

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

Google LLC,
Petitioner

v.

Jawbone Innovations, LLC,
Patent Owner

Case IPR2022-01059

U.S. Patent No. 10,779,080

DECLARATION OF JEFFREY S. VIPPERMAN, PH.D.

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	SUMMARY OF OPINIONS.....	1
III.	BACKGROUND AND QUALIFICATIONS	2
	A. Education.....	2
	B. Experience	3
	C. Compensation.....	6
IV.	MATERIALS CONSIDERED	7
V.	LEGAL STANDARDS	9
	A. Claim Construction	10
	B. Level of Ordinary Skill	10
	C. Obviousness.....	11
VI.	THE '080 PATENT.....	13
	A. Overview of Disclosure.....	13
	B. Prosecution History	18
VII.	ANALYSIS OF GROUNDS OF UNPATENTABILITY.....	19
	A. Claim Construction	19
	B. Ground 1: Claims 1-3, 5-9, 11-14, and 16-20 Are Rendered Obvious over Ikeda in View of McCowan and Kanamori.....	20
	1. Overview of Ikeda.....	20
	2. Overview of McCowan.....	26
	3. Overview of Kanamori	31

4.	The Ikeda-McCowan-Kanamori Combination	35
a.	The Ikeda-McCowan-Kanamori main microphone has the same formulation for signal processing as the virtual microphone V ₁ described in the '080 patent, and the Ikeda-McCowan-Kanamori reference microphone has the same formulation for signal processing as the virtual microphone V ₂ described in the '080 patent.....	52
5.	Simulations of Virtual Microphone Responses	56
6.	Claim 7	67
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;”	67
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,”	72
c.	[7c] “wherein the first virtual microphone has a first linear response to speech and first linear response to noise,”	75
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,”	78
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,”	81
f.	[7f] “wherein the second linear response to noise is substantially similar to the first linear response to noise,”	84
g.	[7g] “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”	85
h.	[7h] “the second linear response to speech is substantially dissimilar to the first linear response to speech,”	87
i.	[7i] “wherein the speech is human speech; and”	88
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,”	88
k.	[7k] “wherein the denoised output signals include less acoustic noise than acoustic signals received at the first and second physical microphones.”	94
7.	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.”	95
8.	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.”	96

9.	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.”	97
10.	Claim 12: “The system of claim 11 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.”	98
11.	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.”	99
12.	Claim 14.....	101
a.	[14a] “A system, comprising: a first virtual microphone comprising a first combination of a first microphone signal and a second microphone signal,”	101
b.	[14b] “the first virtual microphone having a first linear response to speech and a first linear response to noise,”	101
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,”	101
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;”	102

- e. [14e] “a second virtual microphone comprising a second combination of the first microphone signal and the second microphone signal,” 102
- f. [14f] “the second virtual microphone having a second linear response to speech and a second linear response to noise,” 102
- g. [14g] “the second linear response to noise being substantially similar to the first linear response to noise,” 102
- 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” 103
- i. [14i] “the second linear response to speech being substantially dissimilar to the first linear response to speech,” 103
- j. [14j] “wherein the second combination is different from the first combination,” 103
- k. [14k] “wherein the first virtual microphone and the second virtual microphone are distinct virtual directional microphones; and” 103
- 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,” 105
- m. [14m] “wherein the output signal is a denoised acoustic signal.” 107

13.	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.”	107
14.	Claim 17: “The system of claim 16 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.”	107
15.	Claim 18: “The system of claim 14, wherein the acoustic signals from the first virtual microphone, the second virtual microphone or both are delayed.”	108
16.	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.”	110
17.	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.”	113
18.	Claim 1	116
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;”	116
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,”	116
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,”	117
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”	117
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,”	118
f.	[1f] “wherein the denoised output signals include less acoustic noise than acoustic signals received at the microphone array.”	118
19.	Claim 2: “The system of claim 1, wherein the acoustic noise comprises noise content and the acoustic signals comprise speech content.”	119
20.	Claim 3: “The system of claim 2, wherein the speech content comprises human speech.”	119
21.	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.”	120
22.	Claim 6: “The system of claim 5 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.”	121
C.	Ground 2: Claims 4, 10, and 15 are rendered obvious over Ikeda in view of McCowan, Kanamori, and Yang	121
1.	Overview of Yang	121
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.”	122
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.”	124
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.”	124
VIII.	CONCLUSION.....	125

I, Jeffrey S. Vipperman, declare as follows:

I. INTRODUCTION

1. I have been retained as an independent expert by Google LLC (“Petitioner” or “Google”) in connection with an *inter partes* review of U.S. Patent No. 10,779,080 (“the ’080 patent”) (Ex. 1001). I have prepared this declaration in connection with Google’s petition.

2. Specifically, this document contains my opinions about the technology claimed in claims 1-20 of the ’080 patent (“the challenged claims”) and Google’s grounds of unpatentability for these claims.

II. SUMMARY OF OPINIONS

3. This declaration considers the challenged claims of the ’080 patent. Below I set forth the opinions I have formed, the conclusions I have reached, and the bases for these opinions and conclusions.

4. In forming my opinions, I have assumed that the priority date of the ’080 patent is June 13, 2007, which is the filing date of U.S. Provisional Patent Application No. 60/934,551 (“the ’551 application”), as listed on the cover page of the ’080 patent. Ex. 1001, Cover. I understand the ’080 patent claims priority to the ’551 application. Ex. 1001, Cover.

5. Based on my experience, knowledge of the art, analysis of the asserted grounds and references, and understanding a person of ordinary skill in the

art (“POSITA”) would have had of the claims, it is my opinion that the challenged claims of the ’080 patent would have been obvious to a person of ordinary skill in the art as of 2007, based on the asserted grounds.

III. BACKGROUND AND QUALIFICATIONS

6. I believe that I am well qualified to serve as a technical expert in this matter based upon my educational and work experience, which I summarize below. I understand that my curriculum vitae, which includes a more detailed summary of my background, experience, patents, and publications, is attached as Ex. 1004.

A. Education

7. I received my Ph.D. in Mechanical Engineering from Duke University in 1997. Previously, I obtained Master of Science and Bachelor of Science degrees in Mechanical Engineering from the Virginia Polytechnic Institute and State University (“Virginia Tech”) in 1992 and 1990, respectively. My dissertation at Duke was titled “Adaptive Piezoelectric Sensoriactuators for Multivariable Structural Acoustic Control.” My dissertation addressed the development of a hybrid analog/digital circuit and adaptation method to permit piezoelectric transducers to be used simultaneously as a sensor and an actuator. Doing so provides an array of truly “co-located” sensor/actuator pairs with minimum phase, such that stability of the multichannel feedback system is greatly enhanced. These were demonstrated for active structural acoustic control.

B. Experience

8. I am a Professor of Mechanical Engineering, Bioengineering, and Communication Sciences and Disorders. I also currently serve as Vice Chair of the Mechanical Engineering and Materials Science Department at the University of Pittsburgh.

9. I first began research in acoustics and sound systems in 1989 as an undergraduate student. My masters research concerned adaptive feedforward control of broadband structural vibration, and my Ph.D. research concerned the development of arrays of self-sensing piezoelectric transducers that could be used for active structural-acoustic control. I have also developed a number of algorithms for active control of noise and vibration.

10. My acoustics research has included a mix of theory, analytical and numerical modeling, and measurement of acoustic and vibration systems. Aside from the previously mentioned array research, my acoustics research has included transducer and controls development, transducer modeling/fabrication/testing, analog/digital signal processing, embedded systems, active and passive noise and vibration control, development of various types of metamaterials (e.g., phononic crystals, resonant lattices, layered media, and pentamode materials) for acoustical filtering and cloaking, development of noise classifiers to discern types of military noise or for incorporation into surgical devices as surgical aids, development of

thermoacoustic engines, refrigerators, and sensors (e.g., a wireless, “in-core” thermoacoustic sensor that can measure temperature and neutron flux inside a nuclear reactor). Additional topics of my research include developing structural acoustic models (i.e., concerned with sound radiation from vibrating structures) of sound transmission through finite cylinders, various methods of passive and active control of noise, vibration, and structural-acoustic radiation (i.e., controlling sound radiation of a vibrating structure by introducing additional vibrations to make it an inefficient radiator), hearing loss prevention, and modeling of ear response and damage to the inner ear for impulsive and ultrasound sources. During the early stages of the microelectromechanical systems (MEMS) revolution, I worked on producing some of the earliest silicon-on-insulator (MEMS) microphones through electronic fabrication methods.

11. As a professor, I have developed and taught three graduate courses directly related to acoustics and signal processing, including “Measurement and Analysis of Vibroacoustic Systems,” “Fundamentals of Acoustics and Vibration,” and “Measurement and Analysis of Random Data from Dynamical Systems.” The latter two courses cover acoustical arrays. I have also taught three mechanical measurements courses, a dynamic systems and introductory undergraduate and graduate mechanical vibrations course, and an advanced (Ph.D. candidate level) vibrations course, as well as related courses such as controls, undergraduate and

graduate dynamics, kinematics, mechanical measurements, and electrical circuits.

Further, I have developed and given a short course at the American Controls Conference on “Active Control of Sound, Vibration, and Structural Acoustics,” as well as two other short courses for local industry on “Acoustical Theory and Measurements” and “Noise and Vibration Measurements.”

12. I also have a consulting business (Blue Ridge Consulting) and am Vice President of Atlas Medtech, LLC, a University of Pittsburgh licensed startup company.

13. I have worked on Department of Defense (“DoD”) projects as a Principal Investigator and Co-Principal Investigator on projects that involve acoustic arrays. In one project, a microphone array and cross-correlation methods (time difference of arrival or TDOA methods) were used to determine the bearing angle for acoustic plane waves associated with various forms of military and natural noise. Multiple arrays were used to triangulate the location of the noise source. In conjunction, we developed machine learning algorithms to classify the noise source, which provided additional help for noise management programs around U.S. military bases. A corporate partner commercialized the array research into a product. In another project, I helped co-develop a method for localizing sound using small arrays of unidirectional (e.g., “shot-gun”) microphones. The methods worked in both the time and frequency domains. Another military project

funded by DoD involved the development of 2-D and 3-D source parametric arrays for steering heterodyned ultrasound for communications systems.

14. Some of my professional activities include chairing an American National Standards Institute (ANSI) Committee to revise the ANSI S1.1 Acoustical Terminology Standard. I am also a Fellow in the American Society of Mechanical Engineers (ASME) and a former Chair of the Noise Control and Acoustics Division of ASME. I also chaired the Per Bruel Gold Medal in Acoustics Award selection committee for ASME. I have organized nine conference sessions on acoustics and was a Track Organizer (over multiple conference sessions) for nine ASME conferences, as well as Technical Program Chair over all acoustics-related conference sessions at the ASME International Mechanical Engineering Congress and Exposition (IMECE) in 2009. I also participated on a National Research Council (National Academies) panel to evaluate the hearing loss prevention component of the mining program for the National Institute for Occupational Safety and Health (NIOSH) research programs.

15. I have published numerous technical papers, book chapters, reports, and the like related to acoustic sensors and acoustic signal processing.

C. Compensation

16. I am being compensated for services provided in this matter at my usual and customary rate of \$400 per hour plus travel expenses. My compensation

is not conditioned on the conclusions I reach as a result of my analysis or on the outcome of this matter, and in no way affects the substance of my statements in this declaration.

17. I am not aware of any financial interest that I have in the Patent Owner, or any of its subsidiaries or affiliates. Likewise, I am not aware of any financial interest that I have in Petitioner, or any of its subsidiaries or affiliates. I do not have any financial interest in the '080 patent or any proceeding involving the '080 patent.

IV. MATERIALS CONSIDERED

18. In forming my opinions, I have analyzed the following, including the '080 patent, its file history, the prior art listed in this declaration and in the petition grounds, and the materials listed in this declaration.

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. 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”)
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. 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. 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)
Ex. 1028	Excerpt of Alan V. Oppenheim et al., <i>Discrete-Time Signal Processing</i> , Prentice-Hall, Inc. (2nd ed. 1999)

19. My opinions are based on my experience, knowledge of the relevant art, the documents identified above, and the documents discussed in this declaration.

V. LEGAL STANDARDS

20. I am not a lawyer. My understanding of the legal standards to apply in reaching the conclusions in this declaration is based on discussions with counsel for Petitioner, my experience applying similar standards in other patent-related matters, and my reading of the documents submitted in this proceeding. In preparing this declaration, I sought to faithfully apply these legal standards to the challenged claims.

A. Claim Construction

21. I have been instructed that the terms appearing in the '080 patent should be interpreted in view of the claim language itself, the specification, the prosecution history of the patent, and any relevant extrinsic evidence. The words of a claim are generally given their ordinary and customary meaning, which is the meaning that the term would have to a person of ordinary skill in the art at the time of the invention, which I am assuming here is June 13, 2007. While claim limitations cannot be read in from the specification, the specification is the single best guide to the meaning of a disputed term. I have followed these principles in reviewing the claims of the '080 patent and forming the opinions set forth in this declaration.

B. Level of Ordinary Skill

22. I understand a person of ordinary skill in the art is determined by looking at (i) the type of problems encountered in the art; (ii) prior art solutions to those problems; (iii) rapidity with which innovations are made; (iv) sophistication of the technology; and (v) educational level of active workers in the field.

23. In my opinion, 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. Work experience can substitute for formal education and additional formal education can substitute for work experience. I was at least a POSITA as of June 13, 2007.

C. Obviousness

24. I have been told that under 35 U.S.C. § 103, a patent claim may be obvious if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.

25. I have been told that a proper obviousness analysis requires the following: (a) determining the scope and content of the prior art; (b) ascertaining the differences between the prior art and the claims at issue; (c) resolving the level of ordinary skill in the pertinent art; and (d) considering evidence of secondary indicia of non-obviousness (if available).

26. I have been told that the relevant time for considering whether a claim would have been obvious to a person of ordinary skill in the art is the time of invention. For purposes of my analysis, I assumed that the date of invention for the challenged claims is June 13, 2007.

27. I have been told that a reference may be modified or combined with other references or with the person of ordinary skill's own knowledge, if the

person would have found the modification or combination obvious. I have also been told that a person of ordinary skill in the art is presumed to know all the relevant prior art, and the obviousness analysis may take into account the inferences and creative steps that a person of ordinary skill in the art would employ.

28. I have been told that whether a prior art reference renders a patent claim obvious is determined from the perspective of a person of ordinary skill in the art. I have also been told that, while there is no requirement that the prior art contain an express suggestion to combine known elements to achieve the claimed invention, and while a suggestion to combine known elements to achieve the claimed invention may come from the prior art as a whole or individually and may consider the inferences and creative steps a person of ordinary skill in the art would employ, as filtered through the knowledge of one skilled in the art, obviousness grounds cannot be sustained by mere conclusory statements and must include some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness.

29. I have been told that there is no rigid rule that a reference or combination of references must contain a “teaching, suggestion, or motivation” to combine references. But I also have been told that the “teaching, suggestion, or motivation” test can be used in establishing a rationale for combining elements of

the prior art. I have also been told to be aware of distortions caused by hindsight bias, and that reading into the prior art the teachings of the invention at issue is improper.

30. I am aware that a claim may be obvious where the claim represents nothing more than a combination of prior art elements according to known methods that yields predictable results. I am further aware that a claim may be obvious where it merely involves the simple substitution of one known element for another to achieve predictable results. I am aware that a claim may be obvious where the claim involves nothing more than applying a known technique to a known device ready for improvement to yield predictable results. I am additionally aware that it may be obvious to try a particular combination of claim features if selecting them requires merely choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success.

VI. THE '080 PATENT

A. Overview of Disclosure

31. 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.

32. In one embodiment, two virtual directional microphones V_1 and V_2 are formed based on two physical omnidirectional microphones O_1 and O_2 according to the following equations expressed in the z (digital frequency) domain:

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

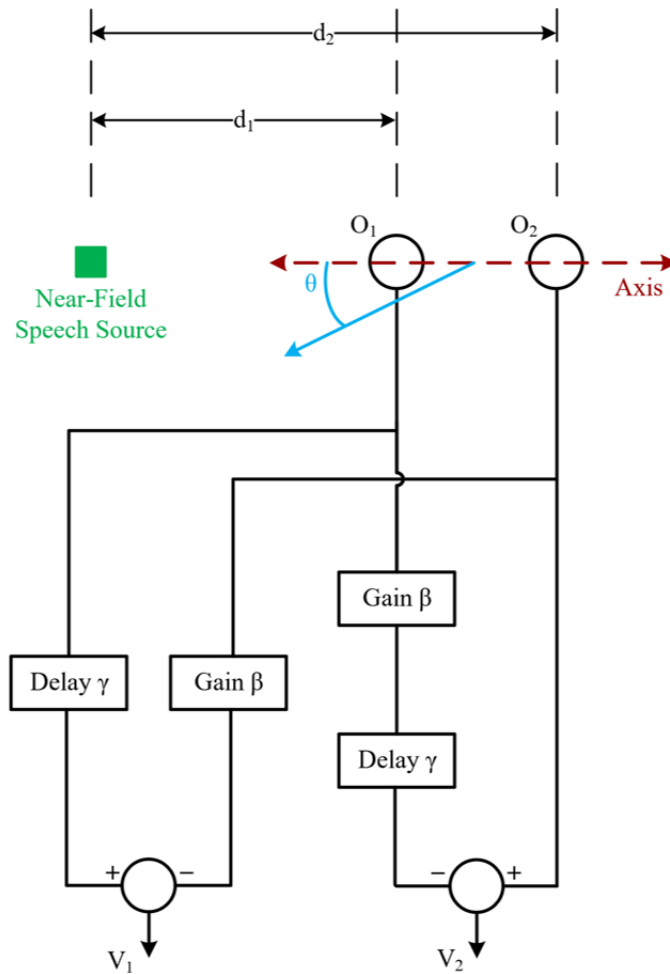
$$V_2(z) = O_2(z) - z^{-\gamma} \beta O_1(z)$$

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

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

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

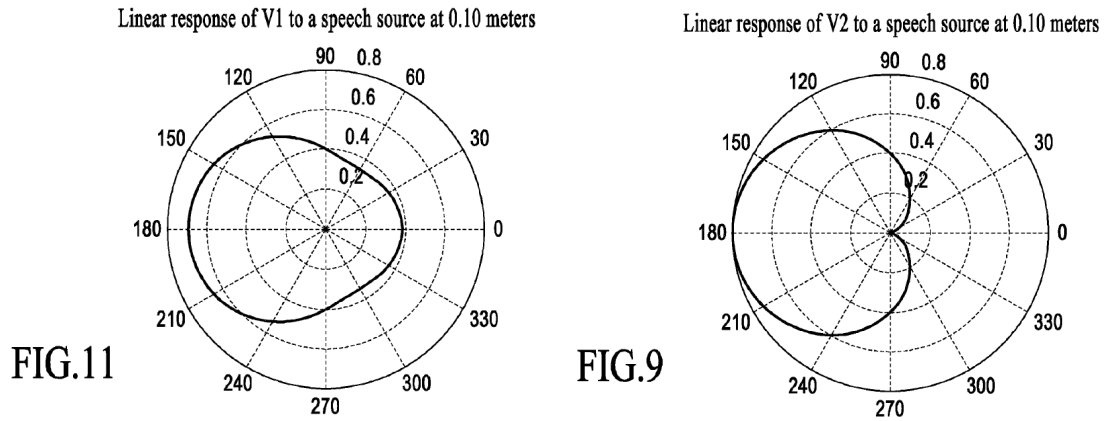
where d_1 and d_2 “are the distance from O_1 and O_2 to the speech source” respectively, c is the speed of sound, and f_s is a sampling frequency. Ex. 1001, 9:54-67. The figure below visually represents the equations for the virtual microphones V_1 and V_2 above. Ex. 1001, 2:28-34, 7:17-38, Fig. 2. This 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 zero degrees. Ex. 1001, 2:28-34, 7:17-38, 9:46-67, 10:30-32, 11:62-65, 16:4-6, Fig. 2.



