

Filed: June 15, 2023

Filed on behalf of Sony Electronics Inc.

By: Gregory S. Gewirtz
Russell W. Faegenburg
LERNER DAVID LLP
Telephone: 908.654.5000
Facsimile: 908.654.7866
Email: ggewirtz@lernerdavid.com
rfaegenburg@lernerdavid.com
litigation@lernerdavid.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SONY ELECTRONICS INC.,
Petitioner(s),

v.

JAWBONE INNOVATIONS, LLC,
Patent Owner.

IPR2023-01065
U.S. Patent No. 10,779,080

**PETITION FOR *INTER PARTES* REVIEW OF
U.S. PATENT NO. 10,779,080 PURSUANT TO
35 U.S.C. §§ 311-319, 37 C.F.R. § 42**

TABLE OF CONTENTS

	Page
TABLE OF AUTHORITIES	x
EXHIBIT LIST	xii
I. STATEMENT OF PRECISE RELIEF REQUESTED FOR EACH CLAIM CHALLENGED	1
II. OVERVIEW OF THE '080 PATENT	2
A. Specification	2
B. Prosecution History.....	7
III. THE LEVEL OF ORDINARY SKILL IN THE ART	8
IV. CLAIM CONSTRUCTION	8
V. SPECIFIC GROUNDS OF UNPATENTABILITY	9
A. Ground 1: Ikeda in View of McCowan and Kanamori Renders Obvious Claims 1-3, 5-9, 11-14, and 16-20	9
1. Overview of Ikeda.....	9
2. Overview of McCowan.....	13
a. McCowan is Prior Art.....	16
3. Overview Of Kanamori.....	17
4. The Ikeda-McCowan-Kanamori Combination	19
a. The Ikeda-McCowan-Kanamori main microphone has the same formulation for signal processing as the virtual microphone V1 described in the '080 patent, and the Ikeda-McCowan-Kanamori reference microphone has the same formulation for signal processing as the virtual microphone V2 described in the '080 patent	34

5. Claim 7	38
a. [7a] “A system, comprising: a first virtual microphone formed from a first combination of a first microphone signal and a second microphone signal, wherein the first microphone signal is generated by a first physical microphone and the second microphone signal is generated by a second physical microphone;”	38
b. [7b] “a second virtual microphone formed from a second combination of the first microphone signal and the second microphone signal, wherein the second combination is different from the first combination,”	42
c. [7c] “wherein the first virtual microphone has a first linear response to speech and first linear response to noise,”	44
d. [7d] “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source located within a predetermined angle relative to an axis of the microphone array and devoid of a null,”	46
e. [7e] “wherein the second virtual microphone has a second linear response to speech that has a single null oriented in a direction toward a source of the speech and a second linear response to noise,”	49
f. [7f] “wherein the second linear response to noise is substantially similar to the first linear response to noise,”	51
g. [7g] “one or both of the first linear response to noise and the second linear response to noise being nonzero in a direction toward a source of noise, and”	52
h. [7h] “the second linear response to speech is substantially dissimilar to the first linear response to speech,”	53
i. [7i] “wherein the speech is human speech; and”	54

- j. [7j] “an adaptive noise removal application coupled to the first and second virtual microphones and generating denoised output signals by forming a plurality of combinations of signals output from the first virtual microphone and the second virtual microphone, by filtering and summing the plurality of combinations of signals in the time domain, and by a varying linear transfer function between the plurality of combinations of signals,” 55
- k. [7k] “wherein the denoised output signals include less acoustic noise than acoustic signals received at the first and second physical microphones.”.....60
- 6. Claim 8: “The system of claim 7 and further comprising: a microphone array, the first and second physical microphones positioned in [sic] the microphone array.”.....61
- 7. Claim 9: “The system of claim 7, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.”.....61
- 8. Claim 11: “The system of claim 7 and further comprising: a communications channel coupled with the processing component and including one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel.” 62
- 9. Claim 12: “The system of claim 11 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.” 63
- 10. Claim 13: “The system of claim 7, wherein the second microphone signal is multiplied by a ratio, wherein the ratio is a ratio of a third distance to a fourth distance, the third distance being between the first physical microphone and the speech source and the fourth distance being between the second physical microphone and the speech source.” 64
- 11. Claim 14.....64

- a. [14a] “A system, comprising: a first virtual microphone comprising a first combination of a first microphone signal and a second microphone signal,” 64
- b. [14b] “the first virtual microphone having a first linear response to speech and a first linear response to noise,” 64
- c. [14c] “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source located within a predetermined angle relative to an axis of a microphone array,” 65
- d. [14d] “wherein the first microphone signal is output from a first physical microphone and the second microphone signal is output from a second physical microphone;” 65
- e. [14e] “a second virtual microphone comprising a second combination of the first microphone signal and the second microphone signal,” 65
- f. [14f] “the second virtual microphone having a second linear response to speech and a second linear response to noise,” 65
- g. [14g] “the second linear response to noise being substantially similar to the first linear response to noise,” 65
- h. [14h] “one or both of the first linear response to noise and the second linear response to noise being non- zero in a direction toward a source of noise, and” 66
- i. [14i] “the second linear response to speech being substantially dissimilar to the first linear response to speech,” 66
- j. [14j] “wherein the second combination is different from the first combination,” 66

- k. [14k] “wherein the first virtual microphone and the second virtual microphone are distinct virtual directional microphones; and” 66
- l. [14l] “a processing component coupled to the first and second virtual microphones, the processing component including an adaptive noise removal application receiving acoustic signals from the first virtual microphone and the second virtual microphone, filtering and summing the acoustic signals in the time domain, applying a varying linear transfer function between the acoustic signals, and generating an output signal,” 67
- m. [14m] “wherein the output signal is a denoised acoustic signal.” 68
- 12. Claim 16: “The system of claim 14 and further comprising: a communications channel coupled with the processing component and including one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel.” 69
- 13. Claim 17: “The system of claim 16 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.” 69
- 14. Claim 18: “The system of claim 14, wherein the acoustic signals from the first virtual microphone, the second virtual microphone or both are delayed.” 69
- 15. Claim 19: “The system of claim 18, wherein the delay is raised to a power that is proportional to a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone.” 71
- 16. Claim 20: “The system of claim 19, wherein the power is proportional to a sampling frequency multiplied by a quantity equal to a third distance subtracted from a fourth distance, the third distance being between a first physical microphone and the speech source, the fourth distance being between a second

physical microphone and the speech source, and the first and second physical microphones are positioned in the microphone array.”	74
17. Claim 1	77
a. [1a] “A system, comprising: a microphone array including a first physical microphone outputting a first microphone signal and a second physical microphone outputting a second microphone signal;”	77
b. [1b] “a processing component coupled to the microphone array and generating a virtual microphone array including a first virtual microphone and a second virtual microphone,”	77
c. [1c] “the first virtual microphone including a first combination of the first microphone signal and the second microphone signal, the second virtual microphone including a second combination of the first microphone signal and the second microphone signal, wherein the second combination is different from the first combination,”	78
d. [1d] “wherein the first virtual microphone and the second virtual microphone have substantially similar responses to noise and substantially dissimilar responses to speech; and”	78
e. [1e] “an adaptive noise removal application coupled to the processing component and generating denoised output signals by forming a plurality of combinations of signals output from the first virtual microphone and the second virtual microphone, by filtering and summing the plurality of combinations of signals in the time domain, and by a varying linear transfer function between the plurality of combinations of signals,”	79

f. [1f] “wherein the denoised output signals include less acoustic noise than acoustic signals received at the microphone array.”	79
18. Claim 2: “The system of claim 1, wherein the acoustic noise comprises noise content and the acoustic signals comprise speech content.”	79
19. Claim 3: “The system of claim 2, wherein the speech content comprises human speech.”	80
20. Claim 5: “The system of claim 1 and further comprising: a communications channel coupled with the processing component and including one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel.”	80
21. Claim 6: “The system of claim 5 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.”	81
B. Ground 2: Ikeda in view of McCowan, Kanamori, and Yang Renders Obvious Claims 4, 10, and 15.....	81
1. Overview of Yang.....	81
2. Claim 10: “The system of claim 7 and further comprising: a voice activity detector (VAD) coupled with the processing component and operative to generate voice activity signals.”	82
3. Claim 15: “The system of claim 14 and further comprising: a voice activity detector (VAD) coupled with the processing component and operative to generate voice activity signals.”	83
4. Claim 4: “The system of claim 1 and further comprising: a voice activity detector (VAD) coupled with the processing component and operative to generate voice activity signals.”	84
VI. INSTITUTION IS APPROPRIATE HERE	84

A.	Co-Pending Litigation Does Not Warrant Discretionary Denial Under § 314(A) (<i>Fintiv</i>).....	84
1.	Factor 1: Potential Stay	84
2.	Factor 2: Proximity of Trial to FWD	86
3.	Factor 3: Investment in Parallel Proceeding	86
4.	Factor 4: Overlapping Issues	87
5.	Factor 5: Parties in Parallel Proceedings	87
6.	Factor 6: Other Circumstances.....	88
B.	The Existence of Previously Filed IPR Petitions Does Not Warrant Discretionary Denial (<i>General Plastic</i>).....	88
C.	Discretionary Denial Under § 325(D) is Also Not Appropriate.....	90
VII.	MANDATORY NOTICES UNDER 37 C.F.R. § 42.8.....	90
A.	Real Party-in-Interest.....	90
B.	Related Matters	90
C.	Lead and Back-Up Counsel	92
D.	Service Information	92
E.	Payment of Fees Pursuant to 37 C.F.R. § 42.103	92

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Apple Inc. v. Fintiv, Inc.</i> , IPR2020-00019, Paper 11 (PTAB Mar. 20, 2020)	84, 89
<i>Cal. Inst. of Tech. v. Broadcom Ltd.</i> , 25 F.4th 976 (Fed. Cir. 2022)	87
<i>Celltrion, Inc. v. Genentech, Inc.</i> , IPR2018-01019, Paper 11 (PTAB Oct. 30, 2018)	88
<i>Central Security Group</i> , IPR2019-01609, Paper 11	88
<i>Ericsson Inc. v. Intellectual Ventures I LLC</i> , IPR2014-00527, Paper 41 (P.T.A.B. May 18, 2015)	17
<i>GAF Materials LLC v. Kirsch Research and Dev., LLC</i> , IPR2021-00192, Paper 14 (PTAB May 25, 2021)	86
<i>Global Tel*Link Corp. v. HLFIP Holding, Inc.</i> , IPR2021- 00444	86
<i>Google</i> , IPR2022-00630, Paper 13	88, 89, 90
<i>In re GPAC Inc.</i> , 57 F.3d 1573 (Fed. Cir. 1995)	8
<i>Huawei Techs. Co. v. WSOU Invs., LLC</i> , IPR2021-00226, Paper 10 (PTAB June 10, 2021)	87
<i>Jawbone Innovations, LLC v. Sony Electronics Inc.</i> , No. 2:23-cv-01161-CCC-LDW (D.N.J.)	<i>passim</i>
<i>Kirsch Rsch. & Dev., LLC v. GAF Materials, LLC</i> , Civil Action No. 20-13683 (JMV), 2021 U.S. Dist. LEXIS 111667 (D.N.J. June 15, 2021)	85

Sony Electronics Inc. v. Jawbone Innovation, LLC
U.S. Patent No. 10,779,080

<i>Mercedes-Benz USA, LLC v. Carucel Invs. L.P.</i> , IPR2019-01404, Paper 12 (PTAB Jan. 22, 2020)	89
<i>Microsoft Corp. v. IPA Techs. Inc.</i> , IPR2019-00811, Paper 44 (P.T.A.B. Oct. 15, 2020).....	17
<i>NetNut Ltd. v. Bright Data Ltd.</i> , IPR2021-00465, Paper 11 (PTAB Aug. 12, 2021).....	89
<i>Samsung Elecs. Am. Inc. v. Snik LLC</i> , IPR2020-01428, Paper 10 (PTAB Mar. 9, 2021).....	87
<i>Sand Revolution II, LLC v. Cont’l Intermodal Grp.-Trucking LLC</i> , IPR2019-01393, Paper 24 (PTAB June 16, 2020)	87
<i>Toshiba Am. Info. Sys., Inc. v. Wallelex Microelecs. Ltd.</i> , IPR2018-01538, Paper 11 (PTAB Mar. 5, 2019).....	89, 90
<i>Toyota Motor Corp. v. Cellport Sys., Inc.</i> , IPR2015-00633, Paper 11 (P.T.A.B. Aug. 14, 2015).....	8
<i>Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.</i> , 200 F.3d 795, 803 (Fed. Cir. 1999)	8

Statutes, Rules & Other Authorities

35 U.S.C. § 102(b)	1, 2
35 U.S.C. § 103	2
35 U.S.C. §325(d)	90
37 C.F.R. § 42.100(b)	8

EXHIBIT LIST

Exhibit	Description
Ex. 1001	U.S. Patent No. 10,779,080 to Burnett (“the ’080 patent”)
Ex. 1002	File History of U.S. Patent No. 10,779,080
Ex. 1003	Declaration of Jeffrey S. Vipperman, Ph.D.
Ex. 1004	Curriculum Vitae of Jeffrey S. Vipperman, Ph.D.
Ex. 1005	Japanese Unexamined Patent Application Publication No. H11-18186A to Ikeda et al. (“Ikeda”)
Ex. 1006	Certified Translation of Japanese Unexamined Patent Application Publication No. H11-18186A to Ikeda et al. (“Ikeda”)
Ex. 1007	Iain A. McCowan et al., <i>Near-Field Adaptive Beamformer for Robust Speech Recognition</i> , Digital Signal Processing, Vol. 12, Issue 1 (2002), 87-106 (“McCowan”)
Ex. 1008	U.S. Patent Application Publication No. 2004/0185804 to Kanamori et al. (“Kanamori”)
Ex. 1009	U.S. Patent Application Publication No. 2002/0193130 to Yang et al. (“Yang”)
Ex. 1010	U.S. Patent No. 5,471,538 to Sasaki et al. (“Sasaki”)
Ex. 1011	U.S. Patent Application Publication No. 2007/0244698 to Dugger et al. (“Dugger”)
Ex. 1012	U.S. Patent Application Publication No. 2008/0152167 to Taenzer (“Taenzer”)
Ex. 1013	U.S. Patent Application Publication No. 2006/0120537 to Burnett et al. (“Burnett”)

Sony Electronics Inc. v. Jawbone Innovation, LLC
U.S. Patent No. 10,779,080

Exhibit	Description
Ex. 1014	U.S. Patent Application Publication No. 2003/0031328 to Elko et al. (“Elko”)
Ex. 1015	U.S. Patent No. 6,370,401 to Baranowski et al. (“Baranowski”)
Ex. 1016	U.S. Patent No. 6,006,115 to Wingate (“Wingate”)
Ex. 1017	U.S. Patent No. 5,590,417 to Rydbeck (“Rydbeck”)
Ex. 1018	U.S. Patent Application Publication No. 2005/0259811 to Kimm et al. (“Kimm”)
Ex. 1019	U.S. Patent Application Publication No. 2003/0179888 to Burnett et al. (“Burnett II”)
Ex. 1020	Excerpt of Bernard Widrow & Samuel D. Stearns, <i>Adaptive Signal Processing</i> , Prentice-Hall (1985) (“Widrow”)
Ex. 1021	RESERVED
Ex. 1022	RESERVED
Ex. 1023	Patent Owner’s Claim Chart for U.S. Patent No. 10,779,080, <i>Jawbone Innovations, LLC v. Google LLC</i> , No. 6:21-cv-00985 (W.D. Tex.)
Ex. 1024	Declaration of Shauna L. Wiest Regarding McCowan
Ex. 1025	Excerpt of Lawrence E. Kinsler et al., <i>Fundamentals of Acoustics</i> , John Wiley & Sons, Inc. (4th ed. 2000)
Ex. 1026	Excerpt of M. P. Norton et al., <i>Fundamentals of Noise and Vibration Analysis for Engineers</i> , Cambridge Univ. Press (2nd ed. 2003)
Ex. 1027	Gary W. Elko et al., <i>A Simple Adaptive First-Order Differential Microphone</i> , Proceedings of 1995 Workshop on Applications of Signal Processing to Audio and Acoustics, IEEE (1995)

Exhibit	Description
Ex. 1028	Excerpt of Alan V. Oppenheim et al., <i>Discrete-Time Signal Processing</i> , Prentice-Hall, Inc. (2nd ed. 1999)
Ex. 1029	Declaration of Shauna L. Wiest Regarding Widrow
Ex. 1030	Declaration of Rachel J. Watters Regarding Widrow
Ex. 1031	Declaration of Shauna L. Wiest Regarding Oppenheim
Ex. 1032	Declaration of Rachel J. Watters Regarding Oppenheim
Ex. 1033	Declaration of Shauna L. Wiest Regarding Elko IEEE Workshop

**I. STATEMENT OF PRECISE RELIEF
REQUESTED FOR EACH CLAIM CHALLENGED**

Petitioner Sony Electronics Inc. (“Sony” or “Petitioner”) petitions for *inter partes* review (“IPR”) of claims 1-20 of U.S. Patent No. 10,779,080 (Ex. 1001, “the ’080 patent”) as unpatentable in view of the following:

Prior Art
Japanese Unexamined Patent Application Publication No. H11-18186A to Ikeda et al. (“Ikeda”) (Ex. 1005, translation at Ex. 1006); published January 22, 1999; prior art under pre-AIA 35 U.S.C. § 102(b).
Iain A. McCowan et al., <i>Near-Field Adaptive Beamformer for Robust Speech Recognition</i> , Digital Signal Processing, Vol. 12, Issue 1 (2002), 87-106 (“McCowan”) (Ex. 1007); published in 2002; prior art under pre-AIA 35 U.S.C. § 102(b).
U.S. Patent Application Publication No. 2004/0185804 to Kanamori et al. (“Kanamori”) (Ex. 1008); published September 23, 2004; prior art under pre-AIA 35 U.S.C. § 102(b).

Prior Art
U.S. Patent Application Publication No. 2002/0193130 to Yang et al. (“Yang”) (Ex. 1009); published December 19, 2002; prior art under pre-AIA 35 U.S.C. § 102(b).

Unpatentability Grounds	
Ground 1	Ikeda in view of McCowan and Kanamori renders obvious claims 1-3, 5-9, 11-14, and 16-20 under pre-AIA 35 U.S.C. § 103.
Ground 2	Ikeda in view of McCowan, Kanamori, and Yang renders obvious claims 4, 10, and 15 under pre-AIA 35 U.S.C. § 103.

II. OVERVIEW OF THE '080 PATENT

A. Specification

The '080 patent describes a “dual omnidirectional microphone array (DOMA)” for adaptive noise suppression. Ex. 1001, 3:49-50, 4:52-53, Abstract. Two omnidirectional microphones are used to “form two distinct virtual directional microphones which are configured to have very similar noise responses and very dissimilar speech responses.” Ex. 1001, 3:49-56, 4:48-51, 7:39-51. The

two virtual microphones are “paired with an adaptive filter algorithm” to reduce noise. Ex. 1001, 3:57-62.

