

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM, INC.,
Petitioner,

v.

JAWBONE INNOVATIONS, LLC,
Patent Owner.

IPR2023-00286
Patent 8,326,611 B2

Before GEORGIANNA W. BRADEN, NORMAN H. BEAMER, and
STEVEN M. AMUNDSON, *Administrative Patent Judges*.

AMUNDSON, *Administrative Patent Judge*.

DECISION
Granting Institution of *Inter Partes* Review
35 U.S.C. § 314
Denying Motion for Joinder as Moot
35 U.S.C. § 315(c); 37 C.F.R. § 42.122

I. INTRODUCTION

Amazon.com, Inc. (“Petitioner”) filed a Petition requesting an *inter partes* review of claims 1–28 in U.S. Patent No. 8,326,611 B2 (Exhibit 1001, “the ’611 patent”) under 35 U.S.C. §§ 311–319. Paper 1 (“Pet.”). Jawbone Innovations, LLC (“Patent Owner”) filed a Preliminary Response. Paper 9 (“Prelim. Resp.”).

Concurrently with the Petition, Petitioner filed a Motion for Joinder seeking joinder as a party in *Apple Inc. v. Jawbone Innovations, LLC*, IPR2022-01085 (“the Apple IPR”) and consolidation with the Apple IPR. Paper 3 (“Mot.”). Patent Owner filed an Opposition to Petitioner’s Motion for Joinder. Paper 7 (“Opp.”). And Petitioner filed a Reply on Petitioner’s Motion for Joinder. Paper 8 (“Reply”).

Under 37 C.F.R. § 42.4(a), we have authority to determine whether to institute an *inter partes* review. We may institute an *inter partes* review only if “the information presented in the petition filed under section 311 and any response filed under section 313 shows that there is a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition.” 35 U.S.C. § 314(a) (2018). The “reasonable likelihood” standard is “a higher standard than mere notice pleading” but “lower than the ‘preponderance’ standard to prevail in a final written decision.” *Hulu, LLC v. Sound View Innovations, LLC*, IPR2018-01039, Paper 29 at 13 (PTAB Dec. 20, 2019) (precedential).

We also have authority to consider Petitioner’s Motion for Joinder. *See* 35 U.S.C. § 315(c); 37 C.F.R. § 42.122 (2022).

After considering the Petition, the Preliminary Response, and the evidence of record, and for the reasons explained below, we determine

Petitioner has shown that there is a reasonable likelihood that it would prevail with respect to at least one of the challenged claims. Thus, we institute an *inter partes* review of claims 1–28 in the '611 patent on all challenges included in the Petition.

Additionally, for the reasons explained below, we deny Petitioner's Motion for Joinder as moot.

II. BACKGROUND

A. Real Parties in Interest

Petitioner identifies the following entities as the real parties in interest: Amazon.com, Inc. and Amazon.com Services LLC (formerly Amazon.com Services, Inc.). Pet. 97. Patent Owner identifies itself as the real party in interest. Paper 5, 2. The parties do not raise any issue about real parties in interest.

B. Related Matters

Petitioner and Patent Owner identify the following civil actions as related matters involving the '611 patent:

- *Jawbone Innovations, LLC v. Samsung Electronics Co., Ltd. et al.*, No. 2:21-cv-00186-JRG (E.D. Tex. filed May 27, 2021);
- *Jawbone Innovations, LLC v. Apple Inc.*, No. 6:21-cv-00984-ADA (W.D. Tex. filed Sept. 23, 2021);
- *Jawbone Innovations, LLC v. Google LLC*, No. 6:21-cv-00985-ADA (W.D. Tex. filed Sept. 23, 2021); and
- *Jawbone Innovations, LLC v. Amazon.com, Inc. et al.*, No. 5:22-cv-06727-NC (N.D. Cal. filed Nov. 29, 2021).

Pet. 97–98; Paper 5, 3.

Petitioner and Patent Owner identify the following Board proceedings as related matters involving the '611 patent:

- *Google LLC v. Jawbone Innovations, LLC*, IPR2022-00604 (PTAB filed Feb. 22, 2022);
- *Google LLC v. Jawbone Innovations, LLC*, IPR2022-00889 (PTAB filed May 16, 2022);
- *Apple Inc. v. Jawbone Innovations, LLC*, IPR2022-01085 (PTAB filed June 3, 2022);
- *Apple Inc. v. Jawbone Innovations, LLC*, IPR2022-01495 (PTAB filed Sept. 2, 2022); and
- *Amazon.com, Inc. v. Jawbone Innovations, LLC*, IPR2023-00285 (PTAB filed Nov. 28, 2022).

Pet. 98; Paper 5, 2.

C. The '611 Patent (Exhibit 1001)

The '611 patent, titled “Acoustic Voice Activity Detection (AVAD) for Electronic Systems,” issued on December 4, 2012, from an application filed on October 26, 2009. Ex. 1001, codes (22), (45), (54). The patent identifies that application as a continuation-in-part of the following applications: (1) U.S. Patent Application No. 11/805,987, filed on May 25, 2007; and (2) U.S. Patent Application No. 12/139,333, filed on June 13, 2008. *Id.* at 1:8–11, code (63). The patent claims priority to U.S. Provisional Patent Application No. 61/108,426, filed on October 24, 2008. *Id.* at 1:6–7, code (60). The patent states that the disclosure relates generally “to noise suppression” and more particularly “to noise suppression systems, devices, and methods for use in acoustic applications.” *Id.* at 1:16–19; *see id.* at code (57).

The '611 patent explains that the “ability to correctly identify voiced and unvoiced speech is critical to many speech applications including speech recognition, speaker verification, noise suppression, and many others.” Ex. 1001, 1:23–26. In “the speaker’s environment,” however, “there may

exist one or more noise sources that pollute the speech signal, the signal of interest, with unwanted acoustic noise.” *Id.* at 1:28–31. The unwanted acoustic noise “makes it difficult or impossible for the receiver, whether human or machine, to understand the user’s speech.” *Id.* at 1:31–32.

The ’611 patent also explains that “[t]ypical methods for classifying voiced and unvoiced speech have relied mainly on the acoustic content of single microphone data, which is plagued by problems with noise and the corresponding uncertainties in signal content.” Ex. 1001, 1:33–36.

According to the patent, non-acoustic methods “have been employed successfully in commercial products,” but “an acoustic-only solution is desired in some cases (e.g., for reduced cost, as a supplement to the non-acoustic sensor, etc.).” *Id.* at 1:41–46.

Toward that end, the ’611 patent discloses “Acoustic Voice Activity Detection (AVAD) methods and systems” that use physical microphones “to generate virtual directional microphones which have very similar noise responses and very dissimilar speech responses.” Ex. 1001, 3:57–62, code (57); *see id.* at 4:21–25, 5:35–52, 17:43–45. “The ratio of the energies of the virtual microphones is then calculated over a given window size and the ratio can then be used with a variety of methods to generate a VAD signal.” *Id.* at 3:62–65, code (57). “The virtual microphones can be constructed using either a fixed or an adaptive filter.” *Id.* at 3:65–66, code (57); *see id.* at 4:27–30.

The ’611 patent defines various terms. Ex. 1001, 17:57–18:32. The patent includes the following definitions:

- The term “speech” means desired speech of the user.

- The term “noise” means unwanted environmental acoustic noise.
- The term “O₁” means a first physical omnidirectional microphone used to form a microphone array.
- The term “O₂” means a second physical omnidirectional microphone used to form a microphone array.
- The term “virtual microphones (VM)” or “virtual directional microphones” means a microphone constructed using two or more omnidirectional microphones and associated signal processing.
- The term “V₁” means the virtual directional “speech” microphone, which has no nulls.
- The term “V₂” means the virtual directional “noise” microphone, which has a null for the user’s speech.

Id. at 18:10–32.

The ’611 patent explains that “V₂ is configured in such a way that it has minimal response to the speech of the user” and “V₁ is configured so that it does respond to the user’s speech but has a very similar noise magnitude response to V₂.” Ex. 1001, 4:21–25. “A further refinement is the use of an adaptive filter to further minimize the speech response of V₂.” *Id.* at 4:27–28; *see id.* at 6:10–11.

The '611 patent's Figure 3 (reproduced below) depicts a block diagram of a virtual microphone V_2 :

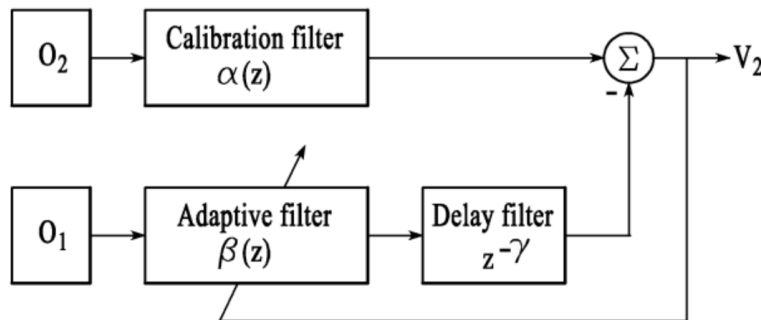


FIG.3

Figure 3 illustrates the following components where “z” denotes a discrete frequency domain and “ γ ” is a fixed delay that depends on the size of the microphone array:

- (1) a physical omnidirectional microphone O_1 providing a signal to an adaptive filter $\beta(z)$;
- (2) the adaptive filter $\beta(z)$ providing a signal to a delay filter $z^{-\gamma}$;
- (3) a physical omnidirectional microphone O_2 providing a signal to a calibration filter $\alpha(z)$;
- (4) a summing circuit (Σ) summing (i) a positive signal from the calibration filter $\alpha(z)$ and (ii) a negative signal from the delay filter $z^{-\gamma}$ to form a virtual microphone V_2 ; and
- (5) the virtual microphone V_2 providing a signal to the adaptive filter $\beta(z)$.

See Ex. 1001, 1:62–63, 4:41, 5:22–33, 6:1–16, Fig. 3.

The '611 patent's Figure 4 (reproduced below) depicts a block diagram of a virtual microphone V_1 :

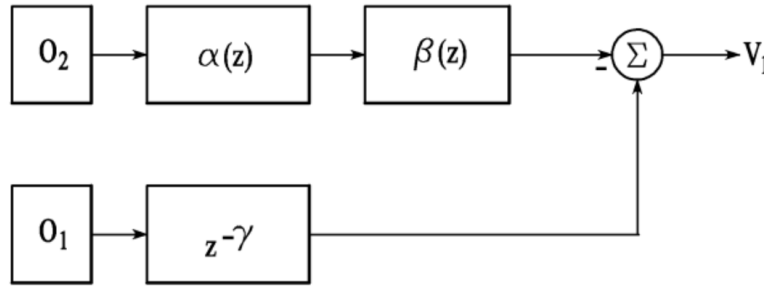


FIG.4

Figure 4 illustrates the following components where “z” denotes a discrete frequency domain and “ γ ” is a fixed delay that depends on the size of the microphone array:

- (1) a physical omnidirectional microphone O_1 providing a signal to a delay filter $z^{-\gamma}$;
- (2) a physical omnidirectional microphone O_2 providing a signal to a calibration filter $\alpha(z)$;
- (3) the calibration filter $\alpha(z)$ providing a signal to a filter $\beta(z)$; and
- (4) a summing circuit (Σ) summing (i) a positive signal from the delay filter $z^{-\gamma}$ and (ii) a negative signal from the filter $\beta(z)$ to form a virtual microphone V_1 .

See Ex. 1001, 1:64–65, 4:41, 5:22–33, 6:17–19, Fig. 4.

The '611 patent provides the following equations for virtual microphones V_1 and V_2 constructed using physical omnidirectional microphones O_1 and O_2 :

$$V_1(z) = -\beta(z)\alpha(z)O_2(z) + O_1(z)z^{-\gamma}$$

$$V_2(z) = \alpha(z)O_2(z) - \beta(z)O_1(z)z^{-\gamma}$$

Ex. 1001, 5:20–28. The patent explains that in these equations (1) the filter $\alpha(z)$ “compensate[s] O_2 ’s response so that it is the same as O_1 ” and (2) the filter $\beta(z)$ “describes the relationship between O_1 and calibrated O_2 for speech.” *Id.* at 5:29–32; *see id.* at 23:11–41. The patent also explains that by “varying the magnitude and sign of the delays and gains, a wide variety of virtual microphones (VMs), also referred to herein as virtual directional microphones, can be realized.” *Id.* at 21:7–10.

The ’611 patent discloses an adaptive filter $\beta(z)$ that “minimize[s] the output of V_2 when only speech is being received by O_1 and O_2 ” and preferably has the following construction:

$$\beta(z) = \alpha(z)O_2(z) / z^{-\gamma}O_1(z)$$

Ex. 1001, 6:1–11, 6:35–38. “Any adaptive process may be used,” such as “a normalized least-mean squares (NLMS) algorithm.” *Id.* at 6:14–15.

With $\beta(z)$ suitably selected, “the ratio for speech should be relatively high (e.g., greater than approximately 2) and the ratio for noise should be relatively low (e.g., less than approximately 1.1).” Ex. 1001, 6:29–32. “The ratio calculated will depend on both the relative energies of the speech and noise as well as the orientation of the noise and the reverberance of the environment.” *Id.* at 6:32–35.

The '611 patent's Figure 5 (reproduced below) depicts a flow diagram of acoustic voice activity detection:

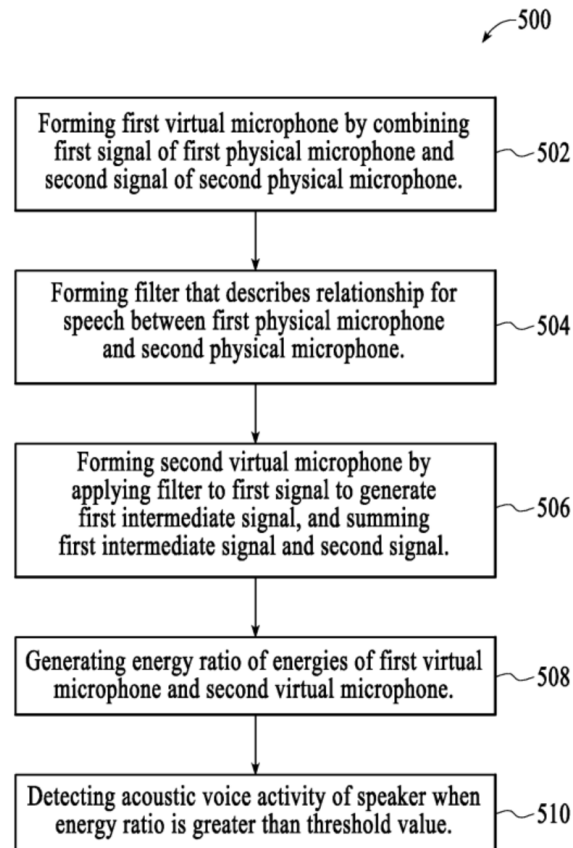


FIG.5

Figure 5 illustrates “acoustic voice activity detection 500” including steps 502 through 510. Ex. 1001, 1:66–67, 6:60–61, Fig. 5.

At step 502, the “detection comprises forming a first virtual microphone by combining a first signal of a first physical microphone and a second signal of a second physical microphone.” Ex. 1001, 6:61–64, Fig. 5. At step 504, the “detection comprises forming a filter that describes a relationship for speech between the first physical microphone and the second physical microphone.” *Id.* at 6:64–67, Fig. 5. At step 506, the “detection comprises forming a second virtual microphone by applying the filter to the

first signal to generate a first intermediate signal, and summing the first intermediate signal and the second signal.” *Id.* at 6:67–7:3, Fig. 5. At step 508, the “detection comprises generating an energy ratio of energies of the first virtual microphone and the second virtual microphone.” *Id.* at 7:3–5, Fig. 5. At step 510, the “detection comprises detecting acoustic voice activity of a speaker when the energy ratio is greater than a threshold value.” *Id.* at 7:5–7, Fig. 5.

D. The Challenged Claims

Petitioner challenges independent method claim 1 and claims 2–28 that depend directly or indirectly from claim 1. Pet. 1, 7–91.

Claim 1 exemplifies the challenged claims and reads as follows (with numbers and letters added for reference purposes):¹

1. [1pre] A method comprising:

[1a] forming a first virtual microphone by combining a first signal of a first physical microphone and a second signal of a second physical microphone;

[1b] forming a filter that describes a relationship for speech between the first physical microphone and the second physical microphone;

[1c] forming a second virtual microphone by applying the filter to the first signal to generate a first intermediate signal, and summing the first intermediate signal and the second signal;

[1d] generating an energy ratio of energies of the first virtual microphone and the second virtual microphone; and

[1e] detecting acoustic voice activity of a speaker when the energy ratio is greater than a threshold value.

¹ We use the same numbers and letters that Petitioner uses to identify the claim language. *See* Pet. List of Challenged Claims 1. These annotations do not impact our analysis.

Ex. 1001, 36:61–37:8.

E. The Asserted References

For its challenges, Petitioner relies on the following references:

Name	Reference	Exhibit
Avendano	US 8,194,880 B2, issued June 5, 2012 (based on an application filed January 29, 2007)	1005
Visser	US 7,464,029 B2, issued December 9, 2008 (based on an application filed July 22, 2005)	1006
Hou	US 7,155,019 B2, issued December 26, 2006 (based on an application filed March 14, 2001)	1008
Byrne	Denis Byrne et al., <i>An International Comparison of Long-Term Average Speech Spectra</i> , The Journal of the Acoustical Society of America, Vol. 96, No. 4, 2108–20 (October 1994)	1009
Bisgaard	US 2011/0103626 A1, published May 5, 2011 (based on an application filed June 25, 2007)	1010
Burnett	US 2002/0198705 A1, published December 26, 2002 (based on an application filed May 30, 2002)	1012
Berglund	Birgitta Berglund et al., <i>Sources and Effects of Low-Frequency Noise</i> , The Journal of the Acoustical Society of America, Vol. 99, No. 5, 2985–3002 (May 1996)	1013

Pet. 1–2, 7–91.

