UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC., SONY ELECTONICS INC., and
LG ELECTRONICS INC.,

Petitioners,

v.

JAWBONE INNOVATIONS, LLC,

Patent Owner.

Patent No. 8,326,611
Filing Date: October 26, 2009
Issue Date: December 4, 2012

Inventor: Nicolas Petit, Gregory Burnett, and Zhinian Jing
Title: ACOUSTIC VOICE ACTIVITY DETECTION (AVAD)
FOR ELECTRONIC SYSTEMS

PATENT OWNER'S RESPONSE

Case No. IPR2023-01128

TABLE OF CONTENTS

Page(s)

I. INTRODUCTION 1

II. THE '611 PATENT 1

III. THE RELEVANT ALLEGED PRIOR ART 5

A. Avendano (U.S. Patent No. 8,194,880) 5

B. Visser (U.S. Patent No. 7,464,029) 8

C. Avendano/Visser Combination 12

IV. CLAIM CONSTRUCTION 13

A. "filter that describes a relationship for speech between the first physical microphone and the second physical microphone" [Claim 1] 14

V. LEVEL OF ORDINARY SKILL IN THE ART 22

VI. ARGUMENT 22

A. Legal Standard 22

B. GROUND 1: The Combination of Avendano and Visser Does Not Render Obvious Claims 1-28 24

1. Petitioner's Combination Does Not Disclose or Render Obvious "forming a filter that describes a relationship for speech between the first physical microphone and the second physical microphone," as Required by All Challenged Claims 24

a) Avendano's Delay Node Does Not Render Obvious Forming the Claimed Filter 24

b) A POSITA Would Not Make the Avendano/Visser Combination 26

2. The Avendano/Visser Combination Does Not Render Claim 7 Obvious Because it Does Not Disclose or Render Obvious "applying a delay to the first intermediate signal" 30

C. GROUND 2: Avendano, Visser, and Bisgaard Do Not Render Obvious Claims 8-16, 23-24 32

1. The Avendano/Visser/Bisgaard Combination Does Not Remedy Ground 1's Deficiencies as to Claim 7.....33

2. The Avendano/Visser/Bisgaard Combination Does Not Disclose or Render Obvious Claim 9. 34

D. GROUND 3: Avendano, Visser, Bisgaard, and Hou Do Not Render Obvious Claims 17-19 36

E. GROUND 4: Avendano, Visser, Bisgaard, Hou, and "Frequency Art" Do Not Render Obvious Claims 20-22 36

VII. CONCLUSION.....37

TABLE OF AUTHORITIES

Page(s)

Cases

Apple Inc. v. Samsung Elecs. Co.,
839 F.3d 1034 (Fed. Cir. 2016) 23

August Tech. Corp. v. Camtek, Ltd.,
655 F.3d 1278 (Fed. Cir. 2011) 21

Forest Lab'ys, LLC v. Sigmapharm Lab'ys, LLC,
918 F.3d 928 (Fed. Cir. 2019) 23

Graham v. John Deere Co. of Kansas City,
383 U.S. 1 (1966).....22, 23

KSR Int'l Co. v. Teleflex Inc.,
550 U.S. 398 (2007).....23

PSN Illinois, LLC v. Ivoclar Vivadent, Inc.,
525 F.3d 1159 (Fed. Cir. 2008) 21

Realtime Data, LLC v. Iancu,
912 F.3d 1368 (Fed. Cir. 2019) 14

Techtronic Indus. Co., v. Int'l Trade Comm'n,
944 F.3d 901 (Fed. Cir. 2019) 18, 19

Ventana Med. Sys., Inc. v. Biogenex Labs., Inc.,
473 F.3d 1173 (Fed. Cir. 2008) 20

Statutes

35 U.S.C. § 103 22

EXHIBIT LIST

Exhibit No.	Description of Document
2001	Joint Motion to Dismiss, *Jawbone Innovations, LLC v. Apple Inc.,* No. 6:21-cv-00984-ADA, Dkt. 110 (W.D. Tex. Dec. 23, 2022)
2002	Declaration of Akbar M. Sayeed, Ph.D.
2003	CV of Akbar M. Sayeed, Ph.D.
2004	Transcript of the Deposition of Thomas W. Kenny, Ph.D., taken April 19, 2024
2005	Second Declaration of Akbar M. Sayeed, Ph.D.

I. INTRODUCTION

The Board should reject Meta's hindsight-based obviousness arguments. Meta's primary argument, based on Avendano, requires imputing into the claims of the '611 Patent the exact time and phase based processing that the '611 Patent inventors rejected and improved upon.

Meta's secondary Visser argument would require a POSITA to take Visser's cross-filters, which explicitly rely on a VAD to adapt, remove the VAD, and insert them into Avendano's system. Without the VAD, the filters will continuously adapt, resulting in a combination with more false readings than Avendano alone. Meta's insertion of VAD-reliant components into Avendano indicates that it manufactured its argument based on a hindsight effort to meet the claims.

Finally, Meta's analysis of the dependent claims is internally inconsistent, and conflicts with the clear language of the claims.

Accordingly, the Board should issue a Final Written Decision finding all the Challenged Claims not unpatentable.

II. THE '611 PATENT

The '611 Patent discloses and claims methods for acoustic voice activity detection in microphone arrays. The claimed inventions "use microphones to generate virtual directional microphones which have very similar noise responses and very dissimilar speech responses." Furthermore, 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. The virtual microphones can be constructed using either an adaptive or a fixed filter". '611 Patent Abstract.; Ex. 2005, ¶ 30.

The claimed acoustic voice activity detection is based on Jawbone's PSAD (Pathfinder Speech Activity Detection) and NAVSAD (Non-Acoustic Sensor Voiced Speech Activity Detection) systems. '611 Patent, 4:27-30, 10:12-52; 11:1-29. The specification describes that one of the key advances of the technology is using the "difference in frequency content between two microphones to calculate a relationship between the signals of the two microphones." *Id.*, 11:19-21. The prior art, in contrast, "attempt[s] to use the time/phase difference of each microphone to remove the noise outside of an 'area of sensitivity.'" *Id.*, 11:22-24; Ex. 2005, ¶ 31

The advance on the prior art is illustrated in Figures 4 and 3 of the '611 Patent. There, the specification provides illustrative embodiments for creating the virtual microphone signals, V_1 and V_2, respectively, from two omnidirectional physical microphone signals (O_1 and O_2). Each embodiment includes: i) a calibration filter (α), ii) a filter that describes a relationship for speech (β), and iii) a delay filter (z).

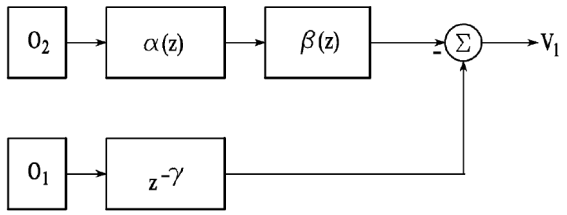


FIG.4

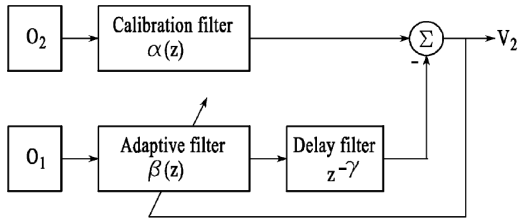


FIG.3

'611 Patent, FIGS. 3, 4; Ex. 2005, ¶ 32. '611 Patent, 5:29-33 ("α(z) is a calibration filter used to compensate O_2's response so that it is the same as O_1, β(z) is a filter that describes the relationship between O_1 and calibrated O_2 for speech, and γ is a fixed delay that depends on the size of the array.").

The '611 Patent explains that "[i]t is assumed that O_1 is closer to the speech source (the user's mouth)." *Id.*, 4:64-65. In the above Figures, α(z) and β(z) represent filters, where β(z) can be fixed or adaptive *Id.* 6:17-18. Ex. 2005, ¶ 33.