33. Based on this formulation of the 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; *see also* Ex. 1001, 9:39-12:7, 15:29-16:55.

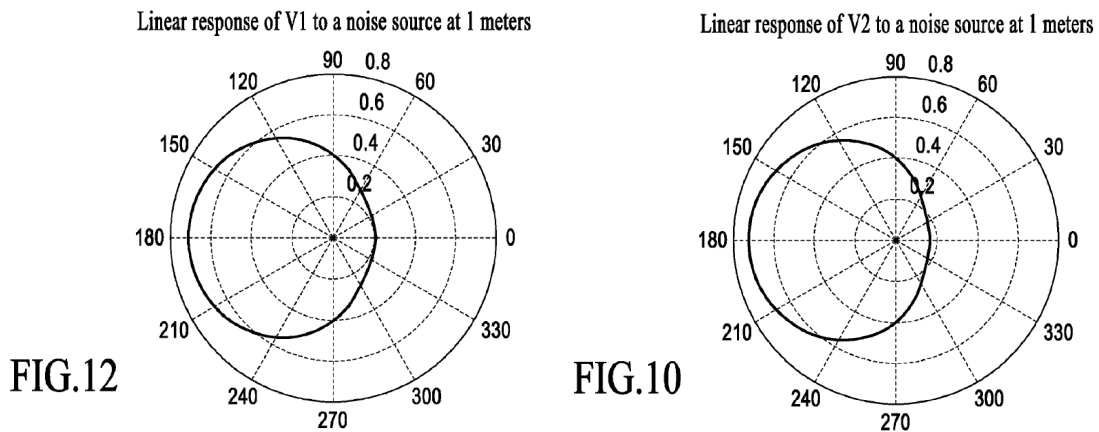
34. These “similar” and “different” responses can be seen when comparing Figures 11 and 9 (different responses) and Figures 12 and 10 (similar responses). Ex. 1001, 2:52-67, 10:18-35, 11:20-39. Figures 11 and 9 (reproduced below) show these “different” linear responses of the virtual microphones V_1 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.

35. Figures 12 and 10 (reproduced below) show these “similar” linear responses of the virtual microphones V_1 and V_2 , 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.

36. The linear response of the virtual microphone V_1 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.

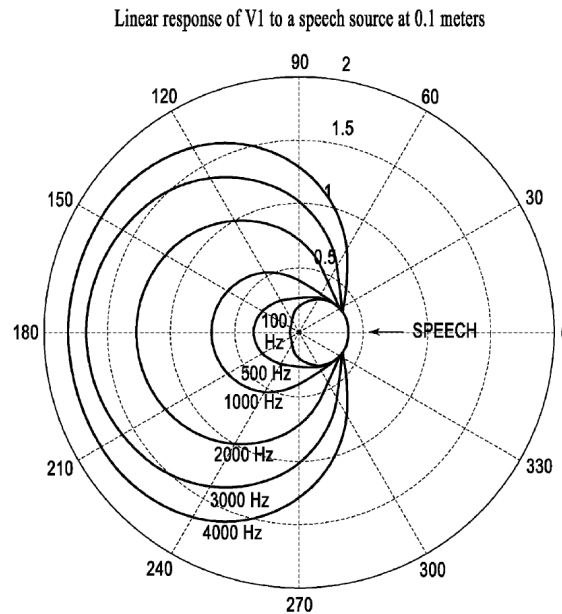


FIG.13

Ex. 1001, Fig. 13.

37. The signals from the virtual microphones V_1 and V_2 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 (“any adaptive filter or noise removal algorithm can be used ... in one or more various alternative embodiments or configurations”), 8:50-9:3, 9:22-38; *see also* Ex. 1001, 4:52-7:16. In an embodiment, the adaptive noise removal system combines the signals from the virtual microphones V_1 and V_2 “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

38. 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 as a continuation of U.S. Patent Application No. 12/139,355, filed on June 13, 2008. Ex. 1001, Cover. The '080 patent also claims priority to U.S. Provisional Patent Application Nos.: 60/934,551, filed on June 13, 2007; 60/953,444, filed on August 1, 2007; 60/954,712, filed on August 8, 2007; and 61/045,377, filed on April 16, 2008. Ex. 1001, Cover.

39. The '160 application was filed with one claim. Ex. 1002, 57. The Examiner rejected the initial claim as anticipated by U.S. Patent No. 8,068,619 to Zhang et al., and based on non-statutory double patenting. Ex. 1002, 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 as being unpatentable over U.S. Patent Nos. 8,494,177 and

8,503,691. 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.

VII. ANALYSIS OF GROUNDS OF UNPATENTABILITY

40. Based on my professional and academic experience, and my review of the prior art, it is my opinion that claims 1-20 of the '080 patent are unpatentable in view of the prior art, including Ikeda (Ex. 1005, translation at Ex. 1006), McCowan (Ex. 1007), Kanamori (Ex. 1008), and Yang (Ex. 1009) as discussed below.

A. Claim Construction

41. I understand that 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. For purposes of my analysis, I have considered each claim term and applied its plain and ordinary meaning consistent with the patent's disclosure and prosecution history.

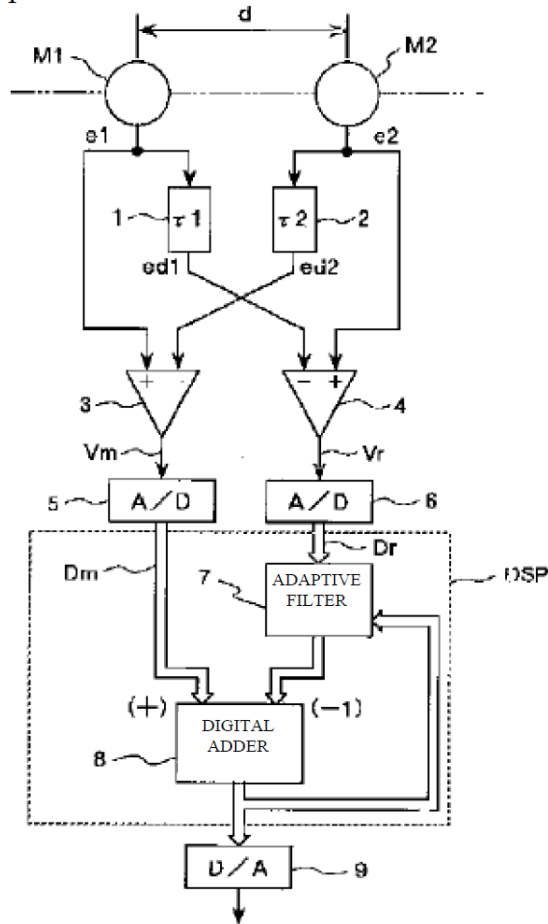
B. Ground 1: Claims 1-3, 5-9, 11-14, and 16-20 Are Rendered Obvious over Ikeda in View of McCowan and Kanamori

42. In my opinion, claims 1-3, 5-9, 11-14, and 16-20 are rendered obvious over Ikeda in view of McCowan and Kanamori.

1. Overview of Ikeda

43. 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,” as shown in Figure 1 (reproduced below). Ex. 1006, ¶ [0018], Fig. 1.

FIG. 1



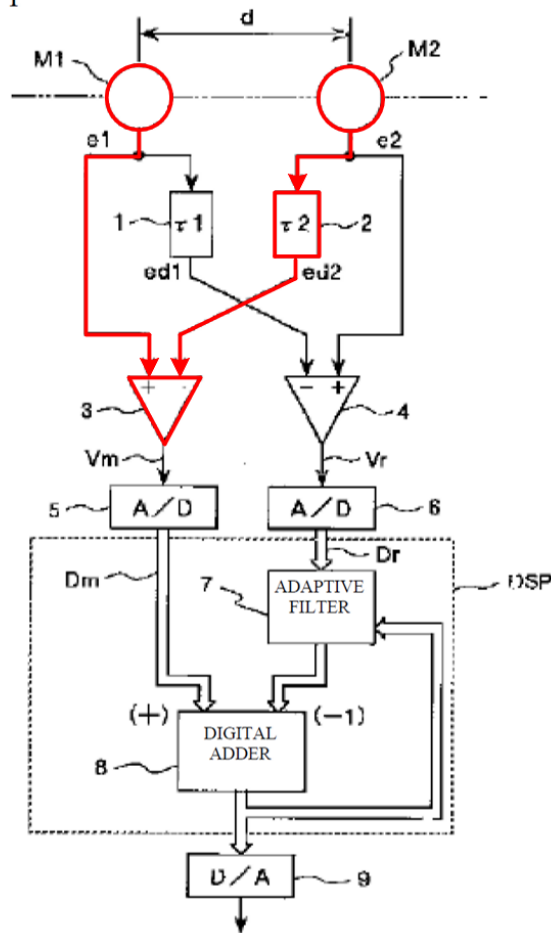
Ex. 1006, Fig. 1.

44. Ikeda discloses forming a **main microphone**, as shown in Figure 1 (annotated below). Ex. 1006, ¶¶ [0002]-[0004], [0018]-[0022], Figs. 2-3. In particular, a delay circuit 2 delays the output of microphone unit M2 and an inversion adding circuit 3 subtracts the output of delay circuit 2 from the output of microphone unit M1 and outputs a main signal Vm. Ex. 1006, ¶¶ [0018], [0020]-[0022], Fig. 1. A signal system constructed using the microphones M1 and M2, the delay circuit 2, and the inversion adding circuit 3 for outputting the main signal

V_m functions as the **main microphone**. Ex. 1006, ¶¶ [0012], [0020]-[0022], Fig. 2.

The **main microphone** has directional characteristics. Ex. 1006, ¶¶ [0002]-[0004], [0021]-[0022], Fig. 3.

FIG. 1

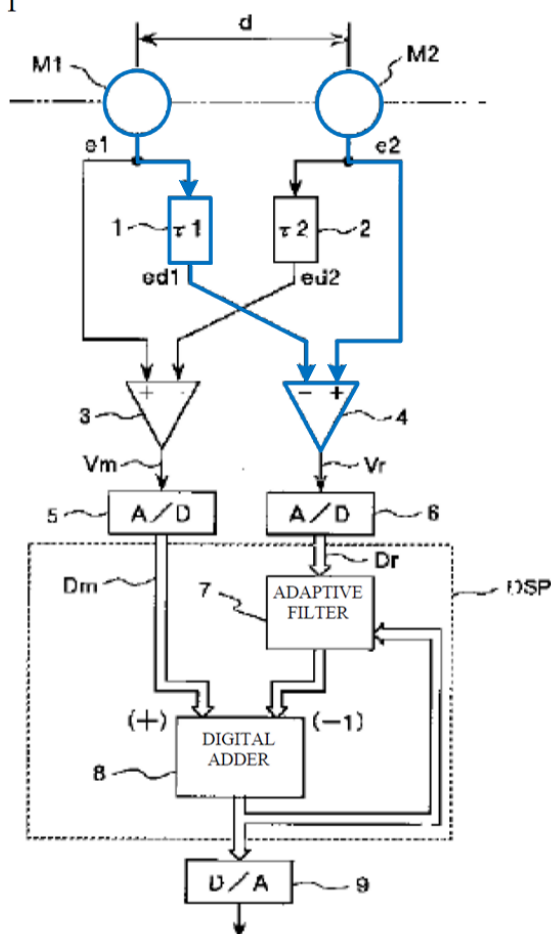


Ex. 1006, Fig. 1 (annotated).

45. Ikeda discloses forming a **reference microphone**, as shown in Figure 1 (annotated below). Ex. 1006, ¶¶ [0018]-[0020], [0023]-[0024], Figs. 4-5. In particular, a delay circuit 1 delays the output of microphone unit M1 and an inversion adding circuit 4 subtracts the output of delay circuit 1 from the output of

microphone unit M2 and outputs a reference signal V_r . Ex. 1006, ¶¶ [0018], [0020], [0023]-[0024], Fig. 1. A signal system constructed using the microphones M1 and M2, the delay circuit 1, and the inversion adding circuit 4 for outputting the reference signal V_r functions as the [reference microphone](#). Ex. 1006, ¶¶ [0012], [0020], [0023]-[0024], Fig. 4. The [reference microphone](#) has directional characteristics. Ex. 1006, ¶¶ [0023]-[0024], Fig. 5.

FIG. 1

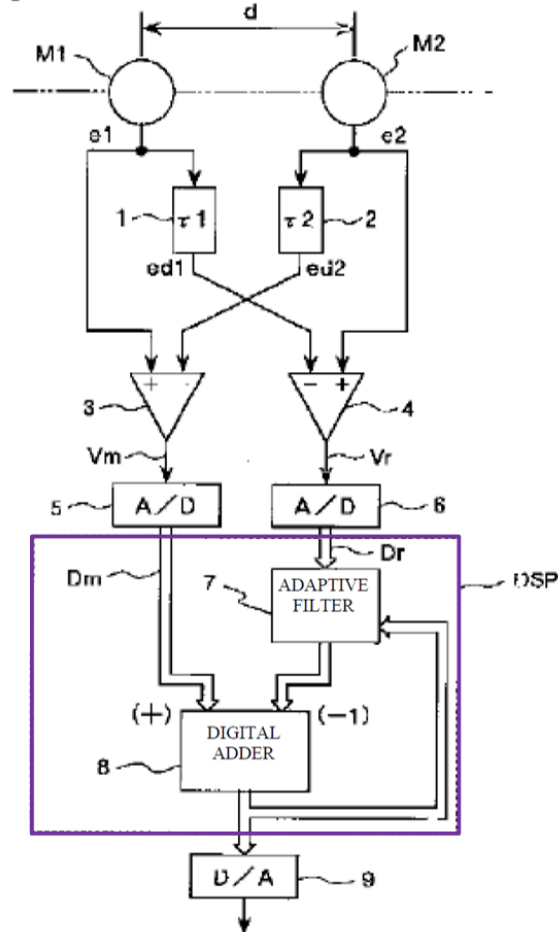


Ex. 1006, Fig. 1 (annotated).

46. 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]; *see also* Ex. 1006, ¶¶ [0012], [0026], Fig. 10, Abstract.

47. The main signal V_m (as converted to a digital signal D_m) and the reference signal V_r (as converted to a digital signal D_r) 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]. Signal D_r is input to adaptive filter 7, and the difference between the output of adaptive filter 7 and signal D_m is calculated by digital adder 8. Ex. 1006, ¶ [0027] (“Main [sic Reference] signal data D_r outputted from the A/D converter 6 is supplied to an adaptive filter 7 ..., and the difference between the filter output of the adaptive filter 7 and the main signal data D_m outputted from the A/D converter 5 is calculated by a digital adder 8.”).

FIG. 1



Ex. 1006, Fig. 1 (annotated).

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

49. 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 $\mathbf{x}(f)$ 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.

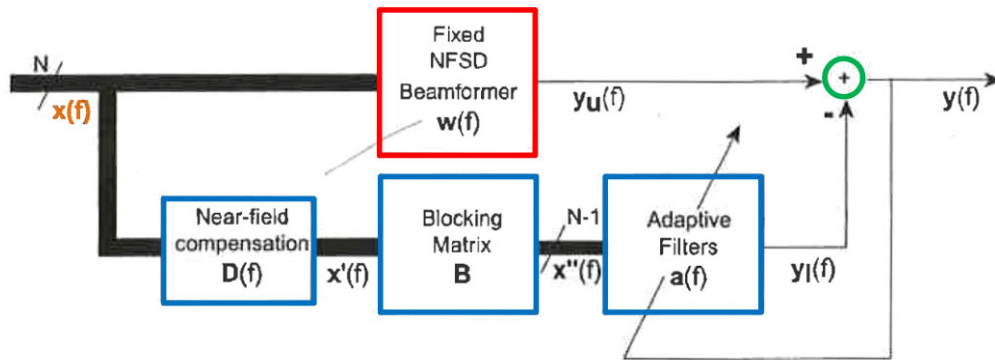


FIG. 3. Near-field adaptive beamformer.

Ex. 1007, Fig. 3 (annotated).

50. The purpose of the **blocking matrix** is “to block the desired signal from the adaptive noise estimate.” Ex. 1007, 93. The blocking matrix does this by causing a signal from a microphone to be subtracted by a signal from another microphone. Ex. 1007, 95 (equations 18, 19). “To ensure complete blocking, the desired signal must both be time aligned and have equal amplitudes across all channels.” Ex. 1007, 93.

51. 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),” which is given by the following:

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

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. In addition, 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 given by the following:

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

Ex. 1007, 90.

52. “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. With the near-field compensation, the output $\mathbf{x}''(f)$ of the blocking matrix equals:

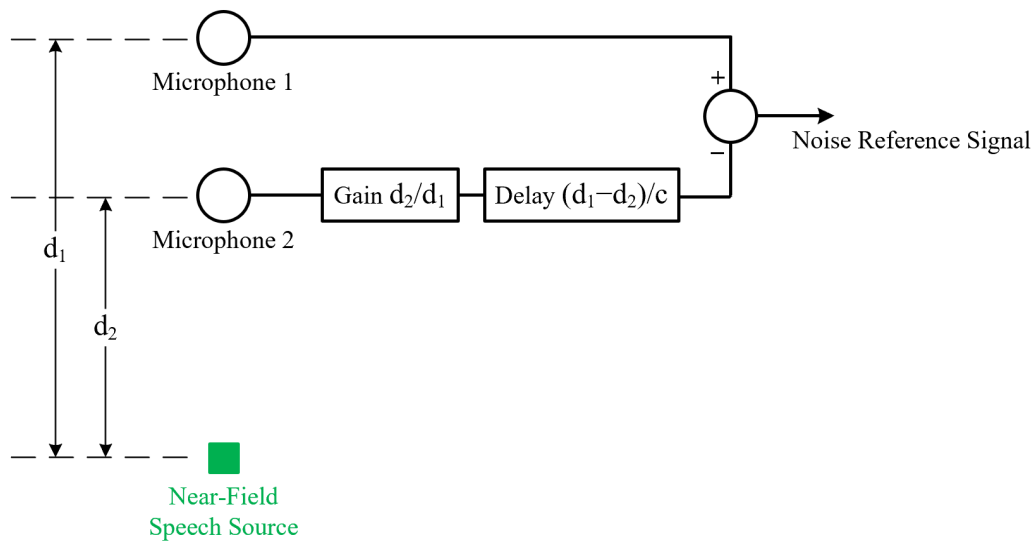
$$\begin{bmatrix} \frac{1}{\alpha_1 e^{-j2\pi f \tau_1}} x_1(f) - \frac{1}{\alpha_2 e^{-j2\pi f \tau_2}} x_2(f) \\ \frac{1}{\alpha_2 e^{-j2\pi f \tau_2}} x_2(f) - \frac{1}{\alpha_3 e^{-j2\pi f \tau_3}} x_3(f) \\ \vdots \\ \frac{1}{\alpha_{N-1} e^{-j2\pi f \tau_{N-1}}} x_{N-1}(f) - \frac{1}{\alpha_N e^{-j2\pi f \tau_N}} x_N(f) \end{bmatrix}$$

where $x_i(f)$ is the signal of the i th microphone. Ex. 1007, 90 (“the signal on the i th microphone is ... $x_i(f)$,” equation 6), 93 (equations 15, 16, 17), 95 (equations 18, 19).

53. “The near-field compensation effectively ensures that a true null exists in the beam-pattern of each blocking matrix row in the direction and distance corresponding to the desired source.” Ex. 1007, 94. For example, the first row of the output $\mathbf{x}(f)$ of the blocking matrix, with $x_1(f)$ being “the desired source at the reference microphone” (Ex. 1007, 90), would be equal to the following:

$$x_1(f) - \frac{d_2}{d_1} e^{-j2\pi f \frac{d_1 - d_2}{c}} x_2(f)$$

Ex. 1007, 90. That is, the signal of microphone 2 is multiplied by d_2/d_1 and delayed by $(d_1 - d_2)/c$, before it is subtracted from the signal of microphone 1, to form a signal system having “a true null ... in the beam-pattern ... in the direction and distance corresponding to the desired source.” Ex. 1007, 94. 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 the microphones 1 and 2. Ex. 1007, 87-91, 93-96, Fig. 3.



