UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,
Petitioner,

v.

SOUNDCLEAR TECHNOLOGIES, LLC,
Patent Owner.

Case No. IPR2025-00345
U.S. Patent No. 9,031,259

DECLARATION OF STUART J. LIPOFF

TABLE OF CONTENTS

I. INTRODUCTION

1. I have been retained by Google LLC ("Petitioner") as an independent expert consultant in this proceeding before the Patent Trial and Appeal Board of the United States Patent and Trademark Office.

2. I am over 18 years of age and, if I am called upon to do so, I would be willing and able to testify as to the matters set forth herein.

3. My compensation is in no way contingent on the nature of my findings, the presentation of my findings in testimony, or the outcome of any related proceeding.

4. I understand that this proceeding involves U.S. Patent No. 9,031,259 ("the '259 Patent") (Ex. 1001). The application for the '259 Patent was filed September 14 2012, as U.S. Patent Application No. 13/617,143.

5. I have been asked by Petitioner to provide my opinion on whether the claims of the '259 Patent would have been anticipated or obvious to a person of ordinary skill in the art ("POSITA") at the time of the earliest claimed priority date of the '259 Patent. In performing my analysis, I have been asked to assume that the priority date is September 15, 2011. My opinions are set forth below.

6. Throughout this declaration, I refer to specific pages, figures, or line numbers of various exhibits. These citations are illustrative and are not intended to suggest that they are the only support for the propositions for which they are cited.

II. SUMMARY OF OPINIONS

7. This declaration considers claims 1-20 of the ’259 Patent. Below, I set forth the opinions I have formed, the conclusions I have reached, and the bases for these opinions and conclusions. I believe the statements contained in this declaration to be true and correct to the best of my knowledge.

8. Based on my experience and knowledge of the art at the time of the earliest claimed priority date of September 15, 2011, it is my opinion that claims 1-20 of the ’259 Patent would have been obvious based on the asserted grounds discussed below. A person of ordinary skill in the art (“POSITA”) would have found it obvious to combine the prior art references I cite and would have been motivated to do so before the priority date.

III. QUALIFICATIONS AND BACKGROUND

9. I believe that I am well qualified to serve as a technical expert in this matter based upon my qualifications, discussed in detail below.

10. I am currently president of IP Action Partners Inc., a consulting practice that serves the telecommunications, information technology, media, electronics, and e-business industries.

11. I earned a Bachelor of Science degree in Electrical Engineering in 1968 and a second Bachelor of Science degree in Engineering Physics in 1969, both from Lehigh University. I earned a Master of Science degree in Electrical

Engineering from Northeastern University in 1974, and then a Master of Business Administration degree from Suffolk University in 1983.

12. I hold a Federal Communications Commission ("FCC") General Radiotelephone License. I also hold a Certificate in Data Processing from the Institute for the Certification of Computing Professionals ("ICCP"), which is supported by the Association for Computing Machinery ("ACM").

13. I am also a registered professional engineer (PE) in the Commonwealth of Massachusetts and in the State of Nevada.

14. I am a fellow of the Institute of Electrical and Electronics Engineers ("IEEE") Consumer Electronics, Communications, Computer, Circuits, and Vehicular Technology Groups. I have been a member of the IEEE Consumer Electronics Society National Board of Governors (formerly known as the Administrative Committee) since 1981, and I was Boston Chapter Chairman of the IEEE Vehicular Technology Society from 1974 to 1976. I served as the 1996-1997 President of the IEEE Consumer Electronics Society, and from 1999 to 2018 I served as Chairman of the Society's Technical Activities and Standards Committee and as Vice President of Publications for the Society. From 2018 to 2023 I served as Vice President of Standards and Industry Activities for the Society and currently serve on the Board of Governors as The Historian for The Society. I have also

served as an Ibuka Award committee member for the IEEE's Award in the field of consumer electronics.

15. I have prepared and presented numerous papers at the IEEE and at other professional meetings. For example, in fall 2000, I served as general program chair for IEEE's Vehicular Technology Conference on advanced wireless communication technology. I have organized sessions at The International Conference on Consumer Electronics, and I was the 1984 program chairman. I conducted an eight-week IEEE-sponsored short course on Fiber Optics System Design. I received IEEE's Centennial Medal in 1984, and I received IEEE's Millennium Medal in 2000.

16. As Vice President and Standards Group Chairman for the Association of Computer Users ("ACU") from 1980 to 1983, I served as the ACU representative to the ANSI X3 Standards Group. From 1976 to 1978, I served as Chairman of the task group on user rule compliance for the FCC's Citizens Advisory Committee on Citizen's Band Radio.

17. Over the last 25 years, I have been a member of the Society of Cable Television Engineers, the Association for Computing Machinery, and The Society of Motion Picture and Television Engineers. From 2001 to 2004, I served as a member of the USA advisory board to the National Science Museum of Israel. In 1998, I presented a short course on international product development strategies as

a faculty member for Technion Institute of Management in Israel. From 2001 to 2003, I served as a member of the board or directors of The Massachusetts Future Problem Solving Program.

18. I am a named inventor on seven United States patents and have several publications on data communications in publications, including Electronics Design, Microwaves, EDN, the Proceedings of the Frequency Control Symposium, Optical Spectra, and IEEE publications.

19. During my professional career dating from 1969 to the present, I have been heavily engaged in the study, analysis, evaluation, design, and implementation of products and technology associated with consumer electronics and electronic appliances. A particular focus of my professional activities has been improving the man-machine interface including voice, speech, and speaker recognition for man-machine interactions. I also have extensive experience in studying foundation technologies and the applications supporting speech signal processing.

20. For approximately three years, from 1969 to 1972, I served as Project Engineer for Motorola's Communications Division, where I had project design responsibilities for paging and wireless communication products. Projects I worked on while employed at Motorola included work on paging systems that included digital voice storage, voice compression, and voice synthesis. I also worked on

projects that interfaced wireless data communications terminals to public safety computer systems for mobile data retrieval and data entry.

21. For approximately four years, from 1972 to 1976, I served as Section Manager for Bell & Howell Communications Company, where I also had project design responsibilities for paging and wireless communication products. The projects I supported included covert audio intelligence systems that recognized speech and activated digital voice compression recording systems. I also led projects for voice-based radio paging systems that recorded speech input, processed the speech to remove silence, processed the speech to digitally compress the speech, and store and forward the speech upon demand from DTMF or computer keyboard retrieval from the servers.

22. For 25 years from 1976 to 2001, I worked for Arthur D. Little, Inc. (ADL), where I became the Vice President and Director of Communications, Information Technology, and Electronics (CIE) and served in that role for 10 years, from 1991 to 2001. At ADL, I was responsible for the firm's global CIE practice in laboratory-based contract engineering, product development, and technology-based consulting. I was also involved in multiple pioneering efforts to identify and explore customer-to-business and business-to-business electronic commerce and transactions information processing opportunities (e-commerce). These projects involved technology assessment and analysis as well as developing

architectures and systems to support multiple applications, and typically involved an information retrieval component.

23. While at ADL, I worked on several projects involving the combination of voice interfaces (including speech recognition and voice audio output) and information retrieval. For example, over the course of three years in the early-1990s, I worked on a project for Bolt Beranek and Newman (BB&N), where I evaluated and benchmarked technology for a voice input/output application that allowed end users (*e.g.*, travel agents) to use speech inputs to interact with airline reservation databases to retrieve information about travel reservation options, where the results were returned to the user in an audible message. This system included a natural language front-end speech-interface module with speech recognition that used pre-defined recognition grammars to convert the end user's speech into structured commands supported by an airline reservation system. As another example, over the course of three years in the mid-1990s, I worked on a project for Texas Instruments that applied a speech-recognition interface for a variety of applications that retrieved information from database servers.

24. Other projects that I worked on at various points in my 25 years at ADL and afterwards that involved speech recognition technologies included the following.

25. Over the course of three years in the early 1990s, I worked on a voice-interface project developing spoken digit telephone number recognition and voiceprint matching for Sprint's long distance alternative access telephone services.

26. Over the course of a year in the late 1980s, I worked on a voice interface project evaluating the processing power needed to perform various voice recognition applications by Rockwell Semiconductor's signal processing technology.

27. Over the course of 15 years starting in the early 1980's, I worked on a project for the United States Postal Service (USPS), where we developed a real-time automated postal teller system that served as an interface between end-users and the USPS's information systems. This system included voice prompts for the vision impaired.

28. I also have extensive experience in public and private network wired and wireless voice telecommunications while employed by Motorola, Bell & Howell, and Arthur D Little, and while self-employed. In the course of these telecommunications projects ranging from 1969 to the present, I have encountered a number of applications where audio input and voice are used to activate devices, for example for the purpose of saving battery power by entering into low power, so-called "sleep" modes. These projects have involved the design of cellular

telecommunications systems that implement industry standard means of entering lower power modes in the absence of voice.

29. My curriculum vitae are provided as Exhibit 1004.

IV. MATERIALS CONSIDERED

30. In forming my opinions, I have reviewed the following documents, as well as other documents cited throughout this declaration:

Exhibit	Description
1001	U.S. Patent No. 9,031,259 to Yamabe ("the '259 patent")
1002	Prosecution History of U.S. Patent No. 9,031,259
1005	U.S. Patent Appl. Pub. No. 2011/0026730 to Li et al ("Li-730")
1006	U.S. Patent Appl. Pub. No. 2012/0027219 to Kale et al. ("Kale")
1007	U.S. Patent Appl. Pub. No. 2011/0135107 to Konchitsky et al. ("Konchitsky")
1008	U.S. Patent Appl. Pub. No. 2010/0111329 to Namba et al. ("Namba")
1009	U.S. Patent Appl. Pub. No. 2012/0197638 to Li et al. ("Li-638")
1010	U.S. Patent Appl. Publ. No. 2012/0128187 to Yamada et al. ("Yamada")
1011	U.S. Patent Appl. Pub. No. 2010/0303267 to Pedersen et al. ("Pedersen")
1012	U.S. Patent Appl. Pub. No. 2012/0022373 to Tateyama ("Tateyama")
1013	U.S. Patent No. 8,615,392 to Goodwin et al. ("Goodwin")
1014	*SoundClear Technologies LLC v. Google LLC*, Case No. 1:24-cv-00729, Complaint for Patent Infringement (E.D. VA May 1, 2024)

31. I have also relied on my education, experience, research, training, and knowledge in the relevant art, and my understanding of legal principles described in this declaration.

32. All of the opinions contained in this declaration are based on the documents I reviewed and my knowledge and professional judgment. My opinions have also been guided by my understanding of how a POSITA would have understood the claims of the ’259 Patent at the time of the earliest claimed priority date.

33. I reserve the right to supplement and amend any of my opinions in this declaration based on documents, testimony, and other information that becomes available to me after the date of this declaration.

V. LEGAL STANDARDS

34. I am an engineer and 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. I have applied these legal standards in preparing this declaration.

35. I have been informed that there are two ways in which prior art may render a patent claim unpatentable. First, I have been informed that the prior art

can "anticipate" a claim. Second, I have been informed that the prior art can render a claim "obvious" to a POSITA. I understand that a claim is patentable if it was not anticipated and would not have been rendered obvious by the prior art at the effective filing date of the patent.

36. I have been informed that a dependent claim is a patent claim that refers back to another patent claim. I have been informed that a dependent claim includes all of the limitations of the claim to which it refers plus its own limitation(s).

37. I have been asked to provide my opinions as to whether the cited prior art discloses or renders obvious claims 1-20 of the '259 Patent from the perspective of a POSITA at the time of the earliest claimed priority date of September 15, 2011, as described in more detail below.

38. I have been informed that in IPR proceedings, such as this one, the party challenging the patent bears the burden of proving unpatentability by a preponderance of the evidence. I understand that a preponderance of the evidence means "more likely than not."

39. For purposes of this declaration, I have been asked to provide my opinions on issues regarding unpatentability. I have been informed of the following legal standards, which I have applied in forming my opinions.

A. Level Of Ordinary Skill

40. I have been informed that a POSITA is determined by considering several factors, including the (i) 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.

41. I have been instructed to assume that a POSITA is not a specific real individual, but rather a hypothetical individual having the qualities reflected by the factors discussed above. A POSITA is assumed to be person of ordinary creativity familiar with the prior art as of the priority date of the patent at issue.

B. Prior Art

42. I have been advised and understand that the information used to evaluate whether an invention was new and not obvious when made is generally referred to as "prior art." I understand that in an IPR proceeding, prior art includes patents and printed publications that existed before the earliest claimed priority date or the earliest filing date of the patent (which I have been informed is also called the "effective filing date"). I have been informed and understand that a patent or published patent application is prior art if it was filed before the earliest filing date of the claimed invention and that a printed publication is prior art if it was publicly available before the earliest filing date.

C. Anticipation

43. I have been informed that under 35 U.S.C. § 102, a patent claim is unpatentable for anticipation if the claimed subject matter was patented or described in a printed publication before the effective filing date of the claimed invention. I have been informed that this is referred to as unpatentability by anticipation. I have been informed that a patent claim is anticipated under § 102 if a single prior art reference discloses all the limitations of the claimed invention. I understand that limitations may be expressed or inherent such that the limitation is essential to the prior art.

D. Obviousness

44. I have been informed that for obviousness under 35 U.S.C. § 103, a patent claim is unpatentable 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 to a POSITA to which said subject matter pertains at the time the invention was made. I have been informed that this is referred to as unpatentability by obviousness.

45. I have been informed that an obviousness analysis includes the following considerations:

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 nonobviousness (if available).

46. I have been informed that the relevant time for considering whether a claim would have been obvious to a POSITA is the time of invention. For my obviousness analysis, counsel for Petitioner instructed me to assume that the date of invention for the challenged claims is September 15, 2011. My opinions would not change if I assumed another, e.g., later, date of invention.

47. I have been informed that a reference may be modified or combined with other references or with a POSITA's own knowledge if the person would have found the modification or combination obvious. I have also been informed that a POSITA is presumed to know all the relevant prior art, and the obviousness analysis may take into account the inferences and creative steps that a POSITA would employ.

48. I have been informed that an obviousness determination must be made from the perspective of a POSITA. I have also been informed that there is no requirement that the prior art contain an express suggestion to combine known elements to achieve the claimed invention, and that a suggestion to combine known

elements to achieve the claimed invention may come from the prior art as a whole or individually. Also, the obviousness analysis may rely on the inferences and creative steps a POSITA would employ, as filtered through his or her knowledge as of the priority date. But I understand that obviousness grounds cannot be sustained by mere conclusory statements and must include some articulated reasoning and rationale to support a legal conclusion of obviousness.

49. In determining whether a prior art reference could have been combined with another prior art reference or other information known to a POSITA, I have been informed that the following principles may be considered:

a. A combination of familiar elements according to known methods is likely to be obvious if it yields predictable results;
b. The substitution of one known element for another is likely to be obvious if it yields predictable results;
c. The use of a known technique to improve similar items or methods in the same way is likely to be obvious if it yields predictable results;
d. The application of a known technique to a prior art reference that is ready for improvement to yield predictable results;

e. Any need or problem known in the field and addressed by the reference can provide a reason for combining the elements in the manner claimed;

f. A person of ordinary skill often will be able to fit the teachings of multiple references together like a puzzle; and

g. The proper analysis of obviousness requires a determination of whether a POSITA would have a "reasonable expectation of success"—but not "absolute predictability" of success—in achieving the claimed invention by combining prior art references.

50. I have been informed that, when a work is available in one field, design alternatives and other market forces can prompt variations of it, either in the same field or in another. I have been informed that if a POSITA could have implemented a predictable variation and would have seen the benefit of doing so, that variation is likely to have been obvious. I have been informed that, in many fields, such as the mechanical or electrical arts, market demand—not scientific literature—may drive design trends. I have been informed that, when there was a design need or market pressure and there are a finite number of predictable solutions, a POSITA would have had a good reason to pursue those known options.

51. I have been informed that the law permits the application of "common sense" in examining whether a claimed invention would have been obvious to a POSITA. For example, I have been informed that combining familiar elements according to known methods and in a predictable way may suggest obviousness when such a combination would yield nothing more than predictable results. I understand, however, that a claim is not obvious merely because every claim element is disclosed in the prior art. A party asserting obviousness must provide a specific motivation to combine or modify the references as recited in the claims and explain why one skilled in the art would have reasonably expected to succeed in doing so.

52. I have been informed 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 understand that the "teaching, suggestion, or motivation" test can be a useful guide in establishing a rationale for combining elements of the prior art. I have been informed that this test poses the question as to whether there is an express or implied teaching, suggestion, or motivation to combine prior art elements in a way that results in the claimed invention, and that it helps to counter the use of hindsight, which is impermissible. Likewise, if a prior art reference "teaches away" from a potential prior art combination, then a motivation to combine may not exist.

53. I am not aware of any evidence of secondary considerations, such as unexpected results, industry skepticism, long-felt unresolved need, commercial success, praise by others, or copying that would alter my opinions set forth below.

54. I have been informed that, in an obviousness analysis, prior art must be analogous art to the patent being considered. I have been informed that a prior art reference is considered to be analogous, or in the same field of art, if the reference is either (1) in the same field of endeavor as the challenged patent, regardless of the problems the challenged patent and the prior art address, or (2) reasonably pertinent to the particular problem being solved by the challenged patent.

VI. THE '259 PATENT (Ex. 1001)

55. The '259 patent was filed on September 14, 2012, issued on May 12, 2015, and claims priority to Japanese Patent Application Nos. JP 2011-201759 and JP 2011-201760, both filed September 15, 2011. Ex. 1001, cover page.

56. The '259 patent generally relates to "a noise reduction apparatus, an audio input apparatus, a wireless communication apparatus, and a noise reduction method." Ex. 1001, 1:17-19. The noise reduction apparatus includes a "main microphone 11, a sub-microphone 12, A/D converters 13 and 14, a speech segment determiner 15, a voice direction detector 16, an adaptive filter controller 17, and an adaptive filter 18," as illustrated in Figure 1 below. Ex. 1001, 3:55-59, FIG. 1.

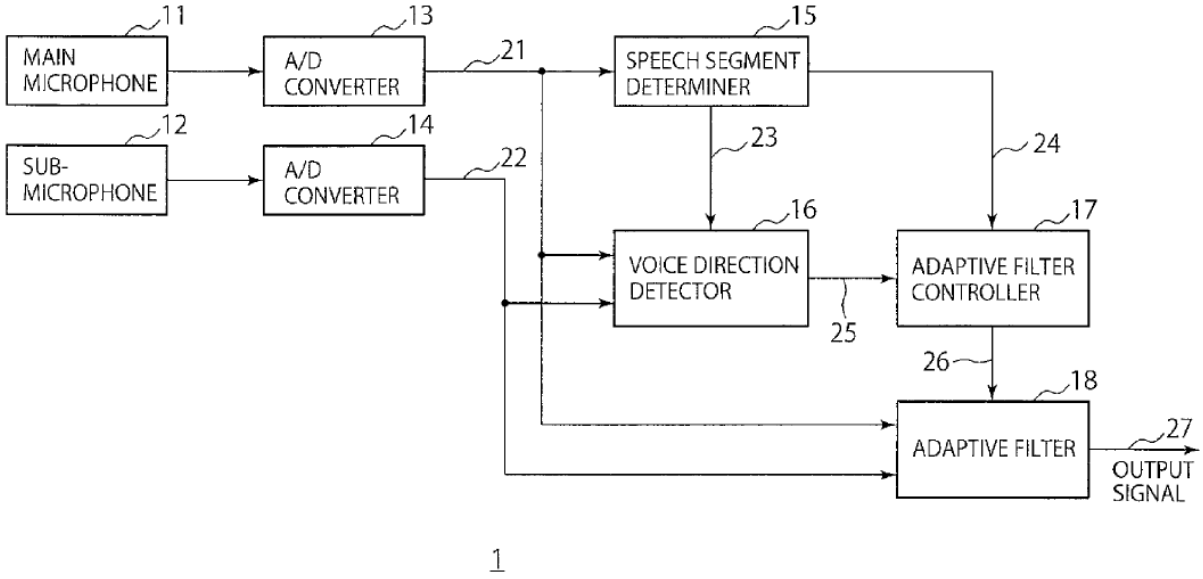


FIG. 1

Ex. 1001, FIG. 1.

57. Sounds received by the main microphone are converted into a first "sound pick-up signal" then input into "speech segment determiner." Ex. 1001, 4:11-20, FIG. 1. The speech segment determiner "determines whether or not [the] sound picked up the main microphone 11 is a speech segment (voice component)" and "outputs speech segment information 23 and 24 to the voice direction detector 16 and the adaptive filter controller 17, respectively." Ex. 1001, 4:40-49, FIG. 1. Sounds received by the sub-microphone are converted into a second "sound pick-up signal" then input into the "voice direction detector." Ex. 1001, 4:41-26, FIG. 1.

