

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ROKU, INC.

Petitioner

v.

ANONYMOUS MEDIA RESEARCH HOLDINGS, LLC.

Patent Owner

Case IPR2024-01054 (U.S. Patent No. 8,510,768)

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

EXPERT DECLARATION OF W. LEO HOARTY

TABLE OF CONTENTS

I.	INTRODUCTION	5
II.	BACKGROUND AND QUALIFICATIONS	5
III.	LEGAL STANDARDS	9
A.	Patentability	9
B.	Prior Art.....	14
C.	Inherency	15
D.	Person of Ordinary Skill In The Art.....	16
E.	Claim Construction.....	17
F.	Written Description	19
IV.	BASES FOR OPINIONS	20
V.	LEVEL OF ORDINARY SKILL IN THE ART	20
VI.	OPINIONS ON CLAIM CONSTRUCTION	22
VII.	OPINIONS ON WRITTEN DESCRIPTION SUPPORT	23
VIII.	OVERVIEW OF HERLEY682	29
IX.	OPINIONS ON VALIDITY	37
A.	Whether Herley682 Discloses or Renders Obvious Certain Limitations of Independent Claims 1 or 9.....	37
1.	“the [audio/video] data samples comprising representations of audio data”	38
2.	“media monitoring device”	45
3.	“representations of audio data captured at a media monitoring	

device”	50
4. “the raw play stream comprising a sequence of content identification results corresponding to the sequence of [audio/video] data samples” (Claim [1.3] / [9.2]).....	53
5. “an indication of the absence of a respective likely match between a respective [audio/video] data sample and an [audio/video] data representation in the electronic database” (Claim [1.5] / [9.4]).....	56
B. Whether the Combination of Herley682 and Seet was Obvious	60
1. Overview of Fundamental Differences between Herley682 and Seet	61
2. Whether it was Obvious to Combine	68
C. Whether the Claims 4 and 12 are Unpatentable.....	72
D. Whether the Claims 7 and 15 are Unpatentable.....	76
X. CONCLUSION.....	79

Exhibit List

Exhibit No.	Description
2001	Expert Declaration of Leo W. Hoarty
2002	<i>Curriculum Vitae</i> of Leo W. Hoarty
2003	Notice of Intent to Issue <i>Ex Parte</i> Reexamination Certificate in Control No. 90/019,375 (Oct. 30, 2024)
2004	PCT Publication WO 2004/030341
2005	PCT Publication WO 2004/044820
2006	PCT Publication WO 2004/040475
2007	MICROSOFT PRESS COMPUTER DICTIONARY (3rd Ed. 1997) (excerpts from)
2008	Hyoung-Gook Kim et al., MPEG-7 AUDIO AND BEYOND, AUDIO CONTENT INDEXING AND RETRIEVAL (2005) (excerpts from)
2009	U.S. Patent No. 7,907,211
2010	Joshua Grossnickle et al. THE HANDBOOK OF ONLINE MARKETING RESEARCH (2001)
2011	Roku, Inc.’s Identification of Proposed Terms for Construction in <i>Anonymous Media Research Holdings, LLC v. Roku, Inc.</i> , Case No. 1:23-cv-01143-DII (served July 3, 2024)

I. INTRODUCTION

1. I have been retained as a technical expert by Anonymous Media Research Holdings LLC (“Patent Owner”) and asked to provide my objective and unbiased opinions concerning U.S. Patent Nos. 8,510,768 (“the ’768 Patent”) and 10,719,848 (“the ’848 Patent”). Specifically, I have been asked to provide my opinions on the patentability of the ’768 and ’848 Patents in view of certain prior art references, the state of the art at the time of the claimed invention, and the knowledge of a person of ordinary skill in the art at the time of the claimed inventions.

II. BACKGROUND AND QUALIFICATIONS

2. In forming the opinions expressed in this Declaration, I have considered and relied upon my education, background, and experience. In addition, I have reviewed and relied upon additional materials in preparation of this Declaration, as described below.

3. I am currently the Principal of WLH Consulting LLC. *My curriculum vitae* is provided (as Exhibit 2002). I studied electronic engineering and computer science at the Ohio State University for two years.

4. I am the founder of six technology startups over the last 30 years. While creating new technology companies over the preceding 30 years, I have raised over \$600M of investment capital from the top venture capital firms as well as from top

technology companies such as IBM, Intel, Cisco, Samsung, among others. The companies that I founded have created jobs for thousands of technology workers.

5. Most of my work has focused on developing set-top technology and digital advertising support systems. Relevant to this matter, an interactive cable television technology company, called ICTV Inc., was founded by me in 1990, and renamed ActiveVideo Networks (AVN) in 2003. The company developed video-on-demand (VOD) as well as virtual cable TV set-top cloud-based systems to deliver video as well as interactive program guide over managed (cable TV) and unmanaged (Internet) networks. My understanding is that AVN is currently providing service to 24 million cable TV set-top boxes in households serviced by Charter and TimeWarner systems throughout the U.S.

6. I founded the Dotcast company in 1999 and served as President and CTO. The company developed a set-top box and service, called MovieBeam, that received both over-the-air (OTA) and broadband (Internet) delivery of feature films stored on an internal hard drive in the set-top box. Movies were delivered slower than real-time on a 24/7 schedule (trickle filled cache). Metadata about each movie was stored in the file header on the internal hard drive. This information included the usual fields of title, actors, director, and importantly, the rating. The set-top featured parental controls that would not play R-rated films without a password, among other useful controls. Consumers selected programming from an internal user

interface guide and purchased movies for rental under similar rules as Blockbuster. Service was live from 2004 to 2007 in the top 50 television markets (DMAs) in the U.S. and was backed by all nine movie studios, led by the Walt Disney Company.

7. In 2015, I was granted a patent for an audio-based automated content recognition (ACR) system and sold it to Vizio, Inc. I was then contracted by the company to assemble a software development team and create the system based on the patent that would run in a smart TV generating fingerprints from TV audio and sending the fingerprints to a cloud-based server system. I also wrote much of server-side software that receives the streaming fingerprints in real-time from the TVs and identifies content currently playing. The Vizio ACR technology is currently running inside of over 20 million smart TVs and is widely used in the television industry for precision daily audience measurement. In some services provided by Vizio using the ACR technology, media related to currently playing programming is returned to the smart TV for display of supplemental information and for targeted advertising.

8. In 2018, I developed and patented a novel technology for video watermarking which places invisible data into the video signal for identification of the video segment. Under contract, I helped develop the software for a dynamic ad insertion system using the video watermarks. The technology is currently in use by major broadcast networks under the name Project OAR. Client software runs on smart TVs that detect and decode video watermarks and then communicates over the

Internet to Content Delivery Network (CDN) servers to download a specific TV ad to substitute for the existing ad based on certain combinations of zip code, demographic profiles, and time of day. Ad swapping actions occur in less than a tenth of a second. Information about the ad, including whether the ad was viewed in its entirety, or whether the person switched the channel, is passed back to the advertiser.

9. Additional technology I developed and patented for Vizio includes a TV ad attribution system using what is colloquially known as “tracking pixels.” The technology allows third-party verification of individual TVs displaying specific TV ads, a service of high value to advertisers.

10. Throughout my career I have received a variety of awards including, for example, the National Cable Television Association (NCTA)—Best New Technology 1996 for ICTV’s interactive TV and VOD service; Consumer Electronics Society (CES)—Best of Show in 2005 and 2006 for the MovieBeam set-top and service; 2023 Emmy Award went to Vizio, Inc for their precision audience measurement service which utilizes my patented audio automated content recognition (ACR) technology for high-resolution monitoring of household TV viewing.

11. I am an inventor on over 30 patents and my patent portfolio has been extensively cited by over a hundred third-party patents. My past work has been cited in many technology articles on cable television as well as on consumer electronics.

12. I am being compensated for services provided in this matter at a rate of \$475.00 per hour for all time spent on it, except for travel time. I am also being reimbursed for reasonable expenses, such as travel expenses. My compensation is not contingent on my opinions, on the outcome of any matter, or on any of the technical positions I explain in this declaration.

13. I have no financial interest in Patent Owner or the '768 or '848 Patents.

III. LEGAL STANDARDS

14. I am not an attorney. For purposes of this Declaration, I have been informed about certain aspects of the law that are relevant to my opinions, as set forth below.

A. Patentability

15. I understand that for an invention claimed in a patent to be valid and patentable, it must be, among other things, new and not obvious in light of what was known and came before it. That which was known or came before the invention claimed is generally referred to as “prior art.”

16. I understand that there are two ways in which prior art may render a patent claim unpatentable. First, the prior art can be shown to “anticipate” the claim.

Second, the prior art can show that the claim was obvious to a person of ordinary skill in the field of the invention at the time the invention was made.

17. I understand that, for a patent claim to be “anticipated” by the prior art, each and every element or requirement of the claim must be found, expressly or inherently, in a single prior art reference. Further, I understand that anticipation requires that the single prior art reference must disclose the elements of the claim arranged or combined in the same way as in the claim.

18. It is my understanding that whether a claimed invention is obvious is a question of law, with underlying factual inquiries that are determined from the perspective of a person of ordinary skill in the art at the time of the claimed invention.

19. It is my understanding that a claimed invention is obvious if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art (a “POSITA”).

20. It is my understanding that obviousness requires consideration of the following underlying factual inquiries:

- determining the scope and content of the prior art
- ascertaining the differences between the claimed invention and the prior art;

- resolving the level of ordinary skill in the art; and
- evaluating objective evidence or “secondary considerations” of non-obviousness, such as commercial success, long-felt but unresolved needs, failure of others, industry praise, and unexpected results.

21. It is my understanding that a reference must be analogous art to the claimed invention in order to be properly considered for use in an obviousness inquiry. It is my understanding that a reference is analogous art to the claimed invention if:

- the reference is from the same field of endeavor as the claimed invention (even if it addresses a different problem); or
- the reference is reasonably pertinent to the problem faced by the inventor (even if it is not in the same field of endeavor as the claimed invention).

22. It is my understanding that whether a reference is reasonably pertinent turns on how the problem faced by the inventor is reflected, either explicitly or implicitly, in the specification.

23. It is my understanding that the following are exemplary rationales that may support a conclusion of obviousness:

- Combining prior art elements according to known methods to yield predictable results;

- Simple substitution of one known element for another to obtain predictable results;
- Use of known techniques to improve a similar method or product in the same way;
- Applying a known technique to a known method or product ready for improvement to yield predictable results;
- Obvious to try – choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success;
- Known work in one field of endeavor may prompt variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations are predictable to one of ordinary skill in the art; and
- Some teaching, suggestion, or motivation in the prior art that would have led one of ordinary skill to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention.

24. In addition to the above, I understand that a person of ordinary skill in the art may resort to logic, judgment and common sense available to him or her at the time of the invention to find a claim feature obvious, even if the prior art relied on does not explicitly state or teach that feature. I understand that when common sense is relied on to find a feature obvious, the reasoning for that conclusion must be articulated with sufficient clarity.

25. I also understand that a patent claim is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art. An obviousness determination requires finding both that a skilled artisan would have been motivated to combine the teachings of the prior art references to achieve the claimed invention, and that the skilled artisan would have had a reasonable expectation of success in doing so. I understand that it is important to be careful not to allow hindsight reconstruction of references to produce the claimed invention, without any explanation as to how or why the references would be combined to produce the claimed invention.

26. Regarding “objective evidence” or “secondary considerations,” I understand that for such evidence to be relevant to the obviousness of a claim, there must be a causal relationship (called a “nexus”) between the claim and the evidence. I also understand that such objective evidence may include: commercial success of products covered by the patent claims; long-felt need for the invention; failed attempts by others to solve the problem addressed by the invention; copying of the invention by others in the field; unexpected results achieved by the invention; praise of the invention; the taking of licenses under the patent by others; expressions of surprise by experts and those skilled in the art at the making or use of the invention; and that the patentee proceeded contrary to the accepted wisdom of the prior art.

27. I understand that a person of ordinary skill is also a person of ordinary creativity. I further understand that the obviousness analysis need not seek out precise teachings directed to the specific subject matter of the challenged claim, but instead can take into account “ordinary innovation” that does no more than yield predictable results, which are inferences and creative steps that a person of ordinary skill in the art would be understood to employ.

28. I understand that sometimes it will be necessary to look to interrelated teachings of multiple patents and publications; the effects of demands known to the relevant field or present in the marketplace; and the background knowledge possessed by a person having ordinary skill in the art. I understand that all these issues may be considered to determine whether there was an apparent reason to combine the known elements in the fashion claimed by the patent at issue.

B. Prior Art

29. I understand that for purposes of this proceeding, prior art includes patents and printed publications in or outside the United States before the effective filing date of the claimed invention. I also understand that prior art includes U.S. patents, published U.S. patent applications, and published international applications (pursuant to the Patent Cooperation Treaty) that themselves have an effective filing date before the effective filing date of the claimed invention, as long as they name another inventor.

