

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ROKU, INC.

Petitioner,

v.

ANONYMOUS MEDIA RESEARCH HOLDINGS LLC.,

Patent Owner.

IPR2024-01055 (Patent No. 8,756,622)
IPR2024-01058 (Patent No. 10,719,849)

EXPERT DECLARATION OF W. LEO HOARTY

EXHIBIT LIST

Exhibit No.	Description
2002	<i>Curriculum Vitae</i> of Leo Hoarty
2003	U.S. Publication 1002/0038456 to Hansen et al. (“Hansen”)

TABLE OF CONTENTS

I.	INTRODUCTION	5
II.	BACKGROUND AND QUALIFICATIONS	5
III.	LEGAL STANDARDS	9
A.	Patentability	10
B.	Prior Art.....	14
C.	Inherency	15
D.	Person of Ordinary Skill In The Art.....	16
E.	Claim Construction.....	17
IV.	BASES FOR OPINIONS	19
V.	LEVEL OF ORDINARY SKILL IN THE ART	21
VI.	OPINIONS ON CLAIM CONSTRUCTION	23
VII.	OPINIONS ON VALIDITY	23
A.	Whether the Wang References disclose or render obvious the claimed <i>play stream</i>	23
1.	The Wang References do not disclose a <i>play stream</i>	23
2.	The proposed combination of Wang777 and Wang060 is not obvious to a POSA.	34
B.	Whether the Wang References disclose or render obvious <i>identifying a channel based on a sequential order of at least two different obtained content identifiers</i>	36
C.	Whether the combination of Pitman, Schrempp, and Wells discloses <i>identifying a channel corresponding to the data samples</i> [Claims 1.3, 6.3, 10.3].....	45
D.	Whether the combination of Pitman, Schrempp, and Wells discloses or renders obvious <i>using one or more computers to utilize a sequential order of at least two different obtained content identifiers in the play</i>	

	<i>stream to identify a channel corresponding to the data samples where the content identifiers identify different media content items</i> [Claims 1.3-4, 6.3-4, 10.3-4].....	47
E.	Whether Pitman, Wells, or Schrempp disclose or render obvious <i>a play stream of content identification results</i> as claimed.....	55
VIII.	CONCLUSION.....	58

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 No. 8,756,622 (“the ’622 Patent”) and U.S. Patent No. 10,719,849 (“the ’849 Patent”). Specifically, I have been asked to provide my opinions on the patentability of the ’622 and ’849 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 also reviewed and relied upon additional materials in preparation of this Declaration, as described below.

3. My experience, qualifications and education are detailed in my curriculum vitae (“CV”), a copy of which is attached as Exhibit 2002. My CV also lists publications on which I am a named author and identifies parties on behalf of whom I have previously provided expert testimony.

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

5. 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.

6. 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.

7. 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.

8. 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.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. I have no financial interest in Patent Owner or the '622 or '849 Patents.

III. LEGAL STANDARDS

15. 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

16. 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.”

17. 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.

18. 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.

19. 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.

20. 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 “POSA”).

21. 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.

22. 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).

23. 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.

24. 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.

25. 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.

26. 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.

27. 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.

28. 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.

29. 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

30. 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.

31. 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

32. 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.

33. 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

34. I understand that “a person of ordinary skill in the art” or POSA 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 POSA is sometimes used synonymously. Such a person would necessarily have the capability of understanding any scientific and engineering principles applicable to the pertinent art.

35. 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.

36. 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.

37. 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.

38. 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

39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

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 '622 Patent (Paper No. 2 in IPR2024-01055, hereafter “the '622 Petition”) and the Petition for *Inter Partes Review* filed against the '849 Patent (Paper No. 2 in IPR2024-01058, “the '849 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.

’622 Pet., 2 / ’849 Pet., 3.

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 26, 2005. I understand this is the priority date 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 understand that claim terms are given their plain and ordinary meaning as understood by a POSA. I have applied the plain and ordinary meaning 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 VALIDITY

A. Whether the Wang References disclose or render obvious the claimed *play stream*.

1. The Wang References do not disclose a *play stream*

53. Each of independent claims 1.1, 6.1, and 10.1¹ recites *a play stream of content identification results corresponding to a sequence of data samples, the data samples comprising representations of [audio/video] data captured at a media monitoring device*. Based on its plain language, this limitation requires that the *play stream* include content identification results from a corresponding *sequence of data samples (plural)* and that those data samples comprise representations of audio/video data captured at *a media monitoring device (singular)*. In other words, the *play stream* must be associated with a sequence of data samples for a single media monitoring device. That *play stream* then serves as the basis for *identify[ing]*

¹ I rely on the same identifiers of claim elements as set forth in Petitioner's claim appendices.

a channel corresponding to the data samples recited later in claim elements 1.3, 6.3, and 10.3.

54. The Petitions contend that Wang777 teaches the claimed *play stream*. See '622 Pet., 45-46 (stating “Wang777 teaches ... capturing a sequence of data samples (e.g., audio samples) comprising representations of audio data (e.g., broadcast radio) captured at a media monitoring device (e.g., audio sampling device 102, 506, blue)”); *Id.*, 47 (stating “Wang777 also teaches generating content identification results (e.g., title, artist(s), and record label) corresponding to the sequence of audio data samples (e.g., the audio samples captured by audio sampling device 102, 506 over time).”); '849 Pet., 45 (“Wang777-Wang060 combination teaches capturing a sequence of data samples (e.g., video samples) comprising representations of video data (e.g., television broadcast) captured at a media monitoring device (e.g., sampling device 102, 506, blue).”); *Id.*, 47 (“Wang777-Wang060 also teaches generating content identification results (e.g., title, artist(s), and record label) corresponding to the sequence of video data samples (e.g., video samples captured by sampling device 102, 506 over time).”). Petitioner is mistaken for the reasons I explain below.

55. Although Wang777 does discuss audio data samples, Wang777 exclusively teaches that its system identifies audio/video content based on an individual (i.e., single) audio data sample from a sampling device (e.g., cellphone).

See Ex. 1005 ¶¶ [0009], [0019]-[0020], [0022], [0047], [0057] (each paragraph referring to “an audio sample” and “the audio sample” (singular)); *Id.*, claims 1-24 (all referring to an “audio sample” (singular) recorded by the user). Wang777 only teaches that it analyzes each audio data sample one-at-a-time to identify each piece of content (e.g., a song). Wang777 does not teach that multiple audio data samples are used together to identify a piece of content or a channel.