58. The voice direction detector determines "which direction a voice sound travels" using either a "phase difference"[1] or "magnitude"[2] based calculation, and "outputs voice incoming-direction information 25 to the adaptive filter controller 17," which controls the adaptive filter. 16:67-17:9. The adaptive filter may use the speech segment information, sound pick-up signals, and the "phase difference" (PD1) or "power difference" (magnitude) (PD2) calculated by the voice direction detector to perform noise reduction. Ex. 1001, 1:57-61, 22:38-50, 23:53-24:52, 25:15-34, 25:44-51, 25:62-26:9

59. For example, if "the phase of a voice component carried by a sound pick-up signal 21 obtained based on a sound picked up by the main microphone 11 is more advanced than the phase of a voice component carried by a sound pick-up signal 22 obtained based on a sound picked up by the sub-

[1] "Phase difference" relates to the "amount of delay" between first and second sound signals. Ex. 1001, 18:41-45.

[2] "[M]agnitude is the maximum amplitude, an integral value of amplitude of the Sound pick-up signals 21 and 22, etc." Ex. 1001, 25:10-12. "The difference and the ratio between the magnitudes of sounds are referred to as a power difference and a power ratio, respectively. Both factors are referred to as power information." Ex. 1001, 17:6-10.

microphone 12" (PD1≧T), "the adaptive filter 18 performs the regular noise reduction process to reduce a noise component carried by the sound pick-up signal (a voice signal) 21 using the sound pick-up signal (a noise-dominated signal) 22 to produce the output signal 27." Ex. 1001, 23:59-24:5. Similarly, if "the magnitude of the sound pick-up signal 21 obtained based on a sound picked up by the main microphone 11 is larger than the magnitude of the sound pick up signal 22 obtained based on a sound picked up by the sub-microphone 12" (PD2≧P), the noise reduction apparatus "reduce[s] a noise component carried by the sound pick-up signal (a voice signal) 21 using the sound pick-up signal (a noise-dominated signal) 22 to produce the output signal." Ex. 1001, 25:19-34.

60. Conversely, if PD1≤-T or PD2≤-P, "the adaptive filter controller 17 treats the sound pick-up signal (a noise-dominated signal) 22 as a voice signal while treats the Sound pick-up signal (a voice signal) 21 as a noise-dominated signal. Then, the adaptive filter controller 17 controls the adaptive filter 18 to reduce a noise component carried by the Sound pick-up signal 22 (a voice signal) 22 using the Sound pick-up signal 21 (a noise-dominated signal) 22 to produce the output signal 27." Ex. 1001, 24:19-27, 25:44-51. Finally, if "the distance from a sound source to the main microphone 11 and the distance from the sound source to the sub-microphone 12 are almost the same as each other" (-T<PD1<T or -P<PD2<P) "the adaptive filter controller 17 controls the adaptive filter 18 to output

either the sound pick-up signal 21 or the Sound pick-up signal 22 as the output signal 27, with no noise reduction process." Ex. 1001, 24:38-52, 25:62-26:9.

VII. LEVEL OF ORDINARY SKILL IN THE ART

61. In my opinion, a person of ordinary skill in the art ("POSITA") in the field of the '259 patent would have had at least a bachelor's degree in electrical engineering or related discipline, and at least two years of experience in voice or audio processing. Relevant work experience can substitute for formal education and additional education could substitute for work experience.

VIII. CLAIM CONSTRUCTION

62. I understand that, in this proceeding, the claims are construed according to their ordinary and customary meaning, in light of the specification and prosecution history, as understood by a POSITA at the time of the invention. I understand that this is the same claim construction approach used in district court litigation.

63. In my opinion none of the claim terms in the '259 require an explicit construction.

IX. THE PRIOR ART

64. As explained in Section X below, it is my opinion that claims 1-20 would have been obvious over combinations of Li-730, Kale, Namba, Li-638, and Yamada, each of which I have been instructed to assume is prior art.

A. Li-730 (Ex. 1005)

65. Li-730 was filed on July 28, 2009, and published on February 3, 2011. Li-730, cover page. I understand that Li-730 is prior art to the ’259 patent.

66. Li-730 discloses an “audio processing apparatus” for “suppressing stationary and non-stationary noises.” Li-730, [0019]. The apparatus includes two microphones (a main and a reference microphone), a voice active detector, a beamformer, and a noise suppressor. Li-730, [0019], FIG. 2A.

67. As illustrated in Figure 2A below, the main and reference microphones receive a “mixture of sounds”—from a “source” (e.g., a “speaker’s mouth”) and a “non-source” (e.g., “noises”) and “convert the mixture of the sounds and the noises [] into a main input M1 and a reference input M2.” Li-730, [0020]. The converted inputs (M1, M2) are transformed into frequency domain representations (S1, S2), calibrated (into signals C1 and C2), and input into the “voice active detector” (VAD). Li-730 [0021]-[0022], [0025], FIG. 2A. The VAD uses the calibrated signals “to generate a voice active signal V1.” Li-730, [0020], [0025], FIG. 2A.

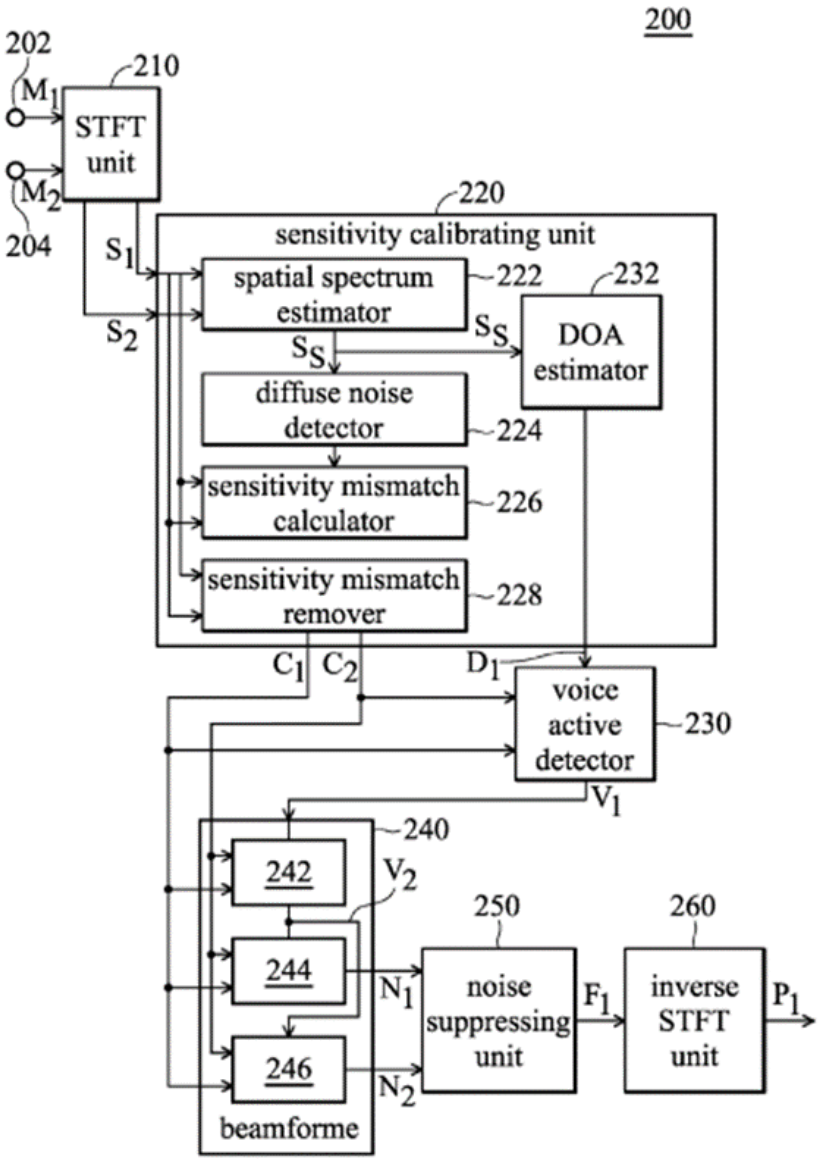


FIG. 2A

Li-730, FIG. 2A.

68. Li-730's beamformer uses the voice active signal V1 to generate and output "steering vector signal V2" which "indicates [the] directional vector" of the source signal (voice signal) when voice is detected—that is, when the "voice active

signal" is "turned on." Li-730, [0026], [0031], claim 13, FIGS. 2A, 3C. The beamformer also generates "main" and "reference" channels N1 and N2 ("corresponding to the sounds received from the source" and "the noises received from non-source," respectively) and outputs the signals to the noise suppressing unit. Li-730, [0026]-[0027], FIG. 2A, 3D. Finally, the noise suppressing unit "suppress[es] stationary and non-stationary noises in the main channel N1 and the reference channel N2 according to the voice active signal V1, and integrate[s] the main channel N1 and the reference channel N2 into a final signal F1." Li-730, [0027], FIGS. 2A, 3A. Figure 3A below illustrates Li-730's noise suppression method.

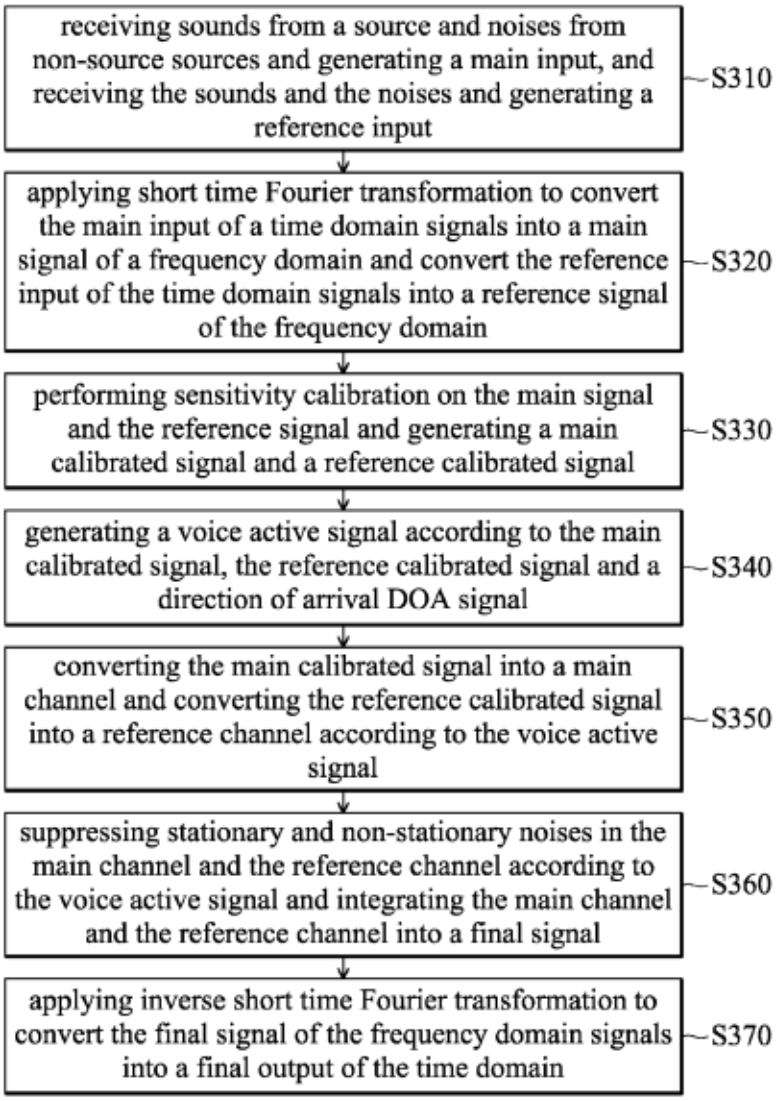


FIG. 3A

Li-730, FIG. 3A.

69. In my opinion, Li-730 is in the same field of endeavor as the ’259 patent: noise reduction. *Compare* Li-730, [0002], *with* Ex. 1001, 1:17-19. It is also my opinion that Li is directed to the same problem as the ’259 patent: improving noise reduction methods to enhance voice sound quality. *Compare* Li-730, [0006],

[0028], *with* Ex. 1001, 1:38-42. Li-730 at least suggests each element of the independent claims because it discloses each of features discussed in Ground 5 below.

B. Kale (Ex. 1006)

70. Kale was filed on July 28, 2010, and published on February 2, 2012. Kale, cover page. I understand that Kale is prior art to the ’259 patent.

71. Kale discloses a “noise cancellation device” and method. Kale, Abstract. The device includes *inter alia* two microphones, computation/FFT modules, a direction of arrival (DOA) module, and a beamformer. Kale, [0028], FIG. 8[3].

[3] In Figure 8, element 830 is erroneously marked “DOA” and should read “IFFT module” or “IFFT.” Kale, [0028].

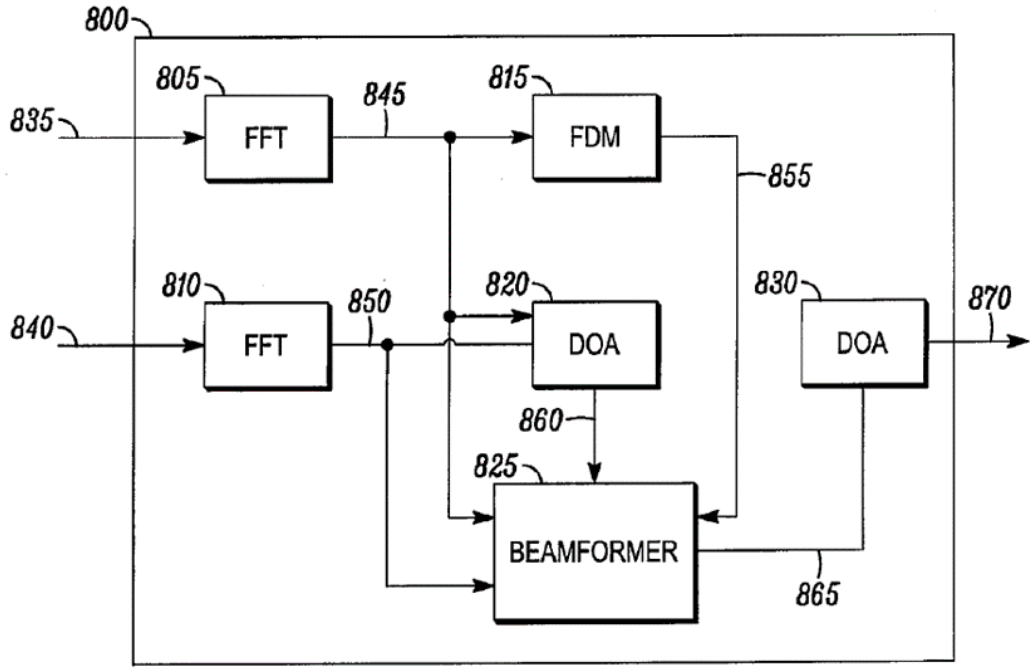


FIG. 8

Kale, FIG. 8.

72. Kale's microphones detect "two [sound] sources," (e.g., "speech" and "noise") and "generate respective raw audio data 835, 845." Kale, [0028] [0043], FIGS. 8-9. The computation/FFT modules "determine which of the raw audio data signals comprises the speech signal" using "factors … such as common formants, formants with known patterns, etc." and output a "speech signal" 845, 850 to the DOA module and beamformer. Kale, [0003], [0043]-[0044], FIG. 8. The DOA module "determin[es] the direction from which the speech data arrived" and outputs a "DOA data signal" 860 containing the direction information to the beamformer. Kale, [0028], [0031], [0046], FIGS. 8-9. Finally, the beamformer

uses the inputted speech signal and direction information to "modify [the] audio signal to isolate a speech therein" or "enhance a speech of an audio signal" by "eliminating" the noise contained in the received audio signal." Kale, [0040], [0047]-[0048], FIGS. 8-9.

73. In my opinion, Kale is in the same field of endeavor as the '259 patent: audio signal processing. *Compare* Kale, Abstract, *with* Ex. 1001, 1:17-19. It is also my opinion that Kale is directed to the same problem as the '259 patent: improving noise reduction methods to enhance voice sound quality. *Compare* Kale, [0016], [0022], *with* Ex. 1001, 1:38-42. Kale at least suggests each element of the independent claims because it discloses each of features discussed in Ground 1 below.

C. Namba (Ex. 1008)

74. Namba was filed on November 3, 2009, and published on May 6, 2010. Namba, cover page. I understand that Namba is prior art to the '259 patent.

75. Namba teaches a sound processing apparatus including, like Kale, a microphone array for receiving a mixture of "noise sound" and "voice," a unit for identifying a voice sound, a distance/direction estimator for determining the direction of the voice, and a unit for suppressing noise in the voice signal. Namba, [0020]-[0023], [0070], table 1. The distance/direction estimator may receive the "volume value sequence of the input sound, spectrum of the input sound and the

like" and "has a function, based on the input, as a positional information calculation unit that estimates the sound source of the input sound or positional information such as direction information and distance information of the sound source from which a dominant sound contained in the input sound originates." Namba, [0070].

76. In my opinion, Namba is in the same field of endeavor as the '259 patent: noise reduction. *Compare* Namba, [0002], [0005]-[0006], *with* Ex. 1001, 1:17-19. It is also my opinion that Namba is directed to the same problem as the '259 patent: improving noise reduction methods to enhance voice sound quality. *Compare* Namba, [0008], *with* Ex. 1001, 1:38-42.

D. Li-638 (Ex. 1009)

77. Li-638 was filed on Dec. 15, 2010, and published August 2, 2012. Li-638, cover page. I understand that Li-638 is prior art to the '259 Patent.

78. In my opinion, Li-638 is in the same field of endeavor as the '259 patent: noise reduction. *Compare* Li-638, [0001] *with* Ex. 1001, 1:17-19. It is also my opinion that Li-638 is directed to the same problem as the '259 patent: improving noise reduction methods to enhance voice sound quality. *Compare* Li-638 [0010], *with* Ex. 1001, 1:38-42.

E. Yamada (Ex. 1010)

79. Yamada was filed on June 16, 2011, and published on May 24, 2012. Yamada, cover page. I understand that Yamada is prior art to the ’259 Patent.

80. Yamada teaches a signal processing method wherein “sound picked up by a microphone” is supplied to a “sound source direction estimation section” “at a sampling frequency of 48 kHz.” Yamada, [0067]-[0069]. In my opinion Yamada is in the same field of endeavor as the ’259 patent: noise reduction. *Compare* Yamada, [0001], *with* Ex. 1001, 1:17-19. Furthermore, Yamada is directed to the same problem as the ’259 patent: improving noise reduction methods to enhance voice sound quality. *Compare* Yamada, [0011], [0013], *with* Ex. 1001, 1:38-42.

X. GROUNDS OF UNPATENTABILITY

81. Based on my review of the materials set forth above, including my application of the knowledge of a POSITA, it is my opinion that claims 1-20 of the ’259 Patent would have been anticipated or obvious to one of ordinary skill in the art as of September 15, 2011.

82. In particular, it is my opinion that claims 1-20 of the ’259 Patent would have been obvious to a POSITA based on the following grounds:

Grounds	Claims	Basis	Prior Art
1	1-3, 5, 6, 8, 13-15, and 17-19	§103	Kale
2	2, 3, 5, 6, 8, 9, 14, 15, and 17-20	§ 103	Kale and Namba
3	2-4, 7, and 14-16	§ 103	Kale and Li-638
4	10-11	§ 103	Kale and Yamada
5	1-3 and 12-15	§103	Li-730

A. Ground 1: Claims 1-3, 5, 6, 8, 13-15, and 17-19 Are Obvious over Kale

1. Independent claim 1

a. [1preamble]: "A noise reduction apparatus comprising:"

83. Kale discloses [1preamble] because it discloses a "noise cancellation device" Kale, [0003], [0028], FIG. 8. Figure 8 below illustrates the noise cancellation device including *inter alia* a "beamformer" configured to "modify an audio signal to isolate a speech therein" or "enhance a speech of an audio signal" by "eliminating" the noise contained in the received audio signal. Kale, Abstract, [0003], [0023], [0034], [0047], FIG. 8.