30. I also understand that prior art may include any public disclosure, including any products in public use, on sale, or otherwise available to the public, before the effective filing date of the claimed invention. I understand that information that constitutes prior art generally must be information that was publicly disclosed before the effective filing date of the claimed invention. I understand that there may be exceptions to what is prior art but have been instructed that they are not relevant for purposes of my opinions here.

C. Inherency

31. I understand that if a prior art reference does not expressly set forth a particular claim element, the prior art may still disclose that element if it is “inherent” in its disclosure. To establish inherency, it is my understanding that evidence must make clear that the missing descriptive matter is necessarily present in the thing described in the reference. It is my understanding that inherency may not be established by probabilities or possibilities. In other words, it is my understanding that the mere fact that a certain thing may result from a given set of circumstances is not sufficient to establish inherency.

32. I understand that under the principles of inherency, if a prior art device, in its normal and usual operation, would necessarily perform a claimed method, then the method will be considered to be anticipated by the prior art device. I also understand that when the prior art device is the same as a device described in the

specification for carrying out the claimed method, it can be assumed the device will inherently perform the claimed process.

D. Person of Ordinary Skill In The Art

33. I understand that “a person of ordinary skill in the art” or POSITA is a hypothetical person who is presumed to have known the relevant art at the time of the invention. I understand that the term “person having ordinary skill in the art” or POSITA is sometimes used synonymously. Such a person would necessarily have the capability of understanding any scientific and engineering principles applicable to the pertinent art.

34. I understand that factors that may be considered in determining the level of ordinary skill in the art include, but are not limited to, the following:

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

35. I understand that a person of ordinary skill in the art is not a real person, but rather is a hypothetical individual having the qualities reflected by the factors

above. This hypothetical person has knowledge of all prior art in the relevant field as if it were arranged on a workshop wall and takes from each reference what it would teach to a person of ordinary skill in the art.

36. I understand that a person of ordinary skill in the art is a person of ordinary creativity, not an automaton. In many cases a person of ordinary skill will be able to fit the teachings of multiple prior art references together like pieces of a puzzle.

37. I understand that a person of ordinary skill in the art can collaborate with one or more other persons of ordinary skill in the art for one or more aspects with which the other person may have superior expertise, experience, and/or knowledge.

E. Claim Construction

38. It is my understanding that claim construction, or the interpretation of the claims of a patent, is ultimately a question of law. I have been informed that claim terms are generally presumed to carry their ordinary and customary meaning. However, I understand that a patent applicant is his own lexicographer and is entitled to provide alternate definitions of words in common usage, but if the patent applicant uses a term that is in common usage without providing an alternate definition, then that term should generally be understood as having its common meaning.

39. I understand that a person of ordinary skill in the art is deemed to read a claim term not only in the context of the particular claim in which the disputed term appears, but in the context of the entire patent specification and the “file history” of the patent developed during examination by the Patent Office.

40. I understand that the claims, the specification, and the file history are referred to as “intrinsic evidence,” and are to be read as a person of ordinary skill in the art at the time of the invention would have done at the time of the alleged invention.

41. In addition to the intrinsic evidence, I understand that the court or tribunal may also consider other relevant evidence, which is called “extrinsic evidence,” concerning relevant scientific principles, the meaning of technical terms, and the state of the art. Such evidence includes the testimony of expert witnesses. Although extrinsic evidence may be used to resolve ambiguities that may exist after reviewing the intrinsic evidence, I understand that extrinsic evidence cannot be used to contradict a claim construction that is clear from the intrinsic evidence.

42. I understand that some claims are independent, and that those independent claims are complete by themselves. Other claims refer to these independent claims, or other dependent claims, and are “dependent” on those claims. The dependent claims include all of the limitations of the claims on which they depend.

F. Written Description

43. I understand that a claimed invention must have sufficient written description support in the underlying specification upon which it relies, which means the specification must reasonably convey to those skilled in the art that the inventor had possession of the full scope of the claimed subject matter as of the filing date based on an objective inquiry into the four corners of the specification. The written description requirement is satisfied when the specification sets forth enough detail to allow a person of ordinary skill in the art to understand what is claimed and to recognize that the inventor invented what is claimed. I also understand that the level of detail required to satisfy the written description requirement varies depending on the nature and scope of the claims and on the complexity and predictability of the relevant technology. Factors used to evaluate the sufficiency of a disclosure include: 1) the existing knowledge in the particular field; 2) the extent and content of the prior art; 3) the maturity of the science or technology; and 4) the predictability of the aspect at issue.

IV. BASES FOR OPINIONS

44. The opinions set forth in this Declaration are my own. In forming the opinions herein, I have considered the Petition for *Inter Partes Review* filed against the '768 Patent (Paper No. 1 in IPR2024-01054, hereafter “the '768 Petition”) and the Petition for *Inter Partes Review* filed against the '848 Patent (Paper No. 1 in IPR2024-01057, “the '848 Petition”) (collectively, “the Petitions”), the exhibits cited therein, and the additional materials cited in this Declaration (including the exhibits to the Patent Owner’s Preliminary Responses cited herein). I also rely on my education, training, and experience.

45. My analysis of the materials relating to this matter is ongoing and I will continue to review any new material that is provided to me. This Declaration represents certain opinions that I have formed to date and the bases for those opinions. I reserve the right to revise, supplement, and/or amend my opinions based on new information presented to me and on my continuing analysis of the materials cited herein or later provided to me.

V. LEVEL OF ORDINARY SKILL IN THE ART

46. I understand that Petitioner has advanced the following definition of a POSA:

A person of ordinary skill in the art (“POSA”) ... would have had a bachelor’s degree in electrical engineering, computer engineering,

computer science, or equivalent degree with at least two years of relevant industry experience, including in digital media content delivery.

'768 Pet., 2 / '848 Pet., 2.

47. For purposes of this Declaration only, I have applied the definition of a person having ordinary skill in the art ("POSITA") proposed by Petitioner but with the two modifications discussed below. All references to a POSITA here are made pursuant to this modified definition. However, my application of that definition is not a reflection of my agreement or disagreement with it. I reserve the right to provide my own opinion on the level of ordinary skill should there be further proceedings.

48. My first modification is that the proposed definition should not require a college degree, but instead may be satisfied with sufficient industry experience in the relevant field as an alternative to a college degree. Petitioner's proposed definition is unduly narrow because it appears to preclude a person without a formal degree from qualifying, even if such a person has considerable industry experience designing and implementing ACR systems.

49. My second proposed modification is that the "relevant industry experience" should require industry experience designing and/or implementing automated content recognition ("ACR") systems, not simply experience with "digital media content delivery." Automated content recognition is a fundamental

aspect of the claimed inventions and the technology described in the Challenged Patents. Indeed, it is the *recognition or identification* of media content that is key, not simply delivering media content.

50. As discussed above, I have two years of college education in the field of electrical engineering, as well as several decades of industry experience designing and implementing ACR systems. I have also been awarded numerous patents covering ACR technologies. My experience and qualifications as described in Section II above exceed that of a POSA and qualify me to testify as to the knowledge of a POSA under any reasonable definition.

51. With these understandings and modifications, my testimony herein relates to the knowledge of a POSA as of May 27, 2004. I understand this is the priority date of the inventions relied on by the Petitioner, and have been asked to use this priority date for the purposes of my testimony. At this time, I take no position on whether this is the correct priority date.

VI. OPINIONS ON CLAIM CONSTRUCTION

52. I have applied the plain and ordinary meaning to a POSA of all claim terms at issue. To the extent I believe an explanation of the plain and ordinary meaning may be helpful, I have provided it where appropriate below.

VII. OPINIONS ON WRITTEN DESCRIPTION SUPPORT

53. I have been asked to offer my opinion on whether U.S. Provisional Application No. 60/545,681 (“Seet Provisional”) provides adequate written description support for claim 9 of U.S. Patent Application Publication No. 2005/0193016 (“Seet”). I have also been informed and understand that for a claim to have adequate written description support, the document relied on for that support must demonstrate that the inventor(s) possessed the full scope of the claimed invention. For the reasons below, it is my opinion that the Seet Provisional does not provide adequate written description support for claim 9 of Seet.

54. The Seet provisional only discloses matching **audio** streams, and in fact specifically defines “media” as being limited to audio as later explained. In contrast, Seet expressly discloses matching audio and/or **video** streams. Consistent with that disclosure, claim 9 of Seet broadly claims matching audio and video media streams by virtue of it claiming “media streams” rather than the narrower term “audio streams”:

9. A method for generating entries for a media content database, the method comprising:

receiving a first **media** stream;

receiving a second **media** stream;

matching at least some portions of the second media stream

to the first media stream to determine repeating content

between the streams; and

presenting the repeating content to a reviewer for identification.

Ex. 1006, 6.

55. Because the claims of Seet plainly encompass matching video data streams (as expressly taught in the specification of Seet) and that concept is not taught or suggested by the Seet Provisional, it is my opinion that the Seet Provisional does not provide written description support for the full scope of claim 9. In other words, the Seet Provisional does not demonstrate that the inventor possessed a system that could match video streams as encompassed by claim 9. Although the Seet Provisional does disclose audio stream matching, that does not support the scope of claim 9 that encompasses video stream matching. I address each of the Seet Provisional and Seet in turn.

56. The Seet Provisional does not mention or discusses using video data from its media streams to match content. The only reference to “video” at all is the statement: “Instead of using a high-quality original master recording (say a CD, or digital video tape, etc), we suggest using the same audio signal that is to be identified....” Ex. 1019, 4. However, it is clear to a POSA that only the audio signal of the “video tape” is actually being referred to. If the video signal was being used, this sentence would refer to “using the same video signal.”

57. Moreover, although “media” would ordinarily be understood to refer to audio and video streams, the Seet Provisional states “Media **will be defined** as **any**

streaming audio source in any form, be it radio, television, cable, satellite, internet stream or the like” where “[a]udio can be music, speech advertisements, promotions or the like.” *Id.* A POSA would thus understand that the “media” in the Seet Provisional is expressly and narrowly defined to refer to “streaming **audio**.” While television is indeed mentioned, it is clear to a POSA that only the audio portion of the television signal is used to match against other audio portions of that same signal, particularly in view of the express definition of “media” quoted above. There is simply nothing in the Seet Provisional that suggests matching video data (e.g., pixel matching) to identify content.

58. The remaining portions of the Seet Provisional are consistent with this understanding. For example, the “Background” describes “One of the biggest challenges for an **audio** identification or ‘**audio** fingerprinting’ company is the development and maintenance of a database of **audio** against which identifications are attempted.” *Id.* It further states that “[i]n order to identify when an ad is playing, the current process is to get the high-quality original ad from the advertiser, extract and normalize **the audio layer**, extract the frequency characteristics (fingerprint **the audio**), then insert the fingerprint and ad information into a database for subsequent retrieval.” *Id.* A POSA would thus understand that the “Background” only discusses matching audio streams.

59. Turning to the section entitled “The Invention,” it likewise makes clear to a POSA that the Seet Provisional is only directed toward analyzing **audio** streams for content identification. For example, it states “[i]nstead of using a high-quality original master recording (say a CD, or digital video tape, etc), we suggest using the same **audio** signal that is to be identified, but from a previous period.” *Id.* That same section explains that this “could be accomplished by dividing up **the recorded audio** for the day into small (~4 second) increments and storing each increment with a unique identifier.”). *Id.*, 4-5. There is simply no mention or suggestion that video streams are or could be matched to identify content. And although the Seet Provisional does mention “samples” without explicitly identifying them as audio samples (e.g., “this database of previous samples”), it is entirely clear from surrounding context discussed above that the samples are audio samples—not video. A POSA would thus find nothing in this section that indicates or suggests the inventor of Seet possessed a system that matches video streams.

60. The Petition identifies various portions of the Seet Provisional as allegedly providing written description support. ’768 Pet., 17-20 / ’848 Pet., 17-20 (claim chart). However, each portion of the Seet Provisional relied on by Petitioner is clearly related to analyzing and matching audio data, not video data, as discussed above. For example, the Petition cites to the title of the Seet Provisional (“Media Content Database Creation and Maintenance Using Time-Delayed Auto-Correlation

of Frequency Characteristics”). ’768 Pet., 18 / ’848 Pet., 18 (citing Ex. 1019, 4). However, as I explained above, the Seet Provisional explicitly defines “media” as being limited to only “streaming audio.” A POSA would thus understand “media” as used throughout the Seet Provisional (including the title) to refer only to streaming audio. Additionally, the portions cited in Petitioner’s chart that refer to “samples” would clearly be understood to refer to audio samples (not video), because audio samples and audio fingerprinting is the only type of content identification described in the Seet Provisional. Ex. 1019, 4-6.

