. Without any disclosures similarly describing ways of dealing with difficulties unique to the video context or explaining how the disclosures for dealing with difficulties in the audio context would have extended to the video context or could be adapted to the video context, it is my opinion that a person of ordinary skill in the art would not have understood the named inventors of the '848 patent to have had possession of the claimed subject matter at the time the '836 provisional application was filed.

65. A detailed review of the claims confirms my understanding that the '836 provisional application does not provide adequate support for the claims. Claim 1 of the '848 patent recites “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.” (EX1001 at 15:15-19.) A “raw play stream comprising a sequence of content identification results corresponding to the sequence of video data samples” is generated, where “the sequence of content identification results is obtained by querying [an] 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.*” (EX1001 at 15:20-33.) Thus, the “content identification results” of claim 1 refer to, e.g., using video data samples to recognize, or match, captured media content by searching a database. But, as I explained above, the '836 provisional application does not disclose using data samples from the visual aspects of video content (e.g., collecting video data and/or generating video data samples from video content that the user is viewing), nor using such video data samples to recognize media content. Thus, a person of ordinary skill in the art would not have understood that the '836 provisional application discloses this limitation of claim 1 of the '848 patent.

66. Claim 1 of the '848 patent also recites, for example, “scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data.” (EX1001 at 15:51-56.) As I noted earlier, the claimed “content of identification results” are obtained using captured video data samples as they include representations of video data and *video data representations in the electronic database*. On the other hand, the '836 provisional application includes a section entitled “MATCHING AND SCRUBBING” (EX1010 at 5) that is entirely premised on “Fingerprint Matching” (EX1010 at 5) that is based on sampling of *audio data*, not video data. In particular, immediately prior to the “MATCHING AND SCRUBBING” section, the '836 provisional application includes sections on “Fingerprint Creation” and “Fingerprint Collection” that describe the fingerprinting process as follows: “Every few seconds, the sampler takes *an audio snapshot* of all detectable media activity within its range, extracts features of *the audio snapshot*, and creates a digital fingerprint identifying the signals present.” (EX1010 at 5.) Thus, to the extent the “MATCHING AND SCRUBBING” section in the '836 provisional application even relates to the claimed scrubbing features, it relates to the use of audio characteristics of a signal and not video data samples. Accordingly, a person of ordinary skill in the art would

not have understood that the '836 provisional application discloses this limitation of claim 1 of the '848 patent.

67. Claims 2-8 each depend from claim 1. (EX1001 at 15:60-16:19.) These claims therefore lack support for at least the same reasons as claim 1.

68. Independent claim 9 of the '848 patent recites features analogous to those recited and discussed above with respect to claim 1, but in the context of “[a] computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers.” (*Compare* EX1001 at 15:13-59 (claim 1) *with id.*, 16:20-67 (claim 9).) Thus, like independent claim 1, independent claim 9 recite features relating to the use of video data samples comprising representations of video data.

69. Accordingly, for the same reasons I articulated above for the limitations of claim 1, the '836 provisional application fails to provide adequate support for independent claim 9. Claims 10-15 depend from claim 9. (EX1001 at 17:1-20.) Therefore, claims 10-15 lack support for at least the same reasons given for claim 1.

b) The '330, '604, '870, and '811 Priority Non-Provisional Applications Do Not Provide Adequate Support For The Challenged Claims

70. The priority non-provisional applications at issue here, *i.e.*, the '330 application, '604 application, '870 application, and '811 application, share substantially the same as-filed specifications. (EX1011 at 577-599 ('330 application

specification), 610-617 ('330 application abstract and drawings), EX1012 at 276-298 ('604 application specification), 311-318 ('604 application abstract and drawings), EX1013 at 21-43 ('870 application specification), 56-63 ('870 application abstract and drawings), EX1014 at 420-442 ('811 application specification), 455 ('811 application abstract), 460-466 ('811 application drawings).) Exhibits 1017, 1018, and 1019 are computer-generated redlines comparing the as-filed specifications of the '604 application, '870 application, and '811 application to the as-filed specification of the '330 application. As shown in exhibits 1017-1019, other than the changes to the "Cross Reference to Related Applications" sections in each of the priority applications, the specifications of these priority non-provisional applications are substantively identical.⁴ Thus, while my opinions below primarily focus on the lack of support for the challenged '848 patent claims in the '330 application's disclosure, it is my opinion that a person of ordinary skill in the art would have understood that the disclosures of the '604 application, '870 application, and '811 application are deficient for the same reasons.

⁴ Because the specifications of the applications are substantively identical, I use the specification of the '330 application as representative of the specifications of the priority applications. Thus, my citations below refer only to the specification of the '330 application.

71. In particular, the disclosures of these priority non-provisional applications do not describe any features that rely on video data samples including representations of video data (e.g., generating a raw play stream, the raw play stream including a sequence of content identification results corresponding to the sequence of video data samples) and consequently, the subsequent related features set forth in the claims. Instead, as I demonstrate below, the disclosures describe only embodiments based on **audio** data samples.

72. The '330 application uses the term “audio” 37 times, including, for instance, in the specification (33 times), abstract (2 times), and figures (2 times). (*See generally* EX1011 at 577-599, 610-617.) It also is replete with terms that are related only to audio (to the exclusion of video): song (25 times, excluding figures), radio (11 times, excluding figures), microphone (10 times, excluding figures), headphones (10 times, excluding figures), album (5 times, excluding figures), sound (2 times), music (2 times). By contrast, the specification of the '330 application uses the term “video” only 4 times. (*See generally* EX1011 at 577-599, 610-617.) The particular claim term “video data sample” appears only once outside of the claims in the '330 application specification:

At the same time, the media measurement context provides opportunities to utilize data exogenous to a particular audio or *video data sample*. Such opportunities have thus far been insufficiently exploited for the purpose of efficiently applying existing content recognition technologies in the media measurement context. Thus, an improved media measurement system and method is needed.

(EX1011 at 578 ([0005]).) Even in this instance, it is unclear if the term “video data sample” is used in the context of identifying content by sampling video data or if the term, as used in the context of the sentence (“the media measurement context provides opportunities to utilize data exogenous to *a particular audio or video data sample*”), is used to reference video content which includes an audio component, and the audio component is sampled for content recognition (since audio-sampling-based content recognition is the only type of content recognition described in the ’330 application). Indeed, the ’330 application does not describe any content recognition using a “video data sample” (e.g., by sampling pixels in a frame as described by Neumeier which I discussed in Section VII.B above) so in my opinion, a person of ordinary skill in the art would not have understood this portion of the ’330 application to sufficiently describe such an embodiment.

73. And while the ’330 application specification uses certain video-related terms (e.g., television, TV, movie, etc.), in my opinion, it is clear in view of the

specification as a whole that these disclosures refer to analyzing and processing the audio data portion of captured video content. In particular, as shown in the reproduced portions of the '330 application specification below, while this portion mentions media played by a DVD player or television (which may include video content) and mentions media content such as “movie, commercial, television show, video game etc.”), the specification only describes capturing the audio sample from this media content:

When the audience member is consuming media from media player 170 *a* without the use of headphones 160 *a* (i.e. if media player 170 *a* has a speaker), monitor 101 captures audio energy acoustically through microphone 105. However, *when the audience member desires to use headphones 160 a to receive media content from media player 170 a* (e.g. an MP3 player, iPod, CD player, **DVD player, television**, radio, computer, or other media player) he or she can plug headphones 160 *a* into headphone port 103 and plug media player 170 *a* into media player port 102. In that case, *monitor 101 a captures audio energy through media player port 102.*

(EX1011 at 580-581 ([0017]).)

As another example, **headphones** 160b are adapted to receive a content signal 171 via wireless transmission 171

from media player 170b. Monitor 101 b is adapted to receive the same content signal 171 from media player 170b. In some embodiments, a syncing process allows monitor IO I b to adapt along with headphones 160b and media player 170b to changes in signal 171 **so that monitor 101b continues to receive the same signal as headphones 160b.** In some embodiments, syncing may be accomplished via a wireless and/or automatic process, such as providing a monitor that responsively changes tuning based on tuning changes between the media player and the **headphones.**

(EX1011 at 581 ([0018]).)

Audio data captured by monitors 101, along with associated monitor data (e.g. device ID) (collectively, “audio data/monitor data” 144) may be uploaded to system 100 through upload port 104, directly, via memory device transfer (e.g. flash memory card, floppy disk, CD, etc.) or through a network such as network 150. **The audio data uploaded from a monitor 101 may be raw audio data, or may be audio data that has undergone one of a variety of levels of processing prior to upload.** For example, 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.** Alternatively, the audio data may comprise pre-calculated

signatures, landmarks, or other data useful for content recognition. On the other hand, in some embodiments, **the uploaded audio data may simply represent a raw signal corresponding to audio energy received at monitor 101.**

(EX1011 at 581 ([0019]).)

Content 141 may include media files associated with a known piece of media content (e.g. a song, *movie*, *commercial*, *television show*, *video game*, etc.). However, it will be understood by those skilled in the art that “content” in the sense of the known content 141 received by system 100 may, in some embodiments, simply refer to data about signals representing the content and does not necessarily refer to a stored version of the media content itself; e.g. “content” may be stored signatures and landmarks *derivable from a song’s audio*, but not necessarily an audio file useable to play the song itself.

(EX1011 at 582-583 ([0022]); *see also id.* at 583 ([0023]) (describing module 121 “divid[ing] **audio data** into samples and determin[ing] which samples should be submitted for further content recognition testing”), 586-587 ([0029]) (describing processing of audio data for content identification).)

At step 210, *characteristics of the audio data are analyzed* to attempt to identify the source of transmission or storage of the media played by an individual based

on audio characteristics of that signal. For example, the process attempts to recognize particular types of encoding or compression, or *identifies sound associated with screen refresh on a television or computer monitor*, or recognizes compression or frequency range of FM or AM radio, or CD or DVD.

(EX1011 at 587 ([0030]); *see also id.* at 587 ([0031]) (describing content identification based on “**audio data** in the tested sample”), 587-588 [0032] (describing adjusting sample control and selection parameters for content identification including increasing resolution “when recent identification results suggest an **extended period of silence or absence of known content**”).)

Play List 560 contains a row for each internally continuous media consumption event. The start log time and end log time for each such event are listed in each row. *If an entire piece of content (e.g. a TV program segment, or an entire TV ad) is listened to without interruption, that would correspond to a row of Play List 560.*

(EX1011 at 594 ([0049]).) Thus, in my opinion, a person of ordinary skill in the art would have understood that none of the embodiments in the '330 application describe, in any way, the video data sample features featured in the claims of the '848 patent.

74. Instead, in my opinion, a person of ordinary skill in the art would have understood that the embodiments disclosed in the '330 application (and the remaining priority non-provisional applications at issue) are solely focused on recognizing content based on audio data (e.g., the use of audio data samples to identify content) and subsequent processing of the content recognition results. (*See generally* EX1011 at 577-599, 610-17.) The '330 application does not disclose any embodiments related to, and thus does not provide any support for, the claimed features of receiving *a sequence of video data samples, the video data samples comprising representations of video data* and generating a raw play stream including a sequence of content identification results corresponding to the sequence of video data samples, let alone using such content identification results generated based on video data samples to carry out the remaining steps of the claim such as scrubbing the raw play stream.

75. Further, in my opinion, a person of ordinary skill in the art would not have understood that the '330 application's disclosures of embodiments relying on audio data would have been interchangeable with embodiments that instead use video data. In my opinion, a person of ordinary skill in the art would have understood that a system that relies on capturing audio would operate in different environments and be subject to different obstructions, challenges, and limitations to a system that relies on capturing video. For example, multiple disclosures in the '330 application

describe peculiarities related to audio content that have no clear analog in the video context.

[A]n audience member may *mute a broadcast* during commercial periods. Thus simply identifying a broadcast channel does not adequately track whether the audience member *listened* to a particular advertisement.

(EX1011 at 577 ([0002]).)

[O]ther embodiments provide a monitoring and measurement system in which a media monitoring device is adapted to *receive a wireless or non-wireless audio signal* from a media player, the audio signal also being received wirelessly by headphones of a user of the monitoring device.

(EX1011 at 578 ([0006]).)

When the audience member is consuming media from media player 170 *a* without the use of headphones 160 *a* (i.e. if media player 170 *a* has a speaker), *monitor 101 captures audio energy acoustically through microphone 105*. However, when the audience member desires to use headphones 160 *a* to receive media content from media player 170 *a* (e.g. an MP3 player, iPod, CD player, DVD player, television, radio, computer, or other media player) he or she can plug headphones 160 *a* into headphone

port 103 and plug media player 170 *a* into media player
port 102.

(EX1011 at 580 ([0017]).)

Other differences corresponding to *content types (e.g. speaking audio versus music audio)* might also be exploited based on prior content identification results to select a most efficient recognition algorithm from a plurality of algorithms.

(EX1011 at 584-585 ([0025]).)

76. The '330 application suffers from the same deficiency described above with respect to the '836 provisional application, namely, it provides no corresponding disclosures describing how difficulties unique to the video context are addressed. (*See* Section VIII.A.1.a, above.) Likewise, nowhere in the '330 application is an explanation for how its disclosures for dealing with difficulties in the audio context would have extended to or could be adapted to the video context. (*Id.*)

c) The Combination of The '836 Provisional Application's Disclosure with the Priority Non-Provisional Applications Does Not Support the Video Data Features of the Challenged Claims

77. I have been asked to consider whether a person of ordinary skill in the art would consider the combination of the disclosure of the '836 provisional application with the disclosure of the priority non-provisional applications (i.e., the

'330, '604, '870, and '811 applications I discussed above) as providing adequate support for the video-related claims.

78. In my opinion, a person of ordinary skill in the art would have understood that the disclosures of the priority non-provisional applications, when considered in combination with the disclosures of the '836 provisional application (including its vague video-data-related paragraph), do not provide support for the video data features and related features of the challenged claims of the '848 patent.

79. As an initial matter, I note that the '330, '604, '870, and '811 priority non-provisional applications do not reference the above-noted vague paragraph from the '836 provisional application concerning video, such that a person of ordinary skill in the art would have consulted that particular paragraph in reviewing and understanding the audio-based disclosures of the priority non-provisional applications.⁵ (*See generally* EX1011 at 577-599 ('330 application specification); EX1012 at 276-298 ('604 application specification); EX1013 at 21-43 ('870 application specification); EX1014 at 420-442 ('811 application specification).) Moreover, in my opinion, a person of ordinary skill in the art would not have understood that the disclosed audio-based embodiments in the priority non-

⁵ As I discussed above in Section VIII.A.1.a, there is no other disclosure in the '836 provisional application that describes using video data.

provisional applications could have been implemented in the video context during the relevant timeframe for all the reasons I explained above. (*See* Section VIII.A.1.a-b, above.) There is simply no guidance provided to a person of ordinary skill in the art in the disclosures of any of the priority non-provisional applications that would have conveyed to such a person that the named inventors were in possession of the claimed invention including the video data features during the timeframe when the priority non-provisional applications were filed (*e.g.*, in the 2005-2014 timeframe).⁶

80. Thus, in my opinion, a person of ordinary skill in the art would have understood that like the '836 provisional application, the priority non-provisional applications (alone or in combination with the '836 provisional application) fail to provide support for the video data features set forth in the claims of the '848 patent.

2. Additional Materials Mentioned in The Priority Non-Provisional Applications Do Not Support the Video Data Features of the Claims of the '848 Patent

81. The '330 application, and similarly the other priority non-provisional applications, purport to be “adapted to implement one or more known content

⁶ As shown in Demonstrative A in Section V and at the outset of this section (*i.e.*, Section VIII), the '330 application was filed in 2005, the '604 application was filed in 2010, the '870 application was filed in 2013, and the 811 application was filed in 2014.

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.” (See, e.g., EX1011 at 584-585 ([0025]).) The ’330 application then mentions several patent applications (which I understand are being submitted as Exhibits 1020-1023) that allegedly describe content recognition methods without specifying any particular portions of the documents as being relevant.

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

**“Improvements in and Relating to Fingerprint
Searching” (WO2004/040475 published May 13, 2004).**

(EX1011 at 584-585 ([0025]).)

82. To the extent the above-noted documents disclose features related to using video data samples for content recognition, none of the priority non-provisional applications make any mention of specific video recognition-related features in these documents or how their use may be incorporated into the disclosed embodiments of the priority non-provisional applications. Instead, the priority non-provisional applications merely generically state that the above-noted documents are examples of content recognition methods and algorithms. (EX1011 at 584-585 ([0025]).) In my opinion, a person of ordinary skill in the art would not have understood from this limited, non-specific disclosure that specific video-data-based content recognition features are what was intended to be considered in these documents, especially since, as I discussed above in Section VIII.A.1.b, the priority non-provisional applications are limited to audio-based content recognition embodiments.

83. In other words, in my opinion, the priority non-provisional applications simply do not identify with any specificity what portions of the above-noted documents (EXS1020-1023) relate to using video data samples and where such features can be found within the documents. In this regard, I note that one of the

references identified in the '330 application, namely U.S. Patent Publication No. 2002/0083060, relates to identification of audio and provides no disclosures related to identification of video content. (*See generally* EX1020.)

84. Indeed, none of the documents (EXS1020-1023) exclusively relate to content identification based on video data samples, so in my opinion, a person of ordinary skill in the art would have understood that only the features of the above-noted documents related to content recognition based on audio data were relevant to the '330 application (particularly because content identification based on audio data samples was the focus of the disclosures of the priority non-provisional applications). Thus, a person of ordinary skill in the art reviewing the documents mentioned in paragraph [0025] of the '330 application specification in light of the disclosure of the priority non-provisional applications would have had to speculate without any guidance as to what teachings should be combined, and how to accomplish such combinations, in order to adapt the disclosed embodiments to the use of video data.