54. This configuration with the gain d_2/d_1 and the 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 output signal of the fixed NFSD beamformer in the upper signal path of the McCowan system. Ex. 1007, 87-91, 93-96, Fig. 3.

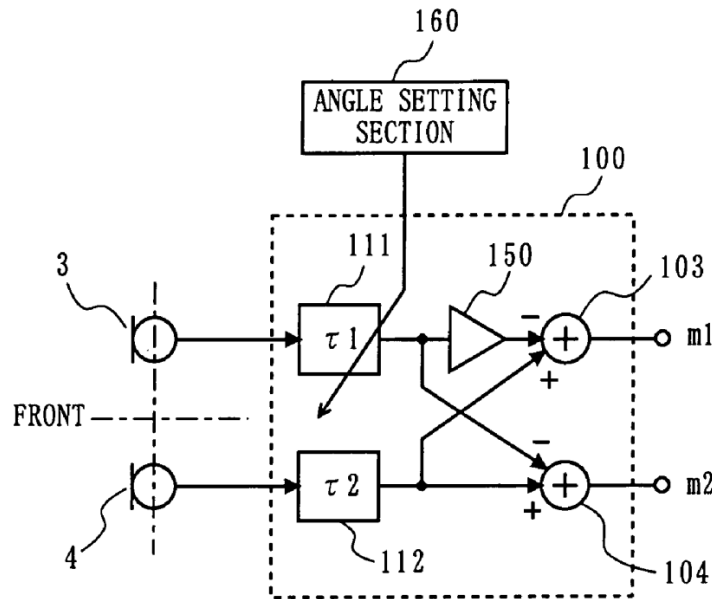
55. The gain d_2/d_1 in the near-field compensation taught by McCowan functions in the same way as the gain $\beta = d_1/d_2$ of the virtual microphone V_2 described in the '080 patent. Ex. 1006, 90, 93-95; Ex. 1001, 7:17-38, 9:50-67, Fig. 2. The swap of the labeling in how the gain is represented using d_1 and d_2 in McCowan and in the '080 patent is due to the different labeling for microphones (microphone 2 and microphone 1 in McCowan, and O_1 and O_2 in the '080 patent). In essence, both the gain in McCowan and the gain in the '080 patent compensate for amplitude decay of sound propagation across microphones because, in both

McCowan and the '080 patent, a signal from a particular microphone (microphone 2 in McCowan, or O_1 in the '080 patent) is multiplied by the gain equal to a ratio of a distance between the near-field speech source and the particular microphone to a distance between the near-field speech source and another microphone (microphone 1 in McCowan, or O_2 in the '080 patent).

3. Overview of Kanamori

56. 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 m_1 and a noise reference signal m_2 are generated using signals from microphone units 3 and 4. Ex. 1008, ¶¶ [0151], [0166], [0172]-[0174].

FIG. 17A

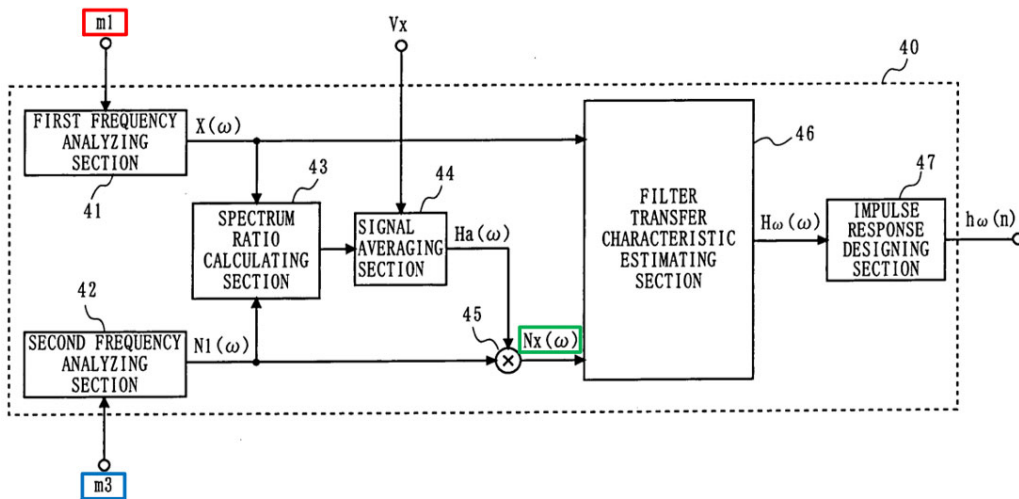


Ex. 1008, Fig. 17A.

57. 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. As discussed in greater detail below, 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]. In the noise suppressing process, Kanamori discloses “calculat[ing] a power spectrum $X(\omega)$ of the signal **m1**, which is the main signal,” and

“calculat[ing] a power spectrum $N1(\omega)$ of the signal **m3**, which is the noise reference signal,” as shown in Figure 4 (annotated below). Ex. 1008, ¶ [0102] (color code added); *see also* Ex. 1008, ¶¶ [0168], [0083]-[0086] (disclosing that the noise reference signal **m3** is generated based on the noise reference signal **m2**), [0093]-[0099] (same).

FIG. 4



Ex. 1008, Fig. 4 (annotated).

58. The spectrum ratio calculating section 43 calculates “a spectrum ratio $H(\omega)=X(\omega)/N1(\omega)$.” Ex. 1008, ¶ [0103]. Based on the spectrum ratio $H(\omega)$, the signal averaging section 44 calculates “a time average $H_a(\omega)$... when the surrounding noise is dominant compared with the target sound.” Ex. 1008, ¶ [0103]. The signal multiplying section 45 “multiplies the power spectrum $N1(\omega)$... by the time average $H_a(\omega)$... [and] outputs the multiplication result as **$Nx(\omega)$** .”

Ex. 1008, ¶ [0103] (color code added). The spectrum $N_x(\omega)$ “represents the noise components included in the spectrum $X(\omega)$ of the main signal ... [and is] referred to as an estimated noise spectrum $N_x(\omega)$.” Ex. 1008, ¶ [0103] (color code added). The spectrum $N_x(\omega)$ is used to suppress “noise components included in the main signal.” Ex. 1008, ¶ [0168]; *see also* Ex. 1008, ¶¶ [0080]-[0083], [0100]-[0115].

59. 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 irrespective of the direction from which noise is coming:”

In order to suppress, based on the noise reference signal, noise components included in the main signal under the circumstances where noise is coming from a plurality of directions, the output of the spectrum ratio calculating section 43 illustrated in FIG. 4 has to be constant irrespective of the direction from which noise is coming.

Ex. 1008, ¶ [0168]. Otherwise, the estimated noise spectrum in the main signal will be correct only “in a specific direction.” Ex. 1008, ¶ [0168] (“if the output of the spectrum ratio calculating section 43 is changed depending on the direction from which noise is coming, only the estimated noise spectrum $N_x(\omega)$ in a specific direction is correctly calculated”). For this reason, Kanamori teaches that it is

preferable that the directivity patterns of the main signal m_1 and the noise reference signal m_2 be as similar as possible in directions from which noise is coming. Ex. 1008, ¶¶ [0167]-[0170].

4. The Ikeda-McCowan-Kanamori Combination

60. 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. And it would have been obvious to a POSITA to combine Ikeda-McCowan-Kanamori as explained below.

61. Ikeda assumes that sources of a target sound (speech) and noise are in the far-field of the microphones M_1 and M_2 . Ex. 1006, ¶¶ [0002]-[0004], [0020]-[0024], [0036], Figs. 2-7; *see also* Ex. 1007, 89-90 (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. 1014 (Elko), ¶¶ [0021]-[0023] (the final expression in Equation (1), which is substantively equivalent to the mathematical expressions representing the signals V_m and V_r in Ikeda, assumes

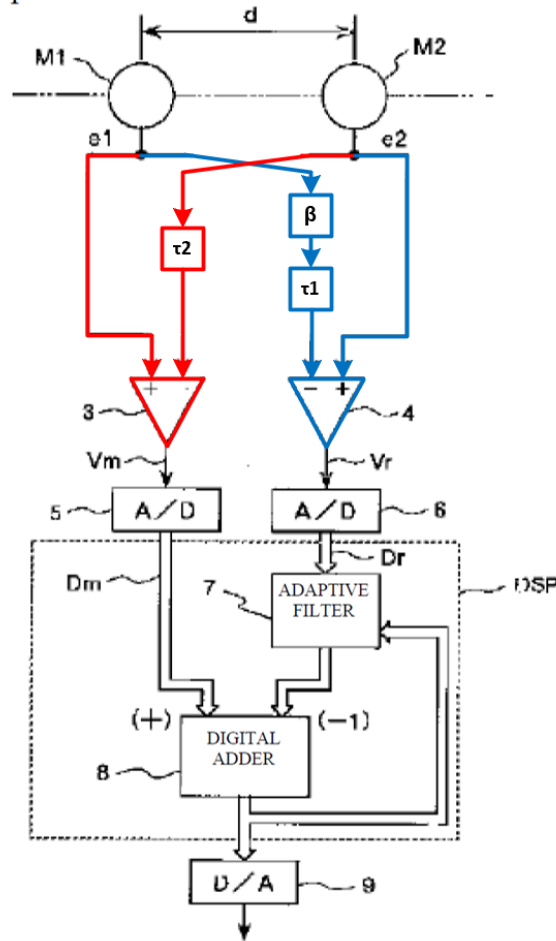
far-field conditions, where sound arriving at a microphone can be treated as a plane wave), Fig. 1, Claim 5; Ex. 1027, 1-2 (Equations (1) and (2), which are substantively equivalent to the mathematical expressions representing the signals V_m and V_r in Ikeda, assume plane-wave sound). However, it would have been obvious to implement the Ikeda system in a device intended to receive near-field speech, such as a headset. For example, Ikeda states: “The present invention can be effectively applied to a wide range of devices ... such as *a voice input microphone of a hands-free telephone unit* or a car navigation system.” Ex. 1006, ¶ [0036] (emphasis added). 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. 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]. 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]; *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”).

62. When the Ikeda 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 the Ikeda system in view of McCowan. In particular, 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.

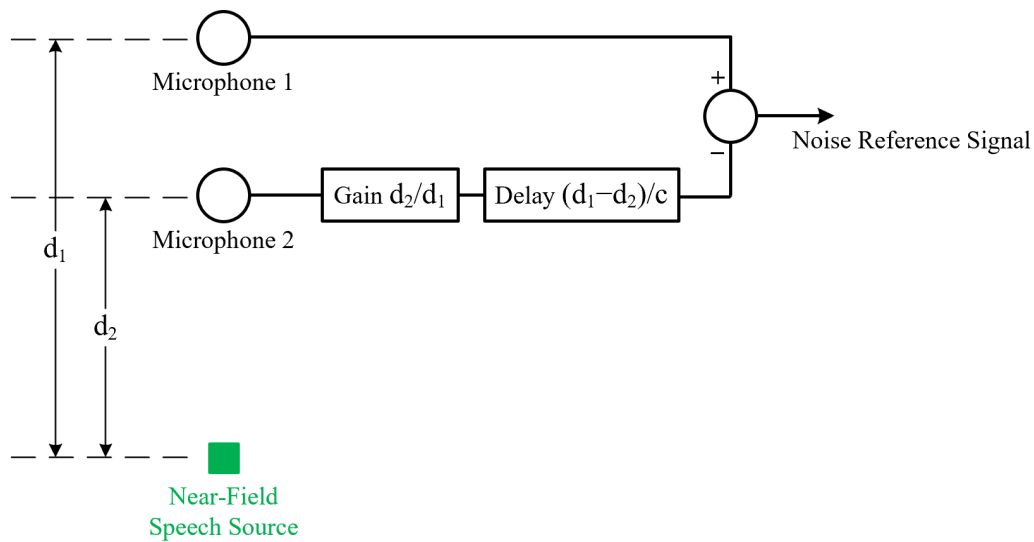
¹ McCowan teaches a gain in the near-field compensation that functions in the same way as the '080 patent's gain β in the virtual microphone V_2 . Ex. 1006, 90, 93-95; Ex. 1001, 9:50-67; § VII.B.2, ¶¶ 53-55. For simplicity, I use the term β to refer to McCowan's gain.

FIG. 1



Ex. 1006, Fig. 1 (modified).

63. A POSITA would have been motivated to make this modification for the following reasons. As discussed in Section VII.B.2, McCowan teaches that the signal of microphone 2 is multiplied by d_2/d_1 and delayed by $(d_1-d_2)/c$, before it is subtracted from the signal of microphone 1, to form a signal system producing a noise reference signal, as shown below (the **near-field speech source** is assumed to be on the axis of the microphones 1 and 2). Ex. 1007, 87-91, 93-96, Fig. 3.



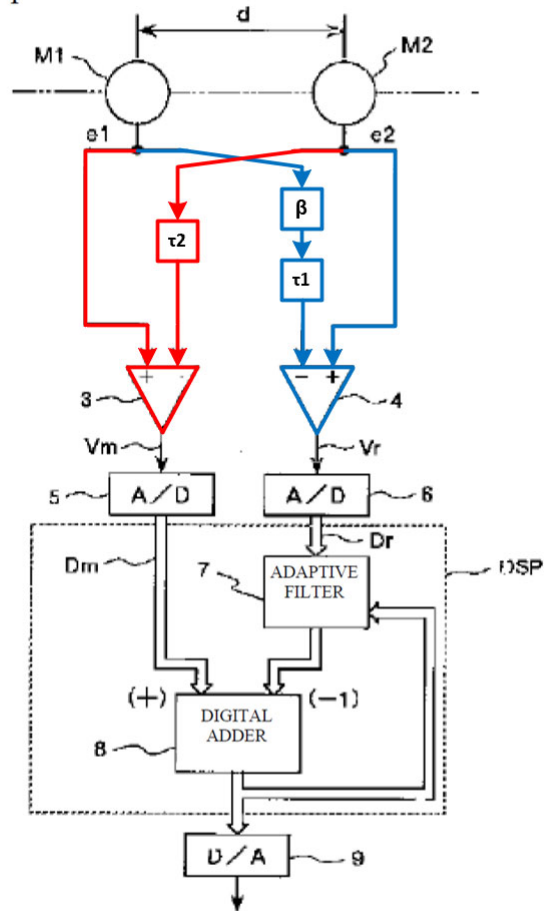
64. McCowan teaches that the gain compensates for the amplitude difference between the microphones 1 and 2 and the delay compensates for the phase difference between the microphones 1 and 2, for the near-field speech source. Ex. 1007, 89-91, 93-96, Fig. 3. The signal system with the gain and the delay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

65. Similar to McCowan, the Ikeda reference microphone outputs a noise reference signal V_r , based on its directional characteristics having a null in the direction of the target sound source. Ex. 1006, ¶¶ [0018]-[0020], [0023]-[0024], Figs. 1, 4-5. A POSITA would therefore have been motivated to apply the McCowan technique to the Ikeda 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. Consistent with Ikeda's disclosures for far-field speech, the **near-field speech source** would have been located on the axis of the Ikeda microphones M1 and M2 as shown below. Ex. 1006, ¶¶ [0019] ("the microphone unit M1 is disposed on the front surface"), [0021], [0023] ("the target sound ... from the front surface side"), Figs. 1-5. Therefore, a POSITA would have been motivated to apply the gain β to the signal of the microphone M1 for forming the reference microphone to compensate for the amplitude difference between the microphones M1 and M2, as taught by McCowan. Ex. 1007, 89-90, 93-96. The gain β is equal to the distance between the near-field speech source and the microphone M1 divided by the distance between the near-field speech source and the microphone M2. Ex. 1007, 89-90, 93-96. Ikeda discloses that the delay τ_1 is equal to the distance between the 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] ("microphone units 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.

■
Near-Field
Speech Source

FIG. 1



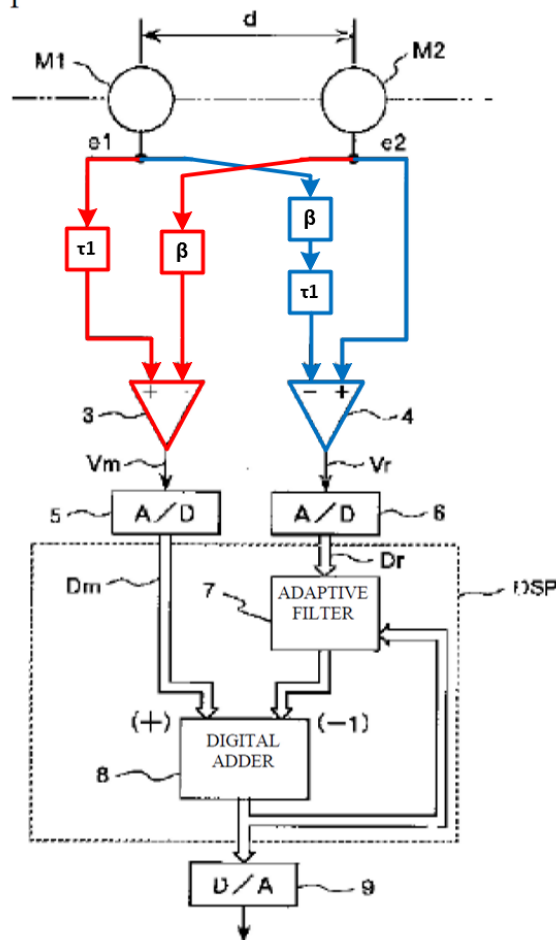
Ex. 1006, Fig. 1 (modified).

66. 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. Adding a gain to the signal path would have been within the knowledge and skill of a POSITA. 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.

67. 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.

FIG. 1



Ex. 1006, Fig. 1 (modified).

68. A POSITA would have been motivated to make this modification for the following reasons. As discussed in Section VII.B.3, Kanamori teaches that it is

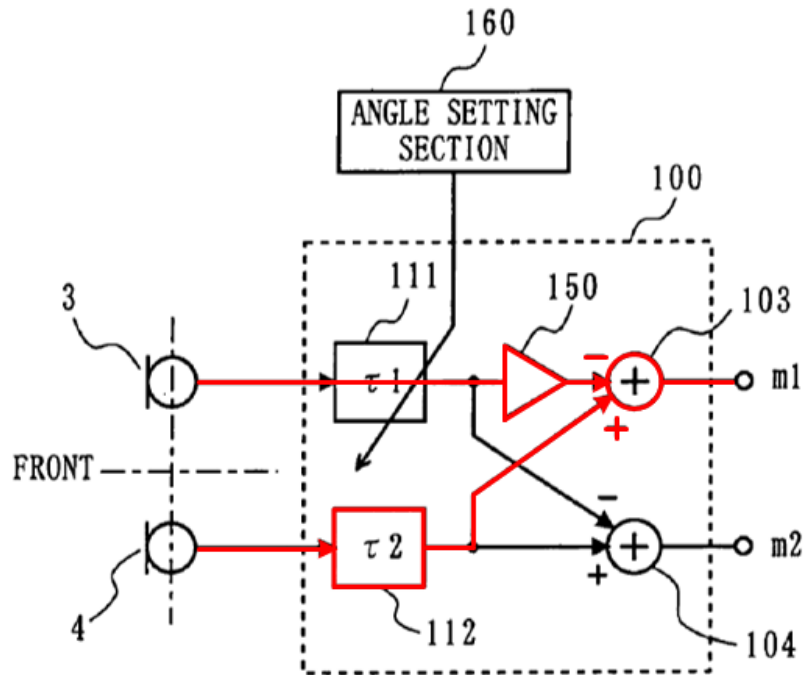
preferable that the directivity patterns of the main signal m_1 and the noise reference signal m_2 be as similar as possible in directions from which noise is coming. Ex. 1008, ¶¶ [0167]-[0170]. For example, Kanamori states that, when “noise is coming from a plurality of directions,” suppressing “noise 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] (“In order to suppress, based on the noise reference signal, noise components included in the main signal under the circumstances where noise is coming from a plurality of directions, the output of the spectrum ratio calculating section 43 illustrated in FIG. 4 has to be constant *irrespectively of the direction from which noise is coming*. That is, if the output of the spectrum ratio calculating section 43 is changed depending on the direction from which noise is coming, only the estimated noise spectrum $N_x(\omega)$ in a specific direction is correctly calculated.”) (emphasis added). A POSITA would have understood that it is preferable that the directivity patterns of the main signal m_1 and the noise reference signal m_2 with respect to noise be as similar as possible so that a more accurate estimation of noise included in the main signal m_1 based on the noise reference signal m_2 can be determined “irrespectively of the direction from which noise is coming.” Ex. 1008, ¶¶ [0101]-[0103], [0168], Figs. 1, 4. The more accurate estimation of noise included in the main signal m_1 is then used to

suppress noise components included in the main signal m_1 . Ex. 1008, ¶¶ [0077]-[0083], [0100]-[0115], [0168], Figs. 1, 4.