61. I next turn to the specification of Seet itself, which expressly discloses matching both audio and video streams. For example, Seet teaches “if the media stream 102 is a radio broadcast, the media items 104n may comprise a song, an advertisement, a segment of a radio show, or any other type of audio content typical of such broadcasts. Alternatively, **if the media stream 102 is a video broadcast**, each media item may comprise a television show or segment, an advertisement, or **any other type of video** or audio/video content typical of such broadcasts.” Ex. 1006

¶ [0026]. Seet further teaches:

A number of embodiments are described above with reference to audio content; however, **it should be understood that various embodiments of the media content database system could be used with any other type of media content or other types of media items.** As used herein, media items and media data may include information used to represent a media or multimedia content, such as all or part of an audio and/**or video file**, a data stream having media content, or a transmission of media content. Media content may include **one or a**

combination of audio (including music, radio broadcasts, recordings, advertisements, etc.), **video** (including movies, video clips, television broadcasts, advertisements, etc.), software (including video games, multimedia programs, graphics software), and **pictures**; however, this listing is not exhaustive. Furthermore, media data, media items, and media content include anything that itself comprises media content, in whole or in part. Media data, media items, and media content can be encoded using **any encoding technology**, and they may also be encrypted to protect their content using an encryption algorithm or any other suitable encryption technique.

Id. ¶ [0050]. A POSA would thus understand Seet to expressly teach analyzing and matching video media streams, including even where no audio is present.

62. Consistent with that understanding, the IPR Petition challenging the '848 Patent—which has almost identical claims to the '768 Patent but that are directed toward “video data samples” rather than “audio data samples”—Petitioner acknowledges that “Seet recognizes that although embodiments are described with respect to audio content, **its ‘media content database system could be used with any other type of media content,’ including a video file.**” '848 Pet., 27 (citing Ex. 1006, ¶¶[0026], [0050]). In other words, Petitioner acknowledges that the language from paragraphs [0026] and [0050] of Seet is disclosing the matching of video streams, not simply audio streams as disclosed in the Seet Provisional.

63. For the reasons discussed above, a POSA would understand Seet to disclose matching video media streams to identify content and that claim 9 broadly encompasses video media stream matching. Because such a concept is not taught or

suggested by the Seet Provisional, a POSA would not have understood the inventors of Seet to have possessed the full scope of claim 9 based on the specification of the Seet Provisional.

VIII. OVERVIEW OF HERLEY682

64. Before turning to the Petition's invalidity theories, it may be helpful to first describe the operation of Herley682. Herley682 is directed to "a system and method for sampling one or more channels of a broadcast spectrum, such as a radio frequency spectrum, identifying and storing content information for each sampled channel, and providing a user interface for allowing interactive user queries and display of the stored content information." Ex. 1005 ¶ [0002]. Stated another way, the system of Herley682 is a broadcast-channel-centric content identification system. This refers to how it receives raw media directly from a broadcast channel source (e.g., radio station) so that it can analyze and identify the content (e.g., songs) being transmitted over a specific broadcast channel (e.g., radio station frequency). Herley682's system is directed toward auditing services, such as verifying ad impressions and song plays for copyright requirements where the identity of the media segment is crucial.

65. A POSA would thus understand that Herley682 is not concerned with tracking the media consumption habit of audience members. Nor is Herley682's broadcast-channel-centric system capable of doing so, as it is limited to cataloging

identified broadcast content by directly tuning to the broadcast source. This does not provide any insights whatsoever into what content audience members are consuming, but only provides insights into what content the broadcast source is distributing. Indeed, if only the source of the broadcast channel content is being tuned to, the habits of the audience consuming that content cannot be monitored.

66. Herley682 may better be understood in reference to Figure 2, which I have reproduced below. It depicts an exemplary architectural diagram of an implementation of the broadcast-channel-centric system:

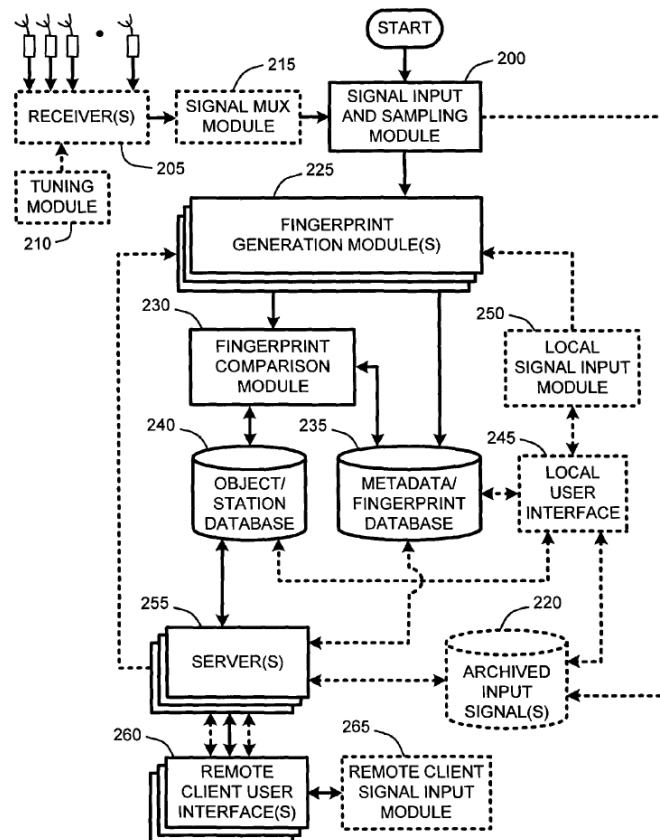


FIG. 2

Herley682 explains that “radio or television broadcasts are captured using one or

more receivers 205.” Ex. 1005 ¶ [0075]. Once the receiver 205 receives one or more radio or television broadcast signals, they are then passed to the signal input and sampling module 200 to generate samples. *Id.* ¶ [0084].

67. Next, “signal samples are continuously provided [by sampling module 200] in real-time to one or more fingerprint generation modules 225.” *Id.* (alterations added). The fingerprints are generated using a “signal fingerprint extraction algorithm, i.e., a ‘fingerprint engine,’ for deriving traces from segments of one more signals.” *Id.* ¶ [0067]. “These traces are referred to as ‘fingerprints’ since they are used to uniquely identify the signal segments from which they are derived.” *Id.* ¶ [0063]. Herley682 further explains that the fingerprints are “derived” from the sample generated by module 200 by using the algorithm to “calculate” the trace fingerprint. *Id.* ¶ [0077] (“[T]he stream might be sampled for a time great enough to calculate a trace fingerprint [I]f one fingerprint is calculated per minute for a repetitive FM radio station....”). In other words, the fingerprints are generated based on (or “derived from” or “calculated” based upon) the samples, but are not themselves part of the samples. This is an important fact to keep in mind, as the manner in which fingerprints are generated is relevant to certain of Petitioner’s mappings of Herley682 to the independent claims.

68. The calculated fingerprints are then provided from fingerprint generation module 225 to fingerprint comparison module 230. *Id.* ¶ [0086].

Fingerprint comparison module 230 searches a metadata/fingerprint database 235 to determine if there is a match between the extracted fingerprint and pre-computed fingerprints stored in database 235. *Id.* “If a match is identified ... then the sampled portion of the input signal from which the fingerprint was extracted is identified as belonging to an object of interest corresponding to the fingerprint and metadata in the metadata/fingerprint database 235.” *Id.* “This information, including the metadata, is stored in an object/station database 240.” *Id.* The database 240 can then be queried by a user of the system to pull statistics and other information regarding the content played on the particular broadcast channel. *Id.* ¶¶ [0125]-[0167].

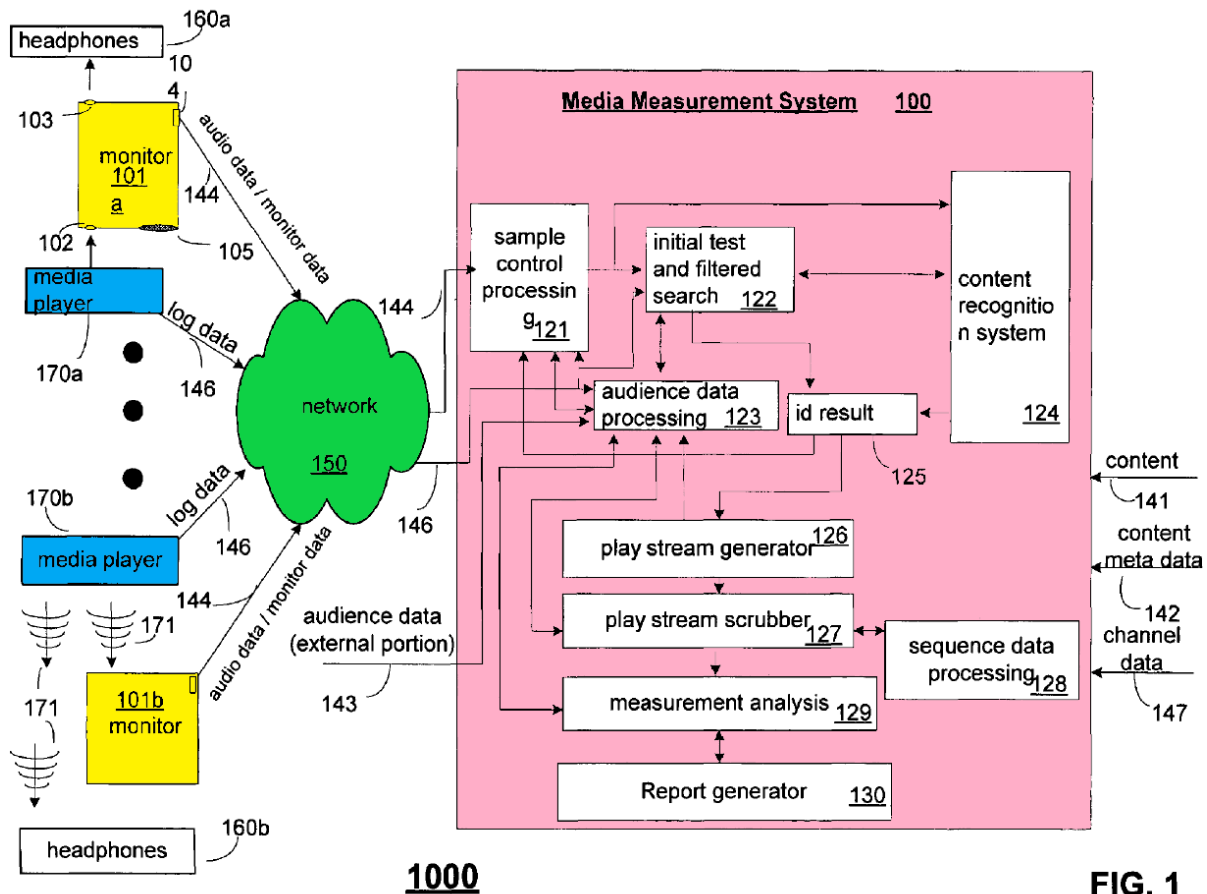
69. A POSA would understand that the broadcast-channel-centric system of Herley⁶⁸² is fundamentally different than the system of the ’768/’848 Patent. The ’768/’848 Patent is directed toward a media usage and monitoring system for tracking the *consumption of media*. As the “Background of the Invention” explains:

Traditional media measurement systems have focused on **directly monitoring** channels being utilized by audience members. However, as media consumption patterns have become more complex, channel-centric media measurement **is inadequate for many purposes**. It may be desirable to track usage of particular media content independent of channel. Furthermore, although “channel” identification in a traditional media measurement system may sometimes be limited to radio or television broadcast station, it is increasingly desirable to track usage of media across several types of media delivery vehicles including

radio, television, CD, DVD, computer download, portable media players (e.g. MP3 players, iPod), and other vehicles.

'768 Patent, 1:14-26 / '848 Patent, 1:22-34. The inventors of the '768/'848 Patent thus recognized that broadcast-channel-centric systems—like Herley682—were “inadequate for many purposes.” The inventors also recognized that existing media content recognition systems had drawbacks, such as reliance on embedded “tags” or on computationally intensive database searches. '768 Patent, 1:32-56 / '848 Patent, 1:40-54. In view of these drawbacks, the inventors determined that “an improved media measurement system and method is needed.” '768 Patent, 1:50-55 / '848 Patent, 1:56-62. To that end, their media measurement system monitors media consumption, while also increasing the accuracy of content identification and reducing computational overhead and data transmission requirements via “scrubbing.” The manner in which the '768/'848 Patent accomplishes this is described below.