Petitioner asserts that:

- (1) Byrne, Burnett, and Berglund qualify as prior art under § 102(b); and
- (2) Avendano, Visser, Hou, and Bisgaard qualify as prior art under § 102(e).

Pet. 2; *see* 35 U.S.C. § 102(b), (e) (2006).²

At this stage of the proceeding, Patent Owner does not dispute that each reference qualifies as prior art. *See, e.g.*, Prelim. Resp. 5–13.

F. The Asserted Challenges to Patentability

Petitioner asserts the following challenges to patentability:

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
1–7, 25–28	103(a)	Avendano, Visser
8–16, 23, 24	103(a)	Avendano, Visser, Bisgaard
17–19	103(a)	Avendano, Visser, Bisgaard, Hou
20–22	103(a)	Avendano, Visser, Bisgaard, Hou, Byrne
20–22	103(a)	Avendano, Visser, Bisgaard, Hou, Burnett
20–22	103(a)	Avendano, Visser, Bisgaard, Hou, Berglund

Pet. 1, 7–91.

G. Testimonial Evidence

To support its challenges, Petitioner relies on the declaration of Thomas W. Kenny, Ph.D. (Exhibit 1003). Dr. Kenny states, “I received a B.S. in Physics from University of Minnesota, Minneapolis in 1983, and a Ph.D. in Physics from University of California at Berkeley in 1989” and “was educated as a Physicist specializing in sensors and measurement.”

² The Leahy-Smith America Invents Act (“AIA”), Pub. L. No. 112-29, 125 Stat. 284 (2011), amended 35 U.S.C. § 102 and § 103 effective March 16, 2013. Because the filing date of the challenged claims predates the AIA’s amendments to § 102 and § 103, this decision refers to the pre-AIA versions of § 102 and § 103.

Ex. 1003 ¶ 3. Dr. Kenny also states, “I have been retained” to “offer technical opinions” about the ’611 patent and “prior art references relating to its subject matter.” *Id.* ¶ 16.

Petitioner also relies on the declaration of June Ann Munford (Exhibit 1014). Ms. Munford states, “I earned a Master of Library and Information Science (MLIS) from the University of Wisconsin-Milwaukee in 2009” and “have over ten years of experience in the library/information science field.” Ex. 1014 ¶ 2. Ms. Munford also states, “I have been retained” to “provide assistance” in “establishing the authenticity and public availability of” certain documents, including Byrne (Exhibit 1009) and Berglund (Exhibit 1013). *Id.* ¶ 12; *see id.* ¶¶ 6–11.

III. MOTION FOR JOINDER

In November 2022, concurrently with the Petition, Petitioner filed a Motion for Joinder seeking joinder as a party in the Apple IPR and consolidation with the Apple IPR. Mot. 1.

In December 2022, Patent Owner filed an Opposition to Petitioner’s Motion for Joinder, requesting that the Board deny the joinder motion because “Apple and Patent Owner have recently settled their dispute and will seek to file a motion to terminate the Apple IPR.” Opp. 1.

In January 2023, the Board terminated the Apple IPR in view of representations by Apple and Patent Owner that they had “reached settlements as to all the disputes in this proceeding and as to the ’611 patent.” IPR2022-01085, Paper 12 at 2; Paper 14 at 2–3.

Later in January 2023, Petitioner filed a Reply on Petitioner’s Motion for Joinder, stating that the “joinder motion should be deemed withdrawn or denied as moot.” Reply 1.

Accordingly, we deny Petitioner’s Motion for Joinder as moot.

IV. PATENTABILITY ANALYSIS

A. Legal Principles: Obviousness

A patent may not be obtained “if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.” 35 U.S.C. § 103(a) (2006). An obviousness analysis involves underlying factual inquiries including (1) the scope and content of the prior art; (2) differences between the claimed invention and the prior art; (3) the level of ordinary skill in the art; and (4) where in evidence, objective indicia of nonobviousness, such as commercial success, long-felt but unsolved needs, and failure of others.³ *Graham v. John Deere Co.*, 383 U.S. 1, 17–18, 35–36 (1966); *Apple Inc. v. Samsung Elecs. Co.*, 839 F.3d 1034, 1047–48 (Fed. Cir. 2016) (en banc). When evaluating a combination of references, an obviousness analysis should address “whether there was an apparent reason to combine the known elements in the fashion claimed by the patent at issue.” *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007).

We analyze the obviousness issues according to these principles.

B. Level of Ordinary Skill in the Art

Factors pertinent to determining the level of ordinary skill in the art include (1) the educational level of the inventor; (2) the type of problems encountered in the art; (3) prior-art solutions to those problems; (4) the rapidity with which innovations are made; (5) the sophistication of the

³ The record does not include evidence or argument regarding objective indicia of nonobviousness.

technology; and (6) the educational level of workers active in the field. *Envtl. Designs, Ltd. v. Union Oil Co.*, 713 F.2d 693, 696–97 (Fed. Cir. 1983). Not all factors may exist in every case, and one or more of these or other factors may predominate in a particular case. *Id.* These factors are not exhaustive, but merely a guide to determining the level of ordinary skill in the art. *Daiichi Sankyo Co. v. Apotex, Inc.*, 501 F.3d 1254, 1256 (Fed. Cir. 2007). Moreover, the prior art itself may reflect an appropriate skill level. *Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001).

Petitioner asserts that a person of ordinary skill in the art at the time of the alleged invention would have had “at least a bachelor of science in electrical engineering, computer engineering, computer science, mechanical engineering, or a related discipline, with at least two years of relevant experience in a field related to acoustics, speech recognition, speech detection, or signal processing.” Pet. 6. Petitioner also asserts that “[a]dditional education or industry experience may compensate for a deficit in the other.” *Id.* Dr. Kenny’s testimony supports Petitioner’s assertions. *See* Ex. 1003 ¶¶ 22–23.

For purposes of the Preliminary Response, Patent Owner “utilizes Petitioner’s proposed level of skill in the art.” Prelim. Resp. 5 (quoting Pet. 6).

Based on the current record and for purposes of analysis, we accept Petitioner’s description of an ordinarily skilled artisan as consistent with the ’611 patent and the asserted prior art.

C. Claim Construction

We construe claim terms “using the same claim construction standard” that district courts use to construe claim terms in civil actions

under 35 U.S.C. § 282(b). *See* 37 C.F.R. § 42.100(b). Under that standard, claim terms “are given their ordinary and customary meaning, which is the meaning the term would have to a person of ordinary skill in the art at the time of the invention.” *Power Integrations, Inc. v. Fairchild Semiconductor Int’l, Inc.*, 904 F.3d 965, 971 (Fed. Cir. 2018) (citing *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–13 (Fed. Cir. 2005) (en banc)). The meaning of claim terms may be determined by “look[ing] principally to the intrinsic evidence of record, examining the claim language itself, the written description, and the prosecution history, if in evidence.” *DePuy Spine, Inc. v. Medtronic Sofamor Danek, Inc.*, 469 F.3d 1005, 1014 (Fed. Cir. 2006) (citing *Phillips*, 415 F.3d at 1312–17).

Petitioner contends that “[n]o formal claim constructions are necessary.” Pet. 6.

Patent Owner “believes that claim construction is not required to resolve any issues.” Prelim. Resp. 5.

Based on the current record, we determine that no claim term requires an explicit construction to decide whether Petitioner satisfies the “reasonable likelihood” standard for instituting trial. “[O]nly those terms need be construed that are in controversy, and only to the extent necessary to resolve the controversy.” *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999); *see Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017).

*D. Alleged Obviousness over Avendano
and Visser: Claims 1–7 and 25–28*

Petitioner contends that claims 1–7 and 25–28 are unpatentable under § 103(a) as obvious over Avendano and Visser. *See* Pet. 1, 7–42. Patent

Owner disputes Petitioner's contentions. *See* Prelim. Resp. 5–12. Below, we provide overviews of Avendano and Visser, and then we consider the obviousness issues. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano and Visser teach the subject matter in claims 1–7 and 25–28.

1. OVERVIEW OF AVENDANO (EXHIBIT 1005)

Avendano is a U.S. patent titled “System and Method for Utilizing Omni-Directional Microphones for Speech Enhancement,” filed on January 29, 2007, and issued on June 5, 2012. Ex. 1005, codes (12), (22), (45), (54). Avendano states that the invention “relates generally to audio processing.” *Id.* at 1:24; *see id.* at code (57).

Avendano discloses systems and methods “for utilizing inter-microphone level differences (ILD) to attenuate noise and enhance speech.” Ex. 1005, 2:5–7, 3:13–17, code (57). Preferably, “the ILD is based on energy level differences of a pair of omni-directional microphones.” *Id.* at 2:7–9, code (57). For example, “a primary acoustic signal is received by a primary microphone, and a secondary acoustic signal is received by a secondary microphone.” *Id.* at 2:14–17, code (57). “The primary and secondary acoustic signals are converted into primary and secondary electric signals for processing.” *Id.* at 2:17–19, code (57).

“A differential microphone array (DMA) module processes the primary and secondary electric signals to determine a cardioid primary signal and a cardioid secondary signal.” Ex. 1005, 2:20–22, code (57). Preferably, “the primary and secondary electric signals are delayed by a delay node.” *Id.* at 2:22–24. “The cardioid primary signal is then determined by taking a difference between the primary electric signal and

the delayed secondary electric signal, while the cardioid secondary signal is determined by taking a difference between the secondary electric signal and the delayed primary electric signal.” *Id.* at 2:24–29.

“The cardioid signals are filtered through a frequency analysis module” that “takes the signals and mimics the frequency analysis of the cochlea (i.e., cochlear domain).” Ex. 1005, 2:34–36, code (57). “Energy levels associated with the cardioid primary signal and the cardioid secondary signals are then computed (e.g., as power estimates) and the results are processed by an ILD module” to obtain the ILD. *Id.* at 2:41–45, code (57). The ILD module may obtain the ILD by “using a non-linear combination,” e.g., by “dividing the power estimate associated with the cardioid primary signal by the power estimate associated with the cardioid secondary signal.” *Id.* at 2:44–48, code (57). The ILD “may then be used” to (1) “suppress unwanted sound sources” and (2) “enhance the speech of the primary acoustic signal.” *Id.* at 2:48–50, code (57).

Avendano's Figure 1a (reproduced below) depicts an environment for practicing embodiments of Avendano's invention:

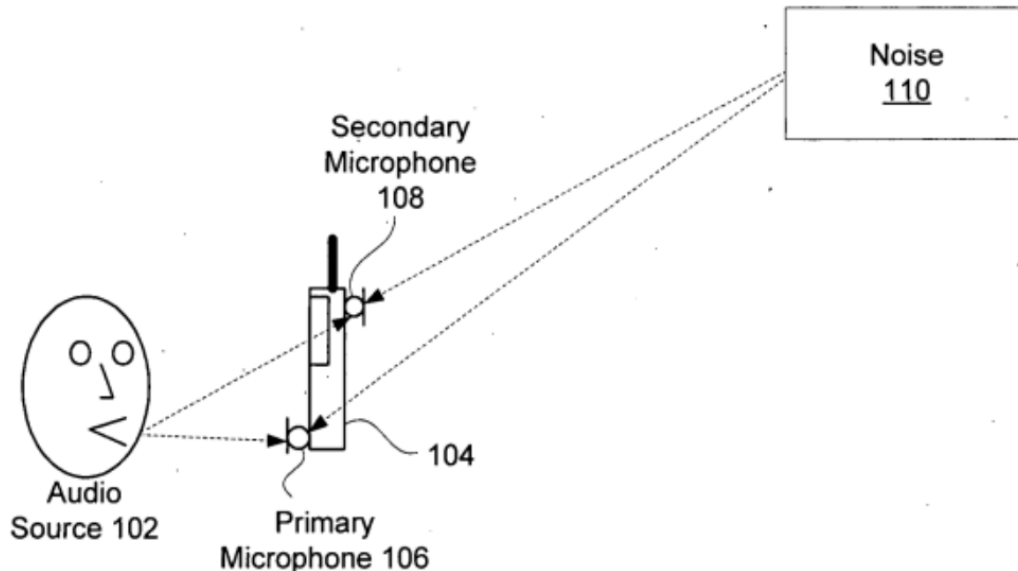


FIG. 1a

Figure 1a illustrates a user who “provides an audio (speech) source 102 to an audio device 104” comprising two microphones: “a primary microphone 106 relative to the audio source 102 and a secondary microphone 108 located a distance, d , away from the primary microphone 106.” Ex. 1005, 3:27–33, Fig. 1a. Preferably, “the microphones 106 and 108 are omni-directional microphones.” *Id.* at 3:33–35.

“While the microphones 106 and 108 receive sound (i.e., acoustic signals) from the audio source 102, the microphones 106 and 108 also pick up noise 110.” Ex. 1005, 3:36–38, Fig. 1a. Noise 110 “may comprise any sounds from one or more locations different than the audio source 102” and “may include reverberations and echoes.” *Id.* at 3:38–42.

Avendano's Figure 2 (reproduced below) depicts an exemplary audio device:

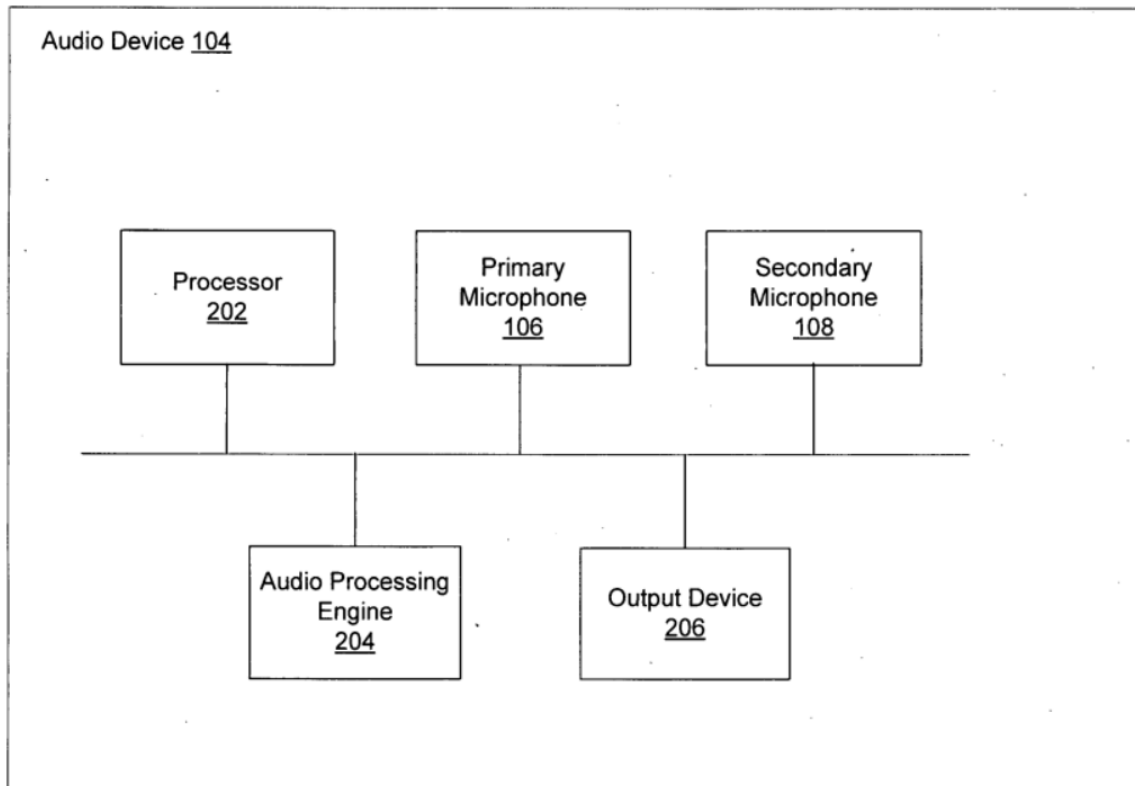


FIG. 2

Figure 2 illustrates audio device 104 including primary microphone 106, secondary microphone 108, processor 202, audio processing engine 204, and output device 206. Ex. 1005, 3:61–66, Fig. 2; *see id.* at 2:57–58. To permit processing by audio processing engine 204, an acoustic signal received by primary microphone 106 (a primary acoustic signal X_1) is converted into a primary electric signal, and an acoustic signal received by secondary microphone 108 (a secondary acoustic signal X_2) is converted into a secondary electric signal. *Id.* at 4:6–8, 4:23–27. Audio processing engine 204 “is embodied within a memory device.” *Id.* at 4:22–23.

Avendano's Figure 4a (reproduced below) depicts an embodiment of audio processing engine 204:

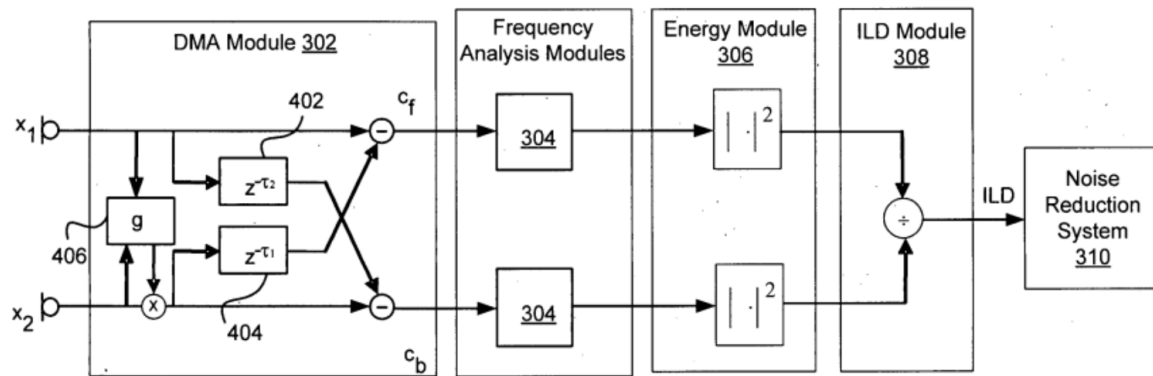


FIG. 4a

Figure 4a illustrates one implementation of audio processing engine 204 including differential microphone array (DMA) module 302, frequency-analysis module 304, energy module 306, ILD module 308, and noise-reduction system 310. Ex. 1005, 2:61–63, 5:15–17, 6:36–40, Fig. 4a; *see id.* at 2:20. As Figure 4a shows, DMA module 302 receives an electric signal x_1 corresponding to the primary acoustic signal X_1 and an electric signal x_2 corresponding to the secondary acoustic signal X_2 . *See id.* at 4:6–8, 4:23–27, 5:17–20, Fig. 4a.