69. Similar to Kanamori, Ikeda also teaches using the adaptive filter 7 to estimate the noise included in the main signal V_m based on the reference signal V_r , and using the estimated noise to cancel noise components included in the main signal V_m . Ex. 1006, ¶¶ [0027]-[0029]. A POSITA would therefore have been motivated to apply Kanamori's teachings to the Ikeda-McCowan system, so that the adaptive filter 7 more accurately estimates the noise included in the main signal V_m based on the reference signal V_r "irrespective of the direction from which noise is coming," resulting in better noise reduction for the main signal V_m . Ex. 1008, ¶ [0168]; *see also* Ex. 1006, ¶¶ [0027]-[0029]; Ex. 1008, ¶¶ [0077]-[0083], [0100]-[0115].

70. Kanamori teaches modifying the main microphone in the Ikeda-McCowan system 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. For example, as shown in Figure 17A (annotated below), the annotated configuration for generating the main signal m_1 , 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 V_m in the Ikeda-McCowan system. Ex. 1008, ¶¶ [0164]-[0176].

FIG. 17A



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

71. A POSITA would have understood that the Ikeda-McCowan main signal V_m , when generated using the configuration annotated above, would have a directivity pattern with respect to noise from the far-field as similar as possible or identical to the directivity pattern of the Ikeda-McCowan reference signal V_r with respect to noise from the far-field, by sharing the same delay configuration of the signals from the 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.

72. Kanamori teaches that the main signal m1 and the noise reference signal m2 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 the microphone 3 is either more delayed or less delayed than the signal from the 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. A POSITA would therefore have understood that using the same relative delay between the signals from the microphones 3 and 4 for each of the two signal systems that generate the main signal m1 and the noise reference signal m2 would allow the directivity patterns of m1 and m2 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].

FIG. 17B

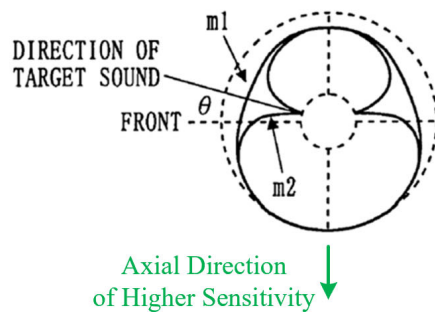
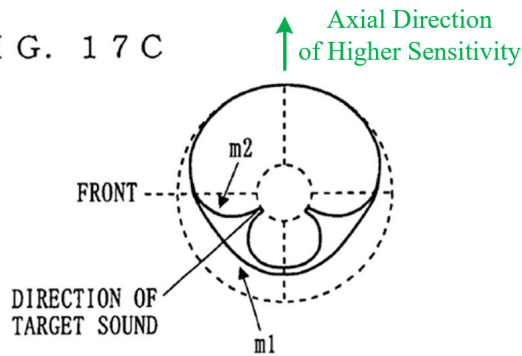


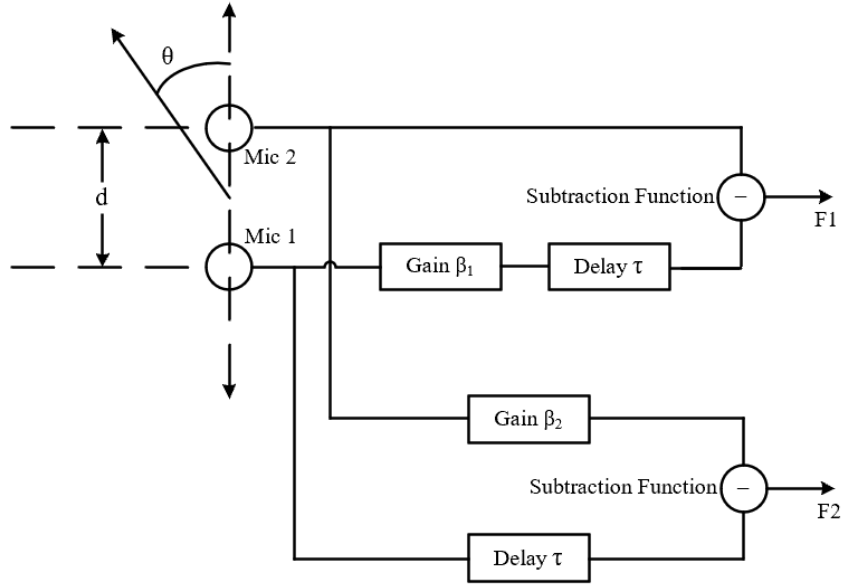
FIG. 17C



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

73. 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. 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 either a relatively delayed signal path or a relatively advanced 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. The following figure shows examples of such differential microphones with the

same delay configuration (i.e., the signal from Mic 1 is delayed by τ for F1 and F2).



74. A POSITA would have understood that, assuming far-field planar wave, the far-field directivity pattern of F1 would be given as follows:

$$D_{F1}(f, \theta) = \left| 1 - \beta_1 \cdot e^{-j2\pi f \left(\tau + \frac{d \cdot \cos(\theta)}{c} \right)} \right|$$

where f is the frequency of soundwave and c is the speed of sound. 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. 1007, 89-90; § VII.B.5. Let x denote $-2\pi f(\tau + d \cdot \cos(\theta)/c)$, a POSITA would have understood that the far-field directivity pattern of F1 would be as follows:

$$D_{F1}(f, \theta) = |1 - \beta_1 \cdot e^{jx}|$$

$$\begin{aligned}
 &= |1 - \beta_1 \cdot (\cos(x) + j \cdot \sin(x))| \\
 &= \sqrt{(1 - \beta_1 \cdot \cos(x))^2 + \beta_1^2 \cdot \sin^2(x)} \\
 &= \sqrt{1 + \beta_1^2 \cdot \cos^2(x) - 2\beta_1 \cdot \cos(x) + \beta_1^2 \cdot \sin^2(x)} \\
 &= \sqrt{1 + \beta_1^2 - 2\beta_1 \cdot \cos(x)}
 \end{aligned}$$

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. 1007, 89-90.

75. A POSITA would have understood that, assuming far-field planar wave, the far-field directivity pattern of F2 would be given as follows:

$$D_{F2}(f, \theta) = \left| \beta_2 - e^{-j2\pi f \left(\tau + \frac{d \cdot \cos(\theta)}{c} \right)} \right|$$

where f is the frequency of soundwave and c is the speed of sound. 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. 1007, 89-90; § VII.B.5. Let x denote $-2\pi f(\tau + d \cdot \cos(\theta)/c)$, a POSITA would have understood that the far-field directivity pattern of F2 would be as follows:

$$\begin{aligned}
 D_{F2}(f, \theta) &= |\beta_2 - e^{jx}| \\
 &= |\beta_2 - (\cos(x) + j \cdot \sin(x))|
 \end{aligned}$$

$$\begin{aligned}
 &= \sqrt{(\beta_2 - \cos(x))^2 + \sin^2(x)} \\
 &= \sqrt{\beta_2^2 + \cos^2(x) - 2\beta_2 \cdot \cos(x) + \sin^2(x)} \\
 &= \sqrt{1 + \beta_2^2 - 2\beta_2 \cdot \cos(x)}
 \end{aligned}$$

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. 1007, 89-90.

76. A POSITA would therefore have been able to understand that the far-field directivity pattern of a differential microphone with a gain component in a relatively delayed signal path and the far-field directivity pattern of a differential microphone with a gain component in a relatively advanced signal path, where the differential microphones have the same delay configuration, are the same function of the gain. That is, 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. 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.

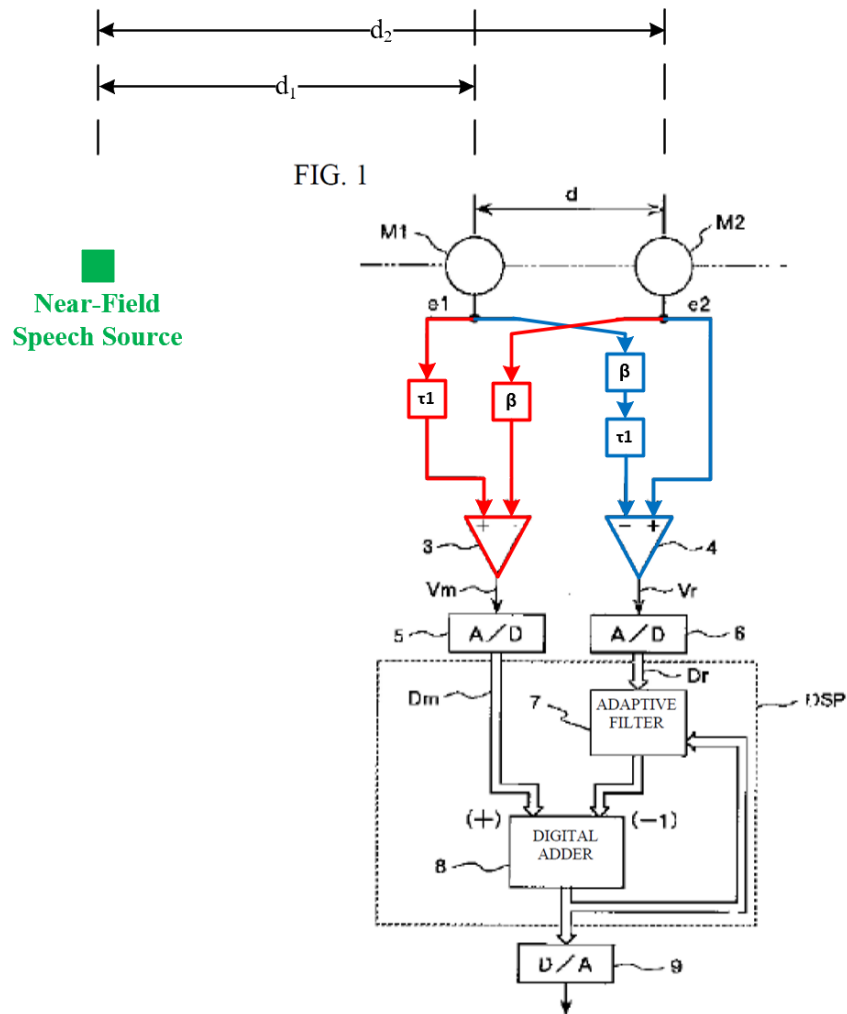
77. A POSITA would therefore have been motivated to modify the Ikeda-McCowan main microphone, using a configuration like the one annotated above in Figure 17A of Kanamori for generating the main signal m_1 , so that the main signal V_m has the same far-field directivity pattern as the Ikeda-McCowan noise reference signal V_r . In particular, a POSITA would have been motivated to cause the output signal of the Ikeda-McCowan microphone M1 to be delayed for the main microphone by the same amount (τ_1) as it is delayed for the reference microphone, as taught by Kanamori. Ex. 1008, ¶¶ [0167]-[0169], [0173]-[0176], Figs. 17A-17C. A POSITA would have been motivated to cause the output signal of the Ikeda-McCowan microphone M2 to be multiplied for the main microphone by the same amount (β) 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. The modification would have resulted in the main signal V_m and the 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. 1008, ¶ [0168].

78. 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. 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. Modifying the signal path with a delay and a gain would have been within the knowledge and skill of a POSITA. 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.

- a. **The Ikeda-McCowan-Kanamori main microphone has the same formulation for signal processing as the virtual microphone V_1 described in the '080 patent, and the Ikeda-McCowan-Kanamori reference microphone has the same formulation for signal processing as the virtual microphone V_2 described in the '080 patent**

79. As discussed above in Section VII.B.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,

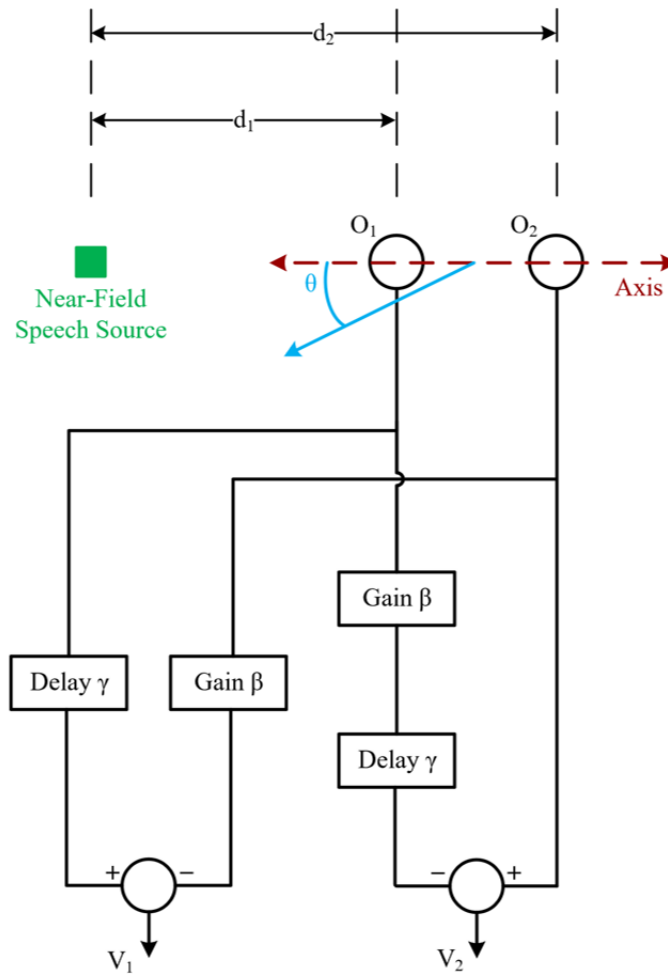
where the **near-field speech source** is located on the axis of the microphones M1 and M2, as taught by Ikeda. The microphones M1 and M2 are omnidirectional microphones. Ex. 1006, ¶ [0018]. The gain β equals d_1/d_2 , and the delay τ_1 equals $(d_2-d_1)/c$, where c is the speed of sound. § VII.B.4.



Ex. 1006, Fig. 1 (modified).

80. As discussed above in Section VI.A, the 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 . The 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. § VI.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 $M1$ and $M2$, 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.



81. As depicted above, the Ikeda-McCowan-Kanamori main signal V_m , which would have been obvious for the reasons explained in Section VII.B.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). Additionally, the Ikeda-McCowan-Kanamori reference signal V_r , which would have been obvious for the reasons explained in Section VII.B.4, and the virtual microphone V_2 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).

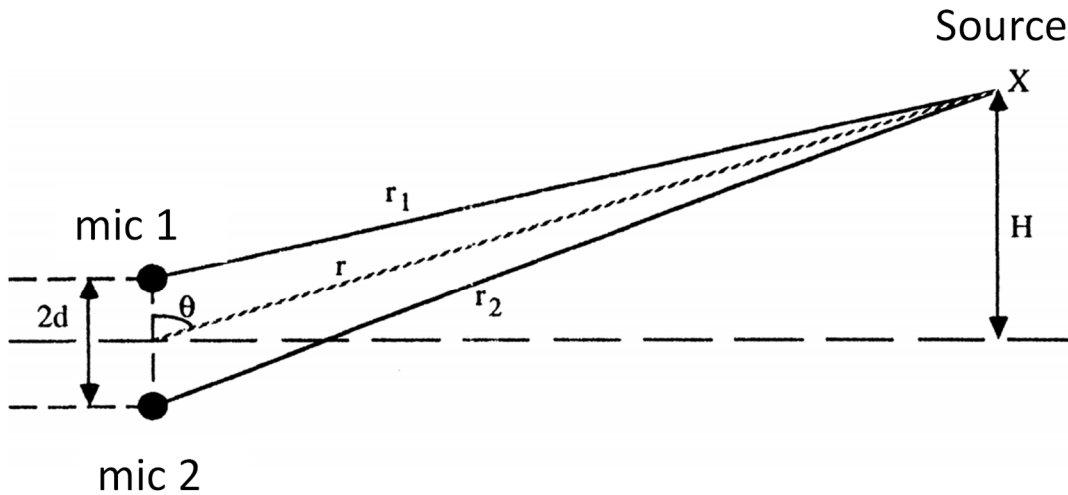
5. Simulations of Virtual Microphone Responses

82. The equation for sound radiation from a spherical source of small dimension is given by

$$p(r, t) = \frac{A}{r} e^{j(\omega t - kr)}, \quad (1)$$

where A is the amplitude of the pressure, r is the distance from the source, $e^{j\omega t}$ is the time dependence, and $\frac{e^{-jkr}}{r}$ is the spherical propagation and divergence of the soundwave. Ex. 1025, 171. For the time dependence, ω is equal to $2\pi f$, where f is the frequency of the soundwave, and t is time. For the spherical propagation of the soundwave, k is equal to $\frac{2\pi f}{c}$, where f is the frequency of the soundwave and c is the speed of sound. Equation (1) was well known to a POSITA and was conventionally used to simulate the response of a spherical source. It was also known that the same equation could be used to model an omnidirectional microphone as a result acoustic reciprocity, where $\frac{e^{-jkr}}{r}$ now represents the spherically divergent reception of acoustic waves. *See* Ex. 1025, 172-176 (explaining the concept of acoustic reciprocity). Note there is no angular or θ dependence in equation (1) because the microphone is omnidirectional. However, when two microphone signals from two omnidirectional microphones are

processed and summed (or differenced) to create a virtual microphone, then the resulting virtual microphone will have a pick-up pattern or linear response that varies with θ by virtue of the spacing between the physical microphones. From the figure below, a two-microphone array can be defined by the microphone spacing ($2d$), average radius, r , and radii between mic 1 and mic 2 and the source location X (at angle θ), given by r_1 and r_2 , respectively. Ex. 1026, 151-152.



Ex. 1026, 152. Distances r_1 and r_2 can be found from the law of cosines as:

$$r_1 = \sqrt{(r^2 + d^2 - 2rd \cos(\theta))}, \quad (2)$$

$$r_2 = \sqrt{(r^2 + d^2 + 2rd \cos(\theta))}. \quad (3)$$

Ex. 1026, 152.

83. For the two-microphone array in the figure above, the pressure from summing (differencing) two identical omnidirectional microphones for a source of strength, A , will be:

$$p(r, \theta, t) = p_1(t) \pm p_2(t) = Ae^{j\omega t} \left(\frac{e^{-jkr_1}}{r_1} \pm \frac{e^{-jkr_2}}{r_2} \right), \quad (4)$$

which now has θ -dependence as a result of equations (2) and (3). The response of the virtual microphone given by equation (4) is simply the sum (difference) of the two physical microphone signals as a result of linear superposition. If the sign is positive, we have a sum array and if negative, a difference array. Removing the time dependence to focus on geometric dependence and substituting equations (2) and (3) into equation (4) gives

$$p(r, \theta) = \left| A \left(\frac{e^{-jk\sqrt{(r^2+d^2-2rd\cos(\theta))}}}{\sqrt{(r^2+d^2-2rd\cos(\theta))}} \pm \frac{e^{-jk\sqrt{(r^2+d^2+2rd\cos(\theta))}}}{\sqrt{(r^2+d^2+2rd\cos(\theta))}} \right) \right|. \quad (5a)$$

Equation (5a) is the linear pressure response of the virtual microphone as a function of (r, θ) . As can be seen from equation (5a), the response is linear because the pressure p of the virtual microphone changes linearly with the amplitude A of the pressure of the sound source. Multiplying equation (5a) by r to normalize the effects of spherical divergence gives the shape of transverse response (just θ dependence at a given r) for the virtual microphone.

