

UNITED STATES PATENT AND TRADEMARK OFFICE

---

BEFORE THE PATENT TRIAL AND APPEAL BOARD

---

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

v.

GRACENOTE, INC.,  
Patent Owner

---

*Inter Partes* Review Case No. IPR2020-00216  
U.S. Patent No. 9,066,114

**PATENT OWNER'S PRELIMINARY RESPONSE  
TO PETITION FOR *INTER PARTES* REVIEW**

## TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	BACKGROUND.....	2
	A.    Summary of the '114 Patent .....	2
	B.    Murphy.....	7
	C.    The Claimed Invention Is Different from Murphy .....	10
III.	LEVEL OF A POSITA .....	12
IV.	CLAIM CONSTRUCTION .....	12
V.	MURPHY DOES NOT TEACH AT LEAST ONE LIMITATION IN EACH OF THE CHALLENGED INDEPENDENT CLAIMS .....	12
	A.    Claims 1, 8, and 10: “a reference time point identified by the reference fingerprint” .....	12
	B.    Claims 1, 8, and 10: Murphy does not teach “reference fingerprints previously derived” from the “content being played back” .....	17
	C.    Claims 11, 20, and 21: Murphy Does Not Teach “an action to be triggered during playback” of the Associated Fingerprinted Segment.....	22
	D.    Claims 11, 20, and 21: Murphy does not teach “accessing segments of content” During Fingerprint Generation.....	28
	1.    Murphy’s System Accesses Different Segments During Playback Than During Fingerprint Generation .....	29
	2.    Murphy Does Not Access the Content During Fingerprint Generation .....	32
VI.	THE CLAIMED INVENTION OF THE '114 PATENT DIFFERENTIATES OVER MURPHY IN VIEW OF BRUNK .....	38

VII. BOTH MURPHY AND BRUNK ARE CUMULATIVE TO PRIOR ART CONSIDERED AND OVERCOME DURING ORIGINAL EXAMINATION .....	41
A. Murphy Is Cumulative with Rhoads, Which Was Cited and Considered During Prosecution .....	41
B. Murphy Is Also Cumulative with Haitsma, Which Was Cited and Considered During Prosecution .....	43
C. Brunk Was Cited and Considered During Prosecution.....	44
VIII. CONCLUSION .....	47

## TABLE OF AUTHORITIES

<b>Cases:</b>	<b>Page:</b>
<i>Belden Inc. v. Berk-Tek LLC</i> , 805 F.3d 1064 (Fed. Cir. 2015)	25
<i>Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.</i> , 868 F.3d 1013 (Fed. Cir. 2017)	12
<i>Vivid Techs., Inc. v. Am. Sci. &amp; Eng’g, Inc.</i> , 200 F.3d 795 (Fed. Cir. 1999)	12
<i>Westinghouse Air Brake Tech. Corp. v. Siemens Ind., Inc.</i> , IPR2017- 00650, Paper 12 (PTAB July 19, 2017)	25
 <b>Statutes:</b>	
35 U.S.C. § 101	9
 <b>Regulations:</b>	
37 C.F.R. § 42.24	49
37 C.F.R. § 42.8	49

## I. INTRODUCTION

Petitioner's sole ground for unpatentability of the independent claims of the '114 Patent is obviousness based on the combination of Murphy in view of Brunk. Contrary to Petitioner's arguments, each independent claim of the '114 Patent recites limitations not taught by Murphy, either alone or in combination with Brunk. As a result of these critical deficiencies, the Petition fails to establish a reasonable likelihood the Challenged Claims are unpatentable, and Patent Owner therefore respectfully requests non-institution of this IPR.

Murphy takes advantage of markers in a broadcast audio signal, namely audio cues such as a program's jingle, inserted by program makers and broadcasters. (Ex. 1011, *Murphy*, 3:10-27). Murphy's system uses these audio cues to prompt an event, such as setting a flag to start/stop recording of a program. (*Murphy*, 13:15-26). Importantly, Murphy relies on just the type of marking the '114 Patent expressly differentiates and improves upon.

The '114 Patent teaches a particular technique of using fingerprints to trigger actions at time points in a multimedia signal without the need to rely on the broadcaster or other actors in the broadcast chain to include information in the content. (Ex. 1001, '*114 Patent*, 1:63–2:17). The '114 Patent discourages the use of Murphy's traditional audio cues because an enhancement service provider is dependent on the program maker or broadcaster inserting the audio cues into the

multimedia signal, a dependency that limits the usefulness of the system. (*'114 Patent*, 2:1-4 (“this requires the broadcaster’s cooperation to insert these triggers thereby making an enhancement service provider dependent on the broadcaster”)).

This fundamental difference between Murphy and the '114 Patent of not having to be dependent on the program maker and broadcaster to insert a marker is reflected in each of the independent claims of the '114 Patent. The '114 Patent’s claims describe using fingerprints to identify a reference time point in content to be played back and taking action at the identified reference time point. This allows the enhancement service provider to cause any number of different actions to be taken at any desired time point in the content without being dependent on where the program maker or broadcaster places audio cues.

The Petition glosses over these important features of the '114 Patent’s claims, often only devoting a single paragraph or even a single sentence to the mapping. As detailed below, each of the independent claims recites at least two limitations not taught or suggested by the combination of Murphy and Brunk.

## **II. BACKGROUND**

### **A. Summary of the '114 Patent**

The '114 Patent is directed to using fingerprints to trigger actions during playback of a multimedia signal at “very accurate and very precise locations” in the multimedia signal without the need to enlist the cooperation of the broadcaster and

other actors in the broadcast chain. (*'114 Patent*, 4:65–5:2, 3:7-12, 2:1-4, 2:14-17; Ex. 2002, Memorandum Opinion, *Gracenote, Inc. v. Free Stream Media Corp. d/b/a Samba TV*, No 1:18-cv-01608-RGA (D. Del. Dec. 11, 2019), ECF No. 25 (“When read as a whole the asserted claims of the [Challenged] Patents are directed to a system and method of implementing accurate fingerprinting in a multimedia stream to avoid the necessity of involving a broadcaster corporation.”)). The *'114 Patent* explains that to facilitate triggering actions during playback of the multimedia signal, “it is necessary to enable a reliable detection of time points” in the signal. (*'114 Patent*, 1:29-63, pinpoint at 1:29-32). For example, specific portions of a nationally broadcast program may be identified and replaced by locally relevant content (such as local ads or local weather forecast). *Id.* Alternatively, supplementary information for the content of a program (such as biographic data for an actor) can be displayed at relevant times (such as when the actor appears on screen). *Id.*

As the *'114 Patent* explains, in prior art solutions to the problem of identifying specific time instants in the signal, it was “*necessary*” for the program maker or broadcaster *to have marked or indicated specific time instants in the multimedia signal for triggering an action*.<sup>1</sup> (*'114 Patent*, 1:63-67, 2:1-29 (describing the

---

<sup>1</sup> All emphases added unless otherwise stated.

drawbacks of prior art trigger-based systems relying on markers in the signal)). This was a problem because “this requires the broadcaster’s cooperation *to insert these triggers* thereby making an enhancement service provider dependent on the broadcaster.” (*’114 Patent*, 2:1-4). The *’114 Patent* describes various known ways for such markers or triggers to be inserted into the multimedia signal to trigger actions, but the *’114 Patent* again emphasizes these known ways “need the cooperation of all actors in the broadcast chain.” (*’114 Patent*, 2:14-17).

The *’114 Patent*’s invention solves these problems by “*relating*” the desired trigger actions to the multimedia signal:

It is an object of the invention to provide a method and corresponding device of *relating one or more trigger actions with a multimedia signal* and corresponding method and device for detecting one or more trigger actions in a multimedia signal that solves the above-mentioned problems.

(*’114 Patent*, 2:36-58). To relate the trigger action with the multimedia signal, a time point is selected in the multimedia signal. (*’114 Patent*, 2:49 (“providing at least one trigger time point”)). This time point, e.g., time point  $T_n$ , in the multimedia signal is the time at which an action is to be triggered during playback of the content. *’114*



*Patent*, 3:51-58.<sup>2</sup> A fingerprint that is or has been *generated* or *derived* based on a segment of the multimedia signal corresponds to the time point  $T_n$ . (*'114 Patent*, 4:15-19, 2:55-58 (“for each given trigger time point deriving a fingerprint on the basis of a segment of the multimedia signal”), 6:2-6 (confirming the time point is determined when the fingerprint is generated, stating “the appropriate time point ( $T_n$ ;  $T_{n+1}$ ) is also determined when the fingerprints (102) have been generated”), 4:61-64 (“According to the present invention, a fingerprint (102) is generated for each time point on the basis of a given segment of the multimedia signal (101) at or near the specific time point.”)). Thus, “the segment of the multimedia signal [] is unambiguously related with the given time point ....” *Id.* “In this way, a given fingerprint (102) is a trigger marker enabling a very accurate and very precise location of a given time point of the signal (101) without using the specific time point but instead using (a segment of) the signal.” (*'114 Patent*, 4:65–5:1).