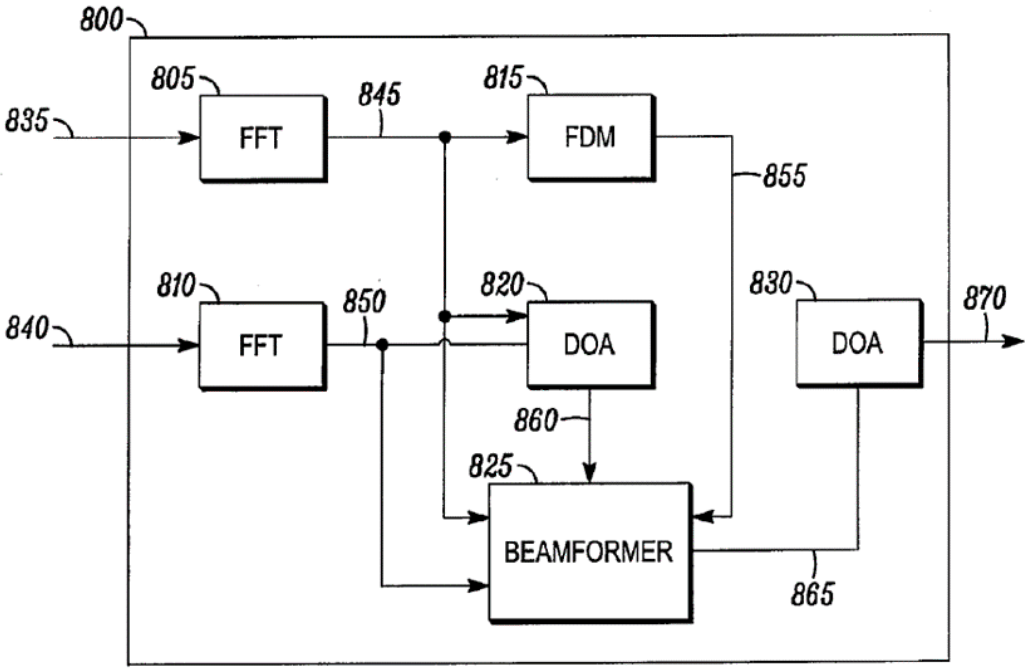


FIG. 8

Kale, FIG. 8.

b. [1a]: "a speech segment determiner configured to determine whether or not a sound picked up by at least either a first microphone or a second microphone is a speech segment and"

84. Kale discloses [1a] because the noise cancellation device includes computation modules (a speech segment determiner), e.g., Fast Fourier Transform (FFT) modules 805, 810 for receiving the raw audio data and generating "speech

data" signals (845, 850) there from. Kale, [0003], [0029]-[0031], FIG. 8, claims 1 and 3[4].

85. ***Kale discloses at least either a first or second microphone for picking up sound as recited in [1a].*** As illustrated in annotated Figure 8 below, Kale's noise cancellation device includes "two microphones" for detecting "two [sound] sources" (e.g., "speech" and "noise") and "generat[ing] respective raw audio data 835, 845." Kale, [0028] [0043], FIGS. 8-9.

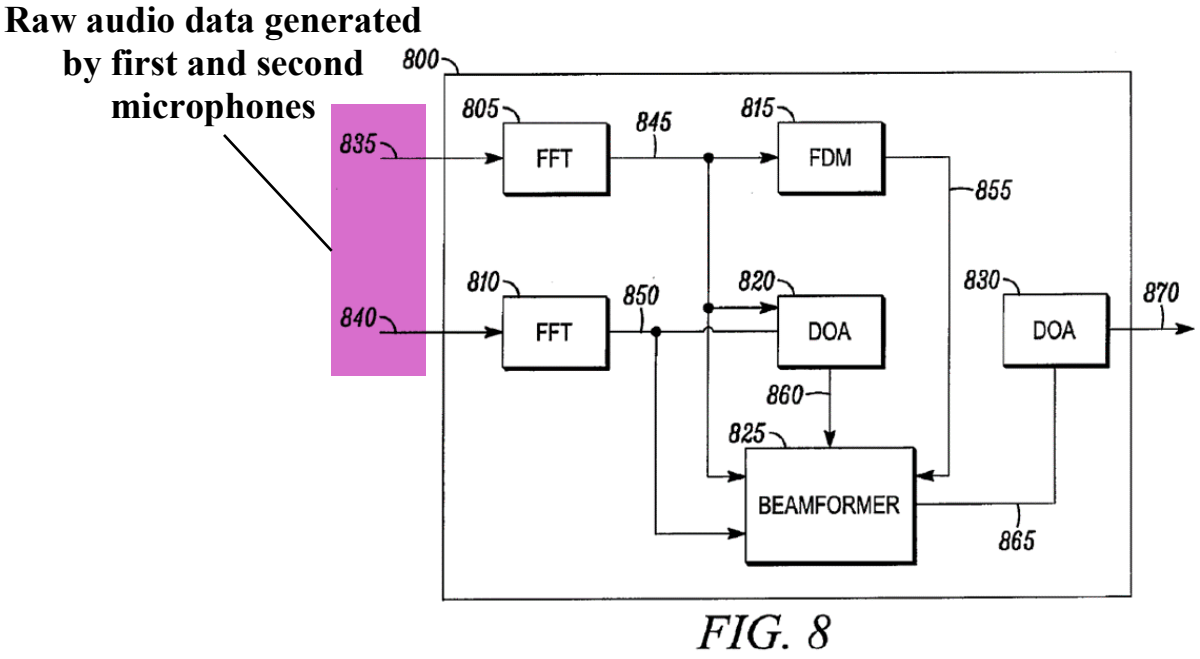


FIG. 8

Kale, FIG. 8 (annotated)

[4] Published claims 1 and 3 were first presented in the as-filed application on July 28, 2010.

86. ***Kale further discloses a computation module (speech segment determiner) for "determin[ing] which of the raw audio data signals comprises the speech signal." Kale, [0003], [0029]-[0031], [0044], FIG. 8, claims 1 and*** 3.

87. The computation module, which may be an FFT module "determine[s] which of the raw audio data signals comprises the speech signal" using "factors … such as common formants, formants with known patterns, etc." Kale, [0043]-[0044]; *see also* [0003] ("A first transformed signal relates to speech data"), claim 1 ("first computation modules receiv[e] raw audio data and generat[e] a respective transformed signal as a function of formants, a first transformed signal relating to speech data), claim 3 ("wherein the first computation modules are Fast Fourier Transform (FFT) modules."). This is similar to the speech segment determiner of the '259 patent which also uses formants to detect speech segments. Ex. 1001, 5:33-36 ("[s]pectra (referred to as formant, hereinafter) represent the feature of a voice and are to be detected in determining speech segments by the speech determination unit 38.").

88. Kale further discloses that "[t]he exemplary embodiments estimate[] formant position and/or maximum speech energy regions in real time using formant tracking algorithms such as Linear Predictive Coding (LPC), Hidden Markov Model (HMM), etc.," all of which are known speech detection techniques. Kale, [0024]. Thus, Kale's computation/FFT module is a speech segment

determiner configured to determine whether or not a sound picked up (raw audio data 835) by at least either a first microphone or a second microphone is a speech segment, as illustrated in annotated Figure 8 below:

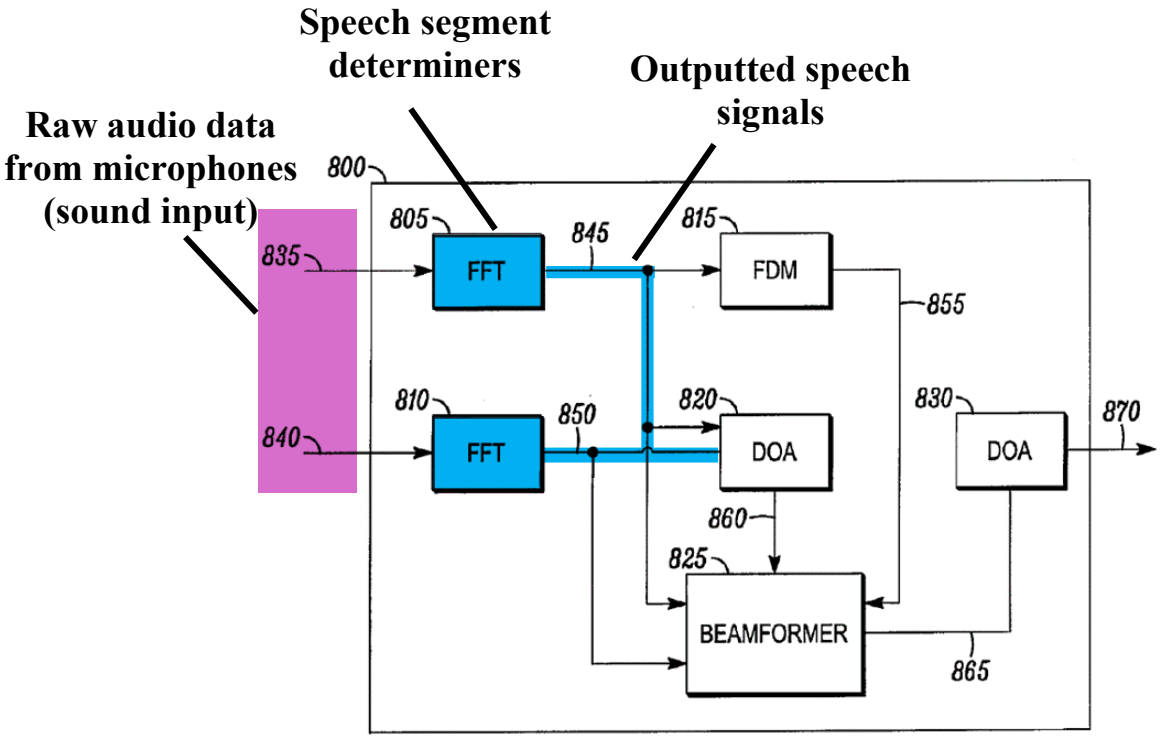


FIG. 8

Kale, FIG. 8 (annotated).

c. [1b]: "to output speech segment information when it is determined that the sound picked up by the first or the second microphone is the speech segment;"

89. Kale discloses [1b] because as explained for [1a], the FFT module (speech segment determiner) outputs "speech data" signals 845, 850 (speech segment information) "[u]pon determining which microphone received the speech signal." Kale, [0029], [0044], FIG. 8

90. The speech segment information (contained in the speech signal) is received by the noise cancelling device's "direction of arrival module" and "beamformer" as discussed in [1c] and [1d], *infra*. Kale, [0044].

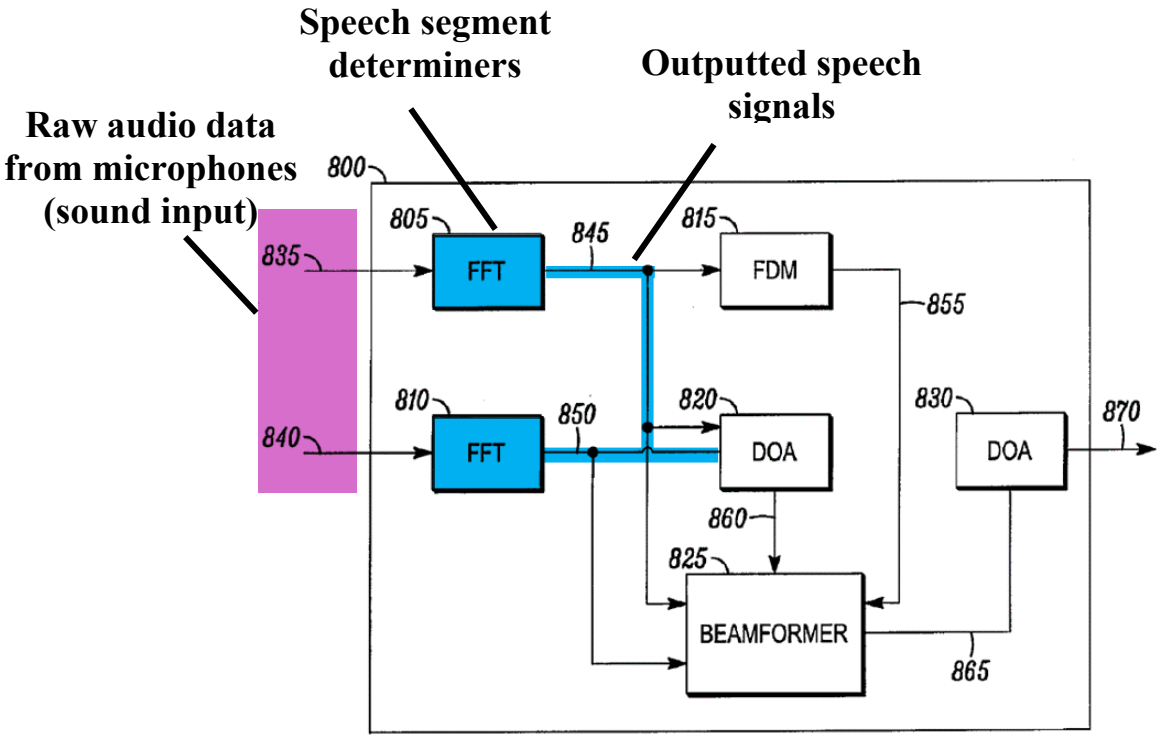


FIG. 8

Kale, FIG. 8 (annotated).

d. [1c]: "a voice direction detector configured, when receiving the speech segment information, to detect a voice incoming direction indicating from which direction a voice sound travels,"

91. Kale discloses [1c] because the noise cancelling device includes a "direction of arrival (DOA) module" for "determining the direction from which the speech data arrived." Kale, [0028], [0031], [0046], FIGS. 8-9. Annotated Figure 8

below illustrates "[t]he direction of arrival module receiv[ing] the first and second transformed signals" (speech segment information) from the FFT module. Kale, [0003], [0046], FIG. 8.

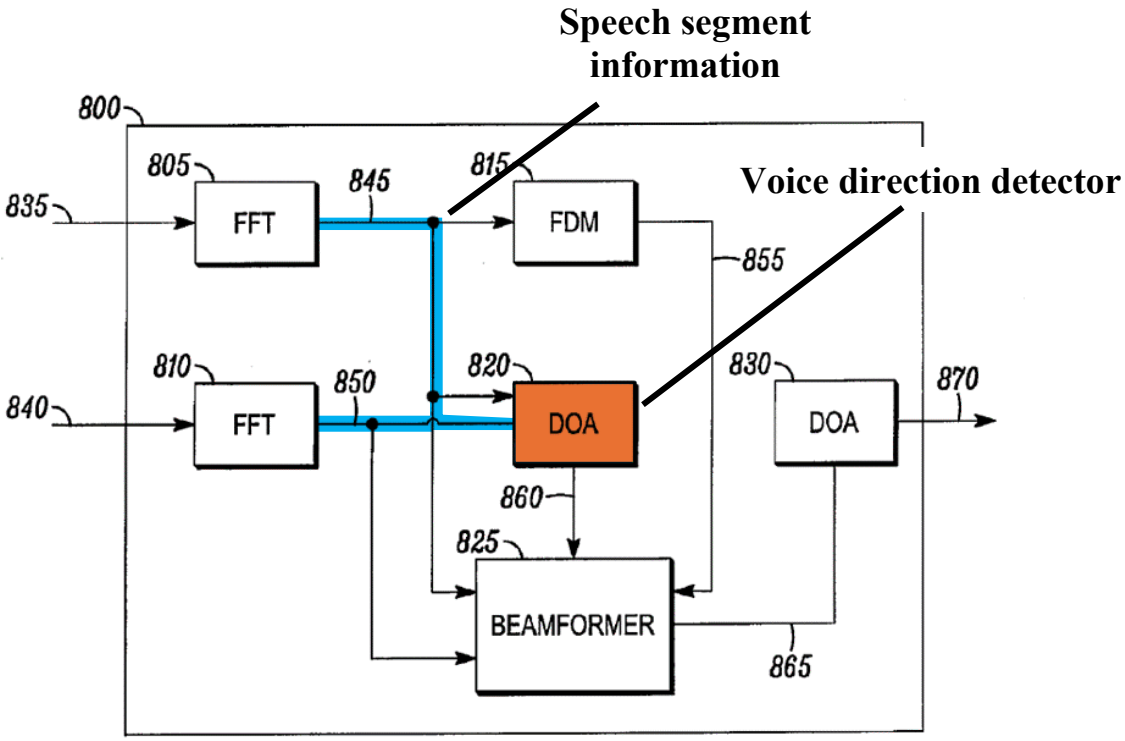


FIG. 8

Kale, FIG. 8 (annotated)

92. As Kale explains, the DOA module calculates the direction of arrival by "determin[ing] a cross-correlation between the first and second transformed signals, and generat[ing] a spatial orientation data signal." Kale, [003]. More specifically, the DOA module calculates the "direction of arrival" of the speech segment by "comput[ing] a cross-correlation between the two [speech and noise] signals 845, 850" to "generate[] a spatial orientation data signal" including "a first

angle relating to the speech data and a second angle relating to the noise data." Kale, [0003], [0031], claim 8. This calculation, as was well understood in the art, indicates the direction at which the speech sound arrived at the first or second microphone. *See e.g.,* Goodwin, 8:15-26, 8:31-47. Thus, like the '259 patent, which also describes an embodiment in which "a voice incoming direction is detected by… measuring the correlation between sounds picked up by the main microphone 11 and the sub-microphone 12," Kale discloses a voice direction detector. Ex. 1001, 19:3-9.

e. [1c.i]: "based on a first sound pick-up signal obtained based on a sound picked up by the first microphone and a second sound pick-up signal obtained based on a sound picked up by the second microphone"

93. Kale discloses [1c.i] because as explained for [1c] and [1a], the cross-correlation (voice direction determination) is based on first and second "speech signals" 845 and 850 corresponding to first and second raw audio data inputs 835 and 840 (sound pick-up signals) picked up by the first and second microphones. Kale, [0003], [0031], [0043]-[0044], claim 8. Annotated Figure 8 below illustrates that "the first and second microphone speech signals 845 and 850" (speech segment information) are based on "raw audio data" 835, 840 (first and second sound pick-up signals) picked up by the first and second microphones. Kale, [0031], [0043]-[0044], FIG. 8.

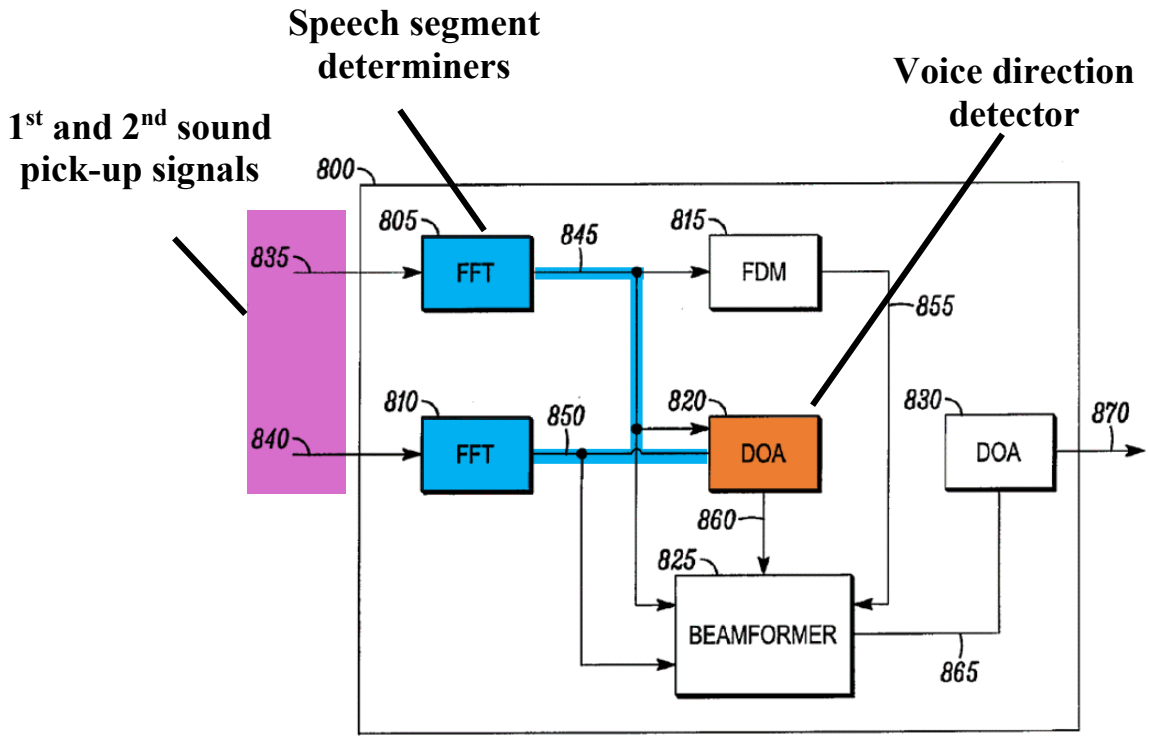


FIG. 8

Kale, FIG. 8 (annotated).

f. [1c.ii]: "and to output voice incoming-direction information when the voice incoming direction is detected; and"

94. Kale discloses [1c.ii] because the DOA module (voice direction detector) outputs a "DOA data signal" (voice incoming-direction information) when the speech is detected. Kale, [0031], [0046], Fig. 8. The DOA data signal is received by beamformer 825, which uses the signal to "place[] a null at the noise frequency direction for the formant range of frequencies, thereby eliminating the

maximum noise in the range" as illustrated in annotated Figure 8 below. Kale, [0033], [0034], [0046]- [0047], FIG. 8.

