

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

FREE STREAM MEDIA CORP.
D/B/A SAMBA TV,
Petitioner

U.S. Patent No. 9,066,114

Case No. IPR2019-00216

DECLARATION OF AHMED H. TEWFIK, SC. D.

TABLE OF CONTENTS

| | Page |
|--|------|
| I. BACKGROUND AND QUALIFICATIONS | 5 |
| II. ASSIGNMENT AND MATERIALS CONSIDERED | 5 |
| III. THE '114 PATENT'S EFFECTIVE FILING DATE | 11 |
| IV. CLAIM INTERPRETATION | 11 |
| A. "A processor" | 11 |
| V. OVERVIEW OF THE '114 PATENT | 15 |
| A. The Problem to Be Solved | 15 |
| B. The '114 Patent's Purported Solution | 16 |
| C. Detailed Description..... | 17 |
| D. Prosecution History of the '114 Patent | 25 |
| VI. THE STATE OF THE ART | 26 |
| A. Level of Skill in the Art..... | 26 |
| B. General Knowledge Known to a Person of Ordinary Skill in the Art..... | 27 |
| 1. It was known to enhance media with time-specific actions and features..... | 27 |
| 2. It was known to trigger time-specific actions at precise times using content identification | 30 |
| 3. Fingerprinting was a well-known, beneficial substitute for watermarking to achieve known, predictable results of time-specific content identification | 34 |
| 4. Algorithms for generating, storing, retrieving, and matching fingerprints for content recognition were well known..... | 37 |
| 5. It was known to associate trigger markers with actions using a database | 41 |
| C. Murphy | 41 |
| D. Brunk | 50 |

TABLE OF CONTENTS
(continued)

| | Page |
|---|-------------|
| VII. UNPATENTABILITY ANALYSIS | 52 |
| A. Detailed Claim Mapping | 54 |
| 1. Claim 1 | 55 |
| 2. Claim 2 | 80 |
| 3. Claim 3 | 80 |
| 4. Claim 4 | 81 |
| 5. Claim 5 | 86 |
| 6. Claim 6 | 88 |
| 7. Claim 7 | 89 |
| 8. Claim 8 | 90 |
| 9. Claim 9 | 94 |
| 10. Claim 10 | 95 |
| 11. Claim 11 | 97 |
| 12. Claim 12 | 104 |
| 13. Claim 13 | 104 |
| 14. Claim 14 | 106 |
| 15. Claim 16 | 106 |
| 16. Claim 18 | 110 |
| 17. Claim 19 | 111 |
| 18. Claim 20 | 112 |
| 19. Claim 21 | 113 |
| 20. Claim 22 | 118 |
| 21. Claim 23 | 118 |
| 22. Claim 24 | 118 |
| VIII. MOTIVATION TO MODIFY OR COMBINE | 118 |

TABLE OF CONTENTS
(continued)

| | Page |
|---|-------------|
| A. Brunk’s Teaching of Storing Actions With Fingerprints in Databases..... | 119 |
| B. Robust Hash Fingerprints of Video..... | 122 |

I, Ahmed Tewfik, Sc. D., hereby declare and state as follows:

I. BACKGROUND AND QUALIFICATIONS

1. I was the Chairman of the Department of Electrical and Computer Engineering at the Cockrell School of Engineering at the University of Texas at Austin from October 2010 through November 2019.

2. I earned a Bachelor of Science degree in Electrical Engineering from Cairo University in Cairo, Egypt in 1982. I earned a Master of Science degree in Electrical Engineering in 1984 and Doctor of Science in 1987, both from the Massachusetts Institute of Technology (“MIT”).

3. I worked as a systems engineer at Alphatech, Inc. in Burlington, Mass. in 1987. I also served as a lecturer at MIT during the spring of 1987.

4. From 1987 to 2010, I was a professor in the Department of Electrical Engineering at the University of Minnesota in Minneapolis, MN. From 1987 to 1992, I served as an Assistant Professor. In 1992, I took the position of Associate Professor which I kept until 1993, when I became the E.F. Johnson Associate Professor of Electronic Communications. In 1996, I took the position of E.F. Johnson Professor of Electronic Communications, which I held until 2010 (including a leave from 1997-2001). In 2010, I became Chairman of the Department of Electrical and Computer Engineering at the Cockrell School of

Engineering at the University of Texas at Austin. I currently hold the Cockrell Family Regents Chair in Engineering in the same Department.

5. Between 1997 and 2001, during my time at the University of Minnesota, I also served as the founder and CEO of Cognicity, Inc., a company that developed and sold watermarking products for the music industry.

6. I have also worked as a consultant for numerous companies, including St. Jude Medical, Visonaire, Macrovision, and Digimarc. A complete list of my prior consulting engagements is included in my concurrently-submitted curriculum vitae. Ex-1010.

7. I have been widely published, and have authored over 100 articles on signal processing and related topics in numerous scientific journals over the course of the career. I have also authored over 350 publications for conferences, and have spoken at conferences, seminars and symposia hundreds of times, and have contributed to multiple treatises on signal processing. A complete list of my publications and presentations is included in my concurrently-submitted curriculum vitae. Ex-1010.

8. I have lectured and presented on signal processing and related topics at multiple prestigious universities, including University of Texas at Austin, Carnegie Mellon University, Cornell University, Boston University, MIT, and many others. I have also been invited to present at many of the world's leading

companies in the field of signal processing, including Texas Instruments, AT&T, Microsoft, Apple, IBM, NEC, Mitsubishi, GE, and several others. A more complete list of my presentations and speaking engagements is included in my concurrently-submitted curriculum vitae. Ex-1010.

9. I am also a named inventor on twenty-five U.S. Patents, several of which relate to signal processing and related topics, including data embedding, watermarking, fingerprinting, and content recognition. A complete list of my granted patents is included in my concurrently-submitted curriculum vitae. Ex-1010.

10. Much of my academic career has focused on signal processing. I received the NSF Research Initiation Award in 1990 and the Taylor Faculty Development Award from the Taylor Foundation in 1992. I was elected an IEEE Fellow in 1996 and was a distinguished lecturer of the IEEE Signal Processing Society from 1997 to 1999. I was awarded the IEEE Third Millennium Award in 2000 and the IEEE Signal Processing Society Technical Achievement Award in 2017 for my research and contributions to the field of signal processing.

11. I am also heavily involved in the IEEE Signal Processing Society (SPS). I served on the Society's Board of Governors from 2006 to 2008 and as the Vice President-Technical Directions from 2010-2012. I am the current President-elect of the SPS and will begin my term in January 2020.

II. ASSIGNMENT AND MATERIALS CONSIDERED

12. I have been retained by Free Stream Media Corp., d.b.a. Samba TV (“Samba”). I have been asked to prepare this declaration in connection with Samba’s Petition for *Inter Partes* Review (“IPR”) of U.S. Patent No. 9,066,114 to Oostveen et al., which I refer to herein as the ’114 patent.

13. I am not an employee of Samba or any affiliate or subsidiary of Samba.

14. I am being compensated at the rate of \$500 per hour for my work, plus reimbursement for expenses. My compensation has not influenced any of my opinions in this matter and does not depend on the outcome of this proceeding.

15. I have been asked to provide analysis and explain the subject matter of the ’114 patent, including the state of the art when the ’114 patent application was filed. I have also been asked to consider, analyze, and explain certain prior art to the ’114 patent, including how that art relates to certain claims of the ’114 patent and to provide my opinions regarding whether that art anticipates or renders obvious the claimed subject matter.

16. The opinions expressed in this declaration are not exhaustive of my opinions regarding the unpatentability of the claims of the ’114 patent. Therefore, the fact that I do not address a particular point should not be understood to indicate

an agreement on my part that any claim complies with the requirements of any applicable statute or other rule.

17. I reserve the right to amend and supplement this declaration in light of additional evidence, arguments, or testimony presented during this IPR.

18. In forming the opinions set forth in this declaration, I have considered and relied upon my education, knowledge of the relevant field, knowledge of scientific and engineering principles, and my experience. I have also reviewed and considered the '114 patent (Exhibit 1001), its prosecution history (Exhibit 1004), and the following additional materials:

| <i>Exhibit</i> | <i>Description</i> |
|-----------------------|--|
| 1001 | U.S. Patent 9,066,114 to Oostveen et al., “Method and Device for Generating and Detecting a Fingerprint Functioning as a Trigger Marker in a Multimedia Signal,” filed Jan. 10, 2014 (“the ’114 Patent”) |
| 1002 | U.S. Patent 9,407,962 to Oostveen et al., “Method and Device for Generating and Detecting a Fingerprint Functioning as a Trigger Marker in a Multimedia Signal,” filed Mar. 16, 2015 (“the ’962 Patent”) |
| 1003 | U.S. Patent 9,479,831 to Oostveen et al., “Method and Device for Generating and Detecting a Fingerprint Functioning as a Trigger Marker in a Multimedia Signal,” filed May 25, 2016 (“the ’831 Patent”) |
| 1004 | Select portions of prosecution history of the ’114 Patent (“the ’114 File History”) |
| 1005 | Intentionally Omitted |
| 1006 | Intentionally Omitted |
| 1007 | Declaration of Dr. Ahmed H. Tewfik regarding U.S. Pat. No. 9,066,114 (“Tewfik Decl.”) |
| 1008 | Intentionally Omitted |

| <i>Exhibit</i> | <i>Description</i> |
|-----------------------|---|
| 1009 | Intentionally Omitted |
| 1010 | <i>Curriculum Vitae</i> of Dr. Ahmed H. Tewfik |
| 1011 | U.K. Published Patent Application GB 2,375,907A to Murphy, “An Automated Recognition System,” filed May 14, 2001, and published November 27, 2002 (“Murphy”) |
| 1012 | U.S. Patent Application Publication 2002/0126872 to Brunk et al., “Method, Apparatus and Programs for Generating and Utilizing Content Signatures,” filed December 19, 2001, and published September 12, 2002 (“Brunk”) |
| 1013 | W. Kate, et al. “trigg&link: A New Dimension in Television Program Making,” in Lecture Notes in Computer Science 1197, vol. 1242, Multimedia Applications, Services and Techniques-ECMAST ’97, (1997) 51-65 (“Kate”) |
| 1014 | J. Oostveen, T. Kalker, and J. Haitsma, “Feature Extraction and a Database Strategy for Video Fingerprinting,” in Recent Advances in Visual Information Systems, 5th International Conference, VISUAL (2002) 117-128 (“Oostveen”) |
| 1015 | S.E. Johnson and P.C. Woodland, “A Method for Direct Audio Search with Applications to Indexing and Retrieval,” IEEE International Conference on Acoustics, Speech, and Signal Processing, June 5-9 2000 (2000) 1427-1430 (“Johnson”) |
| 1016 | U.S. Patent Application Publication No. 2002/0178410 to Haitsma, J., Kalker, A, Baggen, C., and Oostveen J., “Generating and Matching Hashes of Multimedia Content,” filed February 11, 2002, and published November 28, 2002, claiming priority to EP applications 01200505.4 and 01202720.7 (“Haitsma”) |
| 1017 | R. Venkatesan, S.-M. Koon, M. H. Jakubowski, and P. Moulin, “Robust Image Hashing,” 0-7803-6297-7/00, IEEE (2000), 664-666 (“Venkatesan”) |
| 1018 | U.S. Patent 6,637,032 to Feinleib, David, “System and Method for Synchronizing Enhancing Content With A Video Program Using Closed Captioning,” filed Jan. 6, 1997 (“Feinleib”) |
| 1019 | Intentionally Omitted |

| <i>Exhibit</i> | <i>Description</i> |
|----------------|---|
| 1020 | U.S. Patent Application Publication No. 2002/0032864 to Rhoads, Geoffrey B., Levy, Kenneth L. “Content Identifiers Triggering Corresponding Responses,” filed May 14, 2001, and published March 14, 2002 (“Rhoads”) |

III. THE ’114 PATENT’S EFFECTIVE FILING DATE

19. I understand that the ’114 patent claims priority to July 11, 2003. I understand that priority date is not being disputed in Samba’s petition.

20. For purposes of this declaration, I have considered both the understanding and knowledge of one of ordinary skill in the art and unpatentability as of this July 11, 2003, date.

IV. CLAIM INTERPRETATION

21. I understand that during an *inter partes* review, the terms of a patent are given their plain and ordinary meaning as understood by one of ordinary skill in the art in light of the specification.

22. Based on my review of the claims, specification, and prosecution history of the ’114 patent, I have formed opinions concerning the way the following terms should be understood in light of the ’114 patent’s specification.

A. “A processor”

23. Claim 8 of the ’114 patent recites “a processor” that is configured to
- access a database that includes a plurality of reference fingerprints, previously derived from the content, and a plurality of reference actions, each reference fingerprint identifying at least one

corresponding reference action to be performed at a reference time point identified by the reference fingerprint;

- identify the corresponding reference action by obtaining a match in the database between a trigger fingerprint among the plurality of trigger fingerprints and a reference fingerprint among the plurality of reference fingerprints; and
- perform a reference action that corresponds to the reference fingerprint on the playback device.

24. In my opinion, the term “a processor” should be construed as “one or more processors.”

25. I understand that, legally, terms referring to “a” or “an” element are not generally understood to require only a single instance of that element, but may refer to “one or more” instance of that element. I also understand that “a” or “an” may limit the scope of the claim to a single instance of an element where the patent “evinces a clear intent to limit ‘a’ or ‘an’ to ‘one.’” Based on my review of the materials, I have seen no evidence of any such clear intent to limit the term “a processor” to a single processor.

26. Indeed, the ’114 patent’s specification expressly adopts the “one or more” legal rule, explaining that “in the claims,” the use of “the word ‘a’ or ‘an’ preceding an element does not exclude the presence of a plurality of such elements.” Ex-1001, 7:58-63.

27. For that reason, and the reasons explained further below, in my opinion, the term “a processor” as used in claim 8 would also be understood by one

of ordinary skill in the art to refer to “one or more processors.” The “processor” of claim 8 is tasked with accessing the database of reference fingerprints referred to in the specification as data base 203. The ’114 patent explains that the data base 203 is, along with other components constituting the fingerprint generation device 200, “under the control of one or more microprocessors.” Ex-1001, 6:50-54; *see also* 7:24-25 (also describing data base 203 as controlled by one or more microprocessors). Thus, it would be clear to one of ordinary skill in the art reading the specification of the ’114 patent that the “processor” recited in claim 8 is “one or more processors.”

28. Moreover, one of ordinary skill in the art knew that it was often beneficial to implement functionality on more than one processor, even when a single processor could otherwise be utilized. The load distribution benefits obtained by distributing processing operations across multiple processors were well known at the time of the ’114 patent. Thus, a person of ordinary skill in the art would understand that there was no pressing reason why the term “a processor” as used in claim 8 would strictly require only a single processor.

B. “Multimedia”

29. The ’114 patent uses the term “multimedia” in every independent claim. In each instance where the term “multimedia” is recited in the claims, it is used as an adjective to define the type of content to which another claim limitation

pertains. For example, the claims recite terms including “multimedia content,” “multimedia signal,” and “multimedia playback device.”

30. In my opinion, the term “multimedia” should be construed to mean “audio and/or visual media.”

31. Although it is true that some use the term “multimedia” to refer to types of media involving two or more media, one of ordinary skill in the art would understand that the ’114 patent and the relevant prior art use the term more broadly. Indeed, the ’114 patent uses the word “multimedia” multiple times and repeatedly specifies that the term refers to “audio and/or visual media.” *See, e.g.*, Ex-1001, 3:51-55 (describing a multimedia signal as comprising “video and/or audio information/content”); 5:21-25 (describing multimedia as “audio and/or video”); 6:59-62 (same); 7:29-33 (same).

32. Likewise, the ’114 patent also makes clear that “audio clips,” which one of ordinary skill in the art understands to be a solely audio form of media, are considered by the named inventors of the ’114 patent to be a type of “multimedia content.” *Id.*, 4:38-41. This is consistent with the patent’s indication that “multimedia” can include audio and/or visual media.

33. Notably, the prior art (including prior art cited by the ’114 patent) also used the term “multimedia” to refer to audio and/or visual media. *See, e.g.*, Ex-

1016, [0062] (likewise referring to “audio clips” as an example of “multimedia content”).

V. OVERVIEW OF THE '114 PATENT

34. The '114 patent generally relates to a system and method for detecting digital fingerprints during playback of a segment of a multimedia signal, where the fingerprints identify a specified time in the playback of the multimedia content for triggering a particular action at the specified time. The '114 patent accomplishes this operation in part through use of a database that stores reference fingerprints and their corresponding actions. As the signal is played, it is processed to identify fingerprints which are then compared to the stored fingerprints in the database in order to find a match and trigger the appropriate action at the appropriate time. This functionality, the patent explains, can be used to, for example, allow regional commercials or weather reports to be automatically inserted at the proper times in the content of a national broadcast. Ex-1001, 1:53-62, 3:64-4:6.

A. The Problem to Be Solved

35. The '114 patent relates to an admitted “current trend” in the industry at the time the patent was filed, wherein the prior art “enhance[d] passive television viewing and/or music listening on a given playback device by creating more interactive programs and/or listening experiences or by ‘connecting’ external actions to a piece of video and/or audio content.” *Id.*, 1:22-26. The patent

describes examples in the prior art of triggering actions at particular instants in time during playback, such as displaying information or a URL to the consumer. *Id.*, 1:29-67 (referencing Kate (Ex-1013)). The patent explains that those in the industry desired taking such actions at particular times during playback. *Id.*, 1:48-52, 1:63-67.

36. The '114 patent acknowledges that there were several techniques in the prior art for creating a trigger marker that would be used to trigger an action at a desired time during playback. *Id.*, 1:67-2:34. One set of discussed prior art techniques involved embedding data or signals into the broadcast signal to function as trigger markers. *Id.*, 1:67-2:13. The patent criticizes such solutions as being dependent on the actors in the broadcast chain to create such signals, and then to preserve them during the various signal-processing events involved in broadcasting. *Id.*, 2:1-4, 2:13-29. The patent describes another prior art solution involving embedding a watermark into a frame of a signal at the relevant time point to function as a trigger marker. *Id.*, 2:31-33 (discussing commercially available watermark system). The patent notes, however, that embedding the watermark problematically “changes the video/audio.” *Id.*, 2:34-35.

B. The '114 Patent's Purported Solution

37. To address the perceived deficiencies in the prior art, while meeting the art-recognized desire and industry trend of generating actions at specific times

during playback, the '114 patent describes using the prior art technique of fingerprinting—which involved the known process of selecting and identifying specific portions of audio or video content—as the trigger marker, rather than using embedded signals or watermarks. *Compare id.*, Abstract and 2:36-3:16 (discussing the purported invention) with 4:35-44, and 4:53-60 (discussing admitted fingerprinting techniques in prior art, corresponding to Oostveen (Ex-1014) and Haitsma (Ex-1016)). Thus, the specification of the '114 patent presents the purported invention as simply the substitution of one known technique for generating trigger markers (fingerprinting) for another known technique (using embedded watermarks or signals) to obtain the same predictable results.

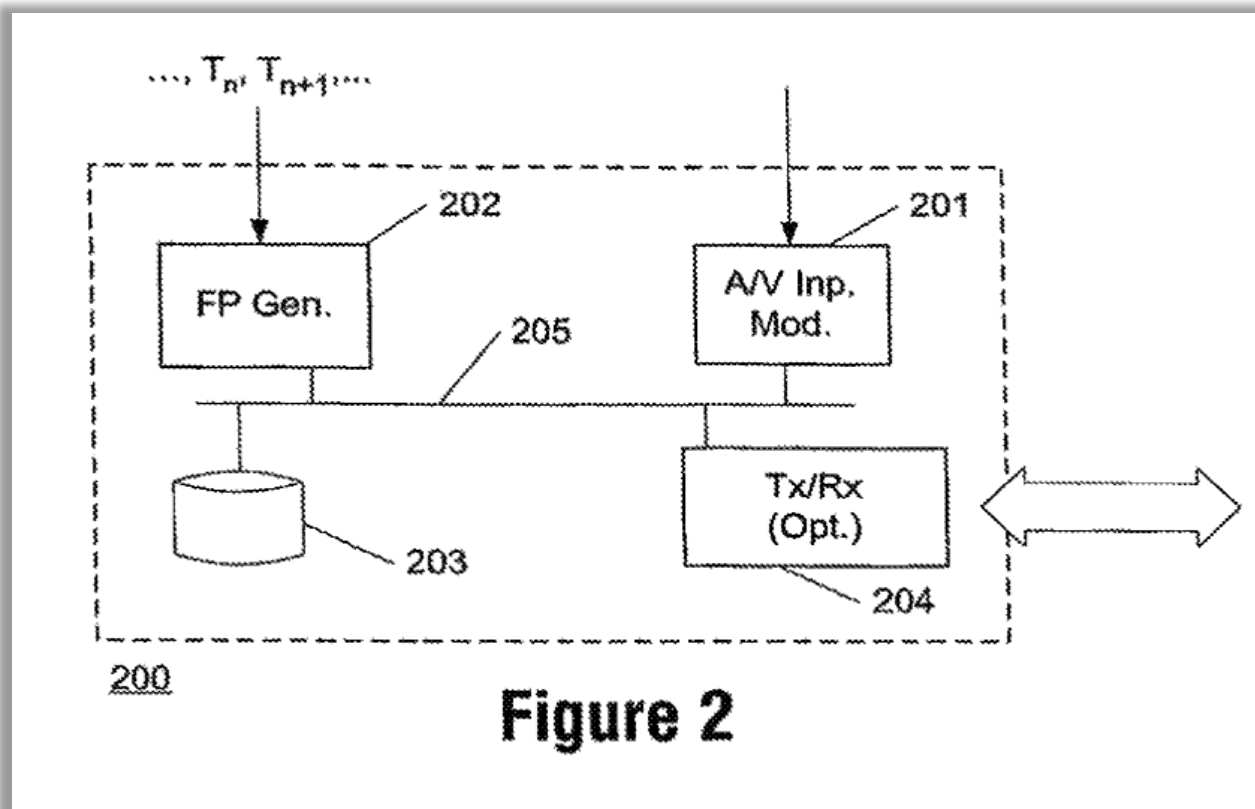
38. As such, the '114 patent and its admitted prior art reveals that it pursued a marginal substitution of one known technique (fingerprinting) for another known technique (watermarking) to improve similar devices (playback devices). The claimed playback devices incorporating this minor substitution work in the same way, to achieve the same predictable results of triggering actions at specific times, in reaction to admitted market forces (the industry trend of providing enhanced user experiences by triggering time-specific actions).