In one embodiment, two virtual directional microphones V1 and V2 are formed based on two physical omnidirectional microphones O1 and O2 according to the following equations expressed in the z (digital frequency) domain:

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

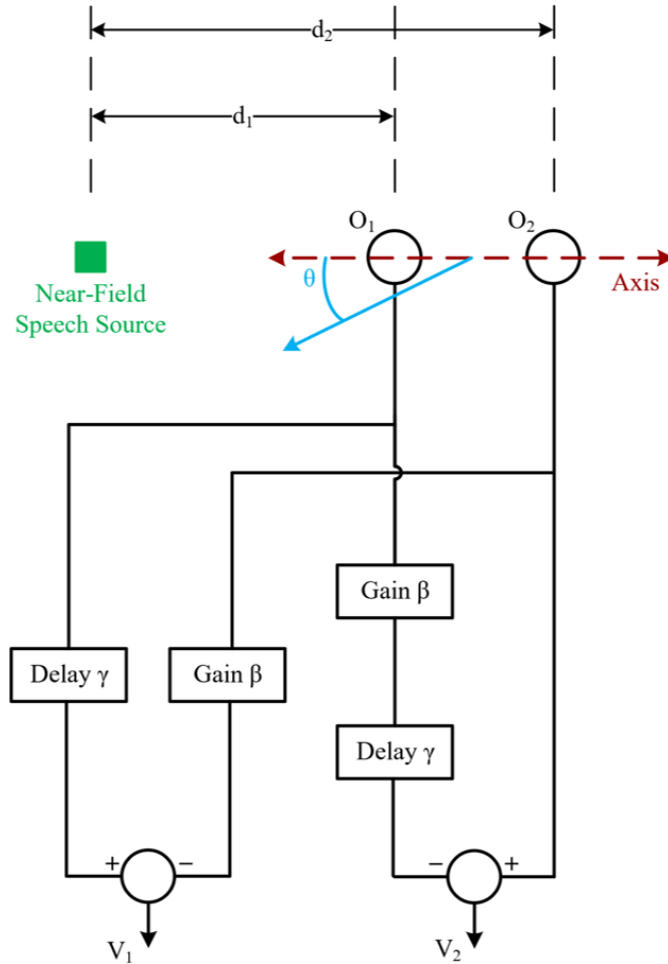
$$V_2(z) = O_2(z) - z^{-\gamma} \beta O_1(z)$$

Ex. 1001, 4:33-36, 4:42-45, 4:66-67, 9:46-53, 11:1-2. In the equations:

$$\beta = \frac{d_1}{d_2}$$

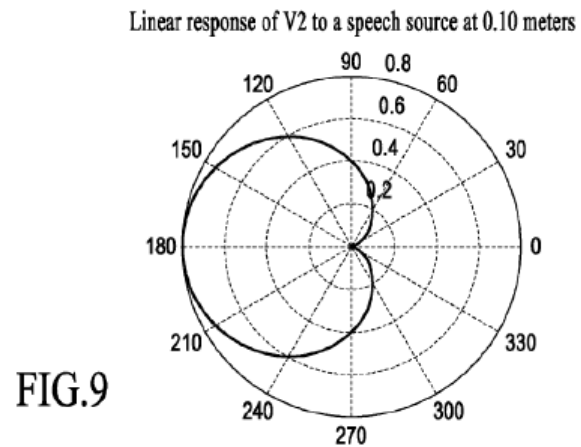
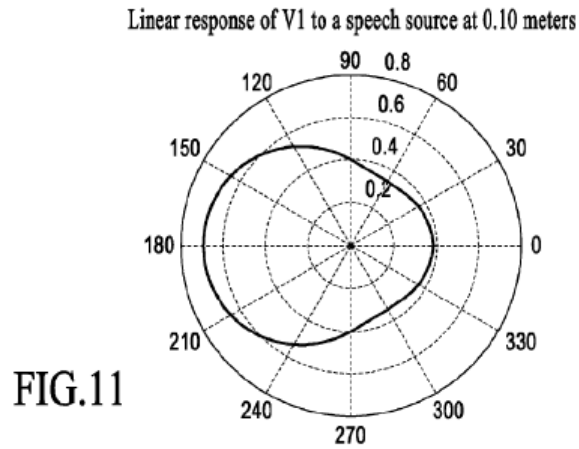
$$\gamma = \frac{d_2 - d_1}{c} \cdot f_s$$

where d_1 and d_2 “are the distance from O_1 and O_2 to the speech source,” respectively; c is the speed of sound; and f_s is a sampling frequency. Ex. 1001, 9:54-67. The figure below visually represents the equations for V_1 and V_2 above. Ex. 1003, ¶¶31-32; *see also* Ex. 1001, 2:28-34, 7:17-38, Fig. 2. This example assumes a **near-field speech source** located on the **axis** of the microphone array of O_1 and O_2 at an angle θ of 0° . Ex. 1001, 2:28-34, 7:17-38, 9:46-67, 10:30-32, 11:62-65, 16:4-6, Fig. 2; Ex. 1003, ¶¶31-32.



Ex. 1003, ¶ 32.

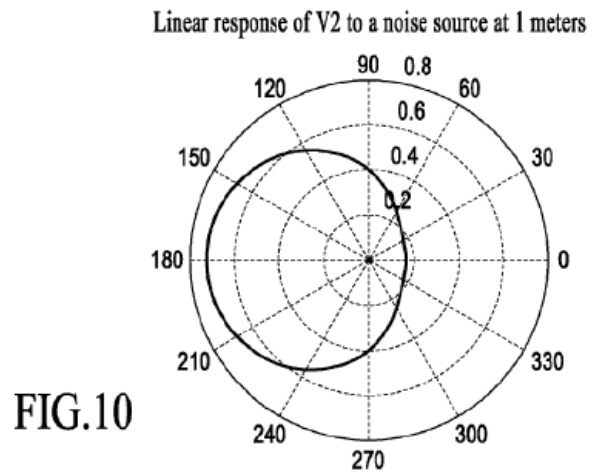
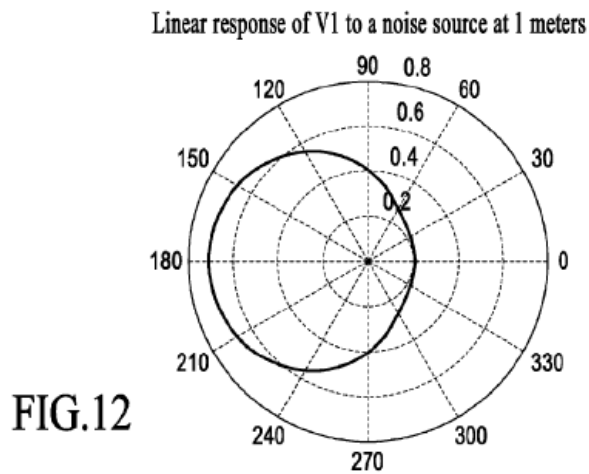
In this formulation of virtual microphones V_1 and V_2 , “the noise (far-field) responses of V_1 and V_2 are very similar while the speech (near-field) responses are very different.” Ex. 1001, 16:4-6; Ex. 1003, ¶¶ 33-35; *see also* Ex. 1001, 9:39-12:7, 15:29-16:55. Figures 11 and 9 (reproduced below) show these “different” linear responses of the virtual microphones V_1 and V_2 , respectively, to a near-field speech source. Ex. 1001, 10:18-22, 11:20-24, 11:62-65, 16:4-6.



Ex. 1001, Figs. 11, 9.

Figures 12 and 10 (reproduced below) show these “similar” linear responses of the virtual microphones V₁ and V₂, respectively, to a far-field noise source.

Ex. 1001, 10:22-27, 11:25-29, 11:62-65, 16:4-6.



Ex. 1001, Figs. 12, 10.

The linear response of the virtual microphone V₁ is “flat for all frequencies within approximately ± 30 degrees of the axis of the array,” as shown in Figure 13 (reproduced below). Ex. 1001, 11:30-35, 11:43-61; Ex. 1003, ¶36.

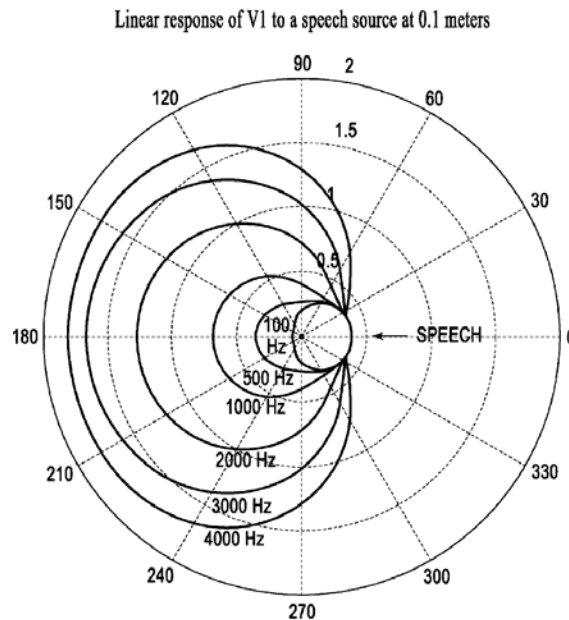


FIG.13

Ex. 1001, Fig. 13.

The virtual microphone signals are input to an adaptive noise removal system, which can use “any adaptive filter or noise removal algorithm.” Ex. 1001, 3:57-62, 6:67-7:3, 8:50-9:3, 9:22-38; Ex. 1003, ¶37; *see also* Ex. 1001, 4:52-7:16. In an embodiment, the adaptive noise removal system combines the signals from virtual microphones V₁ and V₂ “by filtering and summing in the time domain.” Ex. 1001, 7:4-16, 8:50-9:3, 9:22-38. “The adaptive filter generally uses the signal

received from a first microphone ... to remove noise from the speech received from at least one other microphone ..., which relies on a slowly varying linear transfer function between the two microphones for sources of noise.” Ex. 1001, 7:7-12. The system then generates an output signal “in which the noise content is attenuated with respect to the speech content.” Ex. 1001, 7:12-16.

B. Prosecution History

The '080 patent was filed on July 22, 2013, as U.S. Patent Application No. 13/948,160 (“the '160 application”). Ex. 1001, Cover. The '080 patent claims priority to U.S. Patent Application No. 12/139,355, filed on June 13, 2008, and four provisional applications, the earliest of which was filed on June 13, 2007. Ex. 1001, Cover.

During prosecution of the '160 application, the Examiner rejected the initial claim as anticipated by U.S. Patent No. 8,068,619 and based on non-statutory double patenting. Ex. 1002, 57, 98-106. In response, the applicant canceled the claim and added twenty new claims. Ex. 1002, 113-123. The Examiner then rejected the new claims based on non-statutory double patenting. Ex. 1002, 130-136. The applicant filed a terminal disclaimer to overcome the double patenting rejection. Ex. 1002, 142-153, 160-168, 222-234, 248. The Examiner then allowed the claims, stating that “a dual virtual microphone array as claimed including having a linear response to

both speech and noise could not be anticipated nor, in combination, be rendered obvious over the prior art of record.” Ex. 1002, 249-256.

III. THE LEVEL OF ORDINARY SKILL IN THE ART

Based on the relevant factors, *In re GPAC Inc.*, 57 F.3d 1573, 1579 (Fed. Cir. 1995), a person of ordinary skill in the art (“POSITA”) would have had a minimum of 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. Ex. 1003, ¶¶22-23. Work experience can substitute for formal education and additional formal education can substitute for work experience. Ex. 1003, ¶¶22-23.

IV. CLAIM CONSTRUCTION

The claim terms of the ’080 patent should be given their plain and ordinary meaning as understood by a POSITA, consistent with the ’080 patent’s disclosure and prosecution history. 37 C.F.R. § 42.100(b).

For purposes of this Petition, no claim terms require construction. *Toyota Motor Corp. v. Cellport Sys., Inc.*, IPR2015-00633, Paper 11 at 16 (P.T.A.B. Aug. 14, 2015) (citing *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999)).

V. SPECIFIC GROUNDS OF UNPATENTABILITY

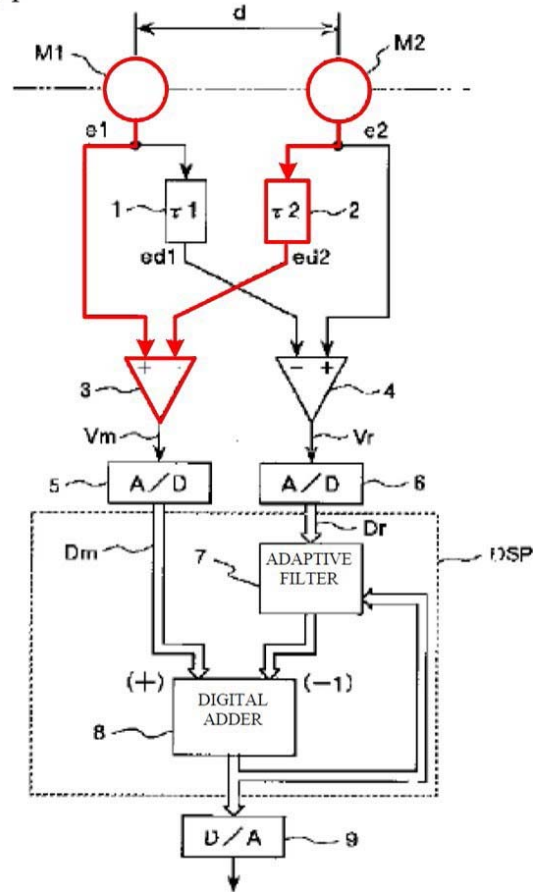
**A. Ground 1: Ikeda in View of McCowan and Kanamori
Renders Obvious Claims 1-3, 5-9, 11-14, and 16-20**

1. Overview of Ikeda

Ikeda discloses a microphone apparatus including an adaptive noise cancellation system. Ex. 1006, ¶¶[0018]-[0024], [0027]-[0030], Abstract, Fig. 1. The microphone apparatus includes “two omnidirectional microphone units M1 and M2.” Ex. 1006, ¶[0018], Fig. 1.

As shown in Figure 1 (annotated below), Ikeda discloses a signal system functioning as a **main microphone** having directional characteristics, where a delay circuit 2 delays the output of microphone M2 and an inversion adding circuit 3 subtracts the output of delay circuit 2 from the output of microphone M1 and outputs a main signal Vm. Ex. 1006, ¶¶[0002]-[0004], [0012], [0018]-[0022], Figs. 2-3; Ex. 1003, ¶¶43-44.

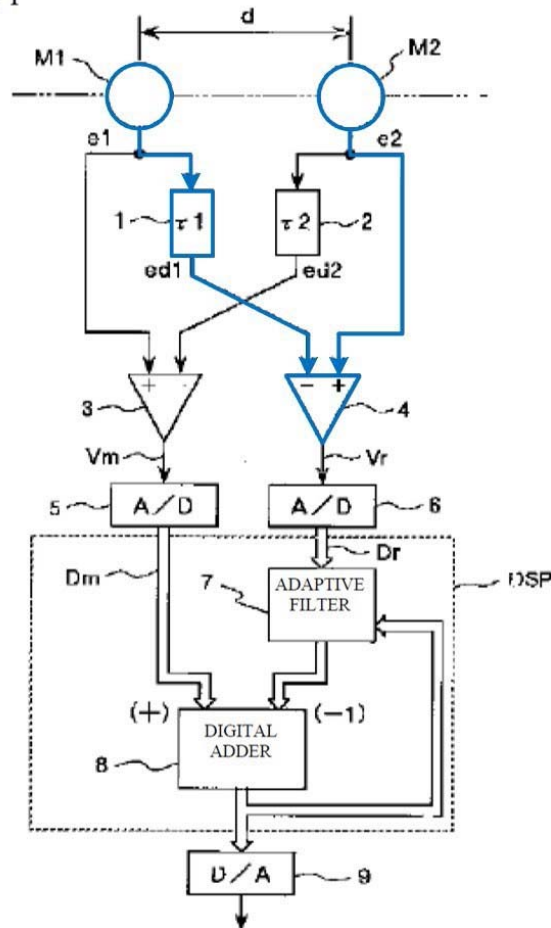
FIG. 1



Ex. 1006, Fig. 1 (annotated).

As shown in Figure 1 (annotated below), Ikeda discloses a signal system functioning as a **reference microphone** having directional characteristics, where a delay circuit 1 delays the output of microphone M1 and an inversion adding circuit 4 subtracts the output of delay circuit 1 from the output of microphone M2 and outputs a reference signal Vr. Ex. 1006, ¶¶[0012], [0018]-[0020], [0023]-[0024], Figs. 4-5; Ex. 1003, ¶45.

FIG. 1



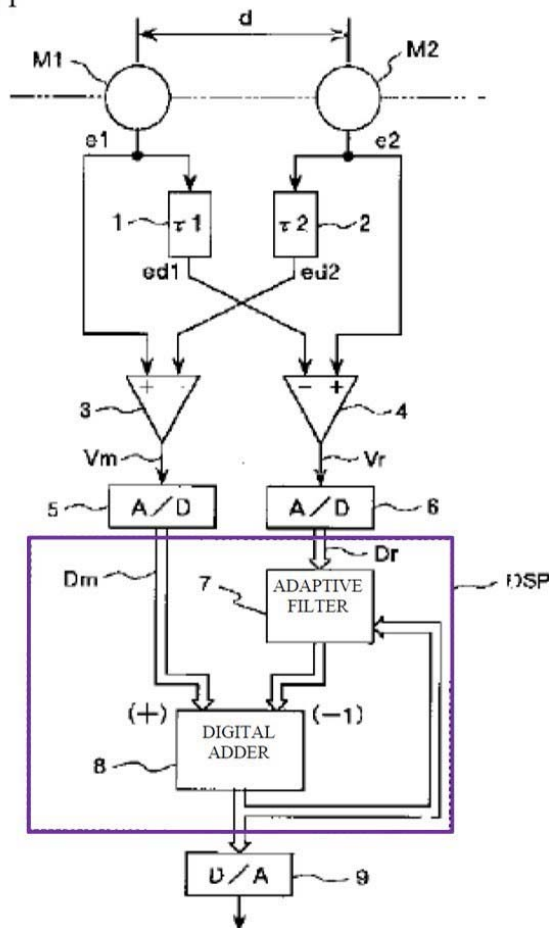
Ex. 1006, Fig. 1 (annotated).

Based on their directional characteristics, the **main microphone** is configured to pick up a target sound and noise, and the **reference microphone** is configured to pick up noise. Ex. 1006, ¶¶[0021], [0023]; Ex. 1003, ¶46; *see also* Ex. 1006, ¶¶[0012], [0026], Fig. 10, Abstract.

Main signal V_m and reference signal V_r (as converted to digital signals D_m and D_r , respectively) are input to the **adaptive noise cancellation system** including

an adaptive filter 7 and a digital adder 8, as shown in Figure 1 (annotated below).
Ex. 1006, ¶¶[0012], [0027]-[0029]; Ex. 1003, ¶47. Signal Dr is input to adaptive filter 7, and the difference between the output of adaptive filter 7 and signal Dm is calculated by digital adder 8. Ex. 1006, ¶ [0027]; Ex. 1003, ¶47.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

The response characteristics of adaptive filter 7 “are determined so as to cancel the respective noise components contained in” signal D_m and signal D_r . Ex. 1006, ¶[0028]; Ex. 1003, ¶48. Ikeda states:

As a result of the response characteristics of the adaptive filter 7 being determined successively in this way, the noise component contained in the main signal V_m is cancelled by the noise component contained in the reference signal V_r due to subtraction by digital adder 8. The output of the digital adder 8 is converted to an analog signal by a D/A converter 9. A signal primarily consisting of the target sound component can be acquired from the D/A converter 9.

Ex. 1006, ¶[0029]; *see also* Ex. 1006, ¶¶[0012], Abstract.

2. Overview of McCowan

McCowan discloses an adaptive noise cancellation system for near-field speech sources. Ex. 1007, 87-89, 97. As shown in Figure 3 (annotated below), the input $x(f)$ including signals from multiple microphones is processed via two signal paths. Ex. 1007, 90, 93 (equation 15), 95-96. The lower path, including a **near-field compensation**, a blocking matrix, and **adaptive filters**, generates a noise estimate, which is used to cancel, at a **subtraction component**, noise in the output signal of a

fixed NFSD (near-field superdirective) beamformer in the upper path. Ex. 1007, 87-91, 93-96; Ex. 1003, ¶49.

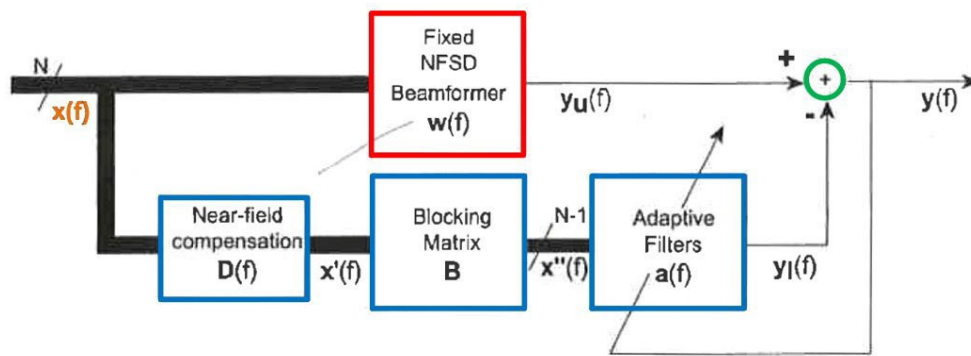


FIG. 3. Near-field adaptive beamformer.

Ex. 1007, Fig. 3 (annotated).

The **blocking matrix** “block[s] the desired signal from the adaptive noise estimate” by subtracting one microphone signal from another. Ex. 1007, 93, 95 (equations 18, 19); Ex. 1003, ¶50. “To ensure complete blocking, the desired signal must both be time aligned and have equal amplitudes across all channels.” Ex. 1007, 93.

For a near-field sound source, phase misalignment across microphones results from “[t]he difference in propagation time to [the i th] microphone with respect to the reference microphone ($i=1$),” defined as:

$$\tau_i = \frac{d_i - d_1}{c}$$

where d_i is the distance between the source and the i th microphone and c is the speed of sound. Ex. 1007, 89-90. Additionally, amplitude differences across microphones are “negligible for far-field sources, but can be significant in the near-field case.” Ex. 1007, 90. The attenuation factor of wave amplitude on the i th microphone, with respect to the amplitude on the reference microphone ($i=1$), is defined as:

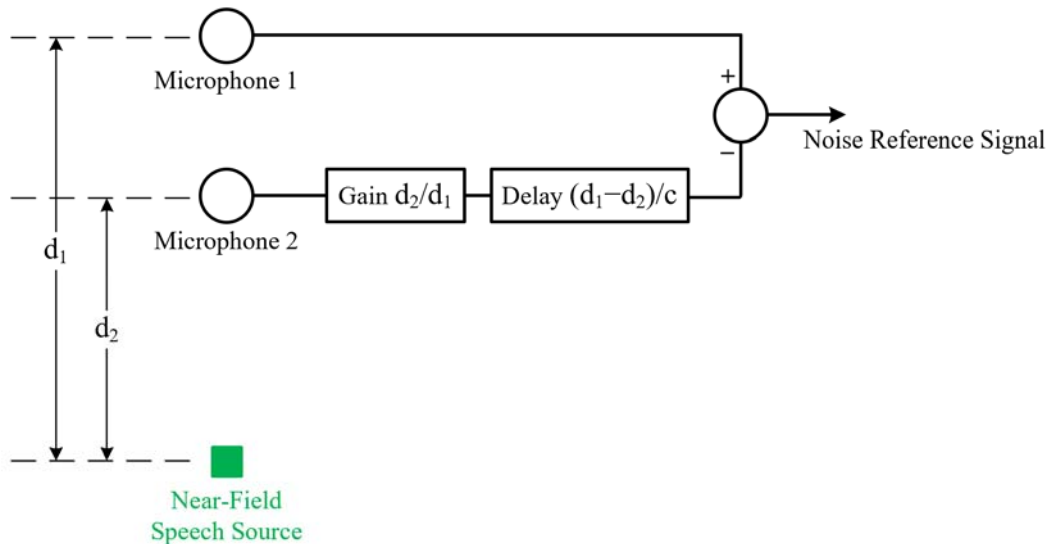
$$\alpha_i = \frac{d_1}{d_i}$$

Ex. 1007, 90; Ex. 1003, ¶51.