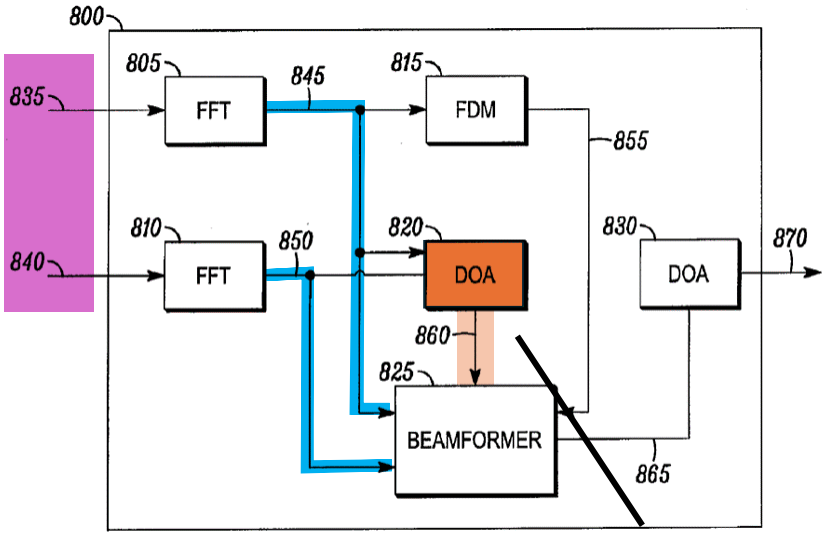


FIG. 8

Kale, FIG. 8 (annotated).

g. [1d]: "an adaptive filter configured to perform a noise reduction process using the first and second sound pick-up signals based on the speech segment information and the voice incoming-direction information."

95. Kale discloses [1d] because the noise canceling device includes a "beamformer" (adaptive filter) configured to "modify an audio signal to isolate a speech therein" or "enhance a speech of an audio signal" by "eliminating" the noise contained in the received audio signal" (perform, a noise reduction process) (1) using the microphone sound pick-up signals, (2) based on the first and second

"microphone speech signal[s]" (speech segment information), and (3) based on "the DOA data signal 860" (voice incoming-direction information). Kale, [0040], [0047]-[0048], FIGS. 8 and 9.

96. Annotated Figure 8 below illustrates the beamformer receiving the DOA data signal 860 (voice incoming-direction information) and the first and speech signals 845, 850 (speech segment information) that are generated using the raw audio data signals 835, 840 (first and second sound pick-up signals).

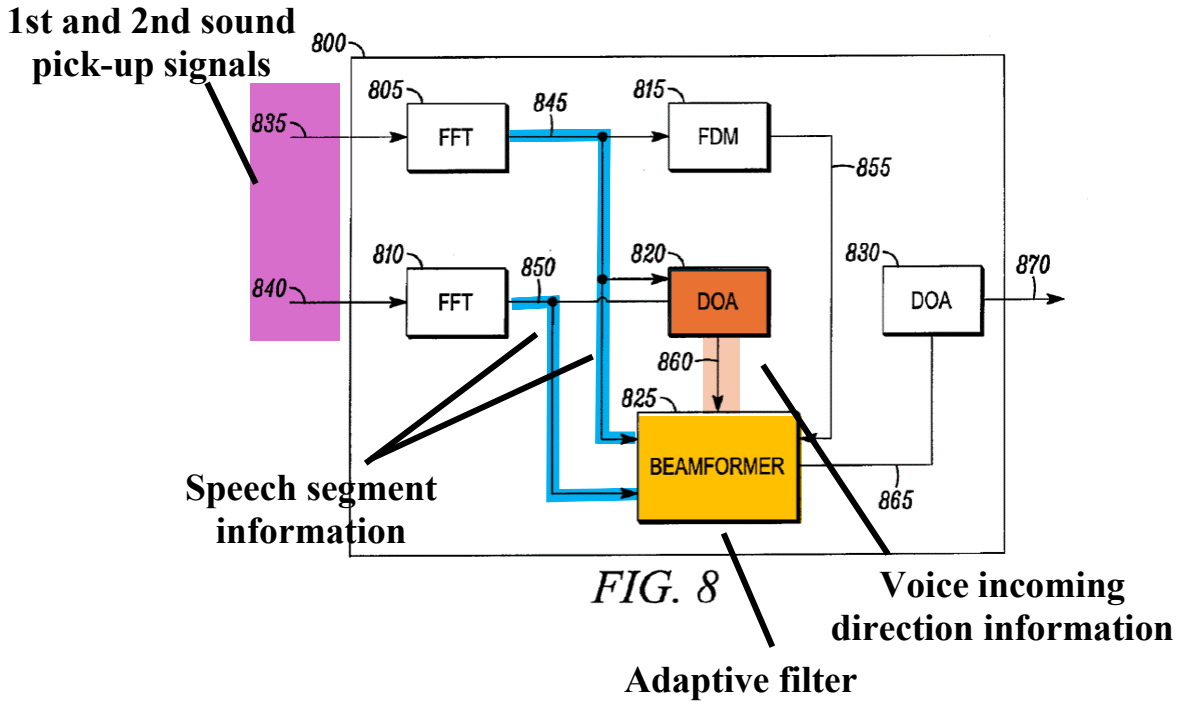


FIG. 8

Kale, FIG. 8 (annotated).

2. Claim 2: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals."

97. Kale at least suggests claim 2 because, as explained for [1c], the DOA module (voice direction detector) detects the arrival direction of the speech by "comput[ing] a cross-correlation between the two [sound] signals 845, 850" that are based on the raw audio data (sound pick-up signals) and determining a "spatial orientation" of the signals." Kale, [0003], [0031].

98. A POSITA would have understood that the computed "cross correlation" pertains to a phase-difference calculation, i.e., the difference between the phase at which the first and second sound pick-up signals are received with respect to a reference microphone. *See e.g.,* Goodwin, 8:15-26 ("The direction estimator module 315 computes the direction d(t) of a source (e.g. user 102) of a speech component within the primary acoustic signal c(t) and/or the secondary acoustic signal f(t) based on a difference between the primary acoustic signal c(t) and the secondary acoustic signal f(t)…"), 8:31-47 ("In the illustrated embodiment, the direction d(t) is determined based on a maximum of the cross-correlation between the primary acoustic signal c(t) and the secondary acoustic signal f(t). A maximum of the cross-correlation between the primary and secondary acoustic signals c(t), f(t) indicates the time delay between the arrival of the acoustic wave

generated by the user 102 at the primary microphone 106 and at the secondary microphone 108. The time delay is dependent upon the distance Δ between the primary microphone 106 and the secondary microphone 108 and the angle of incidence of the acoustic wave generated by the user 102 upon the primary and secondary microphones 106, 108…"). It would have been understood by a POSITA that this phase difference can be used to compute the angle of the direction of arrival with respect to the first and second sound pick-up signals. This is consistent with the '259 patent which also explains that "[t]he phase difference of sounds between the main microphone 11 and sub-microphone 12 can be calculated using the cross correlation function." Ex. 1001, 19:9-12.

99. Moreover, Kale explains that the DOA module's output includes the phases of the sound input signals, e.g., "two angles in degrees, the first being for a desired speech signal while the second is for noise." Kale, [0031]. Output "θN is the direction of noise and [output] θS is the direction of sound." Kale, [0039]. A POSITA would have understood that these angles correspond to the phases of the first and second sound pick-up signals—that is, θN corresponds to the second sound pick-up signal when it contains the noise signal and θS corresponds to the first sound pick-up signal when it contains the voice signal. Thus, Kale discloses that its DOA module detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals.

3. Claim 3: "The noise reduction apparatus according to claim 2, wherein the adaptive filter performs the noise reduction process to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal

whereas the adaptive filter performs the noise reduction process to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

100. Kale discloses claim 3 because the beamformer (adaptive noise filter) reduces the noise component of whichever signal contains the advanced speech signal (either the first or second sound pick-up signal) using the delayed noise signal (the other of the first or second sound pick-up signal). Kale, [0034], [0039].

101. As explained for [1d], Kale's adaptive filter is configured to "modify an audio signal to isolate a speech therein" or "enhance a speech of an audio signal" by "eliminating" the noise contained in the received audio signal" based on the "first microphone speech signal 845, the second microphone speech signal 850" (speech segment information) and based on the DOA data signal 860" (voice incoming-direction information). Kale, [0047]-[0048], FIGS. 8 and 9.

102. And, as explained for claim 2, Kale's DOA module outputs the phases of each sound input signal as "two angles in degrees, the first being for a desired speech signal while the second is for noise." Kale, [0031], [0039] ("θN is the

direction of noise and θS is the direction of sound"). Thus, a POSITA would have understood that the DOA data signal includes data indicating which sound pickup signal includes an advanced phase because the angles indicate which signal was delayed. A POSITA would have also understood that a small angle corresponds to a signal having an advanced phase, while a large angle indicates that the signal was delayed. Li-638, [0030], [0053], [0061]).

103. Kale also explains that "[t]he beamformer [] places a null *at the noise frequency direction* for the formant range of frequencies, thereby eliminating the maximum noise in the range." Kale, [0034] (emphasis added). Accordingly, when "θN is the direction of noise and θS is the direction of sound … the requirement is to *place a null at θN* and unity at θS." Kale, [0039] (emphasis added). That is, when θS (corresponding to the speech signal with the smaller angle of incidence) has an advanced phase relative to θN (corresponding to the noise signal), the noise signal is used for noise reduction.

104. Thus, a POSITA would have understood that Kale's "beamformer" (adaptive filter) reduces a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal received by the first microphone has a more advanced phase (θS) than the second sound pick-up signal received by the second microphone (corresponding to θN).

105. Kale further discloses that either the first or second signal (first sound pick-up signal 845 or second sound pick-up signal 850) may contain the speech signal (corresponding to θS) from which noise is reduced. Kale, [0031]-[0032] (explaining that while "signals 845, 850 are assumed to be speech and noise, respectively…the options may be changed so that the first signal 845 represents noise while the second signal 850 represents speech").

106. Thus, a POSITA would have also understood that Kale's "beamformer" (adaptive filter) reduces a noise component carried by the second sound pick-up signal (speech signal 850) using the first sound pick-up signal (noise signal 845) when the second sound pick-up signal has a more advanced phase **(** θS) than the first sound pick-up signal **(**corresponding to θN).

4. Claim 5: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on magnitudes of the first and second sound pick-up signals."

107. Kale at least suggests claim 5 because the DOA module (voice direction detector) determines the arrival direction ("spatial orientation") of the speech by determining which "peak" in the noise signal or speech signal is more "prominent" (has a greater magnitude) Kale, [0003], [0031].

108. As explained for [1c], the DOA module calculates the direction of arrival by "determin[ing] a cross-correlation between the first and second

transformed signals, and generat[ing] a spatial orientation data signal." Kale, [0003]. For example, the DOA module calculates the "direction of arrival" of the speech segment by "comput[ing] a cross-correlation between the two [speech and noise] signals 845, 850" to "generate[] a spatial orientation data signal" including "a first angle relating to the speech data and a second angle relating to the noise data." Kale, [0003], [0031], claim 8.

109. Kale further explains that the DOA module ascertains the "prominen[ce]" of the "peaks" (magnitudes) associated with each of the first and second spatial orientation signals when determining the direction of arrival. Kale, [0031]. A POSITA would have understood that peak prominence (magnitude) is indicative of speech sound proximity and direction as was well known in the art. *See e.g.,* Li-730, [0023] ("As is well known in the art, a substantially flat curve in the spatial spectrum SS is caused by far field noises, and sharp and *dominant peaks in the spatial spectrum SS is caused by near field sounds of a speaker's voice* and spot noises from the environment." (emphasis added)); Laroche, 4:57-61 ("Because the primary microphone 106 is much closer to the user 102 than the secondary microphone 108, the intensity level is higher for the primary microphone 106, resulting in a larger energy level received by the primary microphone 106 during a speech/voice segment, for example."); *see also* Kale,

[0021], FIGS. 3-4 (describing and illustrating a beampattern wherein noise acquired "directly in front of the microphone" produces a prominent peak).

110. Thus, a POSITA would have understood that a determination by the DOA module that a signal is "prominent" corresponds to a determination that the input is closer, e.g., directly in front of the first microphone; whereas, a determination that the signal is not prominent indicates that the sound source is farther away, i.e., in a direction away from the microphone. Li-730, [0023]; Kale, [0021], FIGS. 3-4.

111. Accordingly, Kale discloses that its DOA module detects the voice incoming direction based on magnitudes of the first and second sound pick-up signals.

5. Claim 6: "The noise reduction apparatus according to claim 5, wherein the adaptive filter performs the noise reduction process to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a greater magnitude than the second sound pick-up signal

whereas the adaptive filter performs the noise reduction process to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a greater magnitude than the first sound pick-up signal."

112. Kale discloses claim 6 because the beamformer (adaptive noise filter) reduces the noise component of whichever signal contains the more prominent

signal (either the first or second sound pick-up signal) using the noise signal (the other of the first or second sound pick-up signal). Kale, [0034], [0039].

113. As explained for [1d], Kale's "beamformer" (adaptive filter) is configured to "modify an audio signal to isolate a speech therein" or "enhance a speech of an audio signal" by "eliminating" the noise contained in the received audio signal" (1) using the microphone sound pick-up signals, (2) based on the first and second "microphone speech signal[s]," and (3) based on "the DOA data signal 860." Kale, [0047]-[0048], FIGS. 8 and 9.

114. Kale further discloses that either the first or second speech signals corresponding to sound pick-up signals 835, 840 may contain the "prominent" speech signal from which noise is reduced. Kale, [0031]-[0032] (explaining that while "signals 845, 850 are assumed to be speech and noise, respectively… the options may be changed so that the first signal 845 represents noise while the second signal 850 represents speech."). And Kale further explains that "[i]f the DOA 820 determines that the second peak of the second signal 850 is not prominent, a null value is provided." Kale, [0031]; *see also* [0034] ("The beamformer [] places a null *at the noise frequency direction* for the formant range of frequencies, thereby eliminating the maximum noise in the range." (emphasis added)), [0039] ("Assuming θN is the direction of noise and θS is the

direction of sound … the requirement is to *place a null at θN* and unity at θS." (emphasis added)).

115. Accordingly, it is my opinion that when the first sound pick-up signal corresponding to the "first microphone speech signal" contains the prominent speech signal, the second sound pick-up signal corresponding to the "second microphone speech signal" is used to reduce the noise component (provide a "null"). Kale, [0031]. Conversely, a POSITA would have understood that when the second sound pick-up signal contains the prominent speech signal, the beamformer reduces the noise contained in the second pick-up signal using the less prominent first sound pick-up signal. Kale, [0031].

6. Claim 8: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals and magnitudes of the first and second sound pick-up signal."

116. Claim 8 recites a combination of claims 2 and 5. Kale renders obvious claim 8 for reasons discussed for claims 2 and 5.

7. Independent Claim 13

a. [13preamble]: "A noise reduction method comprising the steps of:"

117. Kale discloses [13preamble] because it discloses a "method for noise cancellation using multiple microphones that is formant aided." Kale, [0016], FIG. 9.

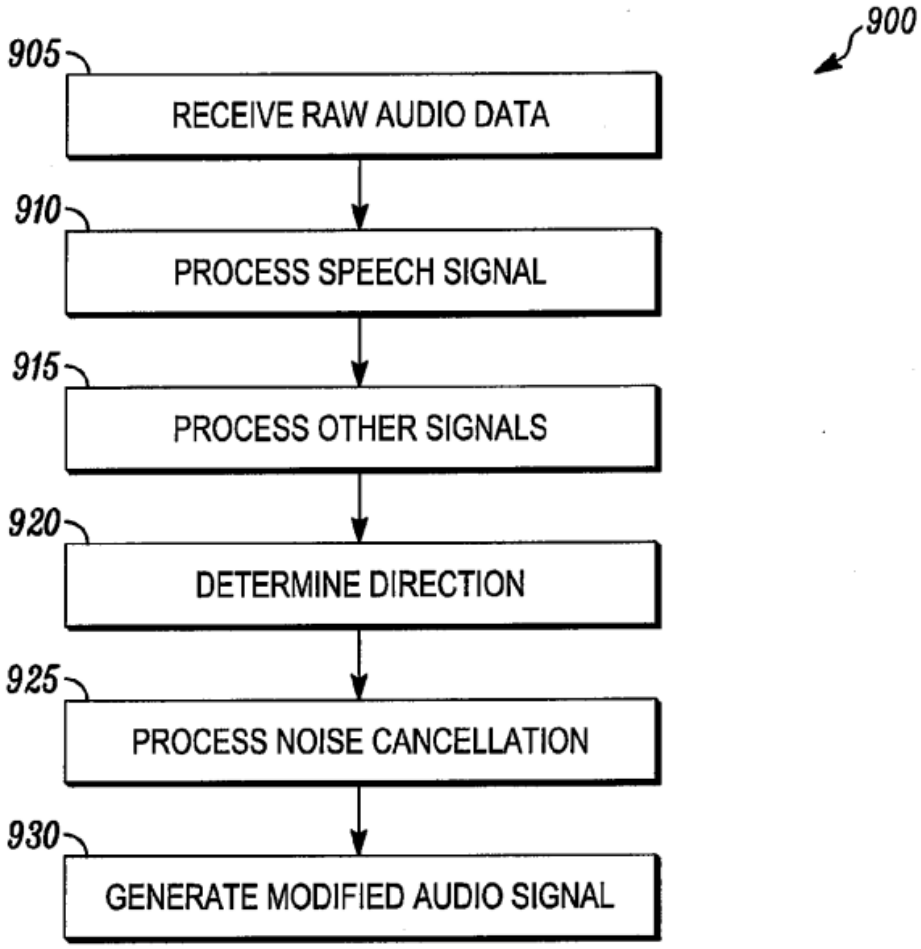


FIG. 9

Kale, FIG. 9.

b. [13a]: "determining whether or not a sound picked up by at least either a first microphone or a second microphone is a speech segment;"

118. [13a] is substantively similar to [1a]. Kale discloses [13a] for reasons discussed for [1a] above.

c. [13b]: "detecting a voice incoming direction indicating from which direction a voice sound travels, based on a first sound pick-up signal obtained based on a sound picked up by the first microphone and a second sound pick-up signal obtained based on a sound picked up by the second microphone, when it is determined that the sound picked up by the first or the second microphone is the speech segment; and"

119. Claim [13b] incorporates the limitations of claims [1b], [1c], and [1c.i]. Kale discloses [13b] for reasons discussed for [1b], [1c], and [1c.i] above.

d. [13c]: "performing a noise reduction process using the first and second sound pick-up signals based on speech segment information indicating that the sound picked up by the first or the second microphone is the speech segment and voice incoming-direction information indicating the voice incoming direction."

120. [13c] is substantively similar to [1d]. Kale discloses [13c] for reasons discussed for [1d] above.

8. Claim 14: "The noise reduction method according to claim 13, the voice incoming direction is detected based on a phase difference between the first and second sound pick-up signals."

121. Claim 14 is substantively similar to claim 2. Kale discloses claim 14 for reasons discussed for clam 2 above.

9. Claim 15: "The noise reduction method according to claim 14, wherein the noise reduction process is performed to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal whereas the noise reduction process is performed to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

122. Claim 15 is substantively similar to claim 3. Kale discloses claim 15 for reasons discussed for clam 3 above.

10. Claim 17: "The noise reduction method according to claim 14, wherein the voice incoming direction is detected based on magnitudes of the first and second sound pick-up signals."

123. Claim 17 is substantively similar to claim 5. Kale discloses claim 17 for reasons discussed for claim 5 above.

11. Claim 18: "The noise reduction method according to claim 17, wherein the noise reduction process is performed to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a greater magnitude than the second sound pick-up signal whereas the noise reduction process is performed to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a greater magnitude than the first sound pick-up signal."

124. Claim 18 is substantively similar to claim 6. Kale discloses claim 18 for reasons discussed for claim 6 above.

12. Claim 19: "The noise reduction method according to claim 14, wherein the voice incoming direction is detected based on a phase difference between the first and second sound pick-up signals and magnitudes of the first and second sound pick-up signals."