C. Detailed Description

39. As was in the admitted fingerprinting prior art (for example, Haitsma (Ex-1016)), the specification describes generating a fingerprint for a segment of

content such as audio and/or video, storing the fingerprint in a database, and later matching the stored fingerprint against a similar fingerprint generated during subsequent playback by a fingerprint detection system, for example when a viewer watches or listens to the content during a broadcast. *See, e.g.*, Ex-1001, Abstract, Claim 1, 3:25-26, 7:29-40. Consistent with the industry trend of using embedded signals and watermarking as trigger markers for actions at particular times, the '114 patent teaches triggering an action, such as displaying information or executing control commands, at a point in time relative to the location of the fingerprint in the content. *See, e.g., id.*, Abstract, 3:9-16, 3:64-4:10, 7:29-40.

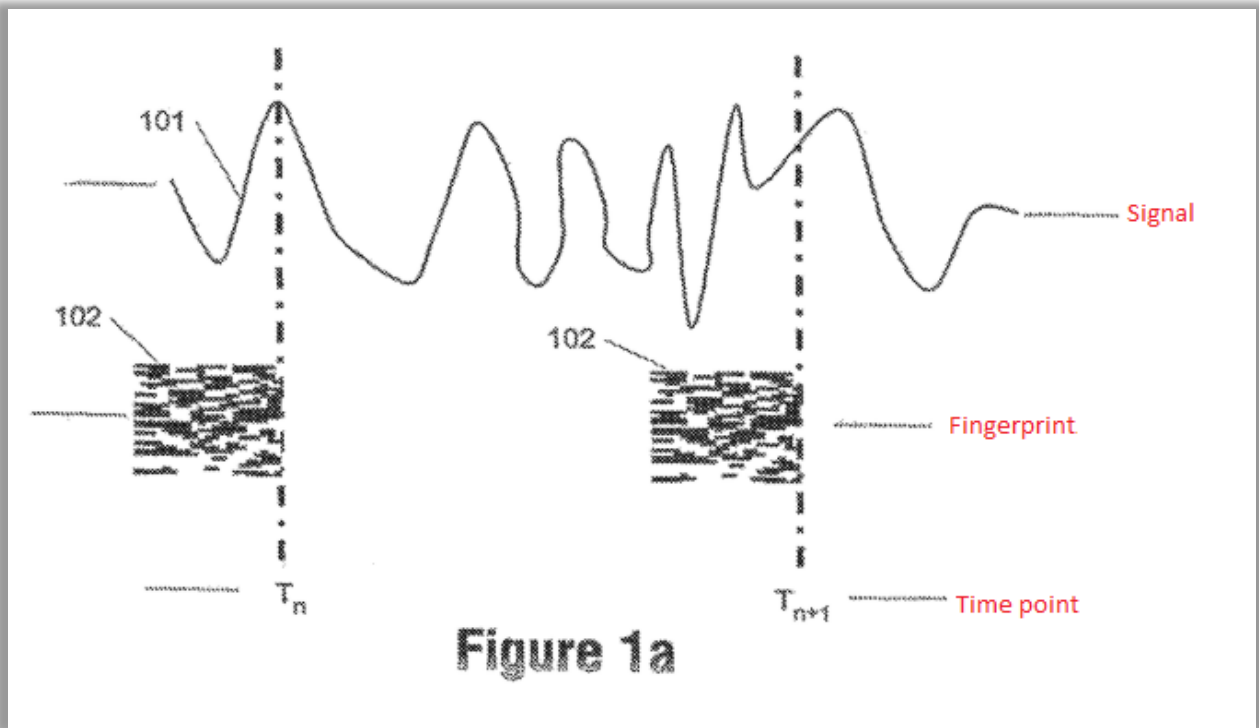
40. More specifically, the '114 patent describes computer modules and components for generating fingerprints, detecting fingerprints, and taking actions at time points identified by matching the detected and generated fingerprints. For example, Figure 2 illustrates a fingerprint generation device (200) comprising a multimedia signal input module (201), a fingerprinting module (202), a database (or memory or the like) (203), and a transmitter and receiver (204):



Id., Fig. 2, 6:48-58.

41. These modules and components are controlled by one or more not-shown microprocessors and communicate within fingerprint generation device (200) via a bus (205). *Id.*, Fig. 2, 6:48-58.

42. With reference to Figures 2 (above) and 1a (below), signal input module (201) receives analog or digital content in signal (101) and feeds the signal (101) to fingerprinting module (202) to compute fingerprints (102):



Id., Figs. 1a (annotated), 2, 3:50-52, 4:15-23, 6:59-64.

43. Fingerprints are generated for each time point (T_n ; T_{n+1}) at which an action is to be associated with signal 101. *Id.*, 3:52-55, 4:15-17. As discussed above, the '114 patent admits that fingerprint generation was already known in the art. *Id.*, 4:34-60.

44. As shown in Figure 4, once generated, the fingerprints may be stored by fingerprinting module (202) in a simple database (203) that associates these “reference” fingerprints (‘FP1’, ‘FP2’, etc.) with one or more actions (‘A1’, ‘A2’, etc.):

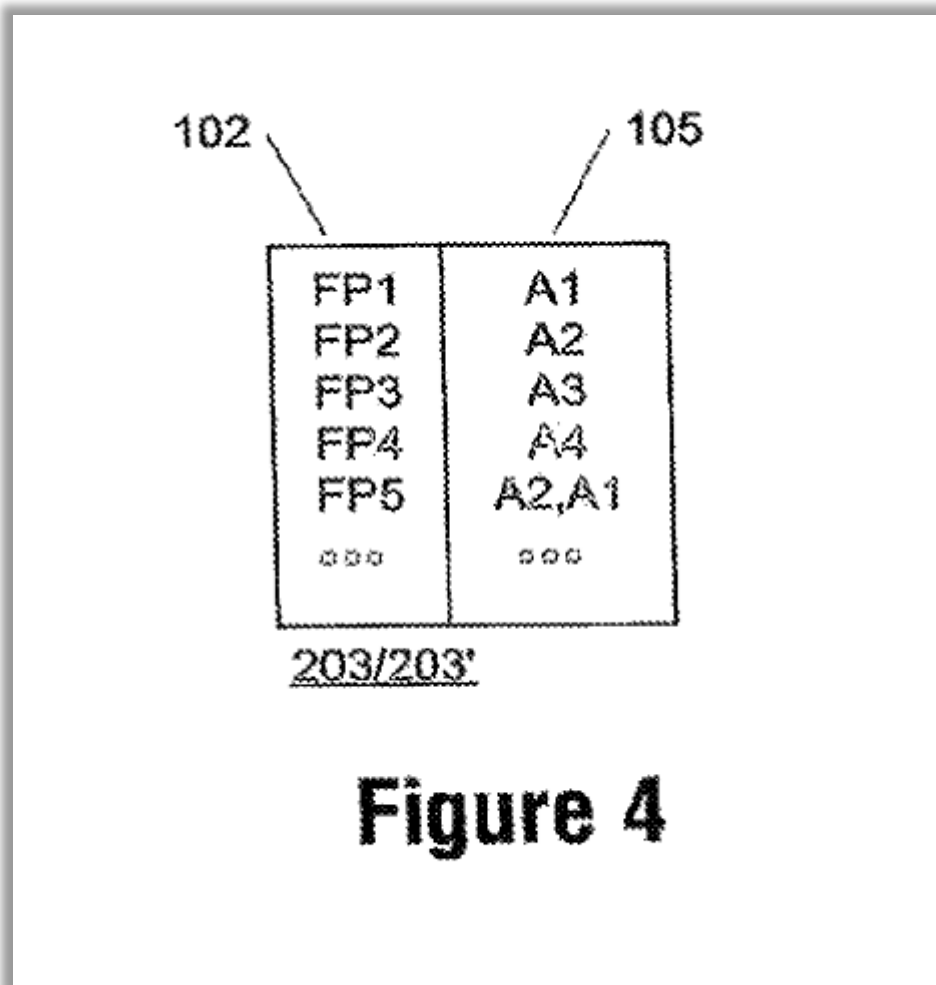


Figure 4

Id., Fig. 4, 7:5-17, *see also* Claim 1.

45. As shown in Figure 3 (below), a playback device (300) looks for matches between the generated, stored reference fingerprints (102, for example, ‘FP1’) and fingerprints detected in subsequent playback of the signal (101). *Id.*, 7:18-20. To match fingerprints, playback device (300) comprises similar modules and components as those within fingerprint generation device (200): a signal receiver (301), a fingerprinting detector (302), a database (or memory or the like) (203’), and a transmitter and receiver (204):

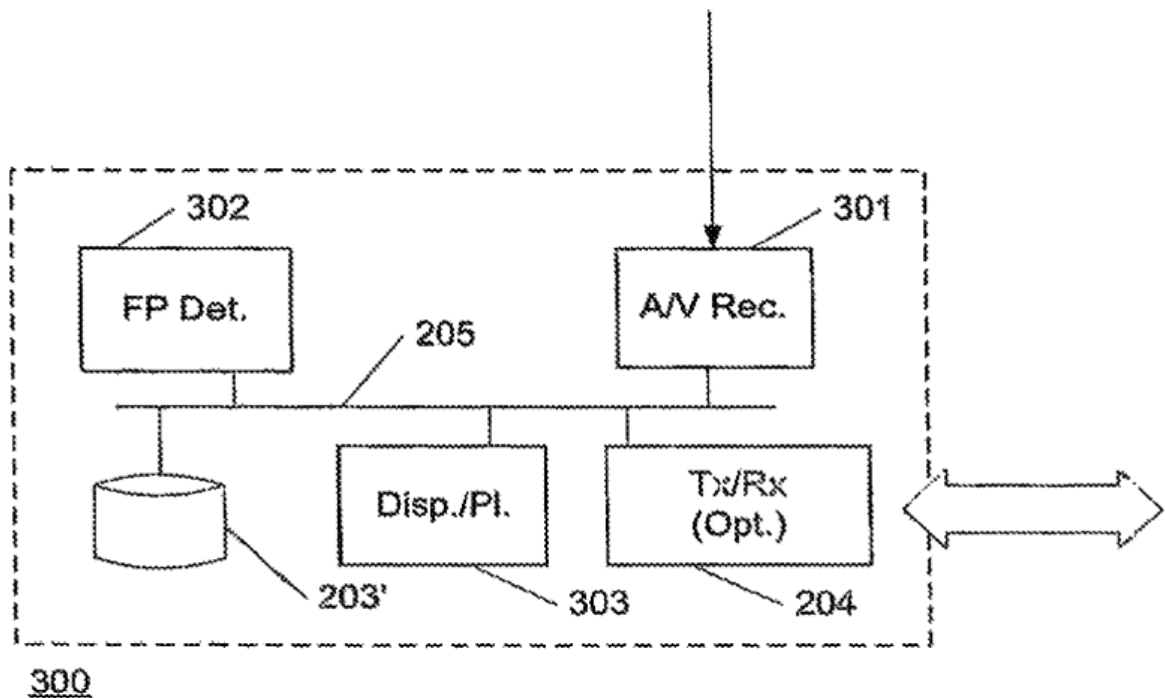


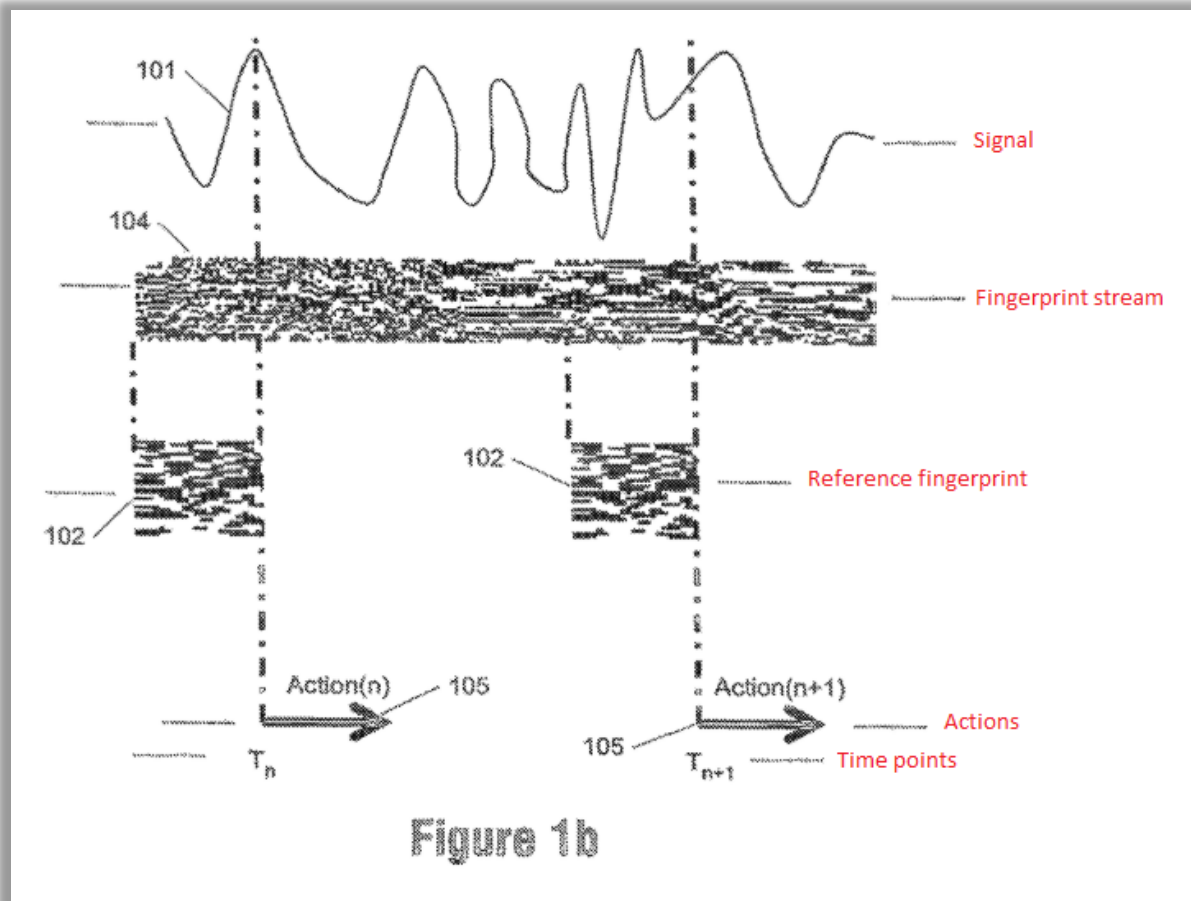
Figure 3

Id., Fig. 3, 7:18-27. The playback device (300) also includes a display/play circuit (303), for displaying and/or playing the received multimedia signal. *Id.*, 7:41-45

46. These modules and components are also operated by “one or more microprocessors (not shown)” and communicate within playback device (300) via a bus (205). *Id.*, 7:23-25.

47. Referring to Figures 3 (above) and 1b (below), signal receiver (301) receives the signal (101) and feeds it to a fingerprinting detector (302) that derives a fingerprint stream (104). Fingerprinting detector (302) further determines if

there is a match between the fingerprint stream (104) and a fingerprint (102) stored in database (203’):



Id., Fig. 1b, 7:28-37.

48. Like the process for generating fingerprints, the process for matching fingerprints was already known in the art. *Id.*, 6:25-42. When a match is found, the action (105) (‘A1’, ‘A2’, etc.) associated with the stored fingerprint (102) (‘FP1’, ‘FP2’, etc.) is executed by the playback device at the associated time point (T_n ; T_{n+1}). *Id.*, 5:65-6:20, 7:37-45.

49. As shown in Figure 1b, the actions are taken at the end of the portion of the signal having the fingerprint, which is at the time that the detection system has read the fingerprint and identified a match to the database. *Id.* The '114 patent states that the action time point could be referenced to the end of the fingerprint, the beginning of the fingerprint or some fixed time period referenced to (i.e. offset from) those points. Ex-1001, 4:15-30. The '114 patent teaches processing the signal at/near the desired action so that the location of the generated fingerprint in the signal becomes the time marker:

According to the present invention, a fingerprint (102) is generated for each time point on the basis of a given segment of the multimedia signal (101) at or near the specific time point.

In this way, a given fingerprint (102) is a trigger marker enabling a very accurate and very precise location of a given time point of the signal (101) without using the specific time point but instead using (a segment of) the signal. Further, this is enabled without changing the signal. For video fingerprinting the localisation is typically frame accurate, at least as long as any distortion of the video signal is not too severe.

Id., 4:61-5:4.

50. However, this type of synchronization between the time marker in the signal (whether it was an embedded data signal, a watermark, or a fingerprint) amounts to no more than the normal and routine synchronization between the time that an action was to be taken (for example, displaying information) to the relevant time in the underlying multimedia content. *Id.*, 1:22-32; 1:36-52 (discussing Ex-1013). That problem was already solved in the ordinary course in the prior art by

those who ensured that time triggers were placed at or near the right point to give the desired “enhanced” listening or viewing experience. *Id.*

D. Prosecution History of the ’114 Patent

51. The application that became the ’114 patent was filed on January 10, 2014, as a continuation of U.S. patent application 13/220,267. Ex-1001, (21, 22, 63). The examiner rejected the claims as anticipated and/or obvious over Haitsma (Ex-1016). The applicant argued that Haitsma was merely used for content identification, and therefore did not teach that “each reference fingerprint identif[ied] at least one corresponding reference action to be performed at a reference time point identified by the reference fingerprint.” Ex-1004, 47, 51-52. The Examiner apparently accepted this position despite the fact that the specification acknowledges that using markers to trigger actions at relevant times was known in the art, and using a fingerprint represented no more than an obvious substitution of one type of marker for another, with entirely predictable results. The prior art references of Murphy (Ex-1011) and Brunk (Ex-1012) were not before the examiner during the prosecution of the ’114 patent, and disclose the limitation—using a fingerprint as a marker to trigger actions at the relevant times—that the Examiner found lacking.

VI. THE STATE OF THE ART

A. Level of Skill in the Art

52. I have been informed and understand that the level of ordinary skill in the relevant art at the time of the invention is relevant to inquiries such as the meaning of claim terms, the meaning of disclosures found in the prior art, and the reasons one of ordinary skill in the art may have for combining different references.

53. I have been informed and understand that factors that may be considered in determining the level of ordinary skill include: (1) the education of the inventor; (2) the type of problems encountered in the art; (3) prior art solutions to those problems; (4) rapidity with which innovations are made; (5) sophistication of the technology; and (6) education level of active workers in the relevant field. I have been further informed and understand that a person of ordinary skill in the art is also a person of ordinary creativity.

54. In my opinion, the ordinarily skilled artisan in the technology field of the '114 patent would have a bachelor's degree in computer science, electrical engineering or a similar field, and two years of work experience in multimedia signal processing, including watermarking, fingerprinting and their applications. A greater degree of education or work experience could substitute for a lesser degree of the other.

55. I note that my qualifications qualify me as a person with knowledge and experience beyond “ordinary skill” in the art because, as explained above in Section I, my educational accomplishments and work experience both exceed this standard.

B. General Information Known to a Person of Ordinary Skill in the Art

56. In my opinion, every aspect of the claimed subject matter recited in the ’114 patent was part of the general knowledge of a person of ordinary skill in the art as of July 11, 2003.

1. It was known to enhance media with time-specific actions and features

57. As of July 11, 2003, it was known to known to enhance media with time-specific actions and features. With the advent of the Internet in the 1990s, media producers and distributors realized that non-interactive, “linear” media content, from over-the-air television and radio broadcasts, could be enhanced to provide an interactive media experience by triggering actions at various specific points during playback. For example, the company I led, Cognicity, developed in 2000 and 2001 enhanced music playing software that relied on embedded watermarks to access relevant information on websites, play related video clips or offer discount coupons for buying music CDs online or in physical retail stores. As another example, as the ’114 patent admits, there was an industry trend in the

prior art to enhance television viewing and music listening by creating more interactive programs and experiences, or by connecting external actions to a piece of video or audio content. Ex-1001, 1:22-26. Likewise, it was known to embed signals or watermarks into a piece of multimedia content to function as trigger markers to trigger actions—such as displaying information or a URL to the consumer that relates to what is currently being viewed or heard by the viewer—at particular times during playback. *Id.*, 1:29-2:35. Indeed, Brunk specifically discloses the use of watermarks and triggering actions to provide a URL. Ex-1012, [0008]-[0009] (watermarks), [0028]-[0029] (URLs).

58. A person of ordinary skill in the art knew that the same triggers could provide additional features and functionality by controlling systems on a playback device. Murphy expressly discusses this. *See, e.g.*, Ex-1011 (discussed *infra*, controlling automated start and stop recording events; controlling the display of additional information to a user).

59. Embedding data or other signals within a piece of audio-visual content, with the embedded data or signal acting as a trigger marker to trigger an action at the relevant time in the content was known in the art. *See, e.g.*, Ex-1001, 1:67-2:13. Murphy specifically teaches using fingerprints to trigger actions (such as starting a recording device) at particular times. Ex-1011, 13:17-25. And as discussed in additional detail below, Kate (Ex-1013), which is acknowledged prior

art in the '114 patent (Ex-1001, 1:33-52), is an exemplary reference for using embedded data or signals as trigger markers.

60. Kate taught using the embedded data to trigger the presentation of an icon or URL to a user, who could then selectively visit a website of interest related to the media content. Ex-1013, 51-52. By doing so, a user could be provided with an interactive media experience. *Id.* Kate taught that it was “obvious to insert that [web]page in the broadcast stream along with the so-called trigger displaying the URL,” because all that was required was to have a website to which the user could be referred. *Id.*, 52.