3. New Disclosures Added to The '204 Application That Issued as The '896 Patent

85. In my opinion, the addition of new disclosures to the as-filed specification of the '204 application, which is another non-provisional application to which the '848 patent claims priority, in an apparent attempt to support the video-

related features of the claims further confirms my understanding regarding the inadequate disclosures in the earlier priority non-provisional applications, *i.e.*, the '330 application, '604 application, '870 application, and '811 application. In particular, I note that the '204 application added the following single generic paragraph regarding video:

In alternative embodiments to what is illustrated and described in FIGS. 1-6, a record is created of individual media usage over time by means of an open-ended measurement system that employs portable and stationary detection devices to generate digital fingerprints of the video signal portion of the media content itself and identifies media played in the environment by means of matching such fingerprints against a database of known media. Such a video-based system may be built upon direct electrical connections and optical sensors (rather than microphones). Thus, various alternatives to the exemplary methods discussed in the context of the audio data-based embodiments illustrated in FIGS. 1-6 rely instead on video data.

(EX1015 (file history of '204 application), 232 (¶[0061]).)

86. While the new paragraph above appears similar to the short paragraph regarding video in the '836 provisional application (EX1010, 4) I discussed above in Section VIII.A.1.a, the reference to the “**FIGS. 1-6**” embodiments and the last

sentence stating that the “various alternatives to the exemplary methods discussed in the context of the **audio data-based embodiments illustrated in FIGS. 1-6 rely instead on video data**” is altogether new. (*See also* EX1028.) Figures 1-6 and their corresponding detailed descriptions in the '204 application were not included in the '836 provisional application, and so in my opinion, extrapolating the application of video data to those detailed embodiments (to the extent a person of ordinary skill in the art would have understood it as such given the unclear language) is not supported by the '836 provisional application. (*See generally* EX1010, 2-9; EX1011, 611-616 (FIGS. 1-6).) The new paragraph above is also not included in any of the priority non-provisional applications. (*See generally* EX1011, 577-599 ('330 application specification); EX1012, 276-298 ('604 application specification); EX1013, 21-43 ('870 application specification); EX1014, 420-442 ('811 application specification).) Thus, in my opinion, there is no support in the '836 provisional application or the other priority non-provisional applications for video data being relied upon in the detailed embodiments of “FIGS. 1-6,” to the extent this new paragraph can be understood in the '204 application.

87. Thus, in my opinion, the '204 application contains new disclosures not present in any earlier priority application.⁷

⁷ To be clear, I take no position regarding whether the newly added disclosures to the specification of the '204 application are sufficient to provide support for the video-related claim features.

IX. THE PRIOR ART DISCLOSES AND/OR SUGGESTS ALL RECITED FEATURES OF CLAIMS 1-15 OF THE '848 PATENT

A. Grounds 1 and 2:

Ground 1: The Combination of Steuer750 and Neumeier Discloses and/or Suggests Claims 1-15

Ground 2: The Combination of Steuer768 and Neumeier Discloses and/or Suggests Claims 1-15

88. **Ground 1** relies on the combination of Steuer750 (EX1005) and Neumeier (EX1007) whereas **Ground 2** relies on the combination of Steuer768 (EX1006) and Neumeier. However, Steuer750 and Steuer768 have substantively identical specifications, as I discussed above in Section VII.A. As such, in the interest of efficiency and clarity, I discuss both grounds collectively, with citations provided to both Steuer750 and Steuer768. That is, I provide citations below to Steuer750 (to support **Ground 1**) and Steuer768 (to support **Ground 2**). In doing so, I am not relying on Steuer750 for **Ground 2** or Steuer768 for **Ground 1**.

89. As I noted above in Section VII.A, Steuer750 and Steuer768 have one difference with respect to their claims. In particular, Steuer768 further includes claims 1-15 which are nearly identical to challenged claims 1-15 here, except the word “audio” in claims 1-15 of Steuer768 is substituted with “video” in the challenged claims. (*Compare* EX1001 at claims 1-15 *with* EX1006 at claims 1-15; *see also* EX1009 (showing comparison of the claims).) Thus, under **Ground 2**, I rely on all the same disclosures in Steuer768 corresponding to the relied-upon disclosures

in Steuer750 under **Ground 1** (and as noted above, provides citations to both Steuer750 and Steuer768), but further rely on claims 1-15 of Steuer768 as they almost verbatim disclose challenged claims 1-15 except the “video” feature. The citations to the claims of Steuer768 are applicable only to **Ground 2**. The reliance on Neumeier is the same in both **Ground 1 and 2**, as is my opinion regarding the rationale for combining Neumeier with Steuer750 or Steuer768.

1. Claim 1

a) [pre] A media measurement method comprising:

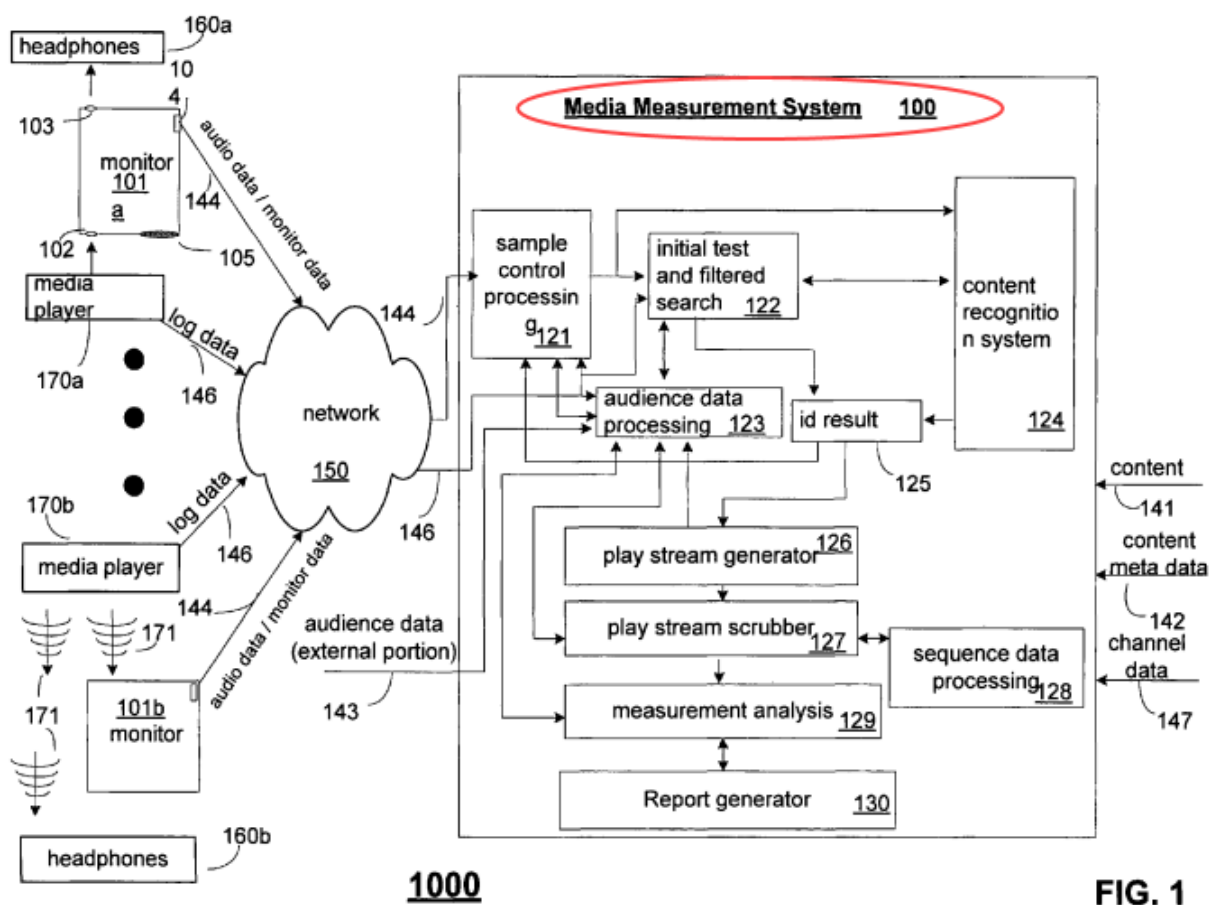
90. I have been asked to assume that the preamble of claim 1 is limiting. Under that assumption, it is my opinion that Steuer750 and Steuer768 both independently disclose the limitation therein. (EX1005 at Title, Abstract, FIGS. 1-2, claims 1, 19, 35, 40, 44, 48, 61-62; EX1006 at Title, Abstract, FIGS. 1-2, claims 1-15.) As an initial matter, both Steuer750 and Steuer768 claim “[a] media measurement method[.]” (EX1005 at claims 1, 19, 35, 40, 44, 48, 61-62; *see also id.* at claim 54 (“A computer reprogram product for measuring media usage”); EX1006 at claim 1.) Steuer750 and Steuer768 also refer to a media measurement method in the specification. For instance, they disclose:

Some embodiments of the present invention 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. Some

embodiments of the present invention provide a **media measurement system and method** that utilizes audience data to enhance content identifications. Some embodiments analyze media player log data to enhance content identification . . . Yet other embodiments provide a method in which play-altering activity of an audience member is deduced from content offset values of identifications corresponding to captured samples.

(EX1005 at [0006]; *see also id.* at Abstract, [0002]; EX1006 at 1:60-2:10; *see also id.* at Abstract, 1:14-32.)

91. Steuer750 and Steuer768 further disclose a “Media Measurement System 100” (Figure 1) for carrying out its methods and “illustrates a media measurement method” in Figure 2 (EX1005 at [0009]; EX1006 at 2:26-28).



(EX1005 at FIG. 1 (annotated); *see also id.* at [0008], [0016]-[0017]; EX1006 at FIG. 1; *see also id.* at 2:23-25, 2:66-3:25.)

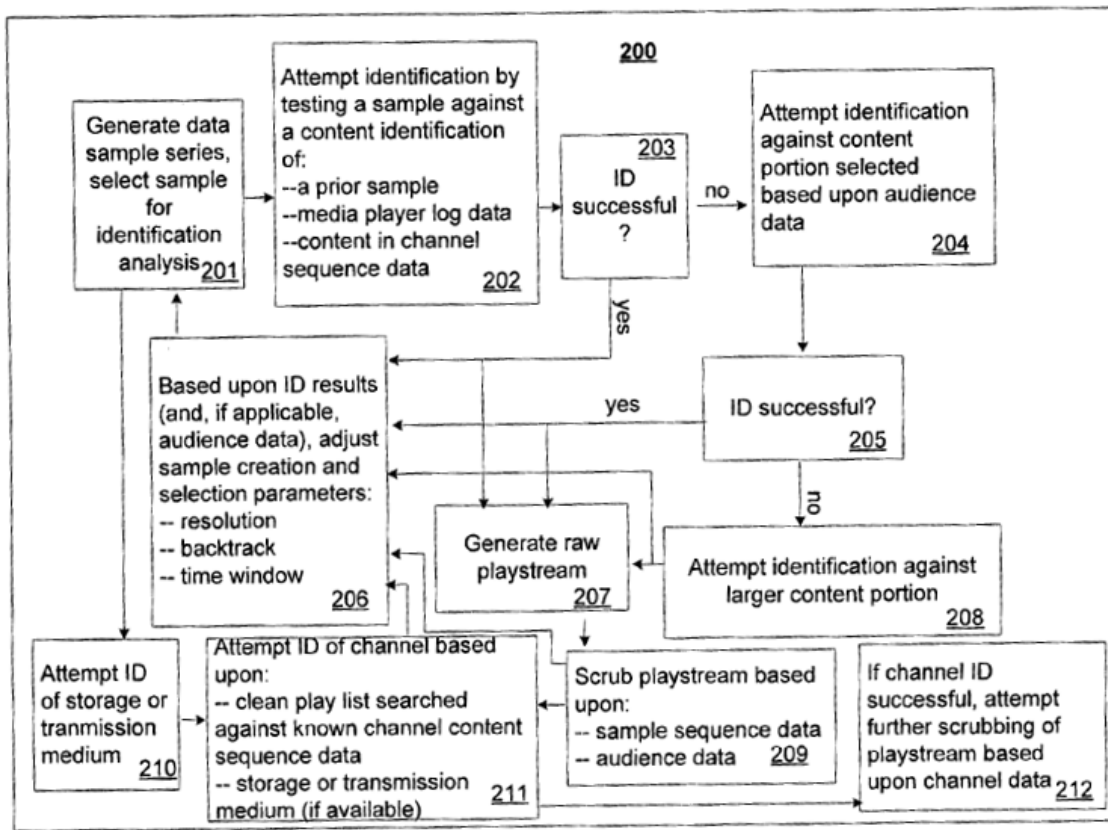


FIG. 2

(EX1005 at FIG. 2; *see also id.* at [0009], [0034]; EX1006 at FIG. 2; *see also id.* at 2:26-28, 6:58-7:3.)

92. Thus, Steuer750 and Steuer768 both disclose “[a] media measurement method,” as claimed. (*See also* Sections IX.A.1.b–f below.)

- b) (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;**

93. In my opinion, the Steuer750 in combination with Neumeier and Steuer768 in combination with Neumeier disclose or suggest this limitation.

94. Steuer750 and Steuer768 disclose receiving over a network 150, using one or more computers (e.g., the media measurement system 100 implemented by computer system 700), data for a sequence of data samples (e.g., audio data samples), the data samples comprising representations of audio data captured at monitors 101, the plurality of audio data samples being submitted over the network 150. (EX1005 at [0017]-[0019], [0023], [0034], [0066] (disclosing that “computer system 700” of Figure 7 may be used to implement “the system 100 of FIG. 1”), FIGS. 1, 2, 4, 7; claims 1, 19, 35; EX1006 at 3:4-54, 4:49-62, 6:58-7:3, 13:54-14:11 (disclosing that “computer system 700” of Figure 7 may be used to implement “the system 100 of FIG. 1”), FIGS. 1, 2, 4, 7;.)

95. For example, similar to claim limitation 1.b, Steuer750 discloses and claims “submitting data samples corresponding to portions of audio data captured at a monitoring device for content recognition against known media content.” (EX1005 at claim 1.)

96. Steuer768's claim 1 recites almost identical language as claim limitation 1.b except it substitutes the claimed video data samples with audio data samples: "receiving over a network, at one or more computers, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network." (EX1006 at claim 1.)

97. Steuer750 and Steuer768 disclose that the sequence of data samples is a sequence of data samples comprising representations of audio data (e.g., audio signatures) captured by one or more monitors 101 ("media monitoring device"), and then transferred ("submitted") to the media measurement system 100 via network 150.

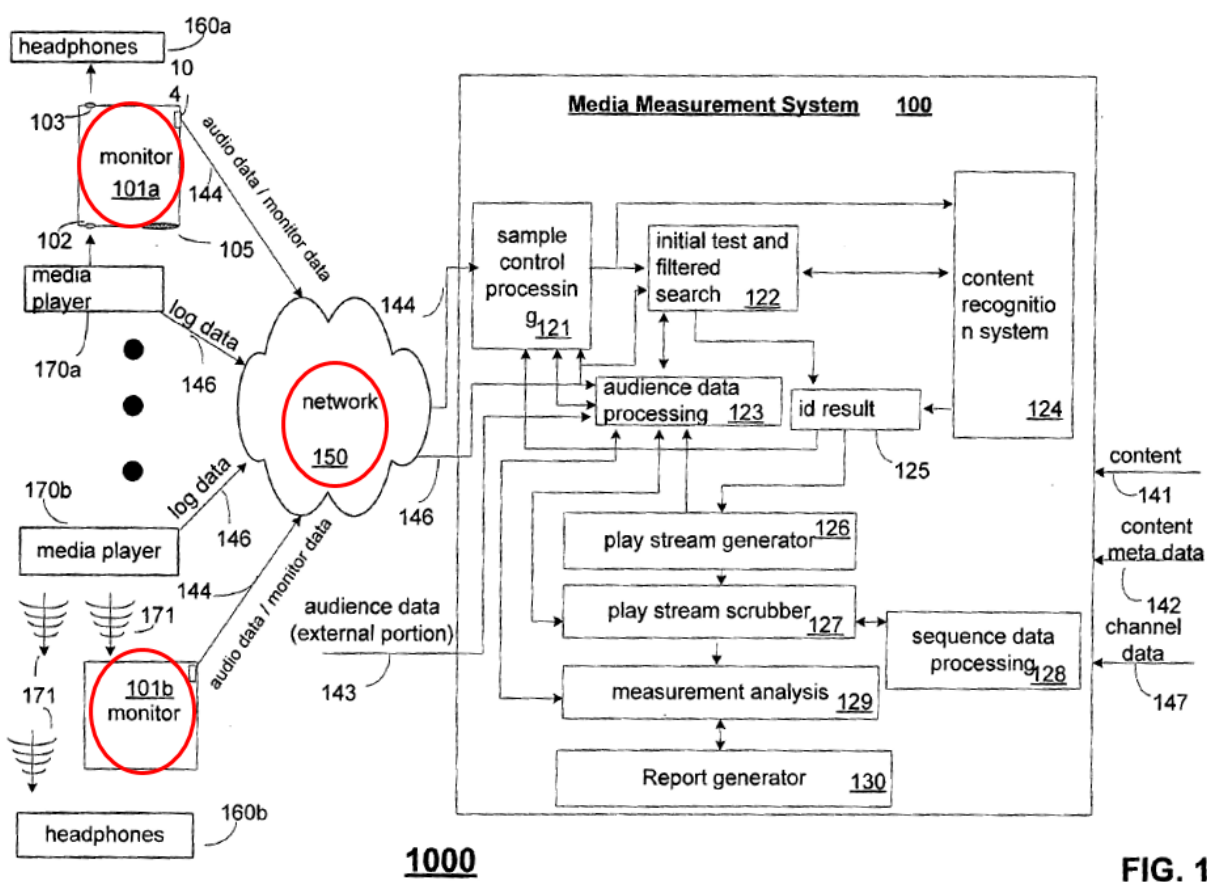
As illustrated, **monitoring device 101a** includes a microphone 105, a media player port 102, a headphone port 103, and a data upload port 104. **A monitored audience member wears monitor 101a during a monitoring period.** When the audience member is consuming media from media player 170a without the use of headphones 160 a (i.e. if media player 170a has a speaker), monitor 101 captures audio energy acoustically through microphone 105 . . .