“For a near-field desired source, to align the desired signal on all channels, a near-field compensation must first be applied to the input channels prior to blocking.” Ex. 1007, 93; *see also* Ex. 1007, 87. The near-field compensation unit “compensate[s] for both phase misalignment and amplitude scaling of the desired signal across sensors.” Ex. 1007, 93; Ex. 1003, ¶¶51-52.

With the near-field compensation, McCowan teaches multiplying microphone 2’s signal by d_2/d_1 , delaying by $(d_1-d_2)/c$, and subtracting the result from microphone 1’s signal to create “a true null ... in the beam-pattern ... in the direction *and* distance corresponding to the desired source.” Ex. 1007, 90, 93-95; Ex. 1003, ¶¶52-53. The following figure illustrates the signal system producing a noise reference signal,

where the near-field speech source is assumed to be on the axis of microphones 1 and 2. Ex. 1007, 87-91, 93-96, Fig. 3; Ex. 1003, ¶¶52-53.



Ex. 1003, ¶ 53.

This configuration with gain d_2/d_1 and delay $(d_1 - d_2)/c$ minimizes leakage of a desired signal of a near-field source into the noise reference signal, which is then used by the adaptive filters to generate the noise estimate for canceling noise in the upper signal path of the McCowan system. Ex. 1007, 87-91, 93-96, Fig. 3; Ex. 1003, ¶54.

a. McCowan is Prior Art

McCowan was publicly available in January 2002 when it published in *Digital Signal Processing*, Volume 12, Number 1, a regularly published periodical journal by Elsevier Science. Ex. 1007, Cover, Title Pages, 87; Ex. 1024, ¶¶12-31.

McCowan was received and stamped by the Library of Congress on or before February 19, 2002, and would have been processed, cataloged, shelved, and made available to the public a few days or weeks after receipt. Ex. 1024, ¶¶12-17, 23-27, Appx. A, B, C.

McCowan was received and stamped by the library of Iowa State University of Science and Technology on or before February 18, 2002, and would have been processed, cataloged, shelved, and made available to the public a few days or weeks after receipt. Ex. 1024, ¶¶12, 18-22, 28-30, Appx. D, E, F.

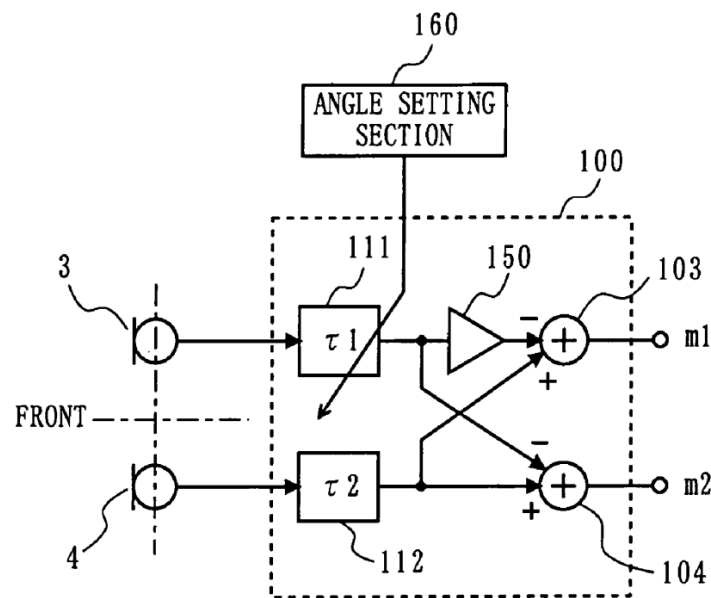
McCowan bears a copyright date of 2002 and indicates that it was published in “Digital Signal Processing 12, 87-106 (2002).” Ex. 1007, Cover, Title Pages, 87. Because *Digital Signal Processing* was a regular periodic journal that was distributed widely and intended to be accessible to the public, McCowan’s copyright date and publication information are evidence of public availability at that date. See *Microsoft Corp. v. IPA Techs. Inc.*, IPR2019-00811, Paper 44 at 142 (P.T.A.B. Oct. 15, 2020); *Ericsson Inc. v. Intellectual Ventures I LLC*, IPR2014-00527, Paper 41 at 10-12 (P.T.A.B. May 18, 2015).

3. Overview Of Kanamori

Kanamori discloses a technique “to generate a main signal and noise reference signal used in a noise suppressing process.” Ex. 1008, ¶[0018]. Figure 17A

(reproduced below) shows such a technique, where a main signal $m1$ and a noise reference signal $m2$ are generated using signals from microphone units 3 and 4. Ex. 1008, ¶¶[0151], [0166], [0172]-[0174].

FIG. 17A



Ex. 1008, Fig. 17A.

Kanamori discloses “suppress[ing], based on the noise reference signal, noise components included in the main signal.” Ex. 1008, ¶[0168]; *see also* Ex. 1008, ¶¶[0076]-[0115], Figs. 1-7. Kanamori discloses a filter coefficient calculating section 40 that calculates the power spectrums of the main and reference signals and, based on the ratio of those power spectrums, calculates the estimated noise spectrum in the main signal, which is used to suppress “noise components included in the main

signal.” Ex. 1008, ¶¶[0100]-[0115], [0168], Fig. 4; *see also* Ex. 1008, ¶¶[0080]-[0086], [0093]-[0099]; Ex. 1003, ¶¶56-58.

Kanamori discloses that, when “noise is coming from a plurality of directions,” suppressing “noise components included in the main signal” requires that the ratio of the power spectrums of the main and reference signals “be constant irrespectively of the direction from which noise is coming.” Ex. 1008, ¶[0168]; Ex. 1003, ¶59. Otherwise, the estimated noise spectrum in the main signal will be correct only “in a specific direction.” Ex. 1008, ¶[0168]; Ex. 1003, ¶59. For this reason, Kanamori teaches that the directivity patterns of main signal m1 and noise reference signal m2 are preferably as similar as possible in directions from which noise is coming. Ex. 1008, ¶¶[0167]-[0170]; Ex. 1003, ¶59.

4. The Ikeda-McCowan-Kanamori Combination

Ikeda discloses or renders obvious many features of the challenged claims, such as first and second virtual microphones, a processing component, and an adaptive noise removal application. *See* Ex. 1006, ¶¶[0002]-[0004], [0018]-[0024], [0026]-[0030], Figs. 1-5. Ikeda, however, does not expressly disclose features relating to the virtual microphones’ linear response to speech or noise, such as those recited in limitations [7c]-[7h]. However, the combined system of Ikeda-McCowan-

Kanamori teaches these features. It would have been obvious to a POSITA to combine Ikeda-McCowan-Kanamori as explained below.

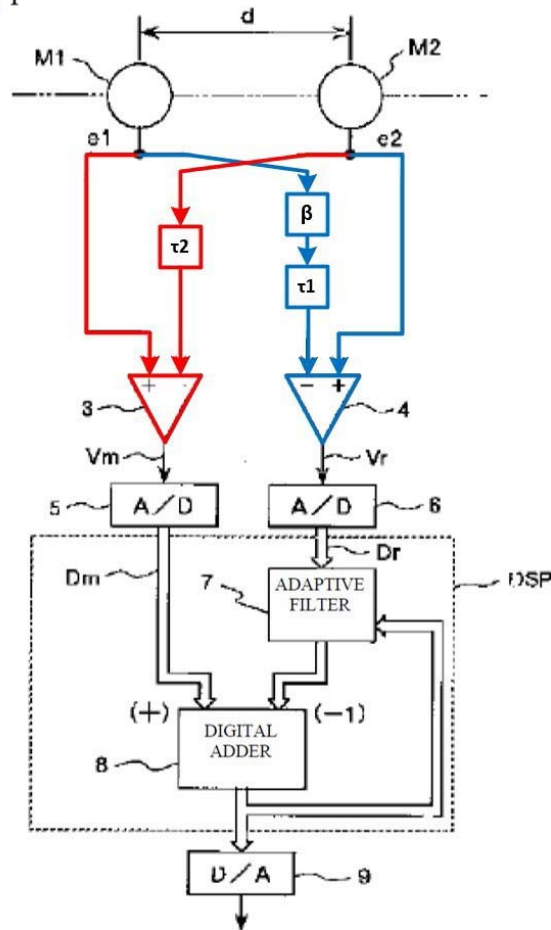
Ikeda assumes far-field speech and noise. Ex. 1006, ¶¶[0002]-[0004], [0020]-[0024], [0036], Figs. 2-7; Ex. 1003, ¶61. However, it would have been obvious to implement Ikeda's system in a device intended to receive near-field speech, such as a headset. Ex. 1003, ¶61. For example, Ikeda discloses applying its invention "to a wide range of devices," including "a voice input microphone of a hands-free telephone unit." Ex. 1006, ¶[0036]. It was well-known that a headset was such "a voice input microphone of a hands-free telephone unit." Ex. 1006, ¶[0036]; Ex. 1011 (Dugger), ¶[0004]; Ex. 1012 (Taenzer), ¶¶[0008]-[0009]; Ex. 1013 (Burnett), ¶¶[0048], [0071], [0126], Abstract; Ex. 1003, ¶61. Such a headset was commonly located 3-5" (approximately 0.08-0.13 meters) away from the user's mouth. Ex. 1012 (Taenzer), ¶¶[0009], [0078]; Ex. 1003, ¶61. For the headset, the speech source is in the near-field and one or more noise sources are in the far-field. Ex. 1012 (Taenzer), ¶[0009]; Ex. 1003, ¶61; *see also* Ex. 1007, 90 (disclosing that "the wavefront amplitude decays at a rate proportional to the distance traveled" and "[t]he resulting amplitude differences across sensors are negligible for far-field sources, but can be significant in the near-field case"); Ex. 1011 (Dugger), ¶[0006]

(disclosing that a noise source (“interferer”) is “in the far field”), ¶[0035] (disclosing that a speech source (“desired speaker”) is “in the near field”).

When Ikeda’s system is implemented in a device intended for environments where the speech source is in the near-field and one or more noise sources are in the far-field, such as a headset, a POSITA would have been motivated to improve Ikeda’s system in view of McCowan. Ex. 1003, ¶¶62-66. A POSITA would have been motivated to modify the Ikeda [reference microphone](#) to add the gain β^1 in view of McCowan to result in a system, for example, as shown below. Ex. 1003, ¶¶62-66.

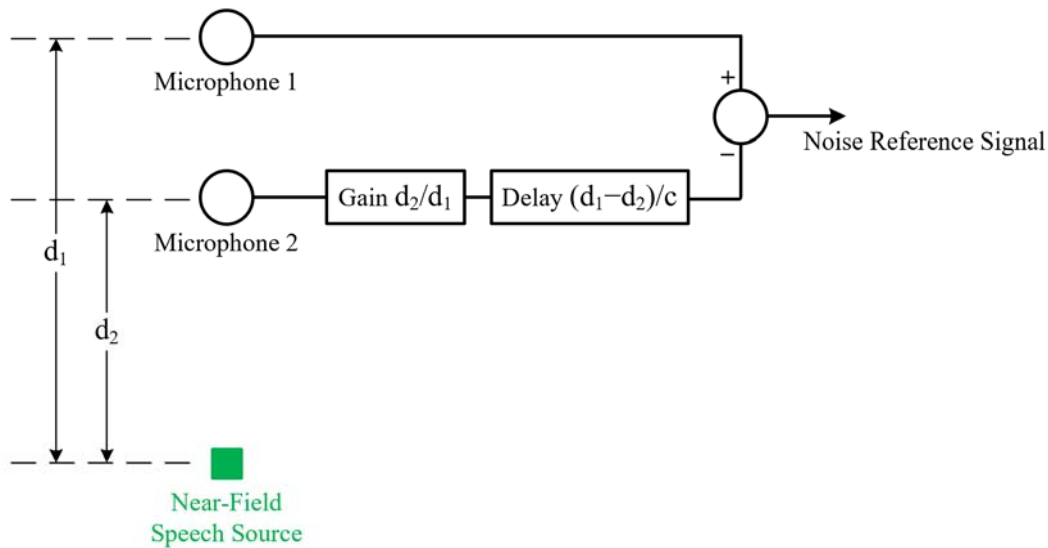
¹ McCowan teaches a gain in the near-field compensation that functions in the same way as the ’080 patent’s gain β in the virtual microphone V₂. Ex. 1006, 90, 93- 95; Ex. 1001, 9:50-67; Ex. 1003, ¶¶53-55, 62 (fn.1). For simplicity, Petitioner uses the term β to refer to McCowan’s gain.

FIG. 1



Ex. 1006, Fig. 1 (modified).

A POSITA would have been motivated to make this modification for the following reasons. As discussed in Section V.A.2, McCowan teaches multiplying microphone 2's signal by d_2/d_1 , delaying by $(d_1-d_2)/c$, and subtracting the result from microphone 1's signal to form a noise reference signal, as shown below (the **near-field speech source** is assumed to be on the axis of microphones 1 and 2). Ex. 1007, 87-91, 93-96, Fig. 3; Ex. 1003, ¶¶52-55, 63.



Ex. 1003, ¶¶53, 63.

McCowan teaches that the gain compensates for the amplitude difference between microphones 1 and 2 and the delay compensates for the phase difference between microphones 1 and 2, for the near-field speech source. Ex. 1007, 89-91, 93-96, Fig. 3; Ex. 1003, ¶64. The signal system with the gain and the delay produces “a true null ... in the beam-pattern ... in the direction and distance corresponding to the desired source.” Ex. 1007, 94; Ex. 1003, ¶64. The signal system with the gain and the delay minimizes leakage of a desired signal of a near-field source into the noise reference signal. Ex. 1007, 87-91, 93-96, Fig. 3; Ex. 1003, ¶64.

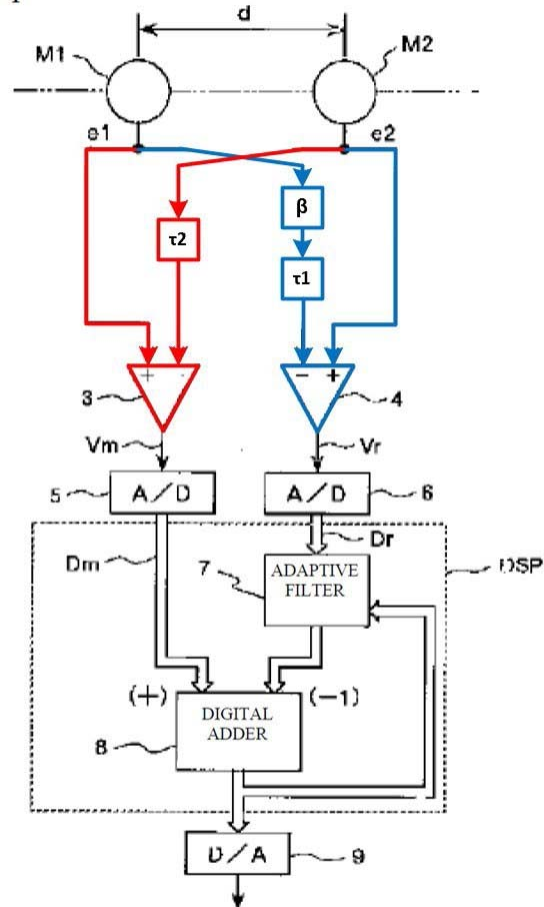
Similar to McCowan, Ikeda’s reference microphone outputs a noise reference signal V_r having a null in the direction of the target sound source. Ex. 1006, ¶¶[0018]-[0020], [0023]-[0024], Figs. 1, 4-5; Ex. 1003, ¶¶45-46, 65. A POSITA

would therefore have been motivated to apply McCowan's technique to Ikeda's reference microphone for a near-field speech source, so that a null exists in the reference microphone's directivity pattern in the "direction *and* distance" corresponding to the near-field speech source. Ex. 1007, 94; Ex. 1003, ¶65. Consistent with Ikeda's disclosures for far-field speech, the **near-field speech source** would have been located on the axis of Ikeda's microphones M1 and M2 as shown below. Ex. 1006, ¶¶[0019] (microphone M1 "is disposed on the front surface"), [0021], [0023] ("target sound ... from the front surface side"), Figs. 1-5; Ex. 1003, ¶65. Therefore, a POSITA would have been motivated to apply the gain β to the signal of microphone M1 for forming the reference microphone to compensate for the amplitude difference between microphones M1 and M2, as taught by McCowan. Ex. 1007, 89-90, 93-96; Ex. 1003, ¶65. The gain β is equal to the distance between the near-field speech source and microphone M1 divided by the distance between the near-field speech source and microphone M2. Ex. 1007, 89-90, 93-96; Ex. 1003, ¶¶53-55, 65. Ikeda discloses that the delay τ_1 is equal to the distance between microphones M1 and M2 divided by the speed of sound, consistent with McCowan's teachings of the delay to compensate for the phase difference between microphones. Ex. 1006, ¶¶[0018] (microphones M1 and M2 "are disposed at a distance d apart from one another;" " τ_1 : amount of delay expressed in terms of distance"), [0023]-

[0024], [0026] (“ $\tau_1=d$ ”), Figs. 1, 4-5, 10; Ex. 1007, 89-90, 93-96; Ex. 1003, ¶¶53-55, 65.

 Near-Field
Speech Source

FIG. 1



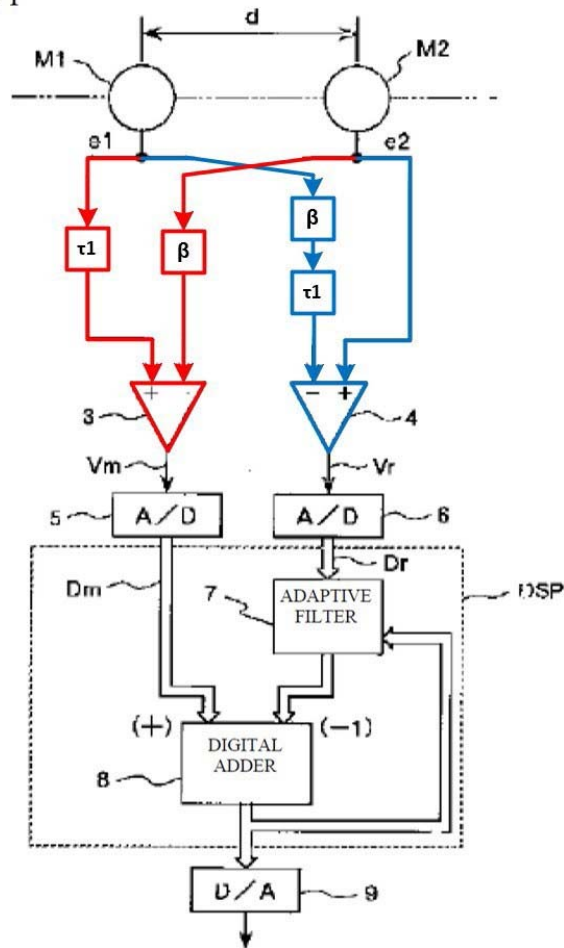
Ex. 1006, Fig. 1 (modified).

A POSITA would have had a reasonable expectation of success to make the modification of Ikeda with McCowan, considering the similarities between Ikeda and McCowan as discussed above. Ex. 1003, ¶66. Adding a gain to the signal path would have been within the knowledge and skill of a POSITA. Ex. 1003, ¶66;

Ex. 1007, 90, 93-95; Ex. 1008, ¶[0169]. As such, the modification would have resulted in predictable results that the reference microphone outputs a noise reference signal where leakage of a desired signal of the near-field speech source into the noise reference signal is minimized. Ex. 1006, ¶¶[0018]-[0020], [0023]-[0024], Figs. 1, 4-5; Ex. 1007, 87-91, 93-96, Fig. 3; Ex. 1003, ¶66.

A POSITA would have been further motivated to improve the Ikeda-McCowan system in view of Kanamori by modifying the main microphone as shown below. Ex. 1003, ¶¶67-78.

FIG. 1



Ex. 1006, Fig. 1 (modified).