The relationship of V_1 and V_2 can be described in the digital frequency domain (z variable) as:

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

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

’611 Patent 5:23-35. In the above equations (and FIG. 3 and FIG. 4 from Ex 1001), “$\alpha(z)$ is a calibration filter used to compensate O_2’s response so that it is the same as O_1, $\beta(z)$ is a filter that describes the relationship between O_1 and calibrated O_2 for speech, and Y is a fixed delay that depends on the size of the array” *Id*. 5:29-33; Ex. 2005, ¶ 34.

Ideally the microphones O_1 and O_2 are assumed to be calibrated so that “their response to a source located the same distance away was identical for both amplitude and phase.” ’611 Patent, 28:55-57. However, this is not always feasible in practice and the calibration filter $\alpha(z)$ can be determined using adaptive algorithms, for example, so that in the presence of speech only signals (with no or low levels of noise), the microphone signals O_1 and O_2 satisfy the relationship $O_1(z) = \alpha(z)\, O_2(z)$. *Id.* 28:61-29:27; Ex. 2005, ¶ 35.

After the calibration filter $\alpha(z)$ is determined, the filter $\beta(z)$ can be determined, either from the geometry of microphones relative to the target speech source or, in other claimed embodiments, by using an adaptive filter algorithm to satisfy the relationship:

$$\tilde{\beta}(z) = \frac{a(z)O_2(z)}{z^{-\nu}O_1(z)}$$

’611 Patent 6:1-11; Ex. 2005, ¶ 36.

The Voice Activity Detection (VAD) is based on computing the ratio R:

$$R = \frac{\|V_1(z)\|}{\|V_2(z)\|} = \sqrt{\frac{\left(-\tilde{\beta}(z)\alpha(z)O_2(z) + O_1(z)z^{-v}\right)^2}{\left(\alpha(z)O_2(z) - \tilde{\beta}(z)O_1(z)z^{-v}\right)^2}}$$

where "double bar indicates norm and again any size window may be used. If $\tilde{\beta}(z)$ has been accurately calculated, 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)." 611 Patent 6:20-32. If the ratio R is above a threshold, speech is present, otherwise speech is absent. *Id.*, FIG. 5 (510). "Many techniques known to those skilled in the art (*e.g.*, smoothing, etc.) can be used to make R more amenable to use in generating a VAD and the embodiments herein are not so limited." *Id.*, 6:38-41; Ex. 2005, ¶ 37.

III. THE RELEVANT ALLEGED PRIOR ART

A. Avendano (U.S. Patent No. 8,194,880)

U.S. Patent No. 8,194,880 ("Avendano") (Ex. 1005) was filed on January 29, 2007, and issued on June 5, 2012.

Avendano describes "[s]ystems and methods for utilizing inter-microphone level differences (ILD) to attenuate noise and enhance speech." Ex. 1005, Abstract. In exemplary embodiments, "primary and secondary acoustic signals are received by omni-directional microphones, and converted into primary and secondary electric signals". *Id.* Further, "[a] differential microphone array module processes the

electric signals to determine a cardioid primary signal and a cardioid secondary signal." *Id.* The cardioid signals are "filtered through a frequency analysis module" and "[e]nergy levels of the signals are then computed, and the results are processed by an ILD module using a non-linear combination to obtain the ILD." In exemplary embodiments, "the non-linear combination comprises dividing the energy level associated with the primary microphone by the energy level associated with the secondary microphone. The ILD is utilized by a noise reduction system to enhance the speech of the primary acoustic signal." *Id.* Ex. 2005, ¶ 60.

FIG. 4a in Avendano represents an illustrative embodiment utilized in the Petition and the Declaration of Dr. Kenny.

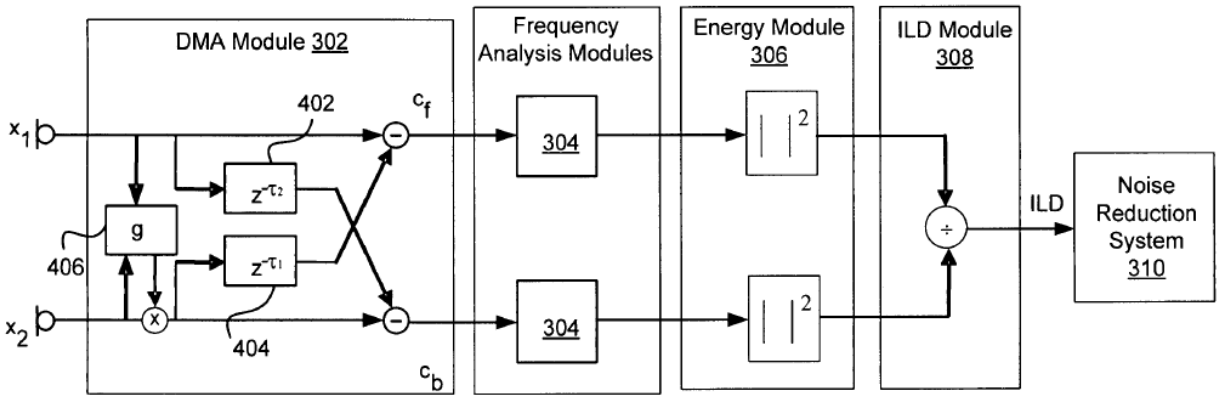


FIG. 4a

'611 Patent, FIG. 4a; Pet., 11; Ex. 1003, ¶ 60; Ex. 2005, ¶ 61.

Avendano uses two "delay nodes" 402 and 404 to delay the primary and secondary acoustic signals. Ex. 1005, 5:15-24. The gain module 406 computes a gain

factor to equalize the signal levels. *Id.*, 5:36-39. The two subtraction units then create cardioid primary (C_f) and cardioid secondary (C_b) signals. *Id.*, 5:25-35; Ex. 2005, ¶ 62. The C_f signal is the cardioid signal in the forward direction. Ex. 2004, 33:24-34:6. The C_b is the cardioid backward signal. Ex. 1005, 34:7-11. As a result of this processing, the C_f and C_b signals are strongly directional, with the C_f having a strong response in the forward direction and a null in the backward direction, and the C_b signal having a strong response in the backward direction and a null in the forward direction. *Id.*, 35:11-36:11. *Id.*, FIG. 6; Ex. 2005, ¶¶ 62-63.

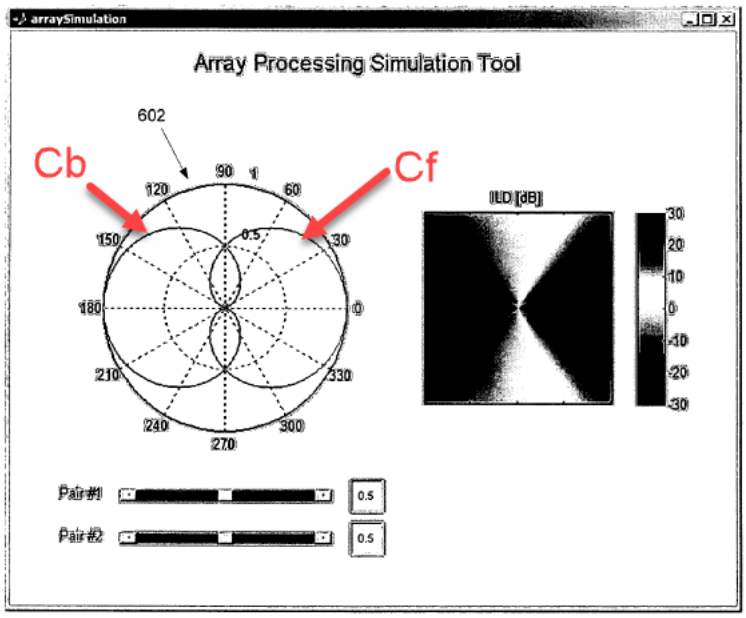


FIG. 6

Ex. 1005, FIG. 6 (annotations added).

Following the DMA Module, the signals are processed through a frequency analysis module 304. *Id.*, 5:40-48. The frequency analysis module may filter the

signals. *Id.* The system then calculates the power estimates associated with each signal, and then uses those estimates to generate the inter-microphone level differences. *Id.*, 6:7-20; Ex. 2005, ¶ 64.