As another example, headphones 160 b are adapted to receive a content signal 171 via wireless transmission 171 from media

player 170 b. Monitor 101 b is adapted to receive the same content signal 171 from media player 170 b . . .

Audio data captured by monitors 101, along with associated monitor data (e.g. device ID) (collectively, “audio data/monitor data” 144) may be uploaded to system 100 through upload port 104, directly, via memory device transfer (e.g. flash memory card, floppy disk, CD, etc.) or through a network such as network 150. The audio data uploaded from a monitor 101 may be raw audio data, or may be audio data that has undergone one of a variety of levels of processing prior to upload. For example, **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.** Alternatively, **the audio data may comprise pre-calculated signatures, landmarks, or other data useful for content recognition.** On the other hand, in some embodiments, the uploaded audio data may simply represent a raw signal corresponding to audio energy received at monitor 101.

(EX1005 at [0017]-[0019], FIGS. 1-2; EX1006 at 3:4-54, FIGS. 1-2.)



(EX1005 at FIG. 1 (annotated); EX1006 at FIG. 1.)

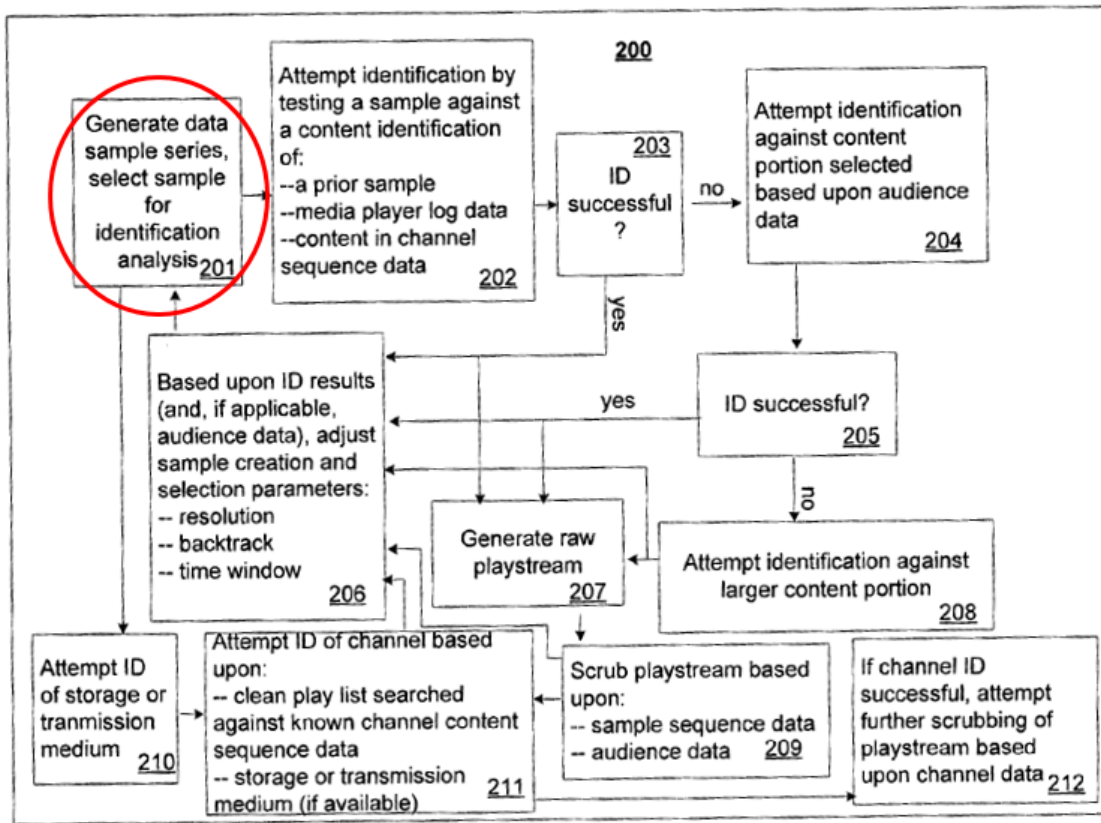


FIG. 2

(EX1005 at FIG. 2 (annotated); EX1006 at FIG. 2.)

98. The disclosures of Steuer750 and Steuer768 are limited to media measurement methods and systems for generating a play stream of content identification results corresponding to a sequence of **audio** data samples and performing subsequent data processing steps based on the content identification results corresponding to the audio data samples. (EX1005 at Abstract, [0006], [0017]-[0019] (describing extracting audio data from various media sources including “an MP3 player, iPod, CD player, DVD player, television, radio, computer, or other media player.”), [0022]-[0023], [0034]-[0035]; *see also id.* at

claims 1-18, [0040]-[0042] (describing subsequent step of scrubbing the play stream), [0052], [0058]-[0059] (describing subsequent step of deducing play-altering actions); EX1006 at Abstract, 1:60-2:10, 3:4-54 (describing extracting audio data from various media sources including “an MP3 player, iPod, CD player, DVD player, television, radio, computer, or other media player.”), 4:13-62, 6:58-7:16; *see also id.* at claims 1-15, 8:1-49 (describing subsequent step of scrubbing the play stream), 10:57-11:9, 12:1-50 (describing subsequent step of deducing play-altering actions), 6:58-7:3.)⁸ As such, in view of the disclosures above, Steuer750 and Steuer768 disclose “receiving over a network, at one or more computers, data for a sequence of [audio] data samples, the [audio] data samples comprising representations of [audio] data captured at a media monitoring device, the plurality of [audio] data samples being submitted over the network.” Steuer750 and Steuer768 do not disclose that the data samples are **video** samples including **representations of video data** captured at a media monitoring device, as claimed. However, as I discuss below, a person of ordinary skill in the art would have been motivated to

⁸ Steuer750’s published disclosure corresponds to the as-filed specification of the ’330 application and is thus limited to audio-based embodiments for all the reasons I discussed above in Section VIII.A.1.b. The same is true for Steuer768 as its disclosure is nearly identical to that of Steuer750. (EX1008.)

modify Steuer750's and Steuer768's disclosed media measurement methods such that data samples including representations of video data, e.g., video data samples, are used instead of audio based on the teachings of Neumeier.

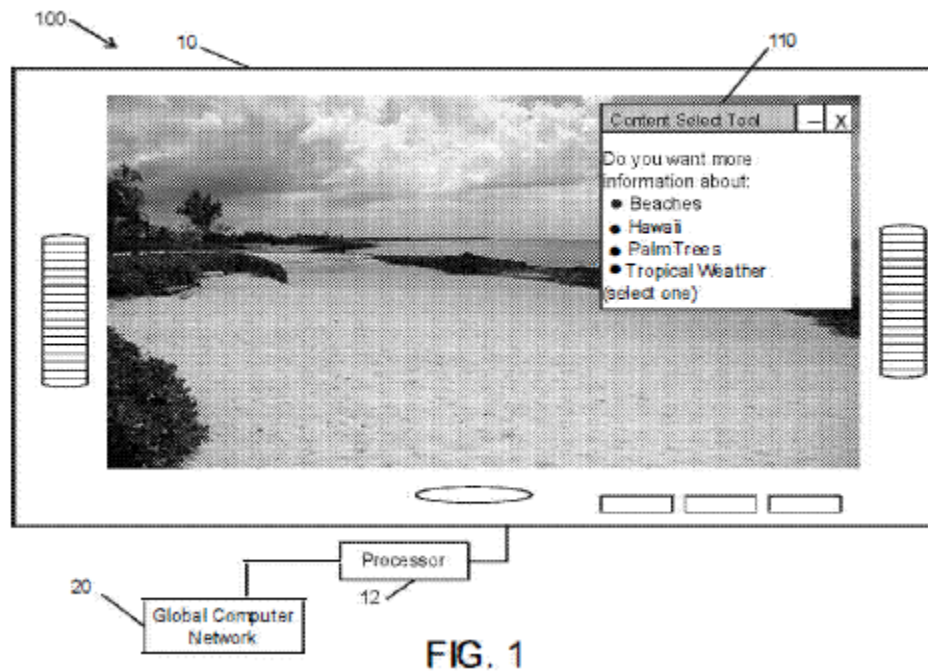
99. Similar to Steuer750 and Steuer768, Neumeier is directed to methods and systems used to identify media content being consumed by a user by matching samples of the consumed media content with a content database, and then performing subsequent data processing steps based on the matching. (*See, e.g.*, (EX1007 at Abstract.) Neumeier discloses that its content identification method can be implemented using video data samples or audio data samples.

The present invention is directed to systems and methods for identifying which video segment is being displayed on a screen of a television system. In particular, the resulting data identifying the video segment being viewed **can be used to extract a viewer's reaction (such as changing the channel) to a specific video segment (such as an advertisement) and reporting the extracted information as metrics.**

In accordance with some embodiments, **the video segment is identified by sampling a subset of the pixel data being displayed on the screen (or associated audio data) and then finding similar pixel (or audio) data in a content database.** In accordance with other embodiments, the video segment is identified by extracting audio or image data associated with such video segment and then finding similar audio or image data in a

content database. In accordance with alternative embodiments, the video segment is identified by processing the audio data associated with such video segment using known automated speech recognition techniques. In accordance with further alternative embodiments, the video segment is identified by processing metadata associated with such video segment.

(EX1007 at [0010]-[0011]; *see also id.* at [0002], [0012]-[0014].)



(EX1007 at FIG. 1; *see also id.* at [0023].)

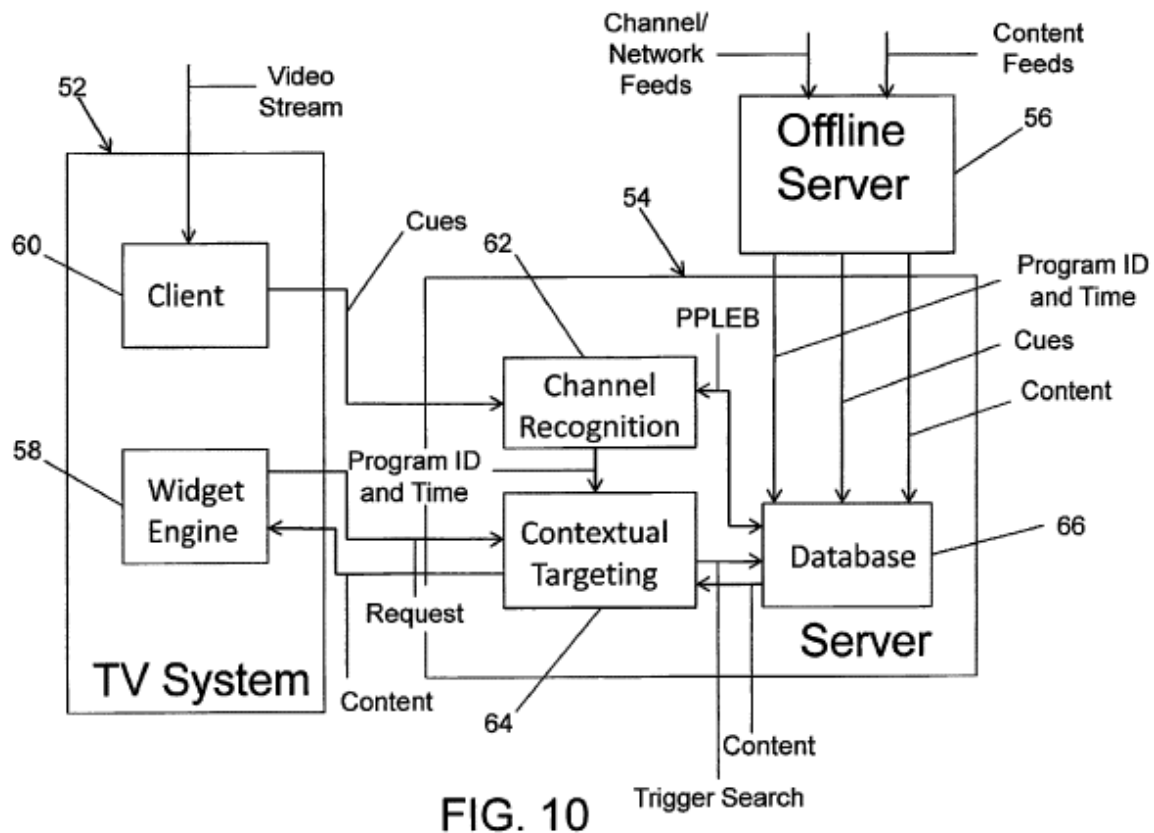


FIG. 10

(EX1007 at FIG. 10; *see also id.* at [0050-[0052].)

100. Thus, Neumeier discloses content identification systems and methods in the same technical field as Steuer750 and Steuer768, including content identification based on captured data samples. Accordingly, a person of ordinary skill in the art would have had reason to consider the teachings of Neumeier when implementing Steuer750 or Steuer768.

101. Neumeier discloses identifying content corresponding to a sequence of sampled pixel data (“sequence of video data samples, the video data samples comprising representations of video data”) captured at a TV or the like (like the

monitoring device of Steuer750/Steuer768) associated with a user (like the monitored audience member of Steuer750/Steuer768). In particular, Neumeier discloses with reference to Figure 10 that “the client module 60 [in television system 52] is programmed to **sample pixel data** and generate an HTTP request addressed to server 54 based on the sampled pixel data,” which the server 54 uses to identify the content being viewed by the user by way of its channel recognition module 62 and database 66. (EX1007 at [0054].) The operations performed by the channel recognition module 62 to identify content (e.g., a program ID) include matching the received sample pixel data with known content stored in database 66.

The channel recognition module 62 comprises a points management submodule and a user management submodule (not shown in FIG. 10). The points management submodule **searches the database 66** in two ways: (1) a search of the entire library for a given set of points, returning all the suspects that are close to a match; and (2) a search for a given set of points and a given suspected location, returning whether the user is indeed where the currently stored data indicates. [A “user” is a unique TV or other device identified by a globally unique ID.]

(EX1007 at [0056]; *see also id.* at [0057].)

As indicated by the arrow labeled “CUES” in FIG. 10, the client module 60 sends the user management submodule of the channel recognition module 62 a regular update with pixel cue information. This communication is done via the

aforementioned HTTP request and the pixel cue information is sent over GET parameters . . .

The parameters contained in this HTTP request are as follows:

The parameter “cueData” is a list of RGB values, e.g., samples of pixel values composed of RGB combinations. The format is R1-G1-B1,R2-G2-G2, . . . |R3-G3-B3,R4-G4-B4, . . . |etc., where each RX-GX-BX indicates a respective RGB Location. . .

(EX1007 at [0059]-[0062]; *see also id.* at [0064], [0068], claims 15, 18.)

In response to receipt of the HTTP request from the client module 10, **the channel recognition module 62 identifies in real time what video segment the cueData in the HTTP request is taken from and in what time offset from the starting time of the segment.** As previously mentioned, the segment and offset together are referred to as the “location” . . .

(EX1007 at [0070]; *see also id.* at [0072].)

Once the most likely location in the database is identified, content stored in the database in association with that location can be retrieved by the contextual targeting module 64 (see FIG. 10). In accordance with the preferred embodiment, by the contextual targeting module 64 receives the program ID and the time offset of the most likely suspect location (i.e., the suspect location having the greatest probability provided that probability exceeds a threshold for success which is presettable) from the

channel recognition module 62 and then uses that information to retrieve the associated enhanced content from the database 66.

(EX1007 at [0076].)

In accordance with the preferred embodiment, the offline server 56 extracts time stamps, pixel cue points and closed captioning from the channel/network feeds. That extracted information is stored as part of the database 66. **More specifically, the database contains the following information for each television program, commercial or other broadcast or video segment: (a) a list of pixel cue points for each video segment; (b) offsets from some fixed point in time, which offsets are respectively associated with the aforementioned pixel cue points, thereby indicating the sequence in time when those pixel cue points occur; and (3) associated metadata (e.g., closed captioning) . . .**

(EX1007 at [0078].)

In accordance with the embodiment depicted in FIG. 10, the database 66 stores pixel cue points and content for a multiplicity of video segments, while the server 54 is programmed to perform the following steps: (a) determining which pixel cue points stored in the database are possible matches to pixel cue points received from a television system 52 via a network; (b) computing a probability distribution for the pixel cue points determined in step (a); (c) retrieving from the database a program identifier and a playing time associated with the pixel cue points

determined to have a greatest probability of matching the pixel cue points received from the television system; (d) retrieving from the database content associated with the program identifier and the playing time retrieved in step (c); and (e) sending the content to the television system via the network.

(EX1007 at [0097]; *see also id.* at [0099], [0101].)

102. Neumeier further discloses additional process steps related to content identification, similar to those disclosed in Steuer750 and Steuer768, such as **filtering deviations in the sample pixel data to identify content, identifying a channel being viewed, and determining a content offset time.** For example, regarding filtering of deviations in the sample pixel data, Neumeier discloses:

The path pursuit algorithm is a probabilistic approach: an exact matching of pixel cue points at all times is not needed. Instead the decision that the result is true is made based on the aggregated evidence. **The algorithm tracks in real time all of the time and is able to handle intermittent pixel cue points in a sequence that deviate from the other pixel data points in that sequence.** For example, the algorithm may only recognize 7 out of 10 frames of a video segment, but is still able to identify the most likely location. The algorithm also responds quickly to the television viewer pausing, changing channels, etc.

(EX1007 at [0073]⁹; *see also id.* at p. 18 (Appendix), Section 2.2.)