61. In an example discussed in Kate, a compact disk of a classical concert could be enhanced by providing multiple triggers within the work that provided information on the composer, conductor, and musical style or link(s) to a live broadcast of the work. Ex-1013, 62-64. This demonstrates that it was known to provide such actions at particular points in time relevant to the content.

62. As another example, Kate taught that one could present the viewer with information about a conductor of a musical performance during periods in which the program was displaying images of the conductor. *Id.*, 63. Further, Kate taught presenting the musical style to the viewer, with the style necessarily reflecting the musical style of the content being played at the time of the trigger. *Id.*

63. Kate further taught that “[i]t [was] obvious that multiple triggers can be presented at the same time” during playback of the media content. *Id.*, 55. The triggered actions could present information or additional media content by (1) wholly replacing the view of the original media content and/or (2) overlaying the original content, such as in a picture-in-picture. *Id.* Notably, Murphy also teaches that a single fingerprint can trigger multiple different actions, including recording and updating program-type data that, like Kate, could be displayed to inform the audience about the content. Ex-1011, 13:17-26.

64. Kate also taught that presenting “context specific information” through enhanced media was “more convenient and direct” than accessing the information independently outside of the content. Ex-1013, 64. Such trigger markers were also often used during national broadcasting to indicate to local broadcasters the time at which a portion of the signal could be replaced with something like a local weather forecast or a local advertisement.

2. It was known to trigger time-specific actions at precise times using content identification

65. The ’114 patent describes that it was known to enhance the playback experience with a set of actions at desired points (for example, displaying information about an actor at the appearance of the particular actor), and then trigger the action to occur precisely when desired, including using the content-recognition technique known as watermarking. Ex-1001, 1:22-32, 1:63-2:5, 2:9,

2:31-34. All of the concepts discussed in the background of the '114 patent from column 1 line 22, through column 2 line 34, were well known. Namely, one of ordinary skill in the art knew that in the late 1990s and early 2000s, there was an industry trend in the prior art of offering enhanced audio-visual content experiences by providing additional information or features at particular times in played content. *See, e.g.*, Ex-1001, 1:22-62.

66. This industry trend created a market incentive for playback systems to generate actions at particular times in the content, as well as several known time-marking techniques to function as a trigger marker, including recognizing a watermark in the audio or video content itself as the relevant time marker. *See, e.g.*, Ex-1001, 1:63-2:1 (including “At these time instants the receiver should be triggered to perform or provide some kind of action.”); 2:5 (“One previously known way of performing time marking” is embedding “marking information” in “fields of the MPEG transport stream structure.”); 2:9-12 (“Another previously known way is” embedding “marking information” in a “blanking interval.”); and 2:30-34 (“Yet another way [of performing time marking] is using watermarking. A watermark may be embedded in the video frames at the relevant time instants.”). Murphy likewise teaches using a fingerprint as a time marker, as Murphy discloses taking actions (for example, to start recording content) upon detection of the

fingerprint (indicating the start of the portion to be recorded) in the live content signal. Ex-1011, 13:17-25.

67. As mentioned above, it was known to use time markers at any time in the broadcast of a content signal, including the beginning, end, or any time throughout. For example, Feinleib (Ex-1018) taught that it was desirable to trigger actions in the middle of the playback of a particular content item. Ex-1018, 1:55-65. Feinleib describes, for example, that it would be beneficial to allow the display of supplemental content to be triggered at specific times in an episode, and goes on to describe triggering the display of a hyperlink to a hat retailer at a specific time in an episode of Gilligan's Island where Gilligan ruins the Skipper's hat. And Feinleib discloses an example of the application of its system whereby the action to display a URL to Levi's Jeans' website is at the point in an episode of the television sitcom "Seinfeld" when Jerry Seinfeld enters the frame wearing a pair of Levi's. *Id.*, 5:45-56.

68. The family history of the '114 patent further shows that it was known to use fingerprints placed at particular times in the content to trigger actions at those particular times. During the prosecution of U.S. Patent No. 8,020,000, a patent to which the '114 patent claims priority, the Examiner rejected some of the claims as anticipated by a reference referred to as "Rhoads" (Ex-1020). Ex. 1019, (Non-final Rejection 7/12/2000) at p. 2. The Examiner explained that because

Rhoads disclosed fingerprints that were derived from particular windows of times in the content, each fingerprint was associated with a window and, thus, each window could be considered a trigger time point. *Id.* Referring to an application involving a song as an example, the examiner explained:

The identification of the song depends on the fingerprint formed from a time window associated with the song (paragraph 21). Since that fingerprint is associated with the time window of the song and is used to trigger a response (paragraph 27), the triggered response corresponds to the time point/time window within the multimedia signal/song since until that time window occurs, the fingerprint cannot be identified to trigger an action.

Id., at p. 3.

69. The Rhoads reference cited by the examiner discusses how a response could be triggered “once a song has been identified in a database” on the basis of a fingerprint. Ex-1020, [0027]. Thus, in the context of a fingerprint based on just a segment of the song, one of ordinary skill in the art reading Rhoads would understand (as the Examiner understood) that the response could be triggered at a point in the middle of the song, if that is where the recognized segment happened to be.

70. These techniques in the prior art required a person of ordinary skill in the art to be able to create the correct playback experience by synchronizing the exact placement of the time marker, for example the watermark, at or near the time that the action should occur, taking into consideration things like the length of the

watermark in the signal and the processing time needed to recognize the watermark during playback.

3. Fingerprinting was a well-known, beneficial substitute for watermarking to achieve known, predictable results of time-specific content identification

71. Various methods of identifying media content were well known and predictable at the time of the priority date of the '114 patent. Watermarking and fingerprinting are two examples.

72. Watermarking involved adding information, typically imperceptible to the content user, into the media content itself. Watermarking had the known disadvantage that it required the original content to have both undergone the watermark modification, and to have done so in a way that did not degrade the quality of the content—neither of which was a given. Ex-1016, [0004]; Ex-1001, 2:34-35. Another known disadvantage was that the watermark could be unintentionally or intentionally distorted beyond recognition. *See* Ex-1017, 666 (“with watermarking, on the other hand, the first move belongs to the embedding algorithm, which commits to a particular watermark that the attacker can subsequently try to remove.”).

73. In contrast, prior art fingerprinting involved processing a known media signal to generate a unique, identifying representation of characteristics of that known signal at a particular point in the signal to serve as a signature, in other

words its “fingerprint.” One or more of these reference signatures were then subsequently compared to fingerprints generated from unknown signals using the same algorithms, with near or exact matches being used to identifying the signal. *See, e.g.*, Ex-1014 (detailing robust hashing algorithms for generating fingerprints of video and audio content); Ex-1016 (same); Ex-1015 (detailing fingerprinting algorithms for audio content); Ex-1017 (discussing robust hashing to fingerprint an image). Murphy discloses a process for generating a covariance matrix signature that can be used as a fingerprint for a particular segment of content. Ex-1011, 7:28-9:8. A covariance matrix is a matrix wherein each element represents the covariance of two scalar variables. In Murphy, as explained in more detail below, the covariance matrices are generated using feature vectors derived from the filtered frequency bank energy of a series of frames of a segment of the content signal. *Id.* Brunk also includes significant disclosures concerning fingerprinting and fingerprinting techniques. *See, e.g.*, Ex-1012, [0045]-[0050].

74. Fingerprinting and watermarking were known to be alternative technologies usable for the same purposes. More specifically, fingerprinting was a well-known, beneficial substitute for watermarking that would achieve the known, predictable results of time-specific content identification for the purpose of precisely identifying when actions should occur. For example, Haitsma (Ex-1016) taught that one could substitute watermarking with fingerprinting for the specific

purpose of generating actions such as presenting a user with information or a link to a URL. *See* Ex-1016, [0004] (“Using a robust hash to identify multimedia content is an alternative to using watermarking technology for the same purpose.”); [0058].

75. One of ordinary skill in the art would have understood that Haitsma’s “robust hash” is a fingerprint, and that Haitsma’s substitution teaching applied generally to fingerprints, regardless of whether a hash was used. Haitsma states that:

The company Digimarc currently has an application called MediaBridge, which is based on watermarking technology. The idea is that a watermark in a piece of multimedia will direct a user to a certain URL on the Internet where he can get some extra information. . . . [A] watermark detector will extract a watermark key that is sent to a database. This database then contains the URL to which the user will be redirected. The same application can work with the use of robust hashing technology.

Ex-1016, [0058]. Notably, Digimarc was identified as the assignee of Brunk.

76. Similarly, Venkatesan taught that fingerprinting “can be used in place of watermarking, with the benefit that nothing is added to images.” Ex-1017, 666. *See also* Ex-1012, [0045] (“One way of identifying audio and video content—apart from digital watermarks—is fingerprinting technology.”).

4. Algorithms for generating, storing, retrieving, and matching fingerprints for content recognition were well known

77. Prior to the '114 patent, techniques for generating, storing, retrieving, and matching fingerprints to identify a particular portion of audio-visual content were well known. One of ordinary skill in the art would know fingerprinting algorithms for audio-visual content such as a portion of an audio clip, an image, or one or more frames of video that provided representation of perceptual features of the portion of the content in question.

78. In the prior art, such fingerprints were sometimes called “signatures” or “hashes.” *See, e.g.*, Ex-1001, 3:19-20 (explaining that fingerprints are also known as “robust hashes”); Ex-1012 [0005] (“The content signature can be derived through a manipulation (e.g., a transformation, mathematical representation, hash, etc.) of the content data. The resulting content signature may be utilized to identify, track, manage, handle, protect the content, link to additional information and/or associated behavior, and etc. Content signatures are also known as ‘robust hashes’ and ‘fingerprints,’ and are used interchangeably throughout this disclosure.”); Ex-1014, 120-123 (describing “algorithm for robust video fingerprinting” using hashing); Ex-1016, [0027]-[0039] (describing similar hashing fingerprinting methods for audio and visual media); Ex-1017, 664-665 (describing “robust image hash” in context of fingerprinting).

79. One category of methods for generating a fingerprint is known as “robust hashing.” Ex. 1001, 3:19-20; Ex-1012 [0005]. What the term “robust” indicates is the desirable feature that permits reliable automated recognition even if the underlying content undergoes modifications that are generally imperceptible to the viewer or listener, for example, through the normal filtering, compression, transcoding, etc. involved in generating and distributing content. *See, e.g.*, Ex-1001, 3:20-24. The same robustness had the known advantage of mitigating the ability of a bad actor to maliciously prevent automated detection of an illicit copy of content.

80. All of this was well known in the prior art. For example, Oostveen (Ex-1014) describes prior art algorithms for creating a fingerprint disclosed as “very robust against all processing which is done on a local basis, like for instance MPEG compression or median filtering.” Ex-1014, 124; *see also* Ex-1016 (same). Haitsma taught the robustness concept as follows:

Signatures of multimedia content that are to a certain extent invariant to data processing (as long as the processing retains an acceptable quality of the content) are referred to as robust signatures or, which is our preferred naming convention, robust hashes. By using a database of robust hashes and content identifiers, unknown content can be identified, even if it is degraded (e.g. by compression or AD/DA conversion). Robust hashes capture the perceptually essential parts of audio-visual content.

Ex-1016, [0003].

81. Venkatesan (Ex-1017) teaches the robustness concept in similar language in the context of image detection:

[D]igital images commonly undergo various manipulations, such as compression, enhancement, scaling, and cropping, just to name a few. Instead, the hash function should take into account the changes in the visual domain and produce hash values based on an image's visual appearance. In other words, the hash values obtained by applying the hash function to visually similar images should remain the same. For visually distinct images, the hash function should produce different hash values.

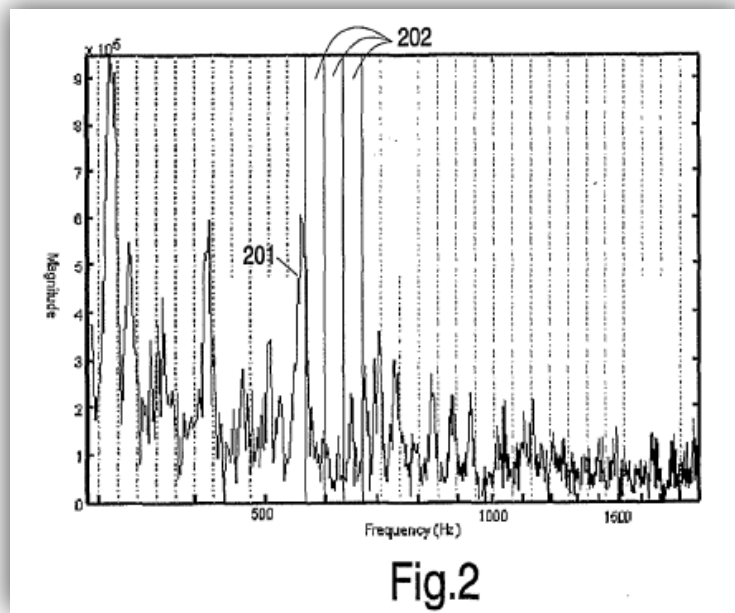
Ex-1017, 1.

82. Venkatesan describes its hash as being robust against “common image processing and malicious attacks,” including “JPEG compression using a quality factor of 10% or more,” “4 x 4 median filtering,” “rotation by up to 2 degrees,” “cropping up to 10% of image area,” “scaling by up to 10%,” “random deletion of up to 5 lines,” and “shifting by up to 5%.” *Id.*, 1, 666.

83. As I discuss below, Murphy (Ex-1011) likewise taught using robust algorithms to reliably identify audio content notwithstanding the types of transcoding, filtering, and similar modifications that change the signal while generally remaining imperceptible to the listener. Ex-1011, 7:28-9:8.

84. Various algorithms were known for generating the fingerprint. A common process involved dividing a media signal into successive, preferably

overlapping, frames or windows. Ex-1016, [0025-28]. The frames were then divided into bands of, for example, frequency:



Id., Fig. 2, [0029].

85. Thereafter, a property such as energy of the band was calculated and a “robust” hash applied. *Id.*, [0030-33]. One or more fingerprints were determined for specific parts of known content, stored in a database, and subsequently compared to the fingerprints of unknown content to identify that portion of content. *Id.*, Fig. 1 (system for fingerprint generation and matching); Fig. 7 (database of fingerprints); Fig. 6 (algorithm for generating fingerprint of an unknown song and matching it to fingerprint in database to identify it); *see also* Ex-1011, 7:28-9:8.

86. Brunk also taught “robust hashing” algorithms for making fingerprints. *See, e.g.*, Ex-1012, [0005], [0030]-[0033].

5. It was known to associate trigger markers with actions using a database

87. It was also known to store triggers (i.e. fingerprints) and associated actions together in a database. For example, Oostveen (Ex-1014) and Haitsma (Ex-1016) taught storing fingerprints in databases for ease of matching during playback. Ex-1014, Abstract, 1-4; Ex-1016, [0003]. Kate (Ex-1013) taught that the trigger markers and the associated actions were stored beforehand, for example, in a database, and multiplexed into a program during broadcast or at an earlier time. Ex-1013, 57. Murphy (Ex-1011) taught using a database to hold signatures of cues generated during pre-processing, as well as related information. Ex-1011, Abstract, 3:32-34, 6:30-7:11, 13:8-14, 14:4-15. Brunk (Ex-1012) specifically taught database structures where actions were stored indexed to fingerprints. Ex-1012, [0006].

C. Murphy

88. UK Patent Application GB 2,375,907 to Murphy (“Murphy,” Ex-1011) was originally published November 27, 2002, and I understand that it therefore qualifies as prior art under 35 U.S.C. § 102(b) (pre-AIA). It describes an “automated recognition system for recognising that a portion of a signal matches a known programme cue associated with a programme of interest contained in the signal,” and further describes taking actions based on the detection of said cues. Ex-1011, 1:4-8; *see also* 12:8-16.

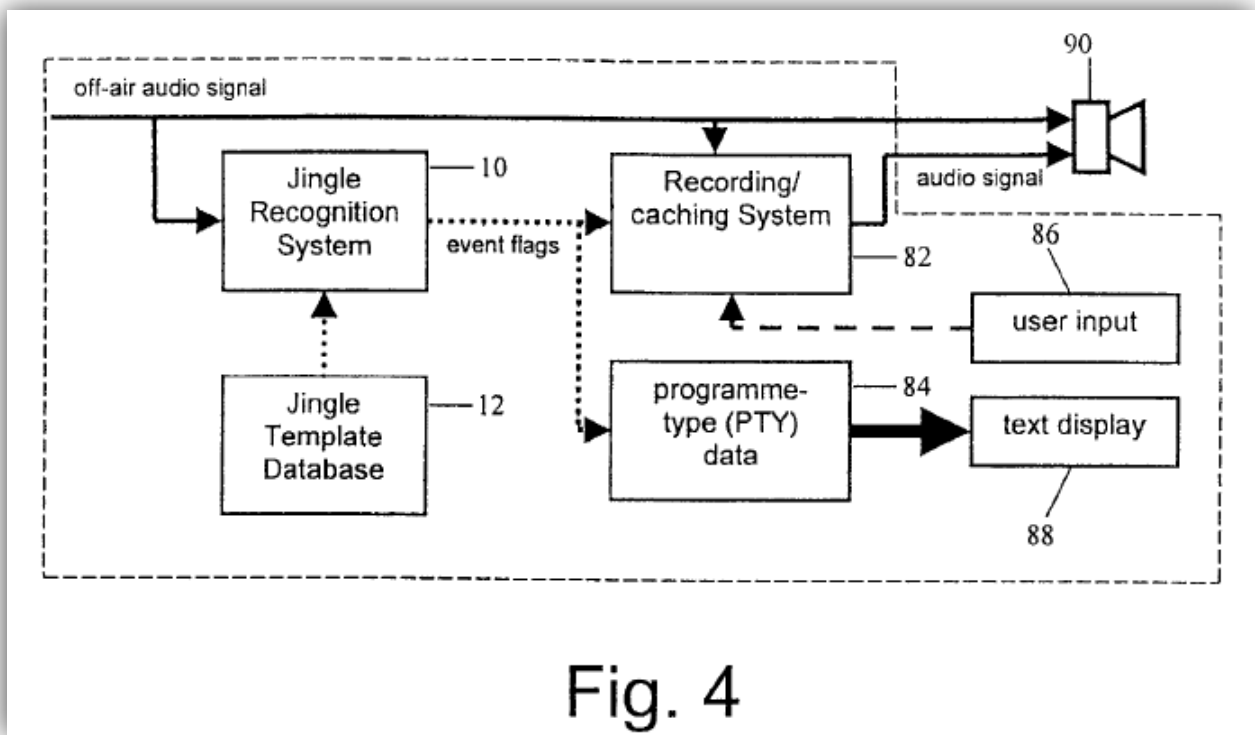
89. Murphy discloses two embodiments: one where the recognition system is situated on the transmitter side and another where the recognition system is on the receiver side. Murphy's disclosure first describes recognition system 10, the component that detects fingerprints, determines matches, and generates event flags to cause the performance of actions, in conjunction with the transmitter side embodiment, and discloses much of the details about how it performs its tasks (such as detecting fingerprints) in that discussion. But it is apparent that Murphy's recognition system 10 operates identically when used in the receiver side embodiment. *Id.*, 13:17-20 (stating, with regard to the receiver embodiment, "[t]he received signal is also fed to the recognition system *where it is processed as described above* and checked for a match with any of the audio cues stored in the template database 12") (emphasis added); *see also, Id.*, 14:33-35 (explaining that combinations among Murphy's features may be used). Accordingly, one of ordinary skill in the art understood that Murphy's disclosures pertaining to the transmitter side embodiment taught the functionality and operation of the receiver-side embodiment as well.

90. In the receiver-side embodiment, the program cues are detected and used to trigger actions at a receiver 80 having a recording/caching system 82 and speaker 90, together serving as a playback device for received content signals. Ex-1011, 13:16-17, Fig. 4. This playback device has a recognition system 10 that is

used to detect “audio cues” in a radio broadcast “audio signal” and match them with pre-stored cues in database 12. Ex-1011, 13:15-26.

91. When receiver 80 finds a match, an event flag is generated that causes one or more of a variety of possible actions to be performed, such as that recording system 82 starts recording and/or programmer-type data 84 displays information on text display 88 (another *action*). *Id.* The displayed information is related to the played audio content. *Id.*, 13:8-14.

92. Fig 4. is reproduced below:



Id., Fig. 4.