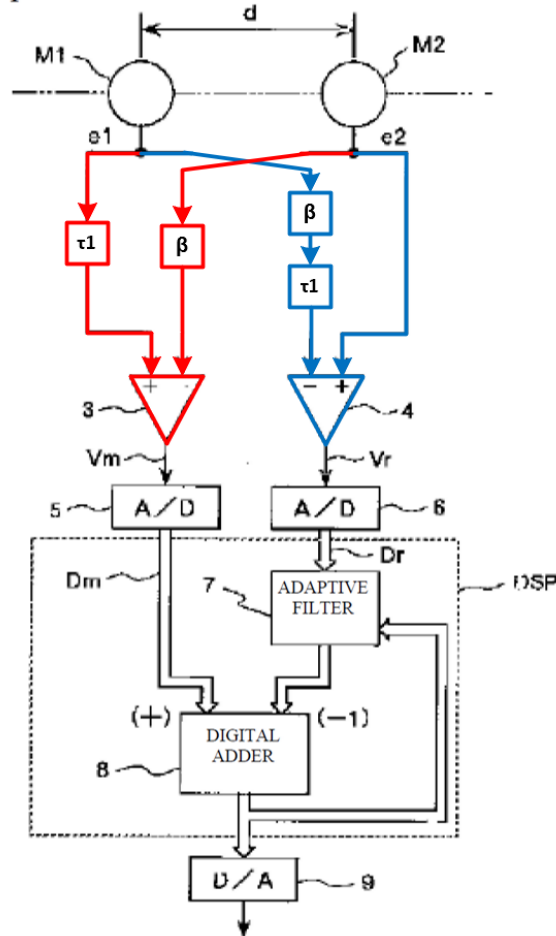
$$p(r, \theta) = r \left| \left(p_1 \frac{e^{-jk\sqrt{(r^2+d^2-2rd\cos(\theta))}}}{\sqrt{(r^2+d^2-2rd\cos(\theta))}} \pm p_2 \frac{e^{-jk\sqrt{(r^2+d^2+2rd\cos(\theta))}}}{\sqrt{(r^2+d^2+2rd\cos(\theta))}} \right) \right|, \quad (5b)$$

where p_1 and p_2 are the pressures measured at mics m_1 and m_2 , respectively. If we again assume that m_1 and m_2 are identical, then we can let $p_1 = p_2 = A$, then divide equation (5b) by A to get the linear response, which would be the response for a unit acoustic pressure source ($A = 1$).

$$p(\theta) = r \left| \left(\frac{e^{-jk\sqrt{(r^2+d^2-2rd\cos(\theta))}}}{\sqrt{(r^2+d^2-2rd\cos(\theta))}} \pm \frac{e^{-jk\sqrt{(r^2+d^2+2rd\cos(\theta))}}}{\sqrt{(r^2+d^2+2rd\cos(\theta))}} \right) \right|, \quad (5c)$$

84. As discussed above in Section VII.B.4, the Ikeda-McCowan-Kanamori system is shown below:

FIG. 1



Ex. 1006, Fig. 1 (modified).

For the above system, the signal output from the microphone M2 is multiplied in the signal path for the **main microphone** by a gain β ; the signal output from the microphone M1 is multiplied in the signal path for the **reference microphone** by a gain β ; the signal output from the microphone M1 is delayed in the signal path for the **main microphone** and in the signal path for the **reference microphone** by time τ_1 , which can be modeled as $e^{-j\omega\tau_1}$. Ex. 1007, 90. Assuming a unit amplitude

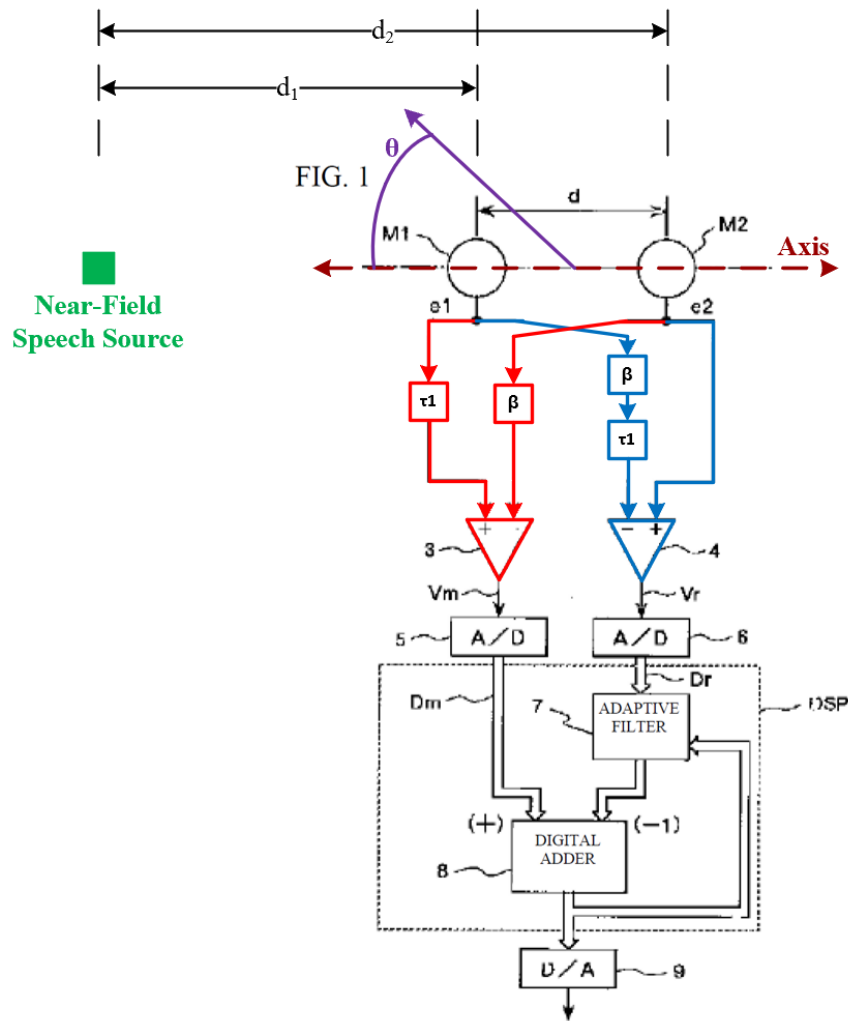
source, the linear response of the two virtual microphones, $V_m(\theta)$ and $V_r(\theta)$ can be computed considering equation (5c), respectively, as:

$$V_m(\theta) = r \left| \underbrace{-\beta \left(\frac{e^{-jkr_2}}{r_2} \right)}_{M2 \text{ portion}} + \underbrace{e^{-j\omega\tau_1} \left(\frac{e^{-jkr_1}}{r_1} \right)}_{M1 \text{ portion}} \right|, \quad (6)$$

$$V_r(\theta) = r \left| \underbrace{\left(\frac{e^{-jkr_2}}{r_2} \right)}_{M2 \text{ portion}} - \underbrace{\beta e^{-j\omega\tau_1} \left(\frac{e^{-jkr_1}}{r_1} \right)}_{M1 \text{ portion}} \right|. \quad (7)$$

In defining the value of θ in equations (6) and (7), the microphone M1 in Ikeda-McCowan-Kanamori corresponds to “mic 1” and the microphone M2 in Ikeda-McCowan-Kanamori corresponds to “mic 2” in the above figure depicting r , r_1 , and r_2 . Ex. 1026, 152.

85. As discussed in Sections VII.B.4 and VII.B.4.a and as shown below, in the Ikeda-McCowan-Kanamori system including the **main microphone** and **reference microphone**, the **near-field speech source** is located on the axis of the microphones M1 and M2 at an angle θ of zero degrees. The 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. § VII.B.4.



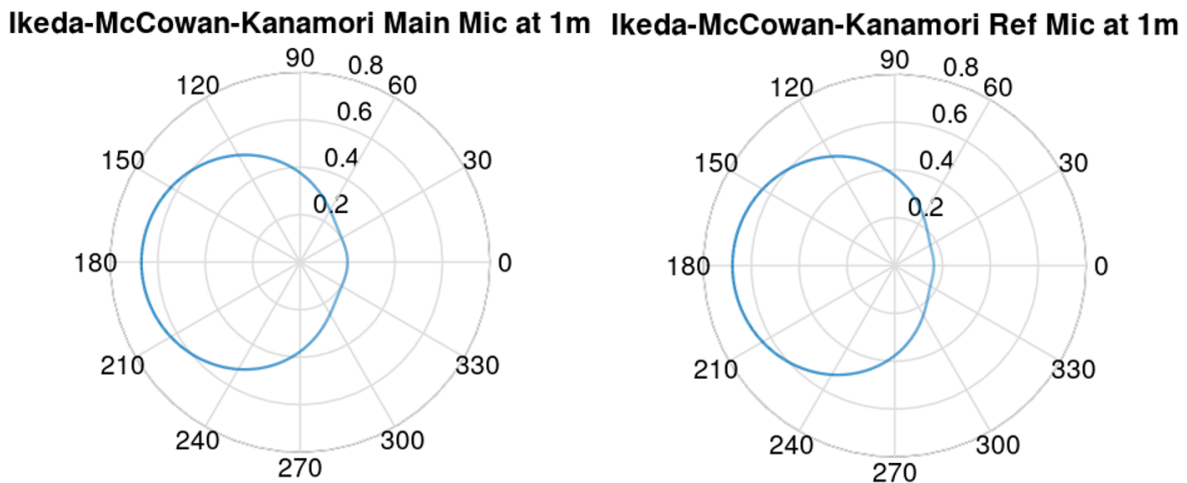
Ex. 1006, Fig. 1 (modified).

86. For purposes of simulating the linear responses of main signal V_m and reference signal V_r , the two physical microphones M1 and M2 are assumed to be 2 cm apart. Ex. 1008, ¶ [0151] (disclosing that “the microphone units are required to be a certain distance (normally, 1 cm to 5 cm) apart from each other”). The **near-field speech source** is typically between 3-5" (about 0.08-0.13 m) from the microphones in a headset, a common device well known before the '080 patent

that uses microphones to receive near-field speech from a user of the headset.

Ex. 1012 (Taenzer), ¶¶ [0009], [0078]; § VII.B.4. For purposes of simulating the linear responses of main signal V_m and reference signal V_r , the **near-field speech source** as shown above is assumed to 10 cm away from the midpoint of the microphone array of the microphones M1 and M2. Accordingly, d_1 is equal to 9 cm, d_2 is equal to 11 cm, β is equal to 9 divided by 11, and τ_1 is equal to 2 cm divided by the speed of sound (which is assumed to 343 m/s).

87. Using equations (6) and (7) and setting the parameters to the above values yields the following linear responses for the main microphone, V_m , (left) and the reference microphone, V_r , (right) when the sound source originates in the far-field.

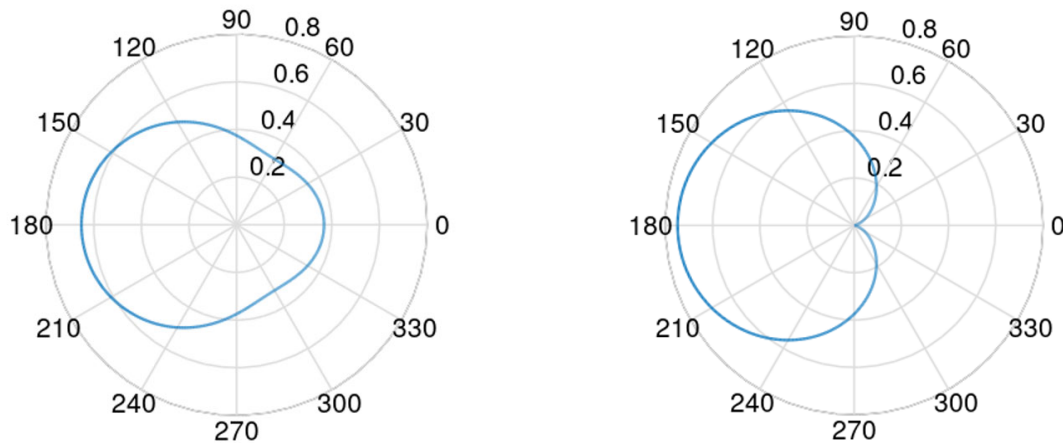


88. In the above simulation, the frequency of the sound signal was assumed to be 1000 Hz, which is within the typical frequency range for audible

sound. Ex. 1007, 87, 95-96. The sound source was plotted across all angles and assumed to be located 1 m from the microphone array, which was chosen because it is far enough from the microphone array to make the amplitude differences caused by amplitude decay due to the distance traveled negligible across the two microphones. Ex. 1007, 90. In other words, the distance from the microphone M1 to the sound source divided by the distance from the microphone M2 to the sound source is very close to 1. Ex. 1007, 90.

89. Using equations (6) and (7) and changing the distance between the microphone array and the sound source from 1 m to 0.1 m yields the following linear responses for the main microphone, V_m , (left) and the reference microphone, V_r , (right) when the sound source originates in the near-field:

Ikeda-McCowan-Kanamori Main Mic at 0.1m Ikeda-McCowan-Kanamori Ref Mic at 0.1m

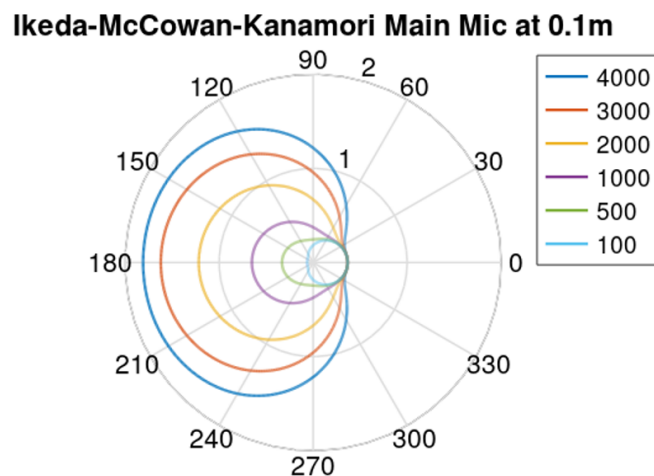


90. In the above simulation, the frequency of the sound signal was assumed to be 1000 Hz, which is within the typical frequency range for audible

sound. Ex. 1007, 87, 95-96. The sound source was plotted across all angles and assumed to be located 0.1 m from the microphone array.

91. Comparing the two near-field (speech) responses above, it can be seen that the linear responses are substantially dissimilar. The reference microphone has a single null pointed in the direction of speech ($\theta = 0^\circ$) that has a response that is lower than any other region of the linear response to speech. In contrast, the two far-field (noise) responses are substantially similar, where both linear responses are devoid of a null.

92. The near-field linear response of the main microphone, V_m , for various frequencies are given below. For a speech source located between $\theta \in [-30^\circ, 30^\circ]$, the linear response is substantially similar across a plurality of frequencies between 100 and 4,000 Hz.



93. These simulations are useful for seeing how a virtual microphone responds to a particular sound source. In a real-world implementation, such as the ones described in Ikeda, McCowan, and Kanamori, the microphone signal produced by each of the omnidirectional microphones is known. Equation (1) above is merely representative of how each omnidirectional microphone responds to a real-world sound source. Similarly, real-world implementations of virtual microphones are often implemented digitally. For example, Kanamori discloses using a “digital process” when performing the directivity combining process. Ex. 1008, ¶ [0151]. Such digital processes rely on discrete, as opposed to continuous, samples from each microphone. *See* Ex. 1008, ¶¶ [0107]-[0108]. Accordingly, while the above simulations assume continuous microphone signals, a POSITA would have understood that a real-world implementation in the digital domain would represent the delay in terms of the number of samples that correspond to the desired time delay, which can be represented as $z^{-\tau f_s}$. This is a function of the desired time delay τ and the sampling frequency f_s . For example, in the above example, the delay τ_1 is set to $(d_2 - d_1)/c$, where $d_2 - d_1 = 2$ cm and $c = 343$ m/s. If the microphones were sampled at 16 kHz, for example, a person of ordinary skill in the art would understand that the signal output from the microphone M1 would be delayed by about 0.93 samples

$(z^{\frac{d_2-d_1}{c}} f_s = z^{\frac{0.02}{343}} 16000 = z^{0.93})$. As the '080 patent acknowledges, “fractional-delay filters” were “well known to those versed in the art” and would have been a conventional technique for implementing such a delay. Ex. 1001, 9:67-10:2. A POSITA would understand that, with a sufficiently high sampling frequency, any differences between the linear responses of a real-world digital implementation of the virtual microphones discussed above and the simulations of those linear responses depicted above assuming continuous signals would be so small as to be negligible.

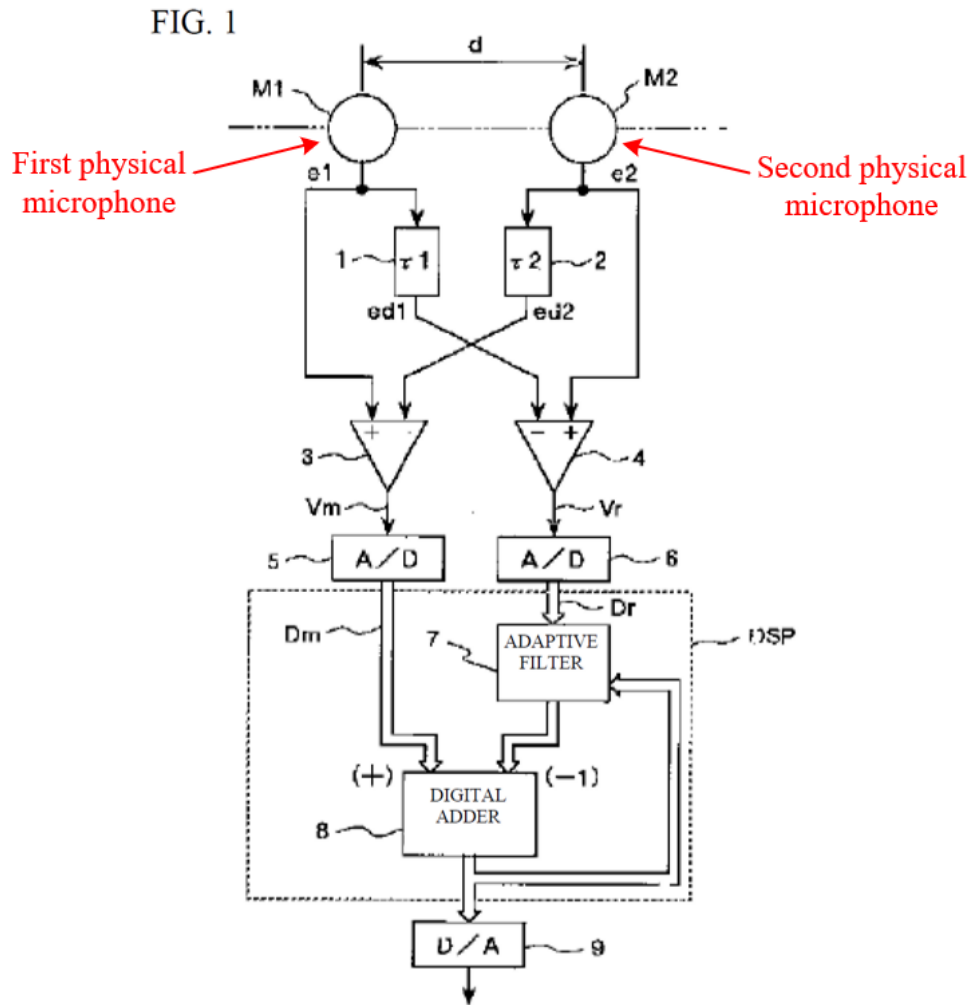
6. Claim 7

94. To simplify the declaration in light of the elements addressed in each claim, I start my analysis with independent claim 7, and then address 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;”**

95. In my opinion, Ikeda discloses, and Ikeda in view of McCowan and Kanamori renders obvious, this feature. Ikeda discloses that “omnidirectional microphone units M1 and M2” as shown in Figure 1 (annotated below) are respectively a first physical microphone and a second physical microphone. Ex.