The '114 Patent also describes how the generated fingerprint and selected time point are associated with each other, stating a “representation of the time point(s) is also supplied to the fingerprinting module (202).” (*'114 Patent*, 6:67–7:4 (note the Certificate of Correction included with the '114 Patent and correcting typographical

---

<sup>2</sup> Note multiple time points can be selected in the multimedia stream, such as  $T_n$  and  $T_{n+1}$ . *'114 Patent*, 3:58-63.

errors), 6:62-64 (describing the fingerprinting module, which receives the time point information per 6:67–7:4, as computing the fingerprint, stating “[t]he fingerprinting module (202) computes a fingerprint on the basis of the received multi-media content.”), 7:5-6 (describing the fingerprinting module supplying the computed fingerprints to the database)). During playback of the multimedia signal, the received multimedia signal is fed to a fingerprint detector, which derives a fingerprint stream and determines if there are any matches with fingerprints stored in the database. (*’114 Patent*, 7:33-40, 5:65–6:6).

Notably, when a match between a segment of the fingerprint stream and a stored fingerprint is determined, the system also determines an appropriate time point for performing an action *associated with the previously-generated and stored fingerprint*. (*’114 Patent*, 6:2-6). The *’114 Patent* states “[i]f a match is found then a representation of the associated action(s) is also retrieved. **The appropriate time point** for the associated action(s) **is given by the matching fingerprint** as described above.” (*’114 Patent*, 7:37-40). Thus, the playback device knows when to perform a stored action because the stored, matching fingerprint **gives** the appropriate time point for performing the action to the playback device:

The playback device simply needs to be aware of the **relationship** between a given **fingerprint** and the **given time point used during generation**.

(*'114 Patent*, 6:18-20, 7:41-45, 5:65–6:2 (“When **a match** between a segment of the fingerprint stream (104) and a given fingerprint (102) in the database **is found** the representation(s) of the **associated action(s)** (105) of that particular fingerprint (102) **is retrieved and executed at the appropriate time point** ( $T_n$ ;  $T_{n+1}$ ).”)).

“For each given trigger time point,” the fingerprint is derived “on the basis of a segment of the multimedia signal,” such that the stored fingerprint identifies the time point. (*'114 Patent*, 2:55-58, 7:37-40). The result is that “the segment of the multimedia signal is unambiguously related with the given trigger time point.” (*'114 Patent*, 2:56-58). And because “[a] fingerprint is derived each time that a trigger action is needed, i.e. for each time point,” the derived fingerprint corresponds to the time point. (*'114 Patent*, 6:67–7:2, 3:9-11 (“A fingerprint thereby serves as a trigger of a particular action, event, etc. associated with a particular point in time of the multimedia signal.”)).

Selecting a time point in the multimedia signal and deriving the fingerprint based on a segment of the multimedia signal at that time point avoids the need to rely on broadcasters, program makers, or other actors in the broadcast chain to have previously marked, indicated, or inserted triggers.

## **B. Murphy**

Murphy describes an “automatic recognition system” in a transmitter for recognizing a program being broadcast to “generate useful ancillary data therefrom”

(for example, to automatically add associated data to a signal being broadcast) or in a receiver to automatically record a program. (*Murphy*, Title, Abst.). *Murphy* requires the program maker or broadcaster to include distinctive and easily recognizable audible jingles or cues placed immediately before or after particular programs, or at the beginnings or ends of particular programs, to prompt an event. (*Murphy*, 3:10-20). Specifically, *Murphy* relies on the “standard practice of programme makers and broadcasters *to place an audio cue* such as a piece of music or jingle immediately before a particular program” and detection of these previously inserted “audio cues” to generate an event, such as generating an event flag to cause a recorder to start recording or to cause text to be displayed on a text display. (*Murphy*, 3:10-27, 5:9-11 (“The audio recognition system is designed to detect a match between an input audio signal and one or more known audio cues stored for recognition purposes.”), 13:21-26). Because these audio cues are placed in the multimedia signal by the program maker or broadcaster, the enhancement service provider is dependent on the broadcaster, including being limited to performing actions only where the audio cues are provided in the broadcast audio signal (i.e., near the beginnings and ends of particular programs). (Ex. 2001, *Moulin Decl.*, ¶¶ 34-35).

*Murphy* relies on the same type of prior art approach the '114 Patent distinguished from its invention—audio cues indicated or inserted by the

broadcaster. (*'114 Patent*, 1:63–2:29, pinpoint at 2:1-4 (disparaging methods that “mak[e] an enhancement service provider dependent on the broadcaster”)). Murphy states its system “does not require the broadcaster to adapt the signal to include additional data which can be detected and used to control the recording system 82 nor does it impose any particular information format onto broadcasters.” (*Murphy*, 14:20-25). But Murphy is still dependent on the program maker or broadcaster to utilize the audio cues, because the program maker or broadcaster must still *place* the audio cue in the multimedia signal. Without placement of the audio cues by the program maker or broadcaster, Murphy’s receiver has no reliable mechanism by which to determine the start or stop of a program to begin/end recording. Murphy’s system simply takes advantage of the audio cues the broadcaster/maker inserts into the program.<sup>3</sup> (*Moulin Decl.*, ¶¶ 34-35). Murphy thus describes the type of system distinguished, and improved upon, by the ’114 Patent.

---

<sup>3</sup> In determining the ’114 Patent’s claimed inventions were not “abstract ideas” under *Alice*’s test for subject matter eligibility under 35 U.S.C. § 101, the district court held the inventions were directed to an improvement in technology that avoided the need for reliance on a broadcaster. (Ex. 2003, Report and Recommendation, *Gracenote, Inc. v. Free Stream Media Corp., d/b/a Samba TV* at 8; Memorandum Opinion, *Gracenote, Inc. v. Free Stream Media Corp., d/b/a Samba TV* at 5).

### C. The Claimed Invention Is Different from Murphy

At least two limitations in independent claims 1, 8, and 10 highlight the differences between the system of the '114 Patent and Murphy (or Murphy combined with Brunk). First, these claims recite the “reference time point identified by the reference fingerprint,” and second, that the reference fingerprint is “previously derived from the content.” In Murphy, which Petitioner relies on for teaching these two limitations, the mapped “reference fingerprint,” i.e., the signature for the audio cue (*see* Paper 1, *Petition*, 28), does not identify an appropriate time point for performing the action.<sup>4</sup> Instead, the action is performed at an arbitrary time point having no relationship with the stored signature for the audio cue. Murphy’s system performs a match, a flag is set, and the system begins/stops recording—all without

---

<sup>4</sup> Murphy’s lack of any time point associated with the fingerprint can be seen in the fact that Murphy takes action *every time* the audio cue signature is matched. In particular, Murphy specifies that actions are taken on the “repeated occurrence of the audio cue.” (*Murphy*, 11:2-3). This means that, in the example given in the *Petition* of the 60 Minutes’ ticking clock (Pet 18), Murphy’s system would start recording at the beginning of the program, then *stop* recording upon return from the first commercial break, then start recording again upon return from the second commercial break, and so forth.

any specified time point identified by the stored audio cue signature. (*Murphy*, 13:17-26). The “very accurate and very precise location of a given time point of the signal” (*’114 Patent*, 4:65–5:1) achieved by the ’114 Patent is not taught or suggested in *Murphy* because *Murphy* does not identify any time point whatsoever in the signature of the audio cue.

Another difference is claims 1, 8, and 10 recite that the fingerprint is “previously derived from the content,” i.e., the multimedia signal. In contrast, *Murphy*’s audio cue signatures (mapped as the claimed “reference fingerprint”) are derived or obtained from samples other than the multimedia signal.

For independent claims 11, 20, and 21, *Murphy* does not teach the steps of “generating a reference fingerprint of the segment,” where the segment is a segment of content, to trigger an action during subsequent playback of the segment; and “assigning the reference fingerprint of the segment to the corresponding action.” Because the segments of content in *Murphy* do not exist prior to broadcast, the actions to be triggered cannot be assigned to a fingerprint in advance.