A POSITA would have been motivated to make this modification for the following reasons. As discussed in Section V.A.3, Kanamori teaches that the directivity patterns of main signal m1 and noise reference signal m2 are preferably as similar as possible in directions from which noise is coming. Ex. 1008, ¶¶[0167]-[0170]; Ex. 1003, ¶¶57-59, 68. For example, Kanamori states that, when “noise is coming from a plurality of directions,” suppressing “noise components included in

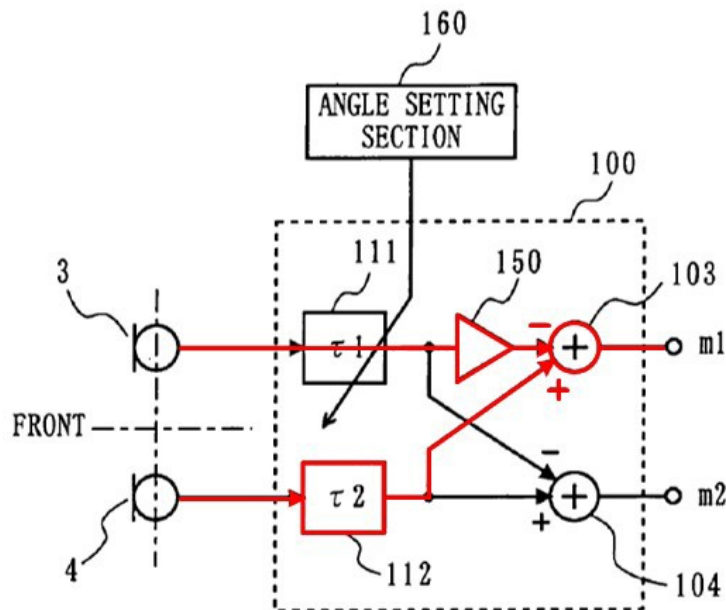
the main signal” requires that the ratio of the power spectrums of the main and reference signals “be constant irrespectively of the direction from which noise is coming.” Ex. 1008, ¶[0168]; Ex. 1003, ¶¶57-59, 68. A POSITA would have understood that the directivity patterns of main signal m1 and noise reference signal m2 with respect to noise are preferably as similar as possible so that a more accurate estimation of noise included in main signal m1 based on noise reference signal m2 can be determined “irrespectively of the direction from which noise is coming.” Ex. 1008, ¶¶[0101]-[0103], [0168], Figs. 1, 4; Ex. 1003, ¶¶57-59, 68. The more accurate estimation of noise included in main signal m1 is then used to suppress noise components included in main signal m1. Ex. 1008, ¶¶[0077]-[0083], [0100]-[0115], [0168], Figs. 1, 4; Ex. 1003, ¶¶57-59, 68.

Similar to Kanamori, Ikeda also teaches using adaptive filter 7 to estimate the noise included in main signal Vm based on reference signal Vr and using the estimated noise to cancel noise components included in main signal Vm. Ex. 1006, ¶¶[0027]-[0029]; Ex. 1003, ¶¶46-48, 69. A POSITA would therefore have been motivated to apply Kanamori’s teachings to the Ikeda-McCowan system, so that adaptive filter 7 more accurately estimates the noise included in main signal Vm based on reference signal Vr “irrespectively of the direction from which noise is coming,” resulting in better noise reduction for main signal Vm. Ex. 1008, ¶[0168];

Ex. 1003, ¶69; *see also* Ex. 1006, ¶¶[0027]-[0029]; Ex. 1008, ¶¶[0077]-[0083], [0100]-[0115].

Kanamori teaches modifying the main microphone in Ikeda-McCowan in the way as shown above, so that the directivity patterns of the main and reference signals with respect to noise from the far-field are as similar as possible to each other. Ex. 1003, ¶¶70-78. For example, as shown in Figure 17A (annotated below), the annotated configuration for generating main signal m1, where the signal delay amount τ_1 of the first signal delaying section 111 is set to zero (Ex. 1008, ¶[0175]), would be readily applicable to generating the main signal Vm in Ikeda-McCowan. Ex. 1008, ¶¶[0164]-[0176]; Ex. 1003, ¶70.

FIG. 17A

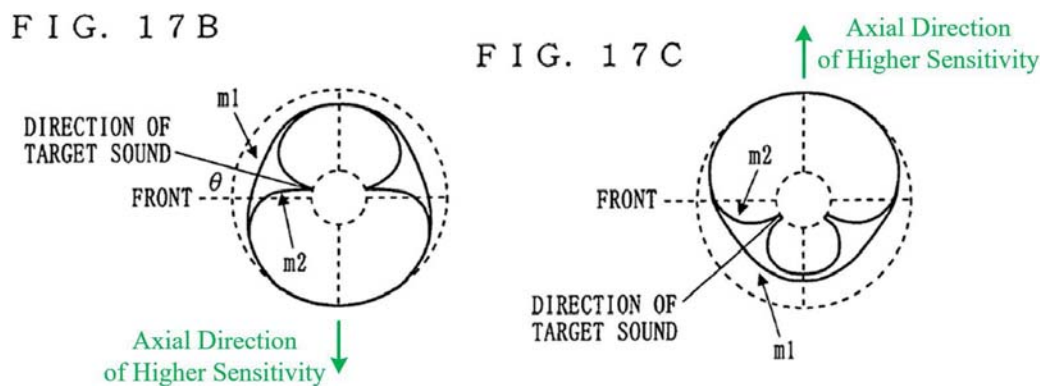


Ex. 1008, Fig. 17A (annotated).

A POSITA would have understood that Ikeda-McCowan's main signal V_m , when generated using the configuration annotated above, would have a directivity pattern with respect to far-field noise as similar as possible or identical to the directivity pattern of Ikeda-McCowan's reference signal V_r with respect to far-field noise, by sharing the same delay configuration of the signals from microphones M1 and M2 for the main and reference microphones, and by setting the gain component of the main microphone to the same value as the gain component of the reference microphone, as discussed below. Ex. 1003, ¶¶71-78.

Kanamori teaches that main signal m_1 and noise reference signal m_2 according to Figure 17A (annotated above) can have the directivity patterns as shown in Figures 17B and 17C (annotated below), when the signal delay amount τ_1 of the first signal delaying section 111 is configured to different values so that the signal from microphone 3 is either more delayed or less delayed than the signal from microphone 4, causing axial directions of higher sensitivity being oriented differently between Figures 17B and 17C. Ex. 1008, ¶¶[0167]-[0169], [0173]-[0176], Figs. 17A-17C; Ex. 1003, ¶72. A POSITA would therefore have understood that using the same relative delay between the signals from microphones 3 and 4 for each of the two signal systems that generate main signal m_1 and noise reference signal m_2 would allow the directivity patterns of m_1 and m_2 to have the same axial

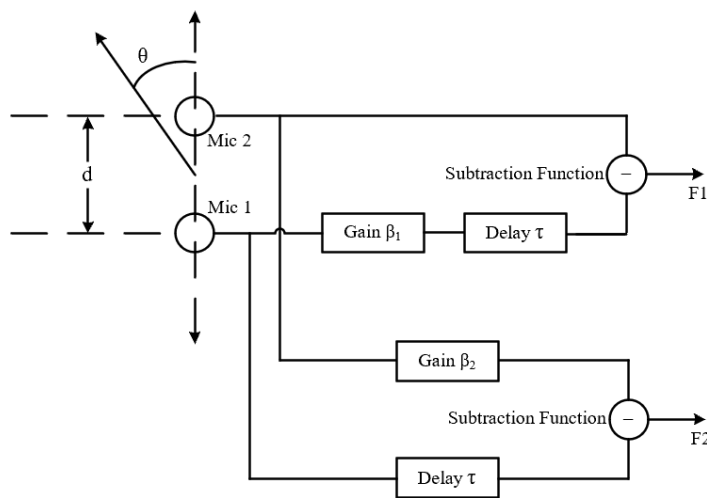
direction of higher sensitivity and would contribute to the directivity patterns of m1 and m2 having similar shapes, although the signal amplifying section 150 causes differences in the directivity patterns of m1 and m2, such as in a region surrounding the direction of target sound. Ex. 1008, ¶¶[0167]-[0169], [0173]-[0176]; Ex. 1003, ¶72.



Ex. 1008, Figs. 17B (annotated), 17C (annotated).

A POSITA would have further understood that, regardless of whether the signal path in which signal amplifying section 150 appears is more delayed or less delayed (by the same amount) than the other signal path, the general shape of the directivity pattern of m1 remains the same as suggested by the similar directivity patterns of m1 in Figures 17B and 17C when observed from their respective axial directions of higher sensitivity. Ex. 1008, ¶¶[0167]-[0169], [0173]-[0176], Figs. 17A-17C; Ex. 1003, ¶¶73-76. Indeed, using tools such as the underlying equations or simulation software, a POSITA would have been able to understand the far-field

directivity pattern of a differential microphone with a gain component in one signal path, such as those shown in Kanamori and in the below examples of F1 and F2. Ex. 1008, ¶¶[0167]-[0169], [0173]-[0176]; Ex. 1006, ¶¶[0002]-[0004], [0020]-[0024]; Ex. 1014 (Elko), ¶¶[0021]-[0023], [0027]-[0028], Figs. 1-2, Claim 5; Ex. 1027, 1-2; Ex. 1033; Ex. 1003, ¶¶73-76.



Ex. 1003, ¶¶ 73-76.

A POSITA would have understood that, when two differential microphones have the same delay configuration (as in the figure above), their far-field directivity patterns are the same if the value of the gain is the same, regardless of the signal path in which the gain component appears. Ex. 1003, ¶¶73-76. It would therefore have been obvious to a POSITA to set the same value for the gain β_1 and the gain β_2 in the examples shown above, so that the far-field directivity patterns of F1 and F2 are identical. Ex. 1003, ¶¶73-76.

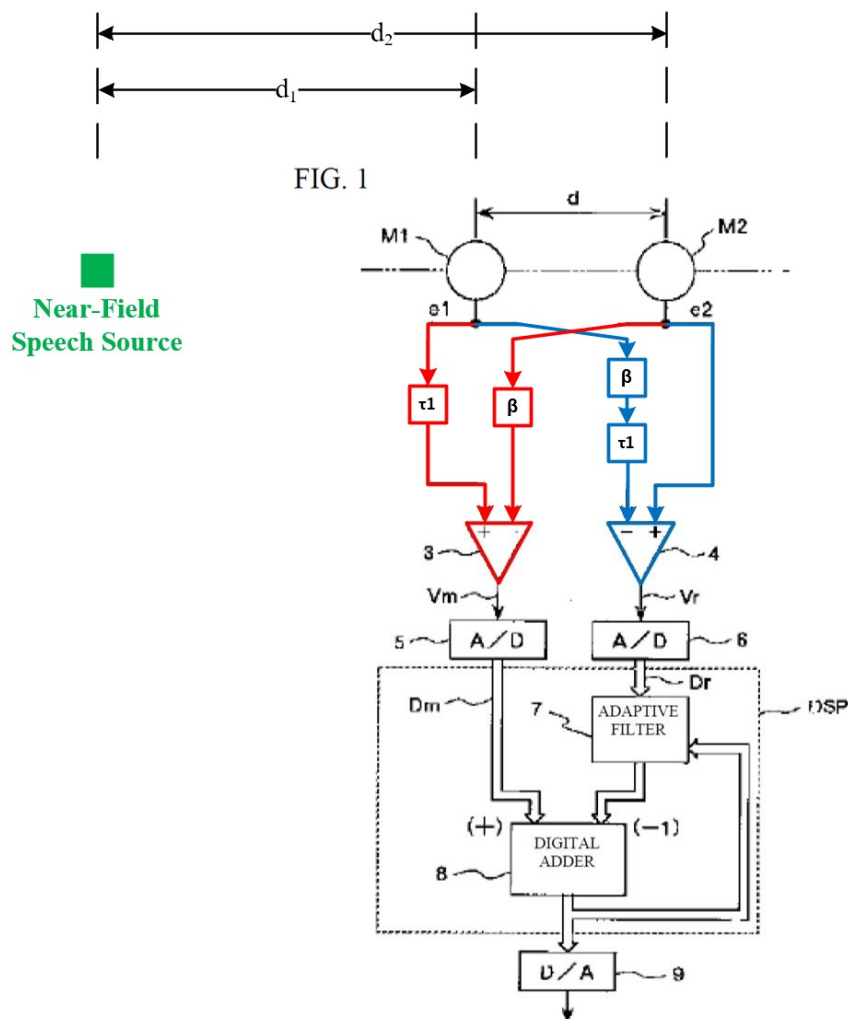
A POSITA would therefore have been motivated to modify the Ikeda-McCowan main microphone, using a configuration like the one annotated above in Figure 17A of Kanamori for generating the main signal m_1 , so that main signal V_m has the same far-field directivity pattern as Ikeda-McCowan's noise reference signal V_r . Ex. 1003, ¶77. A POSITA would have been motivated to delay the output signal of the Ikeda-McCowan microphone M1 by the same amount (τ_1) for both the main microphone and the reference microphone, as taught by Kanamori. Ex. 1008, ¶¶[0167]-[0169], [0173]-[0176], Figs. 17A-17C; Ex. 1003, ¶77. A POSITA would have likewise been motivated to multiply the output signal of the Ikeda-McCowan microphone M2 by the same amount (β) for the main microphone as the output signal of the Ikeda-McCowan microphone M1 is multiplied for the reference microphone, as taught by Kanamori and in view of a POSITA's general knowledge of directivity patterns of differential microphones. Ex. 1008, ¶¶[0167]-[0169], [0173]-[0176], Figs. 17A-17C; Ex. 1006, ¶¶[0002]-[0004], [0020]-[0024]; Ex. 1014 (Elko), ¶¶[0021]-[0023], [0027]-[0028], Figs. 1-2, Claim 5; Ex. 1027, 1-2; Ex. 1003, ¶77. The modification would have resulted in main signal V_m and reference signal V_r having directivity patterns as similar as possible or identical with respect to far-field noise, which Kanamori teaches is preferable. Ex. 1003, ¶77; Ex. 1008, ¶[0168].

A POSITA would have had a reasonable expectation of success to make the modification of Ikeda-McCowan with Kanamori, considering the similarities between Ikeda and Kanamori as discussed above. Ex. 1003, ¶78. Additionally, both Ikeda and Kanamori use signals from two microphones to form a signal system for generating a main signal having directional characteristics, and this similarity would reinforce the reasonable expectation of success to make the modification. Ex. 1006, ¶¶[0018]-[0022], Figs. 1-3; Ex. 1008, ¶¶[0163]-[0176], Figs. 17A-17C; Ex. 1003, ¶78. Modifying the signal path with a delay and a gain would have been within the knowledge and skill of a POSITA. Ex. 1003, ¶78; Ex. 1007, 90, 93-95; Ex. 1008, ¶¶[0163]-[0176]. As such, the modification would have predictably resulted in the main and reference microphones having directivity patterns as similar as possible or identical with respect to far-field noise. Ex. 1006, ¶¶[0027]-[0029], Fig. 1; Ex. 1008, ¶¶[0101]-[0115], [0164]-[0176], Figs. 4, 16A-16B, 17A-17C; Ex. 1003, ¶78.

- a. **The Ikeda-McCowan-Kanamori main microphone has the same formulation for signal processing as the virtual microphone V₁ described in the '080 patent, and the Ikeda-McCowan-Kanamori reference microphone has the same formulation for signal processing as the virtual microphone V₂ described in the '080 patent**

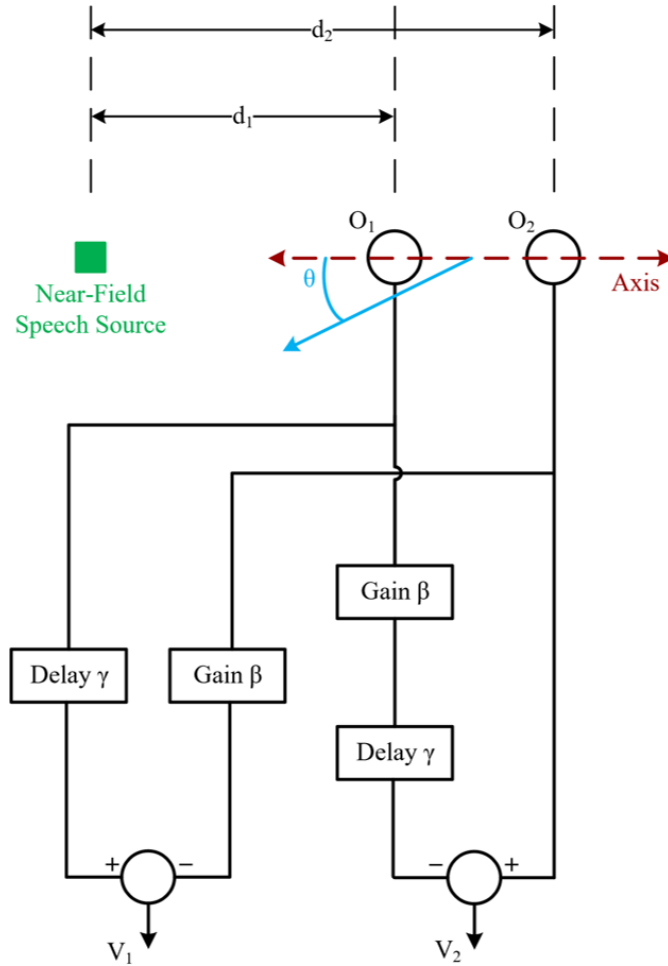
As discussed above in Section V.A.4, the Ikeda-McCowan-Kanamori system including the **main microphone** that generates main signal V_m and the **reference**

microphone that generates reference signal V_r is shown below, where the near-field speech source is located on the axis of microphones M1 and M2, as taught by Ikeda. Ex. 1003, ¶79. Microphones M1 and M2 are omnidirectional microphones. Ex. 1006, ¶[0018]. The gain β equals d_1/d_2 , and the delay τ_1 equals $(d_2-d_1)/c$, where c is the speed of sound. Ex. 1003, ¶79; § V.A.4.



Ex. 1006, Fig. 1 (modified); Ex. 1003, ¶ 79.

As discussed above in Section II.A, virtual microphones V_1 and V_2 described in the '080 patent are shown below, where the near-field speech source is located on the axis of the microphone array of O_1 and O_2 . Ex. 1003, ¶80. Microphones O_1 and O_2 are omnidirectional microphones. Ex. 1001, 4:33-36, 7:17-38, 9:39-43. The gain β equals d_1/d_2 , and the delay γ (in units of samples) equals $((d_2-d_1)/c)*f_s$, where c is the speed of sound and f_s is a sampling frequency. Ex. 1003, ¶80; § II.A. A POSITA would have understood that both the delay γ and the delay τ_1 are intended to account for the time length of sound propagation between the microphones (O_1 and O_2 or M_1 and M_2 , respectively), where the time length is directly used as the delay τ_1 in an analog design, and the time length multiplied by a sampling frequency is used as the delay γ in a digital design. Ex. 1003, ¶¶79-80.



Ex. 1003, ¶ 80.

As depicted above, the Ikeda-McCowan-Kanamori main signal V_m , which would have been obvious for the reasons explained in Section V.A.4, and the virtual microphone V_1 described in the '080 patent are generated by identical signal systems, and therefore would have the same characteristics (e.g., relating to linear responses to speech or noise). Ex. 1003, ¶¶ 79-81. Additionally, the Ikeda-McCowan-Kanamori reference signal V_r , which would have been obvious for the

reasons explained in Section V.A.4, and the virtual microphone V₂ described in the '080 patent are generated by identical signal systems, and therefore would have the same characteristics (e.g., relating to linear responses to speech or noise). Ex. 1003, ¶¶79-81.

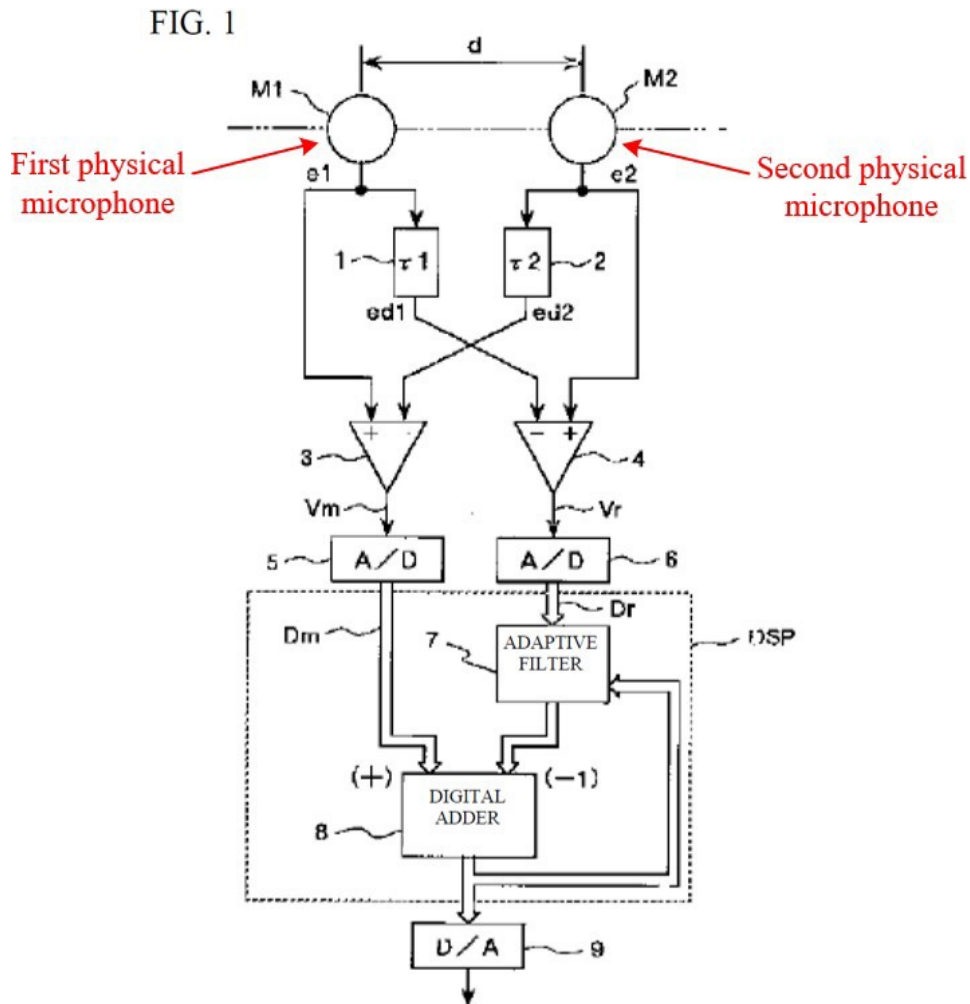
5. Claim 7

To simplify the petition, Petitioner starts its analysis with independent claim 7, and then addresses independent claims 14 and 1.