B. Visser (U.S. Patent No. 7,464,029)

U.S. Patent No. 7,464,029 ("Visser") (Ex. 1006) was filed on July 22, 2005, and issued on December 9, 2008. Visser discusses "a signal separation process [that] is associated with a voice activity detector." Ex. 1006, Abstract. Notably, "[t]he voice activity detector is a two-channel detector, which enables a particularly robust and accurate detection of voice activity." *Id.* Visser uses Independent Component Analysis (ICA) to separate speech signals from noise sources. *Id.*, 3:3-20. Visser's method is also described as blind source separation. *Id.* Ex. 2004, 49:9-50:6; Ex. 2005, ¶ 66.

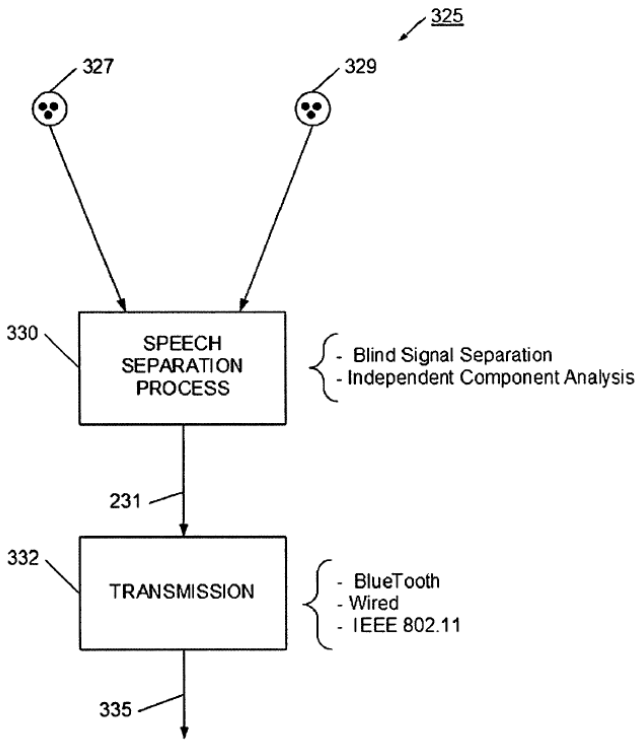


FIG. 5

Visser FIG. 5

While Petitioner mentions Visser's voice activity detector (VAD), it does not rely on the VAD or argue that it would be part of its combination. Pet., 18; Ex. 2004, 61:14-63:6; Ex. 2005, ¶ 67.

Visser explains that its VAD "receives two input signals 105, with one of the signals defined to hold a stronger speech signal." Ex. 1006, 8:31-35. When the VAD detects speech, it "generates a control signal 107." *Id.*, 8:36-38. As Visser explains, the VAD "enables a more stable and effective speech separation process." *Id.*, 8:45-47; Ex. 2005, ¶ 68.

The Petition states that Visser's VAD takes as input "processed signals" that include "(i) a 'speech channel' that 'contains predominantly desired signals,' and (ii) a 'noise channel' that 'contains predominantly noise signals' (*e.g.*, generated by 'ICA or BSS processing')." Pet., 18. Dr. Kenny states that "[t]he 'speech channel' and the 'noise channel' are provided as inputs to the 'VAD 106' (*e.g.*, the 'input signals 105') to determine 'when speech is present.'" Ex. 1003, ¶ 74; Ex. 2005, ¶ 69.

Both the Petition and Dr. Kenny leave out a critical portion of Visser. Visser explains that input signals to the VAD are received directly from the microphones. Ex. 1006, FIGS. 2, 3; 11:27-12:1. The VAD then makes a comparison based on those signals as to whether speech is occurring. *Id.*, FIG. 3; Ex. 2005, ¶ 70.

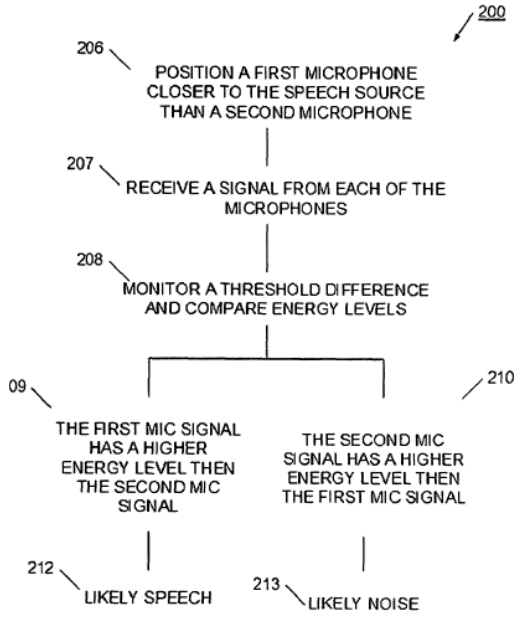


FIG. 3

Ex. 1006, FIG. 3.

After processing, the signal is then compared again, to form a two channel VAD that generates a control signal. *Id.*, 8:48-9:5, FIG. 4; Ex. 2005, ¶ 71.

The control signal is used to control multiple aspects of the remainder of the Visser system. Visser, 9:6-10:51. The control signal controls whether the speech separation process is active (Visser, 8:38-43), when the filters may adapt (Visser, 8:40-43, 9:7-21, 9:29-32), and other aspects of Visser's post processing. Ex. 2005, ¶ 72.

In the relevant embodiment, Visser uses cross-filters w_{21} and w_{12} in a f to separate noise and speech. The filters are adaptive filters. Ex. 1006, 17:29-18:29; Ex. 2005, ¶ 73.

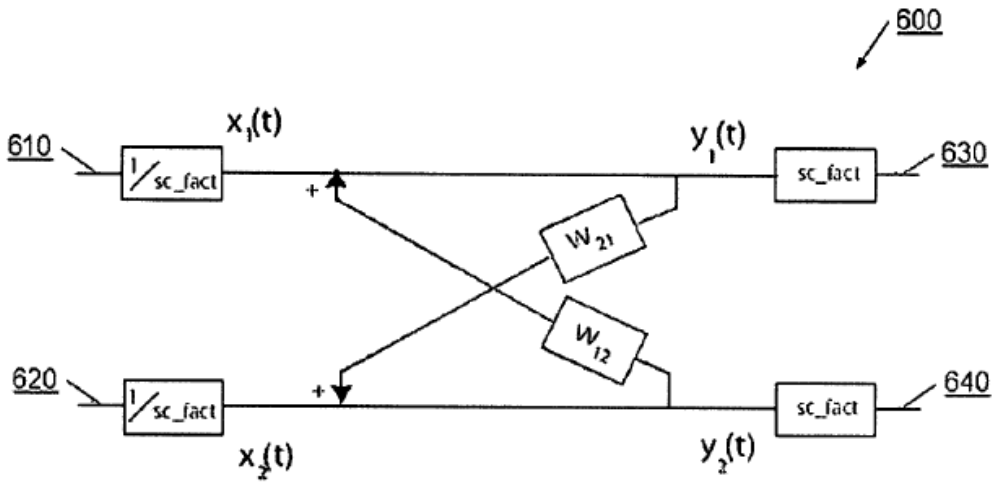


FIG. 10

In FIG. 10, input signals X_1 and X_2 are received from the input channels 610 and 620. Ex. 1006, 17:29-50. The filters w_{21} and w_{12} are applied to each signal to produce the channels of separated signals. *Id.*; Ex. 2005, ¶ 74.

C. Avendano/Visser Combination

Petitioner asserts that a POSITA would "incorporate Visser's 'ICA process' into Avendano's system to further enhance the differences in speech content in Avendano's two processed signals, such that speech is detected with a greater degree of accuracy, 'even in a changing acoustic environment.'" Pet., 19; Ex. 2005, ¶ 75.