Given these crucial differences between *Murphy* and the ’114 Patent claims, the Petitioner fails to establish a reasonable likelihood that at least one claim of the ’114 Patent is unpatentable.

### **III. LEVEL OF A POSITA**

Patent Owner agrees with Petitioner that a person of ordinary skill in the art (“POSITA”) for the ’114 Patent at the time of the invention would have a bachelor’s degree in computer science, electrical engineering, or a related discipline and two years of experience in a relevant technical field such as multimedia signal processing. Additional experience could substitute for formal education, or additional formal education could substitute for experience in the relevant technical field.

### **IV. CLAIM CONSTRUCTION**

Petitioner’s claim constructions provided at the Petition, pages 11-14, do not impact Patent Owner’s arguments provided in this paper. (*Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (quoting *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999)) (noting the PTAB “need only construe terms ‘that are in controversy, and only to the extent necessary to resolve the controversy’”)).

### **V. MURPHY DOES NOT TEACH AT LEAST ONE LIMITATION IN EACH OF THE CHALLENGED INDEPENDENT CLAIMS**

#### **A. Claims 1, 8, and 10: “a reference time point identified by the reference fingerprint”**

Murphy fails to teach “a reference time point identified by the reference fingerprint” at which a reference action is to be performed, as recited in claims 1, 8,



and 10. The stored audio cue signature in Murphy *does not identify a reference time point* at which the corresponding reference action is performed.<sup>5</sup> (*'114 Patent*, Claim 1 (“each reference fingerprint identifying at least one corresponding reference action to be performed at a reference time point identified by the reference fingerprint”); Pet 28 (mapping the stored audio cue signatures to the claimed “reference fingerprint”)). Rather, when the audio cue signature is recognized in Murphy, an event happens. (*Murphy*, 13:15-26; Pet 29). But Murphy does not teach or suggest a specific time point (identified by the audio cue signature) at which the event flag is set or the recording (or other event) is activated.<sup>6</sup> (*Murphy*, 13:15–14:3).

---

<sup>5</sup> Petitioner concedes Brunk does not associate either fingerprints or actions to time points. *See* IPR2020-00217, Paper 1 (*Petition*) at 29 (“Brunk likewise discloses using a plurality of fingerprints to trigger corresponding actions, albeit not expressly at particular trigger time points.”). Petitioner therefore relies upon Murphy alone for allegedly teaching this limitation.

<sup>6</sup> It is unclear to Patent Owner if Petitioner maps the setting of the event flag or, alternatively, the event corresponding to the flag being set (e.g., start/stop recording) to the claimed “reference action.” (*See* Pet 29-30). Regardless, the audio cue signature does not identify a time point for performing either “action.”

Murphy teaches storing an audio cue signature (i.e., the mapped “reference fingerprint”) to be matched to a played audio cue. (*Murphy*, 10:13-19; Pet 29). Murphy further teaches the same audio cue can be played multiple times during the program, e.g., at the beginning or end of the program. Murphy teaches “[o]nce the start flag has been generated, an end flag could be generated on the *repeat occurrence* of the audio cue.” (*Murphy*, 11:1-3). Murphy also teaches recording an audio cue from a previous episode of a program to allow subsequent episodes to be recorded automatically. (*Murphy*, 14:11-17). Thus, actions are taken at different times based on the same audio cue signature (i.e., the mapped reference fingerprint). (*Moulin Decl.*, ¶ 37). When the audio cue associated with a particular program is played at the beginning of the program, the start event flag is set, and the recording system detects the flag and is activated. (*Murphy*, 13:17-34). And upon the “repeated occurrence” of the same audio cue in the program, an end flag is set, which causes the recording system to stop recording. (*Murphy*, 11:1-4 (disclosing setting an end flag on the “repeated occurrence of the audio cue”), 14:1-3 (disclosing the system looking for an end flag to stop recording)).

There is no teaching or suggestion in Murphy that the audio cue signature stored in the Murphy database (i.e., the mapped reference fingerprint) identifies the specific time point at which any action is to be taken. Indeed, the recording system in Murphy is activated without regard to any particular time point identified by the

audio cue signature. Petitioner’s expert, Dr. Tewfik, even admits this: “[b]ecause the analysis of the audio cues is occurring during playback, and because **the process of recognizing the audio cues** takes time and *then causes the immediate triggering of the action....*” (*Tewfik Decl.*, ¶ 170). Thus, even per Dr. Tewfik’s theory, it is not the audio cue signature that identifies the time for performing the action but rather “the process of recognizing the audio cue” itself. Because recognizing the audio cue “**causes**” the triggering of the action without regard for any associated time point, Murphy does not teach the stored audio cue signature identifies a time point at which an action is to be taken. *Id.*

Additionally, Murphy teaches obtaining audio cue signatures from an off-air recording or by downloading the audio cue from a broadcaster. (*Murphy*, 14:4-15). A downloaded audio cue signature that is then stored in the database does not identify a reference time point because the signature has no relationship to the particular broadcast audio signal being played over loudspeaker 90.

Murphy also does not and could not teach the audio cue signature identifying a reference time point because the same audio cue is repeated during the program and matched to the same stored audio cue signature. The stored audio cue signature (i.e., the mapped reference fingerprint) merely enables the system to set an event flag when a match is detected. Because the first appearance of an audio cue in a signal sets the start flag, and the “repeated occurrence” of the same audio cue sets

the end flag (*Murphy*, 11:1-4), the audio cue signature prompts events at *different* time points within the program, thus indicating the audio cue signature does not identify a reference time point. (*Murphy*, 11:1-4; *Moulin Decl.*, ¶ 37).

Petitioner's own mapping highlights the difference between the claimed invention and *Murphy*. Petitioner states "Murphy discloses taking actions (e.g., display information, start recording, stop recording) at a time referenced to the location of the cues in the signal, for example, the time the cue is recognized." (Pet 29). Petitioner thus maps the reference time point to the "time the cue is recognized." But of course, the stored audio cue signature cannot identify a reference time point as the "time the cue is recognized" because the stored signature has no relationship to the time when the cue is recognized. (*Compare to '114 Patent*, 6:18-20 ("The playback device simply needs to be aware of the relationship between a given fingerprint and the given time point used during generation.")).

*Murphy*'s system is limited to taking action only upon recognition of the repeated audio cues already present in the program being broadcast. Because *Murphy*'s audio cue signatures are matched multiple times and prompt different

actions each time they are matched, a signature (reference fingerprint) cannot serve to identify a reference time point at which an action is to be taken.<sup>7</sup>

**B. Claims 1, 8, and 10: Murphy does not teach “reference fingerprints previously derived” from the “content being played back”**

Claim 1 recites the steps of “determining a plurality of trigger fingerprints *from content being played back*” and “accessing a database that includes a plurality of reference fingerprints *previously derived from the content*.” Claims 8 and 10 recite substantially similar limitations. These limitations require the reference fingerprints be derived from the same “content” that is subsequently used to determine trigger

---

<sup>7</sup> Tellingly, after the combination of Murphy and Brunk was disclosed during prosecution of a later family member of the '114 Patent, the Examiner indicated Murphy failed to teach a limitation requiring the action be executed in association with a time point which indicated when in the content such action was to be executed, saying “the fact that the art provides for execution of an action requires speculation as to the time point as recited. For the Examiner to fit the separate pieces of the reference together to render the claims obvious would be resorting to a clear hindsight reconstruction.” (Ex. 2008, '134 *Application Prosecution History*, 227). This lack of a time point is precisely the same deficiency in the prior art noted during the prosecution of the other family members.

fingerprints when “played back.” The Petition relies on Murphy to teach these limitations. (Pet 28). Murphy does not teach or suggest either limitation.

The Petition maps the claimed “content” as the broadcast audio signal taught by Murphy. (Pet 21). The broadcast audio signal in Murphy includes an audio cue associated with a particular program “immediately before” the program broadcast begins. (*Murphy*, 3:10-16, 3:20-23). Murphy’s receiver (mapped to the claimed playback device at Pet 21) checks for a match between the received broadcast audio signal and any of the audio cue signatures stored in the template database. (*Murphy*, 13:17-21). The Murphy audio cue signatures used to populate the template database (i.e., mapped “reference fingerprints”) are derived independently from the content being broadcast (for example, from a different episode of the same program). Thus, the audio cue signature in Murphy is not “previously derived from the content,” as claimed, where the content is the same content that is “played back on a playback device,” as claimed.