70. Figure 1 is provided below for reference. It depicts a media measurement system 1000 comprising media monitoring devices 101a and 101b (blue) that each monitor a respective media player 170a and 170b (yellow). '768 Patent, 2:66-3:25 / '848 Patent, 3:5-32.



Each media monitoring device 101 monitors the media being “consumed” by the audience member(s) of a respective media player 170 and provides “audio data” associated with the media player, along with “monitor data” 144 (e.g., device ID), over network 150 (green). ’768 Patent, 3:6-18, 3:38-46 / ’848 Patent, 3:12-24, 3:45-53. The Specification teaches that the audio data may be “raw audio data” that the monitor receives from the media player, which may further “undergo[] one of a variety of levels of processing” by the monitor, such as pre-calculating signatures (“sometimes referred to as audio fingerprints”). ’768 Patent, 3:43-54 / ’848 Patent, 3:50-62. In other words, the fingerprints or signatures may be generated or captured

by the media monitoring devices 101.

71. A POSA would understand this can help reduce the amount of data needed to be transmitted over the network, by eliminating the need for the monitoring devices 101 to send the raw audio signal over the network to be fingerprinted. For example, assume a digital audio stream being monitored by the system at uncompressed audio data rates of 48,000 samples per second at 16 bits per sample or 768,000 bits per second. Whereas a fingerprint for automated content recognition might, for example, consist of data blocks aggregating between 100 to 600 milliseconds of audio. If the system used a 300-millisecond block of audio data per fingerprint, that would compress 255,740 bits into, for this example, just 280 bits for that 300-millisecond block of audio. This would result in a network data rate of 924 bits per second, an almost three order of magnitude reduction in network traffic.

72. I next turn to Figure 4, which is reproduce below. Once the media measurement system 100 receives the audio data and monitor data from a media monitoring device 101, a raw play stream 410 is generated. '768 Patent, 9:12-31 / '848 Patent, 9:25-45. This involves “itemiz[ing] a sequence of analyzed samples and relates that sequence to corresponding content identification results,” which indicates a likely match with a particular piece of content (e.g., “song 51”) or the absence of a likely match (e.g., “none”):

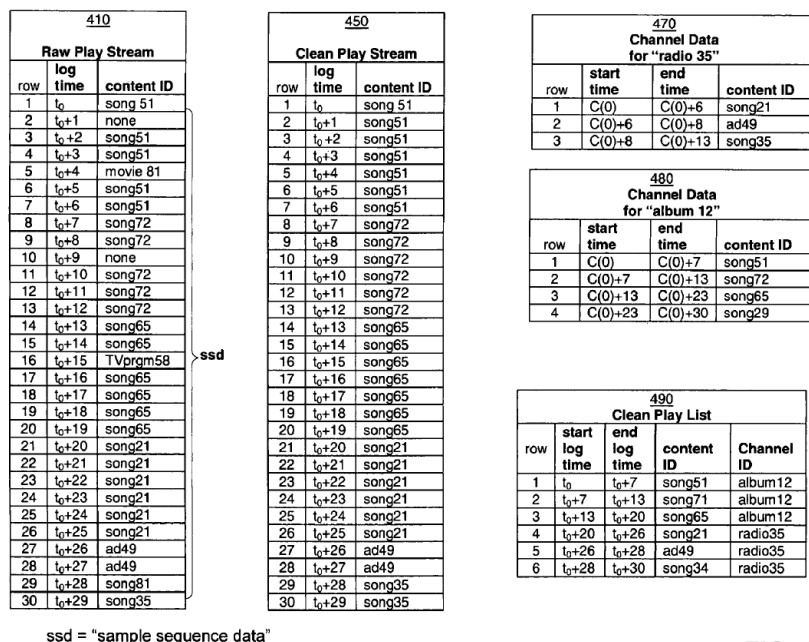


FIG. 4

'768 Patent, 9:12-31 / '848 Patent, 9:25-45.

73. The raw play stream 410 is then "scrubbed" to generate clean play stream 450 by "[b]y analyzing the sequence of data on Raw Play Stream" such that "it is possible to locate and correct any identifications that are likely to be incorrect or to supply missing identifications where the system has not otherwise been able to provide an identification." '768 Patent, 9:32-37 / '848 Patent, 9:46-50.

74. A POSA would understand that the invention of the '768/'848 Patent overcomes the fundamental inadequacy of broadcast-channel-centric systems—like Herley682¹—by observing the media consumption of audience members' media

¹ I note that Herley864 and Seet are both broadcast-channel-centric systems suffering from this same inadequacy. Each of those references likewise tunes directly to a

players using media monitoring devices. In so doing, the invention is able to track audience media consumption habits, as well as “scrub” results to improve content identification accuracy. This is in contrast to the broadcast-channel-centric system of Herley682, which simply receives a raw broadcast channel signal and identifies the content being broadcast by that channel, which provides no insights into audience consumption. Moreover, the system of Herley682 also lacks the ability to “scrub” its content identification results, which is another major improvement of the ’768 Patent over the Herley682 system.

IX. OPINIONS ON VALIDITY

A. Whether Herley682 Discloses or Renders Obvious Certain Limitations of Independent Claims 1 or 9.

75. I understand that for Grounds 1–3, the Petitions rely exclusively on Herley682 as disclosing or rendering obvious most limitations of independent claims 1 and 9. As discussed below, it is my opinion Herley682 that does not disclose or render obvious several of those limitations.

broadcast-channel source to identify content being played over the channel, but provide no insights into audience media consumption habits. Ex. 1007, 1:15-19; Ex. 1019, 4.

1. “the [audio/video] data samples comprising representations of audio data”

76. Claims 1 and 9 each recite: “data for a sequence of [audio/video] data samples, **the [audio/video] data samples comprising representations of [audio/video] data** captured at a media monitoring device.” I understand that Petitioner contends that the samples generated by sampling module 200 are “the [audio/video] data samples” and that the fingerprints generated by fingerprint generation module 225 are “representations of captured audio data.” See *e.g.*, ’768 Pet., 38 / ’848 Pet., 39 (“Sampling module 200 receives and samples these broadcast signals, and sends **the samples** (*data for a sequence of [audio/video] data samples*)”); ’768 Pet., 39 / ’848 Pet., 40 (“Herley682’s **fingerprints** are *representations of captured [audio/video] data* as claimed”); ’768 Pet., 40 / ’848 Pet., 41-42 (“a POSA would have further understood that Herley682’s broadcast signals, **samples** (*data for a sequence of [audio/video] data samples*), and **fingerprints** (*representations of [audio/video] data captured*) are provided respectively to signal input and sample module 200...”); Ex. 1003 (to ’768 Petition) ¶ 173 / Ex. 1003 (to ’848 Petition) ¶ 168 (“Herley682’s sampling module 200 receives and samples these broadcast signals, and sends the **samples** (*data for a sequence of audio data samples*) to fingerprint generation module(s) 225, which in turn extracts fingerprints from the samples In other words, Herley682’s **fingerprints** are *representations of captured [audio/video] data* as claimed.”).

Importantly, however, the plain claim language recites “[audio/video] data samples **comprising** representations of [audio/video] data.” The “representations of [audio/video] data” must thus be **part of** the “[audio/video] data samples” by virtue of the language “**comprising**”. In other words, the “[audio/video] data samples” must include the “representations of [audio/video] data.”

77. The fingerprints (alleged “representations of capture [audio/video] data”) generated by fingerprint generation module 225 are not part of the audio samples (alleged “[audio/video] data samples”) generated by module 200. Nor could they be. Instead, a POSA would understand that the fingerprints are generated based on (or “derived from”) the samples received from module 200. As shown in Figure 2 below, the sampling module 200 (red) generates the samples from a broadcast signal, and the samples are then passed (as indicated by the red arrow) to the fingerprint generation module 225 (green):

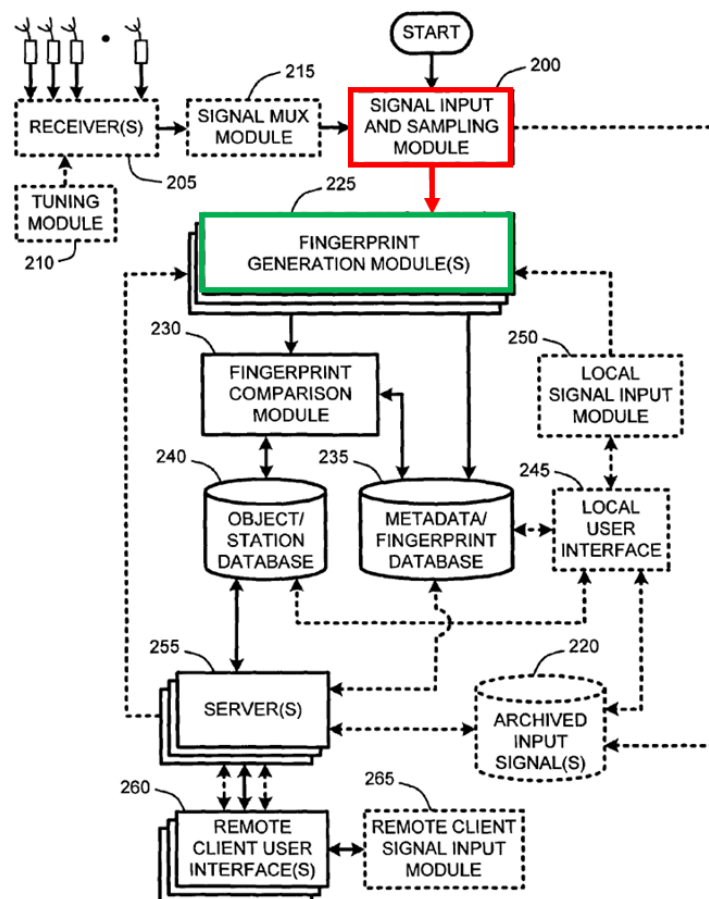


FIG. 2

Ex. 1005 ¶ [0075], [0084].

78. Herley682 goes on to explain that:

One common method for automatically identifying such objects involves analyzing an input signal, or a predefined portion or segment of such a signal [e.g., an audio sample], **to produce a set of parameters or features that are derived from the signal**. These parameters or features are then stored as “fingerprints” that uniquely identify such objects. These fingerprints may then be used to identify subsequent occurrences of objects in a similar signal.

Id. ¶ [0005] (alterations added). Herley682 further teaches that the “interactive signal analyzer uses a signal fingerprint extraction **algorithm**, i.e., a ‘fingerprint engine,’

for **deriving** traces from segments of one or more signals. These traces are often referred to as ‘fingerprints’ since they are used **to uniquely identify** the signal segments **from which they are derived**.” *Id.* ¶ [0013]; *see also e.g., id.* ¶¶ [0015], [0017], [0063]. In other words, Herley682 is expressly teaching that its “fingerprints” are **derived from** a raw signal or samples based on the signal, which means the fingerprints are not part of the samples. Herley682 further explains how the fingerprints are “calculated” (or “computed”) based on the samples using a “signal fingerprint extraction algorithm, i.e., a ‘fingerprint engine,’” which a POSA would understand is how the fingerprints are “derived” from the samples. Ex. 1005 ¶¶ [0063], [0077]; *see also id.* ¶ [0007] (“Once ***computed***, these fingerprints are typically stored in a database of known objects.”).

79. A POSA would readily understand and appreciate that audio and video fingerprints (also referred to as “signatures”) are calculated **based on** (or **derived from**) samples—the fingerprints are not themselves part of, or included within, the samples. While this was common knowledge to any POSA, I briefly discuss several pieces of evidence below demonstrating this knowledge.

80. The Specification of the ’768/’848 Patent discusses PCT International Publication WO 2004/030341 entitled “Fingerprint Extraction,” which accurately states:

A fingerprint, also often referred to as signature or hash, is a sequence

of bits that **is derived from** multimedia content, e.g. an audio song, an image, a video clip, etc. Multimedia fingerprints are used, inter alia, in the field of authentication where it is desired to verify whether received content is original or detect whether the content has been tampered with. Fingerprints are also used to identify media content.

Ex. 2004, 1:6-14 (cited by '768 Patent, 5:52-53 and '848 Patent, 5:60-64).

81. This is consistent with the meaning of “fingerprint” as defined by numerous other sources as well. *See e.g.*, Ex. 2005, 1:6-12 (entitled “Fingerprinting Multimedia Contents” and explaining that “Fingerprints, in the literature sometimes referred to as hashes or signatures, are binary sequences extracted from multimedia contents, which can be used to identify said contents.”) (cited by '768 Patent, 5:49-52 and '848 Patent, 5:60-63); Ex. 2006, 1:6-25 (entitled “Improvements in and relating to fingerprint searching,” and explaining that “Hashes of multimedia content have been referred to as robust hashes ... but are now commonly referred to as multimedia fingerprints.”).