After receiving the electric signals x_1 and x_2 , DMA module 302 determines “sounds and signals in a front and back cardioid region about” audio device 104 by “delaying and subtracting the acoustic signals captured by the microphones 106 and 108.” Ex. 1005, 4:29–33. By doing so, DMA module 302 creates “directional patterns” for the microphones. *Id.* at 4:27–29; *see id.* at 4:40–46.

For example, “one directional pattern created by the DMA module 302 allows sounds generated from an audio source 102 within a

front cardioid region around the audio device 104 to be received,” and “a second pattern allows sounds from a second audio source 102 within a back cardioid region around the audio device 104 to be received.” Ex. 1005, 4:47–52. “Sounds from audio sources 102 beyond these regions may also be received but the sounds may be attenuated.” *Id.* at 4:52–54.

As Figure 4a shows, DMA module 302 delays the electric signal x_1 corresponding to the primary acoustic signal X_1 via delay node 402 with delay $z^{-\tau_2}$, and DMA module 302 delays a gain-adjusted version of the electric signal x_2 corresponding to the secondary acoustic signal X_2 via delay node 404 with delay $z^{-\tau_1}$.⁴ Ex. 1005, 5:20–24, 5:36–37, 8:52–60, Fig. 4a.

DMA module 302 generates a cardioid primary signal C_f and a cardioid secondary signal C_b according to the following equations for C_f and C_b :

$$C_f = X_1 - z^{-\tau_1}gX_2 \qquad C_b = gX_2 - z^{-\tau_2}X_1$$

where the gain factor “g” is “computed by the gain module 406 to equalize the signal levels.” *Id.* at 5:25–37; *see id.* at 8:52–60. The gain factor “g” equalizes the signal levels to prevent a “loss of performance when the microphone signals have different levels.” *Id.* at 5:37–39.

⁴ In contrast to Figure 4a, Avendano’s written description identifies (1) the delay node for the electric signal x_1 corresponding to the primary acoustic signal X_1 as delay node 404 with delay $z^{-\tau_1}$ and (2) the delay node for the electric signal x_2 corresponding to the secondary acoustic signal X_2 as delay node 404 with delay $z^{-\tau_2}$. Ex. 1005, 5:20–24. But the written description includes equations that use the delays $z^{-\tau_1}$ and $z^{-\tau_2}$ consistently with Figure 4a, i.e., delay node 402 with delay $z^{-\tau_2}$ applied to the electric signal x_1 and delay node 404 with delay $z^{-\tau_1}$ applied to the electric signal x_2 after gain adjustment. *Id.* at 5:25–35, Fig. 4a.

Frequency-analysis module 304 “takes the cardioid signals and mimics the frequency analysis of the cochlea (i.e., cochlear domain).” Ex. 1005, 4:55–59; *see id.* at 5:41–48, code (57). To simulate the cochlear domain, frequency-analysis module 304 may employ a filter bank or similar mechanism, e.g., a short-time Fourier transform (STFT), sub-band filter banks, or modulated complex lapped transforms. *Id.* at 2:34–41, 4:55–64, 9:67–10:3. “Because most sounds (e.g., acoustic signals) are complex and comprise more than one frequency, a sub-band analysis on the acoustic signal determines what individual frequencies are present in the complex acoustic signal during a frame (e.g., a predetermined period of time).” *Id.* at 4:64–5:1.

Next, energy module 306 “takes the signals from the frequency analysis module 304 and calculates the power estimates associated with the cardioid primary signal (C_f) and the cardioid secondary signal (C_b).” Ex. 1005, 5:49–52. For example, “the power estimates may be mathematically determined by squaring and integrating an absolute value of the output of” frequency-analysis module 304. *Id.* at 5:52–55.

Avendano discloses the following equations for calculating the energy level $E_f(t, \omega)$ associated with the primary microphone signal and the energy level $E_b(t, \omega)$ associated with the secondary microphone signal by squaring and integrating:

$$E_f(t, \omega) = \int_{\text{frame}} |C_f(t', \omega)|^2 dt' \qquad E_b(t, \omega) = \int_{\text{frame}} |C_b(t', \omega)|^2 dt'$$

Ex. 1005, 5:58–6:7.

Energy module 306 supplies the calculated energy levels $E_f(t, \omega)$ and $E_b(t, \omega)$ to ILD module 308. Ex. 1005, 6:9–10, Fig. 4a. Then, ILD

module 308 determines the ILD “in a non-linear manner by taking a ratio of the energy levels” according to the following equation:

$$\text{ILD}(t,\omega) = E_f(t,\omega)/E_b(t,\omega)$$

Id. at 6:10–13.

“Once the ILD is determined, the signals are processed through a noise reduction system 310.” Ex. 1005, 6:35–36, Fig. 4a. Noise reduction system 310 may include “a noise estimate module 312, a filter module 314, a filter smoothing module 316, a masking module 318, and a frequency synthesis module 320.” *Id.* at 6:36–40, Fig. 3. Noise estimate module 312 may calculate a noise estimate using, among other things, the value $\lambda_1(t,\omega)$ determined as follows:

$$\lambda_1(t,\omega) \approx 0 \text{ if } \text{ILD}(t,\omega) < \text{threshold}$$

$$\lambda_1(t,\omega) \approx 1 \text{ if } \text{ILD}(t,\omega) > \text{threshold}$$

Id. at 6:48–7:6.

Avendano's Figure 6 (reproduced below) depicts “a polar plot of a front-to-back cardioid directivity pattern and ILD diagram”:

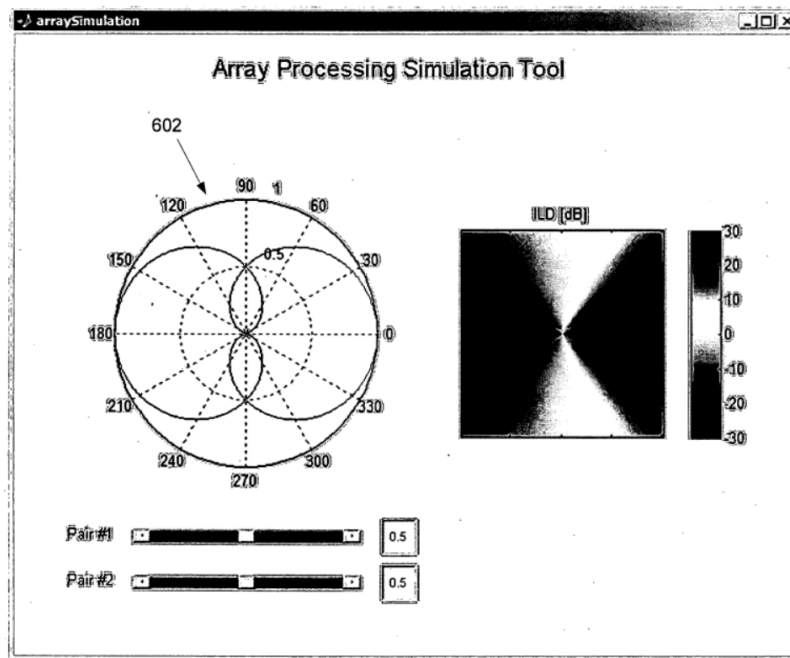


FIG. 6

Figure 6 illustrates “front-to-back cardioid directivity pattern 602” with the directivity “extended in the forward and backward directions (i.e., along the x-axis).” Ex. 1005, 9:29–37, Fig. 6; *see id.* at 3:1–3. “The extension in the forward and backward directions allows significant ILD cues to be obtained from acoustic sources further away from the microphones 106 and 108.” *Id.* at 9:37–39. “As a result, the omni-directional microphones 106 and 108 can achieve acoustic characteristics that mimic those of directional microphones.” *Id.* at 9:39–42.

2. OVERVIEW OF VISSER (EXHIBIT 1006)

Visser is a U.S. patent titled “Robust Separation of Speech Signals in a Noisy Environment,” filed on July 22, 2005, and issued on December 9, 2008. Ex. 1006, codes (12), (22), (45), (54). Visser states that the invention

“relates to processes and methods for separating a speech signal from a noisy acoustic environment.” *Id.* at 1:19–20; *see id.* at 6:57–59, code (57).

Visser discloses methods for separating speech signals from noise sources, including independent component analysis (ICA) and blind signal (or source) separation (BSS). *See, e.g.*, Ex. 1006, 3:3–20, 8:16–19, 14:33–40, 16:6–27, 18:1–4, 18:26–29. As an example of a method for separating speech signals from noise sources, “a signal separation process is associated with a voice activity detector,” e.g., “a two-channel detector.” *Id.* at 6:59–61, code (57). “When speech is detected, the voice activity detector generates a control signal.” *Id.* at 6:63–64, code (57). “The control signal is used to activate, adjust, or control signal separation processes or post-processing operations to improve the quality of the resulting speech signal.” *Id.* at 6:64–67, code (57).

As another example of a method for separating speech signals from noise sources, “a signal separation process is provided as a learning stage and an output stage.” Ex. 1006, 6:67–7:1, code (57). “The learning stage aggressively adjusts to current acoustic conditions, and passes coefficients to the output stage.” *Id.* at 7:1–3, code (57). “The output stage adapts more slowly, and generates a speech-content signal and a noise dominant signal.” *Id.* at 7:3–5, code (57).

Visser explains that a speech-separation process “may be adaptive and learn according to the specific acoustic environment” and “may also adapt to particular microphone placement, the acoustic environment, or a particular user’s speech.” Ex. 1006, 9:8–12. For example, a speech-separation process may include “a set of filters with adaptive filter coefficients.” *Id.* at 16:62–65. As the process operates, “the coefficients are adjusted to

improve separation performance,” and “the new coefficients are applied and used” in the filters. *Id.* at 16:65–17:2, Fig. 9. The “continual adaptation of the filter coefficients enables the process” to “provide a sufficient level of separation, even in a changing acoustic environment.” *Id.* at 17:2–4.

Visser’s Figure 1 (reproduced below) depicts a speech-separation process:

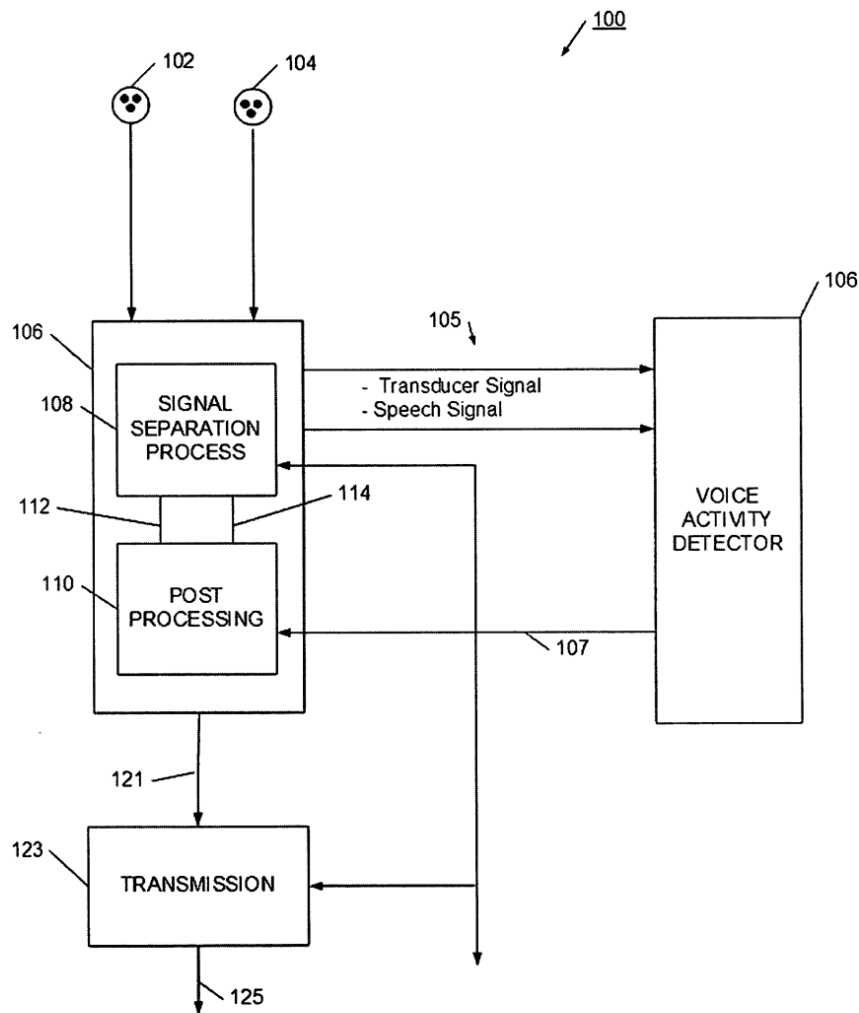


FIG. 1

Figure 1 illustrates speech-separation process 100 receiving “a set of signal inputs (e.g., sound signals from microphones) 102 and 104.” Ex. 1006, 7:36–37, 8:4–7, Fig. 1. For example, “signal input 102 may be from a

microphone arranged to be closest to the speaker's mouth, while signal input 104 may be from a microphone spaced farther away from the speaker's mouth." *Id.* at 8:8–12.

Speech-separation process 100 includes separation process 106 with signal-separation process 108 and post-processing steps 110. Ex. 1006, 8:16–26, Fig. 1. After “the microphones generate a pair of input signals” to signal-separation process 108, process 108 “generates a signal having speech content 112, and a noise-dominant signal 114.” *Id.* at 8:19–23, Fig. 1. Then, post-processing steps 110 “accept these signals, and further reduce the noise to generate an output speech signal 121, which may be transmitted 125 by transmission subsystem 123.” *Id.* at 8:23–26, Fig. 1.

Speech-separation process 100 also includes voice activity detector (VAD) 106 “to activate, adjust, or control selected signal separation, post processing, or transmission functions.” Ex. 1006, 8:27–30, Fig. 1. Upon detecting speech, VAD 106 “generates a control signal 107” that “may be used, for example, to activate the signal separation process only when speech is occurring, thereby increasing stability and saving power.” *Id.* at 8:36–40, Fig. 1. Moreover, post-processing steps 110 “may be controlled to more accurately characterize noise, as the characterization process may be limited to times when no speech is occurring.” *Id.* at 8:40–43.

Visser's Figure 10 (reproduced below) depicts a sub-module for an ICA or BSS processing function:

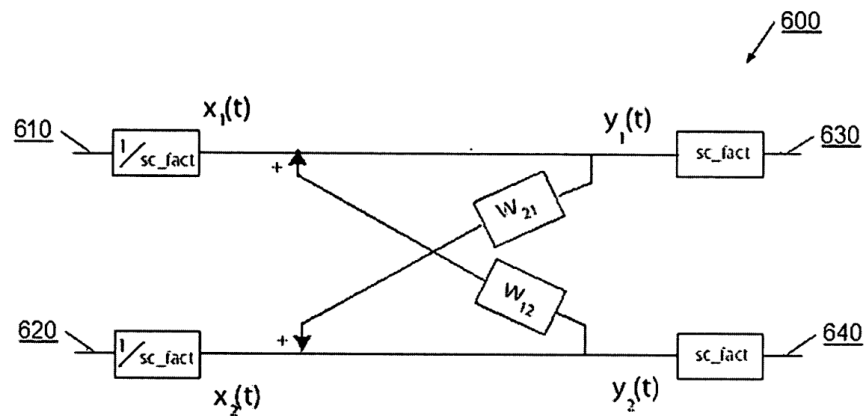


FIG. 10

Figure 10 illustrates processing sub-module 600 receiving input signal X_1 on channel 610 and input signal X_2 on channel 620. Ex. 1006, 7:55–57, 17:29–37, Fig. 10. “Typically, each of these signals would come from at least one microphone.” *Id.* at 17:37–38.

Sub-module 600 includes cross filters w_{12} and w_{21} . Ex. 1006, 17:39–41, 17:56–64, Fig. 10. Cross filters w_{12} and w_{21} facilitate separating speech signals from noise sources. *See id.* at 16:8–27, 17:56–64. Cross filters w_{12} and w_{21} may have adaptive filter coefficients or fixed filter coefficients. *Id.* at 16:62–65, 17:56–64; *see id.* at 4:6–10, 22:66–67, 23:13–25. For instance, in “a most simplified form,” cross filters w_{12} and w_{21} constitute “gain factors with only one filter coefficient per filter,” e.g., “a delay gain factor for the time delay between the output signal and the feedback input signal and an amplitude gain factor for amplifying the input signal.” *Id.* at 17:58–62.

Cross filters w_{12} and w_{21} “are applied to each of the input signals to produce” the following:

- (1) a channel 630 (speech channel) of separated signals U_1 that “contains predominantly desired signals”; and
- (2) a channel 640 (noise channel) of separated signals U_2 that “contains predominantly noise signals.”

Ex. 1006, 17:39–44, Fig. 10.