56. While Wang777 does teach that the length of the individual sample being relied on to identify a piece of content may vary, there is still only an individual sample being analyzed in Wang777. For example, Petitioner states that “the USRTO is a time offset between a beginning of the identified content file from the database and a beginning of the audio sample being analyzed, e.g., a user may record a ten second sample of a song beginning 67 seconds into a song.” ’622 Pet., 48 (citing Ex. 1005, ¶ [0047]). However, this merely refers to how the length of the sample may vary—there is still only a single sample being analyzed one at a time. See Ex. 1005 ¶¶ [0022], [0024], [0047].

57. A POSA would understand that Wang777 is fundamentally premised on analyzing an individual audio sample (e.g., audio from a song that the listener hears on the radio) from each individual user’s sampling device (e.g., a phone 506). In other words, only a single sample is analyzed and identified, one at a time, for each media monitoring device—Wang777 does not rely on a play stream based on

multiple sequential samples from a media monitoring device to do any content or channel identification.

58. Although a listener could submit multiple, separate samples in Wang777, those samples would be analyzed individually to identify their content or broadcast source. It would not be logical for Wang777 to rely on a play stream corresponding to a sequence of audio/video samples for content or channel identification, because the purpose of Wang777 is for a user to record an audio sample spontaneously to identify content (e.g., a song). Ex. 1005 ¶ [0019] (“The user may hear an audio program being broadcast from some broadcast means, such as radio or television, and can record a sample of the audio using the audio sampling device.”). Indeed, nothing in Wang777 teaches or suggests generating a play stream based on multiple sequential audio/video samples and then using those sequential samples to identify content or a channel, which would not make sense given the purpose of Wang777’s system.

59. The claimed invention is different than Wang777, because the invention generates a playlist based on a sequence of multiple audio/video data samples from a monitoring device. The invention is designed to observe the media consumption of an audience member over a longer period of time and identify the channel playing the content. ’622 Patent, 1:22-26 / ’849 Patent, 1:29-33 (“[I]t 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.”). Multiple audio/video data samples are needed because the monitoring period is considerably longer than that needed for monitoring a short portion of content for spontaneous, one-off recognition, as taught by Wang777.

60. The Petition contends that Wang777’s transition-tracking feature (described in paragraph [0072] of Wang777) relies on a play stream corresponding to a sequence of audio data samples (plural) associated with a monitoring device. ’622 Pet., 55 / ’849 Pet., 56 (citing Ex. 1005 ¶ [0072]). But this feature, just like Wang777’s content and channel identification feature, relies exclusively on a **single** sample. Wang777 explains, in reference to this transition-tracking feature, that “to further verify that a user is actually listening to a given broadcast channel,” a sample for content identification can simply be “taken **over a longer period of time**, e.g., longer than a typical audio program, such as over a transition between audio programs on the same channel.” Ex. 1005 ¶ [0072]. What Wang777 discloses is increasing the length of time over which the *single* audio data sample is captured by the listener. *Id.*

61. As Wang777 further explains:

If a match that is considered to be correct is actually the correct channel, a content alignment should be continuously maintained between song transitions. An exception can occur when the user records **an audio**

sample while changing broadcast channels. However, continuity of identity over a program transition may be an indicator that the correct broadcast channel is being tracked. Thus, sample **identity** (e.g., Equation (1)) can be tracked, and user sample **identity** changes can also be tracked. For example, sample **identities** at multiple time periods can be tracked (as shown below in Equations 3-5), and if a first sample identity does not equal a second sample identity from a second time period (as shown below in Equation 5), then the continuity or transition between songs has been tracked. This can provide further confidence that a correct match has been made (e.g., when both the user and the broadcast sources change synchronously).

Ex. 1005 ¶ [0072]. A POSA would understand this to simply teach that a longer user sample is taken (e.g., to capture a song transition) and identities at different points of time in the sample (“sample identities”) are compared to identities at different points in time the broadcast sample(s). In other words, this feature does not involve taking multiple user samples as Petitioner contends, but rather involves taking a longer user sample and identifying the sample at different points in time.

62. This is only confirmed by equations (3), (4) and (5), listed at the end of paragraph [0072]. Equations (3) and (4) track identities (n, n+1) within the sample “where n is the nth sample **in time**,” and compare them to identities (n, n+1) of the broadcast sample(s). *Id.* Consider, for example, a sample that captures a transition from a first song to a second song. Matching the sample pursuant to paragraph [0072] of Wang777 results in the identification of the two songs in the form of two sample

identities, “a first identity of the audio sample” identifying a first time period for the first song (e.g. sample identity [n]), and “a second identity of the audio sample” identifying a second time period for the second song (e.g. sample identity [n+1]). Even though the single user sample has two user sample identity values, it is still a **single** user sample, not multiple user samples.

63. Petitioner appears to rely on the following statement from paragraph [0072] to support its theory that Wang777 relies on multiple sequential samples: “**user samples** can be taken over a longer period of time, e.g., longer than a typical audio program, such as over a transition between audio programs on the same channel.” However, this does not refer to multiple samples being used to verify the channel identification.

64. The statement about “user samples” refers to the fact that the system receives audio samples from various listeners that are separately and discretely analyzed (e.g., to identify the specific song that the user spontaneously recorded a portion of), and that each individually analyzed sample can be longer to as to capture song or program transitions. The statement is not teaching that multiple samples are being analyzed together to identify a particular song or channel. As can be seen by the remainder of paragraph [0072], only a single sample is being used, with multiple identifications being made at different sample time points. See Ex. 1005 ¶ [0072]

(“An exception can occur when the user records an audio sample while changing broadcast channels.”).

65. It does not make sense for Wang777 to teach that discrete audio samples relied on for the transition-tracking feature are “taken **over a longer period of time**, e.g., longer than a typical audio program,” unless the duration of each sample being recorded by each listener is increased (or longer) to ensure that a song transition is captured. Ex. 1005 ¶ [0072]. A POSA would understand that the listener in Wang777 takes a single audio sample for content identification (as repeatedly taught throughout Wang777), and the transition-tracking feature merely identifies song transitions occurring during that sample to provide an increased level of confidence that the channel identification was correct.