Petitioner illustrates its combination with a modified FIG. 4a from Avendano:

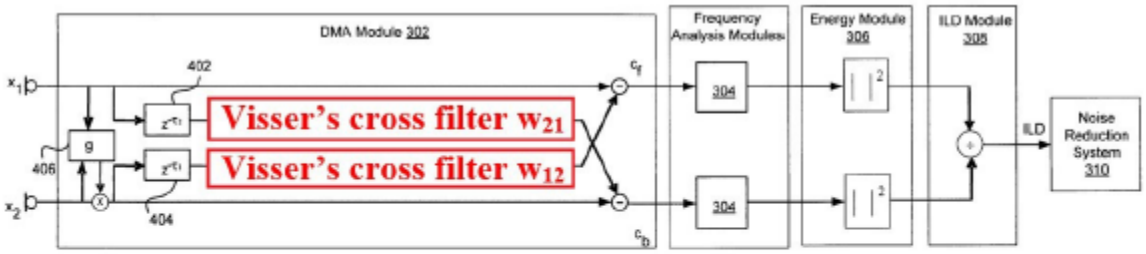


FIG. 4a

Ex. 1005, FIG. 4a (modified)

Pet., 22; Ex. 2005, ¶ 76. Petitioner's expert confirmed that, other than the W_{21} and W_{12} filters and possible connections between them, no other portion of Visser would be incorporated into the combination. Ex. 2004, 61:14-63:6; Ex. 2005, ¶ 77.

In Petitioner's combination, microphones are arranged as they are in Avendano. Ex. 2004, 64:15-24. Thus, the signals are received from two

microphones: one that is nearer to a speaker and one that is further away. Ex. 1005, 3:27-35; Ex. 2004, 37:8-19; Ex. 2005, ¶ 78.

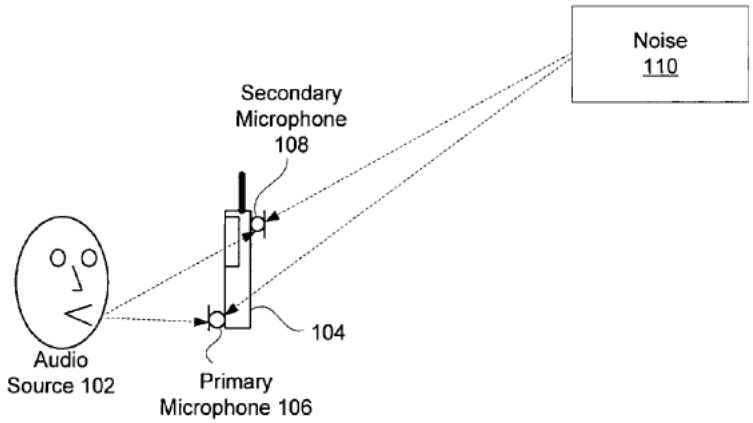


FIG. 1a

IV. CLAIM CONSTRUCTION

Petitioner asserts that no construction is necessary to resolve the issues in this proceeding. Pet., 7. However, as discussed below, the proper construction of the key term "filter that describes a relationship for speech between the first physical microphone and the second physical microphone" is required to resolve Petitioner's invalidity arguments.

A. "filter that describes a relationship for speech between the first physical microphone and the second physical microphone" [Claim 1]

Jawbone's Construction	Petitioner's Construction
"filter that alters a signal based on the non-temporal components of relative responses to speech of the first and second physical microphones"	No construction necessary.

The key dispute as to this term is whether it encompasses a simple "delay filter" or requires more. To that end, Jawbone has revised its construction of the term "filter that describes a relationship for speech between the first physical microphone and the second physical microphone" to "filter that alters a signal based on the non-temporal components of relative responses to speech of the first and second physical microphones." *Realtime Data, LLC v. Iancu*, 912 F.3d 1368 (Fed. Cir. 2019) ("The Board is required to construe 'only those terms . . . that are in controversy, and only to the extent necessary to resolve the controversy.'"). On institution, the Board declined to construe this term[1] because it did not find lexicography or disavowal in the preliminary record. Paper 13 ("Inst. Dec."), 17.

[1] Jawbone's construction in its Preliminary response was "[a] filter that alters the magnitude of a signal based on the relative responses to speech of the first and second physical microphones." Paper 10 (:POPR"), 10. While Jawbone believes its

However, the specification clearly disavows the concept of a delay filter, alone, constituting the "filter that describes a relationship for speech" (hereinafter, "relationship filter"). Instead, the specification explains that an advance over the prior art is avoiding using "the time/phase difference of each microphone." '611 Patent, 11:22-23.

> The NAVSAD and PSAD systems include an array algorithm for speech detection **that uses the difference in frequency content between two microphones to calculate a *relationship between the signals of the two microphones*.** This is in contrast to conventional arrays that attempt to use the time/phase difference of each microphone to remove the noise outside of an "area of sensitivity". ***The methods described herein provide a significant advantage, as they do not require a specific orientation of the array with respect to the signal.***

'611 Patent, 11:18-27 (emphasis added). A POSITA would thus understand that a key advance and advantage is to move away from a delay-based method. Ex. 2005, ¶¶ 41-43.

The specification describes several advantages in moving away from delay-based processing. For example, "[t]he systems described herein are sensitive to noise

---

preliminary construction was and is correct, it proposes this revised construction to narrow the dispute before the Board during trial. Ex. 2005, ¶¶ 39-40.

of every type and every orientation, unlike conventional arrays that depend on specific noise orientations." '611 Patent, 11:28-30. "This results in a robust signal processing system with respect to the type of noise, microphones, and orientation between the noise/signal source and the microphones." *Id.*, 11:35-37; Ex. 2005, ¶ 44.

To achieve its goal of improving on time/phase processing, the specification describes three different filters, a calibration filter ($\alpha(z)$), a delay filter ($z^{-\gamma}$), and the filter that describes a relationship for speech ($\beta(z)$). '611 Patent, 5:20-51, FIG. 3, FIG. 4. The delay γ is "a fixed delay that depends on the size of the array." '611 Patent, 5:32-33; Ex. 2005, ¶¶ 44-45.

FIG. 3 illustrates the relationship:

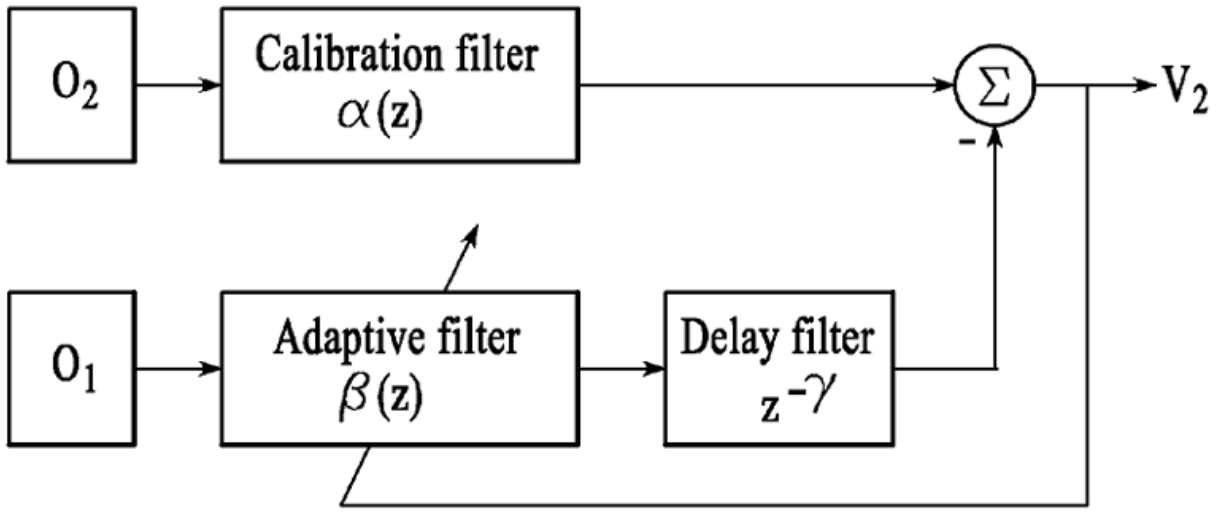


FIG.3

The specification further specifically identifies β(z) as the filter describing the relationship for speech. '611 Patent, 5:31-33 ("β(z) is a filter that describes the relationship between O_1 and calibrated O_2 for speech"). The specification provides two methods to calculate β(z): 1) as a function of the distance between each microphone and the user's mouth (*id.*, 5:52-59); and 2) adaptively (*id.*, 6:1-16). Ex. 2005, ¶¶ 46-47.

$$\beta(z) = \frac{d_1}{d_2} = \frac{d_1}{d_1 + d} \qquad \tilde{\beta}(z) = \frac{a(z)O_2(z)}{z^{-\gamma}O_1(z)}$$

In either case, the β(z) compensates for the difference in response to speech of the two microphones, not the temporal difference. By increasing or decreasing the magnitude of the filtered signal, the filter adjusts the signal to account for the relative strengths of the signal received—*i.e.*, the closer microphone is attenuated to the same speech level received at the farther microphone. Indeed, the ***separate*** delay, which is defined in terms of distance and the speed of sound, compensates for the temporal difference based on the distance between the microphones. '611 Patent, 5:35-45. Ex. 2005, ¶¶ 48-49.

$$\gamma = \frac{d}{c}$$

The constant d is the distance between the two microphones. '611 Patent, 5:43-44; 4:66-67 ("If the microphones and the user's mouth are all on a line, then $d=2d_0$, the distance between the microphones."). The constant c is "the speed of sound in air." *Id.*, 5:43-44; Ex. 2005, ¶ 50.