The “output signals U_1 and U_2 can be further processed by a post processing sub-module, a de-noising module or a speech feature extraction module.” Ex. 1006, 17:64–67. For example, a voice activity detector may use the output signals U_1 and U_2 to determine “when the difference in energy levels between the signals exceeds a predefined threshold.” *See id.* at 8:33–35, 12:11–18.

3. INDEPENDENT CLAIM 1

(a) Preamble [1pre]

Claim 1 recites “[a] method comprising.” Ex. 1001, 36:61.

Petitioner contends that Avendano discloses claim 1’s preamble because Avendano discloses “an exemplary method for utilizing ILD of omni-direction microphones for noise suppression and speech enhancement.” Pet. 22 (quoting Ex. 1005, 9:43–45) (citing Ex. 1003 ¶ 126; Ex. 1005, 3:12–7:9, 9:43–10:25, Figs. 1a–1b, 2–3, 4a–4b, 5–7).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1’s preamble. *See, e.g.*, Prelim. Resp. 5–12. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware, LLC v. Nat’l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015).

Generally, a preamble does not limit a claim. *Allen Eng’g Corp. v. Bartell Indus., Inc.*, 299 F.3d 1336, 1346 (Fed. Cir. 2002). We need not decide whether claim 1’s preamble limits the claim because Petitioner

establishes sufficiently for purposes of institution that Avendano discloses claim 1's preamble. *See* Pet. 22; Ex. 1003 ¶ 126. Specifically, Avendano discloses methods “for utilizing inter-microphone level differences (ILD) to attenuate noise and enhance speech.” *See, e.g.*, Ex. 1005, 2:5–7, 3:13–17, 9:43–45, code (57); *see also* Ex. 1003 ¶¶ 51–60, 126.

(b) Limitation [1a]

Claim 1 recites “[1a] forming a first virtual microphone by combining a first signal of a first physical microphone and a second signal of a second physical microphone.” Ex. 1001, 36:62–64.

Petitioner contends that Avendano discloses limitation [1a] because Avendano's audio device 104 comprises:

- (1) primary microphone 106 generating a primary acoustic signal X_1 that corresponds to a “first signal of a first physical microphone”; and
- (2) secondary microphone 108 generating a secondary acoustic signal X_2 that corresponds to a “second signal of a second physical microphone.”

Pet. 22–24 (citing Ex. 1003 ¶¶ 127–131; Ex. 1005, 3:27–55, 4:20–27, Figs. 1a–1b, 4a–4b). Petitioner also contends that audio device 104 forms a cardioid primary signal C_f corresponding to a “first virtual microphone” by combining the primary acoustic signal X_1 and the secondary acoustic signal X_2 . *Id.* at 24 (citing Ex. 1005, 4:27–6:10, Figs. 4a–4b).

To support its contentions, Petitioner provides an annotated version of Avendano's Figure 4a as reproduced below (Pet. 24):

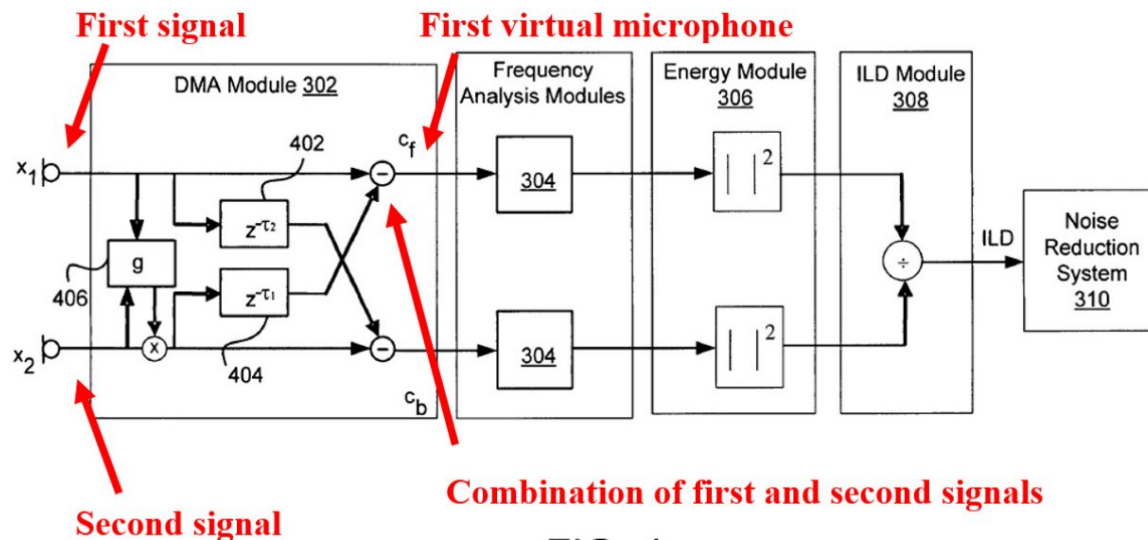


FIG. 4a

Figure 4a illustrates one implementation of audio processing engine 204 including various modules. Ex. 1005, 2:61–63, 5:15–17, 6:36–40, Fig. 4a. Petitioner's annotated version of Figure 4a includes the following annotations:

- (1) the red annotation "First signal" identifying the electric signal x_1 corresponding to the primary acoustic signal X_1 ;
- (2) the red annotation "Second signal" identifying the electric signal x_2 corresponding to the secondary acoustic signal X_2 ;
- (3) the red annotation "Combination of first and second signals" identifying a component that combines the electric signal x_1 corresponding to the primary acoustic signal X_1 with a gain-adjusted and delayed version of the electric signal x_2 corresponding to the secondary acoustic signal X_2 ; and
- (4) the red annotation "First virtual microphone" identifying the cardioid primary signal C_f , i.e., at the output of the component that combines the electric signal x_1 corresponding to the primary acoustic signal X_1 with the

gain-adjusted and delayed version of the electric signal x_2
corresponding to the secondary acoustic signal X_2 .

See Pet. 24; Ex. 1003 ¶ 130.

Patent Owner disputes that Avendano teaches limitation [1a]. *See, e.g.,* Prelim. Resp. 5–8. Specifically, Patent Owner asserts that the “purported first virtual microphone is composed of its ‘first signal’ received from X_1 and the signal received from X_2 after it is processed through ‘delay node 404.’” *Id.* at 6 (emphasis omitted) (citing Ex. 1005, 5:16–24, Fig. 4a; Pet. 24). Patent Owner asserts that “a signal after passing through a delay node is actually an ‘intermediate signal.’” *Id.* Patent Owner also asserts that Petitioner attempts to meet limitation [1a] by “combining an intermediate signal with the ‘first signal.’” *Id.*

Patent Owner further asserts that the Board “has rejected similar arguments” in a related proceeding, i.e., in *Google LLC v. Jawbone Innovations, LLC*, IPR2022-00604 (“the Google IPR”), Paper 12 (PTAB Oct. 6, 2022) (included in the record as Exhibit 2002). Prelim. Resp. 7. According to Patent Owner, the Board determined in the Google IPR that “combining an intermediate signal with another signal (there, the ‘second signal’) could not meet” limitation [1a]. *Id.* (citing Ex. 2002, 12).

Based on the current record, Petitioner establishes sufficiently for purposes of institution that Avendano discloses limitation [1a]. *See* Pet. 22–24; Ex. 1003 ¶¶ 51–60, 127–132. As Petitioner contends, Avendano’s audio device 104 comprises:

- (1) primary microphone 106 generating a primary acoustic signal X_1 that corresponds to a “first signal of a first physical microphone”; and

- (2) secondary microphone 108 generating a secondary acoustic signal X_2 that corresponds to a “second signal of a second physical microphone.”

Ex. 1005, 3:27–55, 4:23–27, Figs. 1a–1b, 4a; *see* Ex. 1003 ¶¶ 52–53, 127–131; Pet. 22–24. Audio device 104 also includes audio processing engine 204 with DMA module 302 that receives an electric signal x_1 corresponding to the primary acoustic signal X_1 and an electric signal x_2 corresponding to the secondary acoustic signal X_2 . Ex. 1005, 2:61–63, 4:6–8, 4:23–27, 5:15–20, Fig. 4a; *see* Ex. 1003 ¶¶ 55, 130–131.

After receiving the electric signals x_1 and x_2 , DMA module 302 generates a cardioid primary signal C_f according to the following equation for C_f :

$$C_f = X_1 - z^{-\tau_1} g X_2$$

where the delay $z^{-\tau_1}$ is introduced by delay node 404 and the gain factor “g” is computed by gain module 406. Ex. 1005, 5:18–30, 5:36–37; *see id.* at 8:52–60; Ex. 1003 ¶¶ 56–57, 132. Hence, as Petitioner contends, Avendano discloses forming a cardioid primary signal C_f corresponding to a “first virtual microphone” by combining the primary acoustic signal X_1 and the secondary acoustic signal X_2 . Ex. 1005, 4:27–5:37, Fig. 4a; *see* Ex. 1003 ¶¶ 56–57, 130–132; Pet. 22–24.

Based on the current record, we disagree with Patent Owner that a signal passing through a delay node or a delay filter becomes an “intermediate signal.” *See* Prelim. Resp. 6. The ’611 patent’s Figure 4 depicts a block diagram of a virtual microphone V_1 where a physical microphone O_1 provides a “first signal” to a delay filter $z^{-\gamma}$ and the delay filter $z^{-\gamma}$ provides the “first signal” to a summing circuit (Σ) to form the

virtual microphone V_1 from the “first signal” and a “second signal.”

Ex. 1001, 1:64–65, 5:22–33, 6:17–19, Fig. 4; *see* Ex. 1003 ¶ 46.

Similar to Avendano’s equation for the cardioid primary signal C_f discussed above, the ’611 patent provides the following equation for the first virtual microphone $V_1(z)$ shown in Figure 4 formed from a first physical microphone $O_1(z)$ that generates a “first signal,” a second physical microphone $O_2(z)$ that generates a “second signal,” a filter $\alpha(z)$, a filter $\beta(z)$, and a fixed delay “ γ ”:

$$V_1(z) = -\beta(z)\alpha(z)O_2(z) + O_1(z)z^{-\gamma}$$

Ex. 1001, 1:64–65, 5:22–25, 6:17–19, Fig. 4.

Regarding the Google IPR, the circumstances there do not parallel the circumstances here. There, Google asserted different prior art than Petitioner. Ex. 2002, 3, 8–11. Furthermore, Google failed to establish sufficiently that the prior art taught both the “first virtual microphone” and the “second virtual microphone” because Google identified identical inputs to the components that allegedly formed those microphones. *Id.* at 11–13. Here, Petitioner does not identify identical inputs to the components that allegedly form the “first virtual microphone” and the “second virtual microphone.” *See* Pet. 22–24, 28–33.

(c) Limitation [1b]

Claim 1 recites “[1b] forming a filter that describes a relationship for speech between the first physical microphone and the second physical microphone.” Ex. 1001, 36:65–67.

Petitioner contends that Avendano discloses limitation [1b] or alternatively that the combined disclosures in Avendano and Visser teach limitation [1b]. *See* Pet. 25–28. Regarding Avendano’s disclosures,

Petitioner contends that there will be a “difference in times of arrival of the signal from a speech source to the microphones,” i.e., primary microphone 106 and secondary microphone 108, “due to a space difference between the microphones.” *Id.* at 25 (citing Ex. 1003 ¶ 135; Ex. 1005, 1:33–36). According to Petitioner, “the ‘relationship for speech’ between Avendano’s two physical microphones is, at least in part, that one microphone will receive speech prior to the other microphone.” *Id.* (citing Ex. 1003 ¶ 136).

Petitioner asserts that Avendano’s audio device 104 includes delay node 402 implemented with a filter. Pet. 25–26 (citing Ex. 1003 ¶ 136; Ex. 1005, 5:15–35, 8:38–51). Petitioner asserts that delay node 402 “delays the first signal, which would otherwise include speech content before the second signal.” *Id.* at 25 (citing Ex. 1005, 5:15–35). Petitioner also asserts that delay node 402 “describes, at least in part, a temporal relationship for speech between the first and second signals” from microphones 106 and 108. *Id.* (citing Ex. 1003 ¶ 136).

Petitioner contends that the combined disclosures in Avendano and Visser teach limitation [1b] because “it would have been obvious to incorporate Visser’s ‘ICA or BSS processing function’ into Avendano’s system.” Pet. 26. Specifically, Petitioner asserts that:

- (1) “Visser’s cross filter w_{12} would have been applied to Avendano’s second signal x_2 ,” e.g., “to further increase speech content in the first virtual microphone C_f ”; and
- (2) “Visser’s cross filter w_{21} would have been applied to Avendano’s first signal x_1 ,” e.g., “to further decrease speech content in the second virtual microphone C_b .”

Id. at 27 (citing Ex. 1003 ¶¶ 140–141); *see id.* at 21 (citing Ex. 1003 ¶ 120).

To support its assertion, Petitioner provides a modified and annotated version of Avendano's Figure 4a as reproduced below (Pet. 27):

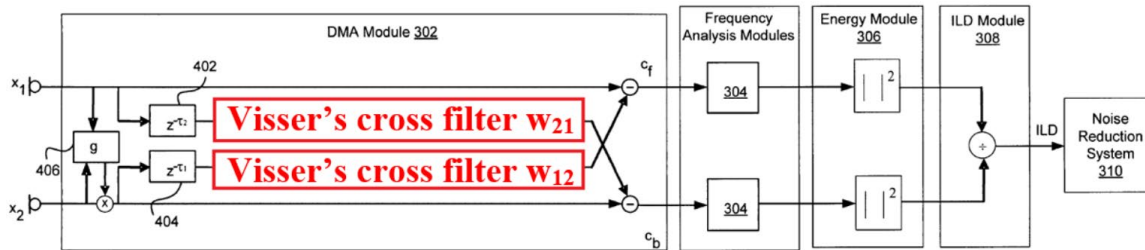


FIG. 4a

Figure 4a illustrates one implementation of audio processing engine 204 including various modules. Ex. 1005, 2:61–63, 5:15–17, 6:36–40, Fig. 4a. Petitioner's modified and annotated version of Figure 4a includes the following modifications and annotations:

- (1) a red box inserted after delay node 402 with delay $z^{-\tau_2}$;
- (2) the red annotation "Visser's cross filter w_{21} " within the red box inserted after delay node 402 with delay $z^{-\tau_2}$;
- (3) a red box inserted after delay node 404 with delay $z^{-\tau_1}$;
and
- (4) the red annotation "Visser's cross filter w_{12} " within the red box inserted after delay node 404 with delay $z^{-\tau_1}$.

See Pet. 27; Ex. 1003 ¶ 141.

Petitioner also asserts that each of Visser's cross filters w_{12} and w_{21} "describes a relationship for speech between first and second physical microphones." Pet. 26–27 (citing Ex. 1003 ¶¶ 140–141).

Patent Owner disputes that Avendano discloses limitation [1b] and that the combined disclosures in Avendano and Visser teach limitation [1b]. See, e.g., Prelim. Resp. 8–12; see also *id.* at 6 n.1. Regarding Avendano's disclosures, Patent Owner asserts that Petitioner fails to show that delay

node 402 describes “a relationship for speech” according to limitation [1b]. *Id.* at 8. Patent Owner asserts that Avendano “does not say” that delay node 402 produces a delay related to a “difference in times of arrival of the signals from a speech source to the microphones” or “related to speech in any way.” *Id.* (quoting Ex. 1005, 1:34–36). Patent Owner also asserts that Avendano “applies a delay to both received signals, which would not be necessary if the delay were only to account for the temporal relationship” as Petitioner suggests. *Id.* at 8–9 (emphasis omitted) (citing Ex. 1005, 5:19–23). Patent Owner further asserts that the ’611 patent describes (1) the filter $\beta(z)$, e.g., an adaptive filter, as “a filter that describes the relationship between O_1 and calibrated O_2 for speech” and (2) the delay filter z as “a fixed delay that depends on the size of the array.” *Id.* at 9–10 (emphases omitted) (citing Ex. 1001, 5:30–33, 6:29–32, Fig. 3).

Regarding the combined disclosures in Avendano and Visser, Patent Owner asserts that “Visser does not suggest or describe that its W_{21} and W_{12} filters describe any relationship for speech between the two microphones.” Prelim. Resp. 10. Patent Owner asserts that Visser’s speech-separation process 180 “relies heavily on Visser’s ‘two channel’ voice activity detector and its control signal 186.” *Id.* at 10–11 (citing Ex. 1006, 8:27–35, 9:1–14, 11:28–12:65). Patent Owner also asserts that Petitioner does not suggest that an ordinarily skilled artisan would have included Visser’s voice activity detector in the proposed combination. *Id.* at 11 (citing Pet. 26–28).

Based on the current record, Petitioner establishes sufficiently for purposes of institution that Avendano discloses limitation [1b] and that the combined disclosures in Avendano and Visser teach limitation [1b]. *See* Pet. 25–28; Ex. 1003 ¶¶ 133–144. Specifically, Avendano discloses

positioning primary microphone 106 closer to audio (speech) source 102 than secondary microphone 108. Ex. 1005, 3:45–47, Fig. 1a; *see id.* at 1:39–41, 3:29–30; Ex. 1003 ¶¶ 53, 135. As Petitioner asserts, Avendano explains that “due to a space difference between the microphones,” the microphones experience a “difference in times of arrival of the signals from a speech source.” Ex. 1005, 1:33–36; *see* Ex. 1003 ¶ 135; Pet. 25.

Avendano also discloses using delay nodes, e.g., implemented with allpass filters, to produce “an alignment of the two channels.” Ex. 1005, 8:45–51; *see* Ex. 1003 ¶¶ 60, 137. For example, DMA module 302 delays the electric signal x_1 corresponding to the primary acoustic signal X_1 via delay node 402 with delay z^{-t_2} before combining the delayed signal with a gain-adjusted version of the electric signal x_2 corresponding to the secondary acoustic signal X_2 . Ex. 1005, 5:19–24, 5:31–37, Fig. 4a; *see* Ex. 1003 ¶ 60.