66. That this feature relies on only an individual sample is only further confirmed by claims 10 and 12 of Wang777. Claim 10 corresponds to how Wang777 identifies a channel by comparing an identity from a single user audio sample to identities from multiple broadcast audio samples. Ex. 1005, 8 (p. 14 of PDF). (reciting “comparing an identity of **an audio sample** with identities of broadcast audio samples,” “comparing a time offset of **the audio sample**,” and “identifying a broadcast channel from which **the audio sample** was recorded). Claim 12, which depends on claim 10, corresponds to paragraph [0072] and recites: “identifying an identity change **within the audio sample**: and comparing a first identify [sic] **of the**

audio sample with identities of the broadcast audio samples, and comparing a second identity **of the audio sample** with identities of the broadcast audio samples.” Claims 10 and 12 confirm that while Wang777 discloses using multiple “identities” from an audio sample, there is still only a single audio sample is being used.²

67. Wang777’s channel-change is multiple identifications (n, n+1, n+2, etc.) within a **single individual sample** for a media monitoring device. While the multiple identifications of the single sample are arguably sequential, they are not *sequential samples* because they are all taken on the same sample.

68. Assuming that it is the intention of the Petition to rely on Wang060’s list of “candidates” (discussed in Ex. 1006 ¶[0081]) to argue that it would have been obvious to modify Wang777 to rely on sequential samples as claimed, the list of candidates does not address this deficiency in Wang777.

69. Like Wang777, Wang060 is also focused on identifying content of a single audio sample. Ex. 1006, Abstract (“A method for recognizing **an audio sample** locates an audio file that most closely matches the audio sample from a database indexing a large set of original recordings.”), ¶¶ [0079]-[0080] (“Using the master index list generated as described above, sound recognition is performed on an exogenous **sound sample**, typically supplied by a user interested in identifying

² Claim 9 is also directed toward paragraph [0072] and likewise confirms that there is only a single listener audio sample being used. Ex. 1005, 8 (p. 14 of PDF).

the sample. For example, the user hears a new song on the radio and would like to know the artist and title of the song.”).

70. This single sample focus is clearly shown in Wang060 Figure 1 (reproduced below), through which the Petition acknowledges that individual single samples are analyzed. ’622 Pet., 25-26 / ’849 Pet., 24 (explaining that “Wang060 is directed to a ‘method for recognizing *an audio sample*’ locates an audio file that most closely matches *the audio sample*”) (citing Ex. 1006, Abstract).

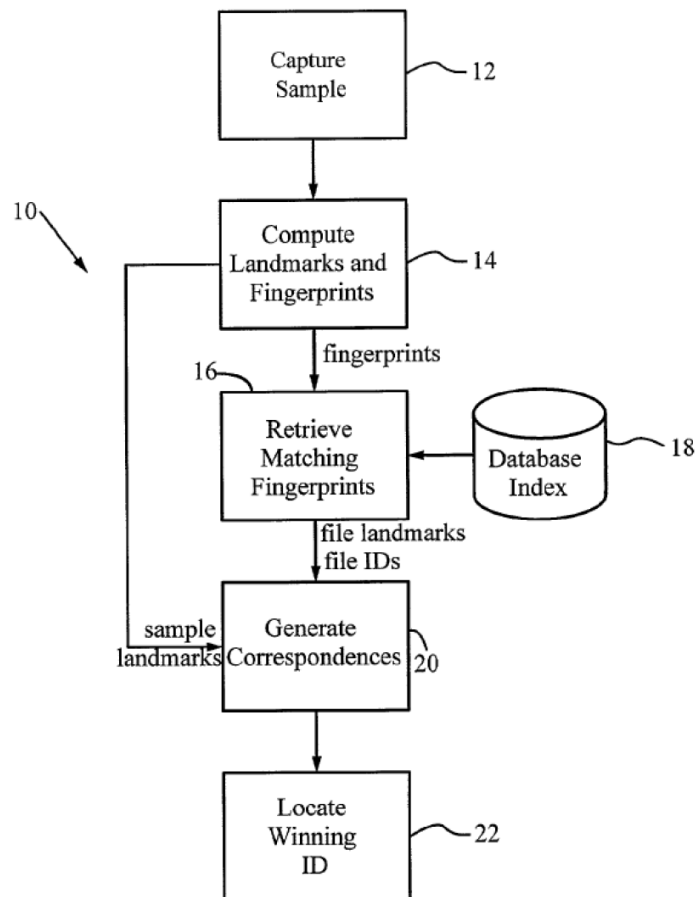
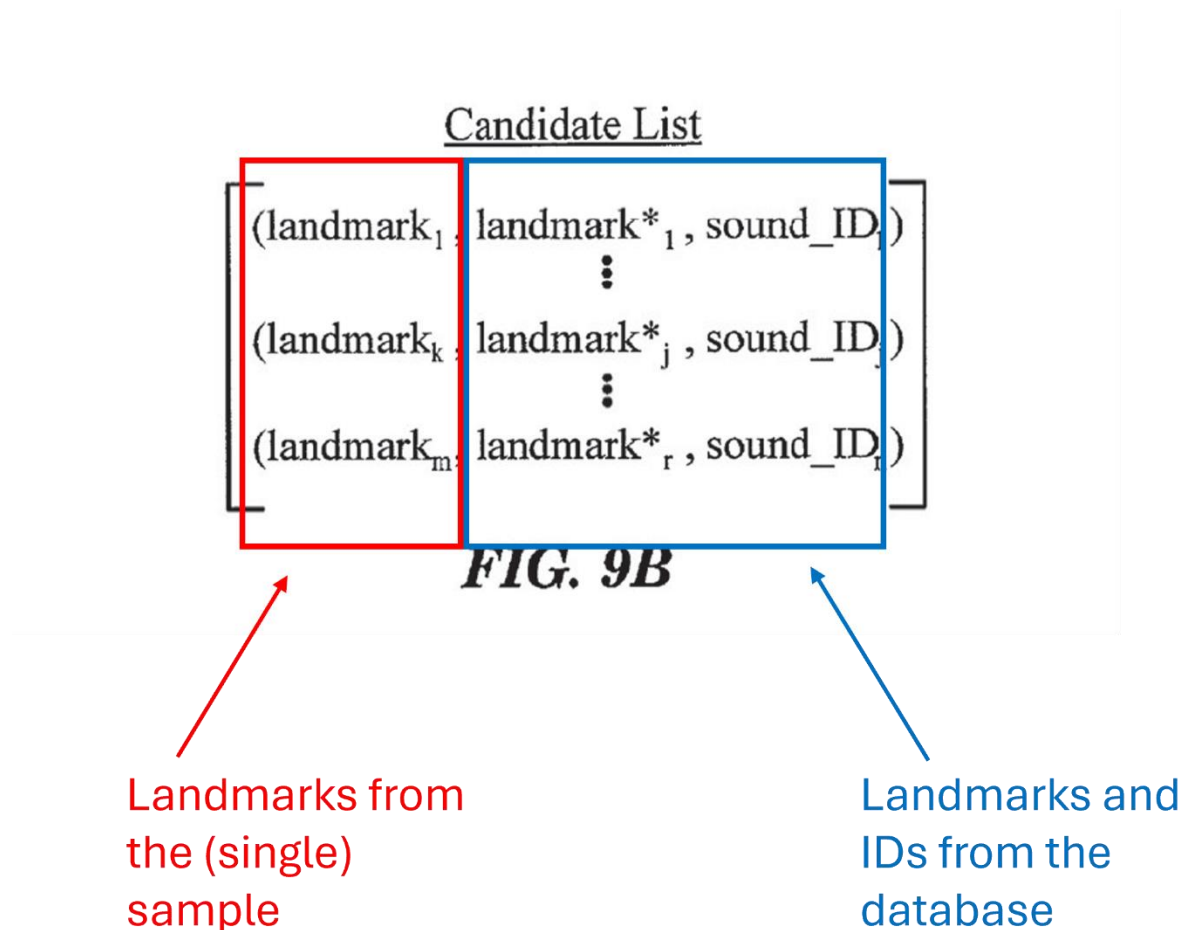