125. Claim 19 is substantively similar to claim 8. Kale discloses claim 19 for reasons discussed for claim 8 above.

B. Ground 2: Claims 2, 3, 5, 6, 8, 9, 14, 15 and 17-20 Are Obvious over Kale and Namba

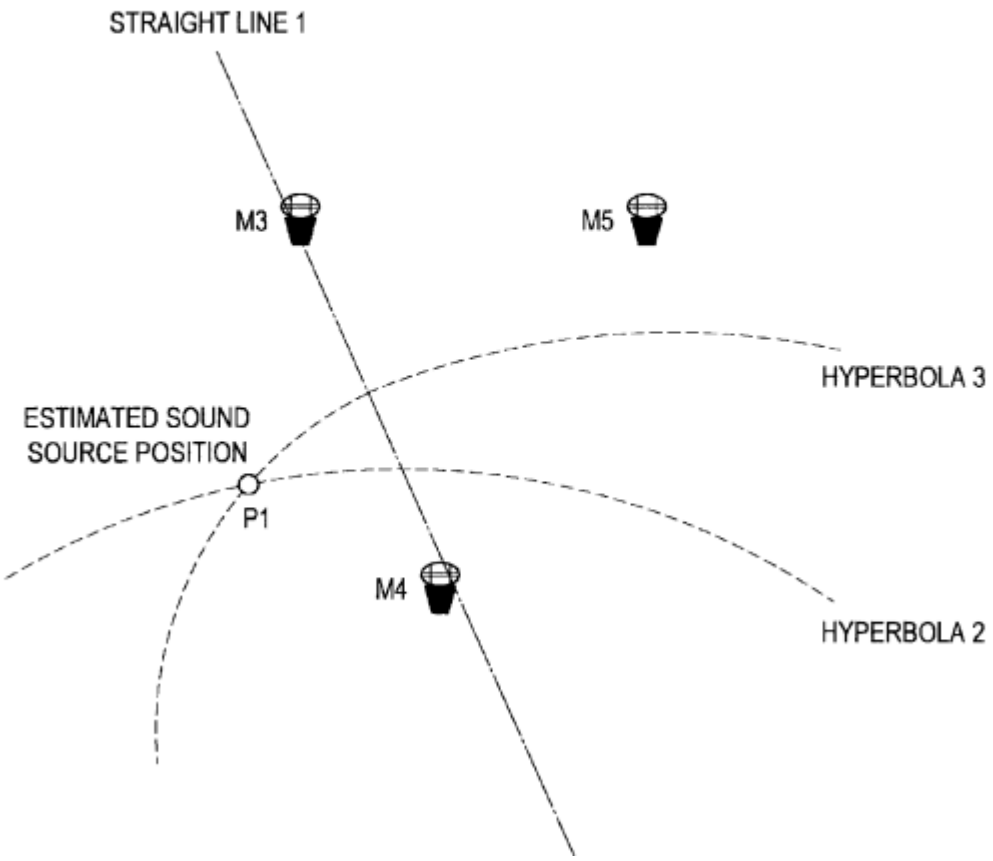
1. Claim 2: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals."

126. I have been asked to consider whether Namba teaches the additional limitations of claim 2 and if the combination of Kale and Namba renders obvious claim 2. In my opinion it does because Namba teaches a sound processing apparatus including a "distance/direction estimator" (voice direction detector) that determines the direction and distance of a "sound source" "based on a phase difference of two input sounds" of the apparatus. Namba, [0070]-[0073], [0074]-[0076], FIGS. 5 and 6.

127. As Namba explains, the distance/ direction estimator measures "the phase of each input sound reaching a microphone" e.g., "M1 and a microphone M2 … and a phase difference of the input sounds" and uses the phases and phase difference to calculate "the sound source position of input sound." Namba, [0071].

For example, and as illustrated in Figure 6, "the distance/direction estimator 144 can determine that the sound source is positioned on the opposite side of the microphone M5 with respect to a straight line 1 linking the microphone M3 and the microphone M4 (front/back determination)" when "phase of input sound arriving at the microphone M5 [is] delayed when compared with that of input sound arriving at the microphone M3 or the microphone M4." Namba, [0076], FIG. 6. The distance/direction determiner can also "a hyperbola 2 on which the sound source could be present based on a phase difference of input sounds arriving at each of the microphone M3 and the microphone M4" and "a hyperbola 3 on which the sound source could be present based on a phase difference of input sounds arriving at each of the microphone M4 and the microphone M5" and "estimate that an intersection P1 of the hyperbola 2 and the hyperbola 3 is the sound source position." Namba, [0077], FIG. 6.

FIG.6



Namba, FIG. 6.

a. Motivation to Combine

128. Namba further explains that the techniques implemented by its distance/direction estimator are particularly useful for determining sound input direction because they *combine* phase *and* volume (magnitude) "estimation methods" to detect sound input direction "even if" the sound input signals (sound pick-up signals) experience "reverberation" or "reflection." Namba, [0070]. The sound signals described in Kale are likewise susceptible to distortions caused by

reverberation and reflection (caused e.g., by bouncing off the sound apparatus itself or objects between the sound apparatus and the user's voice) and would have found it beneficial to integrate phase-based voice direction detection, as taught by Namba to reduce such reverberations or reflections that might otherwise reduce the accuracy of the sound input apparatus's voice direction determination.

129. Accordingly, it is also my opinion that a POSITA would have been motivated to also detect the voice incoming direction based on the phase of the first and second sound pick-up signals to enhance voice direction detection even when the sound input signals experience "reverberation" or "reflection," as taught by Namba. Namba, [0070].

b. Expectation of Success

130. Kale and Namba use similar elements and techniques to achieve similar results (see Overview of Namba above) and a POSITA would have had a reasonable expectation of success in configuring Kale's voice direction detector to implement Namba's phase-based voice direction determination process because it would have merely required routine skill in the art to do so. For example, a POSITA would have known to use the principles and formulas discussed in [0072]-[0073], [0076] of Namba, taking into account the microphone position of the noise suppression device of Kale, to ascertain voice incoming direction based on phase and phase difference.

2. Claim 3: "The noise reduction apparatus according to claim 2, wherein the adaptive filter performs the noise reduction process to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal

whereas the adaptive filter performs the noise reduction process to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

131. Kale-Namba renders obvious claim 3 for the same reasons Kale alone renders claim 3 obvious as explained in Ground 1 above. Modifying Kale to incorporate Namba's phase-based voice direction detection, as described in claims 2 and 14, does not alter this.

3. Claim 5: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on magnitudes of the first and second sound pick-up signals."

132. I have been asked to consider whether Namba teaches the additional limitations of claim 5 and if the combination of Kale and Namba renders obvious claim 5. In my opinion it does because Namba's "distance/direction estimator" (voice direction detector) determines the direction and distance of a "sound source" "based on a volume difference between input sounds recorded by the microphone[s]" of the apparatus. Namba, [0070], [0074]-[0087], FIGS. 5, 7, 9.

133. As illustrated in Table 1 below and discussed further below, the distance/direction estimator uses the volume (amplitude) to calculate source sound direction and distance information.

TABLE 1

Block	Input	Output
Volume detector	Input sound	Volume value sequence (amplitude) in frame
Average volume detector	Volume value sequence (amplitude) in frame	Average value of volume
Maximum volume detector	Volume value sequence (amplitude) in frame	Maximum value of volume
Spectrum detector	Input sound	Spectrum
Sound quality detector	Input sound Average value of volume Maximum value of volume Spectrum	Likeness of human voice Likeness of music Steady or non-steady Impulse property
Distance/direction estimator	Input sound Volume value sequence (amplitude) in frame Spectrum	Direction information Distance information

Namba, Table 1 (annotated). Because "magnitude" as described in the '259 patent includes the "maximum amplitude" of "the sound pick-up signals" and the like ("etc."), a POSITA would have understood that Namba's sound source direction determination based on volume/amplitude of the sound input signals is *a voice direction detector that detects voice incoming direction based on magnitude*, as recited in claim 5. Ex. 1001, 25:10-12.

134. Namba further discloses various equations and mathematical principles that can be used to ascertain sound direction and position based on the volume of the input sounds. Namba, [0074]-[0087]. Figure 5 illustrates a sound source that is determined to be closer to microphone M2 and located on "hyperbola 1" at position (x, y) based on the volume difference (magnitudes of the first and second sound pick-up signals received at M1 and M2). Namba, [0074], FIG. 5.

135. In another example, the "sound source position of an input sound is estimated based on volumes of two input sounds" using *inter alia* "the inverse square law" and "the ratio of volumes arriving at the microphones" to determine the sound source is located on a circle closer to microphone M6. Namba, [0078]-[0083], FIG. 7. In yet another example, similar calculations are performed to determine that the sound source is positioned at the intersection of two circles. Namba, [0086]-[0087], FIG. 8.

a. Motivation to Combine

136. As I explain for claim 2, Namba further explains that the techniques implemented by its distance/direction estimator are particularly useful for determining sound input direction because they *combine* phase *and* volume (magnitude) "estimation methods" to detect sound input direction "even if" the sound input signals (sound pick-up signals) experience "reverberation" or "reflection" caused by the sound pick-up apparatus itself. Namba, [0070]. Thus, a

POSITA would have been motivated to also detect the voice incoming direction based on magnitudes of the first and second sound pick-up signals to enhance direction detection during "reverberation" or "reflection." Namba, [0070].

b. Expectation of Success

137. Kale and Namba use similar elements and techniques to achieve similar results (see Overview of Namba above) and a POSITA would have had a reasonable expectation of success in configuring Kale's voice direction detector to implement Namba's magnitude-based voice direction determination process because it would have merely required routine skill in the art to do so. For example, a POSITA would have known to implement the formulas and questions identified in [0074]-[0083] of Namba, taking into account the microphone position of the noise suppression device of Kale, to ascertain voice incoming direction based on magnitude.

4. Claim 6: "The noise reduction apparatus according to claim 5, wherein the adaptive filter performs the noise reduction process to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a greater magnitude than the second sound pick-up signal

whereas the adaptive filter performs the noise reduction process to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a greater magnitude than the first sound pick-up signal."

138. Kale and Namba render obvious claim 6 because Kale's beamformer (adaptive noise filter) reduces the noise component of whichever signal contains the more prominent signal (either the first or second sound pickup signal) using the noise signal (the other of the first or second sound pick-up signal), as explained in Ground 1, above. Kale, [0034], [0039]. Modifying Kale to incorporate Namba's magnitude-based voice direction detection, as described in claim 5, does not alter this.

5. Claim 8: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals and magnitudes of the first and second sound pick-up signal."

139. Kale and Namba render obvious claim 8 because the combination teaches a voice direction detector that uses volume/amplitude (magnitude) and phase difference to determine voice direction, as explained for claims 2 and 5 above. Namba, [0070].

6. Claim 9: "The noise reduction apparatus according to claim 1, wherein the speech segment determiner detects the speech segment based on the first sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal

whereas the speech segment determiner detects the speech segment based on the second sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

140. Namba renders obvious claim 9 because its "sound type estimation unit" detects voice (a speech segment) based on sound pick-up signal phases (delays) calculated by its "distance/direction estimator." Namba, [0060], [0071], [0089], [0091].

141. Namba discloses a speech segment determiner because it discloses a "sound type estimation unit that *determines whether any sound* originating from the neighborhood of the imaging apparatus such as a voice of an operator of the imaging apparatus or noise resulting from an operation of the operator *is contained*." Namba, [0060] (emphasis added). Namba further discloses that its sound type estimation unit detects the voice (speech segment) based on the sound pick-up signal phases (delays) calculated by its "distance/direction estimator." Namba, [0060] ("The sound type (steady or non-steady, noise or sound) is estimated, for example, from … direction distance information.).

142. For example, and as explained above for claim 2, Namba's distance/direction estimator measures "the phase of each input sound reaching a microphone" e.g., "M1 and a microphone M2 … and a phase difference of the input sounds" and uses the phases and phase difference to calculate "the sound source position of input sound." Namba, [0071]. The direction information is then fed into a "sound estimator" of the sound type unit to determine if the sound is voice (e.g., the device operator's voice). Namba, [0089] ("The sound estimator 146 collectively determines whether any neighborhood sound originating from a specific sound source in the neighborhood of the sound processing apparatus 10 *such as a voice of the operator* … is contained in the input sound based on at least one of the … p*ositional information of input sound*." (emphases added)). If a sound input signal is advanced, the estimator may determine that the input contains voice. Namba, [0091] ("If the position of the sound source of input sound is behind an imaging unit[5] … and the input sound has sound quality that matches or resembles that of human voice, the sound estimator 146 may determine that the voice of the operator is [the] predominant[]…sound input".)

[5] Sound received "behind" the imaging unit (rear microphone) is advanced because it is received closer to another microphone. Namba, [0041], [0077].

143. Namba's direction determination improves sound input detection even when sounds reverberate or reflect off the user device. Namba, [0070]. Thus a POSITA would have been motivated and found it obvious to implement Namba's direction/determination unit functionality to improve Kale's speech detection. A POSITA would have had a reasonable expectation of success in making such a modification because it would have merely required routine skill in the art to do so, as explained for claim 2.

7. Claim 14: "The noise reduction method according to claim 13, the voice incoming direction is detected based on a phase difference between the first and second sound pick-up signals."

144. Claim 14 is substantively similar to claim 2. Namba renders obvious claim 14 for reasons discussed for claim 2 above.

8. Claim 15: "The noise reduction method according to claim 14, wherein the noise reduction process is performed to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal whereas the noise reduction process is performed to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

145. Claim 15 is substantively similar to claim 3. Kale-Namba renders obvious claim 15 for reasons discussed for claim 3 above.

9. Claim 17: "The noise reduction method according to claim 14, wherein the voice incoming direction is detected based on magnitudes of the first and second sound pick-up signals."

146. Claim 17 is substantively similar to claim 5. Namba renders obvious claim 17 for reasons discussed for claim 5 above.

10. Claim 18: "The noise reduction method according to claim 17, wherein the noise reduction process is performed to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a greater magnitude than the second sound pick-up signal whereas the noise reduction process is performed to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a greater magnitude than the first sound pick-up signal."

147. Claim 18 is substantively similar to claim 6. Kale-Namba render obvious claim 18 for reasons discussed for claim 6 above.

11. Claim 19: "The noise reduction method according to claim 14, wherein the voice incoming direction is detected based on a phase difference between the first and second sound pick-up signals and magnitudes of the first and second sound pick-up signals."

148. Claim 19 is substantively similar to claim 8. Namba renders obvious claim 19 for reasons discussed for claim 8 above.

12. Claim 20: "The noise reduction method according to claim 19, wherein the speech segment is detected based on the first sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal whereas the speech segment is detected based on the second sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

149. Claim 20 is substantively similar to claim 9. Namba renders obvious claim 20 for reasons discussed for claim 9 above.

C. Ground 3: Claims 2-4, 7, and 14-16 Are Obvious over Kale and Li-638

1. Claim 2: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals."

150. I have been asked to consider whether Li-638 teaches the additional limitations of claim 2 and if the combination of Kale and Li-638 renders obvious claim 2. In my opinion it does because Li-638 teaches "a noise reduction control device" comprising a "filtering control unit" for "calculating incidence angles of []microphone array signals" (target speech and noise signals) based on the relative time delays of [] frequency bins or sub-bands" and "calculating incidence angles of the microphone array signals *based on the relative time delays* of the frequency bins or sub-bands." Li-638, [0030], [0053], [0061].

151. Like Kale, Li-638's noise reduction control device includes two microphones ("mic_a and mic_b") for receiving "signals s_1 and s_2 … wherein the

signal s_1 is treated as desired speech signal and s_2 as reference signal." Li-638, [0052]. The noise reduction device also includes a filtering control unit (voice direction detector) that determines the "incidence angle" of the "target speech component" (voice incoming direction) by calculating "phase differences between various frequency bins or sub-bands of the microphone array signals and calculating relative time delays of the frequency bins or sub-bands of the microphone array signals from the phase differences." Li-638, [0030], [0053], [0061], FIG. 5.

152. More specifically, the filtering control unit (voice direction detector) includes (1) "a signal delay estimation unit for calculating *phase differences between various frequency bins or sub-bands of the microphone array signals and calculating relative time delays of the frequency bins or sub-bands of the microphone array signals from the phase differences*" and (2) "a signal direction estimation unit for calculating incidence angles of the microphone array signals *based on the relative time delays* of the frequency bins or sub-bands" Li-638, [0030], [0053], [0061], FIG. 5 (emphases added). Li-638 further explains:

> Assuming the target speech signal is incident from 0 degree direction, the signal direction estimation unit converts [the] relative time delays of each of [the] frequency bins or sub-bands of the two signals into their incidence angle [such that] target speech components within [an]

angle of protection and noise components outside the angle of protection may be distinguished according to their incidence angles.

Li-638, [0053].

153. Thus, Li-638 teaches a voice direction detector that detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals.

a. Motivation to Combine

154. Li-638 explains that its "technical solution for noise reduction by means of voice incidence angles" (phase difference-based voice direction determination) "protects the target speech while eliminating noises in the target speech" by allowing the adaptive filter to slow down and eventually stop the filtering processes when "all signals are target speech components" Li-638, [0055], [0065]. This technique is advantageous over other prior art systems, which were unable to adequately account for the fact that "signals received by both microphones contain speech components" and were therefore prone to damaging the desired speech signal during the noise reduction process.

155. Thus, it is my opinion that a POSITA would have been motivated to modify Kale's DOA module to implement (or substitute Kale's DOA module for) Li-638's signal delay estimation, signal direction estimation, and signal component statistic units (voice direction detector), to improve Kale's noise reduction process

by reducing the likelihood of removing desirable speech from the input signals using Li-638's voice direction determination techniques.

b. Expectation of Success

156. Kale and Li-638 use similar elements and techniques to achieve similar results (*see* ¶¶151-153 above) and a POSITA would have had a reasonable expectation of success in configuring Kale's noise suppression device to implement Li-638's voice direction determination process because it would have merely required routine skill in the art to do so. For example, a POSITA would have known to integrate Li-638's voice direction detector such that it receives Kale's first and second sound input signals (speech and noise signals 845, 850) as input into the signal delay estimation, because Li-638's signal delay estimation, like Kale's DOA Module, is configured to receive speech and noise signals from an FFT unit/module. Kale, [0031]; Li-638, [0030].

2. Claim 3: "The noise reduction apparatus according to claim 2, wherein the adaptive filter performs the noise reduction process to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal

whereas the adaptive filter performs the noise reduction process to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

157. Kale and Li-638 also render obvious the additional limitations of claim 3.

158. As explained for claim 2, Li-638's voice direction determiner calculates the "time delay" and "incidence angle" of the noises and targeted speech contained in each of the sound pick-up signals s1, s2 so that its adaptive filter H can reduce noise in the signals until they "contain only target speech components." Li-638, [0064]. The time delay and incidence angle (phase) are indicative of the amount of noise or speech contained in the signal's sub band, such that when the incidence angle is large (with respect to "target speech that is in 0 degree direction"), the sub band contains more noise. Li-638, [0064] [0069] ("The larger the incidence angle is, the more the speech of this frequency bin or sub-band deviates from the target speech that is in 0 degree direction."). Conversely, "[w]hen the incidence angle of the i^{th} frequency bin or sub-band is in the 0 degree

direction" it contains target speech only. Li-638, [0064], [0069]. Thus, a POSITA would have understood that Li-638's adaptive filter reduces the noise component of signals (s1 or S2) having lower incidence angles (advanced phase and less noise) using the noisier signal having the delayed phase (large incidence angle). Li-638, [0064] [0069].

a. Motivation to Combine

159. As explained for claim 2, Li-638's "technical solution for noise reduction by means of voice incidence angles" (phase difference-based voice direction determination) "protects the target speech while eliminating noises in the target speech" by allowing the adaptive filter to slow down and eventually stop the filtering processes when "all signals are target speech components" Li-638, [0055], [0065]. A POSITA would have been motivated and found it obvious to modify Kale's adaptive filter (as described in [1d]) to implement Li-638's phase-based noise suppression process to improve the noise reduction as explained in claim 2.

b. Expectation of Success

160. Kale and Li-638 use similar elements and techniques to achieve similar results (see ¶¶151-153 Li-638 above) and a POSITA would have had a reasonable expectation of success in configuring Kale's noise suppression device to implement Li-638's phase-based adaptive filtering process because it would have merely required routine skill in the art to do so as explained for claim 2. For

example, a POSITA would have known to implement the algorithms identified in [0063]-[0064] of Li-638 to control Kale's adaptive filter to eliminate noise from the desired advanced signal (target speech signal) without harming the speech therein as taught by Li-638.

3. Claim 4: "The noise reduction apparatus according to claim 2, wherein when the phase difference is within a predetermined range, the adaptive filter outputs either the first or the second sound pick-up signal without performing the noise reduction process."

161. Kale and Li-638 (as discussed for claim 2) render obvious claim 4 because Li-638 further teaches that its adaptive filter slows down and eventually stops the filtering processes when "all signals are target speech components" having an incidence angle in the range of -45° to 45°. Li-638, [0055], [0065].

162. For example, Li-638 explains that its adaptive filter slows down as the number of target speech signals in a received set of signals increases, and ceases updating entirely—that is, outputs the target signal without performing noise reduction—when "all signals are target speech components, α=0." Li-638, [0055] ("[T]he more the target signal components, which are within the angle of protection, the smaller α is, the slower the updating of the adaptive filter is. When all signals are target speech components, α=0, the adaptive filter stops updating of weights of the filter in this speech section, thereby protecting speech in the desired speech signal s_1 from being balanced out, thus effectively protecting target speech

from being damaged."). Li-638 explains that its filter stops updating (thereby outputting the target signal without performing noise reduction) to avoid damaging the target speech signal, and a POSITA would have been motivated and found it obvious to do the same in the Kale-Li-638 combination. Li-638, [0055].

4. Claim 7: "The noise reduction apparatus according to claim 5, wherein when a power difference that is a difference between magnitudes of the first and second sound pick-up signals is within a predetermined range, the adaptive filter outputs either the first or the second sound pick-up signal without performing the noise reduction process."