Murphy teaches several methods of obtaining a template signature to be compared to the signature generated from the broadcast audio signal. In a first method, a template database 12 “could be populated by downloading from the particular broadcaster, or a centrally provided database, the signature of the audio cue for the particular programme.” (*Murphy*, 14:4-9). Alternatively, the audio cue itself could be downloaded and a signature derived locally. (*Murphy*, 14:10-11). In

these methods, there is no deriving the template signature of the audio cue *from the content*. Instead, the signature (or cue) is downloaded and stored without any relationship to the particular *episode* (or episodes) of the program later broadcast and recorded using that signature.

Murphy teaches downloading the signature of the audio cue for the particular *program* (the Petition gives the example of “the famous ticking clock from CBS’s ‘60 Minutes’”; Pet 18), but because the audio cue is reused across multiple *episodes* of the program, the downloaded audio cue is not derived from the content (broadcast audio signal) being played, i.e., a particular episode. (*Murphy*, 14:15-17 (recording subsequent episodes using the same cue)). Using Petitioner’s example, the 60 Minutes’ ticking clock has been broadcast thousands of times and is not “previously derived from” any particular episode of 60 Minutes. Indeed, Murphy includes no disclosure of downloading a signature derived from the content actually played (or played back), such as an actual episode of 60 Minutes.

Murphy also teaches that a user can record a sample of the audio cue from one broadcast and use it to generate a signature to allow “subsequent episodes” of the program to be recorded automatically. (*Murphy*, 14:11-17). This disclosure in Murphy also does not teach deriving the reference fingerprint from the content (where such content is the content being played back). Murphy expressly teaches the stored audio cue signature is from one broadcast episode of the program and is

matched against subsequent broadcast episodes of the program. The “subsequent episodes” of the program being played over loudspeaker 90 are not the content from which the audio cue was originally derived. The audio cue is not derived from a single episode of the program that is then played, as in the scenarios discussed above.

The Petition makes no attempt to reconcile these claim limitations. The claimed “content being played back” is mapped to Murphy’s “broadcast audio signal.”<sup>8</sup> (Pet 21). Later, when addressing the limitation requiring that the fingerprints be “previously derived from the content,” the Petition gives the source of the audio cue as either (1) a “preprocessed” and “a clean, noise-free version of the signal” (i.e., one that is not part of the content); or (2) “versions previously broadcast” (i.e., one that is broadcast as part of *entirely different* content, such as a prior episode of the program). (Pet 28). A preprocessed clean version of the audio cue is self-evidently not derived from the broadcast audio signal being played. (*Moulin Decl.*, ¶ 38). And a version of the audio cue obtained from a previous broadcast also is not derived from the broadcast audio signal being played because, as stated, the audio cue is from a *previous* broadcast. Per the Petition’s own

---

<sup>8</sup> The Petition does not contend the broadcast audio signal from which an audio cue is derived is actually “played back” at all, and instead merely contends the system “plays *a* broadcast audio signal (*content*)” (emphasis added). (*See* Pet 21).



admission, the “content being played back” was not used to “previously derive[]” the fingerprints being matched.

Finally, the Petition does not purport to rely on Brunk to teach “reference fingerprints previously derived from the content.” Regarding Claim 1b, Brunk is relied upon solely to remedy Murphy’s lack of a database storing the actions in addition to the fingerprints. (Pet 28-31 (the Murphy/Brunk system characterized as using Murphy’s covariance matrix as the database index pursuant to Brunk), 43 (Brunk’s teaching as to the corresponding limitation in claim 8 characterized as “storage of actions in the database”)). Brunk’s teachings cannot be used to remedy Murphy’s lack of teaching of “reference fingerprints previously derived from the content.” (*See also* IPR2020-00217, Paper 1, *Petition*, 29 (Petitioner admitting, for the ’962 Patent, that Brunk does not teach a fingerprint triggering an action at a trigger time point)).

Because Murphy (alone or in the proposed combination with Brunk) does not teach “reference fingerprints previously derived” from the “content being played back,” Petitioner has not established a reasonable likelihood claims 1, 8, and 10 are unpatentable.

**C. Claims 11, 20, and 21: Murphy Does Not Teach “an action to be triggered during playback” of the Associated Fingerprinted Segment**

Claim 11 recites (twice) an “action to be triggered during playback of the segment.” (*See ’114 Patent*, Claim 11, 9:36-37, 9:41-43). Claims 20 and 21 recite substantially similar limitations. The Petition relies on Murphy for teaching triggering an action *during playback*. (Pet 48). Murphy fails to teach or suggest this limitation.

Murphy’s receiver (either unmodified or modified as proposed in the Petition) does not trigger any actions “*during the playback*” of the segment associated with the fingerprinted segment. Rather (as the Petition concedes elsewhere, *see* Pet 34-35), Murphy’s unmodified system triggers an action *after the segment* in which the signature is detected. Petitioner’s declarant, Dr. Tewfik, expands on this, stating:

Because the analysis of the audio cues is occurring during playback, and *the process of recognizing the cue takes time* and then causes the immediate triggering of the action, one of ordinary skill in the art would understand Murphy to describe the corresponding action being taken at the end of the segment that constitutes the audio cue, *after the entire cue is able to be processed. This is necessarily true* because the system cannot determine which reference action it must take until after the audio cue is recognized.

(Ex. 1007, *Tewfik Decl.*, ¶ 170). Per Dr. Tewfik’s express admission, Murphy’s action is triggered *after the end of the audio cue is played* due to the required

processing time. Thus, Petitioner concedes that Murphy's system "necessarily" cannot trigger actions until *after* the playback of segment corresponding to the audio cue.

Recognizing this deficiency, Petitioner's mapping requires two different modifications to Murphy's system to map triggering actions "during the playback," as recited claim 11. Petitioner's first proposed modification to Murphy's system is modifying the receiver to incorporate the time-delay unit of the transmitter. (Pet 48 (referring back to mapping of limitation 1a); *Tewfik Decl.*, ¶ 100 (Dr. Tewfik contending "it is apparent to one of ordinary skill in the art that the same time delay concept could be employed with regard to that receiver-side embodiment")). Petitioner's second proposed modification is setting the time delay to be "slightly less than the length of the segment." (*Tewfik Decl.*, ¶ 218). With these two modifications, Dr. Tewfik contends the twice-modified Murphy system would then be "enable[ed]" to perform limitation 1a. *Id.*

Petitioner states "[a] POSITA would have understood that Murphy's time delay can be employed to trigger playback of the segment containing the fingerprint that triggers the action." (Pet 48). Elaborating on this, Dr. Tewfik opines "one of ordinary skill in the art would have understood that Murphy's teachings regarding use of a time delay enable the system to trigger an action during the content

containing the fingerprint by setting the delay to slightly less than the length of the segment.” (*Tewfik Decl.*, ¶ 218).

Contrary to Dr. Tewfik’s claim, Murphy does *not* teach any particular delay length for the time-delay system, much less Dr. Tewfik’s “slightly less than the length of the segment” delay setting that would allegedly cause an action to be triggered during the segment. And Murphy does not otherwise teach triggering an action *during* the content containing the fingerprint.

Petitioner’s claim mapping does not establish how the art teaches triggering an action “during playback of the segment.” Dr. Tewfik’s modification to the Murphy system to an (untaught) delay of “slightly less than the length of the segment” to *enable* the system to perform the claimed limitation also does not amount to a teaching of the limitation. Note that Petitioner does not argue it would have been *obvious* to twice modify the Murphy system to perform the claimed step. For example, there is no contention in the Petition or by Dr. Tewfik that modifying the time delay to be “slightly less than the length of the segment” would have been obvious or why a POSITA would have been motivated to make the modification. (Pet 48; *Tewfik Decl.*, ¶ 218). Rather, Dr. Tewfik contends a POSITA would have understood that Murphy’s receiver, modified to include the time delay, can then be *enabled* to perform the step. (*Tewfik Decl.*, ¶ 218). Simply modifying Murphy so that it is *enabled* to perform a step, as Dr. Tewfik contends, does not rise to a teaching

or suggestion that Murphy *actually performs* the step or that such would have been obvious to a POSITA. *Belden Inc. v. Berk-Tek LLC*, 805 F.3d 1064, 1073 (Fed. Cir. 2015) (“[O]bviousness concerns whether a skilled artisan not only *could have made* but *would have been motivated to make* the combinations or modifications of prior art to arrive at the claimed invention.”) (emphases in original); *Westinghouse Air Brake Tech. Corp. v. Siemens Ind., Inc.*, IPR2017-00650, Paper 12 at 16 (PTAB July 19, 2017) (in responding to petitioner’s argument that a POSITA would recognize the combination of elements from the prior art, stating “[t]his amounts to an argument that the individual elements of the claim were known and that combining them would have yielded the claimed invention, but Petitioner does not explain *why* an ordinary skilled artisan would have combined those elements in the claimed manner.”) (emphasis in original).