103. Regarding channel identification, Neumeier discloses that “the TV client 18 may send pixel cues (consisting of batches of pixel data samples) to the server 520, and the channel matching module 530 may identify the channel being viewed by processing ambiguous pixel cues using the technique disclosed in the Appendix hereto.” (EX1007 at [0045]; *see also id.*, [0040]-[0041], [0043], [0056], [0070]; EX1005 at [0041], [0043]-[0044], [0046], FIG. 4; EX1006 at 8:23-29, 8:50-9:11, 9:32-37.)

104. And regarding a content offset time, Neumeier discloses that in addition to identifying video content based on the video data samples, it also identifies the playing time or offset time of the video content being displayed. (EX1007 at [0012]; *see also* EX1005 at [0006], [0053], [0055]; EX1006 at 1:60-2:10, 11:10-26, 11:38-55.) Neumeier discloses that “[t]he terms ‘playing time’ and ‘offset time’ will be

⁹ Neumeier’s disclosure references its Appendix at pp. 13-20 which discloses an algorithm for tracking video transmission using ambiguous cues. (*See* EX1007 at pp. 13-18; *see also id.*, [0021] (referencing Appendix), [0040] (same), [0045] (same), [0070]-[0071] (same), [0075].) Thus, to the extent I cite to portions of Neumeier that reference the Appendix, I also rely on the algorithm disclosed in the Appendix.

used interchangeably herein and refer to a time which is offset from a fixed point in time, such as the starting time of a particular television program or commercial.”

(EX1007 at [0012].)

In accordance with some embodiments, the video segment is identified **and the offset time is determined** by sampling a subset of the pixel data (or associated audio data) being displayed on the screen and then finding similar pixel (or audio) data in a content database. In accordance with other embodiments, the video segment is identified and the offset time is determined by extracting audio or image data associated with such video segment and then finding similar audio or image data in a content database . . .

(EX1007 at [0014]; *see also id.*, [0024], [0032], [0070], [0072], [0076], [0078], [0087].)

Given a large number of video segments, the system must be able to identify in real time what segment a given query video input is taken from **and in what time offset**. The segment and offset together are referred to as the location. **The method is called video pursuit since it must be able to efficiently detect and adapt to pausing, fast forwarding, rewinding, abrupt switching to other segments and switching to unknown segments.**

(EX1007 at p. 13 (Appendix), Abstract.)

105. In view of Neumeier's teachings, it is my opinion that a person of ordinary skill in the art would have been motivated to consider and draw from such teachings when implementing Steuer750's or Steuer768's disclosed methods and systems such that, in addition to, or as an alternative to, using audio data samples, data samples including representations of video data (e.g., video data samples) would have been used to generate content identification results and perform subsequent data processing steps. As I explained above, Steuer750, Steuer768, and Neumeier all concern content identification processes that require collecting data samples, making content identifications based on those data samples, and comparing sequences of content identifications against known sequences.

106. In my opinion, a person of ordinary skill in the art would have been motivated and would have known how to modify Steuer750 and Steuer768 to incorporate Neumeier's teachings as proposed, because as I discussed above, Neumeier discloses a way of identifying content based on matching video data samples with video data stored in a database. Thus, a person of ordinary skill in the art would have recognized that applying Steuer750's or Steuer768's use of audio data to video data could have been implemented based on readily recognizable modifications using known technology, such as sampling video data instead of (or in addition to) sampling audio data, and using the sampled video data to identify content being consumed by a user as taught by Neumeier.

107. In view of the foregoing, a person of ordinary skill in the art would have had the capability and knowledge to implement, and would have been motivated to try, the above modification to include the use of video data instead of (or in addition to) audio data as required by the claim, while ensuring proper operation of the Steuer750/Steuer768 systems and methods. Therefore, a person of ordinary skill in the art would have had a reasonable expectation of success that the combined Steuer750-Neumeier system and the combined Steuer768-Neumeier system would have operated as expected.

108. The above modification would have been a mere combination of known technologies (e.g., the concept of collecting video data samples with Steuer750's/Steuer768's disclosed systems and methods for making content identifications), according to known methods (e.g., sampling pixel data of the video content), to achieve predictable results (e.g., using video data samples to make content identifications by comparing such samples to known content to generate content identification results).

109. Therefore, the Steuer750-Neumeier system and the Steuer768-Neumeier system discloses or suggests claim limitation 1.b.

- c) **(b) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers;**

110. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests this limitation.

111. First, Steuer768's claim 1 recites almost identical language as claim limitation 1.c, albeit in the context of "audio data samples" as noted with respect to limitation 1.b: "using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers." (EX1006 at claim 1.)

112. Further, regarding limitation 1.b, Steuer750 and Steuer768 both disclose using one or more computers (e.g., the media measurement system 100 implemented by computer system 700) to query an electronic database of a plurality of video data representations and corresponding content identifiers. (EX1005 at [0022]-[0031], [0034]-[0036], FIGS. 1-4; EX1006 at 4:13-6:27, 6:58-7:38, FIGS. 1-4; *see also* Section IX.A.1.b above.) In particular, Steuer750 and Steuer768 disclose querying a database of known content 141 of a plurality of audio data representations (e.g., audio signatures) and corresponding content identifiers.

System 100 is also adapted to receive **known content 141 and content meta data 142**. Content 141 may include media files associated with a known piece of media content (e.g. a song, movie, commercial, television show, video game, etc.).

However, it will be understood by those skilled in the art that “content” in the sense of the known content 141 received by system 100 may, in some embodiments, **simply refer to data about signals representing the content and does not necessarily refer to a stored version of the media content itself; e.g. “content” may be stored signatures and landmarks derivable from a song’s audio**, but not necessarily an audio file useable to play the song itself . . .

(EX1005 at [0022]; EX1006 at 4:13-48.)

At step 202, identification of a sample in a series of samples is attempted via initial test methods. In particular, step 202 determines whether the audio data in the tested sample matches data associated with the content ID corresponding to the most recently identified sample. Step 202 may also or instead test the sample against content IDs obtained using data from a media player log file associated with the audience member using the corresponding media monitor. Step 203 determines whether successful identification occurred at step 202. If yes, then the identification result is provided to step 206 and 207. If no, then step 204 searches for a content match against targeted content selected based at least in part upon additional audience data. Step 205 determines whether step 204 obtained an identification. If yes, then the identification result is provided to steps 206 and 207. **If no, then step 208 searches for a content match against a larger portion of the content.** With respect to steps 202, 204, and 208, there are two possible results: an identification has been

made or no identification has been made. No identification at steps 202 and 204 leads to the further identification attempt at step 208. Whatever the result is at step 208 (identification or no identification) that result is passed to steps 207 and 206.

(EX1005 at [0036]; EX1006 at 7:17-38.)

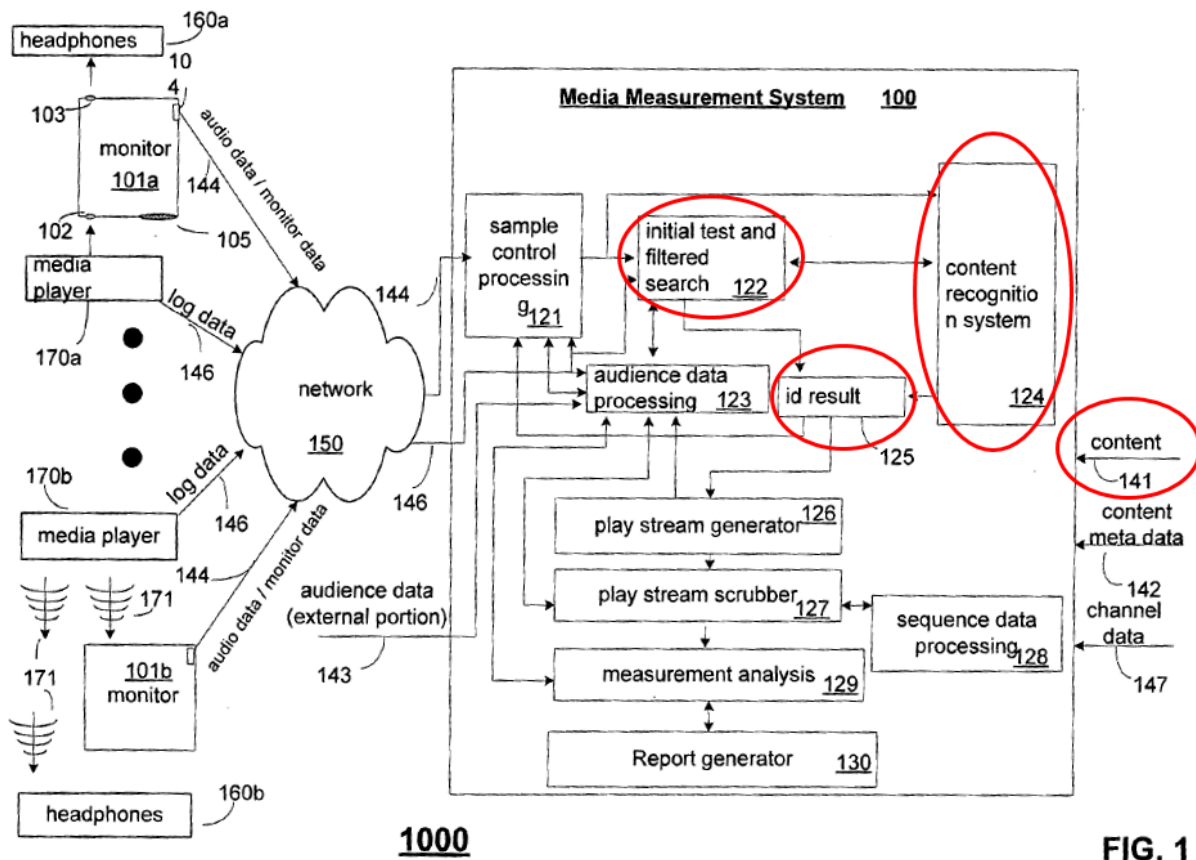


FIG. 1

(EX1005 at FIG. 1 (annotated); EX1006 at FIG. 1.).

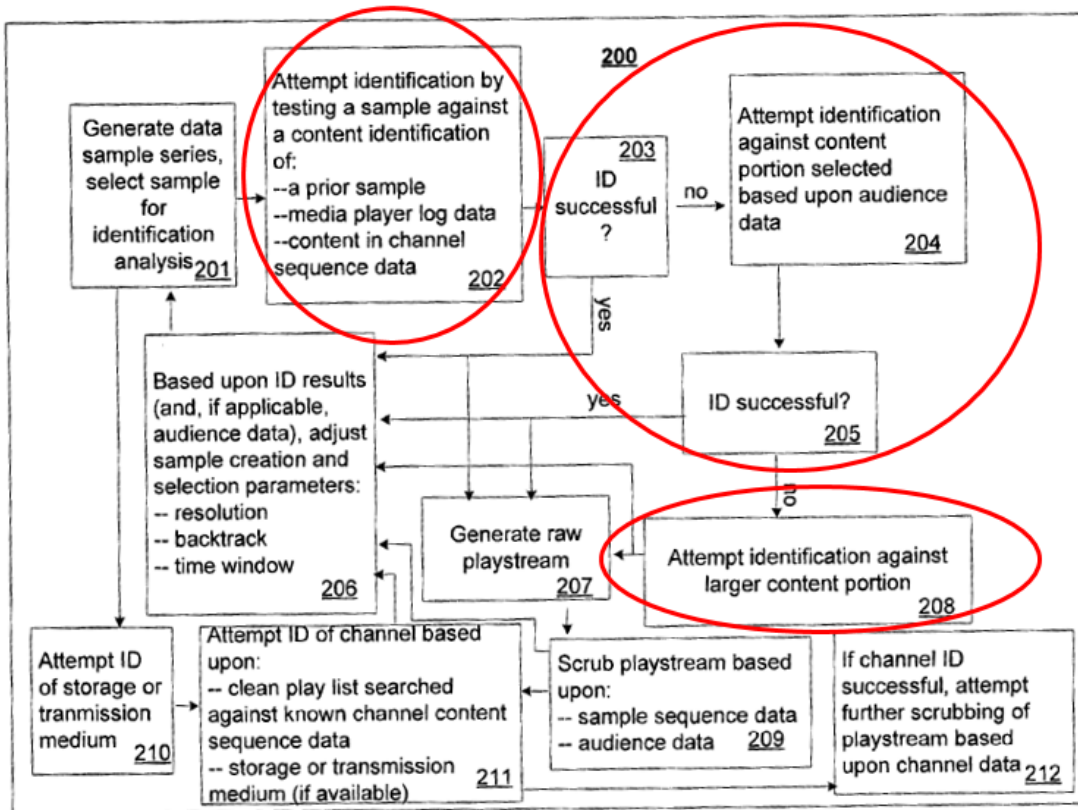


FIG. 2

(EX1005 at FIG. 2 (annotated); EX1006 at FIG. 2.)

113. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 1.c.

d) (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;

114. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggests this limitation.

115. Steuer750 and Steuer768 disclose generating a raw play stream (e.g., using the play stream generator 126), the raw play stream including a sequence of content identification results corresponding to the sequence of video data samples. (EX1005 at [0016]-[0031], [0040], FIGS. 1, 2, 4, 7; claims 35, 40, 44; EX1006 at 2:66-6:27, 8:1-22, FIGS. 1, 2, 4, 7.)

116. For example, similar to claim limitation 1.b, Steuer750 discloses and claims “generating a raw play stream of content identification results corresponding to a sequence of data samples.” (EX1005 at claims 35, 40, 44.) Steuer768’s claim 1 recites almost identical language as claim limitation 1.d, albeit in the context of audio data samples as noted with respect to limitation 1.b: “generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples.” (EX1006 at claim 1.)

117. Steuer750 and Steuer768 disclose additional details regarding generation of the play stream in relation to the play stream generator 126 of Figure 1 (which generates the play stream based on information processed at upstream modules including sample control processing module 121, initial test and filtered

search module 122, audience data processing module 123, content recognition system 124, and ID result module 125).

Turning to the details of system 100, **audio data and monitor data (e.g. device ID) is received at sample control processing module 121**. Module 121 provides monitor data, portions of which might be audience data, to processing module 123. **Module 121 also divides the audio data into samples and determines which samples should be submitted for further content recognition testing . . .**

Initial test and filtered search module 122 conducts one or more initial identification attempts. Several different types of initial identification attempts might be performed. **Module 122 may test a captured sample against content associated with the most recently identified content**. As another possible initial test, module 122 may test the sample against content identified from a media player log file. Another initial identification test may be carried out in a variety of ways. For example, filtered search module 122 might use audience data from audience data processing module 123 to form parameters and then pass those parameters to content recognition system 124 which then uses the parameters to conduct a filtered search of its content . . .

After completing the initial identification attempt, filtered search module 122 passes the identification result to ID result module 125. If an identification has been made, the result is passed to play stream generator 126 and to sample control

processing module 121. If no identification has been made, then content recognition system 124 performs a search using a larger portion of its content to identify the data sample. The result is passed to ID result module 125 which in turn passes the result to play stream generator 126 and to sample control module 121 . . .

Play stream generator 126 generates a raw play stream relating a sequence of samples to content identification results (for an example of a portion of a raw play stream, see FIG. 3 and accompanying text). The raw play stream is then processed by play stream scrubber 127 which analyzes the sequence of sample identification results utilizing sequence data processing module 128. . .

(EX1005 at [0023]-[0031]; *see also id.*, [0040] (describing raw play stream generation in step 207 of Figure 2), [0041]-[0042], [0045] (describing raw play stream 410 generated by step 207 as “itemiz[ing] a sequence of analyzed samples and relat[ing] that sequence to corresponding identification results in the content ID column.”), FIGS. 1-4; EX1006 at 4:49-6:27; *see also id.*, 8:1-22 (describing raw play stream generation in step 207 of Figure 2), 8:23-49, 9:12-37 (describing raw play stream 410 generated by step 207), FIGS. 1-4.)

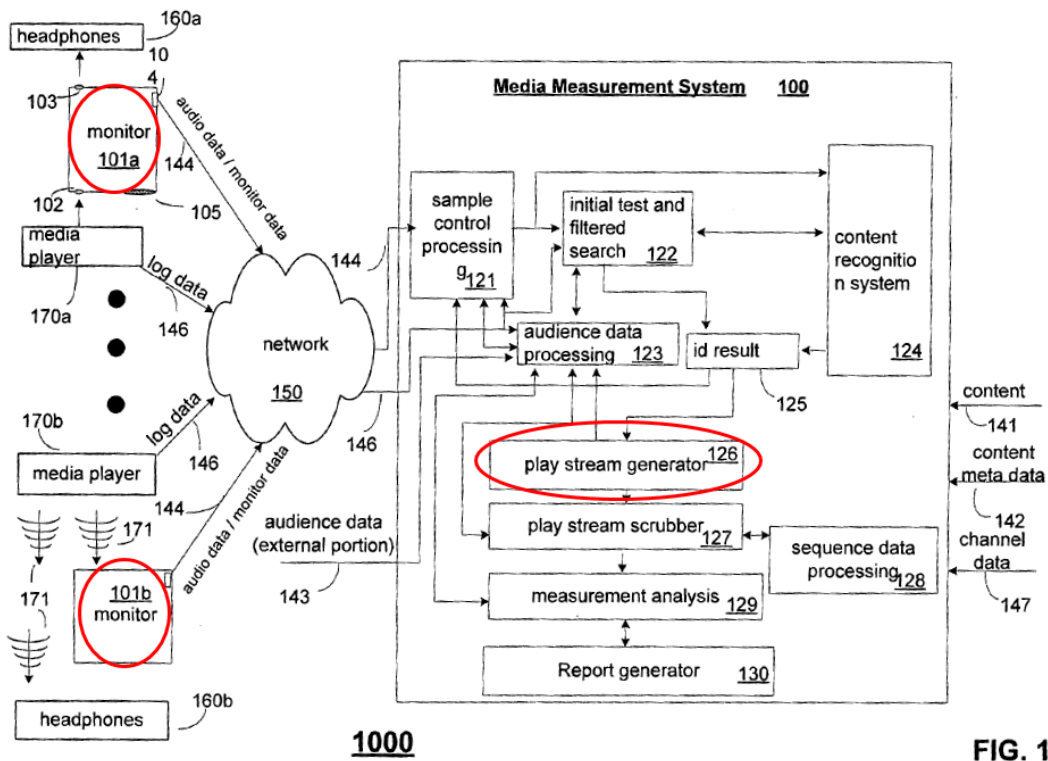


FIG. 1

(EX1005 at FIG. 1 (annotated); EX1006 at FIG. 1.)