- a. [7a] **“A system, comprising: a first virtual microphone formed from a first combination of a first microphone signal and a second microphone signal, wherein the first microphone signal is generated by a first physical microphone and the second microphone signal is generated by a second physical microphone;”**

Ikeda discloses, and Ikeda-McCowan-Kanamori renders obvious, this limitation.

Ikeda discloses that “omnidirectional microphone units M1 and M2,” shown in Figure 1 (annotated below), are first and second physical microphones, respectively. Ex. 1006, ¶¶[0018]-[0019], Fig. 1; Ex. 1003, ¶95. Ikeda discloses that output e1 of microphone M1 is a first microphone signal generated by the first physical microphone, and that output e2 of microphone M2 is a second microphone signal generated by the second physical microphone. Ex. 1006, ¶¶[0018]-[0020], Figs. 1-2, 4; Ex. 1003, ¶95.



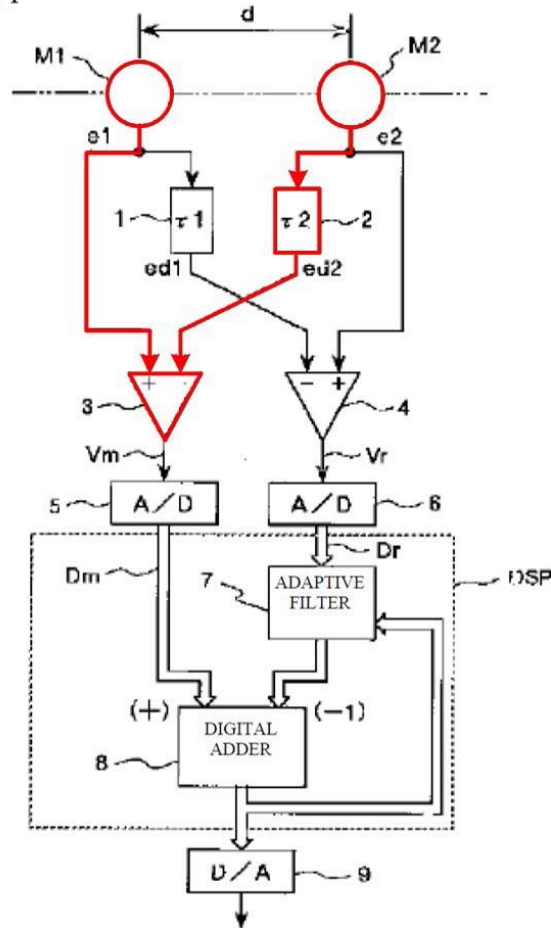
Ex. 1006, Fig. 1 (annotated).

The '080 patent states that a “virtual microphone[]” is “a microphone constructed using two or more omnidirectional microphones and associated signal processing.” Ex. 1001, 4:48-51. The '080 patent uses “virtual microphones” and “virtual directional microphones” interchangeably. Ex. 1001, 4:48-51, 7:32-35, 7:66-8:2. Ikeda discloses a first virtual microphone by disclosing the **main microphone** that constructs main signal V_m , as shown in Figure 1 (annotated below).

Ex. 1003, ¶¶96-97. Ikeda discloses that “[t]he **main microphone** has directional characteristics” and includes two omnidirectional microphones M1 and M2 and associated signal processing (e.g., delay circuit 2 and inversion adding circuit 3).

Ex. 1006, ¶¶[0002]-[0004], [0018]-[0022], Figs. 1-3; Ex. 1003, ¶¶96-97. In the **main microphone**, output e1 and a delayed version of output e2 are combined (a first combination of a first microphone signal and a second microphone signal) at inversion adding circuit 3, which outputs main signal Vm. Ex. 1006, ¶¶[0018]-[0022], [0027]; Ex. 1003, ¶¶96-97.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

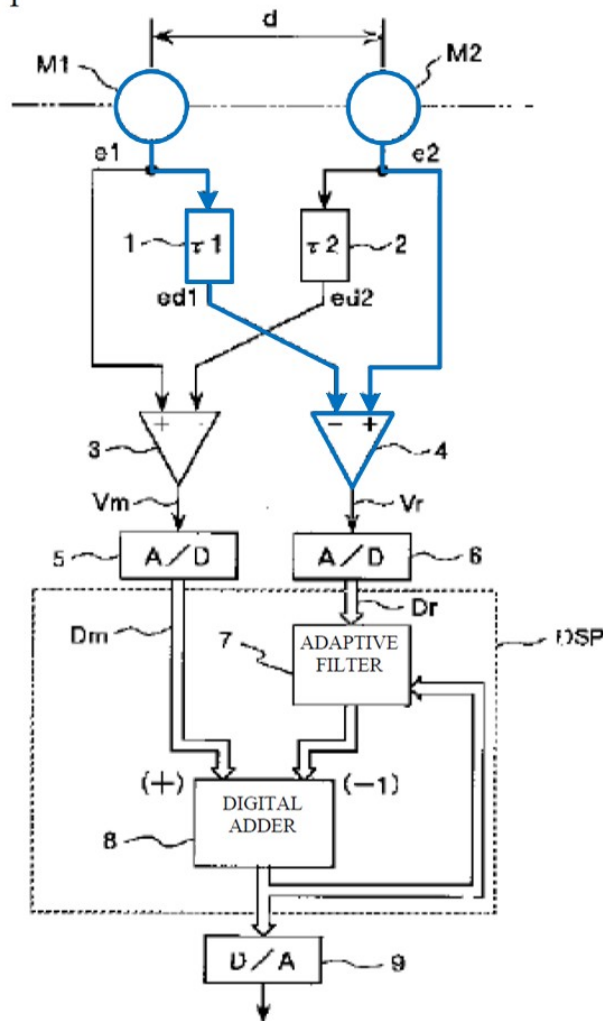
Additionally, the Ikeda-McCowan-Kanamori main microphone that generates main signal V_m (see § V.A.4) also teaches “a first virtual microphone formed from a first combination of a first microphone signal and a second microphone signal” for substantially the same reasons and because the Ikeda-McCowan-Kanamori main signal V_m is constructed using the same signal processing as the virtual microphone V_1 described in the '080 patent. Ex. 1001, 4:42-43, 11:1-2; Ex. 1003, ¶98; § V.A.4.a.

- b. [7b] “a second virtual microphone formed from a second combination of the first microphone signal and the second microphone signal, wherein the second combination is different from the first combination,”**

Ikeda discloses, and Ikeda-McCowan-Kanamori renders obvious, this limitation.

Ikeda discloses a second virtual microphone by disclosing the [reference microphone](#) that constructs reference signal V_r , as shown in Figure 1 (annotated below). Ex. 1003, ¶¶99-100. Ikeda discloses that “[t]he [reference microphone](#) has directional characteristics” and includes two omnidirectional microphones M1 and M2 and associated signal processing (e.g., delay circuit 1 and inversion adding circuit 4). Ex. 1006, ¶¶[0018]-[0020], [0023]-[0024], Figs. 1, 4-5; Ex. 1003, ¶¶99-100. In the [reference microphone](#), a delayed version of output e1 and output e2 are combined (a second combination of the first microphone signal and the second microphone signal) at inversion adding circuit 4, which outputs reference signal V_r . Ex. 1006, ¶¶[0018]-[0020], [0023]-[0024]; Ex. 1003, ¶¶99-100.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

Ikeda discloses that the second combination (as discussed above) is different from the first combination (as discussed in Section V.A.5.a). Ex. 1006, ¶¶[0018]-[0024]; Ex. 1003, ¶101.

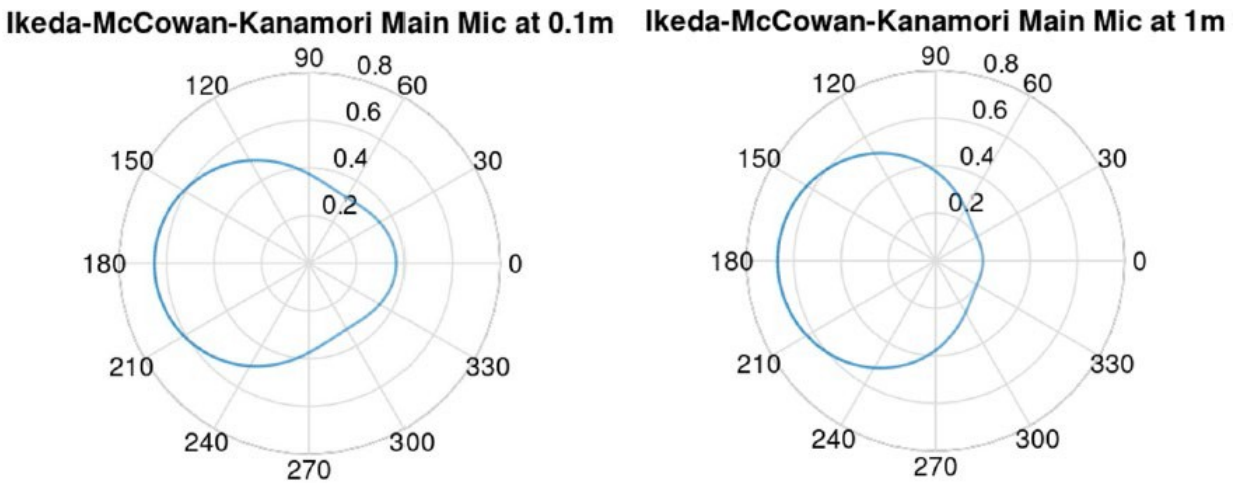
Additionally, the Ikeda-McCowan-Kanamori reference microphone that generates reference signal V_r (see § V.A.4) also teaches “a second virtual

microphone formed from a second combination of the first microphone signal and second microphone signal” for substantially the same reasons and because the Ikeda-McCowan-Kanamori reference signal V_r is constructed using the same signal processing as the virtual microphone V_2 described in the ’080 patent. Ex. 1001, 4:44-45, 9:51-53; Ex. 1003, ¶102; § V.A.4.a. In Ikeda-McCowan-Kanamori, the second combination of the first microphone signal and the second microphone signal is different from the first combination of the first microphone signal and the second microphone signal, because the Ikeda-McCowan-Kanamori main and reference microphones have different formulations for signal processing. Ex. 1003, ¶102; § V.A.4.a.

c. [7c] “wherein the first virtual microphone has a first linear response to speech and first linear response to noise,”

Ikeda-McCowan-Kanamori renders this limitation obvious.

Dr. Vipperman simulated and plotted the linear responses of the Ikeda-McCowan-Kanamori main signal V_m (see §§ V.A.4, V.A.4.a) to a sound source located 0.1 m from the microphone array of microphones M1 and M2 (near-field speech), and to a sound source located 1 m from the microphone array (far-field noise):



Ex. 1003, ¶¶82-93, 103-105 (linear response to speech (left)
and linear response to noise (right)).

These plots are the same types of plots that the '080 patent calls the “linear responses” of the virtual microphone. Ex. 1001, 2:52-3:4, 10:18-27, 11:20-35, Figs. 9-13; Ex. 1003, ¶106. A POSITA would have understood that the Ikeda- McCowan-Kanamori main signal V_m has linear responses because the main signal V_m 's response to a sound source at a particular distance, frequency, and direction is linearly proportional to its input. Ex. 1003, ¶106. For example, as depicted in the above plots, the main signal V_m 's response to a 1000 Hz speech source at 0.1 m and 0° can be determined by multiplying the input amplitude by about 0.36, and the main signal V_m 's response to a 1000 Hz noise source at 1 m and 90° can be determined by multiplying the input amplitude by about 0.38. Ex. 1003, ¶106. Moreover, the Ikeda-McCowan-Kanamori main signal V_m has linear responses just like the virtual

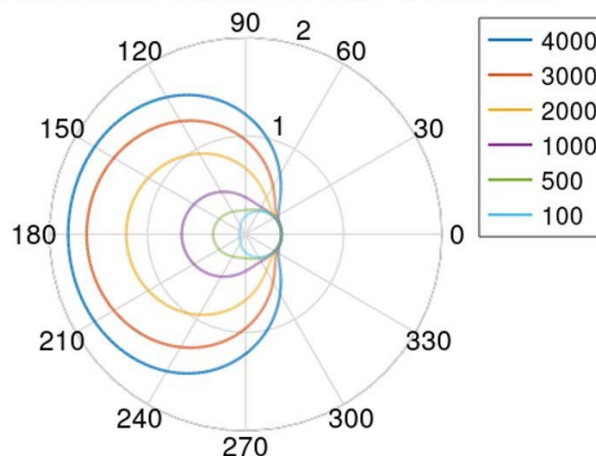
microphone V_1 described in the '080 patent has linear responses, because they are constructed using the same signal processing. Ex. 1001, 2:60-3:4, 4:42-43, 11:1-2, 11:20-35, Figs. 11-13; Ex. 1003, ¶106; § V.A.4.a. Thus, the plots show that the first virtual microphone has a first linear response to speech and a first linear response to noise. Ex. 1003, ¶¶103-106.

- d. [7d] “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source located within a predetermined angle relative to an axis of the microphone array and devoid of a null,”

Ikeda-McCowan-Kanamori renders this limitation obvious.

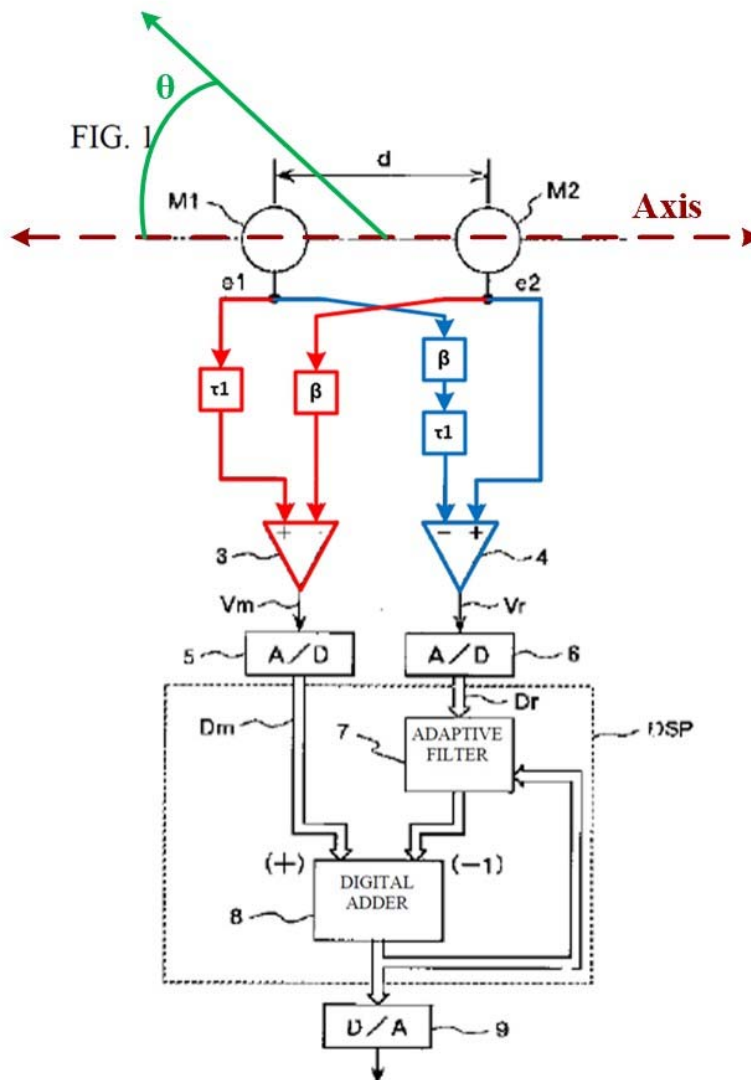
Dr. Vipperman simulated and plotted the linear response of the Ikeda-McCowan-Kanamori main signal V_m to a sound source of various frequencies located 0.1 m from the microphone array of microphones M1 and M2:

Ikeda-McCowan-Kanamori Main Mic at 0.1m



Ex. 1003, ¶¶82-93, 107 (linear response to speech).

In the above depiction of main signal V_m 's linear response to a near-field speech source, the angle corresponds to the value of θ in the figure below. Ex. 1003, ¶108. As depicted above, the linear response to speech is substantially similar across a plurality of frequencies for between approximately ± 30 degrees relative to an axis of the microphone array of microphones M1 and M2, which includes the direction of speech. Ex. 1006, ¶¶[0019] (microphone M1 "is disposed on the front surface"), [0021], [0023] ("target sound ... from the front surface side"), Figs. 1-5; Ex. 1003, ¶108. In other words, for this direction range, the main signal V_m 's response to a speech source between 100 and 4000 Hz at 0.1 m can be determined by multiplying the input amplitude by about 0.36. Ex. 1003, ¶108. Accordingly, the linear response to speech is substantially similar across a plurality of frequencies for a speech source located within an angle of approximately ± 30 degrees relative to an axis of the microphone array. Ex. 1003, ¶¶107-108. As depicted above, the linear response to speech is devoid of a null. Ex. 1001, 4:31-32 (defining the term "null"); Ex. 1003, ¶¶107-108.



Ex. 1006, Fig. 1 (modified).

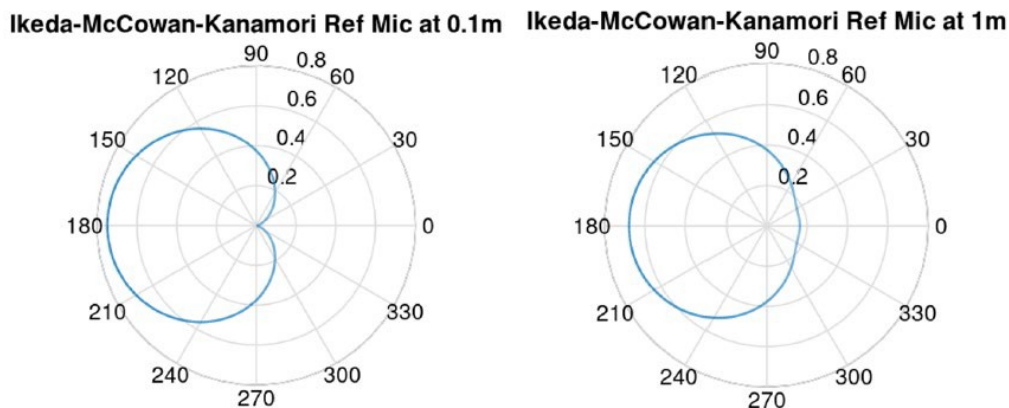
Additionally, the Ikeda-McCowan-Kanamori main signal V_m , which is constructed using the same signal processing as virtual microphone V_1 described in the '080 patent, teaches the features recited in [7d], because virtual microphone V_1

has the features. Ex. 1001, 2:60-63, 3:1-4, 4:42-43, 11:1-2, 11:20-24, 11:30-61, Figs. 11, 13; Ex. 1003, ¶109; § V.A.4.a.

- e. [7e] “wherein the second virtual microphone has a second linear response to speech that has a single null oriented in a direction toward a source of the speech and a second linear response to noise,”

Ikeda-McCowan-Kanamori renders this limitation obvious.

Dr. Vipperman simulated and plotted the linear responses of the Ikeda-McCowan-Kanamori reference signal V_r (see §§ V.A.4, V.A.4.a) to a sound source located 0.1 m from the microphone array of microphones M1 and M2 (near-field speech), and to a sound source located 1 m from the microphone array (far-field noise):



Ex. 1003, ¶¶82-93, 110-112 (linear response to speech (left) and linear response to noise (right)).

These plots are the same types of plots that the '080 patent calls the “linear responses” of the virtual microphone. Ex. 1001, 2:52-3:4, 10:18-27, 11:20-35, Figs. 9-13; Ex. 1003, ¶113. A POSITA would have understood that the Ikeda-McCowan-Kanamori reference signal V_r has linear responses because the reference signal V_r 's response to a sound source at a particular distance, frequency, and direction is linearly proportional to its input. Ex. 1003, ¶113. For example, as depicted in the above plots, the reference signal V_r 's response to a 1000 Hz speech source at 0.1 m and 0° can be determined by multiplying the input amplitude by about 0, and the reference signal V_r 's response to a 1000 Hz noise source at 1 m and 90° can be determined by multiplying the input amplitude by about 0.38. Ex. 1003, ¶113. Moreover, the Ikeda-McCowan-Kanamori reference signal V_r has linear responses just like virtual microphone V_2 described in the '080 patent has linear responses, because they are constructed using the same signal processing. Ex. 1001, 2:52-59, 4:44-45, 9:51-53, 10:18-35, Figs. 9-10; Ex. 1003, ¶113; § V.A.4.a. Therefore, the plots show the second virtual microphone has a second linear response to speech and a second linear response to noise. Ex. 1003, ¶¶110-113.