FIG. 1

71. The Petition contends that “Wang060 teaches generating a candidate list of the identified sound samples that includes a sound_ID for each match and landmark timepoints,” and that this satisfies the *play stream* limitation, although the Petition does not explain how. ’622 Pet., 50-51 / ’849 Pet., 50-51 (citing Ex. 1006, ¶ [0081], FIG. 9B). But this characterization of Wang060’s teachings is incorrect. Wang060 teaches generating a set of triplets associated with **a single sample**. These triplets are then used for matching a **single** audio data sample against a database of multiple, potential matches. The candidate list is itself based on a single sound sample: “Given the pairs for **the sound sample**, the database index is searched to locate potentially matching files.” Ex. 1006 ¶ [0081]. Wang060 is not teaching that multiple sequential audio samples of a monitoring device are being analyzed as part of a play stream. Rather, a **single** audio data sample is being analyzed against a database of multiple, potential matches.

72. As Wang060 explains “[a]s with the construction of index sets for database files, landmarks and fingerprints are calculated for **the sample** using the same algorithm that was used for processing the sound recording database.” *Id.* ¶[0080]. As shown below, part of the candidate list is associated with landmarks from a single sample, and the others are associated with landmarks and sound identifiers from a database:



Ex. 1005, Fig. 9B (annotated). Thus, the candidate list is multiple match “candidates” associated with a single audio sample, and Wang060 also analyzes individual user samples, one at a time, to identify content just like Wang777 does.

2. The proposed combination of Wang777 and Wang060 is not obvious to a POSA.

73. It would not have been obvious to a POSA to “store the identified samples in a known format such as a play stream for later comparison to the identified broadcast samples,” as the Petitions contend. ’622 Pet., 50 / ’849 Pet., 50.

74. In Wang777, as the Petitions concede, individual one-off samples are captured by each media monitoring device (e.g., on a mobile phone). '622 Pet., 23 / '849 Pet., 21 (“sampling device 102 (e.g., a mobile phone, blue) used to ‘record a[n] [audio] sample broadcast or transmitted by a broadcaster 104’”) (citing Ex. 1005 ¶ [0020]). The individual sample is then sent to the Wang777 system for analysis. Ex. 1005 ¶ [0020] (“The user can then cause the audio sampling device 102 to send the audio sample to a sample analyzer 106”). In each instance, only a single sample is identified, one at a time, for each media monitoring device. It does not create *a play stream of content identification results corresponding to a sequence of data samples* (plural), where *the samples compris[e] representations of audio/video data captured at a media monitoring device* (singular).

75. There is no reason for Wang777 to analyze sequential audio data samples captured by a user. The design-purpose of Wang777 is to correctly identify an audio sample spontaneously sent by a user and determine the corresponding channel. There is no motivation or rational for combining a play stream including sequential audio data samples into Wang777, and the Petition provides none. The Petition does not explain why it would have been obvious to modify Wang777 such that it captures multiple audio data samples recorded by a user and then use those multiple samples to perform its content identification. It is not.

76. Nor would it be obvious to a POSA “to store the content identification results of Wang777 . . . in chronological order in a play stream such that each user sample can then be compared to identified broadcast samples having similar timestamps (i.e., in a play stream) in view of Wang060” as the Petition goes on to contend. ’622 Pet., 41 / ’849 Pet., 41. Nothing in Wang777 would make such identifications useful, as it is not tailored toward either the purpose of Wang060 or Wang777. A POSA would see no reason to rely on multiple sequential samples (plural) of a media monitoring device, rather than relying on the single sample actually taught by Wang777. Indeed, Wang777 expressly teaches simply increasing the duration of the single sample as needed, because it is illogical to require the listener to submit separate samples when the user is simply trying to identify content on-the-fly. Ex. 1005 ¶ [0072]; *see also id.*, Cl. 12 (“identifying an identity change within **the audio sample**.”).

B. Whether the Wang References disclose or render obvious *identifying a channel based on a sequential order of at least two different obtained content identifiers*

77. Each of claims 1.3, 6.3, and 10.3 recite that *a channel corresponding to the data samples be identified based on a sequential order of at least two different obtained content identifiers* (e.g., song51 and song81) *in the play stream*. Ex. 1001, claims 1, 6, and 10. Neither Wang Reference discloses this.

78. The Petition’s sole reliance within the Wang References for teaching these limitations is the transition-tracking feature described in paragraph [0072] of Wang777. But this feature does not identify a channel. The transition-tracking feature relies on detecting identity changes that occur during a recorded listener sample to determine whether matching alignment is maintained with a channel that has already been identified. Ex. 1005 ¶ [0072]. As Wang777 explains, “to **further verify** that the user is actually listening **to a given** broadcast channel, and that it is not just a coincidence (such as a user taking a recording from a CD player), user samples can be taken over a longer period of time, e.g., longer than a typical audio program, such as over a transition between audio programs on the same channel.” *Id.* This feature is premised on the channel **already being identified**, and then simply determining whether a transition in the user sample aligns with the broadcast sample.