Patent Owner’s assertion that Avendano “does not say” that delay node 402 produces a delay related to a “difference in times of arrival of the signals from a speech source to the microphones” disregards the principle that an obviousness analysis “need not seek out” the “precise” disclosure of a specific feature in a challenged claim because the analysis “can take account of the inferences and creative steps that a person of ordinary skill in the art would employ.” *See KSR*, 550 U.S. at 418; Prelim. Resp. 8. A person of ordinary skill “must be presumed to know something” about the art “apart from what the references disclose.” *In re Jacoby*, 309 F.2d 513, 516 (CCPA 1962). This issue is better resolved at trial on a more developed record.

Avendano discloses that there is a “difference in times of arrival of the signals from a speech source” at two microphones “due to a space difference

between the microphones.” Ex. 1005, 1:33–36; *see* Ex. 1003 ¶ 135.

Avendano also discloses that a delay node may produce “an alignment of the two channels.” Ex. 1005, 8:45–46. Hence, based on the current record, we are persuaded that Petitioner has shown sufficiently at this stage that delay node 402 produces a delay related to a “difference in times of arrival of the signals from a speech source.” *See* Pet. 25–26.

Regarding Patent Owner’s assertion that Avendano “applies a delay to both received signals, which would not be necessary if the delay were only to account for the temporal relationship,” Avendano employs two delay nodes in a preferred embodiment as follows:

- (1) a first delay node applied to a non-gain-adjusted version of the electric signal x_1 corresponding to the primary acoustic signal X_1 to form the cardioid secondary signal C_b ; and
- (2) a second delay node applied to a gain-adjusted version of the electric signal x_2 corresponding to the secondary acoustic signal X_2 to form the cardioid primary signal C_f .

See Ex. 1005, 2:22–29, 3:29–39, 5:19–35, 9:54–64, Fig. 4a; Prelim.

Resp. 8–9. Avendano does not indicate that the delays introduced by the delay nodes necessarily differ. *See, e.g.*, Ex. 1005, 2:22–29, 3:29–33, 9:54–64. But even if the delays introduced by the delay nodes differ, the different delays may serve different purposes, e.g., one delay to account for a “difference in times of arrival of the signals from a speech source” and another delay to account for a difference in microphone properties. This issue is better resolved at trial on a more developed record.

As noted above, Patent Owner asserts that the ’611 patent describes the filter $\beta(z)$, e.g., an adaptive filter, as “a filter that describes the relationship between O_1 and calibrated O_2 for speech.” Prelim. Resp. 9

(emphases omitted). Based on the current record, we preliminarily determine that Patent Owner improperly seeks to read a limitation from the specification into claim 1. *See In re Van Geuns*, 988 F.2d 1181, 1184 (Fed. Cir. 1993). At this stage of the proceeding, Patent Owner cites nothing in the specification that precludes “a relationship for speech” according to limitation [1b] from encompassing a temporal relationship resulting from a “difference in times of arrival of the signals from a speech source” that a delay filter addresses. *See, e.g.*, Prelim. Resp. 8–12. Instead, the specification discloses using a delay filter to address a difference in distances (and thus a difference in signal arrival times) from a speech source to physical microphones O_1 and O_2 . *See* Ex. 1001, 23:18–41. We invite the parties to address in the Response, Reply, and Sur-reply how the specification affects the claim construction the Board will use in the Final Written Decision when determining claim 1’s patentability.

Regarding the combined disclosures in Avendano and Visser, we are persuaded that “[e]ach of Visser’s cross filters w_{12} and w_{21} describes a relationship for speech between a first physical microphone and a second physical microphone” since each cross filter functions to “distinguish speech from noise.” Ex. 1003 ¶¶ 140, 144; *see* Ex. 1006, 17:29–44, Fig. 10. Specifically, cross filter w_{12} “is configured to filter one microphone signal” containing speech components and noise components to “generate a first processed signal representative of speech” by “combining the filtered signal with another microphone signal.” Ex. 1003 ¶ 115; Ex. 1006, 15:62–65, 16:58–60, 17:36–44, Fig. 10. Additionally, cross filter w_{21} “is configured to filter one microphone signal” containing speech components and noise components to “generate a second processed signal representative of noise”

by “combining the filtered signal with another microphone signal.”

Ex. 1003 ¶ 118; Ex. 1006, 15:62–65, 16:58–60, 17:36–44, Fig. 10.

Because cross filter w_{12} functions to increase speech components from the two microphone signals and cross filter w_{21} functions to decrease speech components from the two microphone signals, each cross filter “describes a relationship for speech between the first physical microphone and the second physical microphone” according to limitation [1b]. *See* Ex. 1003

¶¶ 140–144; Ex. 1006, 15:62–65, 16:58–60, 17:29–44, Fig. 10. Thus, we disagree with Patent Owner that “Visser does not suggest or describe that its W_{21} and W_{12} filters describe any relationship for speech between the two microphones.” *See* Prelim. Resp. 10.

Patent Owner’s assertion that Visser’s speech-separation process 180 “relies heavily on Visser’s ‘two channel’ voice activity detector and its control signal 186” does not address the proposed combination. *See* Pet. 15–22, 26–27; Prelim. Resp. 10–11. Petitioner proposes modifying Avendano’s system to incorporate Visser’s cross filters w_{12} and w_{21} from sub-module 600 depicted in Figure 10. Pet. 20–21, 26–27. Petitioner does not propose modifying Avendano’s system to incorporate Visser’s speech-separation process 180 from communication process 175 depicted in Figure 2. *Id.* at 20–21, 26–27.

(d) Limitation [1c]

Claim 1 recites “[1c] forming a second virtual microphone by applying the filter to the first signal to generate a first intermediate signal, and summing the first intermediate signal and the second signal.” Ex. 1001, 37:1–4.

Petitioner contends that Avendano discloses limitation [1c] or alternatively that the combined disclosures in Avendano and Visser teach limitation [1c]. *See* Pet. 28–33. In particular, Petitioner contends that Avendano discloses limitation [1c] because Avendano’s audio device 104 includes delay node 402 that receives the electric signal x_1 corresponding to the primary acoustic signal X_1 and generates a “first intermediate signal.” *Id.* at 28–29 (citing Ex. 1003 ¶¶ 145–152; Ex. 1005, 5:15–35, Figs. 4a–4b). Petitioner also contends that audio device 104 forms a cardioid secondary signal C_b corresponding to a “second virtual microphone” by combining the “first intermediate signal” with a gain-adjusted version of the electric signal x_2 corresponding to the secondary acoustic signal X_2 . *Id.* at 29–30 (citing Ex. 1003 ¶¶ 145–152; Ex. 1005, 5:15–35, Figs. 4a–4b).

To support its contentions, Petitioner provides an annotated version of Avendano’s Figure 4a as reproduced below (Pet. 29):

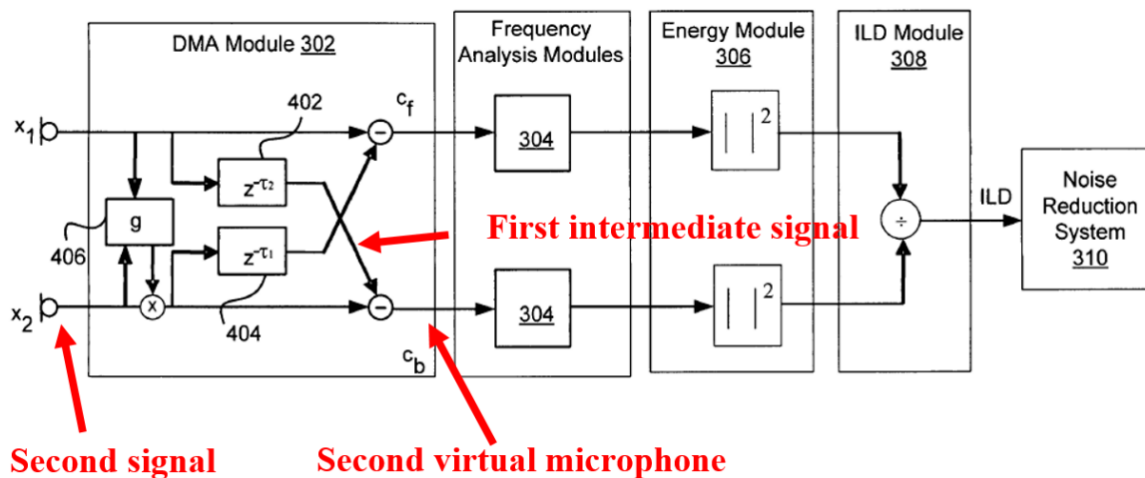


FIG. 4a

Figure 4a illustrates one implementation of audio processing engine 204 including various modules. Ex. 1005, 2:61–63, 5:15–17, 6:36–40, Fig. 4a. Petitioner’s annotated version of Figure 4a includes the following annotations:

- (1) the red annotation “First intermediate signal” identifying a delayed version of the electric signal x_1 corresponding to the primary acoustic signal X_1 , i.e., at the output of delay node 402 with delay $z^{-\tau_2}$;
- (2) the red annotation “Second signal” identifying the electric signal x_2 corresponding to the secondary acoustic signal X_2 ; and
- (3) the red annotation “Second virtual microphone” identifying the cardioid secondary signal C_b , i.e., at the output of a component that combines the delayed version of the electric signal x_1 corresponding to the primary acoustic signal X_1 with a gain-adjusted version of the electric signal x_2 corresponding to the secondary acoustic signal X_2 .

See Pet. 29; Ex. 1003 ¶ 148.

Petitioner contends that the combined disclosures in Avendano and Visser teach limitation [1c] because “[i]t would have been obvious to incorporate Visser’s ‘ICA or BSS processing function’ into Avendano’s system to further enhance the separation between speech and noise.”

Pet. 31–32. Petitioner identifies an example modification where:

- (1) “Visser’s cross filter w_{12} would have been applied to Avendano’s second signal x_2 ,” e.g., “to further increase speech content in the first virtual microphone C_f ”; and
- (2) “Visser’s cross filter w_{21} would have been applied to Avendano’s first signal x_1 ,” e.g., “to further decrease speech content in the second virtual microphone C_b .”

Id. at 32; *see id.* at 21 (citing Ex. 1003 ¶ 120).

At this stage of the proceeding, Patent Owner makes no arguments specific to limitation [1c]. *See, e.g.*, Prelim. Resp. 5–12. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Kenny's testimony, Petitioner establishes sufficiently for purposes of institution that Avendano discloses limitation [1c] and that the combined disclosures in Avendano and Visser teach limitation [1c]. *See* Pet. 28–33; Ex. 1003 ¶¶ 145–162.

(e) Limitation [1d]

Claim 1 recites “[1d] generating an energy ratio of energies of the first virtual microphone and the second virtual microphone.” Ex. 1001, 37:5–6.

Petitioner contends that Avendano discloses limitation [1d] because Avendano's audio device 104 determines the following:

- (1) the energy of the first virtual microphone E_f “by integrating the first virtual microphone over time” using the equation $E_f(t, \omega) = \int_{\text{frame}} |C_f(t', \omega)|^2 dt'$;
- (2) the energy of the second virtual microphone E_b “by integrating the second virtual microphone over time” using the equation $E_b(t, \omega) = \int_{\text{frame}} |C_b(t', \omega)|^2 dt'$; and
- (3) the energy ratio or Inter-Level Difference (ILD) “by dividing the energy of the first virtual microphone by the energy of the second virtual microphone.”

Pet. 33–34 (citing Ex. 1003 ¶¶ 163–166; Ex. 1005, 5:49–6:5, 6:8–34).

At this stage of the proceeding, Patent Owner makes no arguments specific to limitation [1d]. *See, e.g.*, Prelim. Resp. 5–12. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Kenny's testimony, Petitioner establishes

sufficiently for purposes of institution that Avendano discloses limitation [1d]. *See* Pet. 33–34; Ex. 1003 ¶¶ 163–166.

(f) Limitation [1e]

Claim 1 recites “[1e] detecting acoustic voice activity of a speaker when the energy ratio is greater than a threshold value.” Ex. 1001, 37:7–8.

Petitioner contends that Avendano discloses limitation [1e] because Avendano’s audio device 104 “compares the energy ratio (ILD) to a ‘threshold’ to determine the presence or absence of speech.” Pet. 35 (citing Ex. 1003 ¶¶ 167–171; Ex. 1005, 6:35–7:9). According to Petitioner, an energy ratio below the threshold results in setting the value λ_1 to zero “indicating an absence of speech,” while an energy ratio above the threshold results in setting the value λ_1 to one “indicating a presence of speech.” *Id.* at 35–36 (citing Ex. 1005, 6:58–7:3); *see id.* at 16 (citing Ex. 1005, 5:49–6:34, 6:58–7:3).

At this stage of the proceeding, Patent Owner makes no arguments specific to limitation [1e]. *See, e.g.*, Prelim. Resp. 5–12. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Kenny’s testimony, Petitioner establishes sufficiently for purposes of institution that Avendano discloses limitation [1e]. *See* Pet. 35; Ex. 1003 ¶¶ 167–171.

(g) Alleged Reasons for Combining the Teachings of the References

Petitioner identifies reasons that would have prompted an ordinarily skilled artisan to combine Visser’s teachings with Avendano’s teachings in the way Petitioner proposes. *See* Pet. 15–21, 27–28, 32. Additionally,

Petitioner explains that an ordinarily skilled artisan “would have had a reasonable expectation of success” in combining the teachings of the references. *Id.* at 21–22. Dr. Kenny’s testimony supports Petitioner’s positions. *See* Ex. 1003 ¶¶ 101–125, 141–143, 159–160.

For instance, Dr. Kenny testifies that modifying Avendano’s system in view of Visser’s teachings “would have improved Avendano’s system by enabling the system to ‘adapt’ and ‘learn’ to separate speech ‘according to the specific acoustic environment,’ such that the system can accurately separate speech ‘even in a changing acoustic environment.’” Ex. 1003 ¶ 104 (quoting Ex. 1006, 9:9–10, 17:4); *see id.* ¶¶ 111–112. According to Dr. Kenny, “Visser’s filters can be ‘adapt[ed] during operation,’ such that the filters can better separate speech from noise based on the signals received from the physical microphones” to “provide a sufficient level of separation, even in a changing acoustic environment.” *Id.* ¶¶ 111, 143 (alteration by Dr. Kenny) (quoting Ex. 1006, 17:3–4, 21:24–25); *see id.* ¶¶ 112, 158. Dr. Kenny explains that “[s]uch a modification would have further enhanced the differences in speech content in Avendano’s two processed signals” so “speech can be detected with a greater degree of accuracy (e.g., by comparing the energy levels of signals representing speech and noise).” *Id.* ¶ 112 (emphasis omitted).

Dr. Kenny also testifies that an ordinarily skilled artisan would have recognized that:

- (1) “each of Avendano’s system and Visser’s system is configured to process signals from two physical microphones including the step of filtering one microphone signal, in order to generate a first processed signal representative of speech”;

- (2) “Avendano’s system is configured to filter a microphone signal x_2 using a ‘delay node 404,’ and generate a first processed signal representative of speech (e.g., a ‘cardioid primary signal (C_f)’) by combining the filtered signal with another microphone signal x_1 ”;
- (3) “like Avendano’s ‘delay node 404,’ Visser’s ‘cross filter w_{12} ’ is configured to filter one microphone signal (e.g., a ‘channel 620’) in order to generate a first processed signal representative of speech (e.g., a ‘speech channel,’ such as the ‘channel 630’) by combining the filtered signal with another microphone signal (e.g., a ‘channel 610’)”;
- (4) “each of Avendano’s system and Visser’s systems [sic] is configured to process signals from two physical microphones including the step of filtering another microphone signal, in order to generate a second processed signal representative of noise”;
- (5) “Avendano’s system is configured to filter a microphone signal x_1 using a ‘delay node 402,’ and generate a second processed signal representative of noise (e.g., a ‘cardioid secondary signal (C_b)’) by combining the filtered signal with another microphone signal x_2 ”; and
- (6) “like Avendano’s ‘delay node 402,’ Visser’s ‘cross filter w_{21} ’ is configured to filter one microphone signal (e.g., a ‘channel 610’) in order to generate a second processed signal representative of noise (e.g., a ‘noise channel,’ such as the ‘channel 640’) by combining the filtered signal with another microphone signal (e.g., a ‘channel 620’).”

Ex. 1003 ¶¶ 113–118 (emphases omitted).

Dr. Kenny explains that an ordinarily skilled artisan would have

- (1) “applied Visser’s cross filter w_{12} to Avendano’s second signal x_2 to further increase speech content in the first virtual microphone C_f ” and
- (2) “applied Visser’s cross filter w_{21} to Avendano’s first signal x_1 to further

decrease speech content in the second virtual microphone C_b .” Ex. 1003 ¶¶ 120, 141, 160. Dr. Kenny also explains that these modifications would have improved Avendano’s system by “further enhanc[ing] the separation of speech in Avendano’s system,” i.e., “further enhanc[ing] the separation between speech and noise.” *Id.* ¶¶ 119, 159; *see id.* ¶ 154.

Patent Owner disputes that an ordinarily skilled artisan would have been motivated to combine Visser’s teachings with Avendano’s teachings in the way Petitioner proposes. *See* Prelim. Resp. 11–12. Specifically, Patent Owner asserts that Petitioner provides “no reason” why an ordinarily skilled artisan would have configured “the modified Avendano system in the specific way” Petitioner proposes. *Id.* at 11. Patent Owner also asserts that Avendano differs from Visser because (1) Avendano subtracts the signals supplied by delay nodes 402 and 404 and (2) Visser sums the signals supplied by cross filters w_{12} and w_{21} . *Id.* at 11–12. Patent Owner further asserts that Avendano already accounts for “changing acoustic signals” with a “Wiener filter and filter smoothing module.” *Id.* at 12 (citing Ex. 1005, 7:53–8:17).