82. A “hash” or “signature” is a value generated based on underlying data, which in the context of “fingerprinting” acts as a shorthand representation of features of an audio/video signal or sample. Ex. 2007, 4 (defining “hash”). Petitioner even acknowledges that the fingerprints are “used to uniquely identify the signal segments **from which they are derived.**” '768 Pet., 39 / '848 Pet., 40 (citing Ex. 1005 ¶

[0013]); *See also* Ex. 1008 ¶ [0020] (explaining that “each fingerprint **represents a number of features** of the media sample at each location”).

83. The fundamental purpose of fingerprinting is to calculate a value (e.g., a hash or signature) representative of features of the underlying sample, which can then be used for comparison rather than comparing the underlying sample itself. Such fingerprints “can be seen as a summary or perceptual digest of the audio recording” (also called a “signature” or “hash”), which can be compared to other fingerprints. Ex. 2008, 10. Fingerprinting “eliminates the direct comparison of the (typically large) digitized audio waveform as an efficient and effective approach to audio identification.” *Id.* In other words, fingerprinting eliminates the need to compare audio/video samples directly to other audio samples, but instead allows for the comparison of fingerprints **representative of** features of audio/video samples. Ex. 1005 ¶¶ [0005]-[0007] (explaining that fingerprinting involves “produc[ing] a set of parameters or features that are **derived from** the signal” and “[o]nce **computed**, these fingerprints are typically stored in a database of known objects”); Ex. 2009, 2:37-44 (“A fingerprint of a multimedia object/content/signal is a **representation of perceptual features** of the object/content/signal part in question. Such fingerprints are sometimes also known as ‘(robust) hashes.’”).

84. Petitioner and its expert do not account for the “comprising” language at issue here. *See generally*, ’768 Pet., 38-41 / ’848 Pet., 39-42. Petitioner cites to

Herley682's teaching that fingerprint generation module 225 "extracts fingerprints from the samples," perhaps misunderstanding this to mean that the fingerprints are part of the audio/video sample. '768 Pet., 38 / '848 Pet., 39-40. However, a POSA would understand that when Herley682 refers to "extracting" fingerprints from the audio/video samples, it is simply referring to how the fingerprints are "derived from" or "calculated" based on the audio/video samples, as evidenced by the numerous exhibits discussed and cited above.

85. Nothing in the Specification or Herley682 teaches or suggests that a fingerprint is somehow part of the audio/video sample upon which it is based. Nor does the Specification or Herley682 refer to or describe fingerprints (or signatures) as being part of the audio/video sample from which a fingerprint was derived. Indeed, it is simply impossible for the fingerprints of Herley682 identified by Petitioner to be the claimed "representations of [audio/video] data" because those fingerprints are not part of the alleged "[audio/video] data samples" (generated by module 200). Indeed, the '768 / '848 Patent explicitly teaches that fingerprints are calculated or derived. '768 Patent, 3:46-47 / '848 Patent, 3:53-56 (describing audio fingerprints or signatures as "calculated"); '768 Patent, 4:21-24 / '848 Patent, 4:28-34 (explaining that "signatures" and "fingerprints" are "derivable").

86. In conclusion, it is my opinion that Herley682's fingerprints (alleged "representations of [audio/video] data") are not included in, or part of, the identified

audio samples (alleged “[audio/video] data samples”) and thus Herley682 does not disclose or render obvious “the [audio/video] data samples comprising representations of [audio/video] data” as claimed.

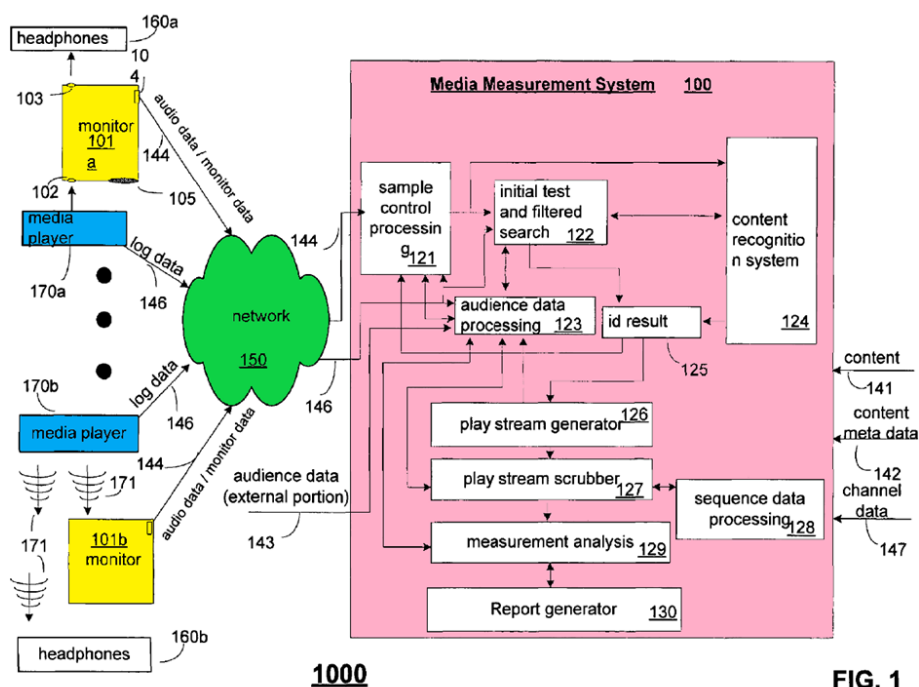
2. “media monitoring device”

87. Claim [1.1] and [9.P] each recite “the [audio/video] data samples comprising representations of [audio/video] data **captured at a media monitoring device.**” Petitioner identifies the receiver 205 in Herley682 as the alleged “media monitoring device” without explanation or citing any intrinsic or extrinsic evidence regarding the meaning of that term. ’768 Pet., 38 / ’848 Pet., 39. Petitioner is mistaken that a POSA would understand the receiver 205 of Herley682 to be a “media monitoring device” as claimed.

88. Turning first to Herley682, it teaches that the receiver 205 is directly tuned to a radio or television broadcast. The reason the system is designed this way is because it is broadcast-channel-centric—it is designed to identify audio content being played on one or more broadcast station channels, not to monitor media consumption. Notably, the ’768/’848 Patent criticizes such channel-centric measurement as “inadequate for many purposes,” and explains “it is increasingly desirable to track usage of media across several types of media delivery vehicles including radio, television, CD, DVD, computer download, portable media players

(e.g. MP3 players, iPod), and other vehicles.” ’768 Patent, 1:14-32 / ’848 Patent, 1:21-39.

89. To that end, in reference to Figure 1 (annotated below) the ’768/’848 Patent discloses media monitoring devices 101a and 101b (yellow) to monitor the media consumption habits of individual audience members’ media players 170a, 170b:



’768 Patent, 3:4-25 / ’848 Patent, 3:10-32. The media monitoring devices 101 not only upload over the network 150 audio data (e.g., fingerprints) associated with audio consumption at the media players 170 of audience members, but also “monitor data” such as device ID. ’768 Patent, 3:38-46, 4:49-55 / ’848 Patent, 3:45-53, 4:58-67. A POSA would understand that by using media monitoring devices 101, the

media consumption habits of audience members are tracked, overcoming a fundamental inadequacy of broadcast-channel-centric systems like that of Herley682.

90. It is important to understand that the term “media monitoring device” is not an industry-recognized term. As mentioned above, the ’768/’848 Patent expressly criticizes channel-centric measurement systems as “inadequate” and explicitly addresses this inadequacy via the use of media monitoring devices 101. The specification also repeatedly and uniformly teaches that media monitoring devices 101 are associated with media players so that the media consumption habits of individual audience members can be observed, thereby overcoming the problem with broadcast-channel-centric systems like Herley682. *See e.g.*, ’768 Patent, abstract, 2:12-15, 2:66-3:55; 6:58-65 / ’848 Patent, abstract, 2:17-22, 3:5-62, 7:1-8. This is true with respect to each and every embodiment in the patent, and there is no teaching or suggestion that a media monitoring device could simply a receiver tuned to a broadcast station.

91. The ’768/’848 Patent also consistently describes the media monitoring devices 101 as providing not only audio/video data (whether raw audio/video or audio/video fingerprints) but monitor data (e.g., device ID). ’768 Patent, 3:39-49, 4:49-51 / ’848 Patent, 3:45-57, 4:58-62. Device ID, for example, could be used to identify the particular media monitoring device among a plurality of media

monitoring devices. In the context of media measurement, identifying the devices associated with many different audience members can be very important (e.g., to help track location, audience member identity, etc.). In light of these exclusive and repeated teachings, a POSA would understand that a *media monitoring device* monitors media consumed by a media player.

92. Moreover, the '768/'848 Patent is exclusively directed towards a “media measurement system.” That is an established, known term in the industry that refers to the monitoring of audience media consumption. Ex. 2010, 11 (“If you are familiar with the Nielsen Ratings, you are familiar with **media measurement**. This type of research is the primary tool of media planners, who use the information to device on the placement of their advertising campaigns. **Media measurement** provides the subscriber with estimates of the number of people **who consume various types of media** (radio, TV, print, online, outdoor, etc.) and **the characteristics of those people**.”). Consistent with that fact, claim 1 recites a “media monitoring device” and a “media measurement method.” In stark contrast, Herley682 never mentions or discusses “media measurement,” but rather is broadcast-channel-centric. Consequently, a POSA would understand and interpret the claimed “media monitoring device” to refer a device that monitors media consumption by a media player.

93. A POSA would not understand receiver 205 of Herley682 to be a “media monitoring device.” The receiver does not monitor media consumed by a media player but rather is tuned directly to one or more broadcast channels, thereby presenting a broadcast-channel-centric system that is inadequate for tracking audience consumption. Notably as well, the receiver 205 of Herley682 is not capable of providing monitor data, but rather only provides raw media data broadcast directly received from a station, essentially just a radio. The Petitioner does not identify anything in the specification that teaches or suggests that a “media monitoring device” could be a basic radio receiver, and a POSA would not refer to a basic radio receiver standing-alone as a “media monitoring device.”

94. Additionally, under Petitioner’s broad interpretation the media measurement system 100 (of Figure 1) would also necessarily have a “media monitoring device” (e.g., a receiver) that receives the media content from the monitoring devices 101 over the network 150. However, the Specification never refers to or suggests that the system 100 has such a device. Instead, the only “media monitoring device” disclosed or referenced by the Specification is a device that captures media consumed by media players, which provides a fundamental advantage over prior art like Herley682.

95. In conclusion, it is my opinion that Herley682 does not teach or suggest the claimed media monitoring device.

3. “representations of audio data captured at a media monitoring device”

96. Claim [1.1] recites “receiving over a network, at one or more computers, data for a sequence of [audio/video] data samples, the [audio/video] data samples comprising *representations of [audio/video] data captured at a media monitoring device.*” Claim [9.P] similarly recites “the one or more computers being coupled to a network to receive, over the network, data for a sequence of [audio/video] data samples, the [audio/video] data samples comprising *representations of [audio/video] data captured at a media monitoring device.*”

97. Petitioner takes the position, in reference to Figure 5 of Herley682 (annotated below), that receiver 205 (red) is the “media monitoring device” and that fingerprints generated by fingerprint generation module 225 (green) constitute “representations of captured [audio/video] data:”