Furthermore, modifying Murphy’s time-delay unit to use such a time-delay setting runs directly counter to Petitioner’s stated motivation for modifying the receiver to incorporate the time-delay unit in the first place:

Such jingles are also recognized by humans and are often part of the program that one would want to record, just as one would begin recording a television program to include the introductory credits and/or theme song. Accordingly, in my opinion, a person of ordinary skill in the art would have appreciated that the time delay would be calibrated in such a way that the jingle itself was included as part of the recording.... In other words, one of ordinary skill in the art would find

it obvious to configure the length of delay unit 74 **to be the same standardized length as the audio cue segment** to allow the recording action start at a time that includes the content that serves as the reference audio cue.

(*Tewfik Decl.*, ¶¶ 175-76 (internal citation omitted)). By modifying the time-delay unit to use a delay “slightly less than the length of the segment,” as proposed for claims 11, 20, and 21, the system would be deliberately configured to record *only a portion* of the “introductory credits and/or theme song.” (*Moulin Decl.*, ¶ 39). But, the result of recording only a portion of the audio cue (i.e., jingle) is directly contrary to Dr. Tewfik’s stated motivation to modify Murphy’s receiver to incorporate the time delay circuit. (*Tewfik Decl.*, ¶ 175 (“Accordingly, in my opinion, person of ordinary skill in the art would have appreciated that the time delay would be calibrated in such a way that the jingle itself was included as part of the recording.”)). Neither Murphy nor the Petition provides an indication of why such a modification would be obvious (or even desirable).

If, as Dr. Tewfik contends, the jingle is “part of the program that one would want to record” (*id.*), then such a modification would also directly contradict Murphy’s stated rationale for providing the time delay circuit in the transmitter embodiment:

**In order to avoid the possibility of losing the initial part of the programme,** the automated recording system 76 may be configured to

record a delayed version of the audio signal provided by a time delay unit 74.

(*Murphy*, 12:23-26). Murphy teaches away from triggering an action (starting recording) during the playback of the associated segment (the jingle) under Dr. Tewfik's theory that the jingle is part of the program because it would "lose the initial part of the programme" (*i.e.*, the portion of the jingle that precedes triggering the action). (*Moulin Decl.*, ¶ 39).

Finally, even assuming that Murphy were modified to incorporate the time-delay unit as Petitioner proposes, the time-delayed content would be the content that is *recorded*, not the content that is *played* over loudspeaker 90. Petitioner contends that, by incorporating a time-delay unit into Murphy's recorder, a recording action could be triggered during "playback" of the segment containing the fingerprint that triggers the action. (Pet 48). However, Petitioner also identifies the alleged "playback" function of Murphy's recorder as the loudspeaker 90, which is connected directly to the off-air audio signal. (Pet 21). Thus, the signal *played over the loudspeaker* in Murphy's (modified) system is not delayed, only the *recorded* version of the signal is delayed. (*See also Murphy*, FIG. 3 (showing the time delay unit in the transmitter embodiment delaying only the recorded version of the audio signal)). Because the content played over the loudspeaker would not be delayed, the action triggered would still take place *after* it is played over the loudspeaker (*i.e.*,

after the alleged “playback”) of the segment, rather than during (as the claims require).

Murphy does not teach or suggest triggering an action during the playback of the segment containing the jingle on which the signature is based, and a POSITA would not have found it obvious to modify Murphy to include a time delay at the receiver that is “slightly less than the length of the segment” to enable the twice-modified Murphy system imagined by Petitioner to perform the claimed limitation.

**D. Claims 11, 20, and 21: Murphy does not teach “accessing segments of content” During Fingerprint Generation**

Claim 11 recites steps of “accessing segments of content, a segment among the segments being associated with a corresponding action to be triggered during playback of the segment,” “generating a reference fingerprint of the segment associated with the corresponding action,” and providing an identifier “causing the playback device to trigger the corresponding action during the playback of the segment.” (*’114 Patent*, Claim 11 (specifically, limitations 11a, 11b, and 11e), 9:35-37, 9:38-39, 9:50-52). Reading these steps collectively, claim 11 recites accessing a segment of content to generate a fingerprint for that segment, where the segment is associated with an action to be triggered during playback of that segment. Claims 20 and 21 recite substantially similar limitations. The Petition relies on Murphy for teaching these limitations. (Pet 48-50, 52). Murphy fails to teach or suggest these limitations.

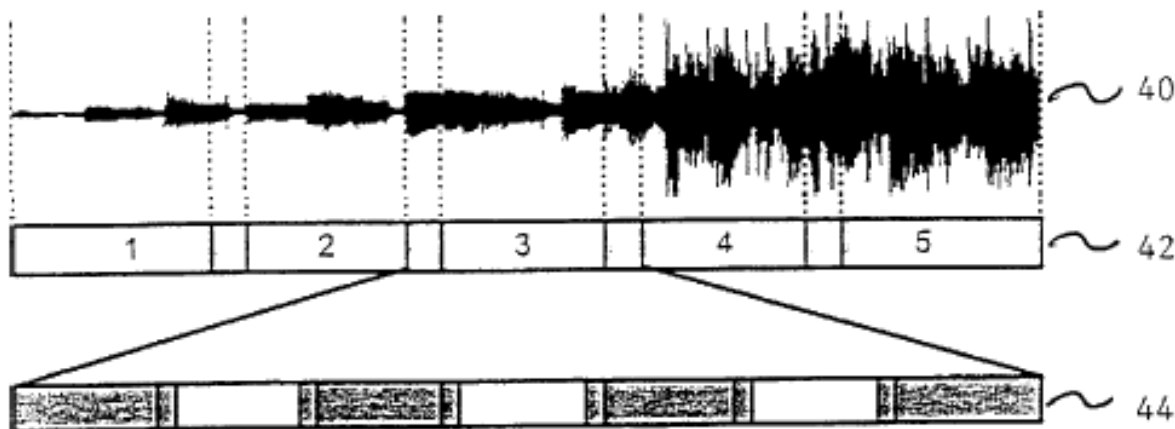


The Petition provides two alternative mappings for the step of “accessing segments of content”: “Murphy discloses accessing such segments of content in at least two contexts: *First*, during playback (as discussed extensively for claim 1); and *second*, in support of the generation of reference fingerprints for storage in a database.” (Pet 48) (emphasis in original). Each of these arguments (accessing during playback and accessing during fingerprint generation) is addressed individually, below.

**1. Murphy’s System Accesses Different Segments During Playback Than During Fingerprint Generation**

The claimed step of “accessing” occurs in order to generate the reference fingerprint, as recited in claim 11b. (*’114 Patent*, 9:38-40 (Claim 11b reciting “generating a reference fingerprint *of the segment* associated with the corresponding action,” where “the segment” is recited in Claim 11a: “a segment among the segments”)). Petitioner contends Murphy accesses segments of content, identifying Murphy’s windows of the sampled audio signal. (Pet 48, citing the mappings for limitations 1a and 1b and *Tewfik Decl.*, ¶¶ 214-215; *Murphy*, 7:28–8:8). The windows of the sampled audio signal 40 illustrated in Fig. 2 of Murphy are used for *matching* the audio cue signatures during playback – not for generating the audio cue signature, as recited in Claim 11b. (*Murphy*, 7:28-33 (referring to the audio cue to be identified and recognized from the audio signal)).

In more detail, the windows (segments of the content) that Murphy accesses during playback are not the same as the audio cue used to generate the matching audio cue signature, at least because the start and end times of the window will not align with the playback of the audio cue. (*Murphy*, FIG. 2 (excerpted below) (showing the received broadcast audio signal being divided into overlapping windows 42)).



Petitioner's identification of Murphy's windows with the segment used for fingerprint generation in the Claims would require that one of these windows 42 align with the audio cue. However, Murphy's system is designed around the assumption that the windows *will not* align with the played audio cue.

Murphy includes specific disclosure about how segments (windows) are sampled during playback:

A window generator 22 segments the sampled audio signal into windows corresponding in length to the length of the audio cues to be

recognised. **The window generator 22 produces overlapping windows 42 from the audio signal 40 as shown in figure 2.** There is a trade off between the amount of overlap between the windows and the amount of processing. Missing the start of the audio cue may result in failing to recognise the audio cue. However, moving the window by a single sample each time results in very high processing overheads and slow processing time. We have found that a suitable overlap is one second.