As also shown above, the Ikeda-McCowan-Kanamori reference signal V_r 's linear response to speech has a single null oriented at 0 degrees, which in the above

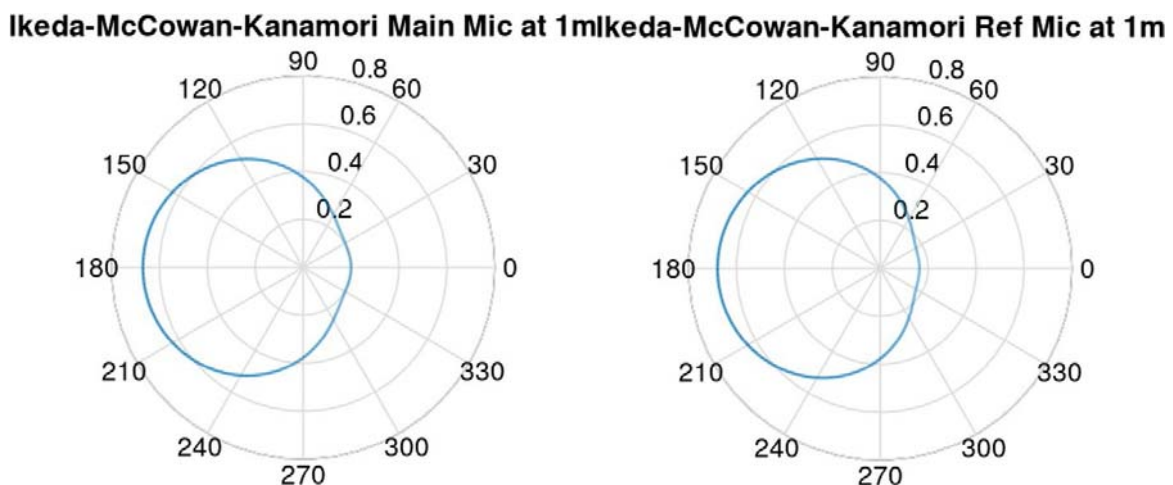
simulation is the direction toward a speech source. Ex. 1001, 4:31-32; Ex. 1003, ¶114; *see also* §§ V.A.4, V.A.5.d.

Additionally, the Ikeda-McCowan-Kanamori reference signal V_r , which is constructed using the same signal processing as virtual microphone V_2 described in the '080 patent, teaches the features recited in [7e], because virtual microphone V_2 has the features. Ex. 1001, 2:52-59, 4:44-45, 9:51-53, 10:18-35, Figs. 9-10; Ex. 1003, ¶¶115; § V.A.4.a.

f. [7f] “wherein the second linear response to noise is substantially similar to the first linear response to noise,”

Ikeda-McCowan-Kanamori renders this limitation obvious.

As discussed in Sections V.A.5.c and V.A.5.e, in the Ikeda-McCowan-Kanamori system, the first linear response to noise and the second linear response to noise are shown below:



Ex. 1003, ¶¶82-93, 116 (linear responses to noise).

As depicted, the second linear response to noise (right) is substantially similar to the first linear response to noise (left). Ex. 1003, ¶¶116-117.

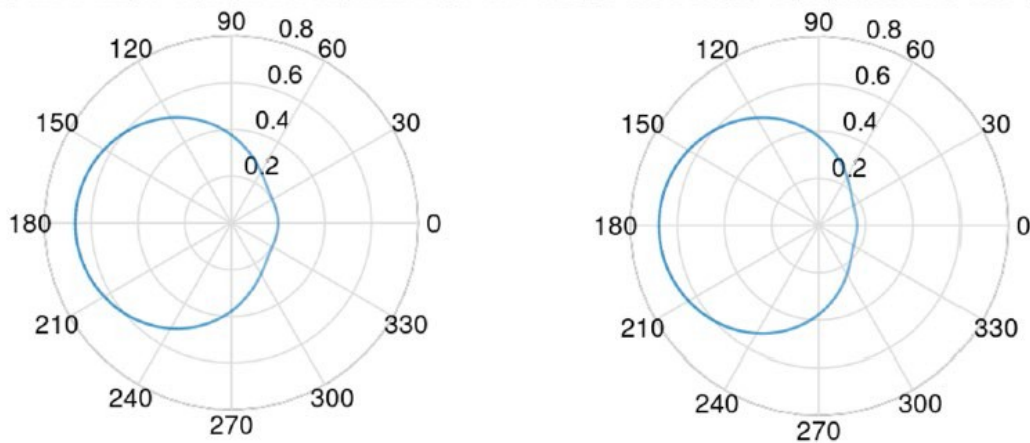
Additionally, the Ikeda-McCowan-Kanamori main and reference signals, which are constructed using the same signal processing as virtual microphones V_1 and V_2 described in the '080 patent, teach the features recited in [7f], because virtual microphones V_1 and V_2 have the features. Ex. 1001, 11:14-16, 16:4-6; Ex. 1003, ¶118; § V.A.4.a; *see also* Ex. 1001, 2:56-59, 2:64-67, 10:22-27, 11:25-29, 11:36-39, Figs. 10, 12.

- g. [7g] “one or both of the first linear response to noise and the second linear response to noise being nonzero in a direction toward a source of noise, and”**

Ikeda-McCowan-Kanamori renders this limitation obvious.

As discussed in Sections V.A.5.c and V.A.5.e, in the Ikeda-McCowan-Kanamori system, the first linear response to noise and the second linear response to noise are shown below:

Ikeda-McCowan-Kanamori Main Mic at 1m Ikeda-McCowan-Kanamori Ref Mic at 1m



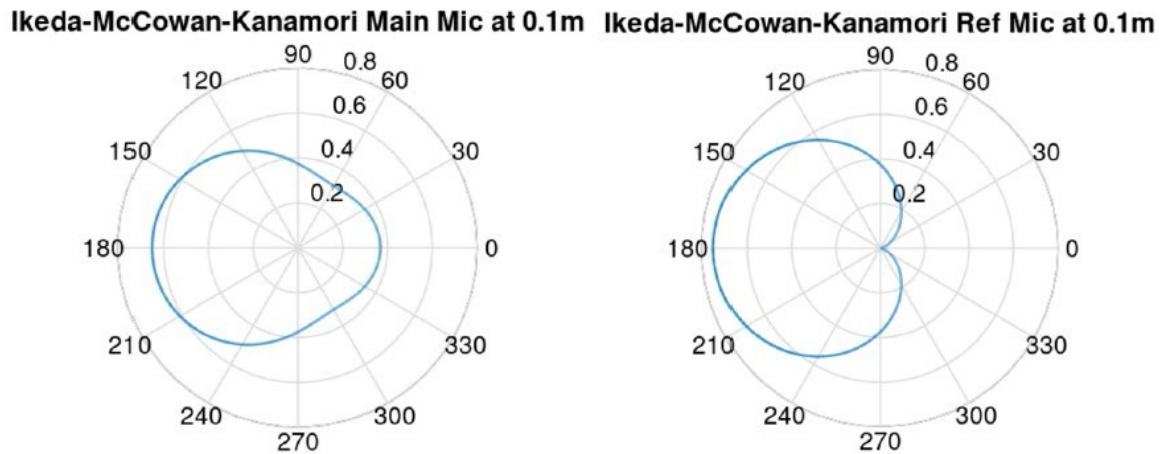
Ex. 1003, ¶¶82-93, 119 (linear responses to noise).

As depicted, both linear responses to noise are non-zero in all directions and therefore are non-zero in a direction toward a source of noise. Ex. 1003, ¶¶119-121.

h. [7h] “the second linear response to speech is substantially dissimilar to the first linear response to speech,”

Ikeda-McCowan-Kanamori renders this limitation obvious.

As discussed in Sections V.A.5.c and V.A.5.e, in the Ikeda-McCowan-Kanamori system, the first linear response to speech and the second linear response to speech are shown below:



Ex. 1003, ¶¶82-93, 122 (linear responses to speech).

As shown above, the second linear response to speech (right) is substantially dissimilar to the first linear response to speech (left), as the second linear response to speech includes a null in the direction of speech and the first linear response to speech does not. Ex. 1003, ¶¶122-123.

Additionally, the Ikeda-McCowan-Kanamori main and reference signals, which are constructed using the same signal processing as virtual microphones V_1 and V_2 described in the '080 patent, teach the features recited in [7h], because virtual microphones V_1 and V_2 have the features. Ex. 1001, 16:4-6; Ex. 1003, ¶124; § V.A.4.a; *see also* Ex. 1001, 2:52-55, 2:60-63, 10:18-32, 11:14-24, 11:36-39, Figs. 9, 11.

i. [7i] “wherein the speech is human speech; and”

Ikeda discloses or renders obvious this limitation.

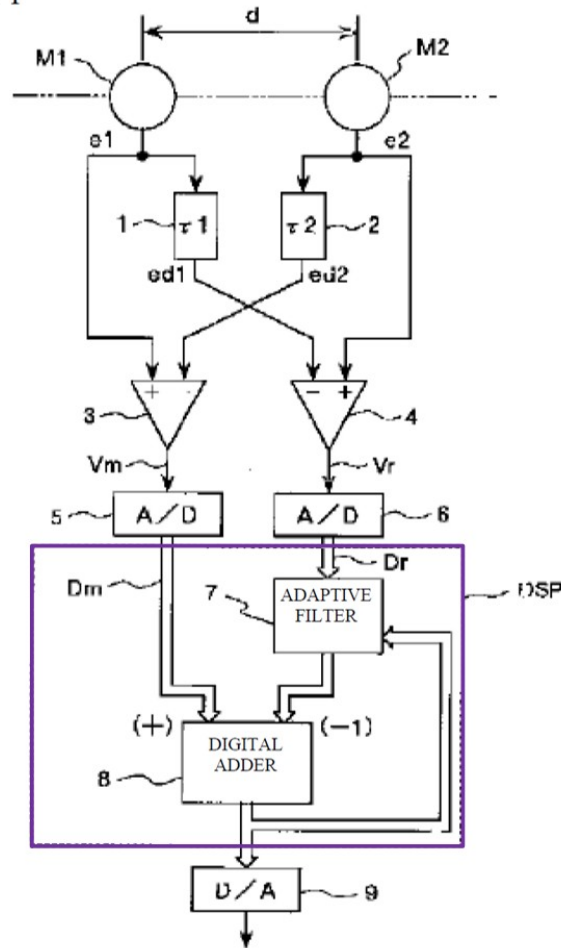
Ikeda discloses applying its teachings “to a wide range of devices,” including “a voice input microphone of a hands-free telephone unit or a car navigation system.” Ex. 1006, ¶[0036]. A POSITA would have understood that “voice input” refers to human speech. Ex. 1006, ¶[0036]; Ex. 1003, ¶125. It would have been obvious to a POSITA to use Ikeda’s system to receive human speech. Ex. 1003, ¶125.

- j. [7j] **“an adaptive noise removal application coupled to the first and second virtual microphones and generating denoised output signals by forming a plurality of combinations of signals output from the first virtual microphone and the second virtual microphone, by filtering and summing the plurality of combinations of signals in the time domain, and by a varying linear transfer function between the plurality of combinations of signals,”**

Ikeda discloses or renders obvious this limitation.

Ikeda discloses that an adaptive noise cancellation system including an adaptive filter 7 and a digital adder 8 as shown in Figure 1 (annotated below) is an adaptive noise removal application. Ex. 1006, ¶¶[0027]-[0029]; Ex. 1003, ¶126. Ikeda discloses that the **adaptive noise cancellation system** is implemented using a digital signal processor and therefore is an “application.” Ex. 1006, ¶¶[0016]-[0017], [0030]; Ex. 1003, ¶126.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

Ikeda discloses that the **adaptive noise cancellation system** is coupled to the main and reference microphones and the corresponding main signal V_m and reference signal V_r (“the first and second virtual microphones”) via A/D converters 5 and 6. Ex. 1006, ¶¶[0027]-[0029], Fig. 1; Ex. 1001, 32:25-32; Ex. 1003, ¶127.

Ikeda discloses that the **adaptive noise cancellation system** generates denoised output signals. Ex. 1006, ¶¶[0028]-[0029]; Ex. 1003, ¶128. Ikeda discloses that “the

noise component contained in the main signal V_m is cancelled by the noise component contained in the reference signal V_r due to subtraction by digital adder 8,” thereby generating denoised output signals. Ex. 1006, ¶¶[0028]-[0029]; Ex. 1003, ¶128.

Ikeda’s **adaptive noise cancellation system** forms a plurality of combinations of signals output from the first virtual microphone and the second virtual microphone because main signal V_m and reference signal V_r are continuously converted to digital signals and input to the **adaptive noise cancellation system** implemented on a digital signal processor DSP so that, over time, they form multiple combinations of signals output from the first and second virtual microphones. Ex. 1006, ¶¶[0027]-[0029]; Ex. 1003, ¶129; *see also* Ex. 1001, 32:61-63 (“Words using the singular or plural number also include the plural or singular number respectively.”). This understanding is consistent with the ’080 patent’s description that an adaptive noise removal system “combin[es] the two microphone signals (e.g., Mic1, Mic2) by filtering and summing in the time domain.” Ex. 1001, 7:4-16; Ex. 1003, ¶129; *see also* Ex. 1001, 8:50-9:3, 9:22-38.

Ikeda’s **adaptive noise cancellation system** filters and sums the plurality of combinations of signals in the time domain because Ikeda discloses that digital versions of reference signal V_r are “supplied to an adaptive filter” (filtering) and the

output of adaptive filter 7 and digital versions of main signal V_m are differenced at digital adder 8 (summing). Ex. 1006, ¶¶[0027]-[0029]; Ex. 1003, ¶130. This understanding agrees with the '080 patent's description that an adaptive noise removal system "combin[es] the two microphone signals (e.g., Mic1, Mic2) by filtering and summing in the time domain." Ex. 1001, 7:4-16; Ex. 1003, ¶130; *see also* Ex. 1001, 8:50-9:3, 9:22-38. A POSITA would have understood that adaptive filter 7 filters in the time domain, and digital adder 8 adds in the time domain because (1) Ikeda's system does not include components for converting signals from the time domain to the frequency domain or from the frequency domain to the time domain, and (2) a signal picked up by a microphone is in the time domain if it is not processed by a component and converted to the frequency domain. Ex. 1006, ¶¶[0027]-[0029]; Ex. 1009, [0042]-[0043], [0058]-[0062]; Ex. 1003, ¶130. Additionally, it would have been obvious to a POSITA to implement Ikeda's system including adaptive filter 7 and digital adder 8 to process signals in the time domain. Ex. 1003, ¶130.

Ikeda's **adaptive noise cancellation system** discloses a varying transfer function between the plurality of combinations of signals because Ikeda discloses that adaptive filter 7's characteristics "are varied in real time" by "a tap coefficient updating operation for updating the tap coefficient based on an error signal obtained as the difference between the filter output and a signal from the target system to

return the filter output,” which is successively repeated to rewrite the tap coefficient. Ex. 1006, ¶¶[0027]-[0029]; Ex. 1003, ¶131. This “error signal is defined as the output of the digital adder 8,” i.e., where the plurality of combinations of signals are summed. Ex. 1006, ¶[0028]; Ex. 1003, ¶131. Ikeda discloses performing this tap coefficient updating operation when “the microphone units M1 and M2 do not pick up the target sound,” which is the time period in which the output of digital adder 8 should be zero (because there is no speech). Ex. 1006, ¶[0028]; Ex. 1003, ¶131. By updating the tap coefficient over time to minimize this error signal, the adaptive filter calculates and applies a varying transfer function between the plurality of combinations of signals that results in the output of adaptive filter 7 equaling the digital signal from main signal V_m at a time when there is no speech. Ex. 1006, ¶[0028]; Ex. 1003, ¶131. This understanding comports with the ’080 patent’s description that “[t]he adaptive filter generally uses the signal received from a first microphone ... to remove noise from the speech received from at least one other microphone ..., which relies on a slowly varying linear transfer function between the two microphones for sources of noise.” Ex. 1001, 7:4-16; Ex. 1003, ¶131; *see also* Ex. 1001, 5:19-44, 6:26-29, 8:50-9:3, 9:22-38.

Ikeda also discloses that this transfer function is linear because it discloses that adaptive filter 7 is “a transversal filter” and “perform[s] an output operation for

obtaining filter output by multiplying the input signal D_r by a tap coefficient and adding the value in multiple tap units.” Ex. 1006, ¶¶[0027]-[0029]; Ex. 1003, ¶132. A POSITA would have understood that this discloses a linear filter. Ex. 1020 (Widrow), 16-17, Fig. 2.2 (“Adaptive linear combiner in the form of single-input adaptive transversal filter”); Ex. 1029; Ex. 1030; Ex. 1028, 366-367; Ex. 1031; Ex. 1032; Ex. 1003, ¶132. It also would have been obvious to a POSITA to use a linear filter to implement adaptive filter 7. It was well-known to use linear filters for adaptive noise cancellation, and it would have been an obvious design choice. Ex. 1010, 4:28-5:15, Fig. 2; Ex. 1003, ¶132.

k. [7k] “wherein the denoised output signals include less acoustic noise than acoustic signals received at the first and second physical microphones.”

Ikeda discloses this limitation.

Ikeda discloses that the output signals of the adaptive noise cancellation system (“the denoised output signals”) include less acoustic noise than acoustic signals received at microphones M1 and M2 (“the first and second physical microphones”), because, unlike the acoustic signals received at microphones M1 and M2, the denoised output signals have had “the noise component contained in the main signal V_m ” canceled and “primarily consist[] of the target sound component.” Ex. 1006, ¶¶[0001], [0011]-[0012], [0029], Abstract; Ex. 1003, ¶133.

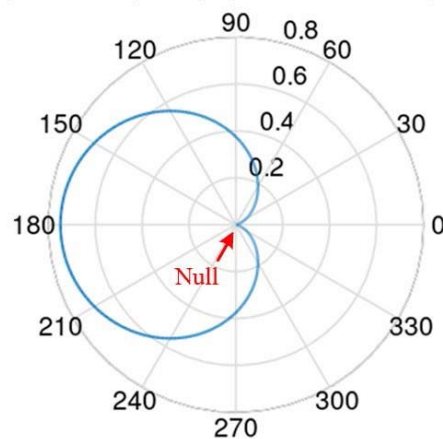
6. Claim 8: “The system of claim 7 and further comprising: a microphone array, the first and second physical microphones positioned m [sic] the microphone array.”

Ikeda discloses a microphone array including microphones M1 and M2 (“the first and second physical microphones”), which are “positioned m [sic] the microphone array.” Ex. 1001, Claim 8; Ex. 1006, ¶¶[0018], Fig. 1; Ex. 1002, 146; Ex. 1003, ¶¶134-135.

7. Claim 9: “The system of claim 7, wherein the single null is a region of the second linear response to speech having a measured response level that is lower than the measured response level of any other region of the second linear response to speech.”

Ikeda-McCowan-Kanamori renders this limitation obvious. As shown in the plot of the Ikeda-McCowan-Kanamori second linear response to speech (annotated below), the single **null** is a region having a measured response level that is lower than the measured response level of any other region. § V.A.5.e; Ex. 1003, ¶¶136-138; *see also* Ex. 1001, 2:52-55, 10:18-32, Fig. 9.

Ikeda-McCowan-Kanamori Ref Mic at 0.1m



Ex. 1003, ¶¶82-93, 137 (linear response to speech).

8. **Claim 11: “The system of claim 7 and further comprising: a communications channel coupled with the processing component and including one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel.”**

Ikeda renders this limitation obvious.

Ikeda discloses a processing component. Ex. 1006, ¶¶[0016]-[0017], [0030] (“the adaptive filter 7, the digital adder 8, and the like can be configured a digital signal processor DSP”); Ex. 1003, ¶¶139-140.

Ikeda discloses applying its disclosures “to a wide range of devices,” including “a voice input microphone of a hands-free telephone unit.” Ex. 1006, ¶[0036]. It would have been obvious to a POSITA that there is a communications channel between the “voice input microphone” to which Ikeda’s system is applied and the “hands-free telephone unit.” Ex. 1003, ¶141. A POSITA would therefore

have understood that the communications channel is coupled with the processing component in Ikeda's system. Ex. 1003, ¶141.

It would have been obvious to a POSITA that the communications channel includes one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel. Ex. 1003, ¶¶141-142. It was well-known to implement such a communications channel using one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel. Ex. 1003, ¶142; Ex. 1015 (Baranowski), 2:58-64, Fig. 2; Ex. 1016 (Wingate), 4:1-5:25, Figs. 3-4; Ex. 1017 (Rydbeck), 3:20-25, 3:50-64, Claim 1, Fig. 4b; Ex. 1018 (Kimm), ¶¶[0029]-[0030]; Ex. 1019 (Burnett II), ¶[0138].

**9. Claim 12: “The system of claim 11 and further comprising:
a communication device wirelessly coupled with the wireless
channel of the communications channel.”**

Ikeda discloses that the “hands-free telephone unit” is a communication device. Ex. 1006, ¶[0036]; Ex. 1003, ¶¶143-144. As discussed in Section V.A.8, the communications channel is between the “voice input microphone” to which Ikeda's system is applied and the “hands-free telephone unit.” Ex. 1003, ¶¶143-144. It would have been obvious to a POSITA that the communication device is wirelessly coupled with the wireless channel of the communications channel. Ex. 1003, ¶¶143-144.

- 10. Claim 13: “The system of claim 7, wherein the second microphone signal is multiplied by a ratio, wherein the ratio is a ratio of a third distance to a fourth distance, the third distance being between the first physical microphone and the speech source and the fourth distance being between the second physical microphone and the speech source.”**

Ikeda-McCowan-Kanamori renders this feature obvious.