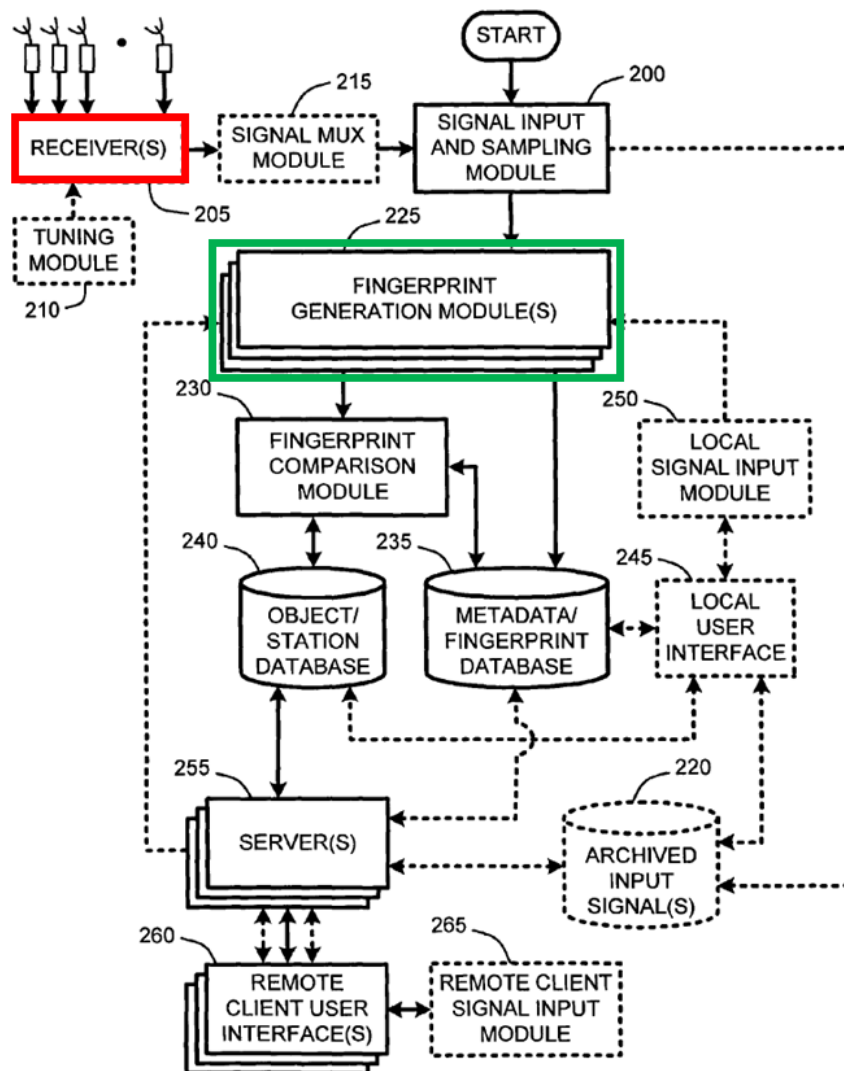


FIG. 2

'768 Pet., 39 / '848 Pet., 40 (“Herley682’s fingerprints are *representations of captured [audio/video] data* as claimed”); Ex. 1003 to '768 Petition ¶¶ 170 (“Thus, the generated fingerprints are ‘representations of audio data captured at a media monitoring device as claimed.’”), 171 (“a POSA would have further understood that Herley682’s fingerprints (the representations of audio data captured) are provided to fingerprint comparison module 230....”); Ex. 1003 to '848 Petition ¶¶ 165-66.

Petitioner is mistaken because the claim language clearly recites “the audio data samples comprising representations of audio data captured at a media monitoring device.” In other words, the “representations of audio data” are what is “captured at a media monitoring device.” The identified fingerprints cannot be the claimed “representations of audio data,” because they were not “captured at” receiver 205 (alleged “media monitoring device,” red).

98. Herley682 tunes receiver 205 to a broadcast radio or television channel, and the raw audio received is passed along to signal mux module 215, then signal input and sampling module 200, and then fingerprint generation module 225. Ex. 1005 ¶¶ [0075], [0084], [0086]. The fingerprints (alleged “representations of audio data”) are thus not “captured **at**” the receiver 205 (alleged “media monitoring device”). Instead, the fingerprints are captured at fingerprint generation module 225. This is in stark contrast to the teachings of the ’768/’848 patent, which teaches that the media monitoring devices 101 may generate the signatures or fingerprints² and then transmit them over the network 150. ’768 Patent, 3:38-51 / ’848 Patent, 3:43-59 (explaining that the audio data provided by monitoring devices 101 may “comprise pre-calculated signatures...”); ’768 Patent, 4:48-57 / ’848 Patent, 4:64-

² The ’768/’848 Patent explains that “signatures” and “fingerprints” have the same meaning. ’768 Patent, 3:46-48 / ’848 Patent, 3:53-57 (“audio signatures (sometimes referred to as audio fingerprints)”). This is consistent with the understanding of a POSA.

67 (explaining that audio data may be “divided into samples ... at monitor 101”); ’768 Patent, 4:65-67 / ’848 Patent, 5:7-10 (explaining that audio samples are “captured”).

99. In conclusion, it is my opinion that Herley682 does not teach or suggest that alleged “representations of audio data” (fingerprints generated at module 225) are “captured at” the alleged “media monitoring device” (receiver 205) as claimed.

4. “the raw play stream comprising a sequence of content identification results corresponding to the sequence of [audio/video] data samples” (Claim [1.3] / [9.2])

100. Petitioner corresponds “the collection of data stored in Herley682’s object database 240” to the claimed “raw play stream” but appears to acknowledge that the collection of data is not in sequential, chronological order. ’768 Pet., 43-44 / ’848 Pet. 43-45; *See also* ’768 Pet., 47 / ’848 Pet., 48 (arguing that “the data in Herley682’s object databases 240 corresponds to the claimed *raw play stream*”). To address this, Petitioner argues that “it would have been obvious ... at a minimum, to sort the data **in Herley682’s object database 240** based on the time data to put it in sequential, chronological order.” ’768 Pet., 44 / ’848 Pet., 45.

101. Petitioner does not provide any explanation as to why a POSA would want to put the data in sequential order, including why there would be any desire by a POSA for the objects in Herley682’s objects database 240 to be sequentially ordered. Regardless, a POSA would have found no reason to do so because the

system's user can simply query the database to generate reports in the desired format. Ex. 1005 ¶¶ [00125]-[00158], [0179]-[0186] (describing comprehensive nature of database query tools and report generation).

102. Although the object entries may have time stamps, that is simply irrelevant to whether there would be any reason to store them in the object database 240 sequentially. The stamps are simply for the user to know when the content was played as part of the query, not for sequentially ordering content within the database. In fact, even though Herley682 teaches that its object database 240 captures "each new occurrence," where there are identical occurrences Herley682 "increments a counter for that object **rather than creating a new entry** for the already played object." Ex. 1005 ¶ [0029]. In other words, Herley682's object database is specifically designed such that it does not even store every occurrence of an identified object. The fact that Herley682's object database does not even store every occurrence of an identified object is further evidence that Herley682 has no reason to store objects in sequential order as part of a "raw play stream."

103. A POSA would understand that the reason the claimed invention requires a "raw play stream" comprising "a **sequence** of content identification results corresponding to the **sequence** of audio data samples," is because the sequence is used to perform the later recited "scrubbing." Herley682 simply has no

need for sequentially ordering its objects because the user can query the database for object information, regardless of whether the objects are stored sequentially.

104. I understand that Petitioner also contends that “it would have been obvious to store the data in sequential order as the data in the stream is processed.” ’768 Pet., 44 / ’848 Pet., 45. I note again that Petitioner provides no articulated rational as to why it would have been obvious to do so. In any event, even if it was obvious to store the data in sequential order as the data in the stream is processed, Petitioner does not explain how this temporary processing data stream could somehow be used for the “scrubbing” required later by the claim. For instance, how would this allegedly obvious sequential data stream later be used by Herley682 to “scrub”? Where would the allegedly obvious sequential data stream be stored in Herley682? Petitioner does not say or explain because, again, Herley682 has no reason to store a sequential order of objects for later use. Thus, even if it was obvious to process the data sequentially as part of a data stream, it was not obvious to store that data sequentially for later use.

105. Petitioner asserts that “Seet provides an express teaching of maintaining the temporal integrity, i.e., chronological order, of the sample data, confirming that a POSA would have been motivated to do so.” ’768 Pet., 44 / ’848 Pet., 45. However, a POSA would readily appreciate that the reason Seet maintains the temporal integrity is not applicable to Herley682. Seet teaches “maintaining the

temporal integrity” because it is relying on the time-aligned sequence to infer matches for unmatched content. Ex. 1019, 4-5. Herley682 has no reason to similarly maintain temporal integrity in its object database because it is not inferring matches as in Seet. In other words, a POSA would not be motivated to store Herley682’s objects sequentially based on Seet, because Seet’s time-aligned chronological ordering is for the purpose of something (inferential matching) not taught or suggested by Herley682. Thus, the mere fact that Seet may maintain its results in chronological order for its own distinct purpose, does not mean that a POSA had any reason to do so with Herley682’s object database, which is for a different purpose.

106. In conclusion, it is my opinion that a POSA would not have been motivated to modify Herley682’s object database to store objects in sequential order as part of a play stream based on Seet.

5. “an indication of the absence of a respective likely match between a respective [audio/video] data sample and an [audio/video] data representation in the electronic database” (Claim [1.5] / [9.4])

107. Petitioner asserts that the “scrubbing” recited in Claim [1.7] / [9.6] would be obvious based on the combination of Herley682 and Seet, where Seet is relied on “to identify and correct **missing data** in order to obtain a complete and accurate data set in the object/station database 240” of Herley682. ’768 Pet., 52-54 / ’848 Pet., 52-55. In other words, Petitioner does not propose that it would be obvious to correct inaccurate matches, but only that it would be obvious to infer

matches for “unmatched objects.” Thus, despite the conjunctive “or” in Claims [1.5] and [9.4], Petitioner’s invalidity theory is premised on showing that the “scrubbing” of Claim [1.7] / [9.6] is based on identifying and correcting missing data (not correcting misidentified data).

108. Petitioner argues, in reference to Figure 2 of Herley682, that the metadata/fingerprint database 235 of Herley682 is the *electronic database*. ’768 Pet., 46 / ’848 Pet., 47.

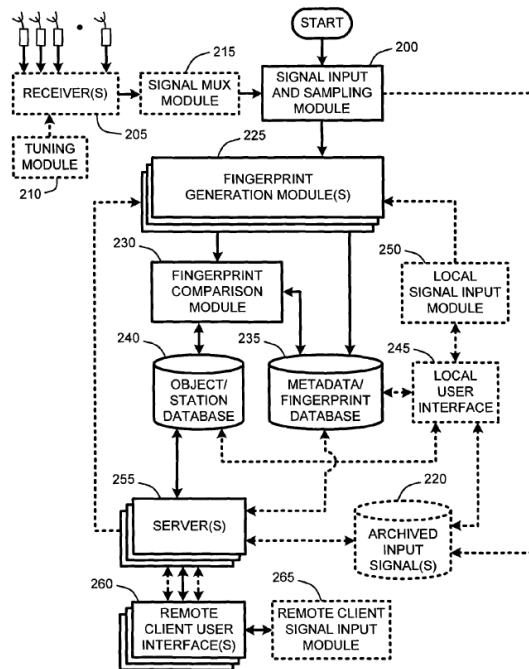


FIG. 2

Petitioner next contends that a “POSA would also have understood that these teachings indicate that Herley682’s **object database 240** would include multiple entries therein, e.g., one for each song played on the station,” and further that a “POSA would have understood or at a minimum found it obvious that, at least most

of the time ... these multiple entries would ... include *at least one content identifier* and an *indication of the absence of a respective likely match*, e.g., **an ‘unknown object’** as discussed above for element [1.5] part (ii).” ’768 Pet., 48 / ’848 Pet., 49. In other words, Petitioner argues that Herley682 discloses or renders obvious that its object database 240 would include “an indication of the absence of a respective likely match” as claimed.

109. Petitioner is mistaken, because Herley682 expressly teaches that “unknown objects” within the object database 235 indicate that a match **has been** made. More specifically, Herley682 teaches that if the fingerprint extracted from a sample does not match any entries in the metadata/fingerprint database 235, that “fingerprint is stored **to the metadata/fingerprint database** as an ‘unknown object entry.’” Ex. 1005 ¶ [0088]; *see also id.* ¶ [0164]. Herley682 then explains that “[i]f any subsequent occurrences of an object having a matching fingerprint are identified, **then both the first instance of the unknown object, and each subsequent instance, will then be added to the object/station database 240** along with any statistical information that is available for that object.” *Id.* ¶ [0088]. Thus, the “unknown object” is only stored in the object database 204 ***after*** it has been matched to a subsequent or prior fingerprint for the same “unknown object” in the metadata/fingerprint database 235. *Id.* In other words, under Petitioner’s mapping, the fact that the “unknown object” is stored in object/station database 240 is an

indication **of a match** (not an absence of a match) between a respective audio data sample and an audio data representation in the electronic database (alleged to be the fingerprint database 235).

110. This is because the “unknown object” is only sent to the object database 240 if it matches a fingerprint in the fingerprint database 235 (the alleged *electronic database*). Thus, this is the opposite of what the claim limitation requires, which is “an indication **of the absence** of a respective likely match.” This is perfectly logical in the context of Herley682, which wants to avoid including in its object database 240 unknown objects that are associated with “noise” that will never recur (e.g., a radio disc jockey talking in between songs). Herley682 instead ensures that there are multiple matched fingerprints indicating the identification of a repeating object of interest within the fingerprint database 235 before sending the unknown object to the object database 240.

111. For the reasons above, Petitioner is mistaken that the “unknown objects” within object database 240 meet this claim limitation. Although Petitioner also argues that a POSA would have “found it obvious,” Petitioner does not explain why it would be obvious to modify this purposeful design feature of Herley682. Herley682 is specifically designed to originally store its “unknown objects” in the separate and distinct fingerprint database 235 and only send them to the object database 240 after they match a prior or subsequent fingerprint in database 235. A

POSA would not have been motivated to modify this feature solely for the sake of modification, because it provides the advantage of preventing unknown objects that are simply “noise” from being sent to the object database for later manual identification by a human reviewer.