Based on the current record, Petitioner establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had reasons to combine Visser’s teachings with Avendano’s teachings in the way Petitioner proposes. *See* Pet. 15–21, 27–28, 32; Ex. 1003 ¶¶ 101–120, 141–143, 159–160. As Dr. Kenny explains, an ordinarily skilled artisan would have (1) “applied Visser’s cross filter w_{12} to Avendano’s second signal x_2 to further increase speech content in the first virtual microphone C_f ” and (2) “applied Visser’s cross filter w_{21} to Avendano’s first signal x_1 to further decrease speech content in the second virtual microphone C_b .”

Ex. 1003 ¶¶ 120, 141, 160. As Dr. Kenny also explains, these modifications would have improved Avendano’s system by “further enhanc[ing] the separation of speech in Avendano’s system,” i.e., “further enhanc[ing] the separation between speech and noise.” *Id.* ¶¶ 119, 159. Hence, combining Visser’s teachings with Avendano’s teachings would have furthered Avendano’s objective to “attenuate noise and enhance speech.” *See* Ex. 1005, 2:3–7, 2:48–50, 3:13–17, 9:43–45, code (57).

Any “need or problem known in the field of endeavor at the time of invention and addressed by the patent can provide a reason for combining” references. *KSR*, 550 U.S. at 420. At this stage of the proceeding, we are persuaded that improving Avendano’s system by “further enhanc[ing] the separation of speech in Avendano’s system” suffices as a reason for combining Visser’s teachings with Avendano’s teachings. *See* Ex. 1003 ¶¶ 119–120, 141, 159–160. The “desire to enhance commercial opportunities by improving a product or process is universal” *DyStar Textilfarben GmbH v. C.H. Patrick Co.*, 464 F.3d 1356, 1368 (Fed. Cir. 2006).

Because Dr. Kenny explains that combining Visser’s teachings with Avendano’s teachings would have improved Avendano’s system by “further enhanc[ing] the separation of speech in Avendano’s system,” i.e., “further enhanc[ing] the separation between speech and noise,” we disagree with Patent Owner that Petitioner provides “no reason” why an ordinarily skilled artisan would have configured “the modified Avendano system in the specific way” Petitioner proposes. *See* Ex. 1003 ¶¶ 119, 159; Prelim. Resp. 11.

Moreover, a motivation to combine “can be found explicitly or implicitly in the prior art references themselves, in market forces, in design incentives, or in ‘any need or problem known in the field of endeavor at the time of invention and addressed by the patent.’” *Forest Labs., LLC v. Sigmapharm Labs., LLC*, 918 F.3d 928, 934 (Fed. Cir. 2019) (quoting *Arctic Cat Inc. v. Bombardier Recreational Prods. Inc.*, 876 F.3d 1350, 1359 (Fed. Cir. 2017)). Here, Dr. Kenny identifies a motivation to combine “in the prior art references themselves,” i.e., in Visser’s disclosures. *See, e.g.*, Ex. 1003 ¶¶ 104, 111–112, 118–120, 143, 158 (citing Ex. 1006, 8:16–18, 9:8–12, 9:20–62, 16:3–28, 16:65–17:4, 17:29–25:9, Figs. 9–13).

Also, “[t]here is a motivation to combine when a known technique ‘has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way,’ using the ‘prior art elements according to their established functions.’” *Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1380–81 (Fed. Cir. 2023) (citations omitted) (quoting *KSR*, 550 U.S. at 417; *Intel Corp. v. Qualcomm Inc.*, 21 F.4th 784, 799–800 (Fed. Cir. 2021)). Here, Petitioner proposes using Visser’s cross filters w_{12} and w_{21} according to their established functions, i.e., to facilitate separating speech signals from noise sources. *See* Ex. 1003 ¶¶ 120, 141, 160; Ex. 1006, 16:8–27, 17:56–64; Pet. 20–21, 27, 32.

Patent Owner’s assertion that Avendano differs from Visser because (1) Avendano subtracts the signals supplied by delay nodes 402 and 404 and (2) Visser sums the signals supplied by cross filters w_{12} and w_{21} disregards the principle that “[c]ombining the *teachings* of references does not involve an ability to combine their specific structures.” *See In re Nievelt*, 482 F.2d 965, 968 (CCPA 1973); Prelim. Resp. 11–12. Obviousness does not depend

on “whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference.” *In re Keller*, 642 F.2d 413, 425 (CCPA 1981).

That Avendano’s system already accounts for “changing acoustic signals” with a “Wiener filter and filter smoothing module,” as Patent Owner asserts, does not undermine modifying Avendano’s system to incorporate Visser’s cross filters w_{12} and w_{21} to “further enhance the separation of speech in Avendano’s system,” i.e., “further enhance the separation between speech and noise.” *See* Ex. 1003 ¶¶ 119–120, 141, 159–160; Prelim. Resp. 12. Visser’s cross filters w_{12} and w_{21} would have improved Avendano’s system regardless whether the cross filters adapted during operation. *See* Ex. 1003 ¶¶ 119–120, 141, 159–160. We are persuaded that Visser’s disclosure about filters adapting during operation and learning to separate speech “according to the specific acoustic environment,” such that the system can accurately separate speech “even in a changing acoustic environment,” provides an independent reason for combining Visser’s teachings with Avendano’s teachings. *See, e.g.*, Ex. 1003 ¶¶ 104, 111–112, 143, 158–159; Ex. 1006, 7:1–5, 9:8–12, 16:8–19, 16:62–17:4, code (57).

(h) Preliminary Conclusion for Obviousness/Nonobviousness of Claim 1

As discussed above, Petitioner’s analysis addresses how the combined disclosures in Avendano and Visser teach claim 1’s subject matter. *See supra* §§ IV.D.3(a)–(f). Additionally, Petitioner provides a reason with rational underpinning as to why an ordinarily skilled artisan would have been motivated to combine Visser’s teachings with Avendano’s teachings in the way Petitioner proposes and would have had a reasonable expectation of

success. *See supra* § IV.D.3(g). Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claim 1 unpatentable under § 103(a) as obvious over Avendano and Visser.

4. DEPENDENT CLAIMS 2–7 AND 25–28

Claims 2–7 and 25–28 depend directly or indirectly from claim 1. Ex. 1001, 37:9–25, 38:20–32.

For claims 2–7 and 25–28, Petitioner provides an analysis supported by Dr. Kenny’s testimony about how the combined disclosures in Avendano and Visser teach the claimed subject matter. *See* Pet. 35–42; Ex. 1003 ¶¶ 172–214.

At this stage of the proceeding, Patent Owner makes no arguments specific to claims 2–7 and 25–28. *See, e.g.*, Prelim. Resp. 5–12. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Kenny’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano and Visser teach the subject matter in claims 2–7 and 25–28. *See* Pet. 35–42; Ex. 1003 ¶¶ 172–214. Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claims 2–7 and 25–28 unpatentable under § 103(a) as obvious over Avendano and Visser.

E. Alleged Obviousness over Avendano, Visser, and Bisgaard: Claims 8–16, 23, and 24

Petitioner contends that claims 8–16, 23, and 24 are unpatentable under § 103(a) as obvious over Avendano, Visser, and Bisgaard. *See* Pet. 1, 43–58. Patent Owner disputes Petitioner’s contentions. *See* Prelim.

Resp. 13. Above, we provided overviews of Avendano and Visser. *See supra* §§ IV.D.1–IV.D.2. Below, we provide an overview of Bisgaard, and then we consider the obviousness issues. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano, Visser, and Bisgaard teach the subject matter in claims 8–16, 23, and 24.

1. OVERVIEW OF BISGAARD (EXHIBIT 1010)

Bisgaard is a U.S. patent application publication titled “Hearing Instrument with Adaptive Directional Signal Processing,” filed on June 25, 2007, and published on May 5, 2011. Ex. 1010, codes (12), (22), (43), (54). Bisgaard states that the invention “relates to a hearing instrument, such as a hearing aid, an implantable hearing prosthesis, a head set, [or] a mobile phone,” with “a signal processor for directional signal processing.” *Id.* ¶ 2; *see id.* at code (57). “The signal processor may be adapted to perform directional signal processing selected from the group consisting of an adaptive beam former, a multi-channel Wiener filter, an independent component analysis, and a blind source separation algorithm.” *Id.* ¶ 27.

Bisgaard explains that a “directional signal processing system combines the electrical signals from the microphones in the array into a signal with varying sensitivity to sound sources in different directions in relation to the array” that results in a “directivity pattern.” Ex. 1010 ¶ 4. “Typically, a directivity pattern has at least one direction wherein the microphone signals substantially cancel each other” called a “null direction.” *Id.* “A directivity pattern may comprise several null directions depending on the number of microphones in the array and depending on the signal processing.” *Id.*

Bisgaard's Figure 2 (reproduced below) depicts signal processing in a hearing instrument:



Figure 2 illustrates a hearing instrument including microphones 20 and 22 “positioned in a housing to be worn at the ear of the user.” Ex. 1010 ¶ 39; *see id.* ¶ 31. “When the housing is mounted in its operating position at the ear of the user, one of the microphones, the front microphone 20, is positioned in front of the other microphone, the rear microphone 22.” *Id.* ¶ 39. A “horizontal line extending through the front and rear microphones defines the front direction, i.e. azimuth=0°, corresponding to the looking direction of the user of the hearing instrument.” *Id.*

Microphones 20 and 22 “convert received sound signals into corresponding electrical sound signals that are converted into digital sound signals 24, 26 by respective A/D converters (not shown).” Ex. 1010 ¶ 41. “Each of the digitized sound signals 24, 26 is input to a respective subtraction circuit 28, and a respective delay 32, 34 with delay D_H .” *Id.* ¶ 42. “Each delay 32, 34 delays the digitized sound signal 24, 26 by the

amount of time used by a sound signal to propagate in the 0° azimuth direction from the front microphone 20 to the rear microphone 22.” *Id.* “Each subtraction circuit 28, 30 subtracts the respective delayed signal 36, 38 from one microphone 20, 22 from the direct signal 26, 24 of the other microphone 22, 20.” *Id.*

“Each of the subtracted signals 40, 42 has a fixed directional pattern 44, 46, a so-called cardioid pattern.” Ex. 1010 ¶ 42. Specifically, “cardioid pattern 44 of the upper branch (a) has a null direction 48 at 180° azimuth, i.e. pointing in the rear direction of the user,” and “cardioid pattern 46 of the lower branch (b) has a null direction 50 at 0° azimuth, i.e. pointing in the front direction of the user.” *Id.*

“The subtracted signal 42 of the lower branch (b) is filtered by an adaptive filter 52 with a transfer function H , and the subtracted signal 40 of the upper branch (a) is delayed by a delay 54 with a delay D_H equal to the group delay of the adaptive filter 52.” Ex. 1010 ¶ 43. Subsequently, “the two signals 56, 58 are subtracted for formation of a combined signal 60 with a directivity pattern 62 with an adaptive null direction θ .” *Id.* “The hatched area of the resulting directivity pattern 62 illustrates the prohibited range of directions which in the illustrated example is symmetrical around 0° azimuth.” *Id.*

2. DIFFERENCES BETWEEN THE CLAIMED INVENTION AND THE PRIOR ART

Claims 8–16, 23, and 24 depend indirectly from claim 1. Ex. 1001, 37:26–53, 38:10–13. As an example, claim 8 depends directly from claim 7 and specifies that “the delay is proportional to a time difference between arrival of the speech at the second physical microphone and arrival of the

speech at the first physical microphone.” *Id.* at 37:26–29. As another example, claim 23 depends directly from claim 12 and specifies that “the filter is a static filter.” *Id.* at 38:10–11.

For claims 8–16, 23, and 24, Petitioner provides an analysis supported by Dr. Kenny’s testimony about how the combined disclosures in Avendano, Visser, and Bisgaard teach the claimed subject matter. *See* Pet. 43–58; Ex. 1003 ¶¶ 215–279.

As an example of how Petitioner relies on Bisgaard, Petitioner asserts for claim 8 that “Bisgaard’s ‘delays’ are used to account for ‘the amount of time used by a sound signal to propagate in the 0° azimuth direction from [a] front microphone . . . to [a] rear microphone.’” Pet. 48 (alterations by Petitioner) (quoting Ex. 1010 ¶ 42); *see* Ex. 1003 ¶ 228. Petitioner also asserts that an ordinarily skilled artisan “would have recognized that, to account for ‘the amount of time used by a sound signal to propagate in the 0° azimuth direction from [a] front microphone . . . to [a] rear microphone,’” the sound signal would be delayed “proportional to a time difference between arrival of the speech at the second physical microphone and arrival of the speech at the first physical microphone” according to claim 8. Pet. 48 (alterations by Petitioner) (quoting Ex. 1010 ¶ 42); *see* Ex. 1003 ¶¶ 229–230.

As another example of how Petitioner relies on Bisgaard, Petitioner asserts that Bisgaard would have rendered a “static filter” according to claim 23 obvious because an ordinarily skilled artisan would have recognized that if “the first and second microphones do not move relative to one another or to the source of sound, then ‘the amount of time used by a sound signal to propagate in the 0° azimuth direction from [a] front

microphone . . . to [a] rear microphone’ would not change.” Pet. 56 (alterations by Petitioner) (quoting Ex. 1010 ¶ 42); *see* Ex. 1003 ¶¶ 272–274.

Patent Owner contends that Petitioner “has not shown that any claim is rendered obvious” over Avendano, Visser, and Bisgaard because (1) “Avendano and Visser fail to disclose or render obvious” limitations in claim 1 and (2) Petitioner “does not argue that Bisgaard remedies these deficiencies.” Prelim. Resp. 13.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Kenny’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano, Visser, and Bisgaard teach the subject matter in claims 8–16, 23, and 24. *See* Pet. 43–44, 47–58; Ex. 1003 ¶¶ 78–81, 223–279.

3. ALLEGED REASONS FOR COMBINING THE TEACHINGS OF THE REFERENCES

Petitioner identifies a reason that would have prompted an ordinarily skilled artisan to combine the teachings of Avendano, Visser, and Bisgaard. *See* Pet. 44–46. Additionally, Petitioner explains that an ordinarily skilled artisan “would have had a reasonable expectation of success” in combining the teachings of the references. *Id.* at 46–47. Dr. Kenny’s testimony supports Petitioner’s positions. *See* Ex. 1003 ¶¶ 215–222.

For instance, Dr. Kenny testifies that an ordinarily skilled artisan “would have looked to Bisgaard for details related to implementing” a “delay in Avendano’s system” because:

- (1) “both of Avendano’s system and Bisgaard’s system are configured to receive signals from two microphones, and process the signals to obtain (i) a cardioid signal directed in a forward direction (e.g., towards a user or other audio source), and (ii) another cardioid signal directed in a

backwards direction (e.g., away from the user or other audio source)”;

- (2) “both of Avendano’s system and Bisgaard’s system are configured to form each of the cardioid signals by delaying the signal from one microphone, and combining it with the signal from the other microphone”; and
- (3) “Avendano’s ‘delay nodes’ and Bisgaard’s ‘delays’ serve the same purpose, namely delaying signals from physical microphones to produce directional cardioid signals (e.g., by combining the delayed microphone signal with another microphone signal).”

Ex. 1003 ¶¶ 217–219 (emphases omitted).

At this stage of the proceeding, Patent Owner makes no arguments against combining the teachings of Avendano, Visser, and Bisgaard. *See, e.g.*, Prelim. Resp. 13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record, Petitioner establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had a reason, e.g., as articulated by Dr. Kenny, to combine the teachings of Avendano, Visser, and Bisgaard. *See* Pet. 44–46; Ex. 1003 ¶¶ 215–220. Petitioner also establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had a reasonable expectation of success in combining the teachings of the references. *See* Pet. 46–47; Ex. 1003 ¶¶ 221–222.

4. PRELIMINARY CONCLUSION ABOUT OBVIOUSNESS/NONOBVIOUSNESS

As discussed above, Petitioner’s analysis addresses how the combined disclosures in Avendano, Visser, and Bisgaard teach the subject matter in claims 8–16, 23, and 24. *See supra* § IV.E.2. Additionally, Petitioner provides a reason with rational underpinning as to why an ordinarily skilled

artisan would have been motivated to combine the teachings of Avendano, Visser, and Bisgaard and would have had a reasonable expectation of success. *See supra* § IV.E.3. Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claims 8–16, 23, and 24 unpatentable under § 103(a) as obvious over Avendano, Visser, and Bisgaard.

F. Alleged Obviousness over Avendano, Visser, Bisgaard, and Hou: Claims 17–19

Petitioner contends that claims 17–19 are unpatentable under § 103(a) as obvious over Avendano, Visser, Bisgaard, and Hou. *See* Pet. 1, 58–70. Patent Owner disputes Petitioner’s contentions. *See* Prelim. Resp. 13. Above, we provided overviews of Avendano, Visser, and Bisgaard. *See supra* §§ IV.D.1–IV.D.2, IV.E.1. Below, we provide an overview of Hou, and then we consider the obviousness issues. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano, Visser, Bisgaard, and Hou teach the subject matter in claims 17–19.

1. OVERVIEW OF HOU (EXHIBIT 1008)

Hou is a U.S. patent titled “Adaptive Microphone Matching in Multi-Microphone Directional System,” filed on March 14, 2001, and issued on December 26, 2006. Ex. 1008, codes (12), (22), (45), (54). Hou states that the invention “relates to multi-microphone sound pick-up systems” and “matching microphone sensitivity in multi-microphone sound pick-up systems.” *Id.* at 1:22–25.