79. To fully understand this feature, it is important to understand how Wang777 **actually** identifies channels. Wang777 teaches:

To identify a broadcast source, the user audio sample and broadcast audio samples are compared to first identify matching sample identities, as shown at block 414, and then to identify matching “relative times” as shown at block 416. If no matches are found, another broadcast channel is selected for comparison, as shown at blocks 418 and 420. If a match is found, the corresponding broadcast information is reported back to the user, as shown at block 422.

Ex. 1005 ¶[0051]. Stated another way, Wang identifies channels by: (1) comparing an identity within a listener's audio sample to identities of a broadcast sample; and (2) identify matching "relative times," which simply refers to identifying a time alignment between the sample identities. Wang777 discloses two equations for identifying a channel:

$$(\text{User sample identity}) = (\text{Broadcast sample identity}) \text{ Equation (1)}$$

$$\text{USRTO} + (\text{ref. time} - \text{UST}) = \text{BSRTO} + (\text{ref. time} - \text{BST}) + \text{delay} \text{ Equation (2)}$$

Ex. 1005 ¶[0052]

80. The first equation corresponds to matching identities as discussed, and the second equation corresponds to matching "relative times." It is this combination of both matching identities and matching times that allow Wang777 to identify channels. *Id.* ¶¶ [0052]-[0056] ("A decision is thus made as to whether the user audio sample originated from a given broadcast source by noting whether Equations (1) and (2) hold."). Wang777 explains that the "the probability of misidentification is low," because "the probability that a user sample is taken from the wrong broadcast channel or non-monitored audio source (such as a CD player) and happens to satisfy Equations (1) and (2) is fairly small." *Id.* ¶ [0055].

81. The Petition does not appear to rely on Wang777's actual channel identification to meet the claim limitation. The Petition instead relies on Wang777's

transition-tracking feature as meeting the claim limitation. The entirety of Wang777's description of this feature is provided below:

In another embodiment, to further verify that the user is actually listening to a given broadcast channel, and that it is not just a coincidence (such as a user taking a recording from a CD player), user samples can be taken over a longer period of time, e.g., longer than a typical audio program, such as over a transition between audio programs on the same channel. If a match that is considered to be correct is actually the correct channel, a content alignment should be continuously maintained between song transitions. An exception can occur when the user records an audio sample while changing broadcast channels. However, continuity of identity over a program transition may be an indicator that the correct broadcast channel is being tracked. Thus, sample identity (e.g., Equation (1)) can be tracked, and user sample identity changes can also be tracked. For example, sample identities at multiple time periods can be tracked (as shown below in Equations 3-5), and if a first sample identity does not equal a second sample identity from a second time period (as shown below in Equation 5), then the continuity or transition between songs has been tracked. This can provide further confidence that a correct match has been made (e.g., when both the user and the broadcast sources change synchronously).

Ex. 1005 ¶[0072].

82. As shown, Wang777 explains that “sample identity (e.g., Equation (1)) can be tracked, and user sample identity changes can also be tracked. For example,

sample identities at multiple time periods can be tracked (as shown in Equations 3-5), and if a first sample identity does not equal a second sample identity from a second time period (as shown below in Equation 5), then the continuity or transition between songs has been tracked.” *Id.* Those equations are reproduced below:

$$\text{User sample identity}[n] = \text{Broadcast sample identity}[n] \quad \text{Equation (3)}$$

$$\text{User sample identity}[n+1] = \text{Broadcast sample identity}[n+1] \quad \text{Equation (4)}$$

$$\text{User sample identity}[n] \neq \text{User sample identity}[n+1] \quad \text{Equation (5)}$$

where $[n]$ is the n th sample in time:

Ex. 1005 ¶ [0072].

83. As paragraph [0072] expressly states, these equations are “to further verify that the user is actually listening to a given broadcast channel.” A POSA would understand that the “given broadcast channel” is the channel identified pursuant to earlier equations (1) and (2).

84. Once that channel has been identified, equations (3) – (5) are used to continue tracking the matching-alignment between different identities in the listener’s audio sample and different identities in the identified broadcast sample. If that matching alignment is maintained (equations (3) and (4)), and the user sample identities n and $n+1$ are not the same (equation (5), meaning not the same song) then the system can assume that there was a song transition during the listener’s sample. If the matching alignment was not maintained, then the system can assume that the

listener changed channels. In either case, the equations are not “identifying a channel” as claimed, but are merely confirming that the song transition is not the result of a channel change. Indeed, the transition-tracking feature does not rely on the matching of relative times pursuant to equation (2), because channel identification is not occurring in equations (3)-(5)—the channel identification already occurred pursuant to equations (1)-(2).

85. While this transition-tracking feature does in fact compare different identities (e.g., song51 and song81) in a longer user sample (e.g., a sample long enough to obtain different identities) against different identities (e.g., song51 and song81 over those same periods of time) of a broadcast sample(s), it does not do so to *identify a channel*. Rather, the channel was previously identified pursuant to equations (1) and (2) and this feature is simply tracking song transitions to determine whether the listener changed channels. Ex. 1005 ¶¶ [0051]-[0061] (discussing how Wang777 identifies channels).

86. Wang777 does not teach that this feature is capable of identifying the correct channel even if a channel change is detected. Wang777 simply teaches that if it is determined that a user has changed channels, the system can go back and attempt to identify a new channel match using equations (1) and (2) described earlier in Wang777. Ex. 1005 ¶ [0073].

87. There is nothing in Wang777 that teaches or suggests that the channel-change feature is used to identify a channel, such as by initially determining the matching channel, correcting an initial identification, or selecting a channel from a candidate list of channels. In fact, Wang777 itself never refers to this feature as performing “identification,” in contrast to how it describes “identifying” with respect to broadcast channel recognition. *Compare e.g.*, Ex. 1005 ¶¶ [0051]–[0052] (explaining that equations (1) and (2) “identify a broadcast source”) *with id.*, ¶ [0072]; *Compare also id.*, cl. 10 (referring to “identifying a channel” based on methodology of equations (1) and (2)) *with id.*, cl. 12 (referring to transition-tracking feature as simply “identifying an identity change within the audio sample,” not identifying a channel).