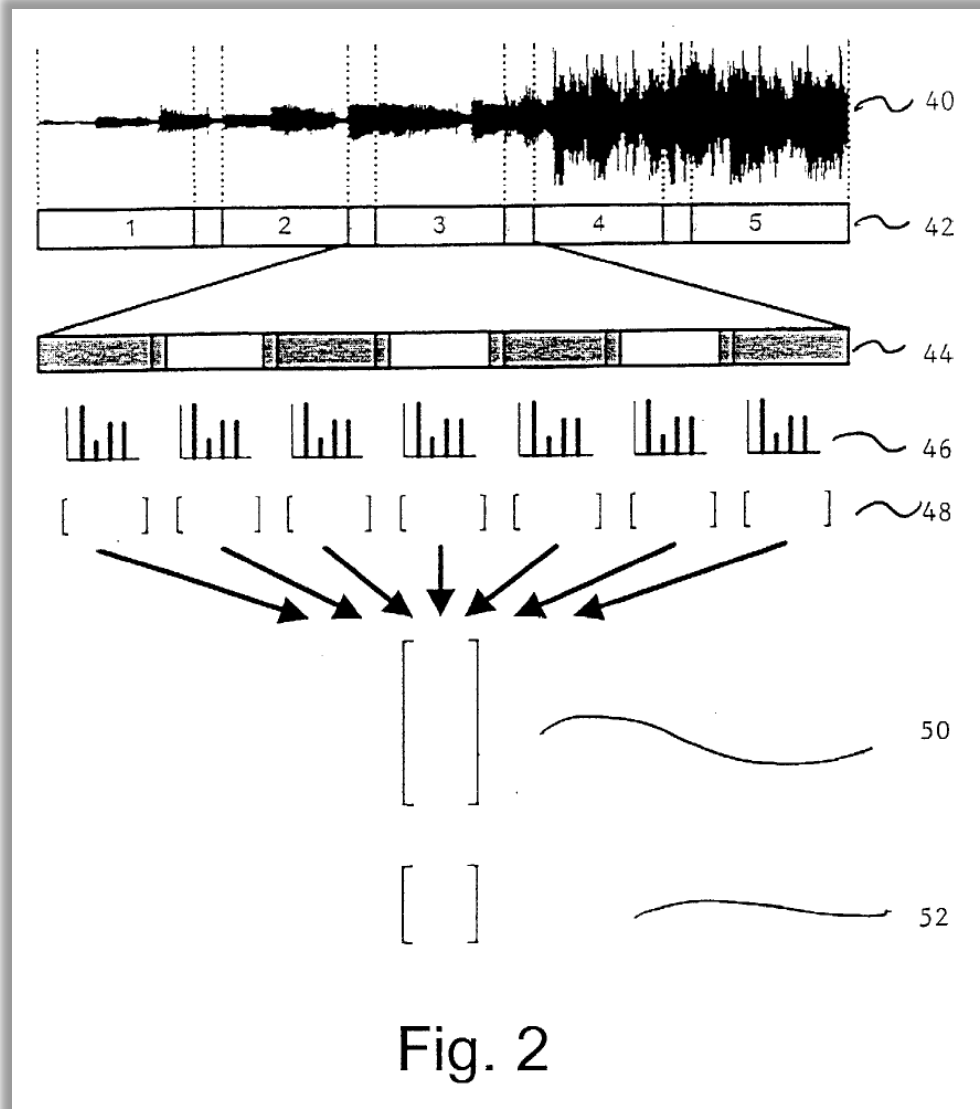
93. The programme cue identifier 10 (referred to in Figure 4 as a “Jingle Recognition System”) comprises (among other components) a library (or database) 12 which is used to store the audio cues that the system is to recognize. *Id.*, 6:27-32. As will be discussed in more detail below, these audio cues are processed to determine their unique covariance matrices that act as a signature (or “fingerprint” in the parlance of the ’114 patent). *Id.*, 6:32-34; 9:5-8. Information concerning the programme is also stored in library 12. *Id.*, 7:9-11.

94. The program cue identifier 10 further comprises a processor 14 that processes the signal 40 received by the recognition system. The signal is first fed into an analogue to digital convertor 20, which converts the received signal 40 to a digital signal at the appropriate sample rate for comparison to the stored cues. *Id.*, 7:17-27. Next, a window generator 22 divides the signal into overlapping windows 42 (or “segments”) of equal length to the standardized length of the stored audio cues. *Id.*, 7:28-8:2. Each window 42 is then further segmented in frames 44 by frame generator 24. *Id.*, 8:9-11.

95. Each frame 44 of each window 42 then undergoes a series of processing steps. First, the frame 44 is fed into a Fast Fourier Transform (FFT) processor 26. The data is then filtered using filter 28 to generate coefficients corresponding to the energy within each frequency bank of each frame. Next, a log processor 30 calculates the logarithm of those coefficients. A frame feature

vector is assembled and matrix generator 32 assembles each of the frame feature vectors for the frames in a particular window into matrix 50. Covariance processor 34 converts the frame feature vector matrix 50 into a covariance matrix 52, which is a representation of the data signal window and may be used as a “fingerprint.” *Id.*, 8:9-9:6.

96. This process for determining covariance matrixes of a particular received signal is depicted in Figure 2:

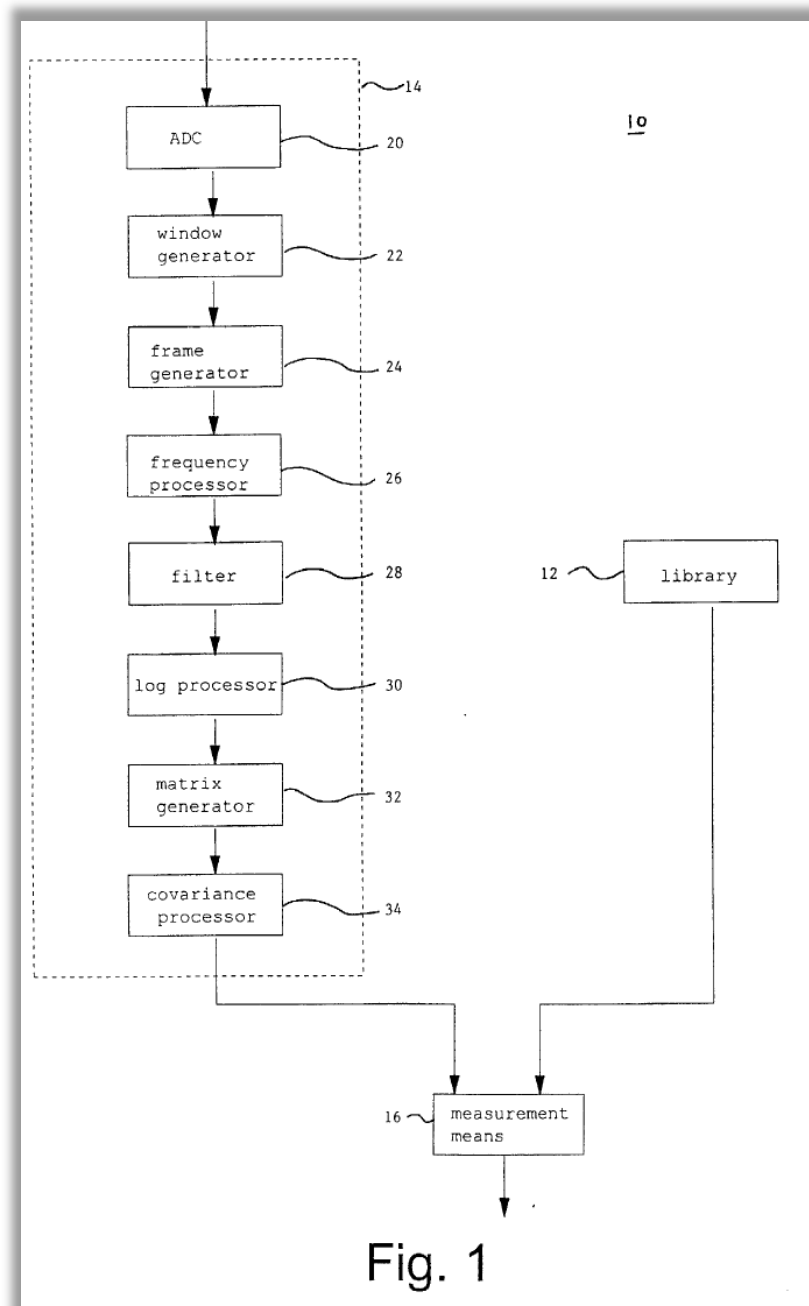


Id., Fig. 2.

97. Because the covariance matrix derived from a window of the received signal and the stored covariance matrix for the audio cue might differ (due to noise or other imperfections), Murphy further discloses using a method of measuring the difference and determining a match has been detected when the difference is below a certain threshold. *Id.*, 9:9-17. Murphy discloses that this process produces

suitable signatures for recognizing audio-cue content of the media signal 40. *Id.*, 9:3-4.

98. Fig. 1 of Murphy depicts the programme cue identifier 10 and its constituent parts, as described above:



Id., Fig. 1

99. When a match in the audio cues is detected, an event flag is generated that causes one or more of a variety of possible actions to be performed, such as recording system 82 starts recording and/or programmer-type data 84 displays information on text display 88. *Id.*, 13:22-26. The displayed information is related to the played audio content. *Id.*, 13:8-14. Murphy also discusses event flags being used in conjunction with the embodiment where the recognition system generates event flags at the transmitter end of the content signal transmission, including setting programme data regarding the content being played or changing the channel. *See, e.g., id.*, 10:21-31, 11:9-15. These actions, including updating the “programme type data” and scheduling information data and triggering automatic recording, are also disclosed as available actions in connection with the Figure 4 receiver embodiment. *Id.*, 13:21-26. Further, there is no functional limitation that would be apparent to one of ordinary skill in the art that would make any of the event flag actions depicted in Fig. 3 only operable if used in a transmitter-side implementation. Accordingly, it would have been apparent to one of ordinary skill in the art that the functionality of recognition system 10 regarding event flags as displayed in Fig. 3 also existed in recognition system 10 as shown in Fig. 4.

100. Murphy also discloses that a time delay can be used to allow for the Murphy system to account for possible delays in the processing of audio cues. *Id.*,

12:17-28. With this functionality, the fingerprints can be detected in a live signal and trigger a recording operation for the content relayed in a delayed signal, thereby ensuring that the recording begins at the appropriate time in relation to the content. *Id.* Figure 3, which depicts an embodiment of Murphy's recognition system being employed on the transmitter side, shows this delayed signal:

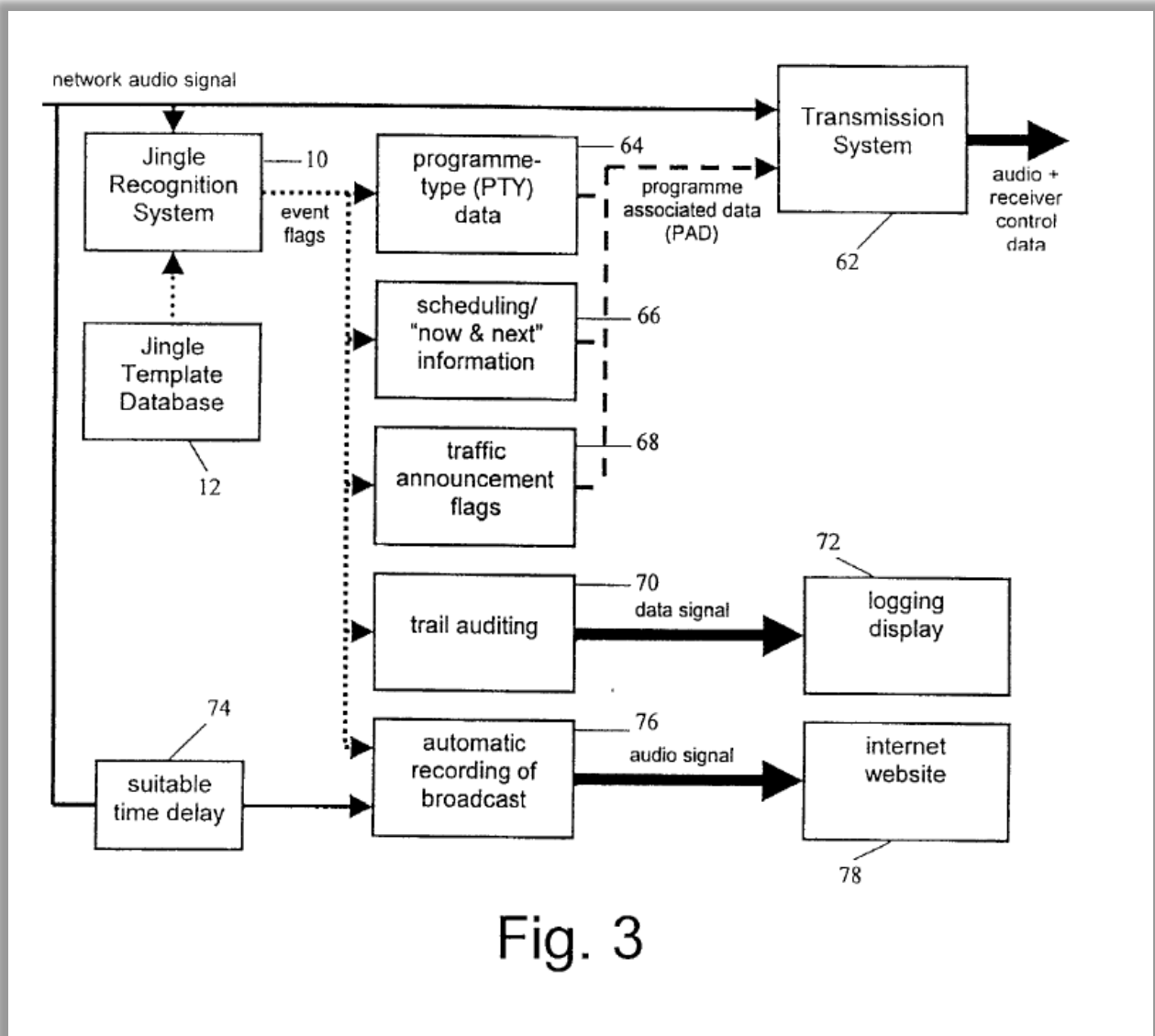


Fig. 3

Id., Fig. 3. Because Murphy’s receiver-side embodiment depicted in Figure 4 likewise involves triggered recording of the content signal, it is apparent from Murphy’s disclosure to one of ordinary skill in the art that the same time delay concept could be employed with regard to that receiver-side embodiment.

101. Although Murphy is focused mostly on audio signals, it also teaches that the system can be implemented using multimedia signals, such as the audio track in a television signal, as well as for video. *Id.*, 15:22-25.

D. Brunk

102. U.S. Patent Application Publication No 2002/0126872 to Brunk et al. (“Brunk,” Ex-1012) was published in September 2002. It is therefore my understanding that Brunk is prior art under at least pre-AIA 35 U.S.C. § 102(b).

103. Like Murphy, Brunk describes a system for generating and detecting “content signatures,” which serve as “fingerprints,” using the parlance of the ’114 patent. Ex-1012, [0020]. As explained in more detail below, Brunk’s disclosure includes significant overlap with Murphy.

104. Brunk teaches that a content signature can be derived from an input signal by dividing the signal into segments, feeding those segments into a frequency transform (such as a Fourier transform domain), and then using the perceptual features of the segments to identify markers for use as signatures. *Id.*,

[0022]-[0023], [0026]. The content signature can be used to uniquely identify the segment or the entire content item. *Id.*, [0026].

105. Brunk further discloses a number of specific hashes that can be used to generate content signatures. For example, Brunk discusses using a generalized Hough transform to generate a signature from either audio or video content. *Id.*, [0025]. Brunk also explains that known hashing algorithms, including MD5, MD2, SHA, and SHA1, can be used to generate content signatures. *Id.*, [0021]. Brunk explains that such hashes can be used to create signatures that are robust. *Id.*, [0033].

106. In one embodiment, Brunk discusses storing content signatures for a particular content item in a database that may be local or remotely accessed. *Id.*, [0028]. The database may also be used to store data that is linked to the signature, including data that specifies actions to be taken. *Id.* Such actions include, for example, providing a URL, providing licensing information, linking to another computer, and/or verifying content and access.” *Id.*, [0028]-[0029]. The stored signature is used as an index to access additional data, including any actions to be performed as a result of matching signatures. *Id.*, [0006].

107. Brunk further describes the use of cue signatures for detecting content signatures during playback. *Id.*, [0066]. A cue signal informs the system that it

should start calculating a content signature in media that is being broadcast. *Id.* [0068].

108. Like Murphy, Brunk's signature recognition and database processing is implemented on computer hardware, including a processor. *Id.*, [0107].

VII. UNPATENTABILITY ANALYSIS

109. I understand that a patent claim is anticipated when a single piece of prior art describes every element of the claimed invention, either expressly or inherently, and is arranged in the same way as in the claim. For inherent anticipation to be found, it is required that the missing descriptive material is necessarily present in the prior art. I understand that, for the purpose of an *inter partes* review, prior art that anticipates a claim can include both patents and printed publications from anywhere in the world.

110. I understand that a patent claim is unpatentable and invalid if the subject matter of the claim as a whole would have been obvious to a person of ordinary skill in the art of the claimed subject matter as of the time of the invention at issue. I understand that, when assessing the obviousness of claimed subject matter, the following factors are evaluated: (1) the scope and content of the prior art; (2) the difference or differences between each claim of the patent and the prior art; and (3) the level of ordinary skill in the art at the time the patent was filed. Different from anticipation, which I understand allows for consideration of only

one item of prior art, I understand that claimed subject matter may be obvious in view of more than one item of prior art. I understand, however, that it is not enough to show simply that all the limitations of the claimed subject matter are spread throughout the prior art. Instead, for claimed subject matter to be obvious over multiple references, there must be some reason or motivation for one of ordinary skill in the art to combine the prior art references to arrive at the claimed subject matter. I also understand that one of ordinary skill in the art must have had a reasonable expectation of success when combining references for claimed subject matter to be obvious. Moreover, I have been informed and I understand that factors referred to as “objective indicia of non-obviousness” or “secondary considerations” are also to be considered when assessing obviousness. I understand that these factors can include: (1) commercial success; (2) long-felt but unresolved needs; (3) copying of the invention by others in the field; (4) initial expressions of disbelief by experts in the field; (5) failure of others to solve the problem the claimed subject matter solved; and (6) unexpected results. I also understand that evidence of objective indicia of non-obviousness must be commensurate in scope with the claimed subject matter.

111. I am aware of no secondary indicia of non-obviousness that might impact the validity of the '114 patent, but reserve the right to address any that are raised by the patent owner at a later date.

A. Detailed Claim Mapping

112. In the below sections, I explain my opinions as to how Murphy, Brunk, and the general knowledge of one of ordinary skill in the art render the claims obvious.

113. In general, Murphy teaches nearly every element of the '114 patent. Murphy, however, lacks an express disclosure teaching that the reference fingerprints and their corresponding actions are both stored in the same database (as discussed in more detail with regard to claim [1.b], below). Brunk does provide that teaching, however, as explained below.

114. For the purposes of my opinion set forth herein, Murphy serves as the primary reference, with Brunk as a supporting secondary reference. That said, Brunk itself discloses many of the limitations of the '114 patent. Accordingly, for several of the limitations below, I explain how both Murphy and Brunk each disclose relevant material that teaches the limitation. This is done both to show that the proffered grounds teach the limitation, but also to demonstrate the significant degree of overlap between the teachings of Murphy and Brunk. This significant overlap is a factor that supports my overall opinion that it would have been obvious for one of ordinary skill in the art to combine the teachings of Murphy and Brunk.

1. Claim 1

[1.preamble] A method comprising:

115. Murphy discloses a method. Murphy discloses at length the method by which its disclosed audio cue recognition system can be employed, including by disclosing and discussing the steps depicted in Figs. 1-4. Murphy recites several method claims. *See* Ex.-1011, 19:9-20:2. Thus, one of ordinary skill in the art would understand that Murphy's disclosure of its system also describes a method for designing and employing that system.

116. Brunk also discloses a method. Ex-1012, [0020].

[1.a] determining a plurality of trigger fingerprints from content being played back on a playback device, each trigger fingerprint identifying a corresponding trigger time point of a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered, the determining being performed by a processor;

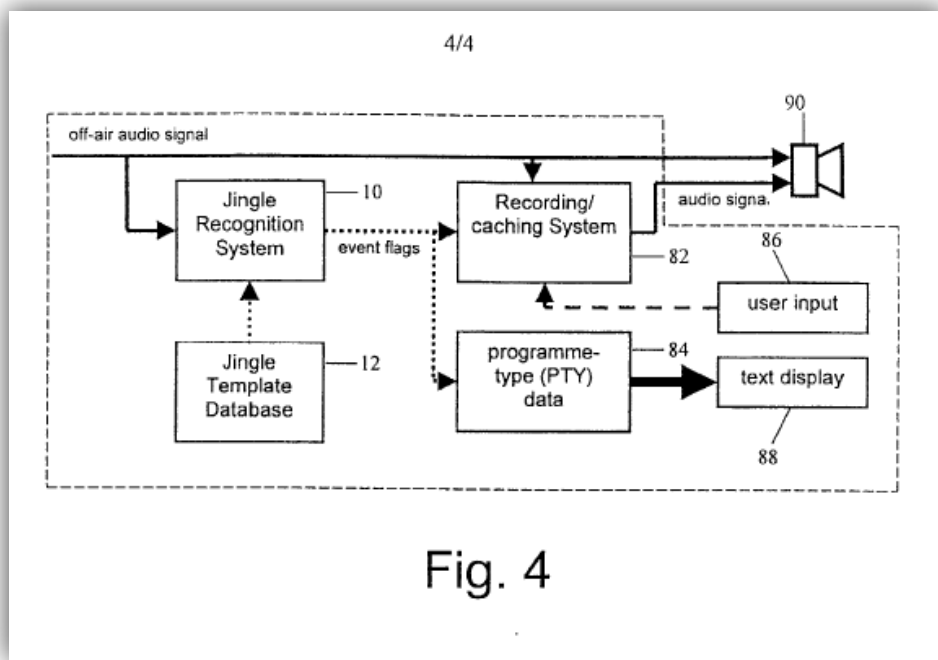
117. Murphy teaches these limitations.

118. My discussion below breaks this lengthy limitation up into parts.

First, I discuss Murphy's disclosure of "determining ... trigger fingerprints" as "performed by a processor." **Second**, I discuss the "content being played on a playback device" limitation in more detail. **Third**, I discuss Murphy's disclosure of "fingerprints." **Fourth**, I discuss Murphy's teachings regarding a "plurality of trigger fingerprints that each identify corresponding trigger time points in a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered."

119. **First**, in my opinion, Murphy discloses a processor that “determin[es]” (or detects) “fingerprints from content being played on a playback device.” In Murphy, during playback, processor 14 of recognition system 10 continually generates signatures (determining fingerprints) of the signal as it plays. Ex-1011, 6:24-30 (“The recognition system is implemented on a processor or as a program is a computer. . . . The programme cue identifier 10 is shown in Figure 1 and comprises ... processor 14”), 5:9-17, 6:33-34. Within processor 14, Murphy further discloses a specific sub-processor responsible for determining the specific exemplary embodiment of a fingerprint in the form of a covariance matrix. *Id.*, Abstract (“A processor 34 forms a covariance matrix which is compared with those in the library 12.”), 9:5-7 (“The covariance matrix of the audio cue is determined ...”). The process of determining the fingerprints from the signal is shown in Figure 2, using the components of processor 14 shown in Figure 1. *Id.*, 7:13-9:31, Figs. 1, 2; *see also* Section VI.C (overview of Murphy).