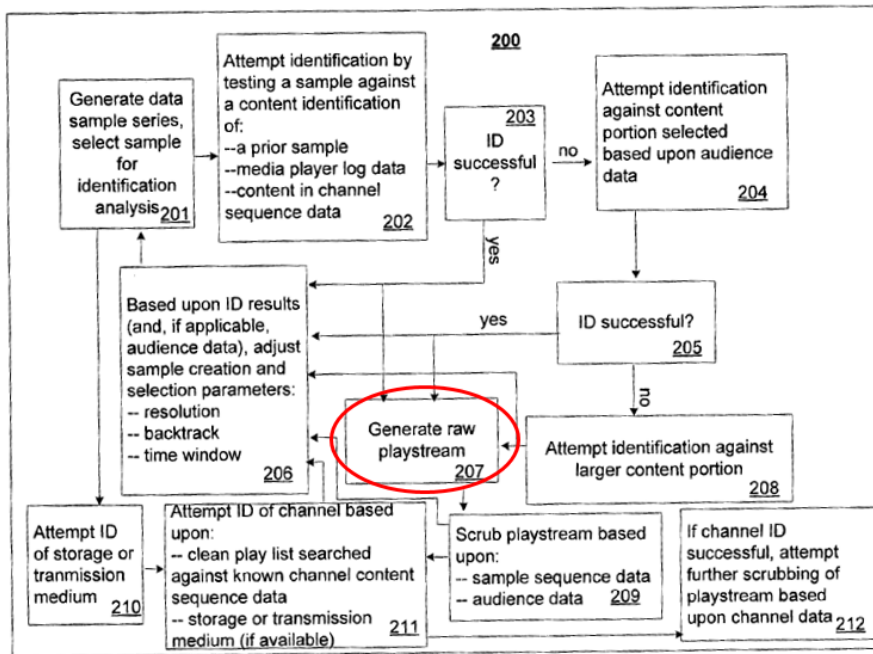


FIG. 2

(EX1005 at FIG. 2 (annotated); EX1006 at FIG. 2.)

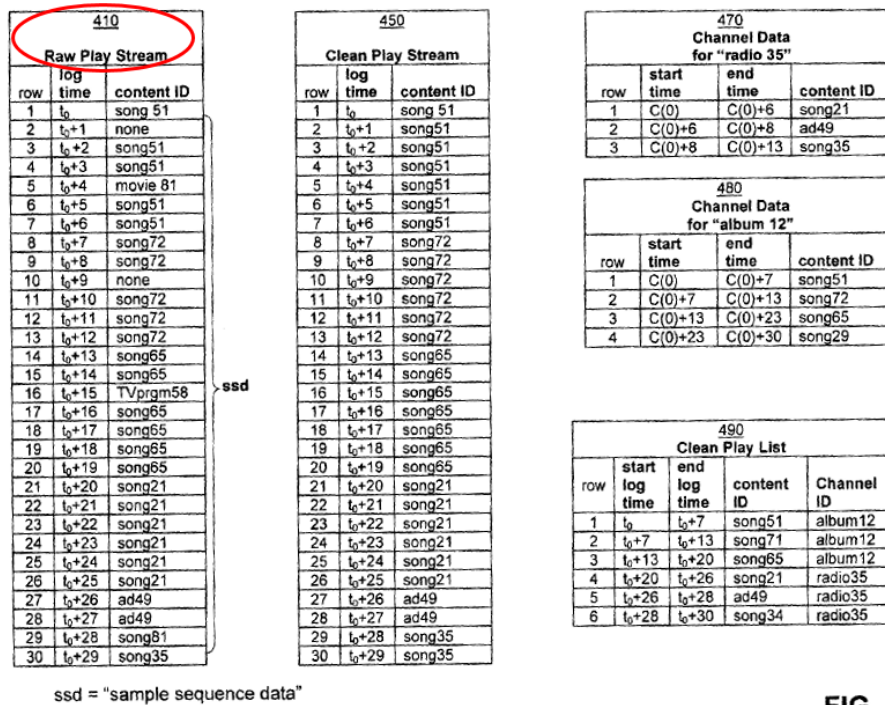


FIG. 4

(EX1005 at FIG. 4 (annotated); EX1006 at FIG. 4.)

118. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 1.d.

- (1) **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;**

119. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons as those given above with respect to limitations 1.c and 1.d. (*See* Sections IX.A.1.c-d, above.)

120. For example, as I discussed above with respect to limitations 1.c and 1.d, Steuer750 and Steuer768 both disclose that obtaining the content identification results (e.g., the content IDs shown in raw play stream 410 in figure 4 by querying the electronic database of known content 41 to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations e.g., audio samples) in the electronic database. (*See* Sections IX.A.1.c-d, above (citing EX1005 at [0016]-[0031], [0034]-[0036], [0040], FIGS. 1-4; EX1006 at 2:66-6:27, 6:58-7:38, 8:1-22, FIGS. 1-4).)

121. Steuer768's claim 1 recites almost identical language as claim limitation 1.d, albeit in the context of audio data samples as noted with respect to limitation 1.b: "the sequence of content identification results is obtained by querying

the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database.” (EX1006 at claim 1.)

122. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750’s and Steuer768’s disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (*See* Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 1.d.1.

- (2) 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; and¹⁰**

123. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons as those given above with respect to limitations 1.c and 1.d. (See Sections IX.A.1.c-d, above.)

124. For example, Steuer750 and Steuer768 both that a content identification result of the sequence of content identification results includes either (i) a content ID (e.g., “song51” or “movie81” in content ID column of raw play stream 410) associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data

¹⁰ Claim limitation 1.d.2 recites alternative language in the form of the content identification result comprising “**either:** (i) . . .; **or** (ii) . . .” While the claim limitation is met by the prior art showing at least one of these alternatives, I map the prior art to both alternatives below.

samples or (ii) an indication of the absence of a respective likely match (e.g., “none” in content ID column of raw play stream 410) between a respective audio data sample and a audio data representation in the electronic database.

At step 202, identification of a sample in a series of samples is attempted via initial test methods. In particular, step 202 determines whether the audio data in the tested sample matches data associated with the content ID corresponding to the most recently identified sample. Step 202 may also or instead test the sample against content IDs obtained using data from a media player log file associated with the audience member using the corresponding media monitor. Step 203 determines whether successful identification occurred at step 202. If yes, then the identification result is provided to step 206 and 207. If no, then step 204 searches for a content match against targeted content selected based at least in part upon additional audience data. Step 205 determines whether step 204 obtained an identification. If yes, then the identification result is provided to steps 206 and 207. If no, then step 208 searches for a content match against a larger portion of the content. **With respect to steps 202, 204, and 208, there are two possible results: an identification has been made or no identification has been made.** No identification at steps 202 and 204 leads to the further identification attempt at step 208. Whatever the result is at step 208 (identification or no identification) that result is passed to steps 207 and 206.

(EX1005 at [0036]; EX1006 at 7:17-38.)

FIG. 4 illustrates Raw Play Stream 410 generated by a process such as step 207 of FIG. 2; Clean Play Stream 450 and Clean Play List 490 generated by a scrubbing step such as step 209 of FIG. 2 or steps 301, 302, or 305 of FIG. 3; Channel Data 470 and Channel Data 490. Referring to FIG. 4, Raw Play Stream 410 itemizes a sequence of analyzed samples and relates that sequence to corresponding identification results in the content ID column. **The content ID result indicates either a reference corresponding to a particular piece of content or indicates that no identification was made.** The example in FIG. 4 assumes that no user play-altering activities took place and this example does not rely on content offset data in scrubbing the raw play stream to obtain a clean play list. However, note that in other examples, such as that of FIG. 5, sample sequence data (“ssd”) includes content offset data. For purposes of presenting and analyzing sample sequence data as described herein, the times in list 410 may be relative to an arbitrary value (e.g. “t0”) rather than representing an absolute time; however, in many instances, a log time will in fact indicate the actual time the sample started to be captured.

(EX1005 at [0045]; EX1006 at 9:12-37.)

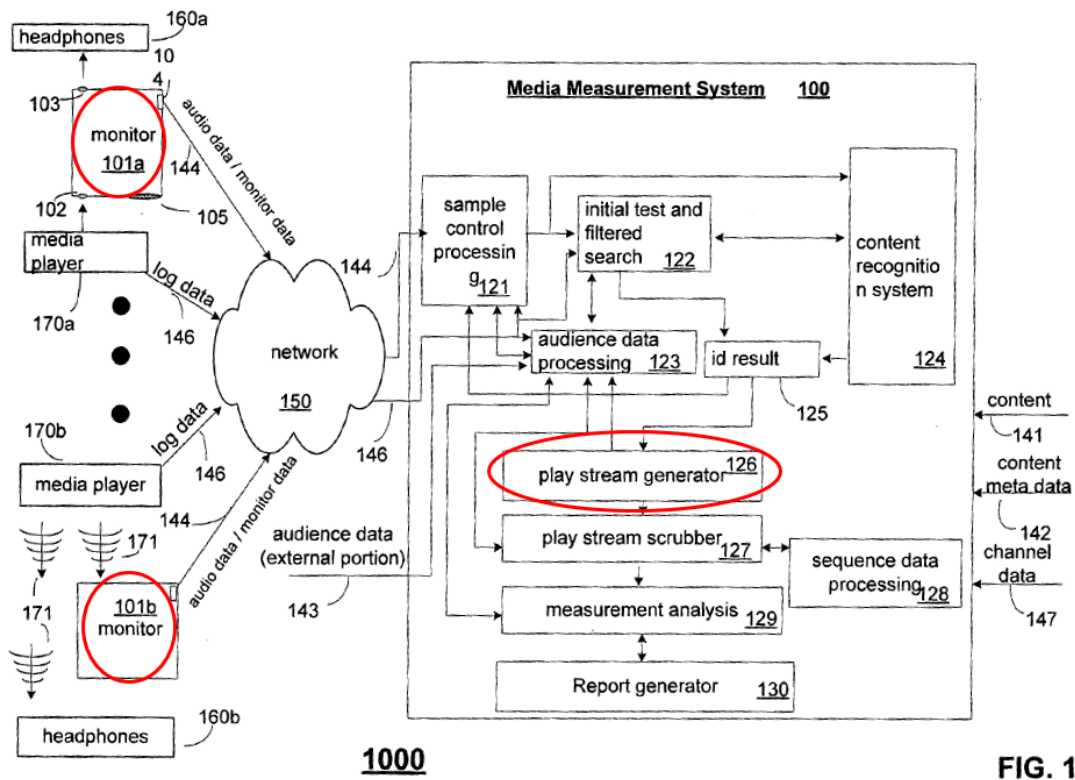


FIG. 1

(EX1005 at FIG. 1 (annotated); EX1006 at FIG. 1.)

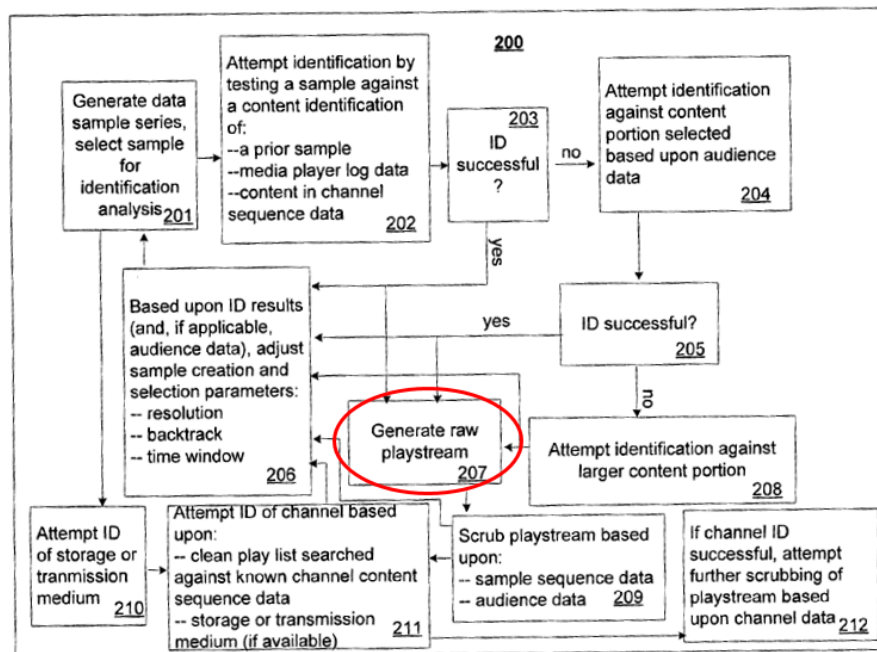


FIG. 2

(EX1005 at FIG. 2 (annotated); EX1006 at FIG. 2.)

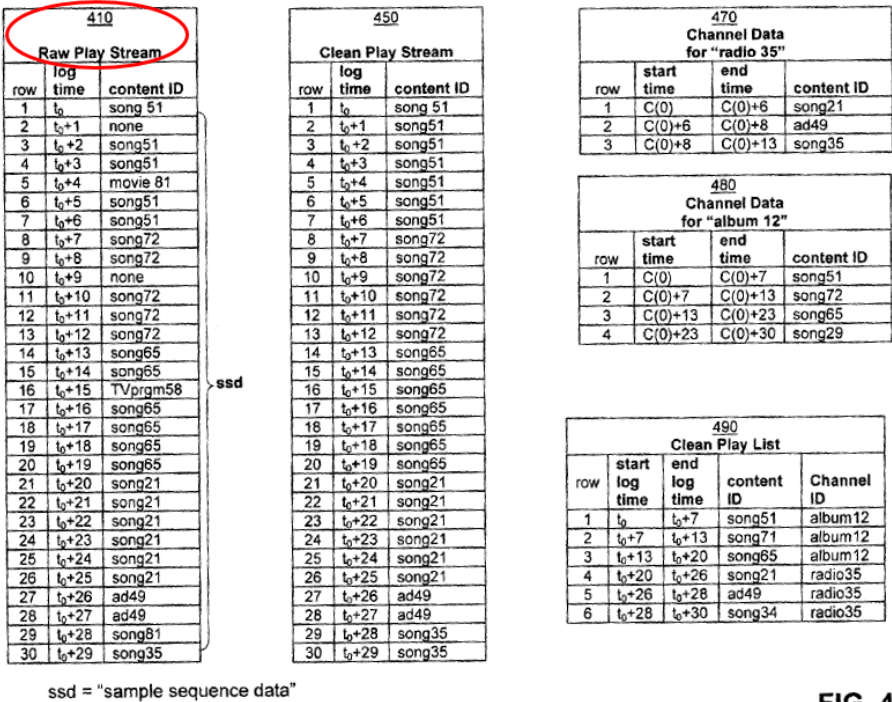


FIG. 4

(EX1005 at FIG. 4 (annotated); EX1006 at FIG. 4.)

125. Steuer768's claim 1 recites almost identical language as claim limitation 1.d, albeit in the context of audio data samples as noted with respect to limitation 1.b: "a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database." (EX1006 at claim 1.)

126. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (*See* Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 1.d.2.

- (3) **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;**¹¹

127. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the

¹¹ Claim limitation 1.d.3 recites alternative language in the form of the raw play stream including “**either:** [i] at least two different content identifiers . . .; **or** [ii] at least one content identifier . . . and at least one indication of the absence of a

same reasons as those given above with respect to limitations 1.c, 1.d, and 1.d.2.
(See Sections IX.A.1.c-d, IX.A.1.d.2, above.)

128. For example, Steuer750 and Steuer768 both disclose that raw play stream 410 includes either at least two different content identifiers obtained from the electronic database (e.g., “song51” and “movie81” in content ID column of raw play stream 410) or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database (e.g., “song51” and “none” in content ID column of raw play stream 410). (See Sections IX.A.1.c-d, IX.A.1.d.2, above (citing EX1005 at [0036], [0045]; EX1006 at 7:17-38, 9:12-37).)

respective likely match.” While the claim limitation is met by the prior art showing at least one of these alternatives, I map the prior art to both alternatives below.

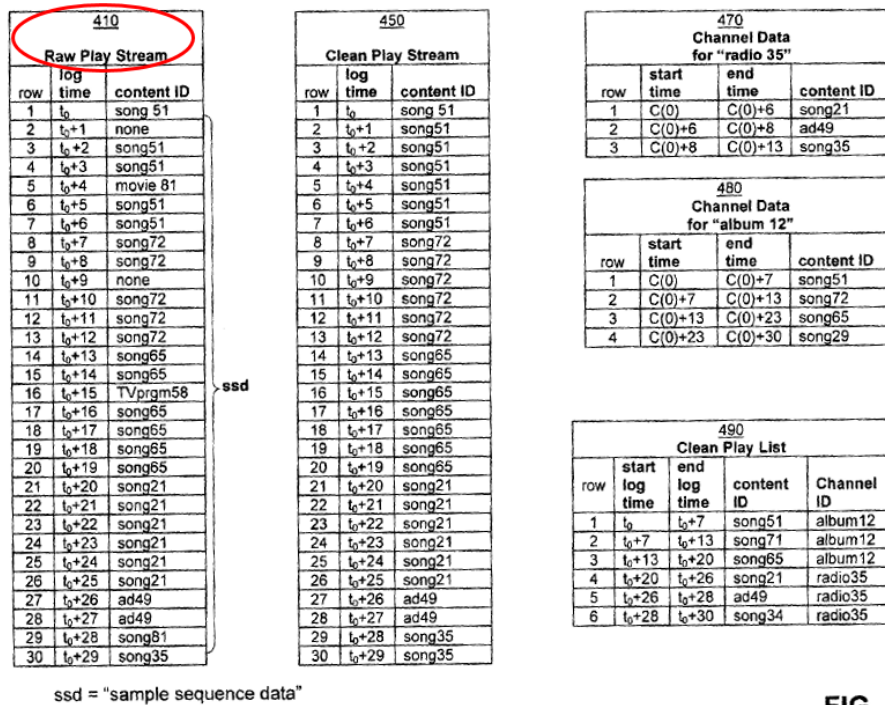


FIG. 4

(EX1005 at FIG. 4 (annotated); EX1006 at FIG. 4.)