112. Petitioner further contends that “this feature would have been obvious in view of Seet” because “Seet expressly teaches a database of samples that include multiple entries, both identified samples and unidentified samples.” ’768 Pet., 48-49 / ’848 Pet., 49-50. As an initial matter, I again note that Petitioner does not explain why a POSA would have been motivated to modify Herley682’s object database 240 to include the “unmatched samples” of Seet when Herley682 purposefully only stores its “unknown objects” in the object database 240 **after a match is made** to a fingerprint in fingerprint database 235. It simply makes no sense for Herley682 to store such unknown objects in object database 240 without first matching them to fingerprints in fingerprint database 235 as expressly taught.

B. Whether the Combination of Herley682 and Seet was Obvious

113. I understand that Petitioner contends that it would have been obvious to modify Herley682 to rely on the alleged “scrubbing” taught by Seet. In attempting to combine Herley682 and Seet, I believe Petitioner overlooks fundamental differences in their operation that renders the combination non-obvious.

1. Overview of Fundamental Differences between Herley682 and Seet

114. It is helpful to first understand the operation of Herley682 and Seet and their fundamental differences. Turning first to Herley682, in reference to Figure 2 the system captures radio or television broadcasts by directly tuning receiver 205 to a broadcast station. Ex. 1005 ¶¶ [0075], [0084]. Once the receiver 205 receives one or more radio or television broadcast signals, they are then passed to the signal input and sampling module 200 to generate samples, which are then passed to fingerprint generation module(s) 225 to create fingerprints. *Id.* ¶¶ [0067], [0084].

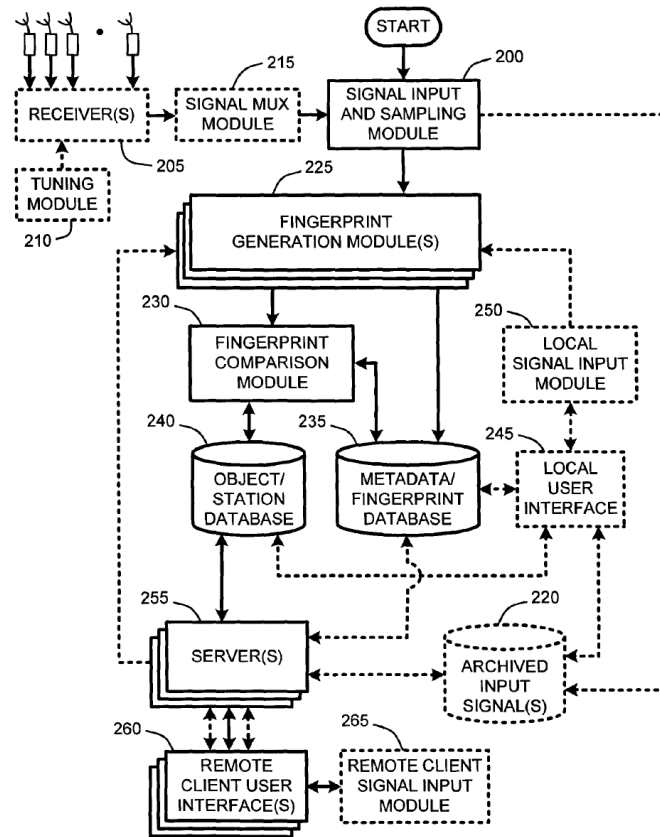


FIG. 2

115. The calculated fingerprints are then provided from fingerprint generation module 225 to fingerprint comparison module 230. *Id.* ¶ [0086]. Fingerprint comparison module 230 searches a metadata/fingerprint database 235 to determine if there is a match between the extracted fingerprint and pre-computed fingerprints stored in database 235. *Id.* “If a match is identified ... then the sampled portion of the input signal from which the fingerprint was extracted is identified as belonging to an **object of interest** corresponding to the fingerprint and metadata in the metadata/fingerprint database 235. This information, including the metadata, is stored in an **object/station database 240.**” *Id.* Herley682 explains, “for example, in the context of an audio signal, such objects include, songs, jingles, advertisements, station identifiers, program ‘signature tunes’, emergency broadcast signals, speech from one or more known speakers, etc.” *Id.* ¶ [0012]; *see also id.* ¶ [0004].

116. Herley682 further teaches that “where trace fingerprints generated from samples do not match any fingerprints in the database, those trace fingerprints are added as fingerprints to the fingerprint database as ‘unknown objects.’” *Id.* ¶¶ [0016]; *see also id.* ¶ [0088]. The boundaries of the “unknown object” (e.g., song) are then determined (*Id.* ¶ [0016]), and when two or more occurrences of the “unknown object” are identified based on a fingerprint match in database 235, that object is then stored in the object database 240 as an “unknown object.” *Id.* ¶¶

[0088]-[0089]. Each “object” in the object database (whether known or unknown) thus corresponds to an object of interest (e.g., a complete song, jingle, or ad).

117. Herley682 explains that the “unknown objects can be manually identified by a user via a local user interface 245, or alternately, via a remote client user interface 260” or a “more authoritative and up-to-date fingerprint database” can be queried. *Id.* ¶¶ [0089], [0017]. Herley682 is thus fundamentally premised on having a database of known objects from which to match its fingerprints. To the extent its matching results in unknown objects, either a human reviewer simply reviews those objects for manual identification or a more robust fingerprint database is queried.

118. Seet operates in a fundamentally different manner than Herley682. In contrast to Herley682, Seet does not rely on a database of fingerprints associated with known content. Instead, the sole purpose of Seet is to create its own database of known audio content to avoid the need to “[g]ain[] access to the source audio and then fingerprinting it and inserting it into a database,” which is described as “a complex process from a licensing and update management process.” Ex. 1019, 4. In other words, Seet wants to avoid having to procure or obtain the database of fingerprints associated with known content, and instead seeks to create its own database from scratch.

119. To accomplish this goal, Seet teaches using the same audio signal that is to be identified, but from a previous period. *Id.* For example, for television or radio monitoring, Seet proposes “that the previous period (i.e. day or day-period) for that particular station (or genre, or group of similar stations), be fingerprinted, and inserted into a fingerprint database.” *Id.*³ During the subsequent period, “matches are then performed against this database of previous samples, and if matched, then the identifier (or identifiers if multiple-matches are found) and a time-stamp are recorded.” *Id.* In other words, Seet is relying on the signal from a prior day to create fingerprints and then match them to the next-day’s fingerprints from the same signal.

120. An additional organization step follows, where Seet “attempts to group audio together by examining relative alignment of identified sequential samples against the previous day samples in the database.” *Id.*, 5. Seet teaches that this establishes a relationship between the previous period’s database and the current period’s database such that it can be assumed that “the portions of the previous period database that were matched are repeated content like advertisements, promotions, songs or the like.” *Id.* This allows a human reviewer to “manually listen to just the series of identified samples to add meta-data (like ad name, or artist, and title information) to the series of samples. *Id.*

³ In the media measurement industry, the term “day-part” is used to refer to various portions of the day being monitored. For example, “morning,” “afternoon,” “prime-time,” and “late night,” are examples of “day-parts.”

121. I have provided Figure 1 of the Seet Provisional below for discussion:

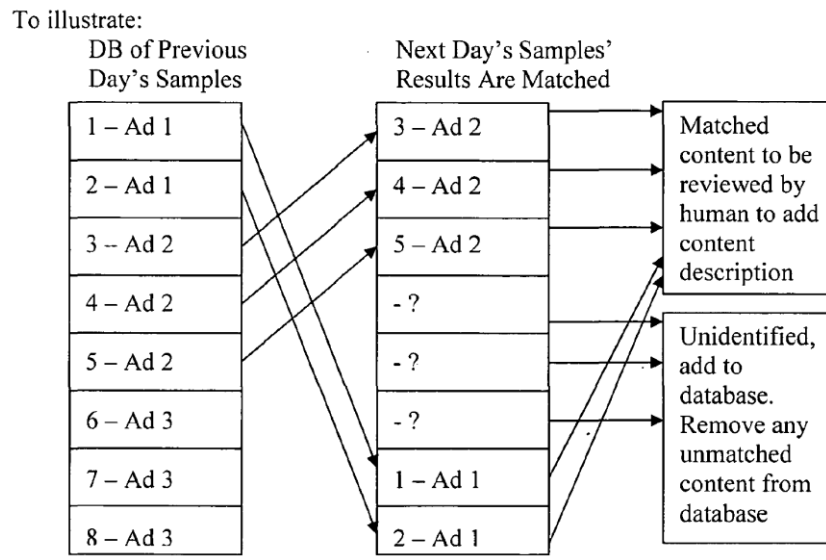


Figure 1. Auto-correlation Example

It is important to note that the Seet Provisional makes clear that there is no database of known content, but rather that the content for the samples is identified by a human reviewer after the relationship of matching samples (without content identifiers) has been established. *See also* Ex. 1006 ¶ [0029] (“Although the system does not initially know the identity of each sample, the identity is shown in the figure for explanation purposes.”). In other words, the labels (Ad #1, Ad #2, Ad #3) in Figure 1, and the associations that they imply, are unknown to Seet’s system. Seet instead matches samples without knowing their content identity, groups them based on their time-aligned relationship, and then a user manually identifies the content of the samples (e.g., by adding metadata to the database).

122. While Seet does refer to “identifiers” when discussing Figure 1, that is simply a reference to the “unique identifiers” assigned to the samples (e.g., the numbers “1-”, “2-”, etc.) not the identifiers of the content of those samples, which is only added later via manual review by a human. Ex. 1019, 4-5 (“This could be accomplished by dividing up the recorded audio for the day into small (~4 second) increments and storing each increment with a unique identifier.”). A POSA would thus understand that the purpose of Seet’s system is to enable the creation of a database of known content by manual review by grouping matched samples (with unknown content), which makes it easier and quicker for the human reviewer to manually identify content that has already been grouped together by samples.

123. Although Seet does teach “inferring” matches, it does not do so in the manner described by the Petition. Figure 2 (reproduced below) may, at first glance, appear to depict inferring content identifiers. However, it is actually only inferring sample matches without knowing the identity of the content, as explained above.

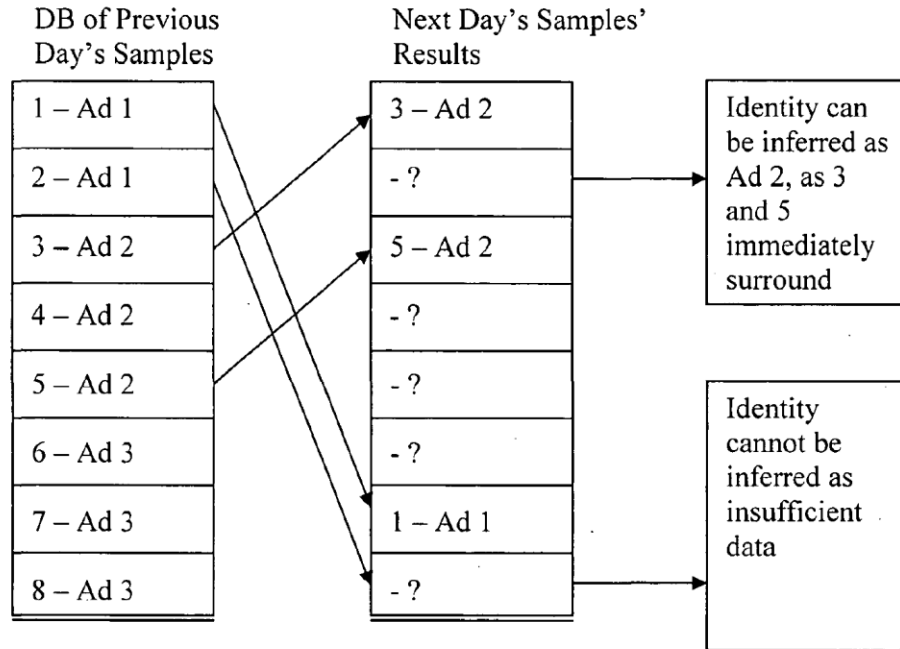


Figure 2. Auto-correlation Identity Inference Example

In other words, in the example shown, Seet is merely inferring that the sample for “?” (between the sample for “3 –” and the sample for “5 –”) matches fingerprint “4 –” from the previous day. It is not inferring the *identity* of the content intervening sample “?” because, as explicitly taught by Seet, it does not know the identity until after a human user manually adds it to the database later. Seet thus does not rely on a database of known content and, even though it does “infer” matches, it does not infer the identity of those matches. Nor does Seet operate at the “object” level, as in Herley682. Instead, Seet operates at the “sample” level and relies on the sampling of repeating signals to manually create a database of known content.