163. Kale and Li-638 render claim 7 obvious because similar to claim 4, Li-638 teaches that its adaptive filter "stops updating, i.e., stops "suppressing noise" when the "energy" (magnitude) difference between the noise signal s3 and speech signal s1 is at a minimum. Li-638, [0045], [0048], [0056]. A POSITA would have been motivated and found it obvious to output the target signal without performing noise reduction when the energy is at a minimum (in a predetermined range), as taught by Li-638, to avoid damaging the target speech signal. Li-638, [0055].

5. Claim 14: "The noise reduction method according to claim 13, the voice incoming direction is detected based on a phase difference between the first and second sound pick-up signals."

164. Claim 14 is substantively similar to claim 2. Kale-Li-638 renders obvious claim 14 for reasons discussed for claim 2 above.

6. Claim 15: "The noise reduction method according to claim 14, wherein the noise reduction process is performed to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal whereas the noise reduction process is performed to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

165. Claim 15 is substantively similar to claim 3. Kale-Li-638 renders obvious claim 15 for reasons discussed for claim 3 above.

7. Claim 16: "The noise reduction method according to claim 14, wherein when the phase difference is within a predetermined range, the adaptive filter outputs either the first or the second sound pick-up signal without performing the noise reduction process."

166. Claim 16 is substantively similar to claim 4. Kale-Li-638 renders claim 16 obvious for reasons discussed for claim 4 above.

D. Ground 4: Claims 10 and 11 Are Obvious over Kale and Yamada

1. Claim 10: "The noise reduction apparatus according to claim 1, wherein signals are supplied to the voice direction detector as the first and second sound pick-up signals at a sampling frequency of 24 KHz or higher and signals are supplied to the adaptive filter as the first and second sound pick-up signals at a sampling frequency of 12 KHz or lower."

167. Kale further discloses supplying the sound pick-up signals to the beamformer (adaptive filter) at a sampling frequency of 12 KHz or lower because Kale discloses supplying 8 KHz signals:

> The FFT 805 may receive a first microphone speech data 835 while the FFT 810 may receive a second microphone speech data 840. With reference to the exemplary rate of 20 ms, speech samples from the first and second microphones in 20 ms frames are computed by the FFTs 805, 810, respectively. According to the exemplary embodiments, the FFTs 805, 810 may compute a 128, 256, and/or 512 point FFT of a 8 kHz signal…

Kale, [0029].

168. Kale does not expressly disclose supplying sound pick-up signals to its DOA module (voice direction detector) at a sampling frequency of 24 KHz or more, but does explain that the sampling rate "may be changed" depending on the "resolution desired" of the output signals. Kale, [0029].

169. Yamada teaches a signal processing method wherein "sound picked up by a microphone" is supplied to a "sound source direction estimation section" "at a sampling frequency of 48 kHz." Yamada, [0067]-[0069].

a. Motivation to Combine

170. A POSITA would have been motivated and found it obvious to supply the sound pickup signal of Kale to the DOA module at a sampling rate of 24 KHz or more, including at 48 KHz, as taught by Yamada, to improve the accuracy of Kale's voice direction determination. A POSITA would have understood that increasing the sampling rate would have been particularly useful for voice direction determination because lower rates, such as 8 KHz, while suitable for

noise cancellation and suppression, would have produced low-resolution signals that would have made it more difficult to pinpoint the voice direction. Furthermore, it was known in the art to adapt sampling rate depending on the desired accuracy of the signal. Pedersen, [0023] ("In general, the sampling frequency is adapted to the application (available bandwidth, power consumption, frequency content of input signal, necessary accuracy, etc.). In an embodiment, the sampling frequency f_s is in the range from 8 kHz to 40 kHz, e.g. from 12 kHz to 24 kHz, e.g. around or equal to 16 kHz or 20 kHz.").

b. Expectation of Success

171. A POSITA would have had a reasonable expectation of success in configuring Kale's noise cancellation device to supply sound input signals to the DOA module at 24 KHz or more because it would have merely required routine skill in the art to do so.

172. For example, a POSITA would have also known to implement e.g., a sample frequency converter or sown sampling unit to continue to supply the sound input signal to Kale's adaptive filter as was known in the art. *See e.g.*, Tateyama, [0009]-[0010] (describing a signal processing method wherein an input signal is down sampled to a lower sampling rate before being provided to an adaptive filter). Maintaining the lower sampling frequency to the adaptive filter would have

reduced power consumption on the adaptive filter by reducing the data input. Tateyama, [0009]-[0010].

2. Claim 11: "The noise reduction apparatus according to claim 1, wherein the speech segment determiner outputs more accurate speech segment information to the voice direction detector than speech segment information to the adaptive filter."

173. Kale and Yamada render claim 11 obvious because, as explained for claim 10, increasing the sampling frequency to 48 KHz, as taught by Yamada, increases the accuracy of the outputted signal relative to the 8 KHz sampling frequency.

E. Ground 5: Claims 1-3 and 12-15 Are Obvious over Li-730

1. Independent claim 1

a. [1preamble]: "A noise reduction apparatus comprising:"

174. Li-730 discloses [1preamble] because it discloses an "audio processing apparatus" for "suppressing stationary and non-stationary noises." Li-730, [0019], [0028], FIGS. 2A-2B, 3A. Figure 2A below illustrates the "audio processing apparatus" including a "noise suppressing unit" 250. Li-730, [0019], FIG. 2A.

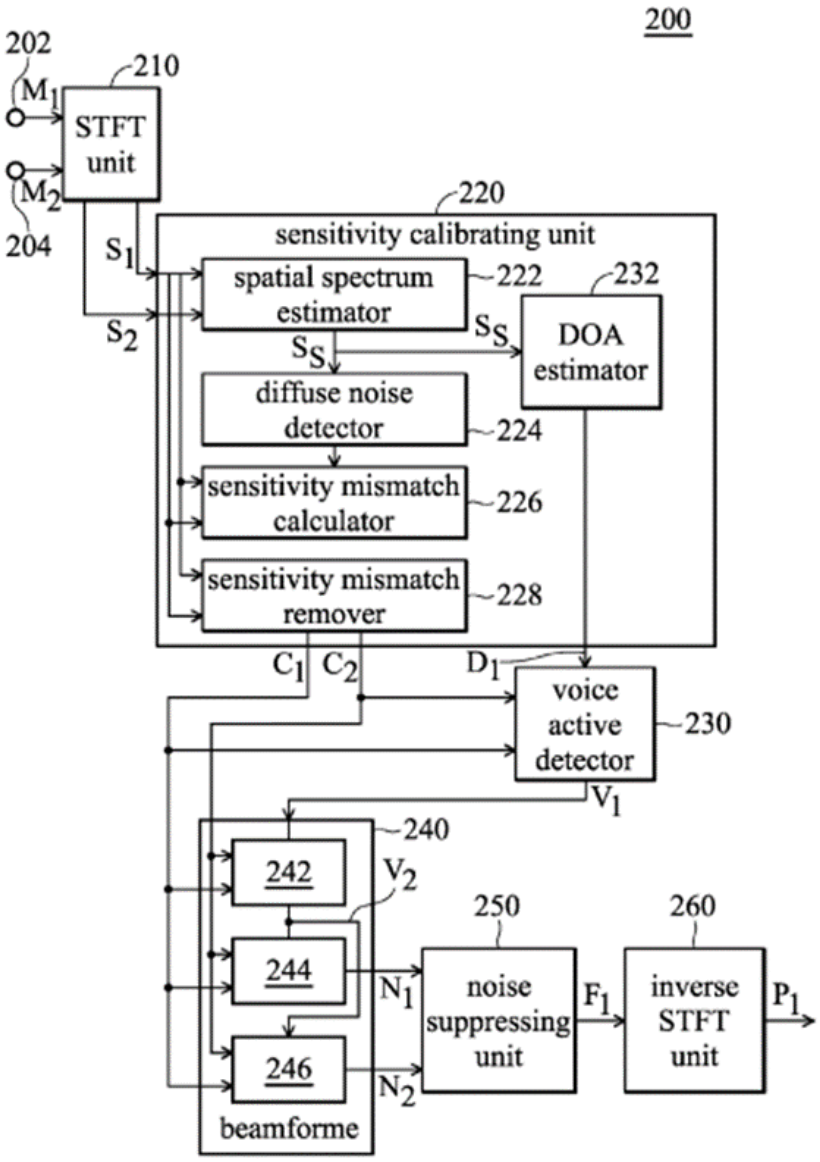


FIG. 2A

Li-730, FIG. 2A.

175. Figure 2B below illustrates a "cell phone" implementing the disclosed audio processing apparatus. Li-730, [0020].

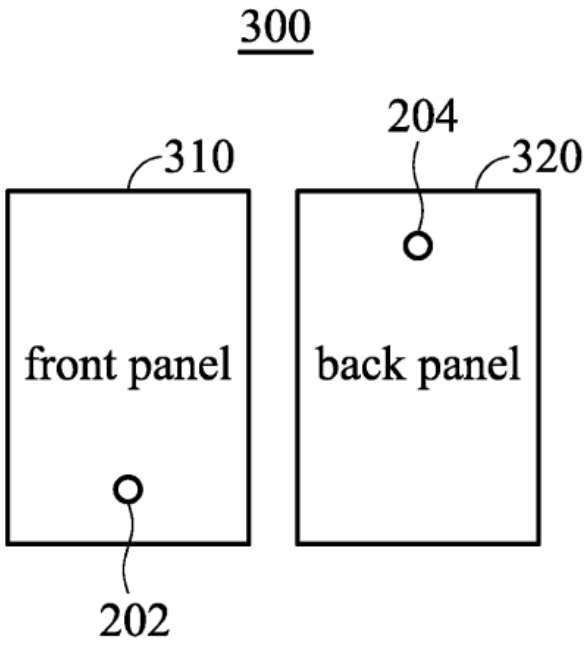


FIG. 2B

Li-730, FIG. 2B.

b. [1a]: "a speech segment determiner configured to determine whether or not a sound picked up by at least either a first microphone or a second microphone is a speech segment and"

176. Li-730 discloses [1a] because its audio processing apparatus includes a "voice active detector (VAD)" (speech segment determiner) "for generating a voice active signal" from calibrated "main" and "reference" sound inputs (M1, M2) picked up by "main" and "reference" microphones 202 and 204, respectively. Li-730, [0007], [0019], [0020], [0025], FIGS. 2A, 3C.

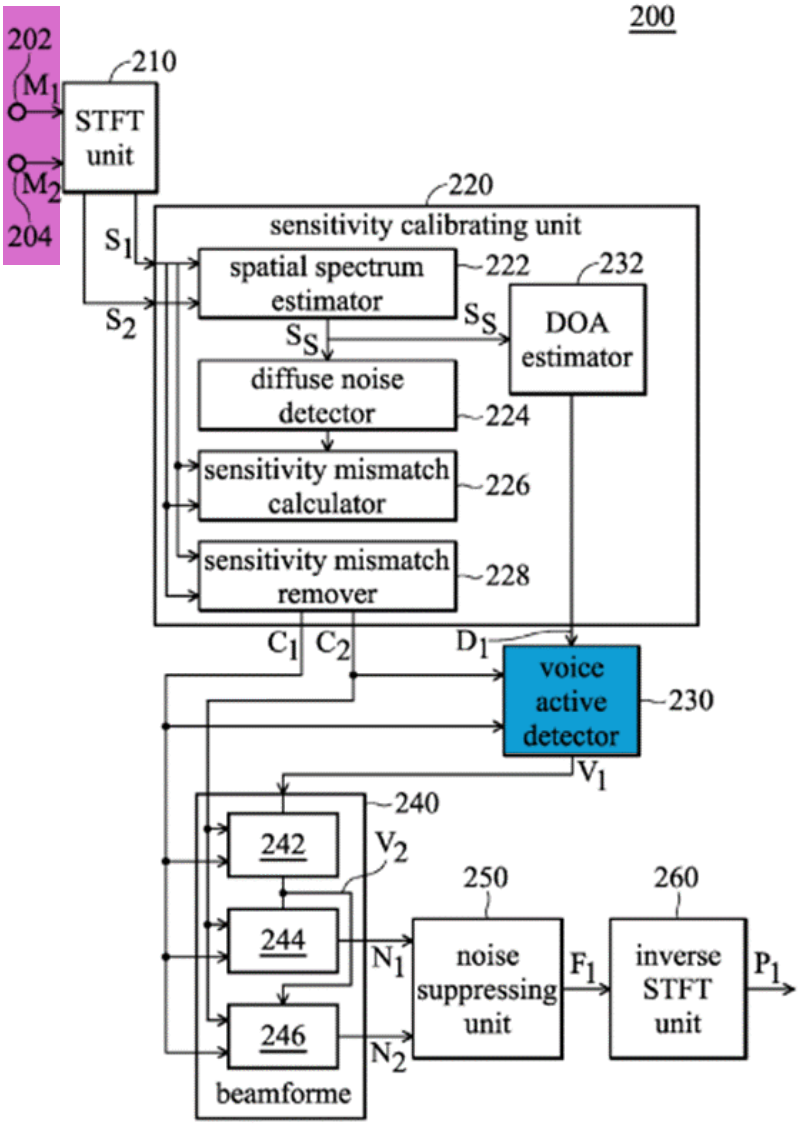


FIG. 2A

Li-730, FIG. 2A (annotated).

177. ***Li-730 discloses at least either a first or second microphone for picking up sound as recited in [1a].*** As illustrated in Figure 2B below, Li-730's apparatus includes a "main microphone" 202 and a "reference microphone" 204

(first and second microphones) which may be disposed along front and back panels of e.g., a cell phone. Li-730, [0020], FIGS. 2A, 2B.

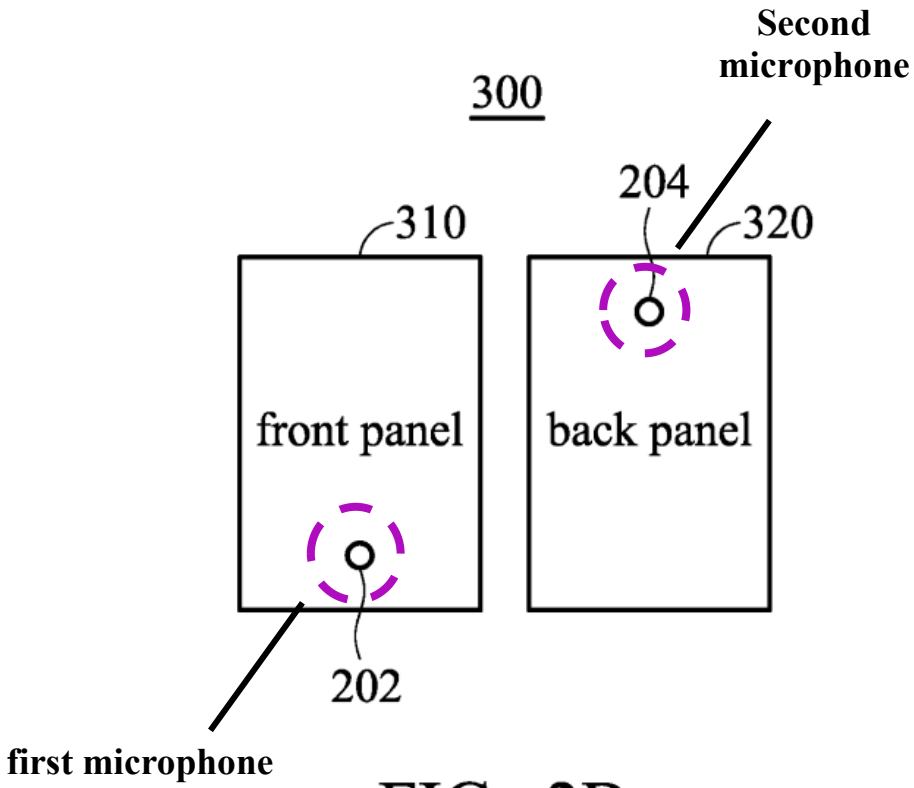
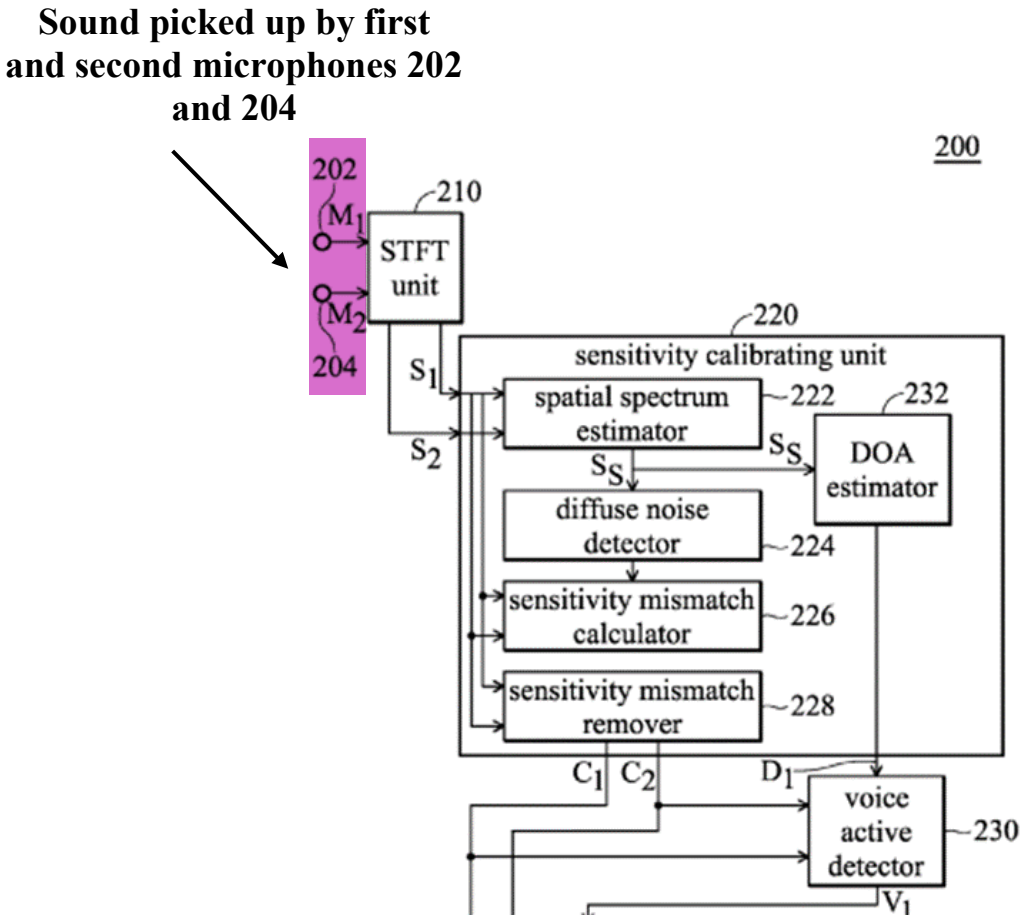


FIG. 2B

Li-730, FIG. 2B (annotated). "The main microphone 202 and the reference microphone 204 are all used to receive sounds"—that is, a "mixture of sounds"—from a "source" (e.g., a "speaker's mouth") and a "non-source" (e.g., "noises"). Li-730, [0020]. These sounds are converted "into a main input M1 and a reference input M2 as shown in FIG. 2." Li-730, [0020], FIG. 2A. Thus, Li-730 discloses sound picked up by at least either a first microphone or a second microphone.



Li-730, FIG. 2A (annotated, excerpted).

178. ***Li-730 discloses a VAD (speech determiner) configured to determine whether or not the sound picked up by the main or reference microphone is a speech segment based on a power ratio associated with the sound.***

179. Li-730 explains that the main and reference sound inputs M1, M2 (picked up sounds) are transformed into frequency domain representations (S1, S2), of the sound inputs M1, M2, calibrated into signals C1 and C2, and input into

the VAD, which uses the calibrated signals "to generate a voice active signal V1" (a determination that the sound is a speech segment) Li-730, [0025].

180. More specifically, Li-730 explains that its short-time Fourier transformation (STFT) unit "converts the main input M1 and the reference input M2 of the time domain signals into a main signal S1 and a reference signal S2 of the frequency domain" and outputs the analog signals to "sensitivity calibrating unit 220." Li-730, [0021]. Once received, the calibrating unit processes the signals (e.g., to generate a "spatial spectrum SS" depicting "the functional relationship between the power distribution and the angles of incident of the main signal and reference signal" and "determine "whether there is a dominant peak in the spatial spectrum SS") and outputs main and reference calibrated signals (C1 and C2 corresponding M1 and M2) to the VAD. Li-730, [0023]-[0025], FIG. 2B.