Thus, as explained above, every embodiment in the specification treats the relationship filter as different from, and performing a different purpose than, the delay filter. Moreover, including functionality beyond mere delay is one of the key advances, and provides the advantage that the array does not require "a specific orientation." '611 Patent, 11:26-30; Ex. 2005, ¶¶ 52-53. The above facts constitute a clear and unmistakable disavowal of using a delay filter as the claimed relationship filter. *Techtronic Indus. Co., v. Int'l Trade Comm'n*, 944 F.3d 901, 907 (Fed. Cir. 2019) ("A disavowal must be clear, but it need not be explicit.").

The *Techtronic* decision is instructive. There, the specification lacked an express disavowal. *Id.* However, the specification's description of a prior art's feature (there, the location of a detector in the head unit of a garage door opener) as a problem to be overcome, plus the patent's own disclosure of overcoming those problems, constituted a disavowal. *Id.* at 907-09. The patent's consistent disclosure of every embodiment lacking that feature further supported the implied disavowal. *Id.*

The situation here is similar. Just as in *Techtronic*, every embodiment uses a β filter that is different from the delay filter. Moreover, the patent describes delay-only processing as deficient, and explains how it overcomes the problems of such a system. Ex. 2005, ¶ 54.

In its Preliminary Reply, Petitioner argued that Jawbone's construction would exclude disclosed embodiments. Paper 11 ("Reply"), 2-3. But the only so-called embodiments Petitioner identified are the delay filter and calibration filter. *Id.* Neither of these filters are embodiments of the ***invention***; they are merely ***components*** of embodiments. Indeed, as discussed above, ***every*** disclosed embodiment includes a β filter. Jawbone's construction does not exclude any embodiments, as it allows for the unclaimed delay or calibration filters. Ex. 2005, ¶¶ 55-56.

> The second problem with BioGenex's argument is that each claim does not necessarily cover every feature disclosed in the specification. ***When the claim addresses only some of the features disclosed in the specification, it is improper to limit the claim to other, unclaimed features.*** Here, the claims of the '861 patent are directed toward automated methods of dispensing that use bar codes to determine which reagents should be dispensed onto which slides. ***Although the preferred embodiments also contain a 'direct dispensing' feature, the inventors were not***

> ***required to claim this feature in the '861 patent and, indeed, did not do so.***

Ventana Med. Sys., Inc. v. Biogenex Labs., Inc., 473 F.3d 1173, 1181 (Fed. Cir. 2008) (emphasis added and citations omitted).

Moreover, dependent claims add in the same delay and calibration that Petitioner argues would be excluded. Claims 7 and 32 each apply a delay to the intermediate signal. Similarly, other claims cover the calibration filter α. For example, dependent Claims 6 and 31 add limitations that cover a calibration applied to at least one of the first and second signals. These are the exact embodiments disclosed in the specification and shown in FIG. 3. Ex. 2005, ¶¶ 56-57.

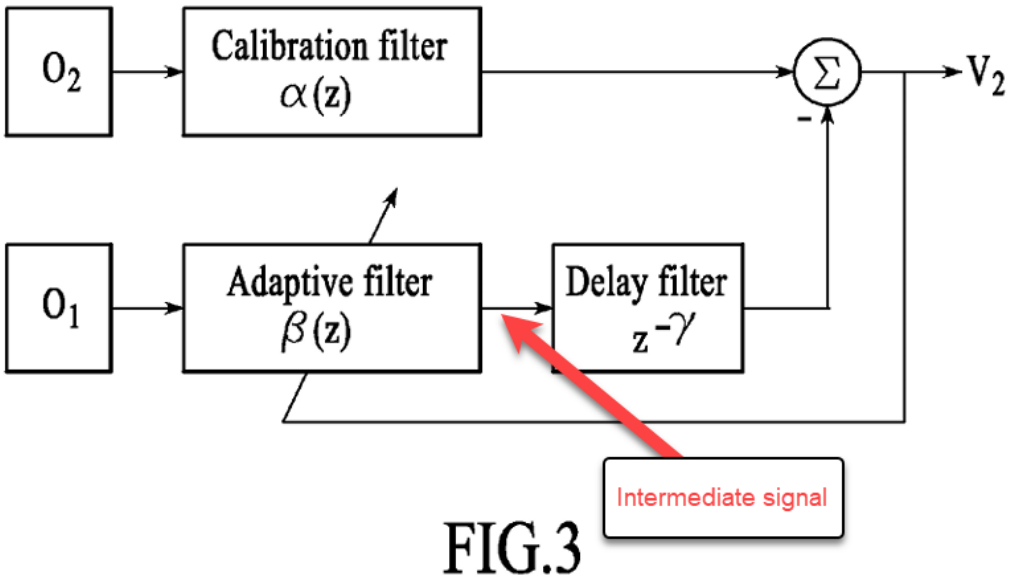


FIG.3

'611 Patent, FIG. 3 (annotations added). In view of these claims, a POSITA would understand that the relationship filter was different than the calibration and delay

filters. *PSN Illinois, LLC v. Ivoclar Vivadent, Inc.*, 525 F.3d 1159, 1166 (Fed. Cir. 2008) ("[C]ourts must recognize that disclosed embodiments may be within the scope of other allowed but unasserted claims. . . . In such cases, unasserted or cancelled claims may provide 'probative evidence' that an embodiment is not within the scope of an asserted claim."). The Federal Circuit has construed claims similarly in these circumstances. For example, in *August Tech. Corp. v. Camtek, Ltd.*, 655 F.3d 1278 (Fed. Cir. 2011), the Federal Circuit reversed a construction of "plurality of wafers" that would have encompassed a single wafer. *Id.* at 1285-86. The specification disclosed both single and multiple wafer embodiments, but "[t]he fact that the claims at issue cover only the latter—a plurality of known good wafers—is little cause for concern." *Id.* at 1285. "This is especially true where, as here, other unasserted claims in the parent patent cover the excluded embodiments." *Id.* The circumstances present a much less difficult question; the claims, under Jawbone's construction, cover all embodiments in the specification. They simply give effect to the inventor's clear distinction between the relationship filter and the calibration or delay filters.

A POSITA, reading the claim in light of the specification, and with the explicit description that the invention improves over delay-based prior art, would therefore understand that the "filter that describes a relationship for speech between the first physical microphone and the second physical microphone" refers to "a filter that

alters a signal based on the non-temporal components of relative responses to speech of the first and second physical microphones." Ex. 2005, ¶ 58. At the very least, a POSITA would understand that a delay filter alone could not meet the limitation.

Accordingly, the Board should adopt Jawbone's construction.

V. LEVEL OF ORDINARY SKILL IN THE ART

A person of ordinary skill in the art would have a bachelor's degree in computer engineering, computer science, electrical engineering, mechanical engineering, or a similar field, and approximately three years of industry or academic experience in a field related to acoustics, speech recognition, speech detection, or signal processing. Work experience can substitute for formal education and additional formal education can substitute for work experience. Ex. 2005, ¶ 27.