88. This reading of Wang777 is only further confirmed by claims 1, 9, 10 and 12 of Wang777. Claim 1 recites “**identifying a broadcast station** from which the audio sample was broadcast” consistent with equation (1). Ex. 1005, 7-8 (pages 13-14 of PDF). And claim 9, from which claim 1 depends, corresponds to paragraph [0072] and recites “**identifying a content alignment** between the transition within the audio sample and at least one transition within a sample taken from the broadcast stations,” not “identifying a broadcast station” as recited in independent claim 1. *Id.*

89. Similarly, claim 10 corresponds to how equations (1) and (2) in Wang777 identify the channel based on the single listener audio sample. Ex. 1005

¶¶ [0051]-[0052] (describing channel identification pursuant to equations (1) and (2)); *Id.*, 8 (p. 14 of PDF) (claim 10 reciting “based on substantially matching identities and substantially matching time offsets, **identifying a broadcast channel** from which the audio sample was recorded”). Claim 12, from which claim 10 depends, corresponds to the equations in paragraph [0072] and merely recites detecting an “identity change” within the individual sample—not “identifying a broadcast channel” as recited in claim 10. Ex. 1005, 8 (p. 14 of PDF).

90. The ’622 /’848 Patent uses the difference in content identifiers, and the sequence of what is identified, as a specific mechanism for identifying the channel:

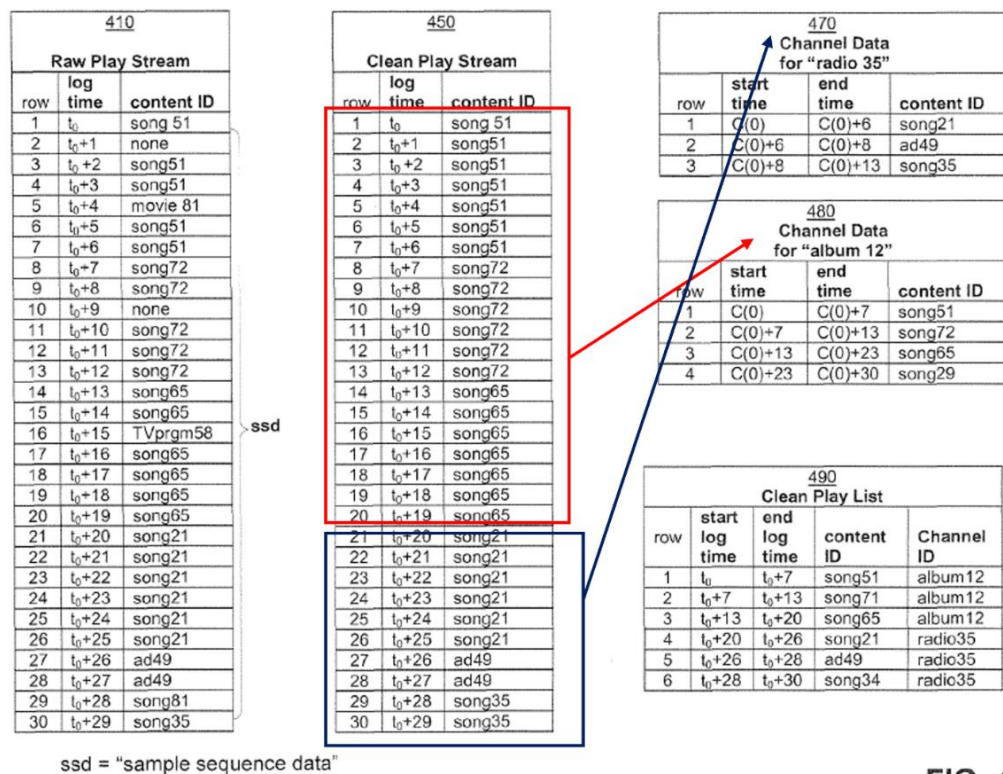


FIG. 4

'622 Patent, 8:9-22, Fig. 4 (annotated); '849 Patent, 8:18-34, Fig. 4.

91. The Specification explains that “[i]nformation about the sequence of content in a play stream can be used to generate a play list and identify a channel.” '622 Patent, 9:52-53 / '849 Patent, 9:66-10:2. As seen in Fig. 4 above, “known channel data includes channel data 470, associated with channel ID ‘radio 35,’ and channel data 480, associated with channel ID ‘album 12,” and the comparison of the sequential list of identified pieces of content from the play stream is used to identify the various, multiple channels the listener was listening to during the collection of sequential samples that resulted in the play stream. '622 Patent, 9:61-63 / '849 Patent, 10:7-10. This is an example of using the *sequential order of two different obtained content identifiers to identify a channel*, which is something Wang777 simply does not do.

92. When the transition-tracking feature relies on two different content identifiers of the single listener sample, it does so only in the context of determining whether a user has changed channels from the already identified channel—it does not itself *identify* the channel. The channel identification of Wang777 is instead performed pursuant to equations (1)-(2), not equations (3)-(5). Indeed, as the very next paragraph of Wang777 makes clear, the transition-tracking methodology makes no further use of the two differing content identifiers. Ex. 1005 ¶ [0073] (“If it is determined that a user has changed channels, the monitoring station 514 **can then**

search for an identify [sic] match for the new identity of the audio sample to verify the new broadcast source to which the user is listening.”).

C. Whether the combination of Pitman, Schrempp, and Wells discloses *identifying a channel corresponding to the data samples* [Claims 1.3, 6.3, 10.3]

93. The Petition contends that Pitman identifies a song, and identifies a channel because its system “identifies one or more products relating to the identified content” (e.g., the name of an album containing the song) using a database of known products corresponding to the song. ’622 Pet., 89-91 / ’848 Pet., 90-91 (citing Ex. 1013 ¶¶ [0030]-[0031], [0035]). But this is mistaken for the reasons discussed below.

94. According to the plain claim language, the identified channel must *correspond[] to the data samples*. Ex. 1001, claims 1.3, 6.3, and 10.3. Pitman’s correspondence of known products associated with identified content (e.g., a song) is fundamentally different.

95. Pitman merely identifies content (e.g., a song) and then uses a database to identify known products associated with that content, “such as a cassette single of the song, an album containing the song, a music video for the song.” Ex. 1013 ¶ [0035]. Pitman has a database of known products linked to identified content and simply provides the listener with all products ever known to be associated with the identified content.

96. This provides the listener with as many purchase options as possible for products containing the content that the listener is interested in buying. It does not identify the channel that the listener originally consumed the content on. Pitman is indifferent to what channel the identified content was played on when the listener heard it. Pitman's purpose is providing the listener with all products associated with the identified content so that the listener has as many purchase options as possible as a way to listen to the song again.