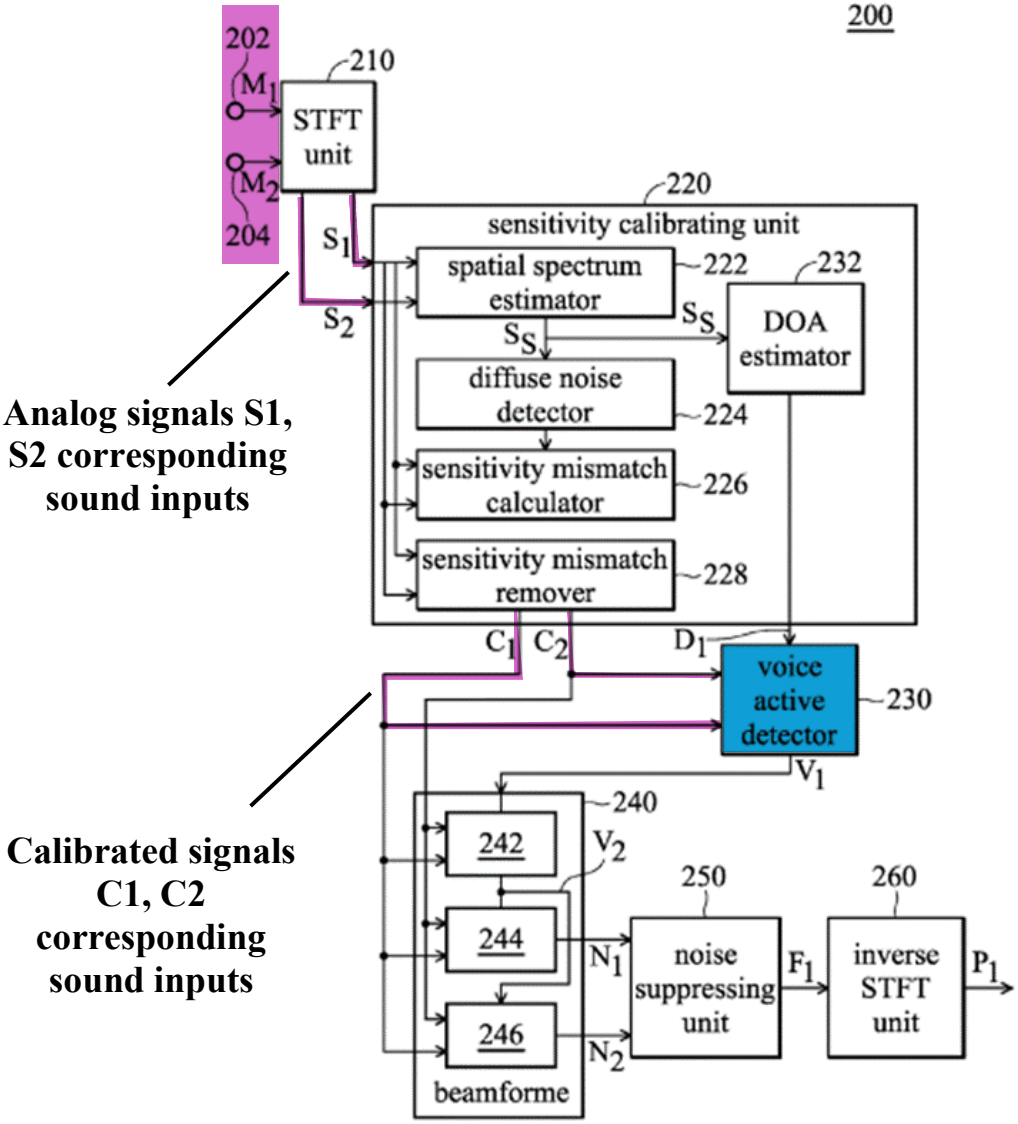


FIG. 2A

Li-730, FIG. 2A (annotated).

181. The VAD then "compares a power ratio between the main calibrated signal C1 and the reference calibrated signal C2 with a predetermined threshold" to determine whether the power ratio exceeds the predetermined threshold for a particular frequency bin. Li-730, [0025]. The results of this comparison indicate

whether a voice (speech segment)[6] is present. If "the power ratio in one bin is smaller than the pre-defined threshold, the signals in that bin may be regarded as noises and may be eliminated." Li-730, [0025]. Conversely, "when the power ratio in one bin is larger than the pre-defined threshold, the signals in that bin may be regarded as *desired signals*," i.e., voice signals, and "may be preserved." Li-730, [0025], [0030], FIG. 3C. As illustrated in step S342 of Figure 3C, this determination also causes a "voice active signal" V1 to turn on. Li-730, [0030], FIG. 3.

[6] The '259 patent indicates that a "speech segment" is a voice input. Ex. 1001, 4:60-66. ("The main microphone 11 and the sub-microphone 12 pick-up a sound including a voice component (speech segment).").

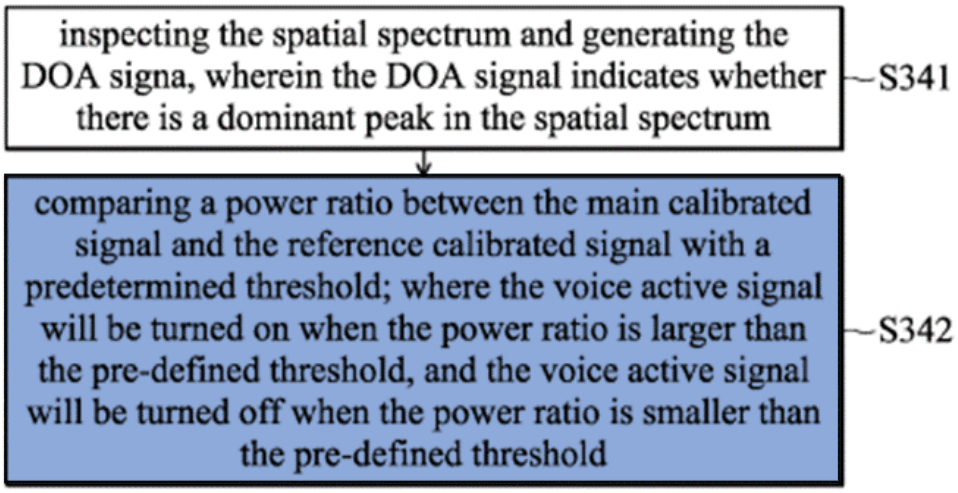


FIG. 3C

Li-730, FIG. 3C (annotated).

182. Li-730's process for identifying a speech segment is similar to the technique used in the '259 patent which also compares signal-to noise ratios (power ratios) to detect a speech segment in certain embodiments. Ex. 1001, 4:59-67 ("[T]echnique I focuses on frequency spectra of a vowel sound that is a main component of a voice sound, to detect a speech segment. In detail, in the speech segment determination technique I, a signal-to-noise ratio is obtained between a peak level of a vowel-sound frequency component and a noise level appropriately set in each frequency band and it is determined whether the obtained signal-to-noise ratio is a specific ratio for a specific number of peaks, thereby detecting a speech segment."); *see also* Ex. 1001, 12:12-20 ("[T]he speech determination

unit 38 []determines whether a per-frame input signal of a target frame is a speech segment, based on the energy ratio.").

183. Accordingly, Li-730 discloses a speech segment determiner as recited in [1a].

c. [1b]: "to output speech segment information when it is determined that the sound picked up by the first or the second microphone is the speech segment;"

184. Li-730 discloses [1b] because as explained for [1a], the VAD (speech segment determiner) outputs a "voice active signal V1" (speech segment information) after determining that at least one of the processed calibrated inputs includes voice (a speech segment). Li-730, [0007], [0019], [0020], [0025], [0030]. FIGS. 2A, 3C. The voice active signal contains information[7] about the speech segment that may be used to ascertain its direction and inform noise suppression, as explained below in [1c]-[1d]. Li-730, [0026]-[0028], [0031].

[7] The '259 patent does not expressly define "speech segment information" but explains that it "indicat[es] that the sound picked up by the first or the second microphone is the speech segment." Ex. 1001, Abstract. Patent Owner also contends in its complaint that speech segment information broadly relates to "details/parameters of the speech segment." Ex. 1014, 9. In my opinion, Li-730's voice active signal V1 satisfies these descriptions.

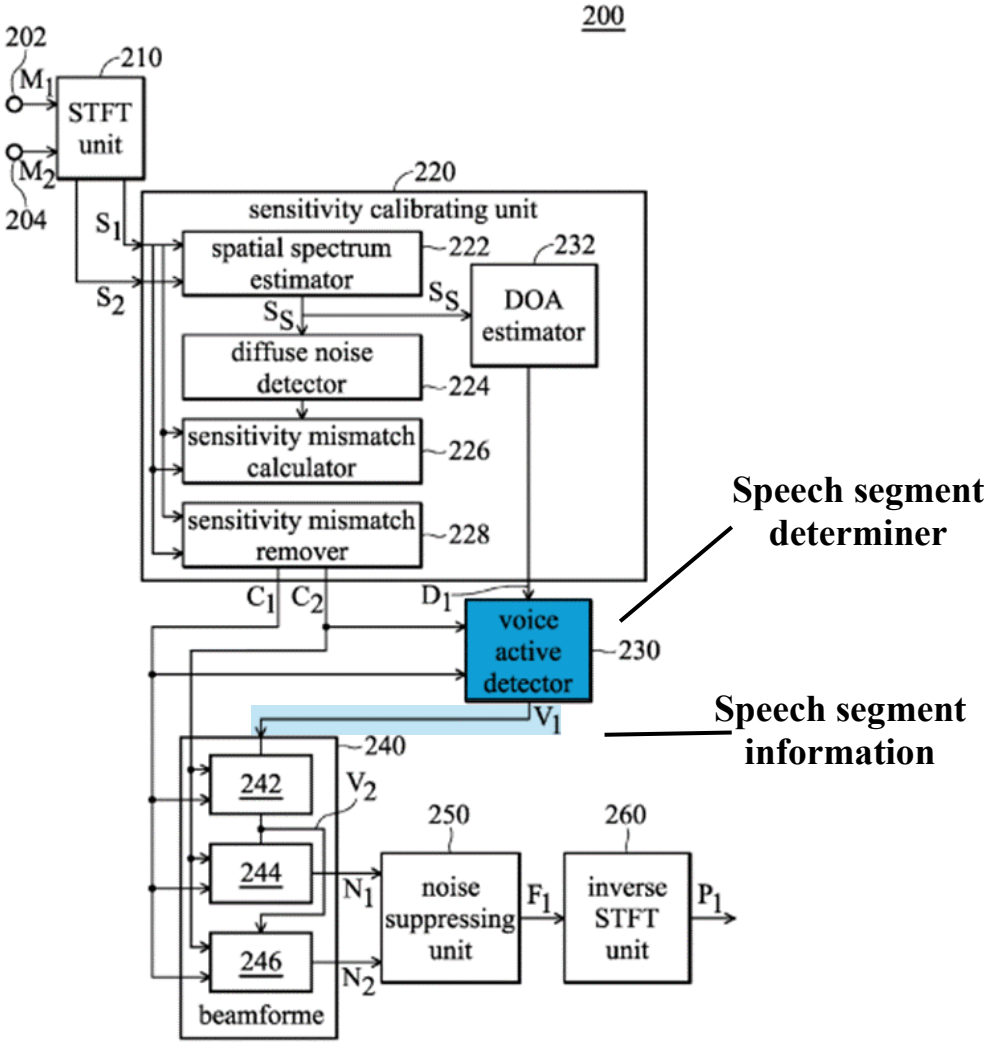


FIG. 2A

Li-730, FIG. 2A (annotated).

d. [1c]: "a voice direction detector configured, when receiving the speech segment information, to detect a voice incoming direction indicating from which direction a voice sound travels,"

185. Li-730 discloses [1c] because its audio processing apparatus further includes a "beamformer" (voice direction detector) configured to determine a

"directional vector" for "each frequency bin according to the voice active signal V1" and apply the directional vector to the voice sound (detect a voice incoming direction indicating from which direction a voice sound travels). Li-730, [0026], FIG. 3D.

186. **First**, Li discloses a beamformer (voice direction detector) configured to receive speech segment information V1, as illustrated in annotated Figure 2A below. Li-730, [025]-[0026], FIG. 2A.

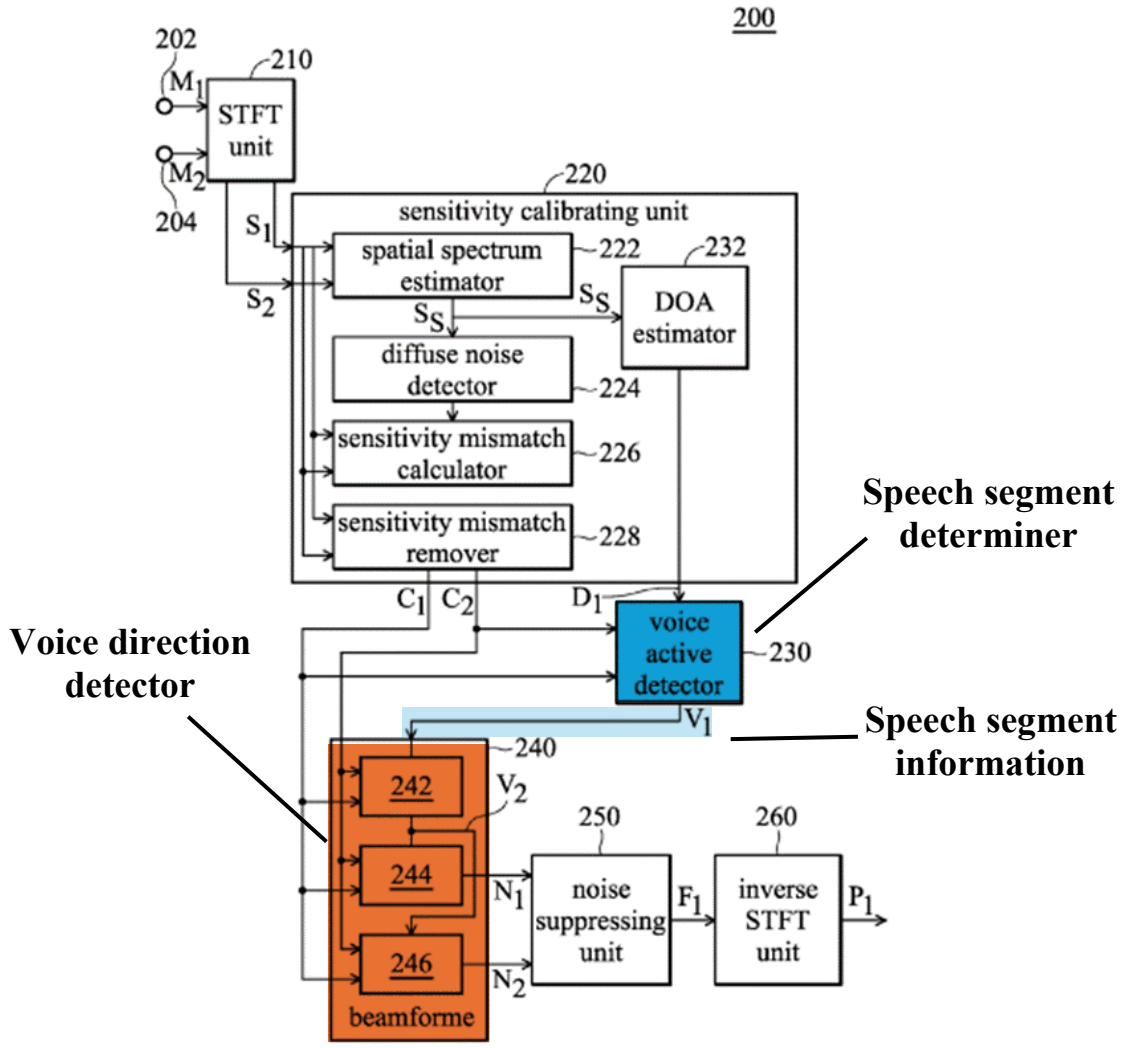


FIG. 2A

Li-730, FIG. 2A (annotated).

187. **Second,** Li-730 discloses that the beamformer uses the speech segment information to generate a steering vector signal V2 that is in turn used to determine the direction from which calibrated signal corresponding to voice sound M1 travels. Li-730, [0026].

188. More specifically, Li-730 explains that the beamformer includes an "array manifold matrix identification unit 242" that implements e.g., a "projection approximation subspace tracking" ("PAST") algorithm to "generate a steering vector signal V2 *according to the voice active signal V1*." Li-730, [0026] (emphasis added), claim 13. This "steering vector signal" V2 "indicates the directional vector" for "each frequency bin according to the voice active signal V1"—in other words, the steering vector determines the direction from which the sounds corresponding to a particular frequency bin travels. Li-730, [0026], [0031].

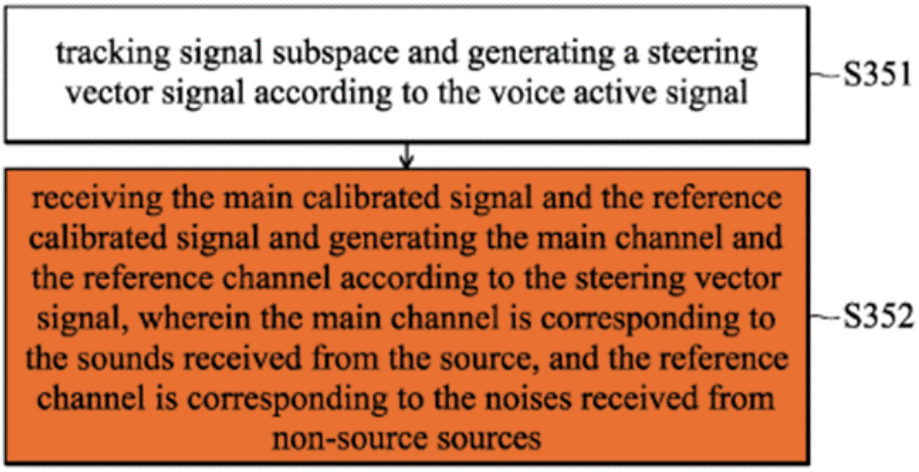


FIG. 3D

Li-730, FIG. 3D (annotated).

e. [1c.i]: "based on a first sound pick-up signal obtained based on a sound picked up by the first microphone and a second sound pick-up signal obtained based on a sound picked up by the second microphone"

189. Li-730 discloses [1c.i] because as explained for [1b] and [1c], the voice incoming direction calculation is based on speech segment information V1, which is based on calibrated signals C1, C2 (first and second sound pick-up signals) corresponding to first and second sounds picked up by the "main" (first) and "reference" (second) microphones. Li-730, [0007], [0026], [0031] FIGS. 2A, 3D.

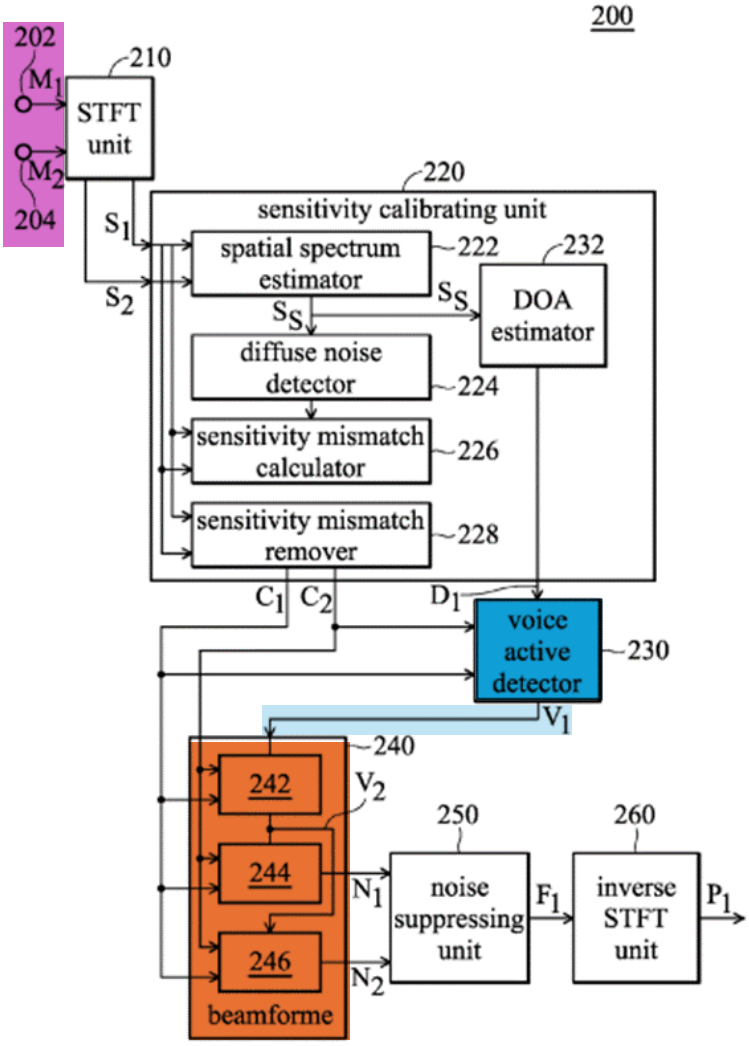


FIG. 2A

Li-730, FIG. 2A (annotated).