(*Murphy*, 7:33–8:8). These overlaps are needed because aligning the window precisely with the start of the audio cue is impossible without knowing in advance where in the played-back broadcast audio signal the audio cue will appear, and sampling every frame results in unacceptable processing requirements. (*Moulin Decl.*, ¶ 40, *citing Murphy*, 8:5-7). *Murphy* also teaches that the system would “fail to operate adequately” due in part to “sample-by-sample synchronization issues” when relying on direct comparison of the stored cue to the sampled audio signal, even assuming the window and audio cue happened to line up. (*Murphy*, 5:17-23). As a result of this overlapped sampling process, each window that contains a portion of the audio cue will also contain a portion of the broadcast audio signal that *does* not correspond to the audio cue (i.e., was not used in *Murphy* to generate the audio cue signature). *Id.* And the windows sampled from the played-back broadcast audio signal (mapped as the claimed “segments”) are distinct from the segment consisting

only of the audio cue used during audio cue signature generation in Murphy. (Pet 47-48).

Murphy thus does not teach the segment accessed to generate the fingerprint is the same as any window (mapped to the claimed segment at Pet 48) used for sampling during play back. This is because any segment window sampled by Murphy necessarily includes only a portion of the audio cue and is different from the segment from which the reference audio cue signature is generated (i.e., a segment that includes the entire audio cue and nothing else).

## **2. Murphy Does Not Access the Content During Fingerprint Generation**

The Petition contends Murphy teaches accessing segments of the content “in support of the generation of reference fingerprints for storage in a database.” (Pet 48). As discussed above for the “reference fingerprints previously derived from the content” for claim 1, the system of Murphy does not generate signatures based on the “broadcast audio signal” mapped by the Petition to the claimed “content” but instead generates signatures based on the audio cue itself. Murphy does not access the *content* during signature generation but rather accesses a copy of the audio cue obtained from a signal other than the broadcast audio signal (i.e., the claimed “content”).

The claimed “content” is mapped by Petitioner as the “broadcast audio signal” in Murphy. (Pet 21). Per Claim 11e, the action is triggered during the playback of a

segment of the content. This same segment of the content, per Claim 11, is the segment that is accessed to perform the fingerprint generation, per Claims 11a and 11b. There is no teaching or suggestion in Murphy that a segment of the broadcast audio signal that is subsequently played back (per Claim 11e) is the segment that is accessed to perform generation of the audio cue signature.

In Murphy, the audio cue signature may be obtained several different ways:

1. The audio cue is pre-processed to obtain the signature, which includes “an individual determining characteristic.” (*Murphy*, 6:30-35).
2. The signature of the audio cue may be processed by the recognition system or generated by a compatible system and downloaded into the recognition system. (*Murphy*, 7:1-5).
3. A clean, noise-free version of the audio cue may be used. (*Murphy*, 7:5-7).
4. An “audio cue *recorded from a **previous** broadcast signal*” may be used. (*Murphy*, 7:5-7).
5. An audio cue signature may be downloaded from the broadcaster or a central database. (*Murphy*, 14:4-9).
6. The audio cue (as opposed to the signature) could be downloaded, pre-processed, and stored. (*Murphy*, 14:10-11).

7. The user could input a sample of the audio cue recorded **off-air**, and the recognition system could process and store the audio cue. (*Murphy*, 14:11-15).

None of these examples of generating the audio cue signature teaches or suggests that the audio cue that is accessed is *from a segment of the broadcast audio signal that is played*. Going through each of the examples in *Murphy* highlights this. For an audio cue that is pre-processed to obtain the signature and stored in the recognition system's database (Examples 1-3 and 6, above), or otherwise downloaded from a central database (Example 5), this signature is not generated from the accessed segment of the broadcast audio signal. Similarly, for an audio cue signature that is recorded **off-air** (Example 7), the signature is not generated from accessing the *current* audio broadcast signal because the signature is obtained from an **off-air** audio cue. In fact, there is no teaching or suggestion in *Murphy* that for these examples, the audio cue signature is taken from *any* broadcast audio signal, let alone the broadcast audio signal being played back. Instead, *Murphy* teaches the audio cue itself is sampled, processed, and stored in the "template database," such that there is no association or relationship of the audio cue signature with a particular broadcast audio signal. (*Murphy*, 14:4-11).

For an audio cue signature that is "recorded from a previous broadcast signal" (Example 4), the signature is not generated from accessing the *current* audio

broadcast signal (or segment thereof). There is **no** discussion in the Petition or supporting Declaration asserting otherwise. (Pet 48; *Tewfik Decl.*, ¶ 215 (Dr. Tewfik discussing storing a pre-processed audio cue but nothing asserting a recorded audio cue from a *previous* broadcast signal satisfies Claim 11)).

The Petition seems to appreciate the deficiency in the Mapping, as the Petition contends Murphy accesses segments of the content during fingerprint generation, but then admits in the next sentence that “Murphy teaches generating fingerprints *by accessing audio cues.*” (Pet 48). Accessing audio cues (such as from central databases or processing and storing) is not the same as generating the signature by *accessing segments of the actual content.* (Pet 48). If Murphy taught or suggested that the content being played back was the same content (or a segment thereof) that was accessed to generate the audio cue signature, Petitioner would have simply relied on that teaching. But Murphy does not include that teaching, leaving Petitioner to gloss over the critical distinction between Murphy’s audio cues and the claimed content.

Petitioner cites to Dr. Tewfik, who opines “Murphy teaches dividing a signal into segments referred to as ‘windows’ and determining whether the covariance matrix ‘fingerprints’ of those segments *match* any fingerprints stored in database 212.” (Pet 48, *citing Tewfik Decl.*, ¶ 213). In this paragraph, Dr. Tewfik is opining that the broadcast audio signal is divided into windows to determine if the signal

includes any fingerprint “matches” to those stored in a database. Therefore, this paragraph is referring to *matching* the audio cue signature – not *generating* the audio cue signature.

Dr. Tewfik goes on to opine “Murphy teaches that the same process of *accessing the content segment by segment occurs when fingerprints are generated for storage in database 12.*” (Pet 48, *citing Tewfik Decl.*, ¶ 215). Dr. Tewfik is incorrect. Murphy does *not* teach the same process of accessing the content is used for signature generation. (*Moulin Decl.*, ¶ 41).

Dr. Tewfik cites Murphy at 6:32–7:5 and 14:4-15. (*Tewfik Decl.*, ¶ 215). Section 6:32–7:5 of Murphy is discussing taking the audio cue signal (not the broadcast audio signal) and preprocessing the audio cue signal to determine an individual determining characteristic or signature. There is no teaching or suggestion in this section (6:32–7:5) that the audio cue signal is taken from the broadcast audio signal. The second section relied on by Dr. Tewfik at 14:4-15 discusses populating a template database by downloading the audio cue signature or downloading the audio cue and processing to obtain the signature. Again, there is no teaching or suggestion in this section (14:4-15) that the audio cue signal is taken from the broadcast audio signal. Thus, there is no teaching or suggestion in the Murphy sections cited by Dr. Tewfik (or in any other section of Murphy for that matter) that



the audio cue signature is generated from a segment of the audio broadcast signal to be played back.

It is insufficient for the audio cue signature to be obtained from a *prior* broadcast of a program, because such prior broadcast is not the *same* broadcast audio signal to be played back. When the audio cue is sampled from a first broadcast, and the Murphy system accesses a second broadcast, the Murphy system is accessing a different segment of the content to trigger the action (per Claim 11e). Claim 11 of the '114 Patent is specific that it is a segment of the *content* that is played back that is accessed to generate the fingerprint and not simply any segment that may be included in any unspecified broadcast audio signal. This claimed feature corresponds to one of the discussed fundamental differences between the claimed invention of the '114 Patent and Murphy, namely generating the fingerprint from the content *to be played back* to allow for very accurate and very precise time markings. ('114 Patent, 4:65-67).

Because Murphy does not teach or suggest accessing a segment of the content to generate the fingerprint, where the accessed segment is played back for triggering the action, Petitioner has not established there is a reasonable likelihood claims 11, 20, and 21 are unpatentable.