129. Steuer768's claim 1 recites almost identical language as claim limitation 1.d, albeit in the context of audio data samples as noted with respect to limitation 1.b: "the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database." (EX1006 at claim 1.)

130. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have

been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (*See* Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 1.e.

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

131. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation.

132. Similar to claim limitation 1.e, Steuer750 discloses and claims “scrubbing the raw play stream by, with respect to samples that are not identified or appear to be misidentified, analyzing sample sequence data to attempt selection of corrected content identifications.” (EX1005 at claim 35.)

133. Steuer768's claim 1 recites almost identical language as claim limitation 1.e, albeit in the context of audio data samples as noted with respect to limitation 1.b: “scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of

the raw play stream compared to an expected pattern of sample sequence data.”
(EX1006 at claim 1.)

134. Further, Steuer750 and Steuer768 disclose scrubbing the raw play stream (e.g., by play stream scrubber 127) by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data. (EX1005 at [0031], [0040] (step 209), [0042] (steps 301 and 302), [0044] (step 307), [0045] (discussing clean play stream 450 and clean play list 490 of figure 4), [0046]-[0049], FIGS. 1-4, claims 7, 35, 39; EX1006 at 6:14-27, 8:1-22 (step 209), 8:30-49 (steps 301 and 302), 9:4-11 (step 307), 9:12-31 (discussing clean play stream 450 and clean play list 490 of figure 4), 9:32-10:27, FIGS. 1-4, claims 1, 9

Play stream generator 126 generates a raw play stream relating a sequence of samples to content identification results (for an example of a portion of a raw play stream, see FIG. 3 and accompanying text). **The raw play stream is then processed by play stream scrubber 127 which analyzes the sequence of sample identification results utilizing sequence data processing module 128. Based on analysis of sequence data and, where appropriate, analysis of audience data via interaction with audience data module 123, play stream scrubber 127 generates a clean play stream and clean play**

list (see FIG. 3 and accompanying text) which can then be used by measurement analysis module 129. Results of the media measurement analysis can be obtained via report generator 130.

(EX1005 at [0031], FIG. 1; EX1006 at 6:14-27, FIG. 1.)

Referring to process 300, depending on available information, either a play stream is made available to step 301 or a play list is made available to step 302. **Step 301 prepares a play stream for further processing necessary for channel matching by scrubbing (i.e. reviewing for missing or incorrect data) the play stream using non-channel data. Step 302 prepares a play list for further processing necessary for channel matching by scrubbing (i.e. reviewing for missing or incorrect data) the play list using non-channel data.**

(EX1005 at [0042], FIG. 1; EX1006 at 8:30-49, FIG. 1.)

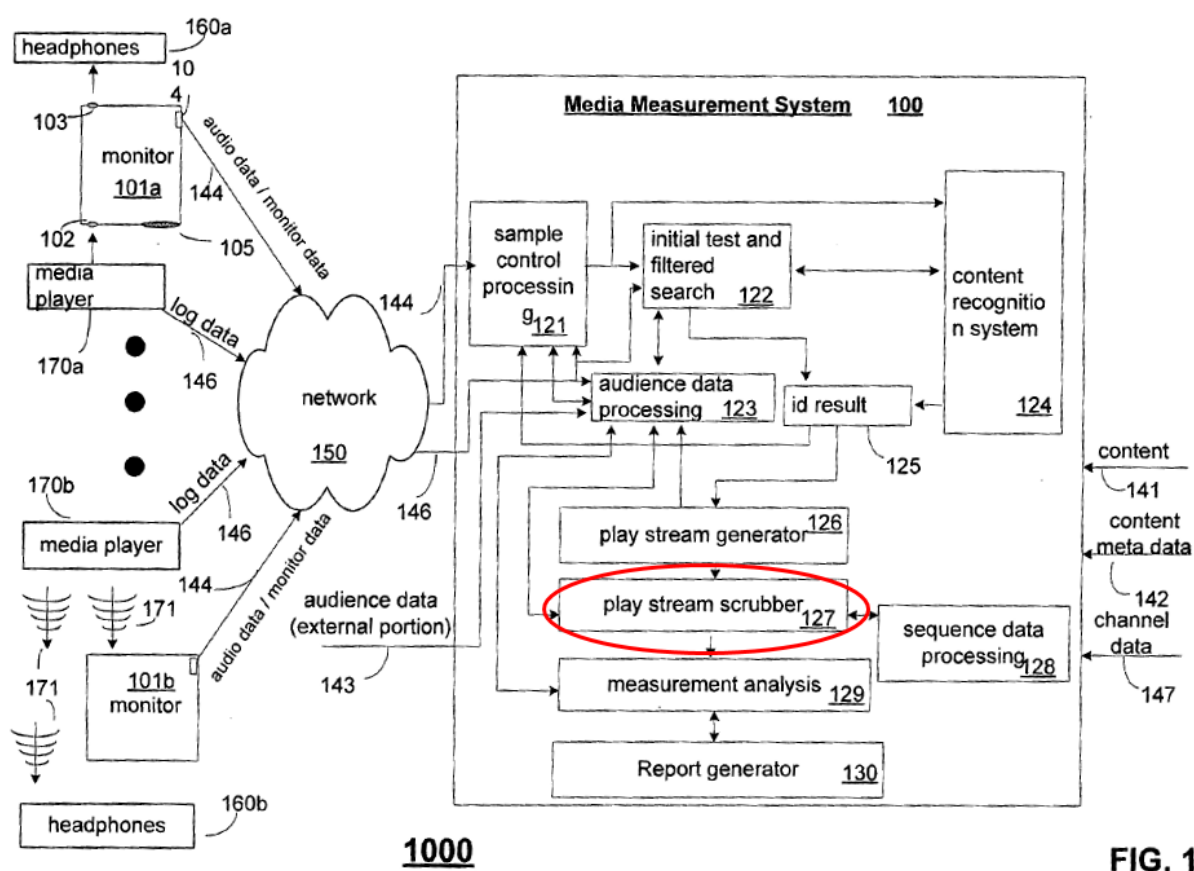
FIG. 4 illustrates Raw Play Stream 410 generated by a process such as step 207 of FIG. 2; **Clean Play Stream 450 and Clean Play List 490 generated by a scrubbing step such as step 209 of FIG. 2 or steps 301, 302, or 305 of FIG. 3;** Channel Data 470 and Channel Data 490 . . .

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 identification.

Sample sequence data can be used to derive portions of Clean Play Stream 450 from raw play list 410 in the following manner: Referring to Raw Play Stream 410, the content IDs associated with samples 1, 3, 4, 6, and 7 all indicate the same media content referenced as “song51.” The content ID value associate with sample 2 indicates that no content has yet been successfully associated with that sample and the content ID value associated with sample 5 indicates “movie81.” 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.

(EX1005 at [0045]-[0047], FIG. 4; EX1006 at 9:12-51, FIG. 4.)



(EX1005 at FIG. 1 (annotated); EX1006 at FIG. 1.)

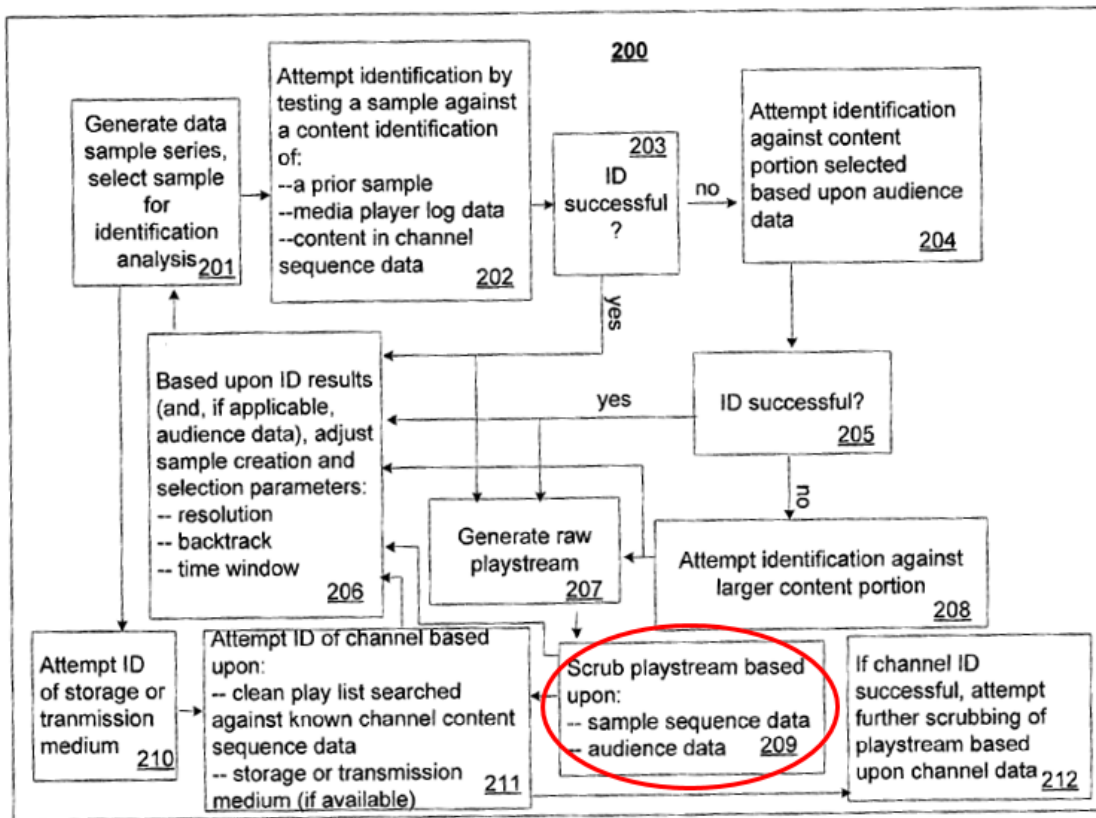


FIG. 2

(EX1005 at FIG. 2 (annotated); EX1006 at FIG. 2.)

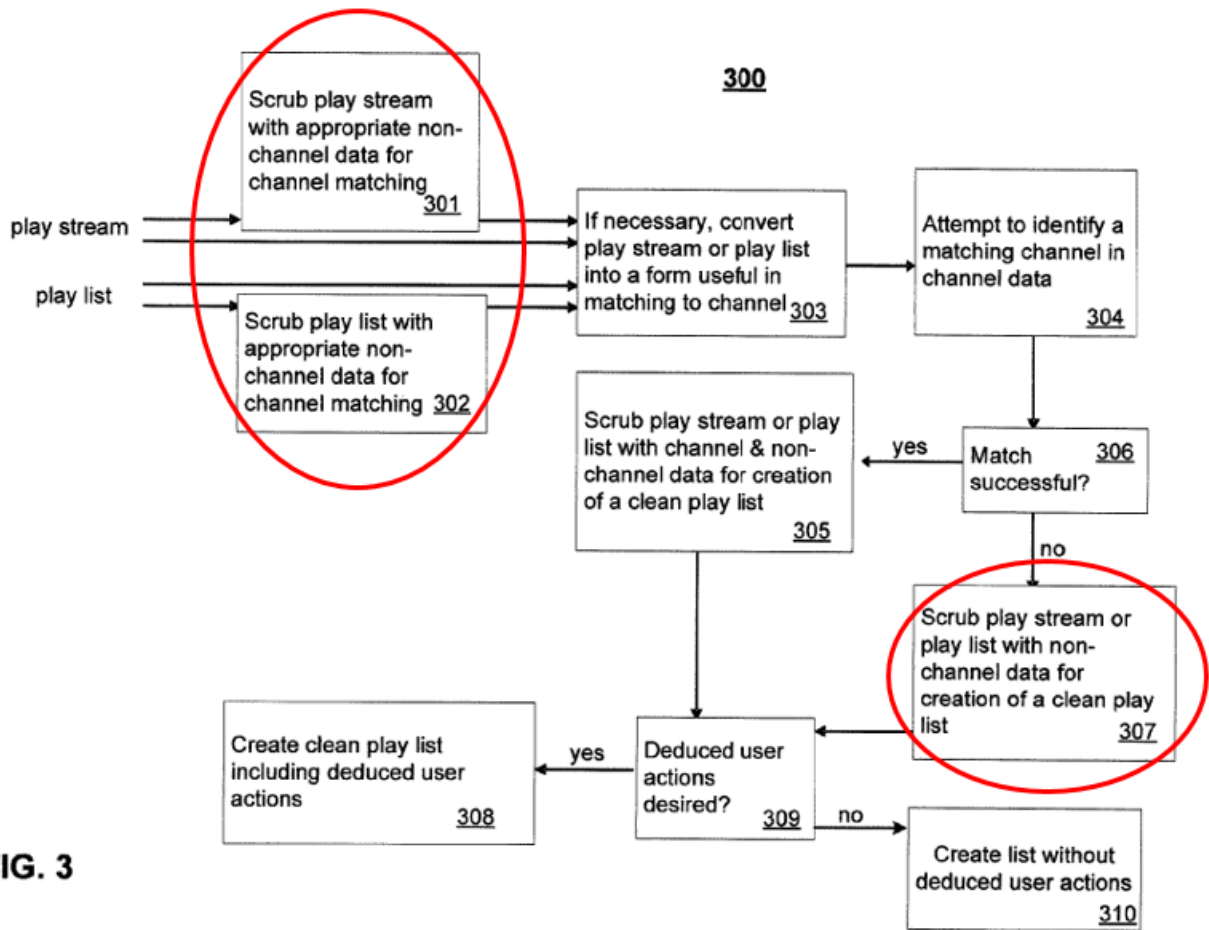


FIG. 3

(EX1005 at FIG. 3 (annotated); EX1006 at FIG. 3.)

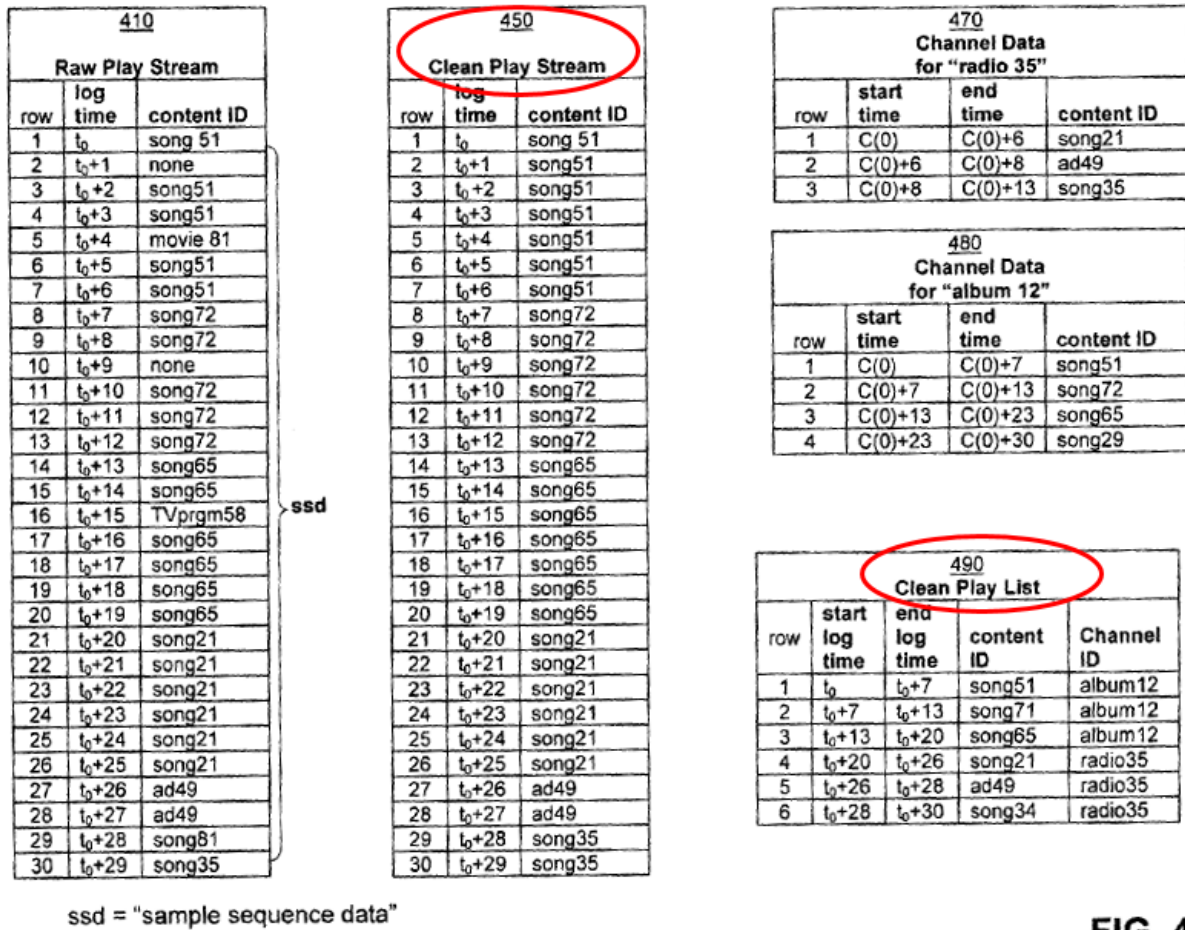


FIG. 4

(EX1005 at FIG. 4 (annotated); EX1006 at FIG. 4.)

135. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier

combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 1.e.

- f) (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.**

136. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggests this limitation for at least reasons similar to those given above with respect to claim limitation 1.e.

137. First, Steuer768's claim 1 recites almost identical language as claim limitation 1.c, albeit in the context of audio data samples as noted with respect to limitation 1.b: "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." (EX1006 at claim 1.)

138. Moreover, as I discussed above, Steuer750 and Steuer768 disclose generating a clean play stream (e.g., clean play stream 450 in figure 4) by making any changes to the raw play stream (e.g., raw play stream 410) that are determined to be made by the scrubbing. (*See* Section IX.A.1.e, above (citing EX1005 at [0031], [0040] (step 209), [0042] (steps 301 and 302), [0044] (step 307), [0045] (discussing clean play stream 450 of figure 4), [0046]-[0049], FIGS. 1-4, claims 7, 35, 39; EX1006 at 6:14-27, 8:1-22 (step 209), 8:30-49 (steps 301 and 302), 9:4-11 (step

307), 9:12-31 (discussing clean play stream 450 of figure 4), 9:32-10:27, FIGS. 1-4, claims 1, 9).)

Sample sequence data can be used to derive portions of Clean Play Stream 450 from raw play list 410 in the following manner: Referring to Raw Play Stream 410, the content IDs associated with samples 1, 3, 4, 6, and 7 all indicate the same media content referenced as “song51.” The content ID value associate with sample 2 indicates that no content has yet been successfully associated with that sample and the content ID value associated with sample 5 indicates “movie81.” **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.**

(EX1005 at [0047], FIG. 4; EX1006 at 9:38-51, FIG. 4.)

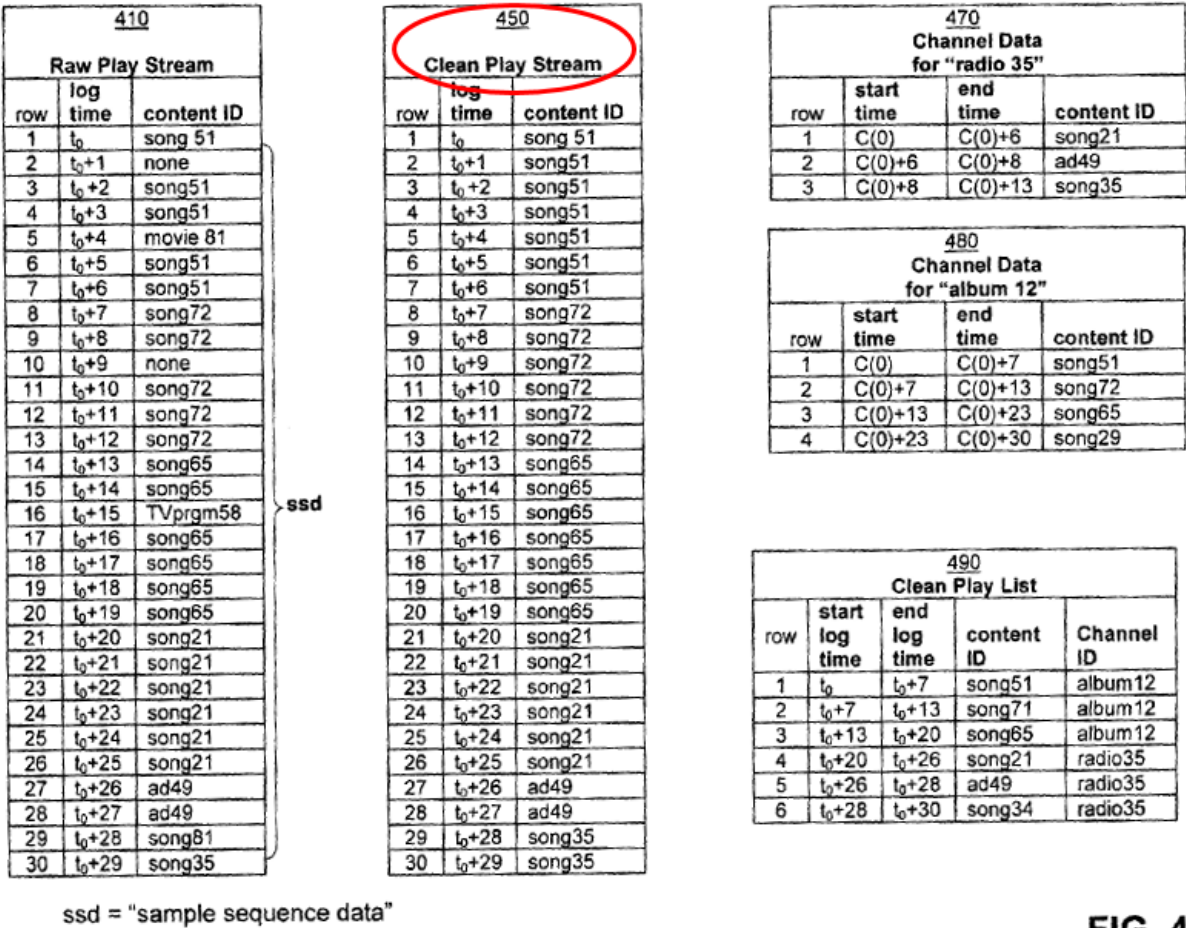


FIG. 4

(EX1005 at FIG. 4 (annotated); EX1006 at FIG. 4.)

139. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier

combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 1.f.

2. Claim 2

- a) The method of claim 1 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.**

140. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation.

141. First, Steuer768's claim 2 recites almost identical language as claim 2, albeit in the context of audio data samples as noted with respect to limitation 1.b: "wherein sample sequence data includes at least one of log times, content identifiers, and content offsets." (EX1006 at claim 2.)

142. Moreover, Steuer750 and Steuer768 disclose sample sequence data includes at least one of log times, content identifiers, and content offsets as they disclose:

However, note that in other examples, such as that of FIG. 5, **sample sequence data ("ssd") includes content offset data.** For purposes of presenting and analyzing sample sequence data as described herein, the times in list 410 may be relative to an arbitrary value (e.g. "t0") rather than representing an absolute time; however, in many instances, **a log time** will in fact indicate the actual time the sample started to be captured.

(EX1005 at [0045], FIGS. 1-5; EX1006 at 9:12-37, FIGS. 1-5.)

Sample sequence data can be used to derive portions of Clean Play Stream 450 from raw play list 410 in the following manner: Referring to Raw Play Stream 410, **the content IDs** associated with samples 1, 3, 4, 6, and 7 all indicate the same media content referenced as “song51.” The content ID value associate with sample 2 indicates that no content has yet been successfully associated with that sample and the content ID value associated with sample 5 indicates “movie81.” **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.**

(EX1005 at [0047], FIG. 4; EX1006 at 9:38-51, FIG. 4.)

Continuing with the description of FIG. 5, **Play Stream 510 presents a series of log times, content IDs, and content offsets.** In the illustrated example, times given in the log time column begins with t_0 and ends with t_0+19 . Each log time indicates a time at which capture of the analyzed sample began, where t_0 simply represents a beginning time for samples on Play Stream 510. Each content ID indicates an identifier of the content that has been identified as corresponding to the captured sample. In the example illustrated, “TV program 27” refers to a TV program segment and “TV program 28” refers to another program segment of the same program. “Ad23” and “ad25,” on

the other hand, refer to separate advertisements. Each content offset value represents a time offset from the beginning of a particular piece of content. For example, row 8 of Play Stream 510 corresponds to a sample whose beginning is located 9 time units from the beginning of TV program 27, thus the content offset is referenced as (0)+9 . . .

Discontinuities in consumption of the identified media may be identified by comparing the progression of sample log times to progression of corresponding content offset times in Play Stream 510. For example, referring to rows 7 and 8 of play stream 510, log times progress from t_0+6 to t_0+7 ; however, the corresponding content offsets progress, with respect to the start of TV program 27, from (0)+3 to (0)+9. Had the media consumption from one sample to the next been continuous, one would expect that the second offset value (in row 8) would have been (0)+4 rather than (0)+9. Therefore, a boundary between media consumption events can be deduced. Thus, rows 3 and 4 of List 560 indicate two consumption events with respect to the same piece of media content (TV program 27). By contrast, row 1 and row 2 of Play Stream 510 are, on Play List 560, collapsed into a single row (row 1) because the content offset progression from row 1 to row 2 of Play Stream 510 suggests continuous consumption of the same piece of media content.

(EX1005 at [0053]-[0055]; EX1006 at 11:10-55.)

510 Clean Playstream			
row	log time	content ID	content offset
1	t_0	TVprgm 27	(0)
2	t_0+1	TVprgm 27	(0)+1
3	t_0+2	none	n/a
4	t_0+3	none	n/a
5	t_0+4	none	n/a
6	t_0+5	TVprgm 27	(0)+2
7	t_0+6	TVprgm 27	(0)+3
8	t_0+7	TVprgm27	(0)+9
9	t_0+8	TVprgm27	(0)+10
10	t_0+9	TVprgm27	(0)+9
11	t_0+10	TVprgm27	(0)+10
12	t_0+11	none	n/a
13	t_0+12	TVprgrm27	(0)+29
14	t_0+13	TVprgrm27	(0)+30
15	t_0+14	ad23	(0)
16	t_0+15	none	n/a
17	t_0+16	ad25	(0)+2
18	t_0+17	TVprgm28	(0)
19	t_0+18	TVprgm28	(0)+1
20	t_0+19	TVprgm28	(0)+2

560 Clean Play List				
row	start log time	end log time	content ID	content offset
1	t_0	t_0+2	TVprgrm27	(0)
2	t_0+2	t_0+5	none	n/a
3	t_0+5	t_0+7	TVprgrm27	(0)+2
4	t_0+7	t_0+9	TVprgrm27	(0)+9
5	t_0+9	t_0+11	TVprgrm27	(0)+9
6	t_0+11	t_0+12	none	n/a
7	t_0+12	t_0+14	TVprgrm27	(0)+29
8	t_0+14	t_0+15	ad23	(0)
9	t_0+15	t_0+16	none	n/a
10	t_0+16	t_0+17	ad25	(0)+2
11	t_0+17	t_0+20	TVprgrm28	(0)

570 Channel Data for TV C4			
row	start time	end time	content ID
1	C(0)	C(0)+31	TVprgrm27
2	C(0)+31	C(0)+34	ad23
3	C(0)+34	C(0)+37	ad24
4	C(0)+37	C(0)+40	ad25
5	C(0)+40	C(0)+50	TVprgrm28

590 Clean Play List with Deduced Actions				
row	log time	end log time	content ID and <deduced action>	content offset
1	t_0	t_0+2	TVprgrm27	(0)
2	t_0+2	t_0+5	<pause for 3 units>	n/a
3	t_0+5	t_0+7	TVprgrm27	(0)+2
4	t_0+7	t_0+7	<skip frwrd 5 units>	n/a
5	t_0+7	t_0+9	TVprgrm27	(0)+9
6	t_0+9	t_0+9	<skip bckwrd 2 units>	n/a
7	t_0+9	t_0+11	TVprgrm27	(0)+9
8	t_0+11	t_0+12	<fast frwrd 18 units>	n/a
9	t_0+12	t_0+14	TVprgrm27	(0)+29
10	t_0+14	t_0+15	ad23	(0)
11	t_0+15	t_0+16	<fast frwrd 7 units>	n/a
12	t_0+16	t_0+17	ad25	(0)+2
13	t_0+17	t_0+20	TVprgrm28	(0)

FIG. 5

(EX1005 at FIG. 5 (annotated); EX1006 at FIG. 5.)

143. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 2.

3. Claim 3

- a) **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.**

144. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation.

145. As I discussed above, Steuer750 and Steuer768 disclose scrubbing the raw play stream by analyzing sample sequence data of the raw play stream. (*See* Section IX.A.1.e, above.) Steuer750 and Steuer768 further disclose that 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.

If step 211 successfully identifies the channel, information such as **content sequence data for the identified channel may be used by step 212 to further scrub or rescrub the play stream**, particularly with respect to any apparently incorrect content identifications that could not be corrected based upon sample sequence data as utilized in step 209.

(EX1005 at [0040]; EX1006 at 8:1-22.)

If yes, **then step 305 uses channel data and non-channel data to scrub or further scrub the play stream or play list for purposes of creating a clean play list**. If the play list or play stream has already been scrubbed with non-channel data in either step 301 and 302, then the play stream or play list may not need

to be further scrubbed with non-channel data at step 305. In that case, step 305 just uses channel data to further scrub the play stream or play list.

(EX1005 at [0043]; EX1006 at 8:58-65.)

Information 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 . . .**

(EX1005 at [0048]; EX1006 at 9:52-10:4.)

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 . . .

(EX1005 at [0049]; *see also id.*, claim 37; EX1006 at 10:5-27.)

146. Further, Steuer768’s claim 3 recites almost identical language as claim 3, albeit in the context of audio data samples as noted with respect to limitation 1.b: “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.” (EX1006 at claim 3.)

147. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750’s and Steuer768’s disclosed content

identification systems using video data samples instead of its disclosed audio-based approach. (*See* Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 3.

4. Claim 4

- a) 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.**

148. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation.

149. As I discussed above, Steuer750 and Steuer768 disclose scrubbing the raw play stream by analyzing sample sequence data of the raw play stream. (*See* Section IX.A.1.e, above.) Steuer750 and Steuer768 further disclose that 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.

Based on analysis of sequence data and, where appropriate, **analysis of audience data** via interaction with audience data module 123, **play stream scrubber 127 generates a clean play stream and clean play list** (see FIG. 3 and accompanying text) which can then be used by measurement analysis module 129.

(EX1005 at [0031]; EX1006 at 6:13-27.)

. . . **Step 209 scrubs the raw play stream** based upon sample sequence data (described further below in the context of FIGS. 3-4 and accompanying text) and **based upon audience data** . . .

(EX1005 at [0040]; EX1006 at 8:1-22.)

Audience data may also be used in the “scrubbing” process.

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.

(EX1005 at [0050]; *see also id.*, claim 38; EX1006 at 10:28-33.)

150. Further, Steuer768’s claim 4 recites almost identical language as claim 4, albeit in the context of audio data samples as noted with respect to limitation 1.b: “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.” (EX1006 at claim 4.)

151. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750’s and Steuer768’s disclosed content identification systems using video data samples instead of its disclosed audio-based

approach. (*See* Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 4.

5. Claim 5

a) The method of claim 1 wherein a clean play list is generated from the clean play stream.

152. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation.

153. For example, similar to claim 5, Steuer750 discloses and claims that “a clean play list is generated from the scrubbed play stream.” (EX1005 at claim 39.)

154. Steuer768’s claim 5 recites almost identical language as claim 5, albeit in the context of audio data samples as noted with respect to limitation 1.b: “wherein a clean play list is generated from the clean play stream.” (EX1006 at claim 5.)

155. Steuer750 and Steuer768 disclose additional details regarding generation of a clean play list from the clean play stream:

Continuing with the description of FIG. 2, step 207 generates a raw play stream indicating a series of samples and corresponding identification results. Step 209 scrubs the raw play stream based upon sample sequence data (described further below in the context of FIGS. 3-4 and accompanying text) and based upon audience data. **The resulting clean play stream may be converted to a play list listing the content captured and indicating the order in which it was captured.** Step 209

attempts to identify the channel by searching known channel data to find an apparent content sequence match to the play list. If the storage or transmission medium of the captured sample has been identified at step 210, that information may be made available for step 211 and allows step 211 to use that information to filter its search of channel data (e.g., only search channel data including CD channels, or only including radio broadcast channels, etc.). If step 211 successfully identifies the channel, information such as content sequence data for the identified channel may be used by step 212 to further scrub or rescrub the play stream, particularly with respect to any apparently incorrect content identifications that could not be corrected based upon sample sequence data as utilized in step 209.

(EX1005 at [0040]; EX1006 at 8:1-22.)

Referring to process 300, depending on available information, either a play stream is made available to step 301 or a play list is made available to step 302. Step 301 prepares a play stream for further processing necessary for channel matching by scrubbing (i.e. reviewing for missing or incorrect data) the play stream using non-channel data. Step 302 prepares a play list for further processing necessary for channel matching by scrubbing (i.e. reviewing for missing or incorrect data) the play list using non-channel data. With respect to steps 301 and 302, the non-channel data selected might be selected particularly for the purpose of better preparing the play stream or play list for a channel matching process. Step 303 receives either a raw (i.e. “not

scrubbed”) play stream or play list directly or receives a scrubbed play stream or play list from step 301 or 302. **If necessary, step 303 converts received data to a format useful in matching to a channel (e.g. converts a play stream to a play list where a play list format is useful for channel matching,** or converts a play stream or play list to a mathematical representation that is useful for channel matching).

(EX1005 at [0042]; EX1006 at 8:30-49.)

Step 304 attempts to match the play stream or play list (or corresponding data format representing the play stream or play list) against known channel data to identify a channel (e.g., an album, samples of an album presented for marketing—e.g. on a web page—, radio broadcast, television broadcast, theater version of a movie, DVD version of a movie, etc.) associated with the elements of the play list or play stream. Step 306 determines whether a channel match was achieved. **If yes, then step 305 uses channel data and non-channel data to scrub or further scrub the play stream or play list for purposes of creating a clean play list.** If the play list or play stream has already been scrubbed with non-channel data in either step 301 and 302, then the play stream or play list may not need to be further scrubbed with non-channel data at step 305. In that case, step 305 just uses channel data to further scrub the play stream or play list. However, it is possible that **scrubbing the play list or play stream** with non-channel data for purposes of creating data useful for channel matching will be somewhat different than

doing so **for purposes of creating a clean play list**, thus non-channel data is referenced again in the illustration of claim 305.

(EX1005 at [0043]; EX1006 at 8:50-9:3.)

If step 306 determines that a channel match was not achieved at step 304, **then step 307 scrubs the play stream or play list with non-channel data to create a clean play list**. Step 309 determines whether including deduced user actions on the clean play list is desired. If yes, then step 308 creates a clean play list including deduced user actions. If no, then step 310 creates the clean play list without including deduced user actions.