120. **Second**, in my opinion, Murphy teaches “content being played back on a playback device.” Specifically, Murphy teaches a playback device in the form of receiver 80, which includes recording/caching system 82 and loudspeaker 90. As Murphy explains, “receiver 80 receives a broadcast audio signal which is outputted to the loudspeaker 90.” Ex-1011, 13:16-17.



Id., Fig. 4. For context, the '114 patent similarly describes “content” as an audio “and/or” video signal and “television viewing and/or music listening” as examples of using a playback device. *See* Ex-1001, 1:22-23, 5:21-25, 6:59-62, 7:25-33, claim 19.

121. **Third**, in my opinion, Murphy also teaches a “fingerprint” as claimed. Murphy uses an identifying “signature” (i.e., a fingerprint) of a cue in the audio signal: “the audio cue is pre-processed to determine an individual determining characteristic, or signature.” Ex-1011, 6:33-34. Murphy’s disclosure broadly encompasses various methods for generating the signature: “Any method of characterising a signal which makes it discernable from a different signal could be used to detect the presence of an audio cue in an audio signal.” *Id.*, 5:9-14.

Murphy teaches that the signature may be generated based on features “discernable from other cues,” for example, “pitch, tempo, or a statistical characterization” (*id.*, 5:9-17), so long as the signature still identifies an unknown version of the signal that has undergone typical signal transformations:

[I]t is not sufficient to simply compare sample values of the audio signal with a stored version of the audio cue because such a system would fail to operate adequately in the presence of changes in amplitude, the use of audio bitrate reduction or sample by sample synchronisation issues between the stored audio cue and the detected audio signal from the programme audio feed.

Id., 5:17-24. A person of ordinary skill in the art would have understood Murphy to be describing the desirable feature that the best fingerprints are “robust.” *See, e.g.*, Ex-1012, [0033] (contrasting “robust” fingerprints from “fragile” fingerprints).

122. Like the ’114 patent, Murphy discusses drawbacks of prior art techniques of adding identifying data to a signal (*i.e.*, watermarking) (Ex-1011, 2:30-3:9), and contrasts that with the benefits of using fingerprinting:

The system does not require the broadcaster to adapt the signal to include additional data which can be detected and used to control the recording system 82 nor does it impose any particular information format onto broadcasters; the added functionality is generated entirely by the receiver using characteristics of the transmitted signal to deliver to the user additional functionality. Thus the receiver 80 may be used to record programmes and indicate information about programmes

from broadcasters who are not equipped to transmit this information or who transmit such information in a format which is not supported by the receiver 80.

Ex-1011, 14:18-31. Murphy goes on to discuss various methods known in the art for generating fingerprints (*id.*, 5:11-6:22) before providing a detailed example of its fingerprint of choice, a covariance matrix (*id.*, 6:23-9:31, Figs. 1 and 2).

123. In the preferred embodiment, Murphy teaches that the exemplary covariance matrix is a suitable signature:

The covariance matrix is unique to the particular window of data signal and can be compared to the covariance matrix of known audio cues to find a match, thereby identifying the window of data as the audio cue. The covariance matrix is therefore a suitable signature of the signal which may be used for recognition purposes.

Ex-1011, 8:35-9:4. Murphy thus teaches a “fingerprint” in at least two ways: *first*, the concept of processing a portion of a signal to develop a signature representation thereof (Ex-1011, 5:9-24, 6:33-34), and *second*, its specific example of a covariance matrix in a preferred embodiment (*id.*, 6:23-9:31, Figs. 1 and 2).

124. Furthermore, based on these citations, a person of ordinary skill in the art would understand Murphy to teach even the preferred embodiment of the fingerprints of the ’114 patent, together with the benefits of fingerprinting that the inventors purportedly first disclosed. Specifically, Murphy broadly teaches the use of a fingerprint that is a representation of perceptual features of the part of the

content or signal in question. And a person of ordinary skill in the art would have understood Murphy to disclose using a fingerprint of a piece of audio that is an identifier which is computed over that piece of audio and which does not change even if the content involved is subsequently transcoded, filtered, or otherwise modified during the distribution process. *Id.* In other words, one of ordinary skill in the art would have understood Murphy to teach a fingerprint that is “robust.”

125. In addition to the fingerprinting methods disclosed in Murphy, it would have been obvious to a person of ordinary skill in the art that Murphy’s fingerprint recognition system could be implemented using any of the “robust hash” algorithms known in the art to reliably identify audio, video, or combined audio and video signals, including the robust hash taught by Brunk (Ex-1012).

126. Specifically, Murphy states that “[a]ny method of characterizing a signal which makes it discernable from a different signal could be used to detect the presence of an audio cue in an audio signal.” Ex-1011, 5:11-14. And Murphy further acknowledges the need for a fingerprint that is sufficiently resistant to signal modification. *Id.*, 5:17-24. One of ordinary skill in the art would understand Murphy to be explaining that it is desirable to have a fingerprint that is robust. Given this teaching, it would have been obvious to one of ordinary skill in the art to accept Murphy’s invitation to use alternative robust fingerprinting methods, like those discussed in Brunk. Ex-1012, [0025], [0042]-[0044].

127. And **fourth**, Murphy also teaches a “plurality of trigger fingerprints that each identify corresponding trigger time points in a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered.”

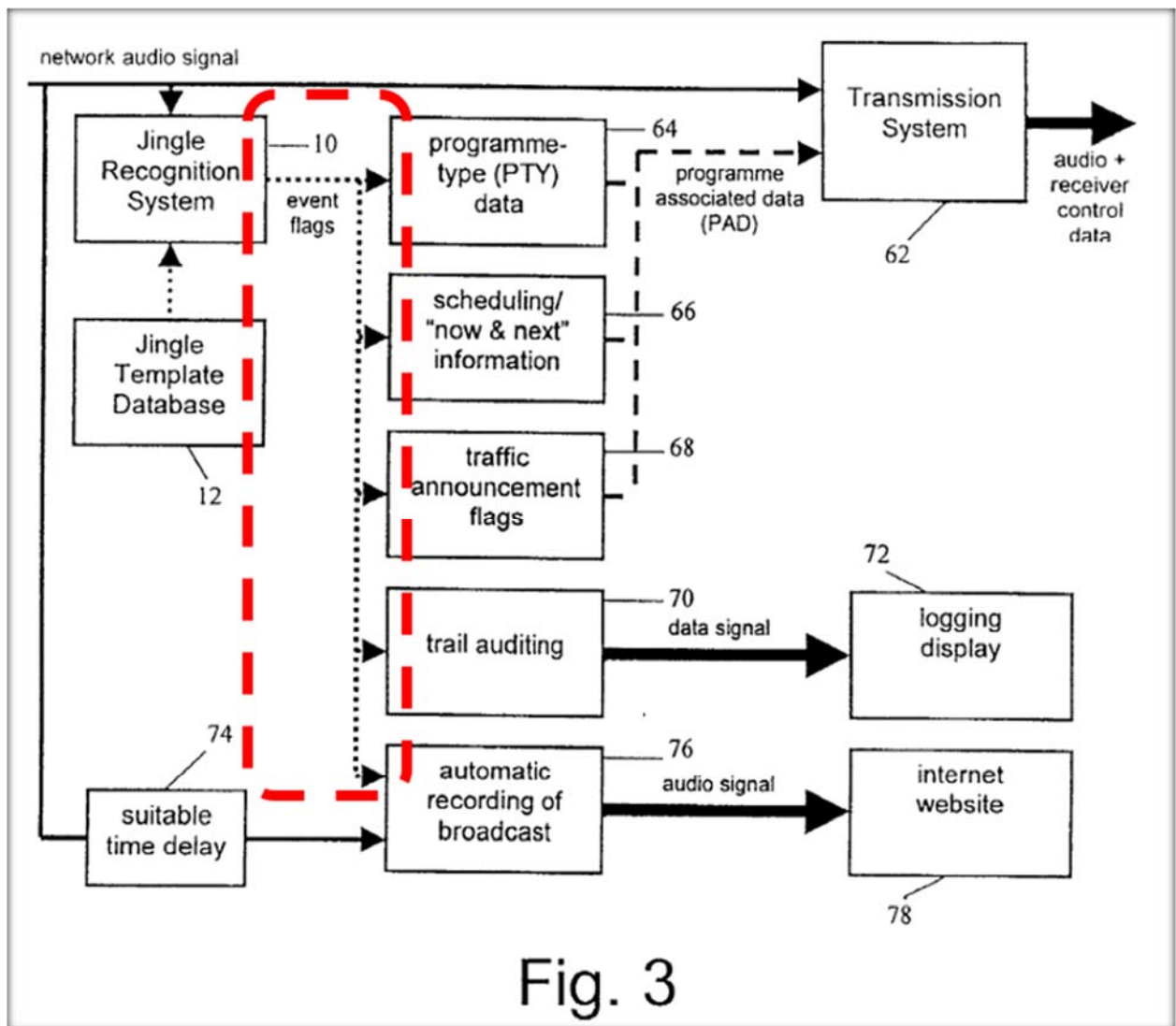
128. To track the significance of the plurality of possible fingerprints, Murphy teaches using a database:

A database of information on the given programmes and the programme cues associated with the given programmes may be created by the user or by a third party and used by the receiver with the recognition system to display information on the current programme.

Ex-1011, 4:16-23 (emphasis added).

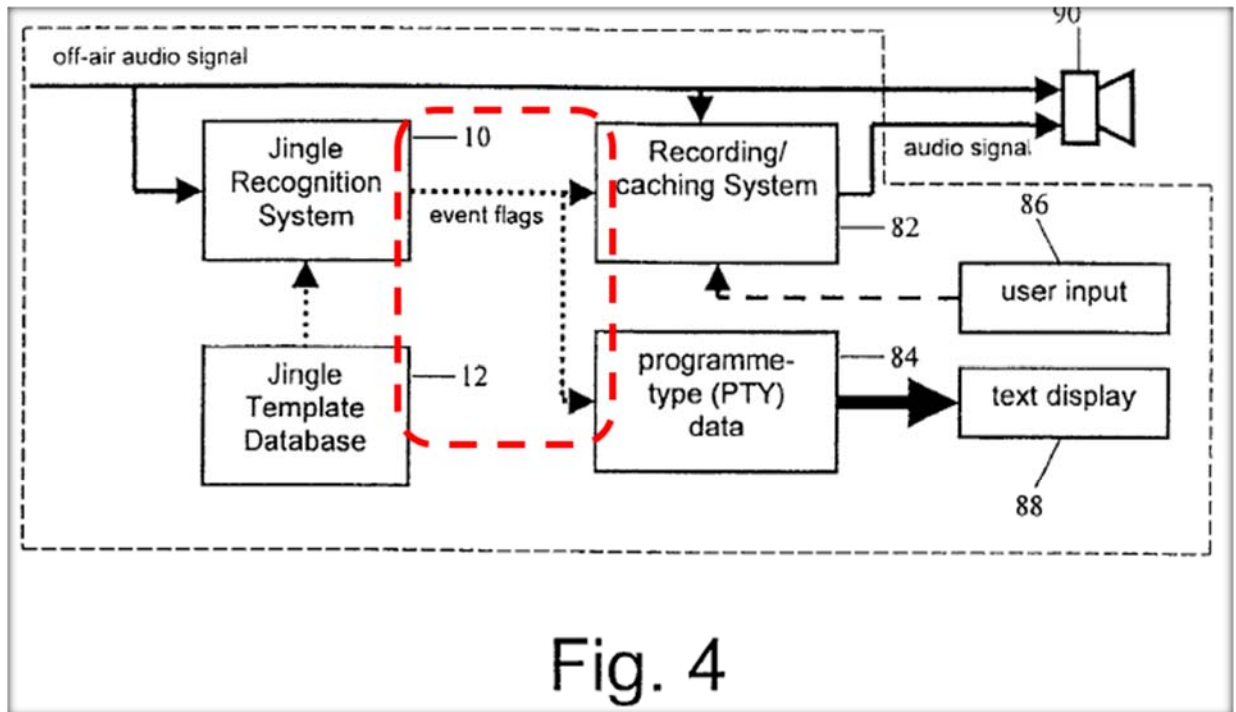
129. When Murphy recognizes the fingerprint of a particular audio cue (*see infra*, claim [1.c-d]), Murphy triggers an action using an action-specific control signal (“event flag”) that initiates corresponding actions, thereby using fingerprints as trigger markers in the same context as the ’114 patent. Ex-1011, 10:1-2, 13:21-26.

130. Taking Murphy’s disclosure as a whole, Murphy broadly teaches associating any number of desired actions upon recognition of a fingerprint. For example, in both Figures 3 and 4, Murphy teaches that fingerprints are used to generate various action-specific “event flags” triggering actions:



Ex-1011, Fig. 3 (annotated), 10:1-12:29 (including fingerprints that initiate actions to (a) identify the start of a particular type of program, such as traffic announcements, (b) generate scheduling information, such as identifying the start and end of programs, (c) identify program type data, (d) update the now and next scheduling information for an electronic program guide (EPG), (e) identify the artist and title information of a song currently being played so that it can be displayed on a display, (f) generate an audit trail of if and when programs were broadcast,

including advertisements and news bulletins, (g) control automated recording systems, and (h) cause the playback device to change channels to receive a different content signal).



Ex-1011, Fig. 4 (annotated), 10:1-2, 13:8-14:3. Although Murphy's disclosure regarding actions and event flags is somewhat more detailed in Murphy's discussion of the transmitter-side embodiment depicted in Figure 3, one of ordinary skill in the art would have understood that the same "jingle recognition system 10" and same "event flags" appear in the receiver-side embodiment depicted in Figure 4 and could be employed to achieve the same functionality in the receiver-side embodiment (which, with speaker 90, forms a playback device) depicted in Figure 4.

131. One such action that Murphy teaches is using fingerprints to display information about programs: “The recognition system may also be used in receivers to display information on a given programme even when the broadcast signal does not include such information.” Ex-1011, 4:25-27; 13:8-14. Murphy teaches that the action is triggered upon detection of the cue. For example, when the fingerprint recognition system detects a match with a particular fingerprint, the system generates a “programme-type data (PTY)” flag that causes text to be displayed on text display 88. *Id.*, 13:22-26.

132. As another example of “identifying a corresponding trigger time point of a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered,” Murphy further teaches controlling a recording system using fingerprint recognition. For example, when Murphy detects a particular fingerprint, the recognition system generates event flags to cause recording system 82 to start or stop recording the signal at specific times marked by the location of the cue in the signal:

- “**When a match** with a particular audio cue **is detected** an event flag is set. The event flag may be a **programme start flag** which, when detected by the recording system 82, **causes the recording system 82 to start** recording the signal ...” Ex-1011, 13:20-26;
- “recording system 82 would be **activated when** it detected an event flag.” *id.*, 13:33-34;

- “recording system 82 may look for an **end flag** to **stop** the recording.” *id.*, 14:2-3.

133. Based on this disclosure, a person of ordinary skill in the art would have understood that the recording system 82 may “look for an end flag” (*id.*) that the receiver generated from an additional fingerprint present in the signal at the end of the program, which is either the second instance of the earlier cue or a specific “play out cue.” For example, Murphy teaches both that “an end flag could be generated on the repeated occurrence of the audio cue” or, alternatively, “on the occurrence of a second, play-out, audio cue.” *Id.*, 11:2-4. Murphy’s reference to end flags at the receiver (*id.*, 14:3) teaches such flags generated by the receiver’s recognition system. *Id.*, 14:18-31, 11:2-4. The locations of these end flags in the content, combined with the locations of the start recording event flags, constitute a “plurality of trigger time points.”

134. Thus, Murphy teaches that the content signal can contain multiple signatures, and each signature (trigger fingerprint) identifies a corresponding trigger time point when one or more corresponding action—such as display information, start recording, stop recording—is to occur. The trigger time point is referenced to the location of the cue in the signal, for example, the time at the end of the cue when it is processed and recognized by the recognition system. *See also, infra*, at claims [3], [4], and [5] (discussing how Murphy teaches and renders obvious various

permutations of the reference time being relative to the start, end, or some offset from the location of the fingerprint in the signal).

135. Put another way, Murphy teaches a “plurality of trigger fingerprints that each identify corresponding trigger time points in a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered” because Murphy teaches using two trigger fingerprints (a plurality) to trigger a start time followed by an end time (both trigger time points in a sequence of trigger time points), which respectively trigger the corresponding action of starting and stopping recording (corresponding actions). Ex-1011, 13:22-25, 14:17-20. This is the precise way one of ordinary skill in the art would have understood Murphy to permit the recording of particular “programmes.”

136. Murphy also teaches a “plurality of trigger fingerprints that each identify corresponding trigger time points in a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered” in the broader context of taking actions in connection with multiple different “programmes” over a longer period of time. Specifically, Murphy also teaches a plurality of trigger fingerprints that each identify corresponding trigger time points and actions in its discussion of the content of the radio signal in Murphy’s preferred embodiment, which is described as containing multiple programs with multiple parts (for example, songs (*id.*, 11:9-15), news bulletins (*id.*, 11:18-19),

advertisements (*id.*), traffic announcements (*id.*, 11:21-24), etc... . Thus, a person of ordinary skill in the art would understand that the content played back at Murphy's receiver would comprise scenarios, for example, an hour-long broadcast having multiple cues used to display information and indicate the starting and stopping of portions of the broadcast. Such a scenario constitutes a sequence of fingerprints and actions as claimed: start recording, followed by sequentially triggered updated display information (for example, artist and title of current song, current type of program, up next information, etc.), followed by a stop recording action. *Id.* One of ordinary skill in the art would readily use Murphy's fingerprint recognition system to allow dynamically identified and displayed information at the relevant times for each section, regardless of whether the program departed from the planned time and makeup, and regardless of the broadcaster's transmission capabilities. *Id.*, 14:21-31 (discussing ability to control recordings and display information without information from broadcaster).

137. Murphy teaches a plurality of types of actions (start recording, stop recording, display information, etc.). It would be implied or obvious to one of ordinary skill in the art to use Murphy's system to arrange any desirable sequence of such actions for a given program or programs to enhance the user experience according to well-known use cases motivated by what the applicant admitted to be an industry trend in the prior art. Ex-1001, 1:22-2:34, and Ex-1011, 5, 51-52, 62-

64. For example, one of ordinary skill in the art could easily configure the system of Murphy to recognize particular audio cues indicating that any of a number of songs by a particular artist was starting to play, and utilize Murphy's fingerprinting system to trigger every song by that artist that played over a series of hours, while not recording songs by other artist.

138. The same theories are equally applicable to Murphy's teaching of implementing its system for television and other video content. Ex-1011, 15:3-25.

139. In my opinion, Brunk teaches most of this limitation as well.

140. **First**, in my opinion, Brunk likewise describes determining the presence of fingerprints during playback. Specifically, Brunk describes software executing on a general-purpose computer, which a person of ordinary skill in the art would understand to be a programmed processor, operating to determine signatures (fingerprints) for content items. Ex-1012, [0029], [0032], [0107]. Brunk discloses making these determinations during playback using cue signals in the played back content to signal the system to begin content signature calculation. *Id.*, [0067]-[0068]. Once the content signature (fingerprint) is recalculated, it is compared to the stored signatures in a database to determine if there is a match. *Id.*, [0083].

141. **Second**, in my opinion, Brunk likewise teaches that content is played back on a playback device. Brunk teaches that software executing on a user

device, such as a computer, PVR, MP3 player, or radio will receive a stream of signatures. Ex-1012, [0029]. A person of ordinary skill in the art would have understood that this user device serves as the claimed “playback device” and is used to playback the content relating to the stream of signatures (which serve as the claimed “fingerprint”). Moreover, Brunk further teaches that a broadcast monitoring application monitoring, for example, a television (together forming a playback device) can use cue signals to perform fingerprint detection during playback of the content on content on the television playback device. *Id.*, [0066]-[0068].

142. **Third**, in my opinion, Brunk also independently teaches the concept of a fingerprint. Specifically, Brunk discloses deriving a statistically unique identifying signature (fingerprint) by applying a hash to a segment of a digital content signal. Ex-1012, [0020]-[0021]. Brunk explains at length how a signal can be segmented and transformed to highlight the segment’s relevant perceptual features, which may be used to form a robust fingerprint. *Id.*, [0022]-[0024]. Brunk further teaches alternative ways to derive fingerprints, including the use of a Hough transform and alternatives for use in images or video content. *Id.*, [0025], [0042]-[0044].

143. In my opinion, Brunk also teaches plurality of fingerprints used to trigger a corresponding action. Brunk discloses storing signatures (which are

equivalent to the '114 patent's claimed fingerprints) in a database along with additional data, including data specifying actions or behavior, such as providing a URL or linking to a website that returns information to the user device. Ex-1012, [0028]-[0029].

144. Brunk differs from the '114 patent and Murphy in that Brunk does not contain an express disclosure of using cue signals to allow for detection of fingerprints to trigger actions at specific times based on the signatures. In other words, Brunk does not expressly disclose using fingerprints to identify a trigger time point of a sequence of trigger time points in the content for triggering an action. *Id.* [0066]-[0068]. But as explained above, Murphy does disclose this aspect of the claims.

145. Besides Murphy and Brunk, it is also my opinion that the '114 patent admits these limitations were known. Specifically, the '114 patent acknowledges that the subject matter recited in these limitations consists essentially of using known fingerprints instead of equally well-known watermarks for performing the known time-marker function. Ex-1001, 2:30-35; 3:7-16.

146. Finally, the '114 patent purportedly invented the use of a fingerprint to perform the time marker function that was previously performed in the industry by embedding data or watermarks at the desired points of time. That invention was not novel. A person of ordinary skill in the art was well aware of the use cases of

generating actions at a sequence of points in content. *Supra* Sections IV.B.1-2 (discussing, for example, applicant's admitted prior art and Kate (Ex-1013, 51-52, 55, 62-64)). Moreover, fingerprints were already known as a beneficial substitute for watermarks to provide the time marking function. *Supra* Section IV.B.3. Thus, even before turning to Murphy or Brunk, a person of ordinary skill in the art would have been well versed in the claimed subject matter.