## **VI. THE CLAIMED INVENTION OF THE '114 PATENT DIFFERENTIATES OVER MURPHY IN VIEW OF BRUNK**

The above sections discuss how Murphy does not teach or suggest various limitations for which Petitioner cites Murphy. This section focuses on why the claimed invention of the '114 Patent presents advantages not provided by Murphy.

Murphy does not offer the advantage of an enhancement service provider having control of where to trigger an action through selecting reference time points in the multimedia signal and creating triggers for those time points by deriving fingerprints of the multimedia signal at those reference time points. (*'114 Patent*, 2:1-4). In Murphy, because the audio cue is placed in the broadcast audio signal by the program maker or broadcaster, an action can only be taken where the audio cues have been added (per the disclosure of Murphy, “immediately before a particular programme” or “marking the beginning or end of a particular programme”). (*Murphy*, 10:32–11:4; 3:12; 6:8-9). As a result, Murphy is useless in the specific scenarios identified by the '114 Patent of replacing an advertisement with a local advertisement or showing biographic data for an actor when they appear on screen. (*'114 Patent*, 1:11-62). The prior art problem the '114 Patent sought to overcome – being limited to triggering actions from triggers marked, indicated, or inserted by program makers or broadcasters – is the very problem inherent in systems such as Murphy.

Additionally, Murphy's signature-matching technique is unsuitable for the use cases of the '114 Patent, where "frame-accurate" (i.e., 1/24-1/60 of a second) detection of fingerprints is possible. (*'114 Patent*, 5:2-4; *Moulin Decl.*, ¶ 42). In Murphy, the recognition of an audio cue may be delayed by as much as three seconds from the beginning of the audio cue being played. (*Moulin Decl.*, ¶ 42). Murphy's system is expressly designed to delay recognition of an audio cue (and the subsequent prompting of an action), in order to allow the audio cue to fully play while still limiting processing demands. (*Murphy*, 7:28–8:8 (discussing the standard audio cue length is 2 seconds, that "missing the start of the audio cue may result in failing to recognise the audio cue," and that sampling the broadcast audio signal with overlapping segments of 1 second addresses the "trade off" between the amount of processing and slow processing time)). Because Murphy may delay recognizing the audio cue and thus, performing the action, by up to 3 seconds from the time the audio cue starts playing, the frame accurate fingerprint matching and taking action made possible by the '114 Patent is not possible.

Because Murphy's system does not generate the fingerprints from the broadcast audio signal being played and does not generate the fingerprint from a reference time point in the signal, Murphy's system is simplistic and limited in its use. ***Petitioner's own example*** illustrates this. The Petition gives the "ticking clock" of 60 Minutes as one example of an audio cue with which Murphy can be used. (Pet

18). The “ticking clock” audio cue, per Petitioner, “appears at the beginning and upon return from commercial break.” *Id.* Murphy expressly teaches that “[o]nce a start flag has been generated, an end flag could be generated on the *repeated* occurrence of the audio cue or on the occurrence of a second, play-out, audio cue.” (*Murphy*, 11:1-4). Petitioner admits that all that happens when the audio cue is matched is that an event flag is generated, which in turn causes start/stop recording. (Pet 18). Because, per the Petition, the same cue is reused “at the beginning and upon return from commercial break,” (Pet 18) a start flag would be generated to begin recording (upon a first recognition of the audio cue), and a stop flag would be generated to stop recording (upon a second recognition of the audio cue). (*Moulin Decl.*, ¶ 43). But the stop flag would be generated “upon return from commercial break” because the same audio cue is used both times, resulting in recording being prematurely stopped. Murphy cannot distinguish between repeated occurrences of the cue used for different purposes (for example, to indicate a return from commercial break or to indicate the end of the program) because Murphy does not generate its audio cues from a reference time point or derive the cues from the broadcast audio signal.

Thus, Petitioner’s own example applying Murphy does not achieve the advantages of the claimed invention. With the Murphy system, the enhancement service provider is dependent on where, in the broadcast audio signal, the program

maker or broadcaster has placed audio cues to otherwise provide the enhancement services. In contrast, the '114 Patent's claimed invention allows the enhancement service provider to identify where in the content the enhancement is to be provided by selecting the reference time point, generating the fingerprint from the reference time point, and playing back the content from which the fingerprint was generated.

## **VII. BOTH MURPHY AND BRUNK ARE CUMULATIVE TO PRIOR ART CONSIDERED AND OVERCOME DURING ORIGINAL EXAMINATION**

### **A. Murphy Is Cumulative with Rhoads, Which Was Cited and Considered During Prosecution**

Murphy is directed to recognizing audio cues in the form of jingles that are placed immediately before or after the beginning of a particular program to identify that program. Murphy describes how flags generated in response to recognizing a program (without regard for any time point) cause certain actions to be taken, such as starting recording, updating program-type data, or generating an audit log. (*Murphy*, 10:32–12:7). Rhoads, which was the primary reference relied upon during the prosecution of the '000 Patent to which the '114 Patent claims priority, operates similarly.

Rhoads uses data derived from audio content to identify the content and thereby trigger actions corresponding to the content. (Ex. 2004, *Rhoads*, Abst.). Like Murphy, Rhoads does not teach or suggest identifying a particular time point (either a reference time point or a trigger time point) in content, but instead acts based on

an identification of the content as a whole. During prosecution of the '000 Patent, Applicant identified Rhoads's lack of specific teachings as to time points as a deficiency in the use of Rhoads to render the claims obvious. (Ex. 2005, '000 *Prosecution History*, pp. 756-57). In particular, Rhoads (like Murphy) teaches sampling the beginning of the content to determine a signature for that content that could be matched against known signatures. (*Rhoads*, ¶ [0014] (sampling first 30 seconds of a song to derive a signature)). Also like Murphy, Rhoads only discloses triggering actions that are relevant to the content as a whole, rather than a particular time point within the content. (*Rhoads*, ¶ [0026] (actions including play/record control limitations, displaying title/artist of the song, and offering user the option to purchase the song))).

As such, Murphy's teaching of triggering program-type information display and starting/stopping recording responsive to identifying a particular program (without regard for any time point) would have been cumulative with the teachings of Rhoads, which were in front of the Examiner during the prosecution of the '000 Patent. For this reason, Murphy's teachings would have been cumulative to the already-of-record teachings of Rhoads, and the Examiner for the '114 Patent would have been aware of the teachings of Rhoads and considered them before allowing the claims of the '114 Patent. Thus, applying the *Becton Dickinson* factors, Murphy

is cumulative to art already of record and is not relevant to the patentability of the claims of the '114 Patent.

**B. Murphy Is Also Cumulative with Haitsma, Which Was Cited and Considered During Prosecution**

During prosecution of the '114 Patent, the Examiner relied upon U.S. Patent Application Publication No. 20020178410 to Haitsma, et al. to reject the claims as anticipated and/or obvious. (Ex. 2009, *'114 Prosecution History*, pp. 205-208). Like Murphy and Rhoads, and unlike the claims of the '114 Patent, Haitsma describes hashing techniques for identifying an *entire piece of multimedia content*, rather than a particular (trigger or reference) time point within a file. (Ex. 1016, *Haitsma*, Abst (“Hashes are short summaries or signatures of data files which can be used to identify the file. ... The disclosed method generates robust hashes for multimedia content, for example, audio clips.”)). Summarizing an Examiner interview, the Examiner stated the following:

Applicant explained the differences between a “time point” versus the intervals/multiple points in time taught by Haitsma, as well as the differences between a fingerprint (does not alter the data) versus a watermark (inserted in the data, thus altering it). At first consideration, the examiner feels these arguments are convincing and will consider them further when submitted formally.

(*'114 Prosecution History*, p. 218). Applicant subsequently presented these arguments formally (characterizing the distinction as “an action performed in

response to identifying an entire song, as described in [Haitsma], in contrast with an action that corresponds to a time point within a piece of content,”) and the prior art rejections were withdrawn. (*'114 Prosecution History*, p. 238, 242-244; 281). Applicant thus again distinguished the claimed invention over a reference that identifies an entire piece of content because the claims required a “time point” within the content.

Murphy, like Haitsma, describes the use of a signature of a piece of multimedia content to *identify that content*. (See, e.g., *Murphy*, 13:30-33 (“The audio cue may be assigned a particular code and when *the program starting with that audio cue is to be recorded*, that code is entered on the recording system 82.”)) Murphy simply uses the distinct jingle at the beginning of a program to recognize the entire program. For this reason, Murphy’s teachings would also have been cumulative to the already-of-record teachings of Haitsma, and the Examiner for the ’114 Patent was aware of the teachings of Haitsma and considered them before allowing the claims of the ’114 Patent. Thus, applying the *Becton Dickinson* factors, Murphy is cumulative to art already of record and is not relevant to the patentability of the claims of the ’114 Patent.