VI. ARGUMENT

A. Legal Standard

The question of obviousness is resolved on the basis of underlying factual determinations, including: (1) the scope and content of the prior art, (2) any differences between the claimed subject matter and the prior art, (3) the level of skill in the art, and (4) where in evidence, so called secondary considerations. *Graham v. John Deere Co. of Kansas City*, 383 U.S. 1, 17–18 (1966). A claim is only unpatentable under 35 U.S.C. § 103 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." *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007) (quoting 35 U.S.C. § 103).

"An invention is not obvious simply because all of the claimed limitations were known in the prior art at the time of the invention. Instead, we ask 'whether there is a reason, suggestion, or motivation in the prior art that would lead one of ordinary skill in the art to combine the references, and that would also suggest a reasonable likelihood of success.'" *Forest Lab'ys, LLC v. Sigmapharm Lab'ys, LLC*, 918 F.3d 928, 934 (Fed. Cir. 2019). "Of course, concluding that the references are within the scope and content of the prior art to be considered for obviousness does not end the inquiry. *Graham* makes clear that the obviousness inquiry requires a determination whether the claimed invention would have been obvious to a skilled artisan." *Apple Inc. v. Samsung Elecs. Co.*, 839 F.3d 1034, 1050 n.14 (Fed. Cir. 2016).

B. GROUND 1: The Combination of Avendano and Visser Does Not Render Obvious Claims 1-28

1. Petitioner's Combination Does Not Disclose or Render Obvious "forming a filter that describes a relationship for speech between the first physical microphone and the second physical microphone," as Required by All Challenged Claims

All Challenged Claims[2] require "forming a filter that describes a relationship for speech between the first physical microphone and the second physical microphone."

a) Avendano's Delay Node Does Not Render Obvious Forming the Claimed Filter

Petitioner first argues that Avendano discloses this limitation via its "delay node 402." Pet. at 25-27. As explained above, under the proper construction, the claim requires "a filter that alters a signal based on the non-temporal components of relative responses to speech of the first and second physical microphones."

[2] Meta did not challenge Claims 3 and 4, but LG and Sony did. LG has been terminated, and as of filing, Sony is subject to a motion to terminate. Jawbone understands that Claims 3 and 4 are therefore no longer in this proceeding. To the extent the Board determines those claims are still under challenge, they are not obvious for the same reasons set forth herein as to Claim 1.

Avendano's delay node, even if implemented as an all-pass filter, does not meet this requirement. Ex. 2005, ¶81.

Indeed, the only relationship Petitioner identifies is "a temporal relationship for speech between the first and second signals." Pet., 26. Petitioner describes this relationship in nearly identical terms to the '611 Patent's delay filter:

> In Avendano, "due to a space difference between the microphones," and because the speed of sound propagation in air is widely known to be approximately 330 m/s, there will be a "difference in times of arrival of the signal from a speech source to the microphones."

Id.; Ex. 2005, ¶ 82.

Simple mathematics shows that the temporal relationship between the microphones is defined as the distance between them divided by the speed of sound in air. That is exactly the relationship handled by the separate delay filter in the '611 Patent. '611 Patent, 5:35-51. Ex. 2005, ¶ 83.

$$\gamma = \frac{d}{c}$$

'611 Patent, 5:41-45 ("where again $d=2d_0$ and c is the speed of sound in air"); 4:66-67 (defining $d=2d_0$ as the "distance between the microphones").

As discussed above, the inventor of the '611 Patent specifically distinguished his speech detection process from those based on delays. Thus, to meet the limitation, the filter must do more than merely delay the signal. Ex. 2005, ¶ 84.

In sum, the delay node merely delays a signal for a predetermined time. Ex. 2004, 27:24-29:22. For the reasons stated above, it is therefore not the claimed "filter that describes a relationship for speech between the first physical microphone and the second physical microphone."[3] Petitioner's obviousness arguments do not obviate this deficiency. Ex. 2005, ¶ 85.

Accordingly, Avendano does not disclose or render obvious this limitation.

b) A POSITA Would Not Make the Avendano/Visser Combination

Petitioner secondarily relies on its Avendano/Visser combination for this limitation. Pet., 28-29. However, a POSITA would not make the combination because it would provide no benefit and, indeed, would likely result in numerous errors. Ex. 2005, ¶ 88. Indeed, Petitioner's combination is based entirely on a hindsight effort to meet the claims.

---

[3] The Petition does not argue that Avendano's gain is the relationship filter. Pet., 25-28. However, Petitioner raised such a mapping in its Preliminary Reply. The gain is not the relationship filter because it does not describe a relationship for speech. Instead, it is simply a gain that matches signal strengths. Ex. 2005, ¶ 85.

At the outset, the proposed Avendano/Visser combination would not function properly. Visser relies on a separate VAD to control the ICA process on which Petitioner relies, particularly the adaptation of the W_{12} and W_{21} filters on which Petitioner relies. Ex. 1006, 9:6-32 ("In this way, the speech separation process only applies its adaptive learning processes when desired speech is likely occurring."); 9:29-32 ("By turning the ICA off when desired speech is not present, the ICA filters do not inappropriately adapt, thereby enabling adaptation only when such adaptation will be able to achieve a separation improvement."); 9:6-43; Ex. 2005, ¶¶ 89-90. Indeed, as Petitioner's expert explained, the W_{12} and W_{21} filters (shown in Visser FIG. 10) are part of Visser's signal separation process as shown in FIG. 1. Ex. 2004, 50:24-51:18.

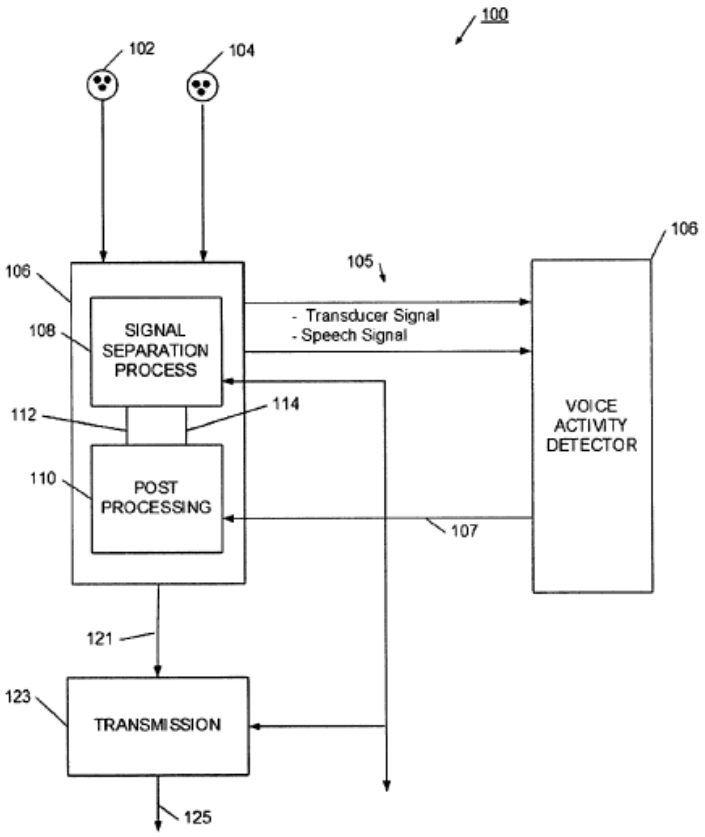


FIG. 1

Ex. 1006, FIG. 1.

The adaptation is controlled based on the input signals received from the microphones. Ex. 2005, ¶¶ 89-90.

Petitioner's combination does not include Visser's VAD. Ex. 2004, 59:5-61:4, 61:14-63:6; Ex. 1006, 9:25-29 ("This 'on' and 'off' state can be monitored and controlled by the voice activity detection module **185** based on comparing energy content between input channels or desired speaker a priori knowledge such as specific spectral signatures.").

Instead of using Visser's VAD, Petitioner proposes to use Visser's filters as part of the Avendano VAD. Pet., 19-23. Ex. 2004, 61:14-63:6; Ex. 2005, ¶ 91

As Dr. Sayeed explains, omitting Visser's VAD will cause the filters to adapt continuously, regardless of whether speech is occurring. This continuous adaptation will result in the W_{21} filter adapting to remove noise, rather than speech, from the first signal. When the system later subtracts the filtered first signal from the second signal, the C_b cardioid signal will include more speech content and less noise content than Avendano alone. Ex. 2005, ¶¶ 92-93.