In Ikeda-McCowan-Kanamori, the output signal of microphone M2 (“the second microphone signal”) is multiplied by the gain β for the main microphone. § V.A.4; Ex. 1003, ¶¶145-147. As discussed in Section V.A.4, the gain β is a ratio of a distance between microphone M1 (“the first physical microphone”) and the near- field speech source to a distance between microphone M2 (“the second physical microphone”) and the near-field speech source. Ex. 1007, 87-91, 93-96; Ex. 1003, ¶¶145-147.

11. Claim 14

- a. [14a] “A system, comprising: a first virtual microphone comprising a first combination of a first microphone signal and a second microphone signal,”**

Ikeda-McCowan-Kanamori renders obvious [14a] for substantially the same reasons discussed in Section V.A.5.a regarding [7a]. Ex. 1003, ¶148.

- b. [14b] “the first virtual microphone having a first linear response to speech and a first linear response to noise,”**

Ikeda-McCowan-Kanamori renders obvious [14b] for substantially the same reasons discussed in Section V.A.5.c regarding [7c]. Ex. 1003, ¶149.

- c. **[14c] “the first linear response to speech being substantially similar across a plurality of frequencies for a speech source located within a predetermined angle relative to an axis of a microphone array,”**

Ikeda-McCowan-Kanamori renders obvious [14c] for substantially the same reasons discussed in Section V.A.5.d regarding [7d]. Ex. 1003, ¶150.

- d. **[14d] “wherein the first microphone signal is output from a first physical microphone and the second microphone signal is output from a second physical microphone;”**

Ikeda-McCowan-Kanamori renders obvious [14d] for substantially the same reasons discussed in Section V.A.5.a regarding [7a]. Ex. 1003, ¶151.

- e. **[14e] “a second virtual microphone comprising a second combination of the first microphone signal and the second microphone signal,”**

Ikeda-McCowan-Kanamori renders obvious [14e] for substantially the same reasons discussed in Section V.A.5.b regarding [7b]. Ex. 1003, ¶152.

- f. **[14f] “the second virtual microphone having a second linear response to speech and a second linear response to noise,”**

Ikeda-McCowan-Kanamori renders obvious [14f] for substantially the same reasons discussed in Section V.A.5.e regarding [7e]. Ex. 1003, ¶153.

- g. **[14g] “the second linear response to noise being substantially similar to the first linear response to noise,”**

Ikeda-McCowan-Kanamori renders obvious [14g] for substantially the same reasons discussed in Section V.A.5.f regarding [7f]. Ex. 1003, ¶154.

- h. [14h] “one or both of the first linear response to noise and the second linear response to noise being non- zero in a direction toward a source of noise, and”**

Ikeda-McCowan-Kanamori renders obvious [14h] for substantially the same reasons discussed in Section V.A.5.g regarding [7g]. Ex. 1003, ¶155.

- i. [14i] “the second linear response to speech being substantially dissimilar to the first linear response to speech,”**

Ikeda-McCowan-Kanamori renders obvious [14i] for substantially the same reasons discussed in Section V.A.5.h regarding [7h]. Ex. 1003, ¶156.

- j. [14j] “wherein the second combination is different from the first combination,”**

Ikeda-McCowan-Kanamori renders obvious [14j] for substantially the same reasons discussed in Section V.A.5.b regarding [7b]. Ex. 1003, ¶157.

- k. [14k] “wherein the first virtual microphone and the second virtual microphone are distinct virtual directional microphones; and”**

Ikeda, or Ikeda-McCowan-Kanamori, teaches [14k] for substantially the same reasons discussed in Sections V.A.5.a and V.A.5.b regarding [7a] and [7b]. Ex. 1003, ¶¶158-159. In either Ikeda or Ikeda-McCowan-Kanamori, the main microphone and its corresponding main signal V_m (first virtual directional microphone) are distinct from the reference microphone and its corresponding reference signal V_r (second virtual directional microphone) because the first and

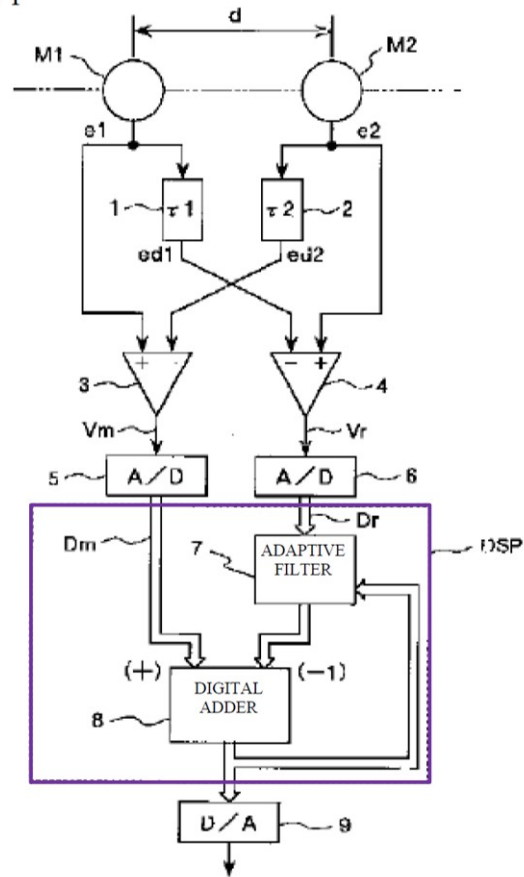
second virtual directional microphones are formed with different combinations of the signals from microphones M1 and M2. Ex. 1006, ¶¶[0002]-[0004], [0018]-[0024], Figs. 1-5; Ex. 1001, 4:48-51, 7:32-35, 7:66-8:2; Ex. 1003, ¶¶158-159; §§V.A.4, V.A.5.a, V.A.5.b, V.A.5.c, V.A.5.e.

- I. [14l] “a processing component coupled to the first and second virtual microphones, the processing component including an adaptive noise removal application receiving acoustic signals from the first virtual microphone and the second virtual microphone, filtering and summing the acoustic signals in the time domain, applying a varying linear transfer function between the acoustic signals, and generating an output signal,”

Ikeda-McCowan-Kanamori renders obvious [14l] for substantially the same reasons discussed in Section V.A.5.j regarding [7j]. Ex. 1003, ¶¶160-162.

Ikeda discloses that a **digital signal processor** DSP as shown in Figure 1 (annotated below) is a processing component including an adaptive noise removal application. Ex. 1006, ¶¶[0016]-[0017], [0027]-[0030]; Ex. 1003, ¶161; § V.A.5.j. The **digital signal processor** DSP is coupled to the main and reference microphones and corresponding main signal V_m and reference signal V_r via A/D converters 5 and 6. Ex. 1006, ¶[0027], Fig. 1; Ex. 1001, 32:25-32; Ex. 1003, ¶161.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

m. [14m]“wherein the output signal is a denoised acoustic signal.”

Ikeda-McCowan-Kanamori renders obvious [14m] for substantially the same reasons discussed in Sections V.A.5.j and V.A.5.k regarding [7j] and [7k].

Ex. 1003, ¶163.

- 12. Claim 16: “The system of claim 14 and further comprising: a communications channel coupled with the processing component and including one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel.”**

Ikeda-McCowan-Kanamori renders obvious this limitation for substantially the same reasons discussed in Section V.A.8 regarding claim 11. Ex. 1003, ¶¶164-165.

- 13. Claim 17: “The system of claim 16 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.”**

Ikeda-McCowan-Kanamori renders obvious this limitation for substantially the same reasons discussed in Section V.A.9 regarding claim 12. Ex. 1003, ¶¶166-167.

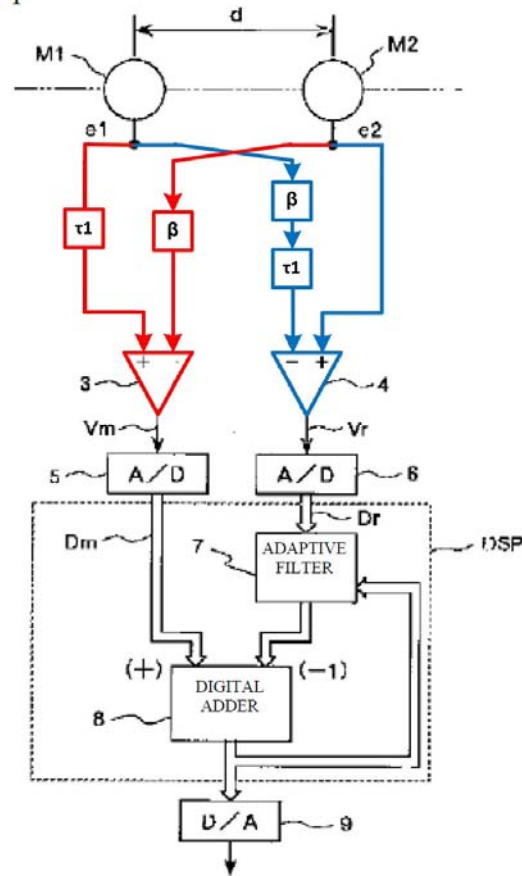
- 14. Claim 18: “The system of claim 14, wherein the acoustic signals from the first virtual microphone, the second virtual microphone or both are delayed.”**

Ikeda discloses this limitation. A POSITA would have understood that main signal V_m , reference signal V_r , or both (“the acoustic signals from the first virtual microphone, the second virtual microphone or both”) are delayed because of the processing by adaptive filter 7 and digital adder 8. Ex. 1006, ¶¶[0027]-[0029]; Ex. 1003, ¶¶168-169; *see also* Ex. 1010, 3:32-36, Fig. 2; Ex. 1008, ¶[0112].

To the extent that “the acoustic signals from the first virtual microphone, the second virtual microphone or both” is interpreted to include “the first microphone

signal,” “the second microphone signal,” or both (Ex. 1023, 97-106; Ex. 1001, 7:17-8:12, 9:46-10:2, 11:1-2, Figs. 3-5), Ikeda-McCowan-Kanamori renders this claim obvious. In Ikeda-McCowan-Kanamori, the output signal of microphone M1 (“the first microphone signal”) is delayed by τ_1 , as shown below. § V.A.4; Ex. 1006, ¶¶0018], Fig. 1; Ex. 1003, ¶170.

FIG. 1



Ex. 1006, Fig. 1 (modified).

15. Claim 19: “The system of claim 18, wherein the delay is raised to a power that is proportional to a time difference between arrival of the speech at the first virtual microphone and arrival of the speech at the second virtual microphone.”

To the extent this claim is not indefinite and “the acoustic signals from the first virtual microphone, the second virtual microphone or both” as recited in claim 18 is interpreted to include “the first microphone signal,” “the second microphone signal,” or both (Ex. 1023, 97-106; Ex. 1001, 7:17-8:12, 9:46-10:2, 11:1-2, Figs. 3-5), Ikeda-McCowan-Kanamori renders claim 19 obvious.

The '080 patent defines a value γ as follows:

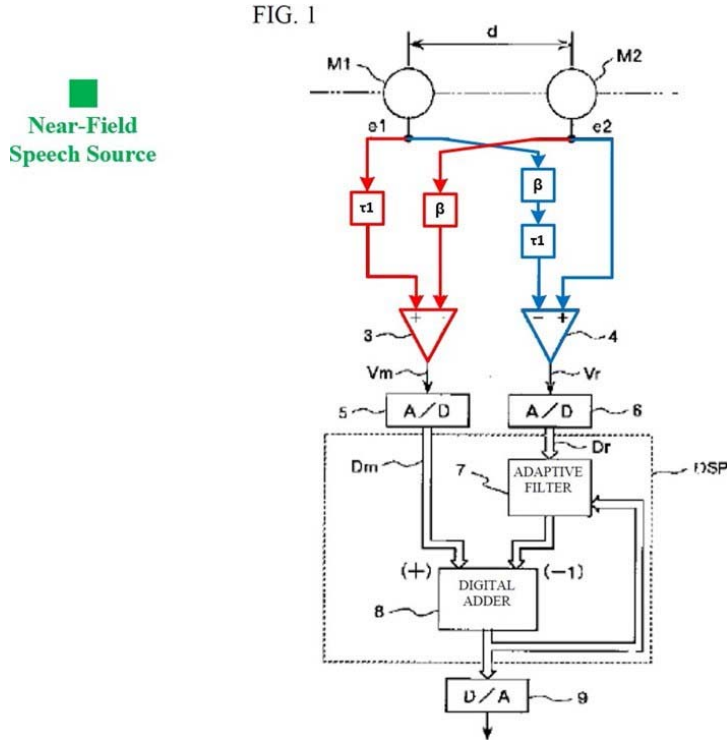
$$\gamma = \frac{d_2 - d_1}{c} \cdot f_s$$

where d_1 and d_2 “are the distance from O_1 and O_2 to the speech source” respectively, c is the speed of sound, and f_s is a sampling frequency. Ex. 1001, 9:54-67. The '080 patent later refers to γ as “the time difference between arrival of speech at V_1 compared to V_2 .” Ex. 1001, 13:66-14:2.

Based on the '080 patent's reference to γ as “the time difference between arrival of speech at V_1 compared to V_2 ” (Ex. 1001, 13:66-14:2), which Petitioner adopts for the purpose of this IPR only, Ikeda-McCowan-Kanamori renders claim 19 obvious.

Ikeda discloses that τ_1 is equal to the distance between microphones M1 and M2 divided by the speed of sound. Ex. 1006, ¶¶[0018] (microphones M1 and M2 “are disposed at a distance d apart from one another;” “ τ_1 : amount of delay expressed in terms of distance”), [0023]-[0024], [0026] (“ $\tau_1=d$ ”), Figs. 1, 4-5, 10; Ex. 1007, 89-90, 93-96; Ex. 1003, ¶¶171-175; *see also* § V.A.4.

In Ikeda-McCowan-Kanamori, the **near-field speech source** would have been located on the axis of microphones M1 and M2 as shown below. Ex. 1006, ¶¶[0019] (microphone M1 “is disposed on the front surface”), [0021], [0023] (“target sound ... from the front surface side”), Figs. 1-5; Ex. 1003, ¶176; *see also* § V.A.4. Thus, the distance between microphones M1 and M2 is a distance between microphone M1 and the speech source subtracted from a distance between microphone M2 and the speech source. Ex. 1003, ¶176.



Ex. 1006, Fig. 1 (modified).

Therefore, τ_1 is a time difference between arrival of the speech at microphone M1 and arrival of the speech at microphone M2. Ex. 1003, ¶177. A POSITA would have understood that, in an equation representing either main signal V_m or reference signal V_r , the delay of the output signal of microphone M1 would be expressed as $e^{-j\omega\tau_1}$, which is “raised to a power that is proportional to” the time difference τ_1 in the exponent. Ex. 1007, 90, 93-95; Ex. 1003, ¶177.

Additionally, it would have been obvious to implement the signal processing for the main and reference microphones digitally. Ex. 1006, ¶¶[0016]-[0017], [0030]; Ex. 1008, ¶[0151]; Ex. 1014, ¶¶[0021], [0026], [0065]; Ex. 1003, ¶178.

With a digital design, the delay of the output signal of microphone M1 would be expressed as z raised to a power proportional to a sampling frequency f_s (for sampling the analog output signal of microphone M1 into a digital signal) multiplied by the time length τ_1 of the delay (*e.g.*, $z^{z-T_1 \cdot f_s}$). Ex. 1003, ¶178; *see also* Ex. 1008, ¶[0108], Fig. 5 (disclosing that z^{-1} refers to delaying by one sample). Therefore, in the digital design, the delay of the output signal of microphone M1 is “raised to a power that is proportional to” the time difference τ_1 , in an equation representing either main signal V_m or reference signal V_r . Ex. 1003, ¶178.

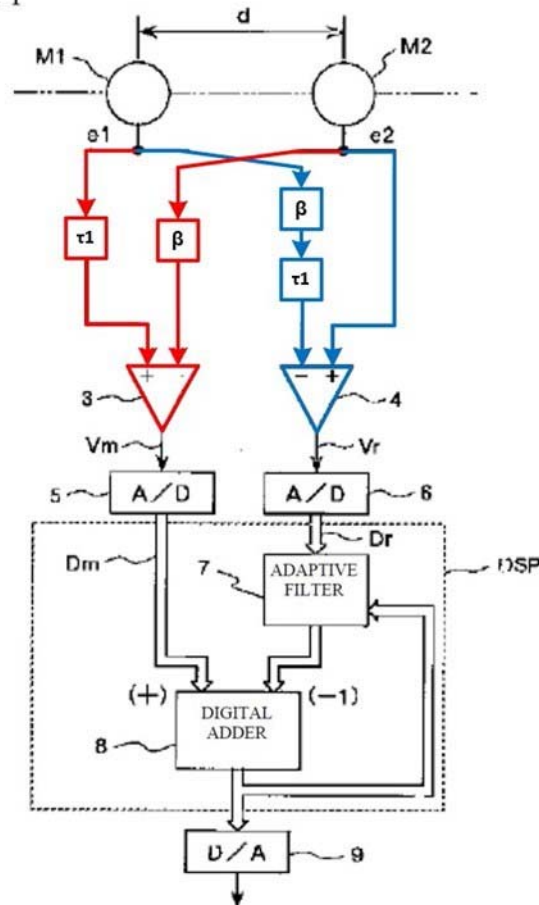
- 16. Claim 20: “The system of claim 19, wherein the power is proportional to a sampling frequency multiplied by a quantity equal to a third distance subtracted from a fourth distance, the third distance being between a first physical microphone and the speech source, the fourth distance being between a second physical microphone and the speech source, and the first and second physical microphones are positioned in the microphone array.”**

To the extent this claim is not indefinite and “the acoustic signals from the first virtual microphone, the second virtual microphone or both” as recited in claim 18 is interpreted to include “the first microphone signal,” “the second microphone signal,” or both (Ex. 1023, 97-106; Ex. 1001, 7:17-8:12, 9:46-10:2, 11:1-2, Figs. 3-5), Ikeda-McCowan-Kanamori renders claim 20 obvious.

As discussed in Section V.A.15 with respect to claim 19, with a digital design, the delay of the output signal of microphone M1 would be expressed as z raised to a power proportional to a sampling frequency f_s (for sampling the analog output signal of microphone M1 into a digital signal) multiplied by the time length τ_1 of the delay, which is the distance between microphones M1 and M2 divided by the speed of sound. Ex. 1003, ¶¶179-181. Because the speech source is located on the axis of microphones M1 and M2 as shown below (see § V.A.15), the distance between microphones M1 and M2 is a distance between microphone M1 and the speech source subtracted from a distance between microphone M2 and the speech source. Ex. 1003, ¶181.



FIG. 1



Ex. 1006, Fig. 1 (modified).

Therefore, in the digital design, the delay of the output signal of microphone M1 is raised to the power (an exponent in an equation representing either main signal Vm or reference signal Vr) that is proportional to a sampling frequency multiplied by a quantity equal to a third distance between a first physical microphone (microphone M1) and the speech source subtracted from a fourth distance between a second physical microphone (microphone M2) and the speech source. Ex. 1006, ¶¶[0018]-[0019], Fig. 1; Ex. 1003, ¶182.

Ikeda discloses that microphones M1 and M2 (“first and second physical microphones”) are positioned in the microphone array. Ex. 1006, ¶[0018], Fig. 1; Ex. 1003, ¶183.

17. Claim 1

- a. [1a] “A system, comprising: a microphone array including a first physical microphone outputting a first microphone signal and a second physical microphone outputting a second microphone signal;”**

Ikeda-McCowan-Kanamori renders obvious [1a] for substantially the same reasons discussed in Sections V.A.5.a and V.A.5.d regarding [7a] and [7d]. Ex. 1003, ¶184.

- b. [1b] “a processing component coupled to the microphone array and generating a virtual microphone array including a first virtual microphone and a second virtual microphone,”**

It would have been obvious to implement the signal processing for the main and reference microphones in Ikeda or Ikeda-McCowan-Kanamori digitally using a processor (“processing component”), which was a well-known technique. Ex. 1006, ¶¶[0016]-[0017], [0030] (“the adaptive filter 7, the digital adder 8, and the like can be configured a digital signal processor DSP”); Ex. 1008, ¶ [0151]; Ex. 1014, ¶¶[0021], [0026], [0065]; Ex. 1003, ¶185. A POSITA would have understood that such a processor is coupled to the microphone array of microphones M1 and M2 and generates main signal Vm and reference signal Vr. Ex. 1006, ¶¶[0018]-[0024],

Fig. 1; Ex. 1003, ¶185. The main signal Vm and reference signal Vr teach “a virtual microphone array including a first virtual microphone and a second virtual microphone” for substantially the same reasons discussed in Sections V.A.5.a and V.A.5.b regarding [7a] and [7b]. Ex. 1003, ¶185.

- c. [1c] **“the first virtual microphone including a first combination of the first microphone signal and the second microphone signal, the second virtual microphone including a second combination of the first microphone signal and the second microphone signal, wherein the second combination is different from the first combination,”**

Ikeda-McCowan-Kanamori renders obvious [1c] for substantially the same reasons discussed in Sections V.A.5.a and V.A.5.b regarding [7a] and [7b]. Ex. 1003, ¶186.

- d. [1d] **“wherein the first virtual microphone and the second virtual microphone have substantially similar responses to noise and substantially dissimilar responses to speech; and”**

Ikeda-McCowan-Kanamori renders obvious [1d] for substantially the same reasons discussed in Sections V.A.5.c, V.A.5.e, V.A.5.f, and V.A.5.h regarding [7c], [7e], [7f], and [7h]. Ex. 1003, ¶187.