[1.b] accessing a database that includes a plurality of reference fingerprints, previously derived from the content, and a plurality of reference actions, each reference fingerprint identifying at least one corresponding reference action to be performed at a reference time point identified by the reference fingerprint;

147. In my opinion, Murphy in view of Brunk teaches these limitations of claim [1.b]. Murphy discloses a database 12 that holds pre-processed signatures:

“The programme cue identifier 10 is shown in Figure 1 and comprises *a database*, or library, **12** a processor 14 and a distance measurement processor 16. The *database 12 stores information on the audio cues* which the system is required to recognise. Rather than store the raw audio cue signal, **the audio cue is pre-processed to determine an individual determining characteristic, or signature**. This helps to minimise the amount [sic] real-time data processing performed by the recognition system.”

Ex-1011, 6:27-7:1 (emphasis added); *see also* 8:30-9:8 (including “The covariance matrix is stored as the signature of the audio cue.”). One of ordinary skill in the art

would have understood these pre-processed signatures to be “reference fingerprints” as disclosed by the ’114 patent.

148. Murphy teaches taking actions based on the detection of a fingerprint. Specifically, as explained above with regard to claim limitations [1.a], Murphy explains how “event flags” are generated at particular times during playback of the content when a match is detected. *See, e.g., id.*, 11:26-29 (discussing use of event flags to indicate “when” particular programs are being played). One of ordinary skill in the art would have understood that Murphy’s event flags can be used as commands to trigger several actions, including, for example, controlling the recording system 76 to start and stop recording at certain times. *Id.*, 12:8-14. Murphy also explains that event flags may be used to take other actions, including displaying information, at a time referenced to the location of the cues in the signal, for example, the time the cue is recognized. *See supra* claim [1.a]. *See* Ex-1011, Fig. 3 (annotated), 10:1-12:29 (including fingerprints that initiate actions to (a) identify the start of a particular type of program, such as traffic announcements, (b) generate scheduling information, such as identifying the start and end of programs, (c) identify program-type data, (d) update the now and next scheduling information for an electronic program guide (EPG), (e) identify for display the artist and title information of a song currently being played, (f) generate an audit trail of if and when programs were broadcast, including advertisements and news bulletins, (g) control automated

recording systems, and (h) cause the playback device to change channels to receive a different content signal).

149. While Murphy discloses that signatures are associated with corresponding actions, as explained above, Murphy does not expressly disclose that the corresponding actions are stored in the same database as the reference fingerprints. Brunk, however, does expressly teach that signatures (fingerprints) and their corresponding action can be stored in the same database. Specifically, Brunk states:

Content signatures can be stored and used for identification of the content item **A content signature may even be utilized to index** (or otherwise be linked to data in) **a related *database***. In this manner, **a content signature is utilized to access additional data, such as ... a desired *action* or behavior**, and validating data.

Ex-1012, [0006] (emphasis added), [0028] (*including* “When stored in a database, a signature may be linked or associated with additional data. Additional data may include ... data specifying actions or behavior (e.g., providing a URL, licensing information or rights, etc.)”), [0029] (“In response, the database ... performs an action related to the signature.”).

150. In my opinion, it would have been obvious to one of ordinary skill in the art that the teachings of Murphy could be combined with Brunk (and further in light of the general background knowledge of a person of ordinary skill in the art) to

implement a Murphy/Brunk system that stored the fingerprints and their corresponding actions in the same database. Specifically, one of ordinary skill in the art would know that if Murphy's recognition system 12 detected a match in the covariance matrix of a segment of content signal 40 and a pre-processed stored covariance matrix in database 12, Murphy's pre-processed covariance matrix fingerprints could be used, pursuant to Brunk's teachings, as an index for Murphy's database 12 to organize and access corresponding actions, and generate event flags that specify such actions for each such fingerprint. *Id.*, [0006]

151. In my opinion, a person of ordinary skill in the art would consider this combination of Murphy and Brunk obvious for multiple reasons. As I explain below, this combination involves little more than minor modifications to adjust software functionality in ways that have been long known, leading to a combination of predictable results with a reasonable expectation of success.

152. First, one of ordinary skill in the art would understand that storing the corresponding reference actions in the same database as the reference fingerprints to which they correspond (database 12) would be a natural and desirable location to store such data. The fingerprints and the actions that they trigger operate similarly to common conditional rules, and one of ordinary skill in the art would expect the conditions and actions of such rules to be stored together as associated data. In other words, storing the fingerprints and their associated actions in the same database

would not just be obvious, it would be the preferred assumption that one of ordinary skill in the art would make regarding the implementation of a system involving actions being taken upon the satisfaction of particular conditions (in this case, the detection of a matching fingerprint).

153. The alternative would be to store the related and associated actions in a separate or external database, which would offer no particular advantages to the Murphy/Brunk system. Moreover, this would disadvantageously require the additional effort to be expended generating, maintaining, and correlating fingerprints to actions in the second data structure or database.

154. Second, one of ordinary skill in the art would be motivated to store Murphy's cues and desired actions in the same database so that they can be easily modified or updated without the need to perform burdensome software updates, additional data structures, data bases and/or memory. Such a design would provide greater flexibility to the designers of the system as well as the users who set the desired fingerprints and actions.

155. In sum, storing the actions (for example, start recording, stop recording, and display information) in the same database with the fingerprint, as in the above-described Murphy/Brunk implementation, is a commonsense design decision that would be readily adopted by one of ordinary skill in the art. In my opinion, it would have been an obvious design choice. *See also, supra* Sections VI.C-D.

[1.c] identifying the corresponding reference action by obtaining a match in the database between a trigger fingerprint among the plurality of trigger fingerprints and a reference fingerprint among the plurality of reference fingerprints; and

156. In my opinion, Murphy teaches these limitations both alone and in view of Brunk. Murphy teaches that the input audio signal (and the covariance matrix fingerprints relating thereto) and the stored audio cues are matched. Ex-1011, 5:9-11. Murphy discloses that such matches can be obtained using a “distance measurement processor” in the fingerprint recognition system. *Id.*, Fig. 1, 6:30-7:1, 8:35-9:31, 9:23-28.

157. Murphy discloses a plurality of trigger fingerprints and reference fingerprints. In operation, Murphy’s recognition system receives the content signal 40 during playback and continually segments and analyzes it to generate potential a plurality trigger fingerprints during playback. *Id.*, 7:11-9:9; 13:16-20. Murphy also teaches that a plurality of reference fingerprints are stored in database 12. *Id.*, 14:6-17. As Murphy discloses using its system to trigger recording of multiple different episodes of content, one of ordinary skill in the art would understand that database 12 could store a plurality of reference fingerprints (indeed, it would be very odd to have a database store only one piece of data). *Id.*

158. Murphy teaches that its preferred way of calculating the difference threshold (and therefore determining matches) between covariance matrices (fingerprints) is to use the arithmetic harmonic sphericity measure. Murphy teaches:

The arithmetic and harmonic mean values of the eigenvalues of the covariance matrix of the window of audio signal is compared with those of the covariance matrices of the known audio cues. When the values differ by less than a threshold value a match is detected between the audio signal and the audio cue

Id., 9:23-28.

159. Once a match is determined between a detected fingerprint and a stored pre-processed fingerprint, the system generates an event flag. *Id.*, 13:21. The matched fingerprint is used to identify an action such as displaying information, start recording, or stop recording (*corresponding reference action*). *Id.*, 9:23-26, 13:22-25. In particular, Murphy teaches that that the appropriate action is taken based on the identity of the matched fingerprint, which one of ordinary skill in the art would have understood to teach identifying the corresponding action, as otherwise the device would not know which action to take.

160. Furthermore, in the Murphy/Brunk implementation of database 12 discussed in connection with limitation [1.b], above, the fingerprints act as an index for the associated actions triggered by the detection of a match to said fingerprint, thereby leading to identification of the correct corresponding reference action.

161. Further, Murphy also teaches that the content signal may contain a plurality of “trigger fingerprints” that match pre-processed and pre-stored “reference fingerprints.” For example, Murphy explains that its system is operable to record a

program using an event flag to start recording and an event flag to stop recording. *Id.*, 13:27-14:3. As Murphy explains, these event flags correspond to segments of the content signal, the fingerprints of which trigger the start and stop recording actions. *Id.*

162. Accordingly, it is my opinion that Murphy teaches that its recognition system identifies the appropriate action to take based on matches between a trigger fingerprint among the plurality of trigger fingerprints and a reference fingerprint among the plurality of reference fingerprints.

163. Independently of Murphy, it is my opinion that Brunk also teaches identifying a reference action by obtaining a match between a trigger fingerprint among the plurality of trigger fingerprints and a reference fingerprint among the plurality of reference fingerprints. Specifically, Brunk discloses performing a matching of the calculated list of signatures with a database of signatures. Ex-1012, Fig. 4, [0069]. Brunk also describes how the stored signature is associated with related information, including actions to be taken based on the matched signature in the database. *Id.*, [0028]-[0029].

[1.d] performing the reference action that corresponds to the reference fingerprint on the playback device.

164. As discussed above, it is my opinion that Murphy teaches these limitations.

165. In my opinion, Murphy teaches performing the reference action that corresponds to the reference fingerprint on the playback device. For example, Murphy teaches that the audio may cause the generation of an event flag that causes a code to be sent to the recorder that commands it to start recording. Ex-1011, 13:30-14:1. Murphy also teaches other actions, including displaying information and stopping the recording. *See supra* claim [1.a].

166. In my opinion, Brunk also teaches performing the corresponding action on the playback device. For example, Brunk teaches actions may be associated with the signature, including, for example, linking the user to another computer, transferring security or licensing information, verifying content and access, etc. that include, for example, directing the user's device to a website. Ex-1012, [0029]. Given that Brunk describes monitoring for signatures during playback, one of ordinary skill in the art would understand that these actions are performed on the playback device. *Id.*, [0068].

167. For all of these reasons, and for the reasons regarding the obviousness of combining Murphy, Brunk and the background knowledge of one of ordinary skill in the art, which I discuss below, it is my opinion that Murphy and Brunk and the background knowledge of a person of ordinary skill in the art render claim 1 obvious.

2. Claim 2

[2] The method of claim 1, wherein: different corresponding actions are to be performed at different trigger time points in the sequence of trigger time points in the content.

168. In my opinion, Murphy discloses these limitations. As discussed in [1.a], one of ordinary skill in the art would have understood Murphy to teach that different event flags can be generated by different audio cues, to take different corresponding actions at different points in the content. For example, Murphy teaches that a program can be recorded using two different actions to start recording the program and stop recording the program. Ex-1011, 13:27-14; *see supra* claim [1.a]. In this case, both actions *are to be triggered at different time points in the sequence of trigger time points in the content* (i.e., the start of the program and the end of the program).

3. Claim 3

[3] The method of claim 1, wherein: the trigger fingerprint is determined from a segment of the content, the segment ending at the trigger time point at which the reference action is to be performed.

169. In my opinion, Murphy teaches these limitations. As explained above, Murphy specifically teaches that fingerprints, including trigger fingerprints, are determined from a segment of the content, which Murphy refers to as a “window.” Specifically, Murphy teaches: “A window generator 22 segments the sampled audio signal into windows corresponding in length to the length of the audio cues to be recognised.” Ex-1011, 7:30-33 (emphasis added).

170. Further, Murphy teaches performing the action when (at a trigger time point) the audio cue is recognized and matched to a preprocessed stored audio cue, such as when Murphy explains that recording actions are taken when the audio cue is matched. *Id.*, 13:20-26, 13:27-14:3. Because the analysis of the audio cues is occurring during playback, and because the process of recognizing the cue takes time and then causes the immediate triggering of the action, one of ordinary skill in the art would understand Murphy to describe the corresponding action being taken at the end of the segment that constitutes the audio cue, after the entire cue is able to be processed. This is necessarily true because the system cannot determine which reference action it must take until after the audio cue is recognized. Because Murphy teaches that the audio cue is a fingerprint that corresponds to a particular “window” or segment, the action that occurs when the fingerprint is recognized (as Murphy describes) would necessarily occur at the end time of the segment.

4. Claim 4

[4] The method of claim 1, wherein: the trigger fingerprint is determined from a segment of the content, the segment beginning at the trigger time point at which the reference action is to be performed.

171. In my opinion, Murphy teaches and renders these limitations obvious. As explained in more detail below, Murphy teaches detecting an audio cue in the live version of the audio signal, while recording a delayed version of the same signal. This delayed version of the signal allows recording (and other actions) to be

triggered to occur at times earlier in the signal than the time corresponding to the end of the audio cue.

172. Murphy teaches the following concerning this time delayed signal:

In order to provide a suitable quality of recording, **a time delay unit 74 may be connected between the audio signal path and the automated recording system 76.** Although the recognition system processes the audio signal rapidly, any delay in the generation of the event flag would result in the beginning of the programme failing to be recorded. **In order to avoid the possibility of losing the initial part of the programme, the automated recording system 76 may be configured to record a delayed version of the audio signal provided by a time delay unit 74.**

Ex-1011, 12:18-26 (emphasis added).

173. Further, Figure 3 of Murphy, as annotated, shows event flags generated by the recognition system based on the content signal 40, but being used to take actions (here, automatic recording of the broadcast) on a separate, time delayed, signal:

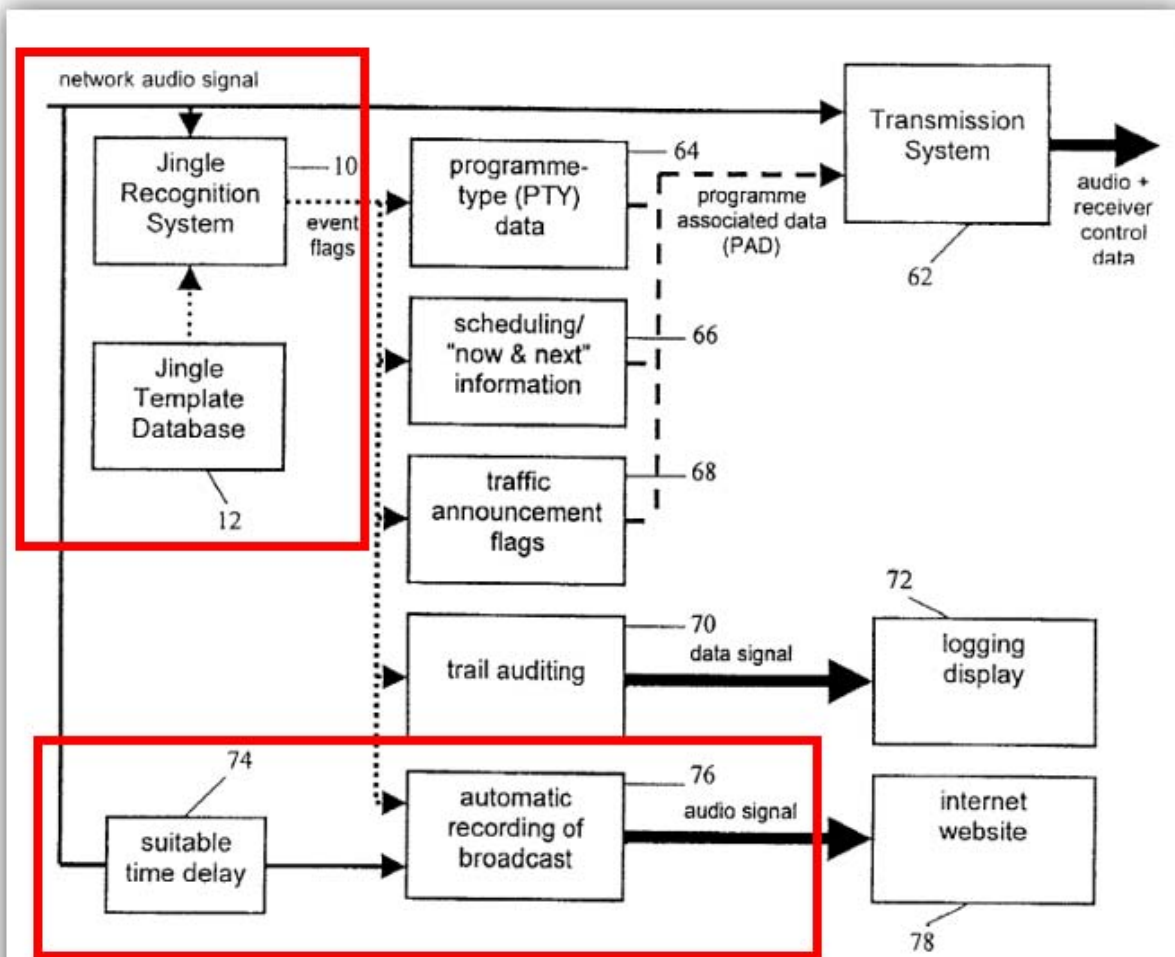


Fig. 3

Id., Fig. 3 (annotated).

174. Murphy teaches that this delay is beneficial to account for the time it may take to process the fingerprints in the received live signal. Ex-1011, 7:29-30. So, for example, when the fingerprint in the live signal is processed, it may specify an action (start recording the delayed signal) as soon as the match is detected. But the delayed version of the signal can be delayed such that the recording begins before

the time point in the content that served as the starting point of the fingerprint segment that triggered the recording action. In other words, the recording of the delayed signal may begin at an earlier time point in the content relative to the time point in the live signal that triggered the recording. *Id.*, 12:18-26.

175. Notably, Murphy discusses the concept of fingerprinting signals and triggering actions in the context of radio broadcasts, and likens its fingerprints to radio jingles. *See, e.g., id.*, Fig. 3 and Fig. 4 (referring to the fingerprint detecting module as a “jingle recognition system”). Such jingles are also recognized by humans and are often part of the program that one would want to record, just as one would begin recording a television program to include the introductory credits and/or theme song. *Id.*, 3:10-17. Accordingly, in my opinion, a person of ordinary skill in the art would have appreciated that the time delay would be calibrated in such a way that the jingle itself was included as part of the recording. That said, one of ordinary skill in the art would understand Murphy was not limited in this manner, and that audio cues and jingles could be located at any point in the content signal.

176. In other words, one of ordinary skill in the art would find it obvious to configure the length of delay unit 74 to be the same standardized length as the audio cue segment to allow the recording action start at a time that includes the content that serves as the reference audio cue. Doing this is very simple. Murphy already teaches utilizing a delayed signal to offset any delay in recording caused by time

needed to process fingerprint and trigger the action. *Id.*, 12:18-26. In order to calibrate the system to begin recording at the beginning of the segment containing the triggering fingerprint, one would merely need to add the standard length of the Murphy audio cue segment to this delay, such that the action would be triggered at a time delayed behind the live signal by the sum of the time needed for processing and the standard length of the audio signal. *Id.*, 7:28-8:2. This type of synchronization between the length of the audio cue and the length of the delay between the delayed and live signals would have been a routine engineering task, well within the grasp of one of ordinary skill in the art utilizing the system of Murphy.

177. Therefore, it is my opinion that one of ordinary skill in the art would have naturally applied common sense to configure the time delay so that “the trigger fingerprint is determined from a segment of the content, the segment beginning at the trigger time point at which the reference action is to be performed.” This configuration would use the taught delay buffer technique to accomplish the desirable goal of recording the content that contains the triggering fingerprint, providing predictable functionality and predictable benefits. It is therefore my opinion that Murphy renders obvious the dependent limitations of claim 4.

5. Claim 5

[5] The method of claim 1, wherein: the trigger fingerprint is determined from a segment of the content, the segment beginning or ending at a predetermined distance before or after the trigger time point at which the reference action is to be performed.

178. Claim 5 includes the word “or” twice and thus recites multiple alternative limitations. Specifically, claim 5 discloses four alternatives: either the beginning *or* ending of the fingerprint segment must be a predetermined distance either before *or* after the time point for performance of the reference action. My understanding is that such a claim is deemed anticipated (or obvious) if at least one of the alternatives is taught by the prior art.

179. Murphy teaches three different scenarios that anticipate and/or render obvious the dependent limitations of claim 5.

180. In the first scenario, Murphy teaches determining a fingerprint for a segment, the segment *ending* at a predetermined distance *before* the trigger time point at which the reference action is to be performed. Specifically, Murphy teaches: “The recording system 82 would be activated when it detected an event flag associated with that code. Either *the length of the programme could be specified by the user* or the recording system 82 may look for an end flag *to stop the recording.*” Ex-1011, 13:27-14:3 (emphasis added). In the configuration where the recording is set to stop based on a predetermined “length of the programme,” the fingerprint

segment of the content ends a predetermined period of time prior to the performance of the “stop recording” action.

181. In a second scenario, Murphy teaches determining a fingerprint for a segment, the segment *beginning* at a predetermined distance *before* the trigger time point at which the reference action is to be performed. This scenario is based on the same situation described above where a particular segment is used to trigger recording, with the “stop recording” action set to occur automatically after a period of time based on the length of the programme.” In that situation, because the length of the triggering segment itself is standardized, the segment would begin a predetermined distance (the length of segment plus the length of the programme) before the time for performance of the stop recording action. *Id.*, 7:28-30.

182. In a third scenario, Murphy renders obvious determining a fingerprint for a segment where the segment *ends* at a predetermined distance *after* the trigger time point at which the reference action is to be performed. Specifically, as discussed above with regard to claim 4, the system of Murphy can be set to “start recording” a delayed version of the content signal. *Id.*, 12:20-26. When this delay is employed, an action to trigger the recording of the delayed signal is triggered at time point marked by the end of the fingerprinted segment in the original content signal. The recording of the delayed signal will begin a time point in the delayed signal that is prior to the end of the fingerprinted segment in the identical but delayed

signal that, in the original signal, triggered the recording action. The triggering fingerprint segment in the delayed signal will thus end a predetermined distance (the delay) from the time point where the recording action begins.

6. Claim 6

[6] The method of claim 1, wherein: the reference action corresponds to multiple reference fingerprints that include the reference fingerprint.