2. Whether it was Obvious to Combine

124. Petitioner proposes modifying Herley682 to rely on Seet’s inferential sample matching, arguing that it would be obvious to infer the identity of “unknown objects” in Herley682’s object database 240. ’768 Pet., 52-53 / ’848 Pet., 54 (“A POSA would have thus been motivated to identify and correct missing data in order to obtain a complete and accurate data set **in the object/station database 240**—a teaching that is explicitly taught by Seet.”).

125. Petitioner overlooks the fact that Herley682 is operating at the “object” level, whereas Seet is operating at the “sample” level. Petitioner instead incorrectly suggests that the objects of Herley682 are analogous or interchangeable with the samples of Seet. ’768 Pet., 32-33 / ’848 Pet., 33-35. However, as I previously explained, in Herley682 an “unknown object” in object database 240 corresponds to an object of interest (e.g., an entire, unknown song), not merely a sample portion associated with part of an object of interest. Ex. 1005 ¶ [0016] (explaining that the boundaries of the object are determined before storing it in the object database). Thus, when Herley682’s “unknown objects” are stored in object database 240, there is no expectation or reason for there to be temporal consistency for those objects in the database. Moreover, given that the “objects” correspond to complete pieces of content (e.g., a complete song), not individual samples associated with a piece of

content, the sample inference methodology of Seet would not work for identifying Herley682's "unknown objects."

126. It is important to recognize that Seet is fundamentally premised on having temporal consistency for its individual, repeating audio samples so that it can infer a sample match based on surrounding sample matches. Seet uses the fact that a repeating song will have a series of fingerprints within it that can be relied upon to repeat in the same sequential order.

127. Even if one were to accept that the objects in Herley682's object database are sequentially ordered,⁴ there would be no expectation that the sequence of objects (e.g., songs) would repeat. Seet is premised on not only having temporal alignment and consistency between the signal being analyzed and the previous day's signal, but also analyzing those signals at the **more specific sample-level** to inferentially match fingerprints. In other words, while the sequence of the surrounding samples in Seet is expected to be the same (because it is analyzing individual samples associated with a single song, for example), the sequence of the surrounding objects (e.g., songs) in Herley682 is not expected to be the same. Indeed, radio stations do not play the same songs, ads, etc. in the same sequence day-by-day, making it impossible to somehow infer "unknown objects" in Herley682

⁴ As explained *supra* IX.A.4, there was no reason or motivation to rely on a play stream of sequentially ordered content identifiers in Herley682.

based on Seet. Likewise, television stations do not broadcast television shows or TV programs in the same sequence day-by-day, likewise making it impossible to somehow infer “unknown” video objects. In short, because the “objects” in Herley682’s object database are not stored at the sample level, Seet’s inferential sample matching methodology would not work.

128. Additionally, POSA would not have been motivated to rely on Seet’s inferential matching with Herley682, because Seet’s methodology is premised on the signal being time-aligned with the signal from the previous day. Petitioner simply assumes that Seet’s inferential matching method would have been recognized as workable with Herley682’s system with virtually no modification. However, Petitioner does not identify anything in Herley682 that indicates or suggests that it is analyzing repeating signals or that, even if it was, those signals are time-aligned as required pursuant to the teachings of Seet. Petitioner overlooks this fact in attempting to combine Herley682 and Seet as if their systems are similar. A POSA would have recognized, however, that Seet is premised on time-alignment of a repeating signal, and that modifying Herley682 to inferentially match as taught by Seet would require a fundamental redesign of Herley682’s system. Petitioner does not address or discuss this redesign, but instead assumes that such a modification would be simple without explanation.

129. Petitioner also does not explain why a POSA would rely on Seet's inferential matching given that it relies on a repeating signal and time-alignment of samples based on the signal. In other words, why would a POSA believe that the inferential matching of Seet—which assumes this time-alignment between individual sequential samples—would help to identify “unknown objects” in Herley682's object database given that those objects are not stored in a time-aligned or sequential fashion? Petitioner does not address this fundamental question.

130. Petitioner also states that “[a] POSA would have been motivated to use Seet's teachings of inferring matches to identify Herley682's unmatched objects at least because a POSA would have been motivated to automate that which is done manually because it would be faster, a stated goal of Herley682.” '768 Pet., 32, 53 / '848 Pet., 33, 54. However, Seet also relies on manual review of unknown content just like Herley682 does for objects that it is unable to identify content for. While Seet is able to infer fingerprint matches (assuming time-aligned sample matching) as discussed above, the identity of the content associated with those fingerprints is still not known. Seet merely infers that an unmatched fingerprint does in fact match a prior day's fingerprint, but it does not infer or know the content identity for that fingerprint. Instead, a human reviewer must manually review those matching fingerprints to identify their content, just as Herley682 relies on manual review for the content of its “unknown objects.” Consequently, a POSA would not have been

motivated to rely on Seet's inferential matching with Herley682 to "automate" the identification of unknown content, because Seet itself also relies on manual review to identify unknown content. Even if Seet's inferential sample matching could be combined with Herley682, it would not solve the purported problem proposed by Petitioner of having to manually review to identify unknown content.

C. Whether the Claims 4 and 12 are Unpatentable.

131. Claims 4 (dependent from claim 1) and 12 (dependent from claim 9) recite wherein "scrubbing the raw play stream further comprises using audience data as part of determining whether to change a result of the sequence of content of identification results." Petitioner asserts that this limitation is taught by the combination of Herley682, Seet and Herley684 in Ground 2. '768 Pet., 71 / '848 Pet., 72. Petitioner acknowledges that the Specification teaches that "audience data" may include "age, gender, and other demographic information **about the audience member,**" "media consumption patterns **of the audience member,**" or "location **of the audience member.**" '768 Pet., 71 / '848 Pet., 72. Petitioner thus appears to acknowledge that "audience data" is, at a minimum, data about an audience member or an audience member's consumption habits.

132. Petitioner then argues, however, that information about the genre of music being played by a broadcast station somehow constitutes "audience data." '768 Pet., 72 / '848 Pet., 73. Petitioner is mistaken, because there is a

fundamental difference between knowing the genre of music *a particular audience member typically listens to* (i.e., an example of “audience data”) and simply knowing the genre of a song being played on the radio. The former is information *about an audience member*, where the latter is simply information about the type of song being broadcast on a channel, which provides no insights about the audience members listening to that song. A POSA would thus not understand the mere identification of a genre of music being played on a radio station to be “audience data.”

133. Herley682, Seet, or Herley864 do not, and cannot possibly teach, using “audience data” to perform “scrubbing,” because all three of those references are broadcast-channel-centric measurement systems. All three focus on identifying content being played by a broadcast channel (e.g., a radio station)—those systems have no insights whatsoever into the audience listening to that content. Although Petitioner cites to portions of Herley682, Seet and Herley684 to support its position that those references teach using “audience data” as claimed, those citations do no support that proposition.

134. Petitioner first discusses Herley682, stating that it: (1) describes identifying its “unknown objects” is by “automatically querying a more authoritative and up-to-date fingerprint database.”; (2) teaches that “when listening to a particular channel, the fingerprints of previously identified objects from that channel are

searched first in the database, thereby speeding the search;” and (3) describes that “when listening to a station that plays pop hits from the 1980’s repetitively it makes sense to search the fingerprints of songs previously identified on that station, before searching the larger database.” ’768 Pet., 71-72 / ’848 Pet., 72-73. Petitioner does not explain how this is somehow using “audience data” as claimed. It clearly is not. At most, these statements merely teach using fingerprints for songs previously played on a station to identify future songs on that same station. This provides no insights whatsoever into the audience consuming that content.

135. Petitioner next argues:

Herley864 similarly teaches limiting searches to results of a particular music genre (as opposed to results from a particular station) as a way to reduce search and comparison times. EX1007, 9:10-20, 12:44-52. With this teaching in mind, and in order to “speed[] the search” when querying the additional database for identifying “unknown objects” in Herley682, it would have been obvious to likewise use Herley682’s teaching of first searching previously identified objects from the channel, as modified by Herley864 to first search objects of a user’s preferred or typical genre (*using audience data*), when querying this additional database. EX1003, ¶268. This would have been the use of known technique to improve similar devices (methods, or products) in the same way. EX1003, ¶268.

’768 Pet., 72-73 / ’848 Pet., 73-74. However, neither Petitioner nor its expert cite any evidence from Herley682 or Herley864 showing that either is capable of relying

on an *audience member's* “preferred or typical genre” to perform the claimed “scrubbing.” Indeed, the “user” referenced by Petitioner would simply be the person operating the media measurement system that is identifying media content—not the audience member consuming the content. Even if that user was selecting “genre” as a search feature or filter, the selection is not based on the media consumption habits of the user. Rather, the selection is based on the user’s desire to determine what content was being played by a particular broadcast source for commercial/advertising research purposes.

136. In conclusion, it is my opinion that the combination of Herley682, Seet and Herley684 do not disclose or render using “audience data” as claimed.

D. Whether the Claims 7 and 15 are Unpatentable.

137. Claims 7 (dependent from claim 1) and 15 (dependent from claim 9) recite “wherein an [audio/video] data representation of the plurality of [audio/video] data representations includes a plurality of signal landmarks.” Petitioner acknowledges that Herley682, Herley864 and Seet do not teach this limitation, and instead argues in Ground 3 that it would have been obvious to include landmarks into the underlying combinations based on Wang060. ’768 Pet., 73-74 / ’848 Pet., 74-75.

138. In reference to Figure 2 of Herley682 (reproduced below), Petitioner corresponds “the plurality of [audio/video] data representations” as claimed to Herley682’s trace fingerprints calculated by fingerprint generation module 225 (green). ’768 Petition, 41 / ’848 Petition, 42.

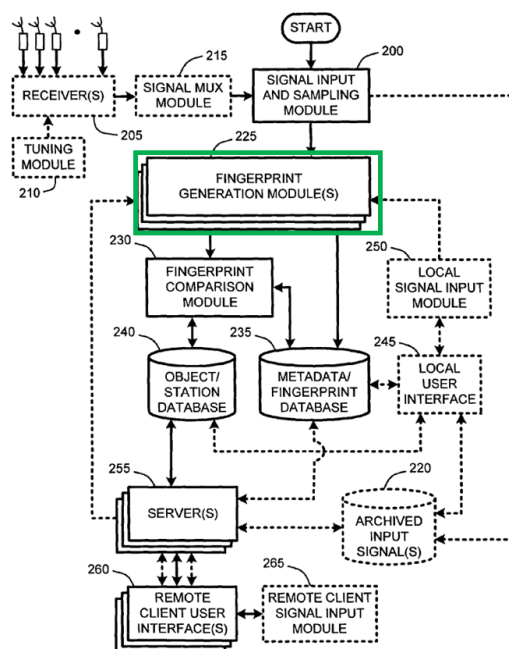


FIG. 2

Petitioner is thus corresponding “an [audio/video] data representation” recited in this claim to one of those fingerprints. Petitioner next argues that it would have been obvious to modify Herley682 to include landmarks as claimed. ’786 Pet., 74-75 / ’848 Pet., 75-76.

139. A fundamental problem with Petitioner’s theory is that Wang060 does not teach or disclose including landmarks *in* fingerprints. The plain claim language recites that “an [audio/video] data representation ... **includes** a plurality of signal landmarks.” Because Petitioner contends that the “audio data representation” is Herley682’s fingerprint, Petitioner must thus show that, under its proposed theory of invalidity, the prior art discloses a fingerprint that “**includes**” landmarks.

Petitioner has not done so because Wang060 does not teach that its landmarks are included in fingerprints.

140. Even accepting that Wang060 teaches signal landmarks, a POSA would understand those signal landmarks are not included in, or part of, Wang060's fingerprints. Wang060 teaches that "Landmarks occur at specific timepoints *of the sound* and have values in time units offset from the beginning of the file, while fingerprints characterize the sound *at or near* a particular landmark." Ex. 1008, ¶ [0049], Fig. 4. Wang060 further explains that "landmarks are first computed, and then fingerprints are computed *at or near* the landmarks." *Id.* ¶ [0048], Fig. 3. Thus, the landmarks are not included in, or part of, the fingerprints. Instead, a landmark is data that is associated with, but external to, the fingerprints. Indeed, Wang060 further explains that "An index set for a given sound recording is a list of pairs of values (fingerprint, landmark)"—the index set is thus what associates the fingerprint with the landmark. Ex. 1008 ¶ [0074]. In other words, the fingerprint and landmark are two separate, distinct things that the index associates with one another.

141. In conclusion, Wang060 does not disclose including landmarks in fingerprints (the alleged "an audio data representation"). Consequently, even if Wang060 was combined with Herley682 as proposed, the proposed combination would not meet the claim requirements.

X. CONCLUSION

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

Dated: November 14, 2024



W. Leo Hoarty