- e. [1e] **“an adaptive noise removal application coupled to the processing component and generating denoised output signals by forming a plurality of combinations of signals output from the first virtual microphone and the second virtual microphone, by filtering and summing the plurality of combinations of signals in the time domain, and by a varying linear transfer function between the plurality of combinations of signals,”**

Ikeda-McCowan-Kanamori renders obvious [1e] for substantially the same reasons discussed in Section V.A.5.j regarding [7j]. Ex. 1003, ¶188.

A POSITA would have understood that Ikeda’s adaptive noise cancellation system is coupled to the processing component that generates the main and reference microphones. Ex. 1006, ¶¶[0027]-[0029], Fig. 1; Ex. 1003, ¶189; § V.A.17.b.

- f. [1f] **“wherein the denoised output signals include less acoustic noise than acoustic signals received at the microphone array.”**

Ikeda-McCowan-Kanamori renders obvious [1f] for substantially the same reasons discussed in Section V.A.5.k regarding [7k]. Ex. 1003, ¶190.

18. **Claim 2: “The system of claim 1, wherein the acoustic noise comprises noise content and the acoustic signals comprise speech content.”**

Ikeda discloses that the acoustic noise comprises noise content. Ex. 1006, ¶¶[0012] (disclosing noise “generated from every direction”), [0018]-[0024], [0027]-[0029], [0036]; Ex. 1003, ¶¶191-192.

Ikeda discloses or renders obvious that the acoustic signals comprise speech content. Ex. 1006, ¶[0036] (“a voice input microphone of a hands-free telephone unit”); Ex. 1003, ¶193; § V.A.5.i.

19. Claim 3: “The system of claim 2, wherein the speech content comprises human speech.”

Ikeda discloses or renders obvious that the speech content comprises human speech. Ex. 1006, ¶[0036]; Ex. 1003, ¶¶194-195; § V.A.5.i.

20. Claim 5: “The system of claim 1 and further comprising: a communications channel coupled with the processing component and including one or more of a wireless channel, a wired channel, and a hybrid wireless/wired channel.”

Ikeda-McCowan-Kanamori renders obvious this limitation for substantially the same reasons discussed in Section V.A.8 regarding claim 11. Ex. 1003, ¶¶196-197. A POSITA would have understood that, in Ikeda or Ikeda-McCowan-Kanamori, the communications channel is coupled with the processing component that generates the main and reference microphones. § V.A.17.b; Ex. 1006, ¶¶[0016]-[0024], [0030], Fig. 1; Ex. 1008, ¶[0151]; Ex. 1014, ¶¶[0021], [0026], [0065]; Ex. 1003, ¶197.

21. Claim 6: “The system of claim 5 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.”

Ikeda-McCowan-Kanamori renders obvious this limitation for substantially the same reasons discussed in Section V.A.9 regarding claim 12. Ex. 1003, ¶¶198-199.

B. Ground 2: Ikeda in view of McCowan, Kanamori, and Yang Renders Obvious Claims 4, 10, and 15

1. Overview of Yang

Yang discloses an adaptive noise reduction system as shown in Figure 4 (annotated below). Ex. 1009, ¶¶[0058]-[0062]. Yang discloses:

The mostly noise signal $x(t)$ is provided to an adaptive filter 450, which filters the noise with a particular transfer function $h(t)$. The filtered noise [sic] $p(t)$ is then provided to summer 434 and subtracted from the filtered speech plus noise signal to provide an intermediate signal $d(t)$ having predominantly speech and some amount of noise.

Ex. 1009, ¶[0059]. The adaptive filter 450 “is adapted during periods of non-speech activity.” Ex. 1009, ¶¶[0060]-[0061]. Yang discloses that a **voice activity detector** 440 “detects the presence of speech activity ... and provides a control signal that enables the adaptation of the coefficients of the [adaptive filter 450] when no speech activity is detected.” Ex. 1009, ¶¶[0060]-[0061].

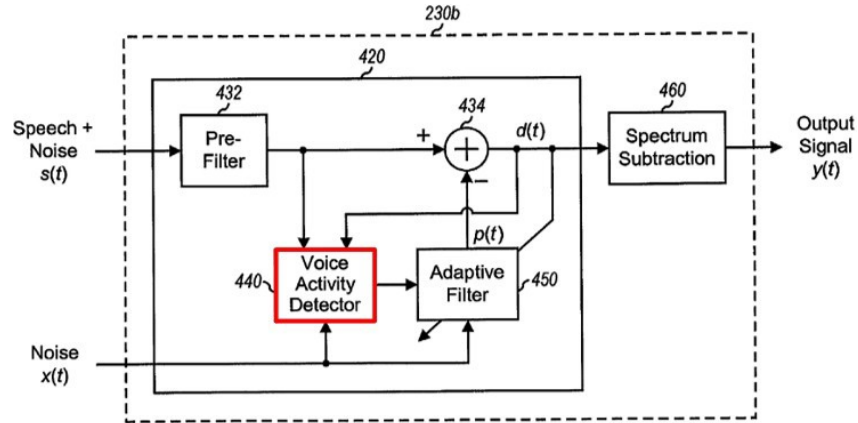


FIG. 4

Ex. 1009, Fig. 4 (annotated).

2. **Claim 10: “The system of claim 7 and further comprising: a voice activity detector (VAD) coupled with the processing component and operative to generate voice activity signals.”**

Ikeda discloses a processing component. Ex. 1006, ¶¶[0016]-[0017], [0030] (“a digital signal processor DSP”); Ex. 1003, ¶¶202-203.

Ikeda states that adaptive filter 7 “adaptively varies its characteristics with respect to changes in the input signal over time,” where “[t]he training sequence timing [for determining the filter response characteristics] is defined as a timing at which the microphone units M1 and M2 do not pick up the target sound.” Ex. 1006, ¶[0028].

It would have been obvious to determine the “timing at which the microphone units M1 and M2 do not pick up the target sound,” such as human speech, using a voice activity detector as taught by Yang. Ex. 1006, ¶¶[0028], [0036]; Ex. 1009,

¶¶[0058]-[0062], Fig. 4; Ex. 1003, ¶¶201, 204-205. Yang teaches: “Voice activity detector 440 detects the presence of speech activity ... and provides a control signal that enables the adaptation of the coefficients of the [adaptive filter 450] when no speech activity is detected.” Ex. 1009, ¶¶[0060]-[0061]; Ex. 1003, ¶¶201, 205. The modification would have been nothing more than combining prior art elements according to known methods to yield predictable results. Ex. 1003, ¶205. A POSITA would have had a reasonable expectation of success to make the modification, because using a voice activity detector to determine the timing when microphones do not pick up human speech would have been within the knowledge and skill of a POSITA. Ex. 1003, ¶205.

Therefore, in Ikeda-McCowan-Kanamori-Yang, a voice activity detector is coupled with the processing component (where adaptive filter 7 is configured) and operative to generate voice activity signals. Ex. 1006, ¶¶[0016]-[0017], [0030]; Ex. 1009, ¶¶[0058]-[0062], Fig. 4; Ex. 1003, ¶206.

3. Claim 15: “The system of claim 14 and further comprising: a voice activity detector (VAD) coupled with the processing component and operative to generate voice activity signals.”

Ikeda-McCowan-Kanamori-Yang renders obvious this limitation for substantially the same reasons discussed in Section V.B.2 regarding claim 10. Ex. 1003, ¶¶207-208.

4. Claim 4: “The system of claim 1 and further comprising: a voice activity detector (VAD) coupled with the processing component and operative to generate voice activity signals.”

Ikeda-McCowan-Kanamori-Yang renders obvious this limitation for substantially the same reasons discussed in Section V.B.2 regarding claim 10. Ex. 1003, ¶¶209-210. A POSITA would have understood that in Ikeda-McCowan-Kanamori-Yang, the voice activity detector is coupled with the processing component that generates the main and reference microphones. § V.A.17.b; Ex. 1009, ¶¶[0058]-[0062], Fig. 4; Ex. 1003, ¶210. For example, the Ikeda-McCowan-Kanamori-Yang voice activity detector receives signals from the main and reference microphones for detecting speech activity. Ex. 1009, ¶¶[0059], [0061], Fig. 4 (the voice activity detector 440 receiving the speech plus noise signal $s(t)$ and the mostly noise signal $x(t)$); Ex. 1003, ¶210.

VI. INSTITUTION IS APPROPRIATE HERE

**A. Co-Pending Litigation Does Not Warrant
Discretionary Denial Under § 314(A) (*Fintiv*)**

Efficiency, fairness, and the merits support institution. *Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11 (PTAB Mar. 20, 2020) (“*Fintiv*”).

1. Factor 1: Potential Stay

On February 28, 2023, Patent Owner (“PO”) sued Petitioner for infringement of the ’080 patent in *Jawbone Innovations, LLC v. Sony Electronics Inc.*, No. 2:23-

cv-01161-CCC-LDW (D.N.J.) (“the Litigation”). There are currently 16 instituted IPRs covering every patent claim asserted against Petitioner in the February 28, 2023 complaint, as follows:

- US10,779,080 (IPR2022-01059, IPR2023-00284);
- US8,019,091 (IPR2022-00649, IPR2023-00253);
- US7,246,058 (IPR2022-00623, IPR2023-00222);
- US8,321,213 (IPR2022-00797, IPR2023-00276, IPR 2023-00279);
- US8,467,543 (IPR2022-01027);
- US8,503,691 (IPR2022-01060, IPR2022-01061);
- US11,122,357 (IPR2022-01124, IPR2023-00251); and
- US8,326,611 (IPR2023-00285, IPR2023-286).

On June 12, 2023, Jawbone amended its complaint to add one additional patent.

The District of New Jersey (“DNJ”) has a “liberal” policy toward staying cases when IPRs have been instituted, to avoid wasting “a court’s time and resources” and because “a USPTO decision could drastically alter the nature of the litigation.” *Kirsch Rsch. & Dev., LLC v. GAF Materials, LLC*, Civil Action No. 20-13683 (JMV), 2021 U.S. Dist. LEXIS 111667, at *5 (D.N.J. June 15, 2021). This is particularly true when, as here, the litigation is at its early stages. *Id.*, at *13.

Given that the Litigation is at its inception and all asserted claims (other than the claims in a single patent added months after the original complaint) are being challenged in instituted IPRs, the likelihood of the grant of a stay weighs against discretionary denial.

2. Factor 2: Proximity of Trial to FWD

The DNJ has not set a trial date. Thus, this factor weighs against discretionary denial. *Global Tel*Link Corp. v. HLFIP Holding, Inc.*, IPR2021- 00444, Paper 14, at 16-19 (PTAB Jul. 22, 2021) (“The lack of a scheduled trial date ... typically weighs against exercising discretion to deny the Petition.”); *GAF Materials LLC v. Kirsch Research and Dev., LLC*, IPR2021-00192, Paper 14, at 12-13 (PTAB May 25, 2021).

3. Factor 3: Investment in Parallel Proceeding

PO filed its complaint on February 28, 2023, and served it on March 16, 2023. (The Litigation, ECF Nos. 1 and 5.) The parties have not conducted any discovery, and an initial scheduling conference will not occur until July 18, 2023. (*Id.*, ECF No. 18.) Thus, the parties have invested very little in the Litigation, and little is expected to be invested in the Litigation over the next several months (if the case is not stayed). Even if a scheduling order subsequently imposes some discovery deadlines over the following 6 months (*e.g.*, fact discovery, claim construction),

much work will still remain, including expert reports, expert discovery, dispositive motions, pretrial motions, and trial. This “remaining investment” significantly outweighs any investment made thus far. *Samsung Elecs. Am. Inc. v. Snik LLC*, IPR2020-01428, Paper 10, at 11 (PTAB Mar. 9, 2021); *Huawei Techs. Co. v. WSOU Invs., LLC*, IPR2021-00226, Paper 10, at 13-14 (PTAB June 10, 2021).

4. Factor 4: Overlapping Issues

If this IPR is instituted and the Litigation is stayed, Petitioner could not pursue in the Litigation any invalidity ground raised or that could have been reasonably raised in this IPR. *Cal. Inst. of Tech. v. Broadcom Ltd.*, 25 F.4th 976, 991 (Fed. Cir. 2022). Thus, duplicative efforts or conflicting decisions are unlikely.

Petitioner hereby stipulates that, if this IPR is instituted but the Litigation is not stayed, Petitioner will not argue in the Litigation invalidity of the '080 patent based on the grounds in this petition. This factor weighs against discretionary denial. *Sand Revolution II, LLC v. Cont'l Intermodal Grp.-Trucking LLC*, IPR2019-01393, Paper 24, at 11-12 (PTAB June 16, 2020) (informative).

5. Factor 5: Parties in Parallel Proceedings

The parties are the same, but trial in the Litigation is highly unlikely to start before this IPR reaches a final written decision, and most likely will occur long after

a final written decision. Thus, this factor is neutral. *See Google*, IPR2022-00630, Paper 13 at 17.

6. Factor 6: Other Circumstances

To Petitioner’s knowledge, other than Google and Amazon’s co-pending IPR petitions on the same grounds that Petitioner is seeking to join in its accompanying motion for joinder, the ’080 patent has never been compared to this combination of references by the PTO, any court, or any jury. The petition’s merits are particularly strong, which favors institution. *Fintiv*, IPR2020-00019, Paper 11, at 18.

Accordingly, the *Fintiv* analysis favors institution.

B. The Existence of Previously Filed IPR Petitions Does Not Warrant Discretionary Denial (*General Plastic*)

The Board has found that the *General Plastic* factors are “not particularly relevant” in this situation, *i.e.*, where a different petitioner files a “me-too” or “copycat” petition with a timely motion for joinder. *Central Security Group*, IPR2019-01609, Paper 11 at e4g8; *Celltrion, Inc. v. Genentech, Inc.*, IPR2018-01019, Paper 11 at 9-11 (PTAB Oct. 30, 2018).

Even if the Board were to consider the *General Plastic* factors, discretionary denial is unwarranted. Sony and the prior petitioners are not the same party and have no significant relationship. They are not co-defendants. They are competitors accused of infringement—in different actions pending in different courts—based on

sales of different products. This weighs against denial. *NetNut Ltd. v. Bright Data Ltd.*, IPR2021-00465, Paper 11 at 9 (PTAB Aug. 12, 2021) (declining to extend *General Plastic and Valve* to different petitioner with no relationship to previous petitioners); *Mercedes-Benz USA, LLC v. Carucel Invs. L.P.*, IPR2019-01404, Paper 12 at 11-12 (PTAB Jan. 22, 2020).

The remaining factors consider timing issues for new challenges in a subsequent petition, *e.g.*, when Petitioner became aware of prior art, whether the PO or Board responded to prior petitions, and the length of time that has passed since an earlier petition. But the present “me too” Petition does not raise new challenges. Furthermore, the sequential IPR filings here were caused by PO’s sequential district court filings and therefore should not be held against Petitioner. Specifically, PO sued Google, Apple and Amazon in 2021 (6:21-cv-00985-ADA, 6:21-cv-00984-ADA (W.D. Tex.), 2:21-cv-00435 (E.D. Tex.)), but waited nearly 18 months before suing Petitioner in February 2023. The Board should not deny Petitioner’s IPR petition merely because PO sued other defendants earlier. *Toshiba Am. Info. Sys., Inc. v. Wallelex Microelecs. Ltd.*, IPR2018-01538, Paper 11 at 21-22 (PTAB Mar. 5, 2019) (PO’s complaint about multiple IPRs “is not persuasive when the volume appears to be a direct result of its own litigation activity”); *id.* at 22 (it was not unfair

for a petitioner to file close to one-year bar date even when co-defendants in related litigation “with distinct interests” and “distinct litigation strategies” filed earlier).

C. Discretionary Denial Under § 325(D) is Also Not Appropriate

The Board has not previously considered “the same or substantially the same prior art or arguments.” 35 U.S.C. §325(d). Here, all grounds rely on Ikeda, which the PTO never considered during original examination of the ’080 patent.

Further, though this Petition presents the same grounds as in Google (IPR2022-01059) and Amazon (IPR 2023-00284), Petitioner is filing this petition to preserve its ability to maintain a petition on the merits in the event that Google and Amazon terminate their involvement. If Petitioner is joined to the Amazon IPR, Petitioner will take an understudy role. Thus, the proposed grounds will not be cumulative of references previously considered by the Board.

VII. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8

A. Real Party-in-Interest

Sony Electronics Inc., Sony Group Corp., Sony Corporation, and Sony Corporation of America are the real parties in interest.

B. Related Matters

To the best knowledge of Petitioner, the ’080 patent is or has been involved in the following district court litigations and petitions for *inter partes* review:

Sony Electronics Inc. v. Jawbone Innovations, LLC
U.S. Patent 10,779,080

Name	Number	Court	Filed
<i>Jawbone Innovations, LLC v. Amazon.com, Inc.</i>	5:22-cv-06727	N.D. Cal.	Nov. 29, 2021 (transferred from E.D. Tex. on Nov. 1, 2022)
<i>Jawbone Innovations, LLC v. Amazon.com, Inc.</i>	2:21-cv-00435	E.D. Tex.	Nov. 29, 2021
<i>Jawbone Innovations, LLC v. Google LLC</i>	6:21-cv-00985	W.D. Tex.	Sept. 23, 2021 (transferred from E.D. Tex. on Feb. 1, 2023)
<i>Jawbone Innovations, LLC v. Google LLC</i>	6:21-cv-00985	W.D. Tex.	Sept. 23, 2021
Petition for <i>Inter Partes</i> Review	IPR2022-01059	PTAB	May 27, 2022
Petition for <i>Inter Partes</i> Review	IPR2022-01320	PTAB	July 27, 2022
Petition for <i>Inter Partes</i> Review	IPR2023-00284	PTAB	Nov. 28, 2022
<i>Jawbone Innovations, LLC v. Meta Platforms, Inc.</i>	6:23-cv-00158	W.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. ZTE Corporation</i>	2:23-cv-00082	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. Panasonic Holdings</i>	2:23-cv-00081	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. Guangdong Oppo Mobile Telecommunications Corp., Ltd.</i>	2:23-cv-00079	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. LG Electronics, Inc.</i>	2:23-cv-00078	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. HTC Corporation</i>	2:23-cv-00077	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. Sony Electronics Inc.</i>	2:21-cv-00186	D.N.J.	Feb. 28, 2023

C. Lead and Back-Up Counsel

Petitioner provides the following designation of counsel.

LEAD COUNSEL	BACKUP COUNSEL
Gregory S. Gewirtz Registration No. 36,522 LERNER DAVID LLP 20 Commerce Drive Cranford, NJ 07016 Telephone: 908.645.5000 Facsimile: 908.654.7866 E-mail:ggewirtz@lernerdavid.com litigation@lernerdavid.com	Russell W. Faegenburg Registration No. 77,876 LERNER DAVID LLP 20 Commerce Drive Cranford, NJ 07016 Telephone: 908.645.5000 Facsimile: 908.654.7866 E-mail:rfaegenburg@lernerdavid.com litigation@lernerdavid.com

D. Service Information

Service information for lead and back-up counsel is provided in the designation of lead and back-up counsel above. Petitioner hereby consents to service by email at the following email address:

litigation@lernerdavid.com

E. Payment of Fees Pursuant to 37 C.F.R. § 42.103

The fee set forth in 37 C.F.R. § 42.15(a) for this petition has been paid. The

undersigned further authorizes payment for any additional fees that may be due in connection with this petition to be charged to Deposit Account 121095.

Respectfully submitted,
LERNER DAVID LLP

Dated: June 15, 2023

/Gregory S. Gewirtz/
Gregory S. Gewirtz (Reg. No. 36,522)
Counsel for Petitioner Sony Electronics Inc.

**CERTIFICATE OF COMPLIANCE
WITH TYPE-VOLUME LIMITATION**

Pursuant to 37 C.F.R. § 42.24(d), the undersigned certifies that foregoing
**PETITION FOR INTER PARTES REVIEW OF UNITED STATES
PATENT NO. 10,779,080 PURSUANT TO 35 U.S.C. §§ 311–319, 37 C.F.R.
§ 42**, exclusive of the parts exempted as provided in 37 C.F.R. § 42.24(a), contains
13,912 words and therefore complies with the type-volume limitations of 37
C.F.R. § 42.24(a).

Respectfully submitted,

LERNER DAVID LLP

Dated: June 15, 2023

/Gregory S. Gewirtz/

Gregory S. Gewirtz (Reg. No. 36,522)

Counsel for Petitioner Sony Electronics Inc.

CERTIFICATE OF SERVICE

The undersigned hereby certifies that a copy of the foregoing **PETITION FOR INTER PARTES REVIEW OF UNITED STATES PATENT NO. 10,779,080 PURSUANT TO 35 U.S.C. §§ 311–319, 37 C.F.R. § 42 and Exhibits 1001-1033** is being served on June 15, 2023, via FedEx Priority Overnight on counsel of record for U.S. Pat. 10,779,080 patent owner Jawbone Innovations, LLC at the address below:

Mark Leonardo
Nutter McClennen & Fish LLP
Seaport West
155 Seaport Blvd.
Boston, MA 02210-2604

Peter Lambrianakos
FABRICANT LLP
411 Theodore Fremd Avenue
Suite 206 South
Rye, NY 10580

LERNER DAVID LLP

Dated: June 15, 2023

/Gregory S. Gewirtz/
Gregory S. Gewirtz (Reg. No. 36,522)
Counsel for Petitioner Sony Electronics Inc.