Because the C_b cardioid signal includes more content from the X_1 microphone, it will no longer be strongly directional as shown in Avendano FIG. 6. This will lead to the signal becoming distorted, which will likely lead to false positives and negatives at the ILD stage. A POSITA would therefore recognize that introducing Visser's filters would result in a less accurate VAD than Avendano alone. Ex. 2005, ¶ 94.

Accordingly, a POSITA would not be motivated to include Visser's filters into Avendano. Ex. 2005, ¶¶ 95-96.

Instead of considering what a POSITA would do at the time of the invention, Petitioner and its expert performed a hindsight-based search to find the '611 Patent's functionality in various references. Ex. 2004, 71:20-73:5. Indeed, the Petition itself notes this hindsight bias when it states that "[t]o the extent that Avendano alone does

not describe each of the features of [1b] [*sic*], it would have been obvious to modify Avendano's system in view of Visser ***to include these features***." Pet., 32 (emphasis added).

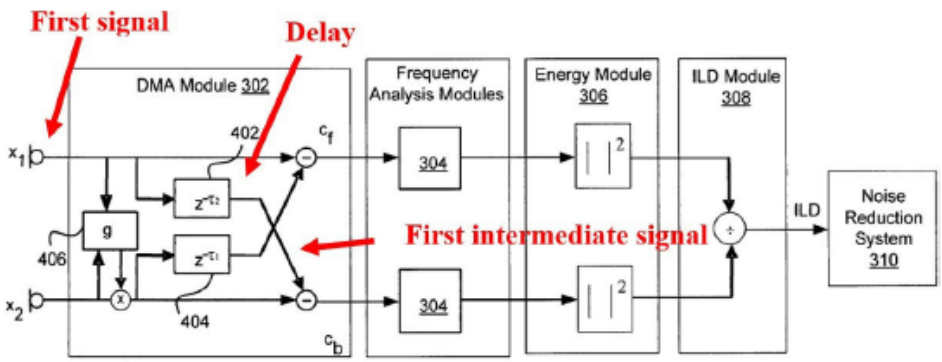
Therefore, Petitioner has not shown that Claim 1 (and its dependent claims) would have been obvious. The Board should issue a Final Written Decision finding all claims not shown unpatentable.

2. The Avendano/Visser Combination Does Not Render Claim 7 Obvious Because it Does Not Disclose or Render Obvious "applying a delay to the first intermediate signal"

Claim 7 depends from Claim 1 and adds the limitation "the method of claim 1, comprising applying a delay to the first intermediate signal." Petitioner argues that this limitation is disclosed by Avendano's delay node 402 or the Visser cross-filters. Pet., 37-38. Petitioner is wrong because neither Avendano alone nor Avendano/Visser apply any delay to the ***first intermediate signal***. Ex. 2005, ¶¶ 97-105.

As recited in Claim 1, the first intermediate signal is generated by "applying the filter to the first signal to generate a first intermediate signal." "The filter" is the relationship filter discussed above. The plain language of the claim states that the intermediate signal exists only ***after applying the filter***. Ex. 2005, ¶ 99.

However, under Petitioner's Avendano theory, the delay node ***is*** the "filter." Thus, the intermediate signal is the signal output from the delay filter and the delay is not applied to it. Petitioner's own annotated Figure demonstrates the contradiction in Petitioner's theory. Ex. 2005, ¶ 100.



Pet., 38.

As Petitioner admits, "Avendano's first intermediate signal is an output of 'delay node 402,' which is generated by applying 'delay node 402' (*i.e.,* a filter) to the first signal." Pet., 37. In other words, the delay is applied to the "first signal" ***not*** the first intermediate signal. Ex. 2005, ¶ 101.

Avendano therefore does not disclose this limitation. Petitioner does not suggest that it would be obvious to introduce a further delay after delay node 402. Pet., 36-38; Ex. 2005, ¶ 102.

Petitioner's Avendano/Visser combination has the same flaw, as shown in Petitioner's own Figure:

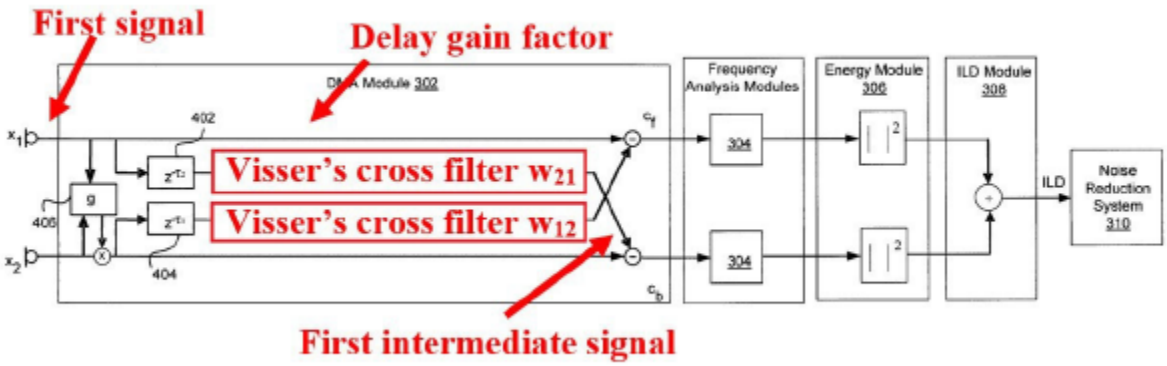


FIG. 4a

Pet., 39; Ex. 2005, ¶ 103.

Just as with Avendano alone, Petitioner has identified the first intermediate signal as the output from the Visser cross-filter. But that signal does not exist until after the cross-filter. Pet., 38-39; *see also* Pet., 33 ("According to the modification, the ***first intermediate signal is an output of 'cross filter w_{21},'*** which is generated by applying 'cross filter w_{21}' to a first signal x_1." (emphasis added)). Ex. 2005, ¶ 104.

Avendano thus does not disclose this limitation. Petitioner does not suggest that it would be obvious to introduce a further delay after the cross filter.

Accordingly, the combination of Avendano and Visser does not disclose Claim 7's limitation and does not render Claim 7 obvious. Ex. 2005, ¶¶ 105-106.

C. GROUND 2: Avendano, Visser, and Bisgaard Do Not Render Obvious Claims 8-16, 23-24

Claims 8-16 and 23-24 are dependent on Claim 7 and its parent Claim 1. The Petition relies on its Ground 1 arguments for the limitations of Claims 1 and 7.

The Petition does not argue that the addition of Bisgaard remedies its Ground 1 deficiencies as to Claim 1; Ex. 2005, ¶¶ 107-108.

1. The Avendano/Visser/Bisgaard Combination Does Not Remedy Ground 1's Deficiencies as to Claim 7

The Petition also does not argue that the addition of Bisgaard remedies its Ground 1 deficiencies as to Claim 7. In particular, Petitioner does not use Bisgaard to add a delay to the intermediate signal, but only to set the amount of delay introduced by either Avendano's delay node 402 or Visser's cross filter ***before the intermediate signal***.

> From this disclosure, and as an example of applying Bisgaard's teaching regarding delays to Avendano's system, **a POSITA would have found it obvious to configure each of Avendano's "delay nodes" to "delay[] the digitized sound signal"** received by each of Avendano's physical microphones "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," **as taught by Bisgaard** (*e.g.*, to account for the delay in the arrival of sound between the two microphones). Ex. 1010, [0042]; Ex. 1011, 5:22-24; Ex. 1003, ¶220.