1006, ¶¶ [0018]-[0019], Fig. 1. For example, Ikeda states: “The microphone is housed in a casing, not illustrated, so that the microphone unit M1 is disposed on the front surface. Ceramic or electrolyte-type [*sic* electret-type] microphone units or the like may be used for the microphone units M1 and M2.” Ex. 1006, ¶ [0019]. Ikeda discloses that “the output e1 of the microphone unit M1” is a first microphone signal generated by the first physical microphone, and that “the output e2 of the microphone unit M2” is a second microphone signal generated by the second physical microphone. Ex. 1006, ¶¶ [0018]-[0020], Figs. 1-2, 4.

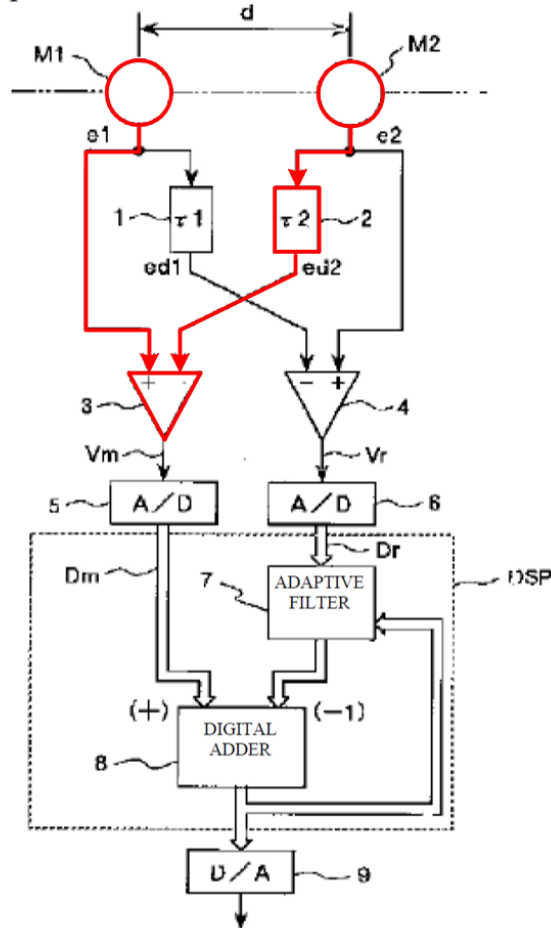


Ex. 1006, Fig. 1 (annotated).

96. 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 Vm, as shown in Figure 1 (annotated

below). 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. 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]; *see also* Ex. 1001, 11:1-2, 7:17-8:12, Figs. 3-5.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

97. A POSITA would have understood that the inversion adding circuit 3 subtracting the output signal of the microphone M2 as delayed from the output signal of the microphone M1 is a “first combination” of the first microphone signal and the second microphone signal. Ex. 1006, ¶¶ [0018]-[0022]. This understanding is consistent with combining microphone signals to form a virtual microphone in a subtraction process as described in the '080 patent. Ex. 1001, 4:42-45, 9:51-53, 11:1-2; *see also* Ex. 1001, 7:17-8:12, Figs. 3-5.

98. Additionally, as discussed in Section VII.B.4, it would have been obvious to modify Ikeda with McCowan and Kanamori. The Ikeda-McCowan-Kanamori main microphone that generates main signal V_m also teaches “a first virtual microphone formed from a first combination of a first microphone signal and a second microphone signal,” 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;

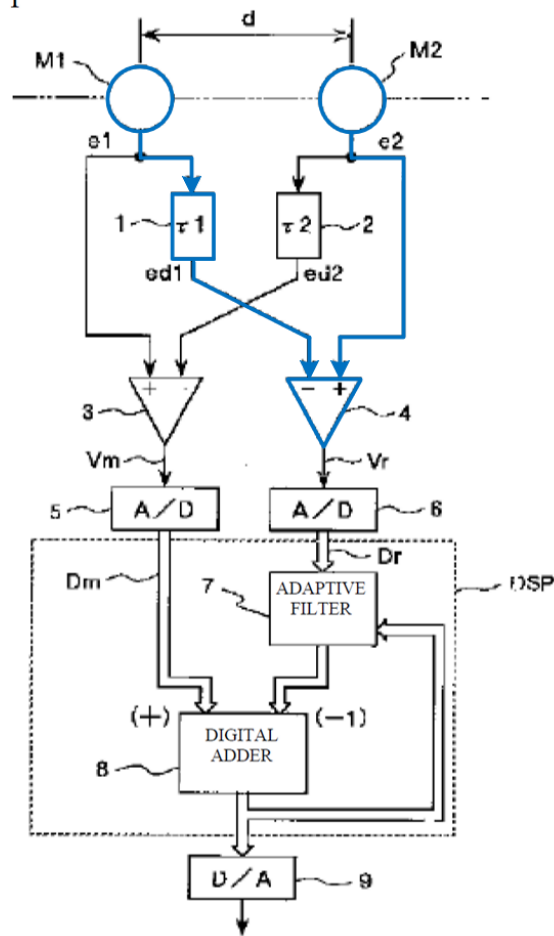
§ VII.B.4.a.

- 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,”**

99. In my opinion, Ikeda discloses, and Ikeda in view of McCowan and Kanamori renders obvious, this feature. 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 second virtual microphone by disclosing the [reference microphone](#) that constructs reference signal V_r , as shown in Figure 1 (annotated below). 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. 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 Vr. Ex. 1006, ¶¶ [0018]-[0020], [0023]-[0024]; *see also* Ex. 1001, 9:51-53, 7:17-8:12, Figs. 3-5.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

100. A POSITA would have understood that the inversion adding circuit 4 subtracting the output signal of the microphone M1 as delayed from the output signal of the microphone M2 is a “second combination” of the first microphone signal and the second microphone signal. Ex. 1006, ¶¶ [0018]-[0020], [0023]-[0024]. This understanding is consistent with combining microphone signals to form a virtual microphone in a subtraction process as described in the '080 patent. Ex. 1001, 4:42-45, 9:51-53, 11:1-2; *see also* Ex. 1001, 7:17-8:12, Figs. 3-5.

101. Ikeda discloses that the second combination (as discussed above) is different from the first combination (as discussed in Section VII.B.6.a), because the second combination is subtracting the output signal of the microphone M1 as delayed from the output signal of the microphone M2, whereas the first combination is subtracting the output signal of the microphone M2 as delayed from the output signal of the microphone M1. Ex. 1006, ¶¶ [0018]-[0024].

102. Additionally, as discussed in Section VII.B.4, it would have been obvious to modify Ikeda with McCowan and Kanamori. The Ikeda-McCowan-Kanamori reference microphone that generates reference signal V_r also teaches “a second virtual microphone formed from a second combination of the first microphone signal and the second microphone signal,” 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,

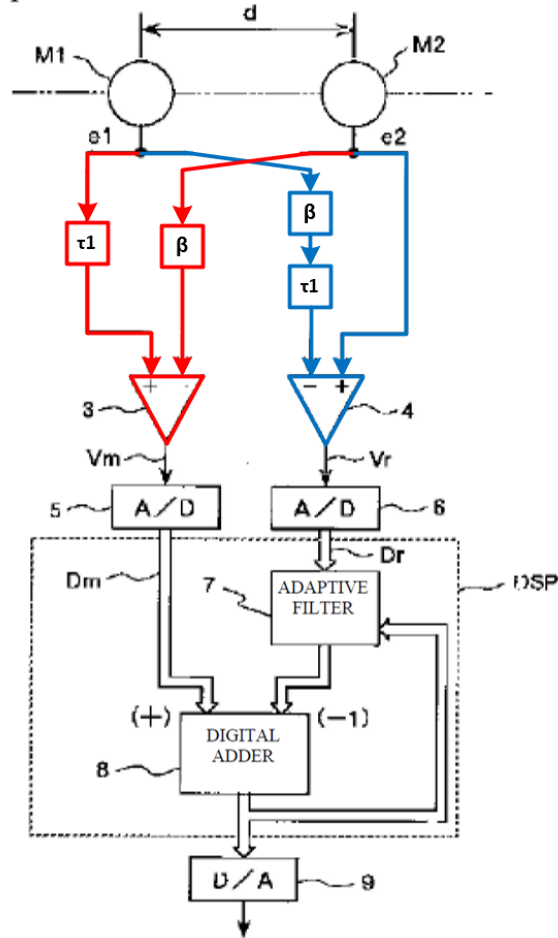
4:44-45, 9:51-53; § VII.B.4.a. In Ikeda-McCowan-Kanamori, the second combination of the first microphone signal and the second microphone signal is different from the first combination of the first microphone signal and the second microphone signal, because the Ikeda-McCowan-Kanamori main and reference microphones have different formulations for signal processing. Ex. 1001, 9:51-53, 11:1-2; § VII.B.4.a.

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

103. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. The '080 patent describes that a virtual microphone's linear response to speech refers to the virtual microphone's response to a speech source in the near-field, and that a virtual microphone's linear response to noise refers to the virtual microphone's response to a noise source in the far-field. Ex. 1001, 10:18-27, 11:20-29, 11:62-65, 16:4-6, Figs. 9-12.

104. As discussed in Section VII.B.4, it would have been obvious to modify Ikeda with McCowan and Kanamori to result in a system as shown below to be used in an environment where a speech source is in the near-field and one or more noise sources are in the far-field.

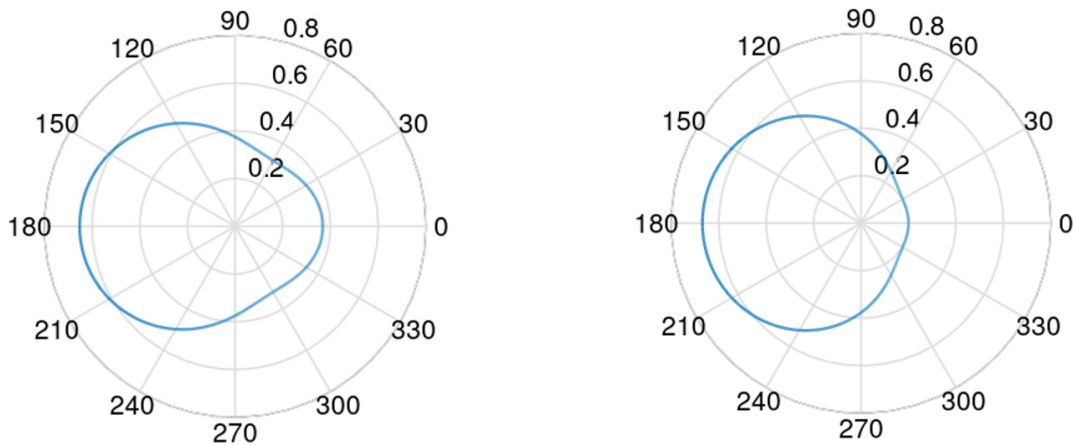
FIG. 1



Ex. 1006, Fig. 1 (modified).

105. As discussed in Section VII.B.5, I simulated and plotted the linear responses of the Ikeda-McCowan-Kanamori main signal V_m to a sound source located 0.1 m from the microphone array of the microphones M1 and M2 (near-field speech), and to a sound source located 1 m from the microphone array (far-field noise):

Ikeda-McCowan-Kanamori Main Mic at 0.1m **Ikeda-McCowan-Kanamori Main Mic at 1m**



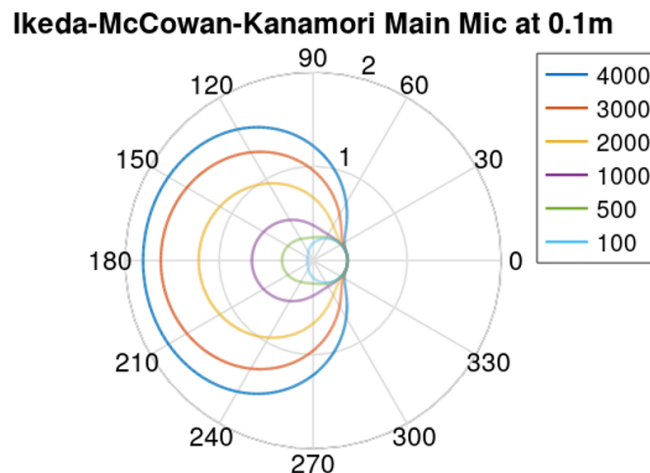
(linear response to speech (left) and linear response to
noise (right)).

106. 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. 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. 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. 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; § VII.B.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.

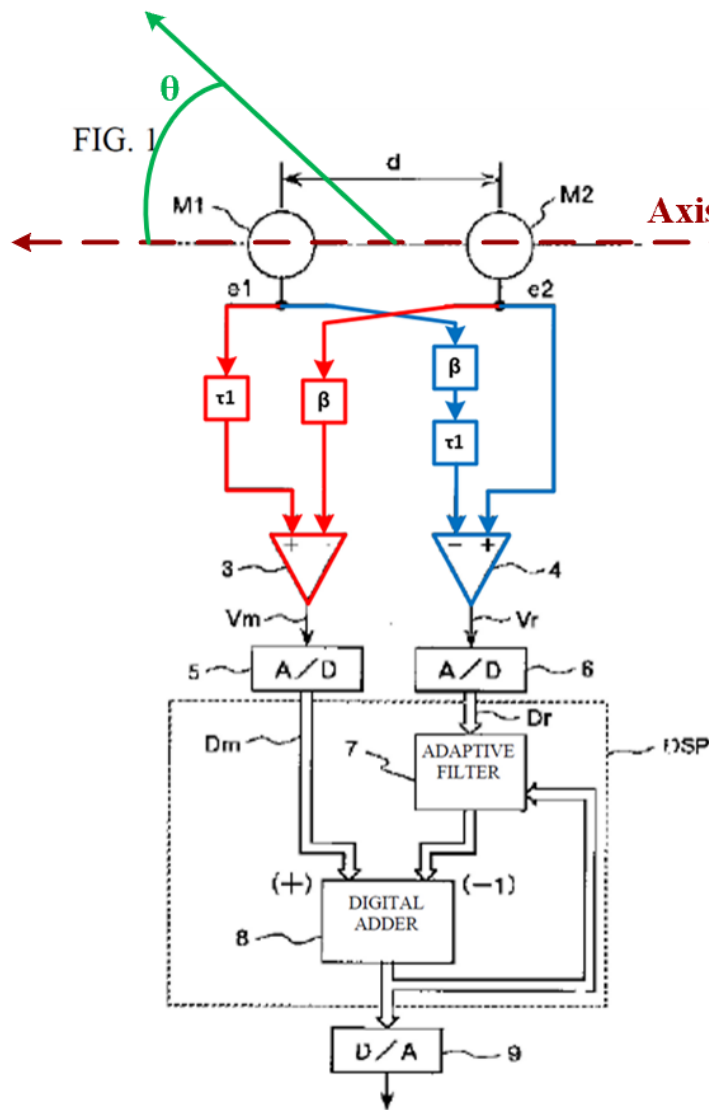
- d. [7d] **“the first linear response to speech being substantially similar across a plurality of frequencies for a speech source located within a predetermined angle relative to an axis of the microphone array and devoid of a null,”**

107. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. As discussed in Section VII.B.5, I 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 the microphones M1 and M2:



(linear response to speech).

108. In the above depiction of the Ikeda-McCowan-Kanamori 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. 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 the microphones M1 and M2, which includes the direction of speech. Ex. 1006, ¶¶ [0019] (“the microphone unit M1 is disposed on the front surface”), [0021], [0023] (“the target sound ... from the front surface side”), Figs. 1-5. 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. 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. As depicted above, the linear response to speech is devoid of a null. Ex. 1001, 4:31-32 (“The term ‘null’ means a zero or minima in the spatial response of a physical or virtual directional microphone.”).



Ex. 1006, Fig. 1 (modified).

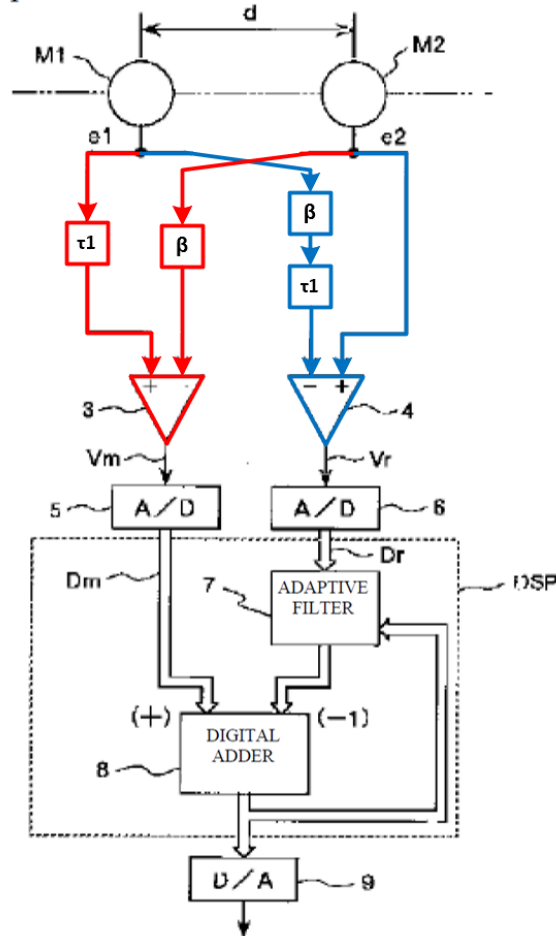
109. Additionally, the Ikeda-McCowan-Kanamori main signal V_m , which is constructed using the same signal processing as the virtual microphone V_1 described in the '080 patent, teaches the features recited in the limitation [7d], because the virtual microphone V_1 has the features. Ex. 1001, 2:60-63, 3:1-4, 4:42-43, 11:1-2, 11:20-24, 11:30-61, Figs. 11, 13; § VII.B.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,”**

110. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. The '080 patent describes that a virtual microphone's linear response to speech refers to the virtual microphone's response to a speech source in the near-field, and that a virtual microphone's linear response to noise refers to the virtual microphone's response to a noise source in the far-field. Ex. 1001, 10:18-27, 11:20-29, 11:62-65, 16:4-6, Figs. 9-12.

111. As discussed in Section VII.B.4, it would have been obvious to modify Ikeda with McCowan and Kanamori to result in a system as shown below to be used in an environment where a speech source is in the near-field and one or more noise sources are in the far-field.

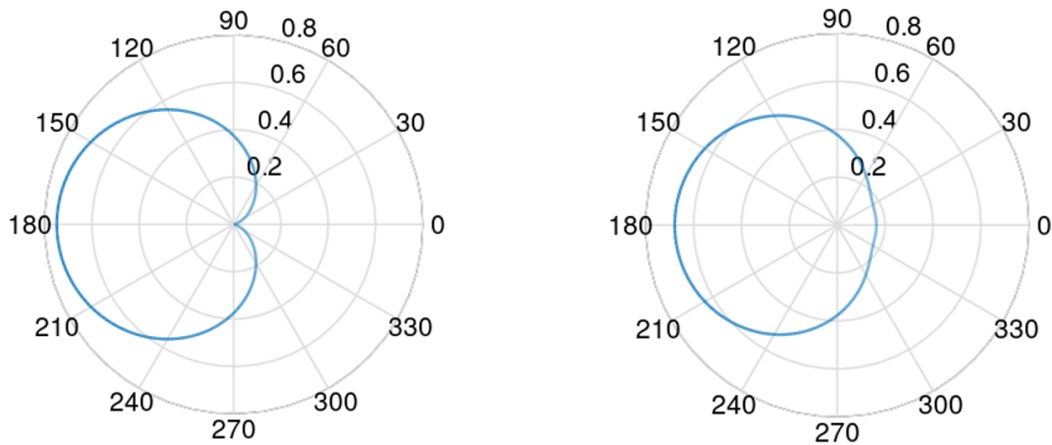
FIG. 1



Ex. 1006, Fig. 1 (modified).