Hou explains that a “multi-microphone array can selectively enhance sounds coming from certain directions while suppressing interference coming from other directions.” Ex. 1008, 1:30–32. To “achieve good

directionality,” however, the sensitivities of the microphones must match. *Id.* at 1:48–50; *see id.* at 1:67–2:2. “When the sensitivities of the microphones are not properly matched, then the directionality is substantially degraded and thus the ability to suppress interference coming from a particular direction is poor.” *Id.* at 1:50–53; *see id.* at 1:53–62, Figs. 2(a)–2(d). The sensitivities of the microphones may not match due to manufacturing differences or “drift over time.” *See id.* at 2:3–10, 2:35–37. “Hence, once matched microphones can become mismatched over time.” *Id.* at 2:10–11.

Hou purports to address that problem by providing “improved approaches to matching sensitivities of microphones in multi-microphone directional processing systems.” Ex. 1008, 2:44–46, 4:41–43, code (57); *see id.* at 2:38–40. Hou’s approaches “adaptively match microphone sensitivities” by adaptively measuring “a sensitivity difference between microphones” and then compensating (or correcting) “an electronic sound signal from one or more of the microphones.” *Id.* at 2:46–48, 4:43–44, 4:50–54, code (57). “As a result of the adaptive processing, the microphones ‘effectively’ become matched and remain matched over time and while in use.” *Id.* at 4:54–56; *see id.* at 2:48–50, 4:45–47, code (57). Hou’s approaches permit “multi-microphone directional processing systems to achieve superior directionality and consistent Signal-to-Noise Ratio (SNR) across all conditions.” *Id.* at 4:57–60, 9:38–41. Accordingly, the “directional noise suppression is robust.” *Id.* at 2:48, 4:44–45, code (57); *see id.* at 9:29–38.

Hou's Figure 5 (reproduced below) depicts a two-microphone directional processing system:

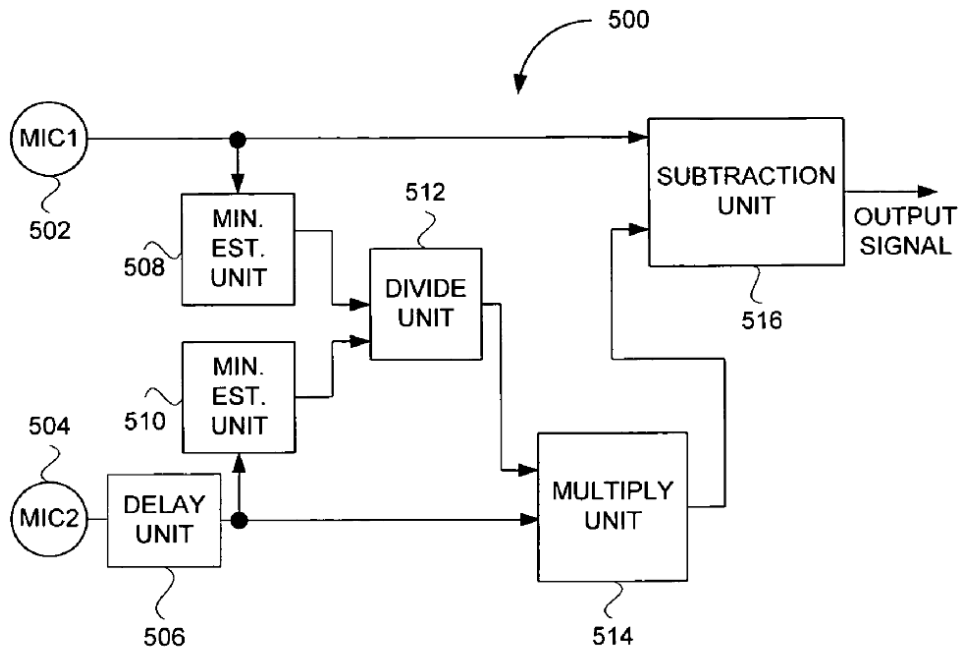


FIG. 5

Figure 5 illustrates two-microphone directional processing system 500 including the following components:

- (1) first microphone 502 that “produces a first electronic sound signal”;
- (2) second microphone 504 that “produces a second electronic sound signal”;
- (3) delay unit 506 that “delays the second electronic sound signal”;
- (4) first minimum estimate unit 508;
- (5) second minimum estimate unit 510;
- (6) divide unit 512;
- (7) multiplication unit 514; and

(8) subtraction unit 516.

Ex. 1008, 4:17–19, 5:26–36, 5:43–45, 5:50–52, Fig. 5.

First minimum estimate unit 508 “estimates the minimum for the first electronic sound signal.” Ex. 1008, 5:36–38. Second minimum estimate unit 510 “estimates the minimum of the second electronic sound signal.” *Id.* at 5:38–39. “Typically, these minimums are measured over a time constant duration, such that the minimum is a relatively long-term minimum.” *Id.* at 5:39–41.

Divide unit 512 “produces a quotient by dividing the first minimum estimate by the second minimum estimate.” Ex. 1008, 5:42–43. The “quotient represents a scaling amount that is sent to” multiplication unit 514. *Id.* at 5:43–45. “The second electronic sound signal is then multiplied with the scaling amount to produce a compensated sound signal.” *Id.* at 5:45–47. “The compensated sound signal is thus compensated (or corrected) for the relative difference in sensitivity between the mismatched first and second microphones 502 and 504.” *Id.* at 5:47–50.

Subtraction unit 516 “subtracts the compensated electronic sound signal from the first electronic sound signal to produce an output signal.” Ex. 1008, 5:50–52, Fig. 5. Thus, “the output signal has been processed by the two-microphone directional processing system 500 to have robust directionality despite a mismatch between the first and second microphones 502 and 504.” *Id.* at 5:52–56.

2. DIFFERENCES BETWEEN THE CLAIMED INVENTION AND THE PRIOR ART

(a) Claim 17

Claim 17 reads as follows:

17. The method of claim 13, wherein the forming of the filter comprises:

generating a first quantity by applying a calibration to the second signal;

generating a second quantity by applying the delay to the first signal;

forming the filter as a ratio of the first quantity to the second quantity.

Ex. 1001, 37:54–61.

(b) Petitioner's Arguments Regarding Claim 17

For “generating a first quantity” according to claim 17, Petitioner asserts that “Hou’s ‘first minimum estimate unit 508’ would receive Avendano’s second signal x_2 (e.g., from Avendano’s second physical microphone 108), and apply a calibration to the second signal x_2 (e.g., estimate a minimum of the second signal) to generate a first quantity.” Pet. 66 (emphases omitted) (citing Ex. 1003 ¶¶ 295–299; Ex. 1008, 5:38–39, Fig. 5).

For “generating a second quantity” according to claim 17, Petitioner asserts that “Hou’s ‘delay unit 506’ would receive Avendano’s first signal x_1 (e.g., from Avendano’s first physical microphone 106), and apply a delay to that first signal x_1 to generate a second quantity.” Pet. 67 (emphasis omitted) (citing Ex. 1003 ¶ 298; Ex. 1008, 5:36–38, Fig. 5).

For “forming the filter” according to claim 17, Petitioner asserts that “Hou’s ‘divide unit 512’ forms a filter as a ratio of the first quantity to the

second quantity” since “divide unit 512 produces a quotient by dividing the first minimum estimate by the second minimum estimate.” Pet. 68 (citing Ex. 1008, 5:42–45, Fig. 5).

To support its assertions, Petitioner provides an annotated version of Hou’s Figure 5 as reproduced below (Pet. 69):

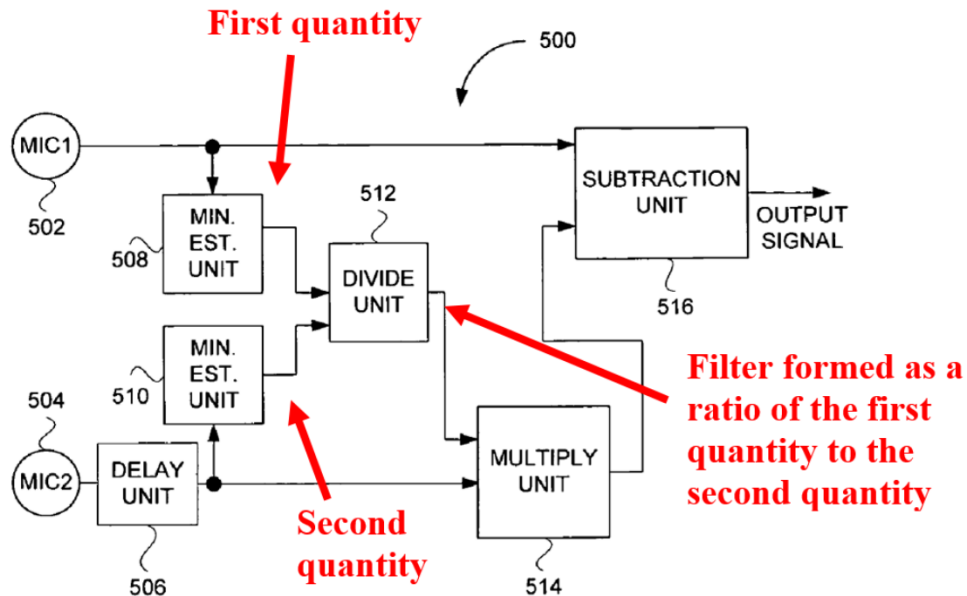


Figure 5 illustrates two-microphone directional processing system 500 including multiple components. Ex. 1008, 4:17–19, 5:26–36, 5:43–45, 5:50–52, Fig. 5; *see* Ex. 1003 ¶ 83. Petitioner’s annotated version of Figure 5 includes the following annotations:

- (1) the red annotation “First quantity” identifying the output of first minimum estimate unit 508, which receives Avendano’s second signal x_2 as an input in the proposed combination;
- (2) the red annotation “Second quantity” identifying the output of second minimum estimate unit 510, which receives a delayed version of Avendano’s first signal x_1 as an input in the proposed combination; and

- (3) the red annotation “Filter formed as a ratio of the first quantity to the second quantity” identifying the output of divide unit 512.

See Pet. 69; Ex. 1003 ¶ 299.

(c) Claims 18 and 19

Claim 18 depends directly from claim 17 and recites “wherein the generating of the energy ratio comprises generating the energy ratio for a frequency band.” Ex. 1001, 37:62–64. Claim 19 depends directly from claim 17 and recites “wherein the generating of the energy ratio comprises generating the energy ratio for a frequency subband.” *Id.* at 37:65–67.

(d) Petitioner’s Arguments Regarding Claims 18 and 19

For claims 18 and 19, Petitioner asserts that “Avendano’s ‘frequency analysis module 304 takes the cardioid signals and mimics the frequency analysis of the cochlea (i.e., cochlear domain) simulated by a filter bank.’” Pet. 69 (quoting Ex. 1005, 4:57–59). Petitioner also asserts that “Avendano’s ‘energy module 306’ calculates the energy ratio (‘ILD’) based on signals that are output by ‘frequency analysis module 304.’” *Id.* at 70 (citing Ex. 1005, 5:3–10). Petitioner further asserts that “the energy ratio is calculated ‘based on bandwidth of the cochlea channel’ (e.g., frequency band(s) or subband(s) for which ‘frequency analysis module 304 . . . mimics the frequency analysis of the cochlea’).” *Id.* (alteration by Petitioner) (emphasis omitted) (quoting Ex. 1005, 4:57–59, 5:6–7) (citing Ex. 1003 ¶¶ 303, 309).

(e) Patent Owner’s Arguments Regarding Claims 17–19

Patent Owner contends that Petitioner “has not shown that any claim is rendered obvious” over Avendano, Visser, Bisgaard, and Hou because (1) “Avendano and Visser fail to disclose or render obvious” limitations in

claim 1 and (2) Petitioner “does not argue that Bisgaard or Hou remedies these deficiencies.” Prelim. Resp. 13.

(f) Analysis

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Kenny’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano, Visser, Bisgaard, and Hou teach the subject matter in claims 17–19. *See* Pet. 58–60, 66–70; Ex. 1003 ¶¶ 82–91, 295–311, 313.

3. ALLEGED REASONS FOR COMBINING
THE TEACHINGS OF THE REFERENCES

Petitioner identifies a reason that would have prompted an ordinarily skilled artisan to combine the teachings of Avendano, Visser, Bisgaard, and Hou. *See* Pet. 61–65. Additionally, Petitioner explains that an ordinarily skilled artisan “would have had a reasonable expectation of success” in combining the teachings of the references. *Id.* at 65–66. Dr. Kenny’s testimony supports Petitioner’s positions. *See* Ex. 1003 ¶¶ 280–294.

For instance, Dr. Kenny testifies that an ordinarily skilled artisan “would have been motivated to modify Avendano’s system to include Hou’s signal compensation or correction components to compensate or correct for the relative difference in sensitivity between mismatched first and second microphones in order to generate a signal ‘hav[ing] robust directionality’ to enhance ‘directional noise suppression.’” Ex. 1003 ¶ 289 (alteration by Dr. Kenny) (quoting Ex. 1008, 2:46–48, 5:54–55).

At this stage of the proceeding, Patent Owner makes no arguments against combining the teachings of Avendano, Visser, Bisgaard, and Hou.

See, e.g., Prelim. Resp. 13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record, Petitioner establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had a reason, e.g., as articulated by Dr. Kenny, to combine the teachings of Avendano, Visser, Bisgaard, and Hou. *See* Pet. 61–65; Ex. 1003

¶¶ 280–292. Petitioner also establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had a reasonable expectation of success in combining the teachings of the references. *See* Pet. 65–66; Ex. 1003 ¶¶ 293–294.

4. PRELIMINARY CONCLUSION ABOUT OBVIOUSNESS/NONOBVIOUSNESS

As discussed above, Petitioner’s analysis addresses how the combined disclosures in Avendano, Visser, Bisgaard, and Hou teach the subject matter in claims 17–19. *See supra* § IV.F.2. Additionally, Petitioner provides a reason with rational underpinning as to why an ordinarily skilled artisan would have been motivated to combine the teachings of Avendano, Visser, Bisgaard, and Hou and would have had a reasonable expectation of success. *See supra* § IV.F.3. Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claims 17–19 unpatentable under § 103(a) as obvious over Avendano, Visser, Bisgaard, and Hou.

G. Alleged Obviousness over Avendano, Visser, Bisgaard, Hou, and Either Byrne, Burnett, or Berglund: Claims 20–22

Petitioner contends that claims 20–22 are unpatentable under § 103(a) as obvious over Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund. *See* Pet. 1, 71–91. Patent Owner disputes Petitioner’s contentions. *See* Prelim. Resp. 13. Above, we provided overviews of

Avendano, Visser, Bisgaard, and Hou. *See supra* §§ IV.D.1–IV.D.2, IV.E.1, IV.F.1. Below, we provide overviews of Byrne, Burnett, and Berglund, and then we consider the obviousness issues. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund teach the subject matter in claims 20–22.

1. OVERVIEW OF BYRNE (EXHIBIT 1009)

Byrne is a paper titled “An International Comparison of Long-Term Average Speech Spectra” by Denis Byrne et al., published in *The Journal of the Acoustical Society of America*, Volume 96, Number 4 (October 1994). Ex. 1009, 2108;⁵ *see* Ex. 1014 ¶¶ 6–8.

Byrne describes a study that investigated the “long-term average speech spectrum (LTASS) and some dynamic characteristics of speech” for twelve languages: “English (several dialects), Swedish, Danish, German, French (Canadian), Japanese, Cantonese, Mandarin, Russian, Welsh, Singhalese, and Vietnamese.” Ex. 1009, 2108; *see id.* at 2110–11. “The research strategy was to record and analyze samples of a wide variety of languages and dialects” according to a standardized protocol for “at least ten male and ten female” speakers of each language and dialect. *Id.* at 2110–11. The study determined that “the LTASS is very similar over the wide range of languages that were analyzed” and that “there is no single language or group of languages which could be regarded as being markedly different from the others.” *Id.* at 2114.

⁵ For Exhibit 1009 (Byrne), we follow Petitioner’s practice and cite to the page numbers that appear in the publication rather than the page numbers that Petitioner applied to the exhibit.