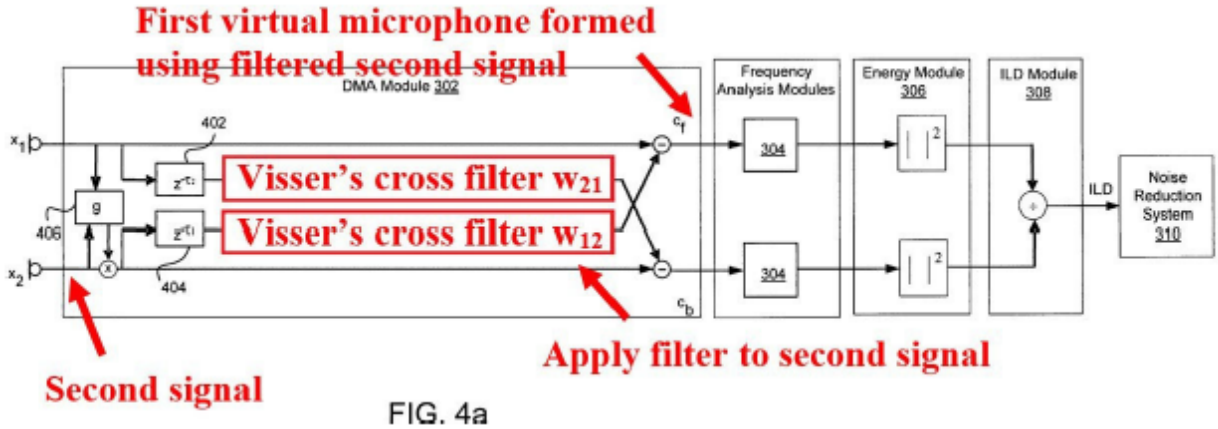
Pet., 44-45 (emphasis added). Dr. Kenny confirmed this under cross-examination. Ex. 2004, 75:14-78:2.

Therefore, introducing Bisgaard does not remedy the lack of a delay applied to the intermediate signal and Petitioner has not shown that Claim 7 would have been obvious. Ex. 2005, ¶¶ 109-110.

2. The Avendano/Visser/Bisgaard Combination Does Not Disclose or Render Obvious Claim 9.

Claim 9 depends from Claim 8 and adds the limitation: "The method of claim 8, wherein the forming of the first virtual microphone comprises applying the filter to the second signal."

Petitioner argues that its combination meets this limitation through Visser's cross filter W_{12} as incorporated into Avendano. Pet., 48. However, W_{12} is not "the filter" in Petitioner's combination. Ex. 2005, ¶¶ 111-22.



Pet., 48.

The claim requires "applying ***the filter*** to the second signal." As discussed above and as recited in Claim 1, "the filter" is the "filter that describes a relationship for speech between the first physical microphone and the second physical

microphone." In Petitioner's combination, the filter is applied to the first signal to generate the first intermediate signal. Claim 1 states this explicitly: "forming a second virtual microphone by applying the filter to the first signal to generate a first intermediate signal." Ex. 2005, ¶¶ 113-115.

Claim 9's additional limitation—"wherein the forming of the first virtual microphone comprises applying the filter to the second signal"—therefore requires that the ***same filter*** be applied to ***both*** the first and second signals. Ex. 2005, ¶ 116.

Under Petitioner's mapping, the W_{21} filter is applied to the first signal. Pet., 33 ("According to the modification, the first intermediate signal is an output of 'cross filter w_{21},' which is generated by ***applying 'cross filter w_{21}' to a first signal x_1.***" (emphasis added)). However, the separate W_{12} filter is applied to the second signal. Pet., 48 ("In the combination, Visser's cross filter w_{12} would be applied to Avendano's second signal x_2."). Ex. 2005, ¶¶ 117-118.

As a POSITA would recognize, the W_{12} and W_{21} filters are individual filters with separate purposes. Petitioner states that the W_{21} filter acts "to further ***decrease*** speech content in the second virtual microphone C_b." Pet., 22. In contrast, the W_{12} filter acts "to further ***increase*** speech content in the first virtual microphone C_f." *Id.* (emphasis added). A POSITA would understand that the two filters (even if interconnected as IIR filters) are separate filters and would have different coefficients. *See* Ex. 1006, 17:28-67; Ex. 2005, ¶ 119.

Petitioner makes no arguments that it would be obvious to apply the same filter to both signals.

Therefore, the Avendano/Visser/Bisgaard combination does not render Claim 9 obvious. In sum, no claims are rendered obvious by the Avendano/Visser/Bisgaard combination. Ex. 2005, ¶¶ 120-121.

D. GROUND 3: Avendano, Visser, Bisgaard, and Hou Do Not Render Obvious Claims 17-19

Claims 17-19 are dependent, through various parent claims, on Claims 9, 7, and 1. The Petitioner relies on Grounds 1-2 for the limitations of those claims. As shown above, however, Avendano and Visser fail to disclose or render obvious Claims 1 and 7, and Avendano/Visser/Bisgaard fails to render obvious Claim 9. Petitioner does not argue that Hou remedies these deficiencies. *See* Pet. at 57-68. Accordingly, Petitioner has not shown that any claim is rendered obvious under Ground 3. Ex. 2005, ¶¶ 123-126.

E. GROUND 4: Avendano, Visser, Bisgaard, Hou, and "Frequency Art" Do Not Render Obvious Claims 20-22

Claims 20-22 are dependent, through various parent claims, on Claims 9, 7, and 1. The Petition relies on its Grounds 1-3 arguments for the limitations of those claims.

As shown above, however, Avendano and Visser fail to disclose or render obvious Claim 1 and 7, and Avendano/Visser/Bisgaard fails to render obvious claim

9. Petitioner does not argue that the "Frequency Art" remedies these deficiencies. *See* Pet. at 68-89. Accordingly, Petitioner has not shown that any claim is rendered obvious under Ground 4. Ex. 2005, ¶¶ 127-130.

VII. CONCLUSION

As discussed above, Petitioner's combinations fail to disclose or render obvious key limitations of each Challenged Claim and, therefore, do not anticipate or render obvious any claims. The Board should therefore issue a Final Written Decision finding all Challenged Claims not shown to be unpatentable.

Respectfully submitted,

Dated: May 2, 2024

By: /*Peter Lambrianakos* /

Peter Lambrianakos (Reg. No. 58,279)
Lead Counsel for Patent Owner
FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Tel. 212-257-5797
Fax. 212-257-5796
Email: plambrianakos@fabricantllp.com

CERTIFICATE OF WORD COUNT

The undersigned hereby certifies that the portions of the above-captioned PATENT OWNER'S RESPONSE specified in 37 C.F.R. § 42.24 has 6,115 words in compliance with the 14,000 word limit set forth in 37 C.F.R. § 42.24. This word count was prepared using Microsoft Word for Office 365.

Respectfully submitted,

May 2, 2024 By: */Peter Lambrianakos /*

Peter Lambrianakos (Reg. No. 58,279)
Lead Counsel for Patent Owner
FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Tel. 212-257-5797
Fax. 212-257-5796
Email: plambrianakos@fabricantllp.com

CERTIFICATE OF SERVICE

A copy of the foregoing Patent Owner's Response and Exhibits 2004 and 2005 have been served on Petitioner's counsel of record as follows:

Grace I. Wang
Email: PH-Meta-Jawbone-IPR@paulhastings.com
PAUL HASTINGS LLP
200 Park Avenue
New York, New York 10166

Lisa K. Nguyen
Eric E. Lancaster (*Pro Hac Vice* to be filed)
Email: PH-Meta-Jawbone-IPR@paulhastings.com
PAUL HASTINGS LLP
1117 South California Avenue
Palo Alto, California 94303

David M. Tennant
Alan M. Billharz
Email: PH-Meta-Jawbone-IPR@paulhastings.com
PAUL HASTINGS LLP
2050 M Street, N.W.
Washington, D.C. 20036

Attorneys for Meta Platforms, Inc.

Gregory S. Gewirtz
Email: ggewirtz@lernerdavid.com
Email: litigation@lernerdavid.com
Russell W. Faegenburg
Email: rfaegenburg@lernerdavid.com
Orville R. Cockings
Email: ocockings@lernerdavid.com
James A. Vezeris
Email: jvezeris@lernerdavid.com

LERNER DAVID LLP
20 Commerce Drive
Cranford, New Jersey 07016

Attorneys for Sony Electronics Inc.

May 2, 2024

By: */Peter Lambrianakos /*

Peter Lambrianakos (Reg. No. 58,279)
Lead Counsel for Patent Owner
FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Tel. 212-257-5797
Fax. 212-257-5796