112. As discussed in Section VII.B.5, I simulated and plotted the linear responses of the Ikeda-McCowan-Kanamori reference signal Vr to a sound source located 0.1 m from the microphone array of the microphones M1 and M2 (near-field speech), and to a sound source located 1 m from the microphone array (far-field noise):

Ikeda-McCowan-Kanamori Ref Mic at 0.1m Ikeda-McCowan-Kanamori Ref Mic at 1m



(linear response to speech (left) and linear response to
noise (right)).

113. 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. A POSITA would have understood that the Ikeda-McCowan-Kanamori reference signal V_r has linear responses because reference signal V_r 's response to a sound source at a particular distance, frequency, and direction is linearly proportional to its input. For example, as depicted in the above plots, 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 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. Moreover, the Ikeda-McCowan-Kanamori reference signal V_r has linear responses just like the 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; § VII.B.4.a. Therefore, the plots show the second virtual microphone has a second linear response to speech and a second linear response to noise.

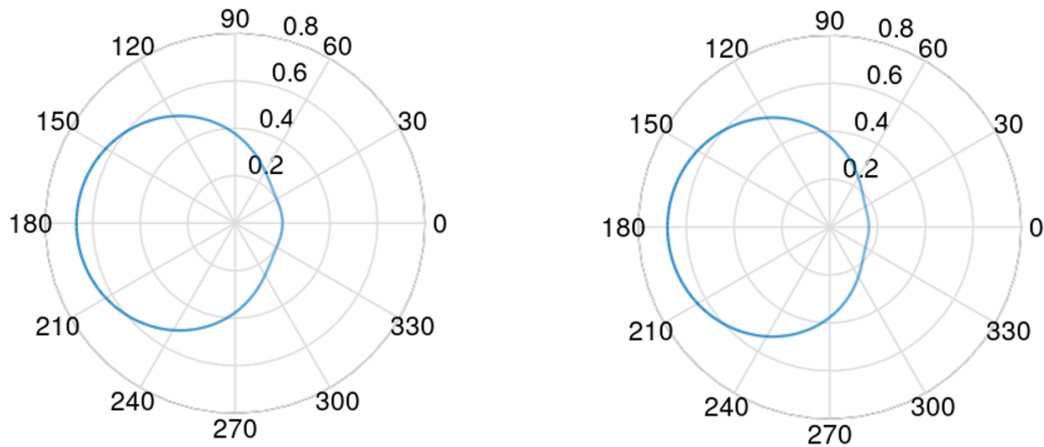
114. As also shown above, the Ikeda-McCowan-Kanamori reference signal V_r 's linear response to speech has a single null oriented at 0 degrees, which in the above simulation is the direction toward a source of the speech. Ex. 1001, 4:31-32 (“The term ‘null’ means a zero or minima in the spatial response of a physical or virtual directional microphone.”); *see also* §§ VII.B.4, VII.B.6.d.

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

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

116. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. As discussed in Sections VII.B.6.c and VII.B.6.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



(linear responses to noise).

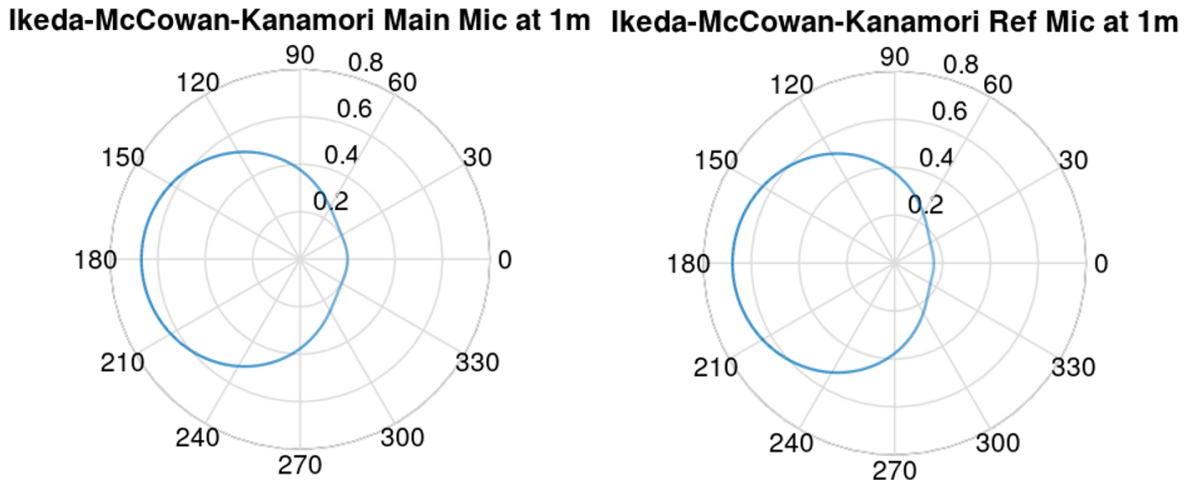
117. As shown above, Ikeda-McCowan-Kanamori teaches that the second linear response to noise (right) is substantially similar to the first linear response to noise (left).

118. Additionally, the Ikeda-McCowan-Kanamori main and reference signals, which are constructed using the same signal processing as the virtual microphones V_1 and V_2 described in the '080 patent, teach the features recited in the limitation [7f], because the virtual microphones V_1 and V_2 have the features. Ex. 1001, 11:14-16, 16:4-6; § VII.B.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 non-zero in a direction toward a source of noise, and”

119. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. As discussed in Sections VII.B.6.c and VII.B.6.e, in the Ikeda-

McCowan-Kanamori system, the first linear response to noise and the second linear response to noise are shown below:



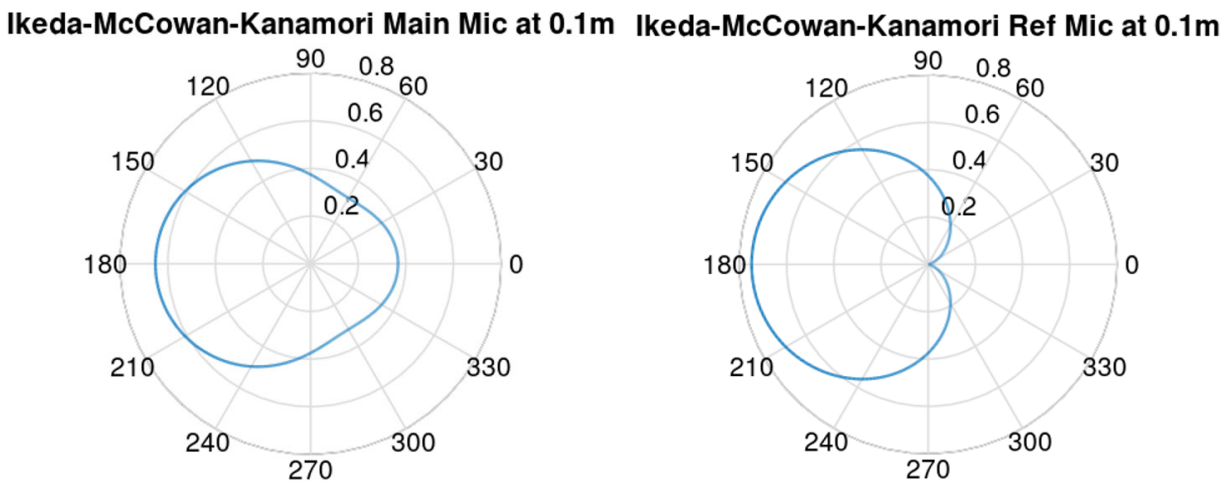
(linear responses to noise).

120. As shown above, Ikeda-McCowan-Kanamori teaches one or both of the first linear response to noise (left) and the second linear response to noise (right) being non-zero in a direction toward a source of noise because both the first and second linear responses to noise are non-zero in all directions.

121. Additionally, the Ikeda-McCowan-Kanamori main and reference signals, which are constructed using the same signal processing as the virtual microphones V_1 and V_2 described in the '080 patent, teach the features recited in the limitation [7g], because the virtual microphones V_1 and V_2 have the features. Ex. 1001, 2:56-59, 2:64-67, 10:22-35, 11:25-29, Figs. 10, 12; § VII.B.4.a.

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

122. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. As discussed in Sections VII.B.6.c and VII.B.6.e, in the Ikeda-McCowan-Kanamori system, the first linear response to speech and the second linear response to speech are shown below:



(linear responses to speech).

123. As shown above, Ikeda-McCowan-Kanamori teaches that 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.

124. Additionally, the Ikeda-McCowan-Kanamori main and reference signals, which are constructed using the same signal processing as the virtual microphones V_1 and V_2 described in the '080 patent, teach the features recited in

the limitation [7h], because the virtual microphones V_1 and V_2 have the features.

Ex. 1001, 16:4-6; § VII.B.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”

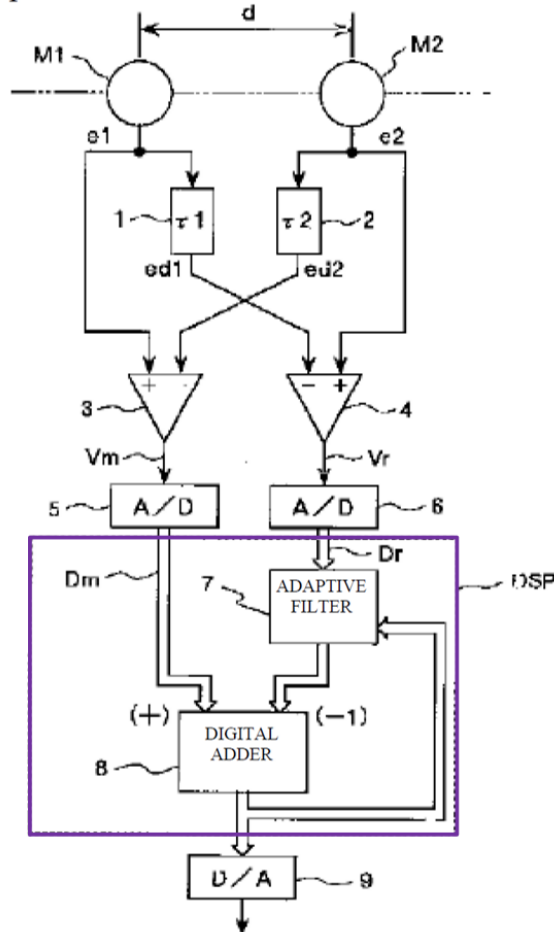
125. In my opinion, Ikeda discloses or renders obvious this feature. Ikeda states: “The present invention can be effectively applied to a wide range of devices ... such as a *voice input microphone of a hands-free telephone unit* or a car navigation system.” Ex. 1006, ¶ [0036] (emphasis added). A POSITA would have understood that “voice input” refers to human speech. Ex. 1006, ¶ [0036]. It would have been obvious to a POSITA to use the Ikeda system to receive human speech.

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,”

126. In my opinion, Ikeda discloses or renders obvious this feature. 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]. 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].

FIG. 1



Ex. 1006, Fig. 1 (annotated).

127. 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”). Ex. 1006, ¶¶ [0027]-[0029]. For example, Ikeda states: “The output V_m of the inversion adding circuit 3 is converted to a digital signal D_m by an A/D converter 5, and the

output V_r of the inversion adding circuit 4 is converted to a digital signal D_r by an A/D converter 6. Main signal data D_r outputted from the A/D converter 6 is supplied to an adaptive filter 7 such as a transversal filter, and the difference between the filter output of the adaptive filter 7 and the main signal data D_m outputted from the A/D converter 5 is calculated by a digital adder 8.” Ex. 1006, ¶ [0027]. A POSITA would have understood 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 via the A/D converters 5 and 6. Ex. 1006, ¶ [0027]; Ex. 1001, 32:25-32.

128. Ikeda discloses that the **adaptive noise cancellation system** generates denoised output signals. Ex. 1006, ¶¶ [0028]-[0029]. For example, Ikeda states: “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]. A POSITA would have understood that the output signals (in a digital or analog form) of the **adaptive noise cancellation system** based on “the noise component contained in the main signal V_m [being] cancelled by the noise component contained in the reference signal V_r ” are denoised output signals. Ex. 1006, ¶¶ [0028]-[0029].

129. Ikeda discloses that the **adaptive noise cancellation system** generates denoised output signals by forming a plurality of combinations of signals output from the first virtual microphone and the second virtual microphone because main signal Vm and reference signal Vr 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]; *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 description in the ’080 patent 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; *see also* Ex. 1001, 8:50-9:3, 9:22-38.

130. Ikeda discloses that the **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 Vr are “supplied to an adaptive filter” (filtering) and the output of adaptive filter 7 and digital versions of main signal Vm are differenced at digital adder 8 (summing). Ex. 1006, ¶¶ [0027]-[0029]. A POSITA would have understood that when the adaptive filter 7 and the digital adder 8 process main signal Vm and reference signal Vr, this processing

discloses the claimed feature “filtering and summing the plurality of combinations of signals in the time domain.” Ex. 1006, ¶¶ [0027]-[0029]. This understanding agrees with the description in the ’080 patent 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; *see also* Ex. 1001, 8:50-9:3, 9:22-38. A POSITA would have understood that the adaptive filter 7 filters in the time domain, and the digital adder 8 adds in the time domain. Both operate in the time domain because (1) the Ikeda 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]. Additionally, it would have been obvious to a POSITA to implement the Ikeda system including the adaptive filter 7 and the digital adder 8 to process signals in the time domain.

131. A POSITA would have understood that the **adaptive noise cancellation system** discloses a varying transfer function between the plurality of combinations of signals because Ikeda discloses that the characteristics of adaptive filter 7 “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]. 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]. 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 the digital adder 8 should be zero (because there is no speech). Ex. 1006, ¶ [0028]. By updating the tap coefficient over time to minimize this error signal, such as by using well known techniques like least mean squares, 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 Vm at a time when there is no speech. Ex. 1006, ¶ [0028]. This understanding comports with the description in the '080 patent 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; *see also* Ex. 1001, 5:19-44, 6:26-29, 8:50-9:3, 9:22-38.

132. Ikeda discloses that the transfer function is linear. Ex. 1006, ¶¶ [0027]-[0029]. For example, Ikeda discloses that the 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]. 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. 1028, 366-367 (“Because of the chain of delay elements across the top of the diagram [of Figure 6.31], this structure is also referred to as a *tapped delay line* structure or a *transversal filter* structure. As seen from Figure 6.31, the signal at each tap along this chain is weighted by the appropriate coefficient (impulse-response value), and the resulting products are summed to form the output”), Fig. 6.31. It also would have been obvious to a POSITA to use a linear filter to implement the 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.

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

133. In my opinion, Ikeda discloses this feature. 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 the 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 Vm” canceled and “primarily consist[] of the target sound component.” Ex. 1006, ¶¶ [0001], [0011]-[0012], [0029], Abstract.

7. 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.”

134. Claim 8 recites: “The system of claim 7 and further comprising: a microphone array, the first and second physical microphones positioned m [sic] the microphone array.” Ex. 1001, Claim 8.

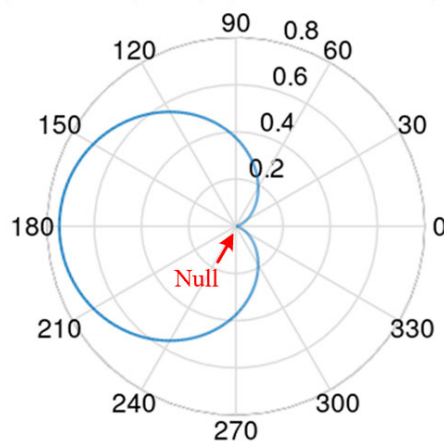
135. In my opinion, Ikeda discloses this feature. Ikeda discloses a microphone array including “two omnidirectional microphone units M1 and M2 ... disposed at a distance d apart from one another.” Ex. 1006, ¶ [0018], Fig. 1. Ikeda discloses that the microphones M1 and M2 (“the first and second physical microphones”) are “positioned m [sic] the microphone array.” Ex. 1001, Claim 8; Ex. 1006, ¶ [0018], Fig. 1; Ex. 1002, 146 (corresponding then-pending claim 9 reciting “the first and second physical microphones positioned in the microphone array”).

8. **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.”**

136. Claim 9 recites: “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.” Ex. 1001, Claim 9.

137. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. As discussed in Section VII.B.6.e, in the Ikeda-McCowan-Kanamori system, the second linear response to speech that has the single **null** is shown below:

Ikeda-McCowan-Kanamori Ref Mic at 0.1m



(linear response to speech).

138. As shown above, the Ikeda-McCowan-Kanamori combination teaches that 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. Ex. 1001, 2:52-55, 10:18-32, Fig. 9; § VII.B.5.

9. 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.”

139. Claim 11 recites: “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.” Ex. 1001, Claim 11.

140. In my opinion, Ikeda renders this feature 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”).

141. Ikeda states: “The present invention can be effectively applied to a wide range of devices ... such as *a voice input microphone of a hands-free telephone unit* or a car navigation system.” Ex. 1006, ¶ [0036] (emphasis added). It would have been obvious to a POSITA that there is a communications channel between the “voice input microphone” to which the Ikeda system is applied and the “hands-free telephone unit.” A POSITA would therefore have understood that the

communications channel is coupled with the processing component in the Ikeda system.

142. 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. 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. 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].

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

143. Claim 12 recites: “The system of claim 11 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.” Ex. 1001, Claim 12.

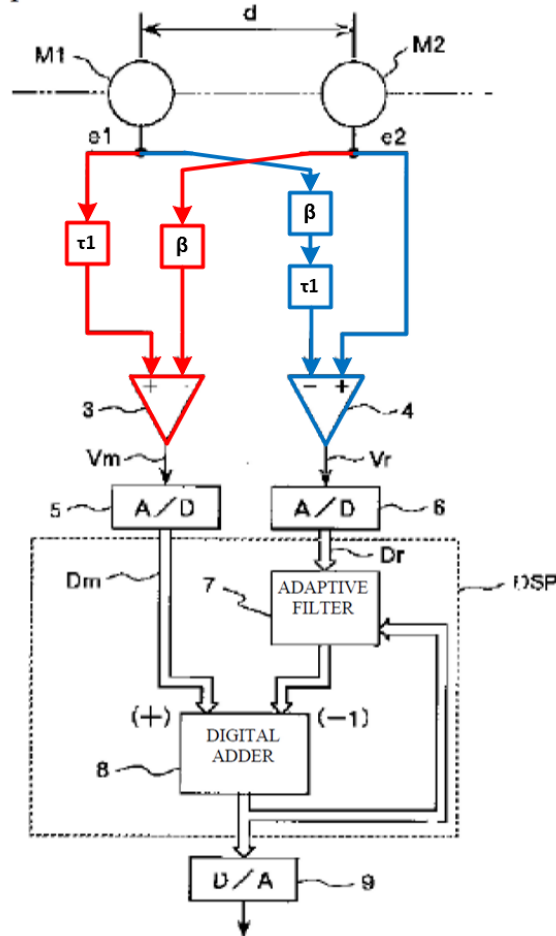
144. In my opinion, Ikeda discloses that the “hands-free telephone unit” is a communication device. Ex. 1006, ¶ [0036]. As discussed in Section VII.B.9, the communications channel is between the “voice input microphone” to which the Ikeda system is applied and the “hands-free telephone unit.” It would have been obvious to a POSITA that the communication device is wirelessly coupled with the wireless channel of the communications channel.

- 11. 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.”**

145. Claim 13 recites: “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.” Ex. 1001, Claim 13.

146. In my opinion, Ikeda in view of McCowan and Kanamori renders this feature obvious. As discussed in Section VII.B.4, it would have been obvious to modify Ikeda with McCowan and Kanamori to result in a system as shown below.

FIG. 1



Ex. 1006, Fig. 1 (modified).

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

12. 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,”**

148. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.a with respect to the limitation [7a].

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

149. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.c with respect to the limitation [7c].

- 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,”**

150. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.d with respect to the limitation [7d].