Byrne's Figure 1 (reproduced below) depicts study data for male and female speakers of five samples of English dialects, i.e., Australia, New Zealand, USA Columbus, USA Memphis, and United Kingdom:

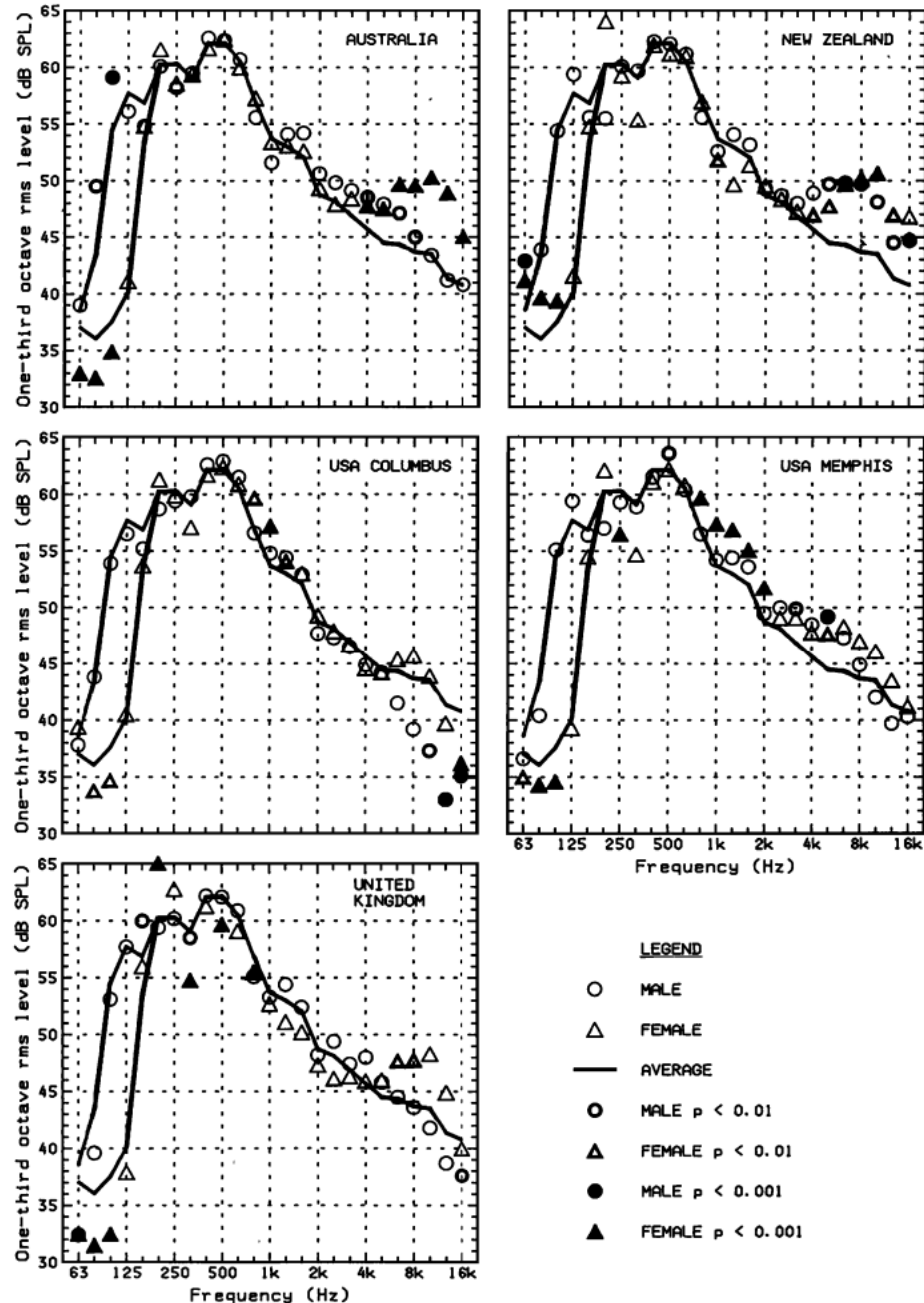


FIG. 1. Male and female long-term average speech spectrum (LTASS) values for five samples of English. Solid line shows LTASS average across 17 speech samples (all samples except Arabic), males and females separately for frequencies below 160 Hz, combined for higher frequencies.

Figure 1 illustrates male and female long-term average speech spectrum (LTASS) values in decibels for five samples of English dialects for frequencies between about 63 Hz and about 1600 Hz. Ex. 1009, 2113.

Figure 1 shows that the LTASS values peaked at approximately 500 Hz. *Id.*

2. OVERVIEW OF BURNETT (EXHIBIT 1012)

Burnett is a U.S. patent application publication titled “Detecting Voiced and Unvoiced Speech Using Both Acoustic and Nonacoustic Sensors,” filed on May 30, 2002, and published on December 26, 2002. Ex. 1012, codes (12), (22), (43), (54). Burnett states that the “disclosed embodiments relate to the processing of speech signals.” *Id.* ¶ 2.

Burnett discloses systems for discriminating or separating “voiced and unvoiced speech” from “background acoustic noise.” Ex. 1012 ¶¶ 18, 63, code (57). The systems “receive acoustic signals at two microphones, and generate difference parameters between the acoustic signals received at each of the two microphones.” *Id.* at code (57). The systems “identify information of the acoustic signals as unvoiced speech when the difference parameters exceed a first threshold, and identify information of the acoustic signals as voiced speech when the difference parameters exceed a second threshold.” *Id.* Additionally, the systems may “bandpass the information from Mic 1 and Mic 2 so that it is possible to see which bands in the Mic 1 data are more heavily composed of noise and which are more weighted with speech.” *Id.* ¶ 63.

Burnett discloses the following example frequency bands:

- (1) 500 Hz to 4000 Hz corresponding to “k” in “kick”;
- (2) 1700 Hz to 4000 Hz corresponding to “sh” in “she”;
- (3) 300 Hz to 2500 Hz corresponding to “/i/ (‘ee’)”; and

(4) 900 Hz to 1200 Hz corresponding to “/a/ (‘ah’).”

Ex. 1012 ¶ 64.

3. OVERVIEW OF BERGLUND (EXHIBIT 1013)

Berglund is a paper titled “Sources and Effects of Low-Frequency Noise” by Birgitta Berglund et al., published in The Journal of the Acoustical Society of America, Volume 99, Number 5 (May 1996).

Ex. 1013, 2985;⁶ *see* Ex. 1014 ¶¶ 9–11.

Berglund describes a study that investigated the “sources of human exposure to low-frequency noise” and the “physical, physiological, and psychological effects on humans.” Ex. 1013, 2985. Berglund explains that the “effects of low-frequency noise are of particular concern because of its pervasiveness due to numerous sources, efficient propagation, and reduced efficacy of many structures (dwellings, walls, and hearing protection) in attenuating low-frequency noise compared with other noise.” *Id.*

Berglund’s Figure 1 (reproduced below) depicts a frequency spectrum from 0 Hz to more than 20,000 Hz:

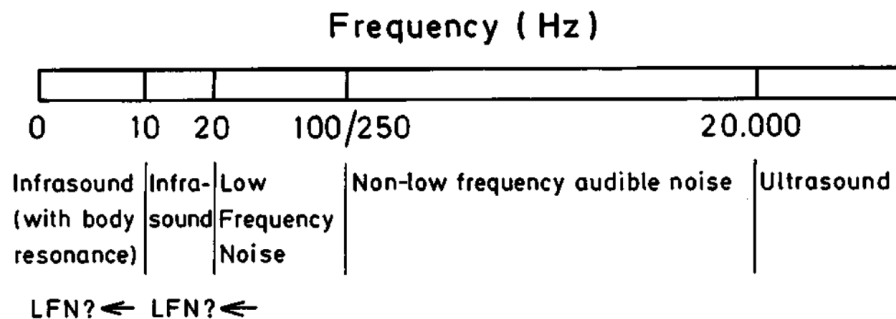


FIG. 1. The frequency spectrum of sound and its nomenclature.

⁶ For Exhibit 1013 (Berglund), we follow Petitioner’s practice and cite to the page numbers that appear in the publication rather than the page numbers that Petitioner applied to the exhibit.

Figure 1 illustrates the frequency spectrum of sound and its nomenclature, i.e., starting on the left with infrasound, then low-frequency noise, then non-low-frequency audible noise, and ending on the right with ultrasound.

Ex. 1013, 2986. Figure 1 includes a dividing line between low-frequency noise and non-low-frequency audible noise at “100/250” Hz. *Id.* Berglund treats “low-frequency noise as including what is normally taken to be infrasound.” *Id.* at 2985.

Berglund’s Figure 3 (reproduced below) depicts study data for low-frequency noise from road transportation vehicles:

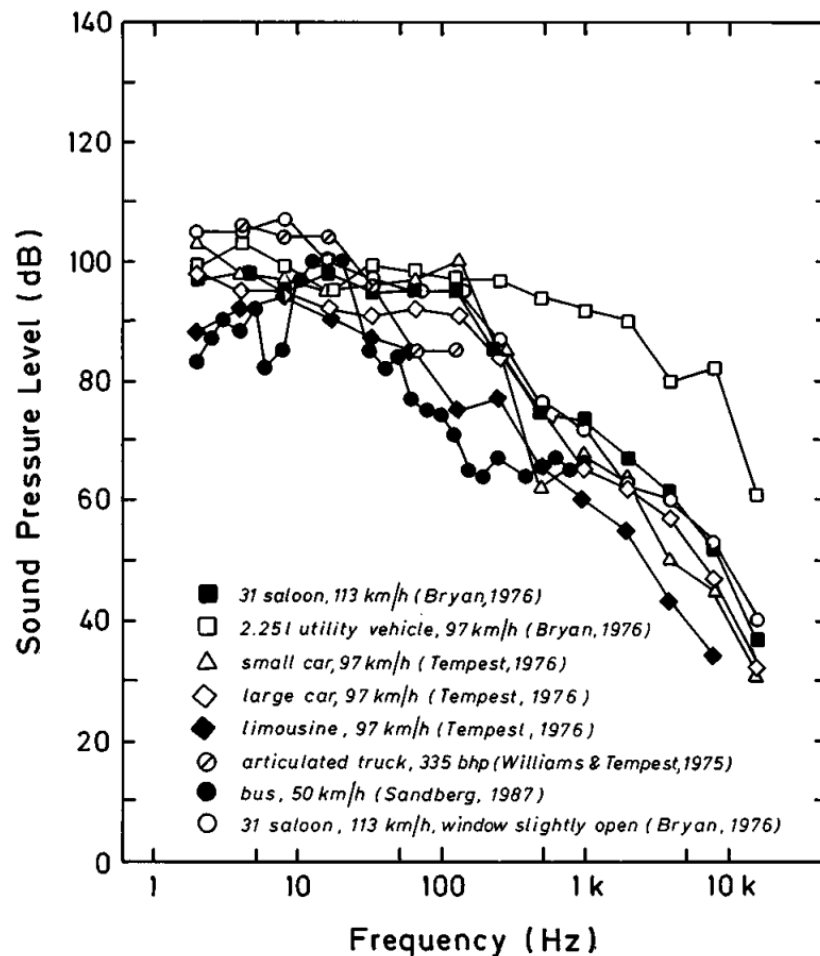


FIG. 3. Passenger noise exposure in road transportation vehicles as a function of frequency.

Figure 3 illustrates passenger noise exposure or sound-pressure level (in dB) as a function of frequency (in Hz) from road transportation vehicles, e.g., a small car, a large car, an articulated truck, and a bus. Ex. 1013, 2987.

4. DIFFERENCES BETWEEN THE CLAIMED INVENTION AND THE PRIOR ART

Each of claims 20–22 depends directly from claim 19 and specifies frequencies for the frequency subband. Ex. 1001, 38:1–9. Claim 20 recites “wherein the frequency subband includes frequencies higher than approximately 200 Hertz (Hz).” *Id.* at 38:1–3. Claim 21 recites “wherein the frequency subband includes frequencies in a range from approximately 250 Hz to 1250 Hz.” *Id.* at 38:4–6. Claim 22 recites “wherein the frequency subband includes frequencies in a range from approximately 200 Hz to 3000 Hz.” *Id.* at 38:7–9.

Petitioner asserts that an ordinarily skilled artisan “would have found it obvious to perform the processes described by Avendano, Visser, Bisgaard, and Hou in frequency range(s) known to have significant spectral components of speech, while filtering out frequency range(s) that are known to have significant spectral components of noise.” Pet. 77; *see id.* at 75–76, 85–91.

Regarding spectral components of speech, Petitioner asserts that Byrne teaches that “the spectral components of English speech are predominantly in a frequency range greater than 200 Hz,” such as ranges with “frequencies between 250-1250 Hz” and “frequencies between 200-3000 Hz.” Pet. 75, 78–79, 85–86, 88–89 (citing Ex. 1003 ¶¶ 314, 326, 337–338, 352; Ex. 1009, 2113 (Fig. 1)); *see id.* at 71–72. Petitioner also asserts that these ranges “include the frequencies of sound expected to have

the highest intensities for human speech,” e.g., “frequencies of and around 500 Hertz.” *Id.* at 85, 88 (citing Ex. 1003 ¶¶ 337–338, 352; Ex. 1009, 2113 (Fig. 1)).

Petitioner asserts that Burnett teaches that “voiced and unvoiced speech” includes “frequencies higher than approximately 200 Hz,” such as speech “in a frequency range of 250-1250 Hz” and speech “in a frequency range of 200-3000 Hz.” Pet. 80–81, 86, 90 (citing Ex. 1012 ¶ 64); *see id.* at 73. For speech “in a frequency range of 250-1250 Hz,” Petitioner identifies Burnett’s disclosure that the sound “ah” has a frequency range of 900-1200 Hz. *Id.* at 86 (citing Ex. 1012 ¶ 64). For speech “in a frequency range of 200-3000 Hz,” Petitioner identifies Burnett’s disclosure that the sound “ee” has a frequency range of 300-2500 Hz. *Id.* at 90 (citing Ex. 1012 ¶ 64).

Regarding spectral components of noise, Petitioner asserts that Berglund teaches that “the spectral components of common noises experienced by passengers of ‘road transportation vehicles’ is highest in the ‘low frequency noise’ range,” e.g., “less than approximately 250 Hz.” Pet. 80 (citing Ex. 1003 ¶ 327; Ex. 1013, 2985–87); *see id.* at 73–74, 87, 90 (citing Ex. 1003 ¶¶ 98–100, 341, 356; Ex. 1013, 2985–87).

Patent Owner contends that Petitioner “has not shown that any claim is rendered obvious” over Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund because (1) “Avendano and Visser fail to disclose or render obvious” limitations in claim 1 and (2) Petitioner “does not argue” that Bisgaard, Hou, Byrne, Burnett, or Berglund “remedies these deficiencies.” Prelim. Resp. 13.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Kenny's testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund teach the subject matter in claims 20–22. *See* Pet. 71–74, 77–91; Ex. 1003 ¶¶ 92–100, 324–359.

5. ALLEGED REASONS FOR COMBINING THE TEACHINGS OF THE REFERENCES

Petitioner identifies a reason that would have prompted an ordinarily skilled artisan to combine the teachings of Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund. *See* Pet. 75–77. Dr. Kenny's testimony supports Petitioner's positions. *See* Ex. 1003 ¶¶ 312–323, 329, 342, 351, 357.

For instance, Dr. Kenny testifies that an ordinarily skilled artisan “would have recognized that the processes described by Avendano (and in combination with Visser, Bisgaard, and Hou) are specifically configured to analyze audible speech in specific frequency range(s).” Ex. 1003 ¶ 314. Dr. Kenny also testifies that an ordinarily skilled artisan “would have found it obvious to”:

- (1) perform these processes “specifically in frequency range(s) that are known to have significant spectral components of speech (e.g., as taught by Byrne)”;
- (2) perform these processes “specifically in frequency range(s) known to have ‘voiced and unvoiced speech’ in order to improve the separation of ‘voiced and unvoiced speech from background acoustic noise’” as taught by Burnett; and

- (3) “filter out frequency range(s) that are known to have significant spectral components of noise (e.g., as taught by Berglund).”

Id. ¶¶ 314–315, 329, 351 (emphases omitted) (quoting Ex. 1012 ¶ 63).

Additionally, Dr. Kenny testifies that an ordinarily skilled artisan “would have recognized that the time and/or computation resources required to perform these processes can be reduced by limiting analysis to a subset of all frequencies,” such as “a frequency range between approximately 250 Hz and 1250 Hz” or “a frequency range between approximately 200 Hz and 3000 Hz,” to “improve the speed, accuracy and/or efficiency in detecting voice activity.” Ex. 1003 ¶¶ 342, 357.

At this stage of the proceeding, Patent Owner makes no arguments against combining the teachings of Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund. *See, e.g.*, Prelim. Resp. 13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record, Petitioner establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had a reason, e.g., as articulated by Dr. Kenny, to combine the teachings of Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund. *See* Pet. 75–77; Ex. 1003 ¶¶ 312–323, 329, 342, 351, 357.

6. PRELIMINARY CONCLUSION ABOUT OBVIOUSNESS/NONOBVIOUSNESS

As discussed above, Petitioner’s analysis addresses how the combined disclosures in Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund teach the subject matter in claims 20–22. *See supra* § IV.G.4. Additionally, Petitioner provides a reason with rational underpinning as to why an ordinarily skilled artisan would have been motivated to combine the

teachings of Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund. *See supra* § IV.G.5. Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claims 20–22 unpatentable under § 103(a) as obvious over Avendano, Visser, Bisgaard, Hou, and either Byrne, Burnett, or Berglund.

V. CONCLUSION

Based on the arguments and evidence presented by the parties, we determine that there is a reasonable likelihood Petitioner would prevail with respect to at least one claim challenged in the Petition. Hence, we institute an *inter partes* review of all challenged claims on all challenges included in the Petition. *See SAS Inst. Inc. v. Iancu*, 138 S. Ct. 1348, 1355 (2018) (noting that the language of 35 U.S.C. § 314(b) “indicates a binary choice—either institute review or don’t”); *see also PGS Geophysical AS v. Iancu*, 891 F.3d 1354, 1360 (Fed. Cir. 2018) (interpreting the statute as requiring “a simple yes-or-no institution choice respecting a petition, embracing all challenges included in the petition”).

VI. ORDER

Accordingly, it is

ORDERED that, pursuant to 35 U.S.C. § 314(a), an *inter partes* review of claims 1–28 in the ’611 patent is instituted on all challenges included in the Petition;

ORDERED that Petitioner’s Motion for Joinder (Paper 3) is denied as moot; and

FURTHER ORDERED that, according to 35 U.S.C. § 314(c) and 37 C.F.R. § 42.4, notice is hereby given of the institution of a trial that commences on the entry date of this Decision.

IPR2023-00286
Patent 8,326,611 B2

For PETITIONER:

Colin B. Heideman
Joseph R. Re
Daniel P. Hughes
Nathan D. Reeves
Logan P. Young
KNOBBE, MARTENS, OLSON & BEAR, LLP
2cbh@knobbe.com
2jrr@knobbe.com
2dph@knobbe.com
2ndr@knobbe.com
2lpy@knobbe.com

For PATENT OWNER:

Peter Lambrianakos
Vincent J. Rubino, III
Enrique W. Iturralde
Richard Cowell
FABRICANT LLP
plambrianakos@fabricantllp.com
vrubino@fabricantllp.com
eiturralde@fabricantllp.com
rcowell@fabricantllp.com