### **C. Brunk Was Cited and Considered During Prosecution**

Brunk, even more so than Murphy, is cumulative with art cited during the prosecution of the family of the ’114 Patent. The Petition, in its “Overview of Brunk”



section, states: “[n]either Brunk, nor any corresponding application publishing its disclosure, was cited during prosecution of the ’114 Patent or its parents. Ex-1004.” (Pet 19). This is incorrect. As discussed below, a continuation-in-part of Brunk, incorporating its disclosure, was cited by Applicant during the prosecution of a parent of the ’114 Patent and explicitly discussed, in combination with Rhoads, during an Examiner Interview.

U.S. Patent Application Publication No. 2003/0021441 to Levy, *et al.* is a continuation-in-part of Brunk<sup>9</sup> and Brunk was explicitly incorporated by reference into Levy.<sup>10</sup> (Ex. 2007, *Levy*, ¶ [0001] (“**The present application is also a continuation in part of U.S. patent application Ser. No. 10/027,783**, filed Dec. 19, 2001, which claims the benefit of U.S. Provisional Application No. 60/263,490, filed Jan. 22, 2001. **Each of these U.S. patent documents is herein incorporated by reference.**”)). *Levy* was cited by Applicant in an Information Disclosure

---

<sup>9</sup> Brunk, identified in the Petition as U.S. Patent Application Publication No. 2002/0126872, was filed as Application No. 10/027,783. (Ex. 1012, *Brunk*, (21)).

<sup>10</sup> Levy does not explicitly include the disclosure of Brunk; however, in addition to incorporating Brunk by reference, Levy independently renders Brunk cumulative with prior art considered during prosecution. (*See Levy* ¶ [0054] (fingerprint used as identifier to index database); ¶ [0065] (identifier associated with action in database)).

Statement during prosecution of the '000 Patent to which the '114 Patent claims priority, and appears on the face of both the '000 Patent and '114 Patent. ('000 *Prosecution History*, pp. 77; Ex. 2006, '000 *Patent*, (56); '114 *Patent*, (56)).

Furthermore, Levy was explicitly discussed during an Examiner interview as a reference that would need to be overcome, and *was* ultimately overcome. ('000 *Prosecution History*, pp. 712-715). As such, Levy (and, by incorporation, the disclosure in Brunk) was before the examiner during the prosecution of the '114 Patent. In that discussion, Levy was considered in combination with Rhoads, over which Murphy is cumulative as discussed above. ('000 *Prosecution History*, at p. 715 (“The prior art was discussed with respect to the inventive concept applicant was trying to capture with the claim language. However, in further discussion with Mr. Wang, the examiner agreed **that the concept the applicant is trying to capture is different from Rhoads and Levy.**”)). For this reason, Brunk’s teachings would have been cumulative to (and incorporated within) the already-of-record teachings of Levy, and the Examiner for the '114 Patent would have been aware of the teachings of Levy and considered them before allowing the claims of the '114 Patent. Thus, applying the *Becton Dickinson* factors, Brunk is cumulative to art already of record and is not relevant to the patentability of the claims of the '114 Patent. Furthermore, Levy (incorporating Brunk) *in combination with Rhoads* (with which Murphy is cumulative, as discussed above) was in front of the Examiner during

prosecution, and the Examiner allowed the claims over the combination. For this reason, Murphy in view of Brunk is cumulative to art already of record and is not relevant to the patentability of the claims of the '114 Patent.

## VIII. CONCLUSION

Because all independent claims recite limitations not taught or suggested by the proposed grounds of unpatentability, Petitioner has not established a reasonable likelihood the Challenged Claims are unpatentable, and Patent Owner respectfully requests non-institution of this IPR.

Respectfully submitted,

/Jennifer C. Bailey/

Jennifer C. Bailey, Reg. No. 52,583

Robin A. Snader, Reg. No. 66,085

Jason R. Mudd, Reg. No. 57,700

Megan J. Redmond (*pro hac vice* to be filed)

ERISE IP, P.A.

7015 College Boulevard, Suite 700

Overland Park, Kansas 66211

Jennifer.Bailey@EriseIP.com

Robin.Snader@EriseIP.com

Jason.Mudd@EriseIP.com

Megan.Redmond@EriseIP.com

PTAB@EriseIP.com

James L. Lovsin, Reg. No. 69,550

MCDONNELL BOEHNEN HULBERT &  
BERGHOFF LLP

300 South Wacker Drive

Chicago, IL 60606

lovsin@mbhb.com

*ATTORNEYS FOR PATENT OWNER*

## APPENDIX OF EXHIBITS

<b>Exhibit 2001</b>	Declaration of Dr. Pierre Moulin
<b>Exhibit 2002</b>	Memorandum Opinion, <i>Gracenote, Inc. v. Free Stream Media Corp.</i> , No 1:18-cv-01608-RGA (D. Del. Dec. 11, 2019), ECF No. 25
<b>Exhibit 2003</b>	Report and Recommendation, <i>Gracenote, Inc. v. Free Stream Media Corp., d/b/a Samba TV</i> , No. 1:18-cv-01608-RGA (D. Del. Nov. 1, 2019), ECF No. 22
<b>Exhibit 2004</b>	U.S. Patent No. 7,185,201 to Rhoads, et al.
<b>Exhibit 2005</b>	Prosecution History for U.S. Patent No. 8,020,000
<b>Exhibit 2006</b>	U.S. Patent No. 8,020,000 Patent to Oostveen, et al.
<b>Exhibit 2007</b>	U.S. Patent Application Publication No. 2003/0021441 to Levy, et al.
<b>Exhibit 2008</b>	Prosecution History for U.S. Patent Application 16/264,134
<b>Exhibit 2009</b>	Prosecution History for U.S. Patent 9,066,114

## CERTIFICATION OF WORD COUNT

The undersigned certifies pursuant to 37 C.F.R. §42.24 that the foregoing Patent Owner's Preliminary Response To Petition For Inter Partes Review, excluding any table of contents, mandatory notices under 37 C.F.R. §42.8, certificates of service or word count, or appendix of exhibits, contains 10,704 words according to the word-processing program used to prepare this document (Microsoft Word).

Dated: March 10, 2020

Respectfully submitted,

/Jennifer C. Bailey/

Jennifer C. Bailey, Reg. No. 52,583

Robin A. Snader, Reg. No. 66,085

Jason R. Mudd, Reg. No. 57,700

Megan J. Redmond (*pro hac vice* to be filed)  
ERISE IP, P.A.

7015 College Boulevard, Suite 700

Overland Park, Kansas 66211

Jennifer.Bailey@EriseIP.com

Robin.Snader@EriseIP.com

Jason.Mudd@EriseIP.com

Megan.Redmond@EriseIP.com

PTAB@EriseIP.com

James L. Lovsin, Reg. No. 69,550

MCDONNELL BOEHLEN HULBERT &  
BERGHOFF LLP

300 South Wacker Drive

Chicago, IL 60606

lovsin@mbhb.com

*ATTORNEYS FOR PATENT OWNER*

## CERTIFICATE OF SERVICE

I hereby certify that a true and correct copy of the foregoing PATENT OWNER'S PRELIMINARY RESPONSE TO PETITION FOR INTER PARTES REVIEW was served on March 10, 2020, by email on the following counsel of record for Petitioner:

Don Daybell at D2DPTABDocket@orrick.com  
Alyssa Caridis at A8CPTABDocket@orrick.com

Dated: March 10, 2020

Respectfully submitted,

/Jennifer C. Bailey/

Jennifer C. Bailey, Reg. No. 52,583  
Robin A. Snader, Reg. No. 66,085  
Jason R. Mudd, Reg. No. 57,700  
Megan J. Redmond (*pro hac vice* to be filed)  
ERISE IP, P.A.  
7015 College Boulevard, Suite 700  
Overland Park, Kansas 66211  
Jennifer.Bailey@EriseIP.com  
Robin.Snader@EriseIP.com  
Jason.Mudd@EriseIP.com  
Megan.Redmond@EriseIP.com  
PTAB@EriseIP.com

James L. Lovsin, Reg. No. 69,550  
MCDONNELL BOEHNEN HULBERT &  
BERGHOFF LLP  
300 South Wacker Drive  
Chicago, IL 60606  
lovsin@mbhb.com

*ATTORNEYS FOR PATENT OWNER*