(EX1005 at [0044]; EX1006 at 9:4-11.)

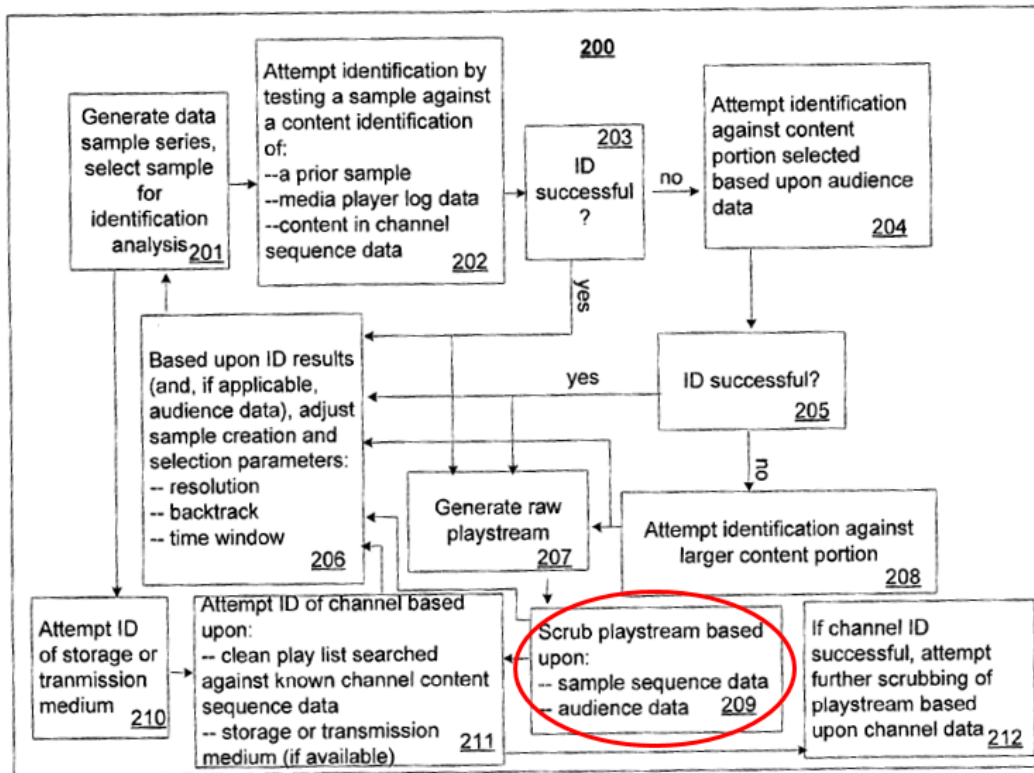
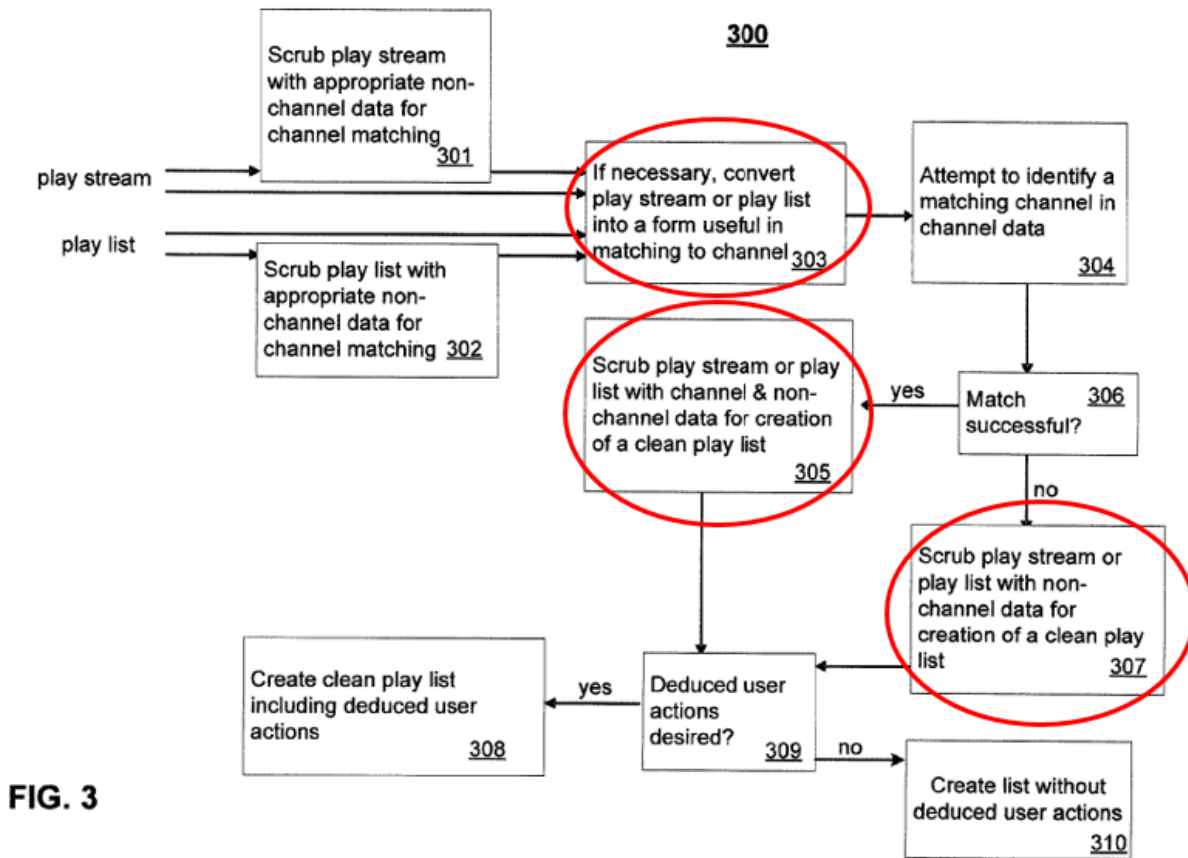


FIG. 2

(EX1005 at FIG. 2 (annotated); EX1006 at FIG. 2.)



(EX1005 at FIG. 3 (annotated); EX1006 at FIG. 3.)

156. Steuer750 and Steuer768 disclose these features in the context of its audio-based content identification systems and methods, but as I discussed above with reference to claim limitation 1.b, a person of ordinary skill in the art would have been motivated to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 5.

6. Claim 6

- a) The media measurement method of claim 1 wherein the plurality of video data representations include a plurality of video signatures.**

157. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation.

158. For example, as I discussed above, Steuer750 and Steuer768 disclose querying a database of known content 141 of a plurality of audio data representations (e.g., audio signatures). (*See* Section IX.A.1.c, above (citing EX1005 at [0022]; EX1006 at 4:13-48.)

159. Steuer768's claim 6 recites almost identical language as claim 6, albeit in the context of audio data representations: "wherein the plurality of audio data representations include a plurality of audio signatures." (EX1006 at claim 6.)

160. As I discussed above for claim limitation 1.b, Neumeier discloses that "the video segment is identified by sampling a subset of the pixel data being displayed on the screen (or associated audio data) and then finding similar pixel (or audio) data in a content database." (*See* Section IX.A.1.b, above; EX1007 at [0011]; *see also id.*, [0054]-[0057], [0059]-[0062].) Neumeier further discloses using audio signatures for content recognition. (*Id.*, [0025], [0040], [0045].) Using video signatures for content recognition was also known at the time of the alleged invention. (EX1016 at Abstract, [0015] (describing derivation of content signature),

[0045] (describing content signature calculation for video).) Thus, in my opinion, for the reasons I discussed above for claim limitation 1.b, a person of ordinary skill in the art would have been motivated and would have known how to implement Steuer750's and Steuer768's disclosed content identification systems using video data samples (and thus matching the video data samples with video data representations including a plurality of video data signatures) instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 6.

7. Claim 7

- a) 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.**

161. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation.

162. For example, Steuer750 and Steuer768 disclose querying a database of known content 141 of a plurality of audio data representations (e.g., signal landmarks). (EX1005 at [0022] ("However, it will be understood by those skilled in the art that "content" in the sense of the known content 141 received by system 100 may, in some embodiments, simply refer to data about **signals representing the**

content and does not necessarily refer to a stored version of the media content itself; e.g. “content” may be stored signatures and **landmarks** derivable from a song's audio, but not necessarily an audio file useable to play the song itself . . .”); EX1006 at 4:13-48.)

163. Steuer768’s claim 7 recites almost identical language as claim 7, albeit in the context of audio data representations: “wherein an audio data representation of the plurality of audio data representations includes a plurality of signal landmarks.” (EX1006 at claim 7.)

164. As I discussed above for claim limitation 1.b, Neumeier discloses that “the video segment is identified by sampling a subset of the pixel data being displayed on the screen (or associated audio data) and then finding similar pixel (or audio) data in a content database.” (*See* Section IX.A.1.b, above; EX1007 at [0011]; *see also id.*, [0054]-[0057], [0059]-[0062].) Neumeier further discloses using audio signatures for content recognition. (*Id.*, [0025], [0040], [0045].) Using video signatures for content recognition was also known at the time of the alleged invention. (EX1016 at Abstract, [0015] (describing derivation of content signature), [0040] (describing combination of watermarks and content signatures), [0045] (describing content signature calculation for video).) Thus, in my opinion, for the reasons I discussed above for claim limitation 1.b, a person of ordinary skill in the art would have been motivated and would have known how to implement

Steuer750's and Steuer768's disclosed content identification systems using video data samples (and thus matching the video data samples with video data representations including a plurality of signal landmarks) instead of its disclosed audio-based approach. (See Section IX.A.1.b, above.) Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 7.

8. Claim 8

- a) 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.**

165. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least reasons similar to those given above with respect to claim limitations 1.b-1.d.1. (See Sections IX.A.1.b-IX.A.1.d.1, above.)

166. As I discussed above for claim limitation 1.b, Neumeier discloses that “the video segment is identified by sampling a subset of the pixel data being displayed on the screen (or associated audio data) and then **finding similar pixel (or audio) data in a content database.**” (See Section IX.A.1.b, above; EX1007 at [0011]; *see also id.*, [0054]-[0057], [0059]-[0062].) Neumeier specifically discloses sampling pixel data at the client side and searching for corresponding pixel data in a

database on the server side. (EX1007 at [0054]-[0057], [0059]-[0062].) The sampled pixel information is a signal usable for generating video energy corresponding to video energy captured at the media monitoring device because the captured pixel values at the client side represent, e.g., the light intensity at a given location in the video frame. (*See* EX1016 at [0045].)¹² Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 8.

¹² Moreover, Steuer768's claim 8 recites almost identical language as claim 8, albeit in the context of audio data representations: "wherein an audio data representation of the plurality of audio data representations includes a signal usable for generating audio energy corresponding to audio energy captured at the media monitoring device." (EX1006 at claim 8.) Steuer750 discloses that the uploaded audio data may represent a raw signal corresponding to audio energy received at monitor 101. (EX1005, [0019].) Thus, in my opinion, a person of ordinary skill in the art would have been motivated to implement the content identification techniques of Steuer750 and Steuer768 based on audio data representations in the context claimed for the same reasons discussed above with respect to claim limitation 1.b.

9. Claim 9

- a) **[pre] A computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers, the one or more computers being coupled to a network to receive, over the network, data for a sequence of 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:**

167. I have been asked to assume that the preamble of claim 9 is limiting. Under that assumption, it is my opinion that the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests this limitation.

168. For example, Steuer750 and Steuer768 disclose with reference to figure 7 “an example of a computer system that may be used to execute instruction code contained in a computer program product, the computer program product being in accordance with an embodiment of the present invention.” (EX1005 at [0014], FIGS. 1-2, 7; EX1006 at 2:47-50, FIGS. 1-2, 7.) Steuer750 and Steuer768 further disclose:

FIG. 7 shows an example of a computer system 700 that may be used to execute instruction code contained in a computer program product 760 in accordance with an embodiment of the present invention. Computer program product 760 comprises executable code in an electronically readable medium that may instruct one or more computers such as computer system 700 to perform processing that implements

the system 100 of FIG. 1 and/or accomplishes the exemplary method 200 of FIG. 2. The electronically readable medium may be any medium that either stores or carries electronic signals (including signals referred to as electrical signals and signals referred to as electromagnetic signals) and may be accessed locally or remotely, for example via a network connection. **The executable instruction code in an electronically readable medium directs the illustrated computer system 700 to carry out various exemplary tasks described herein.** The executable code for directing the carrying out of tasks described herein would be typically realized in software. However, it will be appreciated by those skilled in the art, that computers might utilize code realized in hardware to perform many or all of the identified tasks without departing from the present invention. Those skilled in the art will understand that many variations on executable code may be found that implement exemplary methods within the spirit and the scope of the present invention.

The code or a copy of the code contained in computer program product 760 may be stored in memory 710 for execution by processor 720. Computer system 700 also includes I/O subsystem 730 and peripheral devices 740. I/O subsystem 730, peripheral devices 740, processor 720, and memory 710 are coupled via bus 750.

Those skilled in the art will appreciate computer system 700 illustrates just one example of a system in which a computer

program product in accordance with an embodiment of the present invention may be implemented. To cite but one example of an alternative embodiment, execution of instructions contained in a computer program product in accordance with an embodiment of the present invention may be distributed over multiple computers, such as, for example, over the computers of a distributed computing network.

(EX1005 at [0066]-[0067]; *see also id.*, [0068], FIGS. 1-2, 7; EX1006 at 13:54-14:25.)

169. Moreover, Steuer768's claim 9 includes an identical preamble: "A computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers, the one or more computers being coupled to a network to receive, over the network, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network, the instructions comprising instructions for: . . ." (EX1006 at claim 9.)

170. Similarly, Steuer750 includes claims directed to "[a] computer program product for measuring media usage, the computer program product comprising executable instruction code in an electronically readable medium for at least: . . ." (EX1006 at claims 54-60.)

171. Moreover, as I discussed above with respect to limitation 1.b, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests that 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. (*See* Section IX.A.1.b, above.)

172. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest the claim 9 preamble. (*See also* Sections IX.A.9.b–e, below; Section IX.A.1.a, above.)

b) (a) using the one or more computers to query an electronic database of a plurality of video data representations and corresponding content identifiers;

173. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons I discussed above for claim limitation 1.c. (*See* Section IX.A.1.c, above; *see also* EX1006 at claim 9 (reciting almost identical language albeit in the context of “audio data representations”).)

174. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 9.b.

- c) (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;**

175. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons I discussed above for claim limitation 1.d. (*See* Section IX.A.1.d, above; *see also* EX1006 at claim 9 (reciting almost identical language albeit in the context of “audio data samples”).)

176. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 9.c.

- (1) 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;**

177. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons I discussed above for claim limitation 1.d.1. (*See* Section IX.A.1.d.1, above; *see also* EX1006 at claim 9 (reciting almost identical language albeit in the context of “audio data samples”).)

178. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 9.c.1.

- (2) 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**

179. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons I discussed above for claim limitation 1.d.2. (*See* Section IX.A.1.d.2, above; *see also* EX1006 at claim 9 (reciting almost identical language albeit in the context of “audio data samples”).)

180. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 9.c.2.

- (3) 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;**

181. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons I discussed above for claim limitation 1.d.3. (*See* Section IX.A.1.d.3, above; *see also* EX1006 at claim 9 (reciting almost identical language albeit in the context of “audio data samples”).)

182. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination discloses or suggests claim limitation 9.c.3.

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

183. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons I discussed above for claim limitation 1.e. (*See* Section IX.A.1.e, above; *see also* EX1006 at claim 9 (reciting almost identical language albeit in the context of “audio data samples”).)

184. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 9.d.

- e) **(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.**

185. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least the same reasons I discussed above for claim limitation 1.f. (*See* Section IX.A.1.f, above; *see also* EX1006 at claim 9 (reciting almost identical language albeit in the context of “audio data samples”).)

186. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 9.e.

10. Claim 10

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

187. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least reasons similar to those given above with respect to claim 2. (*See* Section IX.A.2, above; *see also* EX1006 at claim 10 (reciting almost identical language albeit in the context of “audio data samples”).)

188. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim 10.

11. Claim 11

- a) 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.**

189. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least reasons similar to those given above with respect to claim 3. (*See* Section IX.A.3, above; *see also* EX1006 at claim 11 (reciting almost identical language albeit in the context of “audio data samples”).)

190. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 11.

12. Claim 12

- a) 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.**

191. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least reasons similar to those given above with respect to claim 4. (*See* Section IX.A.4,

above; *see also* EX1006 at claim 12 (reciting almost identical language albeit in the context of “audio data samples”).)

192. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 12.

13. Claim 13

- a) 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.**

193. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least reasons similar to those given above with respect to claim 5. (*See* Section IX.A.5, above; *see also* EX1006 at claim 12 (reciting almost identical language albeit in the context of “audio data samples”).)

194. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 13.

14. Claim 14

- a) The computer program product of claim 9 wherein the plurality of video data representations include a plurality of video signatures.**

195. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least reasons similar to those given above with respect to claim 6. (*See* Section IX.A.6,

above; *see also* EX1006 at claim 14 (reciting almost identical language albeit in the context of “audio data representations”).)

196. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 14.

15. Claim 15

- a) 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.**

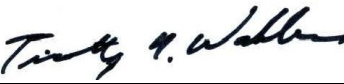
197. In my opinion, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest this limitation for at least reasons similar to those given above with respect to claim 7. (*See* Section IX.A.7, above; *see also* EX1006 at claim 15 (reciting almost identical language albeit in the context of “audio data representations”).)

198. Thus, the Steuer750-Neumeier combination and the Steuer768-Neumeier combination disclose or suggest claim limitation 15.

X. CONCLUSION

199. I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Dated: August 26, 2024

By: 
Timothy Wahlers