97. This forward-looking approach of where the user could listen to the song, rather than a backward-looking approach of where they did listen to the song, is exemplified by the lack of any broadcast channels in the list developed by Pitman. A historic list of songs played by a broadcast channel does not further the goal of Pitman to provide a way for the user to buy a product to listen to the song again.

98. Nor do the results generated by Pitman correspond to the captured data samples that the song identification was based on. The system of Pitman cannot identify the channel that the underlying content was played on, because it simply has no mechanism for doing so. Instead, Pitman simply identifies known content (e.g., a song) from a user, without any insights as to where the user obtained the content, and then provides a list of all known places (e.g., CDs, movies, etc) that the content has ever been played on. As Petitioner acknowledges, Pitman is "not [] able to identify the album that the user is actually listening to." '622 Pet., 93 / '849 Pet., 94.

D. Whether the combination of Pitman, Schrempp, and Wells discloses or renders obvious *using one or more computers to utilize a sequential order of at least two different obtained content identifiers in the play stream to identify a channel corresponding to the data samples* where the content identifiers *identify different media content items* [Claims 1.3-4, 6.3-4, 10.3-4]

99. Pitman does not use a sequential order of at least two different obtained content identifiers in the play stream to identify the associated products. However, the Petition contends that it would have been obvious to modify Pitman based on Wells to reach this claim limitation. Petitioner is mistaken for the reasons set forth below.

100. Wells teaches waiting for consecutive candidate fingerprint matches with a song before identifying a piece of unknown content as being that song. Ex. 1015 ¶¶ [0259]-[0264]. Consequently, the portion of Wells relied on by Petitioner is not related to identifying a channel at all.

101. Even if it were obvious to a POSA to combine Pitman and Wells (as I later explain, it was not), the combination would result in Pitman's content identification (e.g., song identification) relying on consecutive fingerprint matches before matching the content to a particular song. That modification would have no relevance or effect on the way in which Pitman identifies known products associated with the song. This is because the combination would rely on consecutive fingerprint matches to identify a song (as taught by Wells), while still relying on the same database of known content to provide the user with associated products for purchase

(as taught by Pitman). Thus, the alleged “channel identification” of Pitman would still not rely on multiple different content identifiers to identify a channel even if modified based on Wells.

102. Channel identification, as claimed, is based on two different obtained content identifiers where the content identifiers identify different media content items. Wells does not teach or suggest this. Wells’ consecutive fingerprint matching feature is premised on matching to the same song consecutively. Ex. 1015 ¶ [0264] (“For example, a match is recorded if P consecutive candidate fingerprints match the same song in the database.”). It would be more accurate to say that Wells in essence teaches the opposite of what this claim limitation recites. The claim recites using different obtained content identifiers (e.g., song51 and song81) to identify a channel, whereas Wells teaches using the same content identifier (e.g., song 51) obtained consecutively to identify a song.

103. The reason for this distinction is apparent to a POSA—the ’622/’848 Patent relies on analyzing a sequence of different identified content (e.g., different songs or movies) because those differences can be used to help identify a channel by analyzing the sequence of different content within the channel. In contrast, Wells relies on consecutive fingerprint matches for the same content for purposes of identifying that content. Consequently, even if one inserts Wells’ consecutive fingerprint matching feature into Pitman (for the fundamentally different purpose of

identifying channels rather than identifying underlying content), the combination still would not teach or suggest this claim feature.

104. None of the Petitions' suggested references, Pitman, Schrempp, or Wells (either alone or in combination), teach or suggest identifying a channel based on a sequential order of at least two different obtained content identifiers in the play stream, the at least two different obtained content identifiers identifying different media content items.

105. The Petition asserts the following:

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

'622 Pet., 92-93 / '849 Pet., 93-94. Petitioner appears to contend that, despite the combination not teaching channel identification as claimed, this feature would have been obvious to a POSA based on Pitman's teaching of identifying a song based on consecutive fingerprint matches to the song. This contention is mistaken.

106. The Pitman-Wells-Schrempp combination as proposed by the Petition does not discuss or describe **identifying a channel** using **two different content identifiers**. Moreover, the combination does not have the means to determine if songs were sequential such that different songs could even be used to identify a channel. Nor does it appear that the proposed combination would provide any benefit to any of the systems it combines, since all of them are focused on simply identifying the underlying content (e.g., song).

107. The Petition further contends that a POSA would have understood Pitman’s system to “potentially frustrat[e] the user and lead[] them to not purchase a product” because Pitman provides “every possible album or movie to the user for purchase.” ’622 Pet., 93 / ’849 Pet., 94. But providing “every possible album or movie to the user for purchase” is a specific, intended feature of Pitman. Pitman is directed toward a system for “selling products” (e.g., albums, movies, etc.) by providing a user with all products associated with a particular identified content (e.g., song). Ex. 1013, Abstract, ¶¶ [0020], [0030], [0035]. Pitman is thus focused on providing the listener with as many purchase options as possible—it is not concerned with identifying the particular channel that the user was listening to.³

³ I note that Petitioner provides no logical explanation as to why the listener in Pitman would even need to be told the channel he or she is listening to. Presumably the listener knows the identity of the source of the consumed media content—what is unknown to the listener of Pitman is the identity of that content and the products that can be purchased containing that content.

108. Pitman is also focused on identifying discrete, individual pieces of content (e.g., a song), and would have no need to analyze sequential songs. Pitman does not teach and has no need to use multiple pieces of content played sequentially to identify a channel, because Pitman is focused on identifying the discrete piece of content, and then simply relying on the database of known product information to provide associated products to the user.

109. The Petition suggests that Pitman teaches a user submitting multiple sequential songs in its system. '622 Pet., 92 / '849 Pet., 93 (citing Ex. 1013 ¶¶ [0030]-[0031]). But this is an inaccurate characterization of what Pitman teaches. The portion cited by Petitioner merely states that all songs on an album or in a movie may be registered in the audio content database for presentation to a user for sale. Ex. 1013 ¶¶ [0030]-[0031]. The cited passage does not suggest that the user is submitting multiple sequential songs to be analyzed together—the user is simply submitting a single piece of content for identification and then the system provides all known products associated with that content based on the content registered in the database. The system of Pitman thus does not rely on sequential songs to identify content—the system is only designed to identify a single piece of content at a time and provide the user with all known products associated with that content.