183. In my opinion, Murphy teaches the dependent limitations of claim 6. Murphy teaches that the same action can be associated with multiple different fingerprints. For example, Murphy teaches that the actions to start and stop recording can be utilized to record a particular program. Ex-1011, 13:27-14:3. And Murphy also teaches that the system could be configured to record subsequent episodes of programmes automatically, or otherwise record multiple programmes. *Id.*, 14:15-7; 27-31. As one of ordinary skill in the art would have understood, recording several different programs would necessarily require taking both the “start recording” and “stop recording” actions multiple times, based on multiple different fingerprints corresponding to the different segments of content (at the beginning and end of the desired programmes to be recorded).

7. Claim 7

[7] The method of claim 1, wherein: multiple reference actions correspond to the reference fingerprint and are to be performed at the trigger time point; and the performing of the reference action is part of performing the multiple reference actions that correspond to the reference fingerprint.

184. In my opinion, Murphy teaches the dependent limitations of claim 7. Murphy teaches that multiple different actions may be triggered by the detection of a single fingerprint.

185. For example, Murphy teaches that a single detection of an audio cue could generate multiple event flags: “When a match is detected, appropriate event flags are generated ... ” Ex-1011, 10:16-17. As another example, Murphy specifically describes using its system to “record programmes and indicate information about programmes.” *Id.*, 14:27-29. One of ordinary skill in the art would have understood that achieving this result would require a single fingerprint corresponding the start of the programme, that could be used to trigger both an action to start recording the programme, and a second action to update the programme-type data so that it may be indicated as described. Indeed, Murphy describes several different types of actions being based upon a “start flag,” including updating programme-type data and updating “now and next schedule” data to allow for the name of the current song, for example, to be shown on a display. *Id.*, 11:1-15, Fig. 4.

8. Claim 8

186. Claim 8 is a system version of claim 1. In my opinion, claim 8 does not significantly differ from claim 1 and is unpatentable for substantially the same reasons.

[8.preamble] A system comprising:

187. In my opinion, Murphy discloses a system. Murphy describes a fingerprint recognition system, including the playback device of Murphy's Figure 4.

188. Brunk likewise discloses a system and is implemented on a computer. Ex-1012, [0107]. *See also, supra* claim [1.preamble].

[8.a] a fingerprint detector configured to determine a plurality of trigger fingerprints from content being played back on a playback device, each trigger fingerprint identifying a corresponding trigger time point of a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered; and

189. In my opinion, Murphy teaches these limitations. Claim [8.a] essentially recites "a finger detector configured to" perform the functionality recited in the limitations of claim 1.a, and so it is taught by the prior art for the same reasons. *See supra* claim [1.a].

190. Murphy teaches a fingerprint detector as recognition system 10, which is implemented on a processor and makes fingerprint determinations as recited in claim 1. *See supra* claim [1.a].

191. If the Board determines that the claim term "fingerprint detector configured to ..." should be construed as a means-plus-function phrase, I understand

that the corresponding structure that the Patent Owner identified during prosecution of the '114 patent is “fingerprint detector 302,” which the patent describes as a processor programmed to perform the function claimed. Ex-1001, Fig. 3, 7:20-24 and 7:33-40.

192. Murphy teaches a “fingerprint detector” with the structure identified by the Patent Owner during prosecution. Specifically, Murphy teaches that processor 14 and distance measurement processor 16, both of which are part of fingerprint recognition system 10 in receiver 80, perform the “determining” function recited in the claim and are thus the same or equivalent to the corresponding embodiment of fingerprint detector 302 recited in the '114 patent. *See supra* claim [1.a].

193. In my opinion, Brunk also discloses a fingerprint detector. Brunk describes that software operates as the claimed fingerprint detector, as explained above for claim 1.a. Ex-1012, [0032], [0068]; *see supra* claim [1.a]. However, unlike Murphy, Brunk does not expressly disclose identifying corresponding trigger time points.

[8.b] a processor configured to: access a database that includes a plurality of reference fingerprints, previously derived from the content, and a plurality of reference actions, each reference fingerprint identifying at least one corresponding reference action to be performed at a reference time point identified by the reference fingerprint;

194. In my opinion, Murphy combined with Brunk teaches the limitations of claim [8.b] for the same reasons as discussed above for [1.b].

195. Murphy's "recognition system is implemented on a processor." Ex-1011, 6:24-26, Figs. 2, 4. As discussed above with regard to claim 1.b, the processor is configured to access database 12, which, when modified in light of the teaching of Brunk regarding storage of corresponding actions in database 12, teach the limitations of this claim. Ex-1012, [0028]; *see supra* claim [1.b].

[8.c] [the processor configured to] identify the corresponding reference action by obtaining a match in the database between a trigger fingerprint among the plurality of trigger fingerprints and a reference fingerprint among the plurality of reference fingerprints; and

196. In my opinion, Murphy alone and in view of Brunk teaches these limitations for the same reasons discussed above with regard the claim 1.c. *See supra* claim [1.c]. As described above, In the Murphy/Brunk system, Murphy's processor is configured to identify the corresponding action in database 12, using the fingerprints as an index as described in Brunk. *Id.*

[8.d] [the processor configured to] perform the reference action that corresponds to the reference fingerprint on the playback device.

197. In my opinion, for the reasons discussed in claim [1.d] above, Murphy and Brunk each teach these limitations. Murphy teaches taking actions on receiver 80 that correspond to the reference fingerprint. *See supra* claim [1.d]. A person of ordinary skill in the art would have understood that Murphy's receiver 80 utilizes a computer processor that is configured to perform those actions, including start recording and stop recording. Murphy confirms the presence of this processor by,

for example, teaching that the recording system is programmable: “The recording system 82 may be programmed by the user” Ex-1011, 13:27.

198. Brunk also teaches a processor performing an action. Specifically, Brunk teaches that actions associated with the content signature can include accessing “a web site that returns information to the user device.” Ex-1012, [0029]. And Brunk teaches that the system is implemented via a processor. *Id.*, [0107]. *See supra* claim [1.d].

199. Each of elements [8.b-d] are performed by “a processor.” In my opinion, it is not necessary that the same processor be used to perform each of the recited functions. As I describe above, in Section IV.A, the term “a processor” is properly construed to mean “one or more processors.” And the specification of the ’114 patent does not disclose any specific single processor that performs the claimed functions.

200. If the Board determines that only one overall processor must perform all of the functions in [8.b-d], then Murphy still discloses such a processor. Murphy explains that the fingerprint recognition system “is implemented on a processor” and comprises at least “a processor 14 and a distance measurement processor 16.” Ex-1011, 6:23-30, Fig. 1. And processor 14 is itself described as including at least three processors; the frequency processor, log processor and covariance processor. *Id.*,

Fig. 1. Further, Murphy describes all of these processors as being part of a larger overall processor. *Id.*, 6:24-26, Fig. 1.

201. Accordingly, Murphy emphasizes the fact, known to those in the art, that defining the scope of a “processor” is often an arbitrary box-drawing exercise, as several processors can be understood to be part of a larger overall processor for performing a larger overall task. For this same reason, it is my opinion that a person of ordinary skill in the art would likewise understand that the Murphy processor performing the actions may be considered part of the same overall processor that includes the fingerprint-matching and action-identifying functions, and that the larger overall processor is equivalent to the processor recited in claim 8.

202. Moreover, in my opinion, a person of ordinary skill would have found it obvious to implement Murphy’s receiver with a central processing unit in place to control and execute all of its various functions, including playback, recording, and operating text display. This was the typical way in which consumer electronics were designed at the time of the invention and it would have been obvious and expected for Murphy’s receiver to share the same central processor design.

9. Claim 9

[9] The system of claim 8, wherein: the reference action corresponds to multiple reference fingerprints that include the reference fingerprint.

203. In my opinion, Murphy teaches the dependent limitations of claim 9 for the same reasons set forth above with regard to claim 6. *See supra* claim [6].

10. Claim 10

204. Claim 10 is a “computer readable medium” version of claims 1 and 8, and, in my opinion, its limitations are taught by the prior art for substantially the same reasons as I explained above with regard to claims 1 and 8.

[10.preamble] A non-transitory machine-readable storage medium comprising instructions that, when executed by one or more processors of a machine, cause the machine to perform operations comprising:

205. In my opinion, Murphy teaches the preamble of claim 10. *See supra* claim [1.pre], [8.pre]. Murphy teaches that its “recognition system is automated” (Ex-1011, 3:32) and “implemented on a processor or as a program [in] a computer” (Ex-1011, 6:24-26, Fig. 1). A computer sophisticated enough to implement the software functionality described in Murphy would necessarily also comprise a non-transitory storage medium comprising instructions, and this fact would have been apparent to one of ordinary skill in the art.

206. Brunk explains that its system may be implemented in a “programmable computer running executable software.” Ex-1012, [107]. As with Murphy, one of ordinary skill in the art would have understood that a programmable computer running executable software necessarily includes a processor to run the software and a non-transitory storage medium to store the software. Further, those of ordinary skill in the art would have understood the software itself to be the claimed “instructions.”

[10.a] determining a plurality of trigger fingerprints from content being played back on a playback device, each trigger fingerprint identifying a corresponding trigger time point of a sequence of trigger time points in the content and at which trigger time point at least one corresponding action is to be triggered;

207. In my opinion, Murphy teaches these limitations for the same reasons set forth above with regard to claims [1.a] and [8.a]. *See supra* claim [1.a], [8.a].

[10.b] accessing a database that includes a plurality of reference fingerprints, previously derived from the content, and a plurality of reference actions, each reference fingerprint identifying at least one corresponding reference action to be performed at a reference time point identified by the reference fingerprint;

208. In my opinion, Murphy combined with Brunk teaches these limitations for the same reasons set forth above with regard to claims [1.b] and [8.b]. *See supra* claim [1.b], [8.b].

[10.c] identifying the corresponding reference action by obtaining a match in the database between a trigger fingerprint among the plurality of trigger fingerprints and a reference fingerprint among the plurality of reference fingerprints; and

209. In my opinion, Murphy alone and in view of Brunk teaches these limitations for the same reasons set forth above with regard to claims [1.c] and [8.c]. *See supra* claims [1.c], [8.c].

[10.d] performing the reference action that corresponds to the reference fingerprint on the playback device.

210. In my opinion, both Murphy and Brunk teaches these limitations for the same reasons set forth above with regard to claims [1.d] and [8.d.] *See supra* claims [1.d], [8.d].

11. Claim 11

211. In my opinion, claim 11 is a slight variation of the method recited in claim 1 and it is unpatentable for substantially the same reasons as set forth for claim 1. *See supra* claim [1].

[11.preamble] A method comprising:

212. Murphy and Brunk disclose a method, as discussed further below. *See also, supra* claim [1].

[11.a] accessing segments of content, a segment among the segments being associated with a corresponding action to be triggered during playback of the segment;

213. In my opinion, Murphy discloses this limitation. As I explained above with regard to claim limitation [1.a], Murphy teaches dividing a signal into segments referred to as “windows” and determining whether the covariance matrix “fingerprints” of those segments match any fingerprints stored in database 12. *See supra* claim [1.a].

214. In Murphy, the system “accesses” segments in two different contexts. The first context is to detect fingerprint matches during playback, which I have discussed above with regard to claim 1. *See supra* claim [1.a].

215. The second context is during the generation of the pre-stored fingerprints in database 12. Specifically, Murphy teaches that the same process of accessing the content segment by segment occurs when fingerprints are generated for storage in database 12. Specifically, Murphy states:

Rather than store the raw audio cue signal, the audio cue is pre-processed to determine an individual determining characteristic, or signature Preprocessing may be performed by the recognition system itself or the signature of the audio cue may be generated by a compatible system and simply downloaded into the store of the recognition system.

Ex-1011, 6:32-7:5; *see also* 14:4-15.

216. Once signatures are stored, Murphy teaches that the pre-processed segments in database 12 can be accessed both remotely and locally. *See supra* claim [1.a and 1.b]. *See also* Ex-1011, 6:30-7:9; 14:4-15.

217. And, as explained above with regard to claim 1, the segments that are accessed (either during playback or in database 12) are associated with a corresponding action to be triggered during playback of the segment. *See supra* claim [1.a].

218. Further, Murphy teaches that corresponding actions can be triggered during the segment containing the fingerprint that triggers the actions. Specifically, one of ordinary skill in the art would have understood that Murphy's teachings regarding use of a time delay enable the system to trigger an action during the content containing the fingerprint by setting the delay to slightly less than the length of the segment. Ex-1011, 12:20-26.

[11.b] generating a reference fingerprint of the segment associated with the corresponding action, the generating being performed by a processor;

219. In my opinion, Murphy and Brunk each teach these limitations. *See supra* claim [1.a-b]. As discussed above, Murphy teaches that reference fingerprints can be generated (as a covariance matrix) that corresponds with a particular segment associated with a corresponding action, such as actions to start recording, stop recording, and display information. *See supra* claim [1.a-b].

220. As just discussed in [11.a], Murphy teaches several contexts involving preprocessing to generate the reference fingerprints. Ex-1011, 6:30-7:9, 14:4-15. Murphy teaches that the recognition system of the receiver of Figure 4 is used to process the received content signal and store the derived audio cue information in the database. *Id.*, 14:13-15. Murphy further teaches that this *generating* step is performed by a processor, as it discloses that “[p]reprocessing may be performed by the recognition system itself.” *Id.*, 7:1-2. As discussed above with regard to claim [1.a], recognition system 10 includes processor 14 (including its sub-processors), which are used to analyze segments of a signal and generating the reference fingerprints. *See, e.g., id.*, 7:1-2, 7:12-9:8, 9:5-7 (“The covariance matrix of the audio cue is determined in a similar manner except that windowing of the audio cue is not required.”).

221. In my opinion, Brunk also discloses the limitations of [11.b]. Brunk discloses that its system is carried out by a processor. Ex-1012, [0107]. Like

Murphy, Brunk teaches that the content item is divided into segments and a signature (*reference* fingerprint) is generated that represents a particular segment. *Id.*, [0020]; *see supra* claim [1.a-b]. Brunk further discloses that its system can take actions that correspond to signatures, which based on Brunk's disclosure, includes signatures based on particular segments. Ex-1012, [0029]; *see supra* claim [1.a-b].

[11.c] assigning the reference fingerprint of the segment to the corresponding action that is to be triggered during playback of the segment of the content;

222. In my opinion, Murphy alone and in view of Brunk teaches these limitations. *See supra* claim [1.b]. As discussed in [1.b], Murphy explains that actions can be assigned to stored reference fingerprints. *See supra* claim [1.b]. The actual act of "assigning" is also taught by Murphy. One of ordinary skill in the art would understand that Murphy's receiver assigns the corresponding actions to their particular reference fingerprints to a corresponding action. Specifically, Murphy teaches triggering the correct event flags when a matching fingerprint is recognized, a result that one of ordinary skill in the art would have understood to inherently require some association being made in Murphy's programming between the fingerprints that are to be detected and the actions they are to trigger. *See supra* claim [1.b]. The creation of this necessary association constitutes *assigning the reference fingerprint of the segment to the corresponding action that is to be triggered during playback of the segment of the content*.

223. In addition, Murphy explains how receiver 80 can be programmed to record particular programs, which is accomplished by setting particular fingerprints to triggering the particular corresponding recording system during playback. Ex-1011, 13:27-14:1. One of ordinary skill in the art would understand that this programmability necessarily requires assigning the reference fingerprint to the corresponding action.

224. In my opinion, it would alternatively be obvious to combine the teachings of Murphy and Brunk to practice these limitations. As explained above in 1.b with regard to the Murphy/Brunk system, it would be obvious to modify Murphy's database 12 using the teachings of Brunk to use the stored fingerprints as an index for references actions assigned to particular fingerprints. One of ordinary skill in the art would understand that the act of populating this database would necessarily require performing the claimed "assigning" step.

[11.d] providing access to the reference fingerprint of the segment to a playback device configured to match the reference fingerprint to a playback fingerprint derived from the segment by the playback device; and

225. In my opinion, Murphy teaches these limitations. As I explain above for claim limitations [1.a-c], Murphy describes a receiver that includes a playback device that itself includes a distance measurement processor 16. This processor is configured to determine if any signature of an audio cue in database 12 matches any

covariance matrix generated by processor 14 during playback of the content. *See supra* claim [1.a-c].

226. Murphy further teaches “providing access to the reference fingerprint in the segment of the playback devices. For example, Murphy teaches that the reference fingerprints can be downloaded to populate database 12. Ex-1011, 14:4-15. This is consistent with Murphy’s discussion of utilizing a local database 12 at the playback device with access to, for example, “a centrally provided database” having the fingerprints. *Id.*, 14:8.

227. In addition, accessing the reference fingerprint also occurs when the playback device accesses database 12 to determine if any of the reference fingerprints match the covariance matrices detected in the received content signal, as described, for example, for [1.b] and [1.c].

228. In my opinion, Brunk also teaches these limitations. Specifically, Brunk describes a broadcast monitoring application that acts as the claimed playback device and determines a content signature after the detection of a cue signal. Ex-1012, [0068]. The application then compares the fingerprint to the fingerprints that are stored in the database (*reference fingerprints*) to identify a match. *Id.*, [0068]-[0069]. One of ordinary skill in the art would recognize this this step necessarily requires providing access to the reference fingerprint as claimed.

[11.e] providing the identifier of the corresponding action to the playback device based on the reference fingerprint matching the playback fingerprint, the provided identifier causing the playback device to trigger the corresponding action during the playback of the segment.

229. In my opinion, Murphy alone and in view of Brunk teaches these limitations. *See supra* claim [1.b], [1.c].

230. Murphy's event flags are the claimed identifier. Once a match is determined between a detected fingerprint and a stored pre-processed fingerprint, the system generates an event flag. *Id.*, 13:21. The matched fingerprint is used to identify an action such as displaying information, start recording, or stop recording (corresponding reference action). *Id.*, 9:23-26, 13:22-25. In particular, Murphy teaches that the event flag indicating the appropriate action is generated based on the identity of the matched fingerprint, which one of ordinary skill in the art would have understood to teach providing the identifier of the "providing the identifier of the corresponding action to the playback device based on the reference fingerprint matching the playback fingerprint, the provided identifier causing the playback device to trigger the corresponding action during the playback of the segment."

231. In the Murphy/Brunk modified system described above in [1.b], database 12 has reference actions linked to reference fingerprints and, so, database 12 "provides an identifier of the corresponding action to the playback device" when a stored audio cue is matched to a fingerprint detected during playback. When this occurs, an event flag is generated that identifies the action to be performed (such as

recording or displaying information) by the playback device, thereby “triggering the corresponding action during the playback of the segment.” *Id.* See *supra* claim [11.c] and [1.b],

12. Claim 12

[12.preamble] The method of claim 11, wherein:

[12.a] the assigning of the reference fingerprint to the identifier of the corresponding action includes storing the reference fingerprint with the identifier of the corresponding action in a database; and

[12.b] the playback device is configured to search the database to match the reference fingerprint to the playback fingerprint.

232. In my opinion, Murphy alone and in view of Brunk teaches the dependent limitations of claim 12. See *supra* claim [11.c, e].

13. Claim 13

[13] The method of claim 11, wherein: the playback device is configured to trigger the corresponding action at a time point that is a predetermined distance from a start of the segment of the multimedia content.

233. In my opinion, Murphy renders obvious the dependent limitations of claim 13. See *supra* claims [11], [5]. As I discuss above in connection with claim 5, one of ordinary skill in the art would have found it obvious to use Murphy’s receiver to cause a stop recording action to occur at the predetermined program length after the start of the segment having the fingerprint. See claim 5.

234. Further, as explained below in connection with claim 14, Murphy discloses triggering actions at a time point that is a predetermined distance from the

end of a segment. Because the segment lengths disclosed by Murphy are standard, and therefore predetermined, triggering an action at a time point that is a predetermined distance from the end of a segment also necessarily triggers the action at a time point that is a pre-determined distance from the beginning of the segment. Ex-1011, 7:28-8:2.

235. I also note that this claim (and claims 14, 23, and 24, *infra*) refer to “multimedia” content. The ’114 patent uses the term “multimedia signal” to refer to signals that are audio only, video only, or both. Ex-1001, 3:51-52. Others in the art at this time also recognized that audio signals could be “multimedia” content. *See, e.g.,* Ex-1016, Abstract (“The disclosed method generates robust hashes for multimedia content, for example, audio clips.”). Thus, Murphy teaches the “multimedia” limitation.

236. Even if one were to construe “multimedia” to require something more than a bare audio signal, Murphy would still teach the limitation. Specifically, Murphy teaches that its fingerprint recognition system can be used not only with an audio signal but also with the audio track of a television signal, or a video signal. Ex-1011, 15:17-25.

14. Claim 14

[14] The method of claim 11, wherein: the playback device is configured to trigger the corresponding action at a time point that is a predetermined distance from an end of the segment of the multimedia content.

237. In my opinion, Murphy teaches the dependent limitations of claim 14. *See supra* claims [11], [5], [13]. As I discuss above regarding claim 5, Murphy teaches that the playback device can automatically trigger the *stop recording* action at the predetermined program length of time *from* when it detects the fingerprint. As explained above, one of ordinary skill in the art would have understood that Murphy detects the fingerprint at the *end* of the segment having the fingerprint. *See* claim 5. Murphy thus teaches this limitation.

238. In my opinion, alternatively, Murphy also renders these claim limitations obvious. *See supra* claims [11], [5], [13]. As I explained above in connection with claim 5, using Murphy's time delay to trigger the "start recording" action renders this claim obvious. Ex-1011, 12:20-26. *See supra* claim [5].

15. Claim 16

[16] The method of claim 11, wherein: the corresponding action for the reference fingerprint and for the segment of the content includes retrieving and displaying on a display additional information selected from a group consisting of: biography data for an actor that appears in the segment of the content, a link to a website, a local weather forecast, and a local commercial.

239. Claim 16 requires displaying information selected from a group containing several alternative types of information. My understanding is that this

group limitation is deemed to be taught by the prior art if any of the available options are disclosed in the prior art.