- 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;”**

151. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.a with respect to the limitation [7a].

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

152. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.b with respect to the limitation [7b].

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

153. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.e with respect to the limitation [7e].

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

154. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.f with respect to the limitation [7f].

- 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”**

155. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.g with respect to the limitation [7g].

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

156. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.h with respect to the limitation [7h].

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

157. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.b with respect to the limitation [7b].

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

158. In my opinion, Ikeda discloses that the main microphone and its corresponding main signal V_m are distinct from the reference microphone and its corresponding reference signal V_r for substantially the same reasons discussed in Sections VII.B.6.a and VII.B.6.b with respect to the limitations [7a] and [7b].

Ikeda further discloses that main signal V_m and reference signal V_r are distinct because they are formed with different combinations of the signals from the microphones M1 and M2, including a delay on the second microphone signal for V_m and on the first microphone signal for V_r . Ex. 1006, ¶¶ [0002]-[0004], [0018]-[0024], Figs. 1-5; Ex. 1001, 4:48-51 (“The term ‘virtual microphones (VM)’ or ‘virtual directional microphones’ means a microphone constructed using two or more omnidirectional microphones and associated signal processing.”), 7:32-35, 7:66-8:2; *see also* §§ VII.B.6.a, VII.B.6.b.

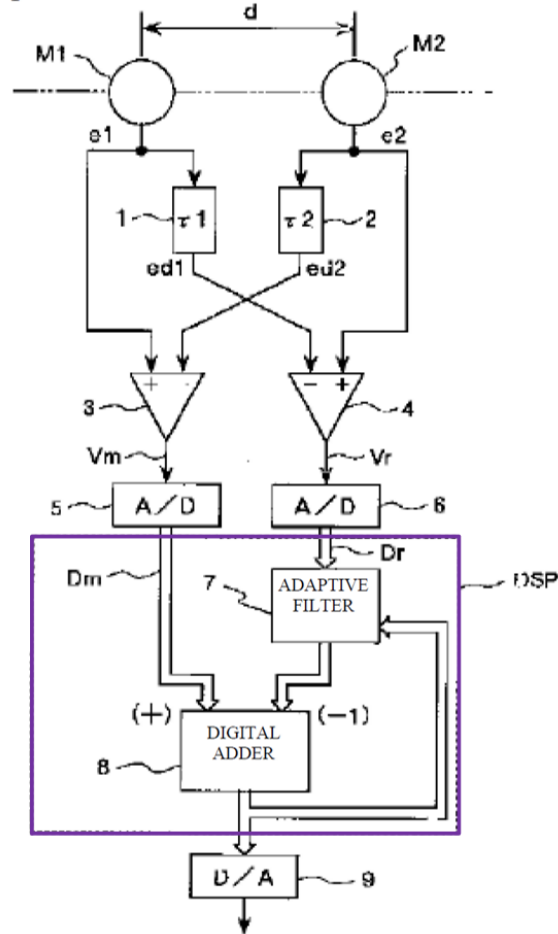
159. Additionally, in Ikeda-McCowan-Kanamori, main signal V_m and reference signal V_r are virtual microphones for substantially the same reasons discussed in Sections VII.B.6.a and VII.B.6.b with respect to the limitations [7a] and [7b]. The Ikeda-McCowan-Kanamori main signal V_m and reference signal V_r are distinct because they are formed with different combinations of the signals from the microphones M1 and M2, including a gain applied to the second microphone signal for V_m and applied to the first microphone signal for V_r . §§ VII.B.4, VII.B.6.a, VII.B.6.b, VII.B.6.c, VII.B.6.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,”

160. Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.j with respect to the limitation [7j].

161. 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]-[0029], [0030] (“the adaptive filter 7, the digital adder 8, and the like can be configured a digital signal processor DSP”); § VII.B.6.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 the A/D converters 5 and 6. Ex. 1006, ¶ [0027], Fig. 1; Ex. 1001, 32:25-32.

FIG. 1



Ex. 1006, Fig. 1 (annotated).

162. Ikeda discloses or renders obvious the “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,” for substantially the same reasons discussed in Section VII.B.6.j with respect to the limitation [7j].

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

163. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Sections VII.B.6.j and VII.B.6.k with respect to the limitations [7j] and [7k].

13. 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.”

164. Claim 16 recites: “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.” Ex. 1001, Claim 16.

165. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.9 with respect to claim 11.

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

166. Claim 17 recites: “The system of claim 16 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.” Ex. 1001, Claim 17.

167. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.10 with respect to claim 12.

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

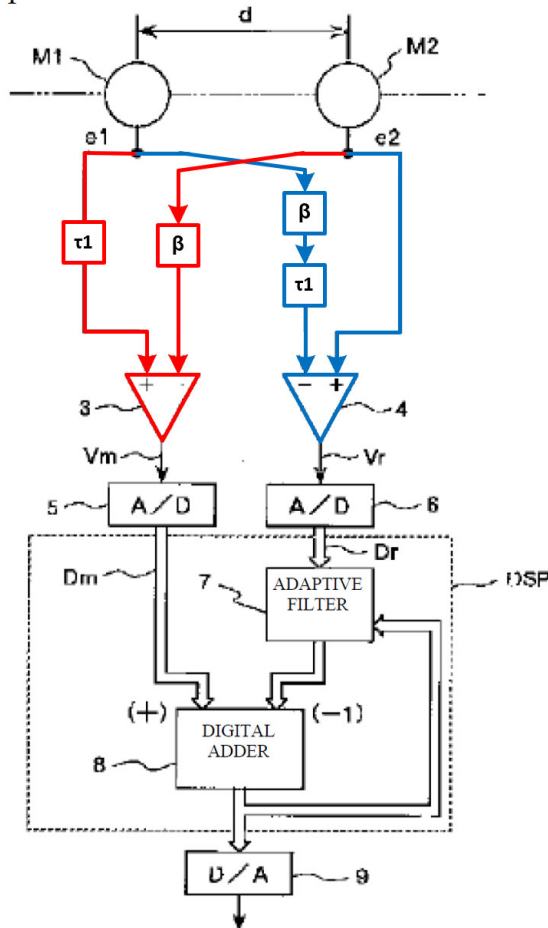
168. Claim 18 recites: “The system of claim 14, wherein the acoustic signals from the first virtual microphone, the second virtual microphone or both are delayed.” Ex. 1001, Claim 18.

169. In my opinion, Ikeda discloses this feature. 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 the adaptive filter 7 and the digital adder 8. Ex. 1006, ¶¶ [0027]-[0029]; *see also* Ex. 1010, 3:32-36 (“The delay circuit 14 is provided to compensate a time delay required by the adaptive filter circuit 24 for adaptive processing and a propagation time in the filter.”), Fig. 2; Ex. 1008, ¶ [0112].

170. I have been asked to assume 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). In my

opinion, Ikeda in view of McCowan and Kanamori renders this claim obvious under this interpretation. In the Ikeda-McCowan-Kanamori combination, the output signal of the microphone M1 (“the first microphone signal”) is delayed by τ_1 , as shown below. Ex. 1006, ¶ [0018], Fig. 1.

FIG. 1



Ex. 1006, Fig. 1 (modified).

16. 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.”

171. Claim 19 recites: “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.” Ex. 1001, Claim 19.

172. I have been asked to assume for the purposes of this declaration that “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). In my opinion, Ikeda in view of McCowan and Kanamori renders claim 19 obvious under this interpretation.

173. 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.

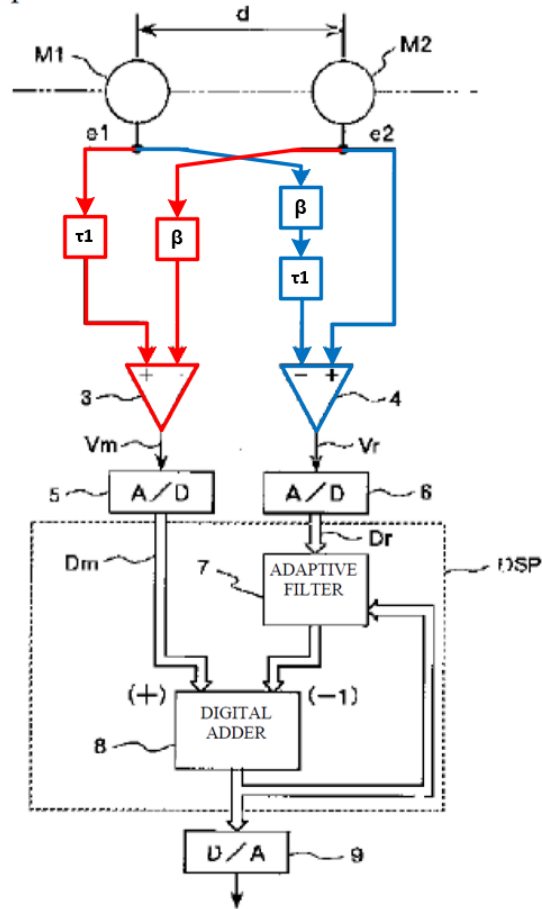
174. Based on the '080 patent's reference to γ as “the time difference between arrival of speech at V_1 compared to V_2 ” (Ex. 1001, 13:66-14:2), which I adopt for the purpose of this declaration, Ikeda in view of McCowan and Kanamori renders claim 19 obvious.

175. Ikeda discloses that τ_1 is equal to the distance between the microphones M1 and M2 divided by the speed of sound. Ex. 1006, ¶¶ [0018] (“microphone units 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; *see also* § VII.B.4.

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

■
Near-Field
Speech Source

FIG. 1



Ex. 1006, Fig. 1 (modified).

177. Therefore, τ_1 is a time difference between arrival of the speech at the microphone M1 and arrival of the speech at the microphone M2. A POSITA would have understood that, in an equation representing either the main signal V_m or the reference signal V_r , the delay of the output signal of the 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.

178. 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]. With a digital design, the delay of the output signal of the 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 the microphone M1 into a digital signal) multiplied by the time length τ_1 of the delay, e.g. $z^{-\tau_1 \cdot f_s}$. Ex. 1008, ¶ [0108], Fig. 5 (disclosing that z^{-1} refers to delaying by one sample); *see also*, § VII.B.5. Therefore, the delay of the output signal of the microphone M1 in the digital design is “raised to a power that is proportional to” τ_1 , which is the time difference between arrival of the speech at the microphone M1 and arrival of the speech at the microphone M2, because in an equation representing either the main signal V_m or the reference signal V_r , the delay z is raised to a power that is proportional to τ_1 .

- 17. 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.”**

179. Claim 20 recites: “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.” Ex. 1001, Claim 20.

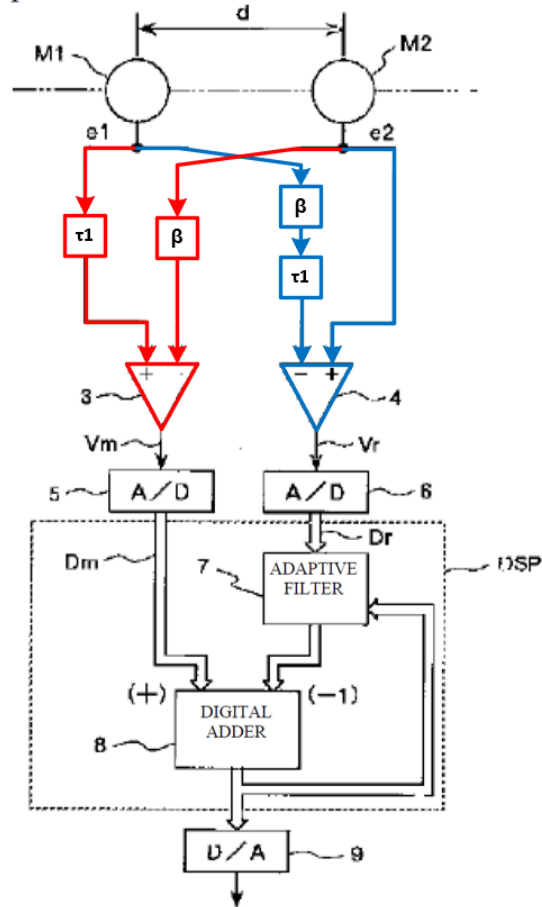
180. I have been asked to assume for the purposes of this declaration that “the acoustic signals from the first virtual microphone, the second virtual microphone or both” as recited in claim 18 is interpreted to include “the first microphone signal,” “the second microphone signal,” or both (Ex. 1023, 97-106; Ex. 1001, 7:17-8:12, 9:46-10:2, 11:1-2, Figs. 3-5). In my opinion, Ikeda in view of McCowan and Kanamori renders claim 20 obvious under this interpretation.

181. As discussed in Section VII.B.16 with respect to claim 19, with a digital design, the delay of the output signal of the 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 the microphone M1 into a digital signal) multiplied by the time length τ_1 of the delay, which is the distance between the microphones M1 and M2 divided by the speed of sound. Because the speech source is located on the axis of the microphones M1 and M2 as shown below (*see* § VII.B.16), the distance between the microphones M1 and M2 is a distance

between the microphone M1 and the speech source subtracted from a distance
between the microphone M2 and the speech source.

■
Near-Field
Speech Source

FIG. 1



Ex. 1006, Fig. 1 (modified).

182. Therefore, in the digital design, the delay of the output signal of the microphone M1 is raised to the power (an exponent in an equation representing either the main signal Vm or the reference signal Vr) that 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 (the microphone M1) and the speech source, the fourth distance being between a second physical microphone (the microphone M2) and the speech source. Ex. 1006, ¶¶ [0018]-[0019], Fig. 1.

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

18. 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;”**

184. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Sections VII.B.6.a and VII.B.6.d with respect to the limitations [7a] and [7d].

- 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,”**

185. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature. 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]. A POSITA would have understood that such a processor is coupled to the microphone array of the microphones M1 and M2 and generates main signal Vm and reference signal Vr. Ex. 1006, ¶¶ [0018]-[0024], Fig. 1. 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 VII.B.6.a and VII.B.6.b with respect to the limitations [7a] and [7b].

- 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,”**

186. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Sections VII.B.6.a and VII.B.6.b with respect to the limitations [7a] and [7b].

- 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”**

187. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Sections

VII.B.6.c, VII.B.6.e, VII.B.6.f, and VII.B.6.h with respect to the limitations [7c], [7e], [7f], and [7h].

- 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,”**

188. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.j with respect to the limitation [7j].

189. A POSITA would have understood that the Ikeda adaptive noise cancellation system is coupled to the processing component that generates the main and reference microphones. Ex. 1006, ¶¶ [0027]-[0029], Fig. 1; § VII.B.18.b.

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

190. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.6.k with respect to the limitation [7k].

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

191. Claim 2 recites: “The system of claim 1, wherein the acoustic noise comprises noise content and the acoustic signals comprise speech content.” Ex. 1001, Claim 2.

192. In my opinion, Ikeda discloses that the acoustic noise comprises noise content. Ex. 1006, ¶¶ [0018]-[0024], [0027]-[0029], [0036]. For example, Ikeda discloses noise “generated from every direction.” Ex. 1006, ¶ [0012].

193. In my opinion, Ikeda discloses or renders obvious that the acoustic signals comprise speech content. Ex. 1006, ¶ [0036] (“The present invention can be effectively applied to a wide range of devices ... such as *a voice input microphone of a hands-free telephone unit* or a car navigation system.”) (emphasis added);
§ VII.B.6.i.

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

194. Claim 3 recites: “The system of claim 2, wherein the speech content comprises human speech.” Ex. 1001, Claim 3.

195. In my opinion, Ikeda discloses or renders obvious that the speech content comprises human speech. Ex. 1006, ¶ [0036] (“The present invention can be effectively applied to a wide range of devices ... such as *a voice input*

microphone of a hands-free telephone unit or a car navigation system.”) (emphasis added); § VII.B.6.i.

21. 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.”

196. Claim 5 recites: “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.” Ex. 1001, Claim 5.

197. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.9 with respect to claim 11. 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.

§ VII.B.18.b; Ex. 1006, ¶¶ [0016]-[0024], [0030], Fig. 1; Ex. 1008, ¶ [0151]; Ex. 1014, ¶¶ [0021], [0026], [0065].

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

198. Claim 6 recites: “The system of claim 5 and further comprising: a communication device wirelessly coupled with the wireless channel of the communications channel.” Ex. 1001, Claim 6.

199. In my opinion, Ikeda in view of McCowan and Kanamori renders obvious this feature for substantially the same reasons discussed in Section VII.B.10 with respect to claim 12.

C. Ground 2: Claims 4, 10, and 15 are rendered obvious over Ikeda in view of McCowan, Kanamori, and Yang

200. In my opinion, claims 4, 10, and 15 are rendered obvious over Ikeda in view of McCowan, Kanamori, and Yang.

1. Overview of Yang

201. 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].

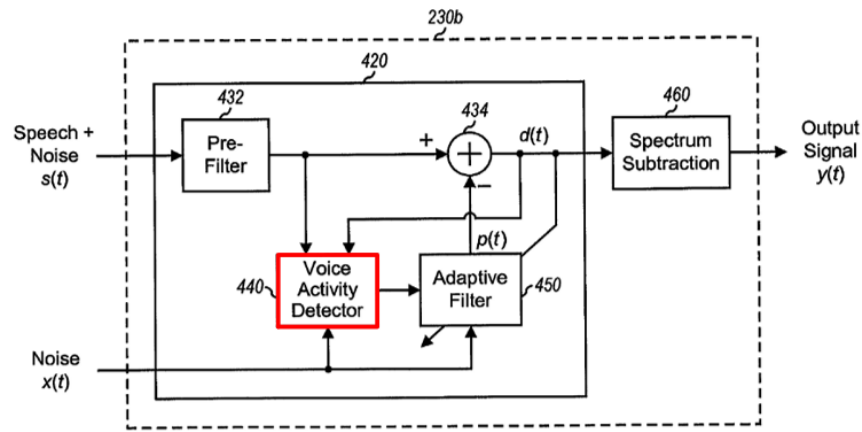


FIG. 4

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.”**

202. Claim 10 recites: “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.” Ex. 1001, Claim 10.

203. In my opinion, 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”).

204. Ikeda states: “The adaptive filter 7 is a filter that adaptively varies its characteristics with respect to changes in the input signal over time. ... The

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*. Therefore, the filter response characteristics of the adaptive filter 7 are determined so as to cancel the respective noise components contained in the main signal data Dm and the reference signal data Dr at the training sequence timing.” Ex. 1006, ¶ [0028] (emphasis added).

205. In my opinion, it would have been obvious to a POSITA 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] (“The present invention can be effectively applied to a wide range of devices ... such as a *voice input microphone* of a hands-free telephone unit or a car navigation system.”) (emphasis added); Ex. 1009, ¶¶ [0058]-[0062], Fig. 4. 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]. The modification would have been nothing more than combining prior art elements according to known methods to yield predictable results. 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.

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

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.”

207. Claim 15 recites: “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.” Ex. 1001, Claim 15.

208. In my opinion, Ikeda in view of McCowan, Kanamori, and Yang renders obvious this feature for substantially the same reasons discussed in Section VII.C.2 with respect to claim 10.

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.”

209. Claim 4 recites: “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.” Ex. 1001, Claim 4.

210. In my opinion, Ikeda in view of McCowan, Kanamori, and Yang renders obvious this feature for substantially the same reasons discussed in Section VII.C.2 with respect to claim 10. 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. § VII.B.18.b; Ex. 1009, ¶¶ [0058]-[0062], Fig. 4. 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 $x(t)$).

VIII. CONCLUSION

211. For the reasons set forth above, it is my opinion that claims 1-20 of the '080 patent are unpatentable in view of the prior art.

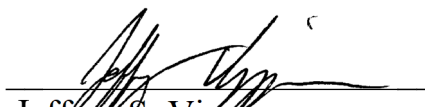
* * *

212. In signing this declaration, I understand that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I acknowledge that I may be subject to cross-examination in the case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

213. I declare that all statements made herein of my knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. § 1001.

214. I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Executed on May 25, 2022.

By: 
Jeffrey S. Vipperman