110. The Petition goes on to contend:

Based on the teachings of Wells, a POSA would have been motivated

to modify Pitman-Schrempp to incorporate Wells' teaching of identifying a song when multiple sequential song segments correspond to the same song and **also identifying an album or movie when multiple sequential songs match the album or movie**. Doing so would have reduced the possibility of an incorrect song identification, and by extension an incorrect album (or movie in which the songs appears), as taught by Wells. Modifying Pitman-Schrempp to incorporate Wells' teaching would have required minor modifications (e.g., modifying Pitman's software).

'622 Pet., 79-80 / '849 Pet., 79-80.

111. A fundamental deficiency in this obviousness contention is that Wells does not teach or suggest using multiple sequential song matches to identify an album or movie. Wells teaches that the song is identified only if "consecutive candidate fingerprints match with the same song in the DB' to reduce the possibility of a mismatch." Ex. 1015 ¶ [0264]. In other words, Wells merely teaches identifying an individual song using consecutive fingerprints that match the song. Nothing in Wells teaches or suggests using sequential song matches to identify a movie or album.

112. Petitioner's additional arguments here are also mistaken. Relying on the identification of sequential songs to identify an album would not "reduce the possibility of an incorrect song identification." '622 Pet., 79-80 / '849 Pet., 79. Wells

already reduces that possibility by using consecutive fingerprint matches to the same song before making the identification.

113. Nor does Pitman have a need to reduce the likelihood of an “incorrect album” identification as the Petition contends. ’622 Pet., 79-80 / ’849 Pet., 79. Pitman is not concerned with identifying the particular channel that the identified song was consumed over by the user. Pitman only cares about providing a list of all products (e.g., albums, movies, etc.) that the song exists on, which provides the user all potential options for purchasing the identified content. There was thus no concern with “incorrect album” identification, because Pitman already wants to provide a listing of all albums associated with a song, allowing the user to select which of those albums he or she wants to purchase.

114. Petitioner also underestimates the major changes to Pitman that would be required to impart the channel identification that Petitioner wishes to incorporate into Pitman. Petitioner contends that this would have “required minor modifications (e.g., modifying Pitman’s software)” (’622 Pet., 79-80 / ’849 Pet., 79) but this overlooks that a fundamental premise of Pitman’s operation is that a user is attempting to identify a single, discrete piece of content (e.g., a single song). To modify Pitman to rely on channel identification would require the user to record multiple pieces of content over a much longer period of time (e.g., multiple songs played within a multi-hour movie), increasing the burden to the user to submit the

content. This is a solution to a problem that does not exist in Pitman and in fact would hinder its simplicity and ease of use for its stated purpose.

115. The Petition is also mistaken that modifying Pitman to perform channel identification as claimed “would have required minor modifications (e.g., modifying Pitman’s software).” As Pitman teaches, “the listener supplies an audio content image to a retailer’s audio identification system,” which then “analyzes the audio content image that was supplied in step 304 using an algorithm that generates audio content identifying information like the information stored in the audio content database” Ex. 1013 ¶¶ [0032]-[0033]. For Pitman to identify a channel based on sequentially identified songs from which the listener obtained the sample (as the Petition proposes), the audio content database would have to include each and every song from each and every album, movie, etc. the listener could have potentially taken their sample from (including all sources of the song the Pitman system does not have available to sell to the listener) in sequential order to perform the proposed channel identification.

116. As an example, for Pitman to be able to identify the specific radio station from which a listener consumed the submitted content using multiple content identifiers (as proposed by Petitioner), it would have to already have stored in the database the sequence of content played on all radio stations to perform channel identification. This would require the database to be exponentially larger and require

significantly more processing power to analyze multiple sequential songs (rather than only a single song) against the much larger database of content containing sequences of content from countless sources. Since Pitman is not concerned with identifying channels that the listener's submitted images were consumed on, there would have been no motivation to make such a fundamental modification.

E. Whether Pitman, Wells, or Schrempp disclose or render obvious *a play stream of content identification results* as claimed

117. The Petition contends that it would have been obvious to modify Pitman to rely on the claimed *play stream* based on Schrempp. '622 Pet., 74-76, 86-88 / '849 Pet., 73-76, 86-88. Petitioner is mistaken for the reasons explained below.

118. There simply is no motivation to combine a play stream with Pitman. The Petition contends that "POSA, however, would have known that this information could have been presented in any number of formats and would have looked to prior art, such as Schrempp, which discloses as much." '622 Pet., 74 / '849 Pet., 73-74. However, even if Pitman's information "could have been" presented in any number of formats and other formats were known, that does not constitute or provide a motivation for a POSA to modify Pitman as proposed.

119. The Petition contends that "[b]ased on the teachings of Schrempp, **it would have been obvious** to a POSA to arrange the identified songs in Pitman in a play stream of content identification results corresponding to a sequence of data

samples.” ’622 Pet., 76 / ’849 Pet., 75. However, the Petition does not provide any explanation as to why or how Schrempp’s teachings would motivate a POSA to make this proposed combination.

120. A POSA would not have been motivated to combine Pitman and Schrempp as proposed. Pitman is directed toward identifying content (e.g., a song) for a user and providing that user with products containing the identified content (e.g., an album) to the user. Ex. 1013 ¶¶ [0032], [0035]. For example, in Pitman a user may hear a song that he or she likes but does not know the identity of, in which case the user can submit a sample of the song to identify it and any products associated with the song. *Id.*

121. Schrempp is directed toward monitoring broadcast content (e.g., a radio broadcast), identifying all content played by the broadcast, and “keep[ing] a record of all matches made” and generating a play list. Ex. 1014 ¶ [0005]-[0006]. Schrempp explains “a playlist is the documentation of the performance of one or more works at a particular time over a particular medium,” and that playlists may need to be generated for artists to “verify royalty payments.” *Id.* Pitman is thus directed towards analyzing individual, one-off samples provided by a user while Schrempp is directed towards monitoring all broadcast content being played over a radio station to generate a generate a playlist.

122. Given the disparate purpose and design of Pitman, a POSA would have no reason to modify Pitman to include a play stream. To the extent Schrempp relies on a play stream at all, it is only because it is creating a playlist of all content being played on a particular broadcast station. Pitman, however, only analyzes unknown one-off samples provided by a user to identify content. Pitman thus has no reason to generate a play stream at all.

VIII. CONCLUSION

123. 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

A handwritten signature in black ink, reading "W. Leo Hoarty", is written over a horizontal line.

W. Leo Hoarty