240. With that understanding, it is my opinion that Murphy in view of Brunk teaches the dependent limitations of claim 16. Brunk teaches that a stored signature may be associated with actions that include “providing a URL” and linking to “a website that returns information to the user device.” Ex-1012, [0028]-[0029]. Brunk further discloses that the device carrying out these actions can be user devices that are known in the art to include displays. Ex-1012, [0029]. Providing a URL or linking to a website on such a user device with a display would necessarily, or at least obviously, require displaying “a link to a website” as the claim suggests, as this is how conventional browser applications present links/URLs to user devices.

241. Furthermore, in my opinion, it would be obvious to display Brunk’s URL or linked website on the playback device’s display in light of Murphy. Murphy teaches that recognition system 10 may be used to “display information on a given programme” and one of ordinary skill in the art would find it obvious, particularly in light of Brunk’s teachings discussed above, to use this display functionality to provide a link to a website. Ex-1011, 4:25-32.

16. Claim 17

[17] The method of claim 11, wherein: the corresponding action for the reference fingerprint and for the segment of the content includes retrieving and playing via a speaker additional information selected from a group consisting of: biography data of an actor that appears in the segment of the content, a local weather forecast, and a local commercial.

242. Claim 17 requires that the corresponding action be selected from a group containing several alternative types of corresponding actions. My understanding is that this group limitation is deemed to be taught by the prior art if any of the available options are disclosed in the prior art.

243. In my opinion, Murphy teaches these limitations because Murphy teaches that one of the actions available to be taken by the recognition system includes causing the device to change channels to a traffic announcement. This teaching renders it obvious to one of ordinary skill in the art that the Murphy could just as easily be employed to play a local weather announcement instead.

244. Specifically, Murphy teaches that one action that can be taken by the fingerprint recognition system includes changing the channel to play a traffic announcement when an audio corresponding to a segment of a signal played at the start of a traffic announcement is recognized by its fingerprint. Ex-1011, 10:21-31. When the cue is recognized an event flag is sent to the transmission system that includes a Traffic Programme Identification Code. *Id.*, 10:21-26. The receiver recognizes this code, and uses it to display additional information (retrieving and

displaying additional information on a display), or alternatively to change the channel to listen to a traffic announcement being broadcast on another channel.

245. As I explained in Section VI.C *supra* there is no reason why the action to take change the channel would be limited to the transmitter embodiment described in Murphy, and one of ordinary skill in the art would understand the receiver embodiment was capable of being configured to perform the same action.

246. In my opinion, the teachings of Murphy to trigger an action to change the channel to one playing a traffic announcement render it obvious to change the channel to one playing a local weather announcement. One of ordinary skill in the art would understand that the functionality of the system is does not depend on the type of content being played in the triggering fingerprint or the channel to which the playback device is changed.

247. One of ordinary skill in the art would have also been motivated by Murphy's teachings to use this functionality to play local weather announcements. From Murphy, it would have been apparent to one of ordinary skill in the art that the purpose of changing the channel to a traffic announcement rather than merely providing the traffic announcement on the same channel that included the traffic-announcement-signaling audio cue was to allow for different traffic announcements to be played based on the different locality of the broadcaster. Traffic information is local in nature, and so it makes sense that a national broadcast would want to

signal when in its broadcast the local traffic information should be substituted. This concept is described in the '114 patent, and acknowledged as a known example of when triggered actions are useful. Ex-1001, 1:53-62. One of ordinary skill in the art would have understood that the local weather announcements, which are also often provided on radio broadcasts, would be another obvious use case for this channel changing functionality.

248. Further, Murphy teaches that such a local weather announcement would be played using a speaker. The playback device of Murphy includes a speaker 90, which is used to play the received audio content. *See supra* claim [1.a].

17. Claim 18

[18] The method of claim 11, wherein: the corresponding action for the reference fingerprint and for the segment of the content is selected from a group consisting of: replacing the content with different content for a period of time, pausing playback of the content, stopping playback of the content, executing a control command, and preparing the playback device to receive user input.

249. Claim 18 requires that the corresponding action be selected from a group containing several alternative types of corresponding actions. My understanding is that this group limitation is deemed to be taught by the prior art if any of the available options are disclosed in the prior art.

250. In my opinion, Murphy teaches the dependent limitations of claim 18 in two different ways. First, as discussed above with regard to claim 1.a, Murphy teaches using fingerprints to start or stop recording. Murphy describes the setting of

event flags when a fingerprint match is detected, which are used to command recording system 82 to start recording the signal or stop recording the signal. Ex-1011, 13:20-26; 14:2-3. In this context, one of ordinary skill in the art would have understood that the event flags are used to control the operation of recording system 82 and constitute executing a control command. *See supra* claim [1.a].

251. Second, Murphy teaches “replacing the content with different content for a period of time.” As explained in claim 17, Murphy discloses that one of the actions that may be triggered by a fingerprint’s recognition is changing the channel to play a traffic announcement being broadcast on another channel. Ex-1011, 10:21-31. In this scenario, the action that is taken replaces the content (the content of the original channel) with different content (the traffic announcement on the other channel) for a period of time.

18. Claim 19

[19] The method of claim 11, wherein: the content includes at least one of audio content or video content.

252. In my opinion, Murphy discloses both of the alternatives referenced in this claim. As discussed above in connection with claim limitation [1.a], Murphy teaches content that includes audio content. *See supra* claim [1.a]. Murphy also teaches using the audio track associated with video in a television broadcast, and further teaches using video itself. Ex-1011, 15:18-25.

253. In my opinion, Brunk also teaches that the content item may comprise both audio and video signals, as it expressly says so. Ex-1012, [0025].

19. Claim 20

254. Claim 20 is a “computer readable medium” version of claim 11 and, in my opinion, it is taught by the prior art for substantially the same reasons that I provided in connection with claim 11. *See supra* claims [11], [1].

[20.preamble] A non-transitory, machine-readable storage medium comprising instructions that, when executed by one or more processors of a machine, cause the machine to perform operations comprising:

255. In my opinion, Murphy and Brunk teach this preamble. *See supra* claims [11.preamble], [10.preamble], [1].

[20.a] accessing segments of content, a segment among the segments being associated with a corresponding action to be triggered during playback of the segment;

256. In my opinion, Murphy teaches these limitations. *See supra* claims [11.a], [1.a].

[20.b] generating a reference fingerprint of the segment associated with the corresponding action;

257. In my opinion, Murphy and Brunk teach these limitations. *See supra* claims [11.b], [1.a-b].

[20.c] assigning the reference fingerprint of the segment to the corresponding action that is to be triggered during playback of the segment of the content;

258. In my opinion, Murphy alone and combined with Brunk teaches these limitations. *See supra* claims [11.c], [1.b].

[20.d] providing access to the reference fingerprint of the segment to a playback device configured to match the reference fingerprint to a playback fingerprint derived from the segment by the playback device; and

259. In my opinion, Murphy teaches these limitations. *See supra* claims [11.d], [1.a-c].

[20.e] providing the identifier of the corresponding action to the playback device based on the reference fingerprint matching the playback fingerprint, the provided identifier causing the playback device to trigger the corresponding action during the playback of the segment.

260. In my opinion, Murphy alone and in view of Brunk teaches these limitations. *See supra* claims [11.e, c], [1.b].

20. Claim 21

261. Claim 21 is the system version of claim 11 and, in my opinion, it is taught by the prior art for substantially the same reasons that I provided in connection with claim 11. *See supra* claim [11].

[21.preamble] A system comprising:

262. In my opinion, Murphy and Brunk each disclose a system, as further discussed below. *See also, supra claim* [11].

[21.a] an input module configured to access segments of content, a segment among the segments being associated with a corresponding action to be triggered during playback of the segment; and

263. In my opinion, Murphy teaches these limitations. *See supra* claims [11.a], [1.a]. Murphy teaches *accessing segments of content, a segment among the segments being associated with a corresponding action to be triggered during*

playback of the segment in multiple contexts, as discussed above in connection with claim limitation 11.a. *See supra* claim [11.a].

264. This accessing functionality is performed by what one of ordinary skill in that art would understand to be an input module. For example, the recognition system depicted in Figure 4 generates reference fingerprints through use of processor 14, (as shown in Figure 1). One of ordinary skill in the art would have understood processor 14 to comprise “an input module” that includes one or all of analog to digital converter 20, window generator 22, and frame generator 24 to access the segments of content. Fig. 1, 7:17-8:13. In the embodiment involving use of a “compatible system” to access the segments, one of ordinary skill in the art would have understood from Murphy’s disclosure that similar components would be used. Ex-1011, 7:1-5.

265. If the Board determines that the claim term “input module configures to... .” should be construed as a means-plus-function phrase, I understand that the corresponding structure that the Patent Owner identified during prosecution of the ’114 patent is “AV input module 201,” which the patent describes as essentially a processor programmed to receive the multimedia content in the form of an analog or digital signal and provide it to the fingerprinting module 202. Ex-1001, Fig. 3, 6:59-62. In my opinion, one of ordinary skill in the art would recognize that processor 14, including analog to digital converter 20, window generator 22 and

frame generator 24, constitutes a structure that is equivalent to the '114 patent's "input module" in that both are processors that receive input of the signal for fingerprinting. Ex-1001, Fig. 2, 6:50-54, 6:59-7:10.

[21.b] a fingerprint module implemented in a processor and configured to: generate a reference fingerprint of the segment associated with the corresponding action;

266. In my opinion, Murphy alone and in view of Brunk teaches these limitations. *See supra* claims [11.b], [1.a-b]. As I explain above in connection with claim limitation [11.b], Murphy teaches a fingerprint recognition system that is implemented in a processor and configured to perform the claimed generating of a reference fingerprint. *See supra* claim [11.b].

267. One of ordinary skill in the art reading Murphy would have understood that the components of Murphy perform the tasks of populating database 12 with fingerprints and their corresponding actions, providing access to that database with distance measurement processor 16, and identifying the actions to the playback device, together constitute a fingerprint module: specifically, components 26-34 of Murphy's processor 14 shown in Figure 1. Ex-1011, Fig. 1. In the embodiment where, when the reference fingerprint is generated by a remote "compatible system" (*see*), similar components would constitute the *fingerprint module*. *See supra* claim [11.b]; Ex-1011, 6:30-7:9.

268. If the Board determines that the claim term “fingerprint module” should be construed as a means-plus-function phrase, I understand that the corresponding structure that the Patent Owner identified during prosecution of the ’114 patent is “fingerprint generator 202,” which the patent describes as essentially a processor programmed to generate fingerprints. Ex-1001, Fig. 2, 6:50-54, 6:59-7:10. In my opinion, one of ordinary skill in the art would recognize that Brunk’s processor and Murphy’s processor 14 constitute structures that are equivalent to the ’114 patent’s “fingerprint module” in that they too are processors that generate fingerprints for the recognition system. Ex-1001, Fig. 2, 6:50-54, 6:59-7:10.

269. In a Murphy/Brunk implementation, the system uses a processor to pre-process the content signal to generate the set of fingerprints and also to store those fingerprints in Murphy’s database 12 along with the actions to which they correspond. Therefore, one of ordinary skill in the art would recognize that the Murphy/Brunk system has a processor that carries out the same functions claimed in [21.b-e], and is at least equivalent to fingerprint generator 202 of the ’114 patent.

270. Brunk also discloses a processor that divides a content item into segments and generates a signature for the particular segments, where that signature corresponds with an action. Brunk thus discloses a processor that generates fingerprints as a fingerprint module for the same reasons set out for the

corresponding limitations of claims 1 and 11. Ex-1012, [0020], [0029], [0107]; *see supra* claims [1.a-b], [11.b].

[21.c] [the fingerprint module configured to] assign the reference fingerprint of the segment to the corresponding action that is to be triggered during playback of the segment of the content;

271. In my opinion, Murphy alone and in view of Brunk teaches these limitations. *See supra* claims [11.c], [1.b], [21.b].

[21.d] [the fingerprint module configured to] provide access to the reference fingerprint of the segment to a playback device configured to match the reference fingerprint to a playback fingerprint derived from the segment by the playback device; and

272. In my opinion, Murphy alone and in view of Brunk teaches these limitations. *See supra* claims [11.d], [1.a-c], [21.b].

[21.e] [the fingerprint module configured to] provide the identifier of the corresponding action to the playback device based on the reference fingerprint matching the playback fingerprint, the provided identifier causing the playback device to trigger the corresponding action during the playback of the segment.

273. In my opinion, Murphy alone and in view of Brunk teaches these limitations. *See supra* claims [11.e], [1.b], [21.b].

21. Claim 22

[22] The system of claim 21, wherein: the system further comprises a database; the fingerprint module is configured to store the reference fingerprint with the identifier of the corresponding action in the database; and the playback device is configured to search the database to match the reference fingerprint to the playback fingerprint.

274. In my opinion, Murphy in view of Brunk teaches these limitations. *See supra* claims [12], [11.c, e], [21.b].

22. Claim 23

[23] The system of claim 21, wherein: the corresponding action is to be triggered by the playback device at a time point that is a predetermined distance from a start of the segment of the multimedia content.

275. In my opinion, Murphy renders these limitations obvious. *See supra* claims [13], [5].

23. Claim 24

[24] The system of claim 21, wherein: the corresponding action is to be triggered by the playback device at a time point that is a predetermined distance from an end of the segment of the multimedia content.

276. In my opinion, Murphy teaches and renders obvious these limitations. *See supra* claims [14], [5], [21.b].

VIII. MOTIVATION TO MODIFY OR COMBINE

277. Having analyzed the claim limitations in view of Murphy, Brunk and the background knowledge of one of ordinary skill in the art, it is my opinion that such a person of ordinary skill in the art would have found it obvious to combine the teachings of the prior art to arrive at the claimed subject matter of the Challenged

Claims of the '114 patent. This combination would require only using the teachings of the art in highly predictable ways with a reasonable expectation of success. As set forth in more detail below, I believe this combination is obvious for multiple reasons.

278. One such reason that serves as a general overarching principle is that Murphy and Brunk are very similar and include significant overlap in their teachings. Both references teach methods of generating and utilizing fingerprints of multi-media content to trigger actions. In addition, as set forth in more detail in the claim mapping above, both references independently teach numerous limitations claimed by the '114 patent. This reason, along with other factors discussed below, indicates to me that one of ordinary skill in the art examining the prior art would consider both Murphy and Brunk and readily combine their teachings to create a single system reflecting the system recited in the Challenged Claims.

A. Brunk's Teaching of Storing Actions With Fingerprints in Databases

279. As discussed above, Murphy teaches using the detection and matching of fingerprints during playback of a content signal to trigger a time-specific action during playback. This functionality is supported by a database that stores the reference fingerprints.

280. Murphy does not explicitly disclose that the actions corresponding the reference fingerprints are stored in the same database 12 along with the reference

fingerprints. However, for the reasons discussed below, a person of ordinary skill in the art would have found it obvious to store corresponding actions in the same database as the reference fingerprints by combining Murphy with Brunk's teaching of using fingerprints as an index for a database that also contained the associated actions to be performed. Ex-1012, [0006], [0028], [0029].

281. First, in my opinion, one of ordinary skill in the art would have been motivated to implement such a system by following Brunk's own teachings, suggestions, and motivations. Brunk explains that an "advantage" of content signatures is that they may be used to "enable actions or specifying behavior" and that using the content signature "to index ... a related database" would allow one to realize this advantage. *Id.*, [0006]. Brunk also explains that implementing the indexed database described results in greater speed for identifying the action: "From that content signature, the correct action or web link can *quickly* be found with CAM or traditional RAM based databases." *Id.* [0095], [0096] (same) (emphasis added). Therefore, Brunk itself provides significant motivations that one of ordinary skill in the art would readily recognize as a reasoned basis for adopting Brunk's database design. And one of ordinary skill in the art would further recognize that utilizing that design in connection with Murphy's database 12 would predictably provide the same such advantages.

282. In other words, Brunk's express teaching of storing the actions in the database with the fingerprint would improve Murphy's fingerprint detection system in the same ways that it improved Brunk's similar system (i.e., by improving the speed with which the system could determine which actions to take). A person of ordinary skill in the art would therefore be more than capable of making these minor improvements to Murphy with entirely predictable results, and an appropriately reasonable expectation of success.

283. Second, the database designed from the combination of Murphy and Brunk would be obvious because it represents no more than a simple design decision incorporating techniques well known in the art. Those of ordinary skill in the art were well informed about fingerprinting, using fingerprinting databases, detecting fingerprints, and triggering actions based on fingerprints. *See supra* Section VI.B. The Murphy/Brunk design requires nothing beyond this basic knowledge: storing data (actions) in a (fingerprint) database, as taught by Brunk, to improve the very similar devices taught by Murphy.

284. More specifically, a person of ordinary skill in the art understood that a database's purpose is to allow for the organization of data into useful associations and, absent any design constraint that would prevent such an action, would naturally store all information in the same database. It would therefore be readily apparent to a person of ordinary skill in the art that the database of Murphy, which is already

described as including the reference fingerprints that trigger actions, could also be beneficially used to associate those reference fingerprints with their corresponding actions. Storing actions with the fingerprints that trigger them in the database of Murphy would be little more than a trivial improvement that would already be a matter of common sense based on Murphy alone. When combined with the express teaching of Brunk, a person of ordinary skill in the art would immediately recognize that the combined database design was readily suitable for Murphy (if they had not already determined that on their own). As such, a person of ordinary skill in the art would have been motivated and fully able to combine Brunk's identification of the triggered action in Murphy's fingerprint database.

B. Robust Hash Fingerprints of Video

285. It is also my opinion that it would have been obvious to implement Murphy's fingerprint recognition system using any of the "robust hash" algorithms known in the art to reliably identify audio, video, or combined audio and video signals, including those taught by Brunk. I do not read the claims to require any particular type of "robustness" to the claimed fingerprints. To the extent the Patent Owner argues or the Board finds that any of the claims (or an amendment thereto) do require robust hashing, my opinion is that such a system would have been obvious over Murphy, and Brunk (Ex-1012), in view of the general knowledge and common sense of a one of ordinary skill in the art, (*see* Sections IV.B.3-4 above).

286. Regarding robust hashes, Murphy itself provides the teaching, suggestion, and motivation for using such a fingerprint. Murphy expressly teaches and suggests that “[a]ny method of characterising a signal which makes it discernable from a different signal could be used” and further teaches using *any* “individual determining characteristic, or signature” with its fingerprint detection system. Ex-1011, 5:9-14, 6:33-34. Murphy therefore acknowledges the advantages of “robustness,” and a person of ordinary skill in the art would be motivated to employ any of the robust fingerprint algorithms disclosed in Brunk or otherwise known in the art to realize this advantage. *See* Sections IV.B.3-4 (discussing Oostveen (Ex-1014), Haitzma (Ex-1016), and Brunk (Ex-1012). Using such a robust hash fingerprint would achieve predictable results with reasonable expectations of success given the state of the art, and it merely requires the use of known techniques employed to achieve the same results in improving similar devices. Sections IV.B.3-4.

287. To provide a specific example, a person of ordinary skill in the art would have found it obvious to implement Murphy’s fingerprint recognition system using Brunk’s teaching of using a “robust hash” as the fingerprinting algorithm. Brunk includes a detailed disclosure concerning the use of hashing algorithms to create a fingerprint and describes the advantages of basing a signature on a large number of perceptible features to improve the robustness of the signature. Ex-1012,

Fig. 1, [0021], [0026], and [0033]. Accordingly, like Murphy, Brunk also provides a teaching, suggestion, and motivation to use robust hashing.

288. Given the aligned disclosures concerning the desirability of robust fingerprints found in both Murphy and Brunk, a person of ordinary skill in the art would readily appreciate that the more specific hashing teachings of Brunk would provide the desired level of robustness, as explained by Murphy. This modification to use Brunk's robust hashing algorithms in Murphy's fingerprint recognition system would require only the use of known techniques from Brunk to improve a similar device in the same way, to predictably achieve the same robustness in Murphy.

289. I also do not read the claims as specifically requiring use of a video content signal. However, to the extent the Patent Owner argues or the Board finds that any of the claims (or an amendment thereto) do require the use of a video signal, I believe that too would be obvious over Murphy, Brunk, and the knowledge of one skilled in the art.

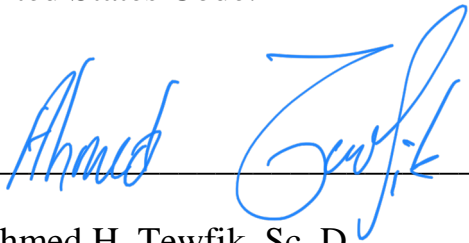
290. As I explain above in connection with the claims utilizing the term "multimedia," Murphy teaches that its fingerprint recognition system can additionally be used with the audio track of an associated television signal (Ex-1011, 15:17-22) or, alternatively, video cues in video signals (*id.*, 15:22-25). Murphy itself therefore provides the teaching, suggestion, and motivation for using the appropriate fingerprinting techniques known in the art for the particular signal with which the

system is to operate (audio, visual or a combination thereof). One of ordinary skill in the art would have known that, by using the appropriating fingerprinting technique (such as those disclosed in Brunk or otherwise known to those of ordinary skill in the art), the same approach described in Murphy could be used to generate fingerprints from audio, image, and/or video content. As indicated by Murphy, one of ordinary skill in the art would be readily able to implement that change as a matter of routine.

291. Brunk too explains that its fingerprinting system can be used with video content, and specifically describes applying hash algorithms to video signals to generate content signatures. Ex-1012, [0021], [0025]. In addition to Murphy's express teaching of utilizing its fingerprint system for video content, one of ordinary skill in the art would have realized that the hashing algorithms disclosed by Brunk were suitable for generating fingerprints based on video content, and would have found it obvious to accept Murphy's invitation to use any available fingerprinting method to implement the specific video hashing algorithms discussed in Brunk. Indeed, Brunk makes clear that several such algorithms were known in the art and discloses them merely by reference (for example, at *id.*, [0021]), further indicating that one of ordinary skill in the art would have been aware of these methods and capable of implementing them in the Murphy system.

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

Date: 12/5/19



Ahmed H. Tewfik, Sc. D.