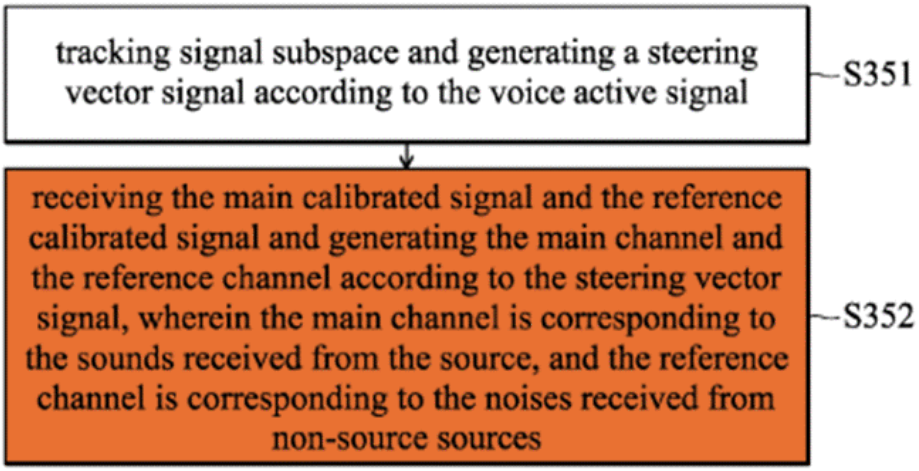


FIG. 3D

Li-730, FIG. 3D (annotated).

f. [1c.ii]: "and to output voice incoming-direction information when the voice incoming direction is detected; and"

190. Li-730 discloses [1c.ii] because the beamformer (voice direction detector) outputs main and reference channels (voice incoming direction information) based on "*steering vector signal V*2" which "indicates [the] directional vector" of the source signal (voice signal). Li-730, [0026], [0031], claim 13, FIGS. 2A, 3C (emphasis added).

191. As illustrated in Figure 2A, the steering vector signal V2 is output to a "main channel calibrator" 244 and "reference channel calibrator 246" where it is then used to generate main channel and reference channel signals N1 and N2.

Li-730, [0026], [0031], FIG. 2A. More specifically, the beamformer applies the steering vector (and a "minimal variance distortion response (MVDR) algorithm") to calibrated voice signal C1 (as described in [1a]) to "generate" a "main channel [signal] N1" "*corresponding to the sounds received from the source*" (voice). Li-730, [0026], FIGS. 2A, 3D (emphasis added); *see also* [0020] (explaining that the "source" originates from "a speaker's mouth").

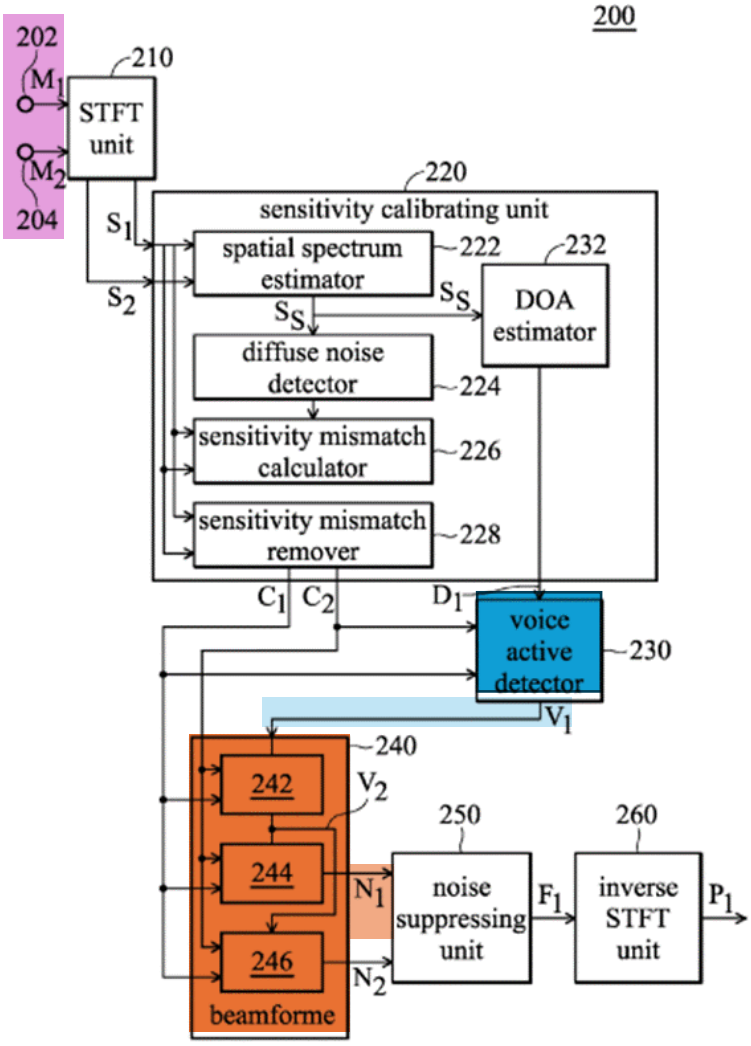


FIG. 2A

Li-730, FIG. 2A (annotated).

g. [1d]: "an adaptive filter configured to perform a noise reduction process using the first and second sound pick-up signals based on the speech segment information and the voice incoming-direction information."

192. Li-730 discloses [1d] because its apparatus includes a "noise suppression unit 250" (an adaptive filter such as a "Weiner post filter") configured

to "suppress stationary and non-stationary noises" (perform a noise reduction process) using "the main calibrated signal C1 and the reference calibrated signal C2" (the first and second sound pick-up signals), "according to the voice active signal V1" (and based on the speech segment information), and the main and reference channels N1 N2 generated using the steering vector V2 (based on the voice incoming direction information). Li-730, [0027], FIG. 2A.

193. ***Li-730's adaptive filter performs noise reduction using the first and second sound pick-up signals.***

194. As discussed above in element 1c.i, Li-730's signals C1 and C2 correspond to the claimed first and second sound pickup signals, respectively. The noise suppression unit 250 uses those signals to reduce noise after some preprocessing. Li-730, [0027]-[0028]. As explained in [1c.ii], the beamformer (voice direction detector) applies a "steering vector" and a "minimal variance distortion response (MVDR) algorithm" to "the main calibrated signal C1 and the reference calibrated signal C2" (the first and second sound pick-up signals) to "generate the main channel N1" and "the reference channel N2." Li-730, [0026]. Main channel N1 "correspond[s] to the sounds received from the source," i.e., the "speaker's voice" and reference signal N2 "corresponding to the noises received from non-source sources." Li-730, [0026], FIGS. 2A, 3D; *see also* [0020] (explaining that the "source" originates from "a speaker's mouth"). Both N1 and

N2 are output to the noise suppression unit (adaptive filter) which "further suppress[es] stationary and non-stationary noises in the main channel N1 and the reference channel N2 according to the voice active signal V1, and integrate[s] the main channel N1 and the reference channel N2 into a final signal F1." Li-730, [0027]-[0028], [0031], FIG. 3A.

195. Thus, as illustrated in annotated Figure 2A below, Li-730's adaptive filter performs noise reduction using the first and second sound pick-up signals (C1, C2) because they are used to generate the main and reference channels N1 and N2 and "[t]he noise suppressing unit 250 is used to further suppress stationary and non-stationary noises in the main channel N1 and the reference channel N2." Li-730, [0027], FIG. 2A, 3A.

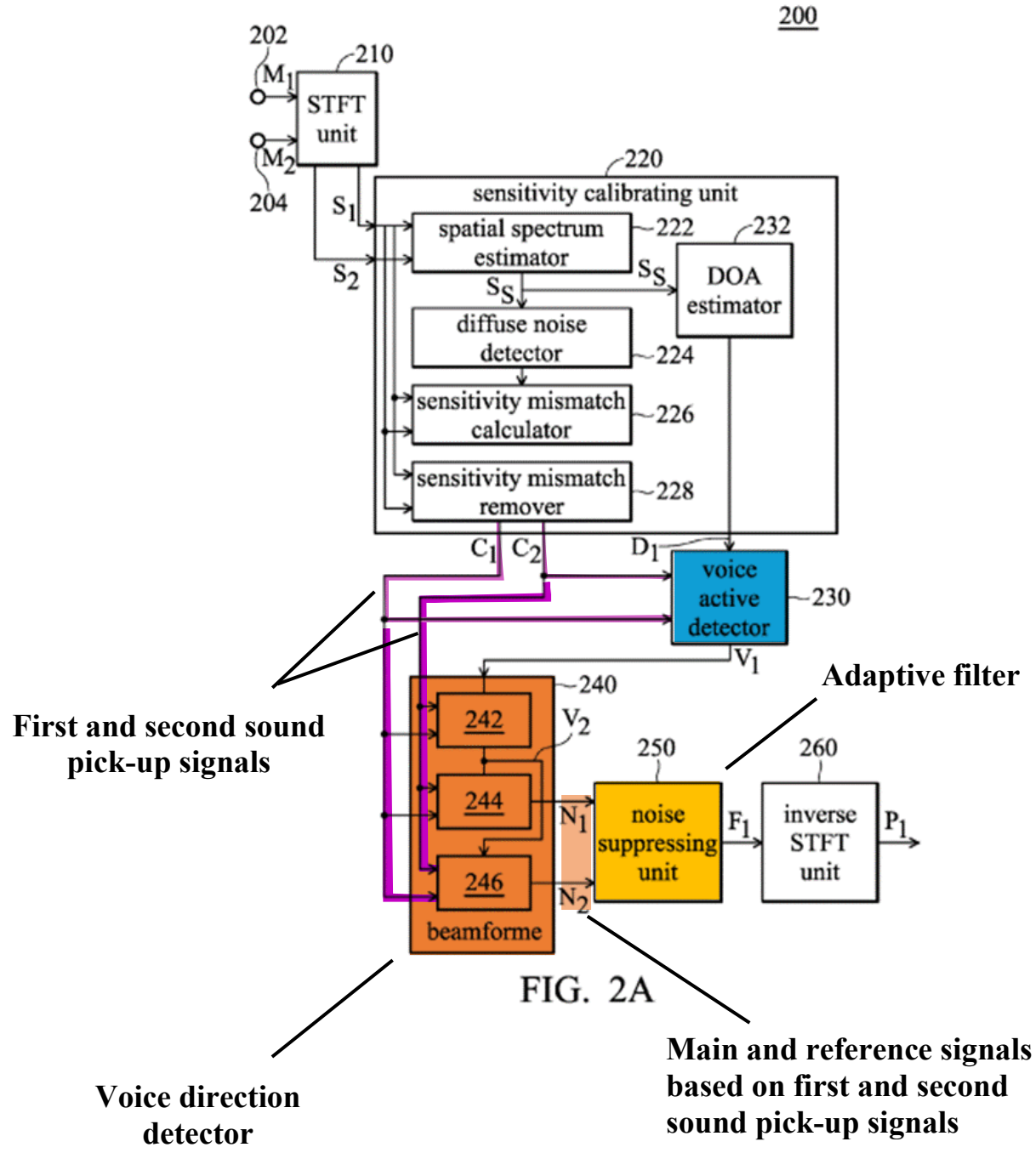


FIG. 2A

Li-730, FIG. 2A (annotated).

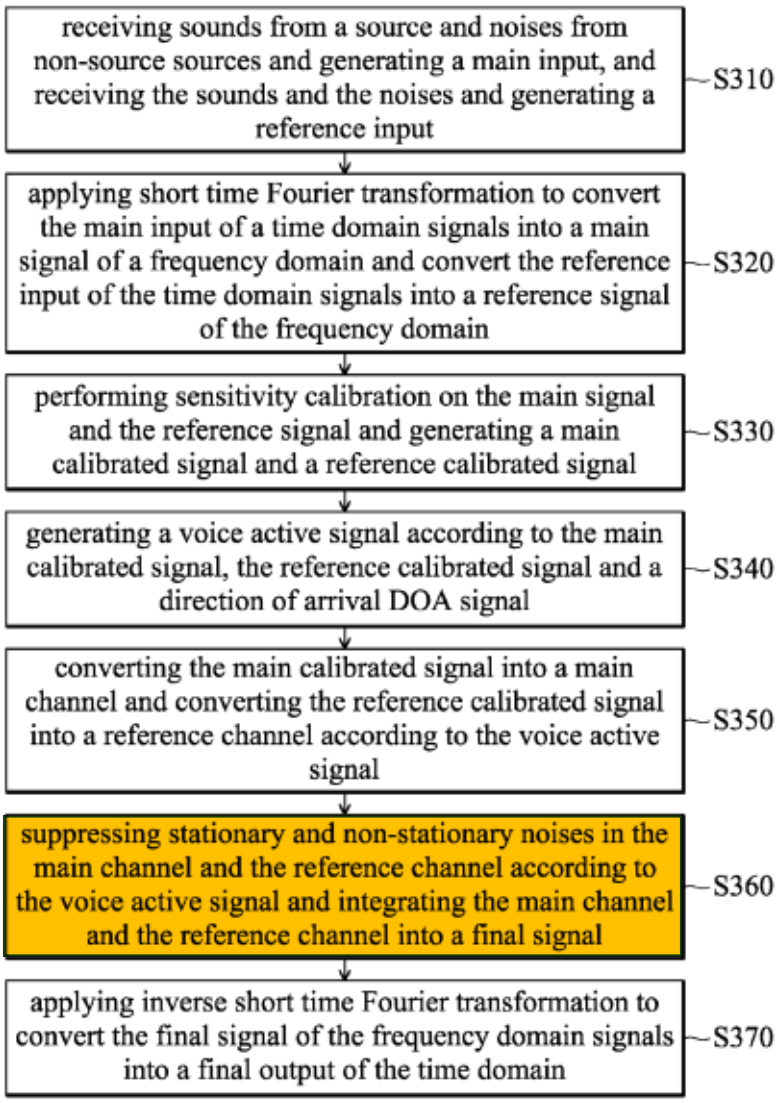


FIG. 3A

Li-730, FIG. 3A (annotated).

196. Noise suppression is also based on first and second pick-up signals C1 and C2 because they are used to generate the speech segment information V1 (as explained in [1a]) on which noise suppression is based, as explained below.

Li-730, [0026]-[0027], FIG. 2A.

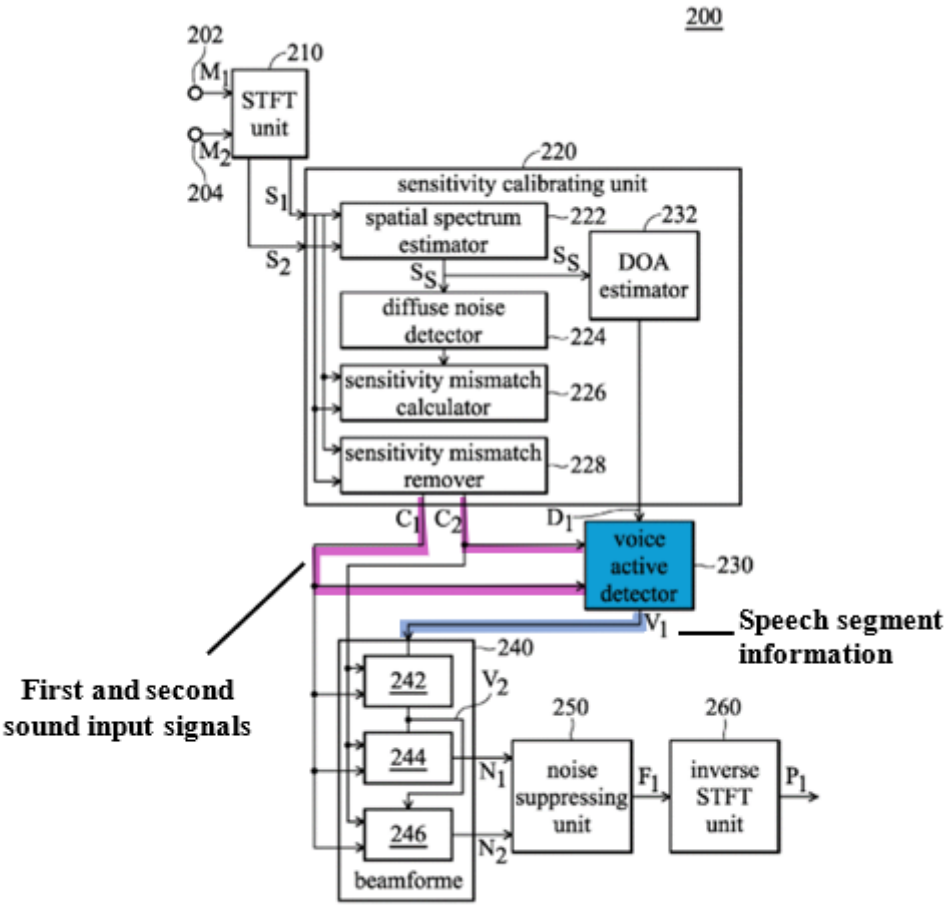


FIG. 2A

Li-730, FIG. 2A (annotated).

197. ***Li-730's adaptive filter performs noise reduction based on the speech segment information V1.***

198. Li-730 further discloses that its noise suppression unit suppresses the stationary and non-stationary noises "according to V1" (the speech segment information as described in [1a]). Li-730, [0026]-[0027], FIG. 2A, 3A.

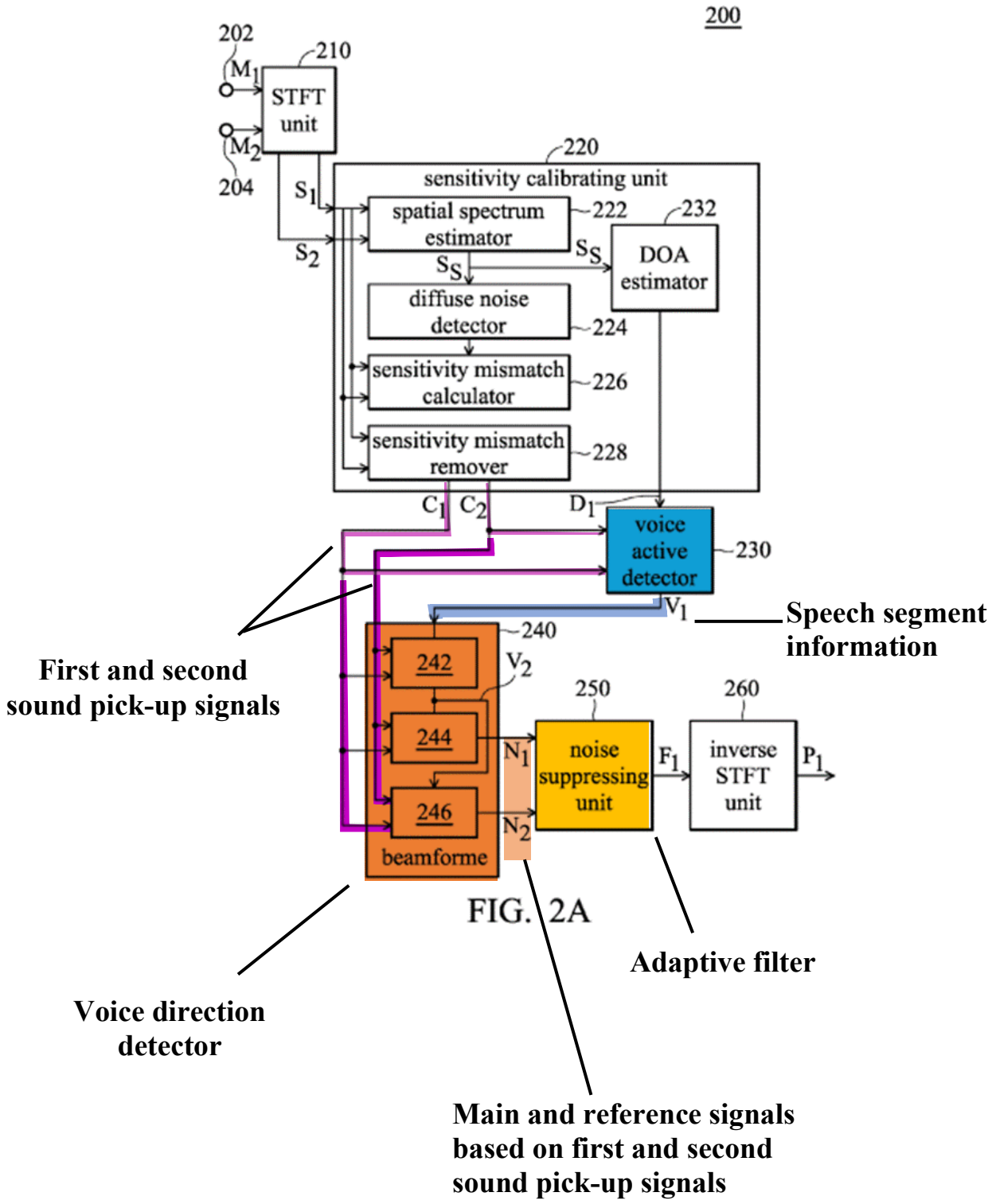


FIG. 2A

Li-730, FIG. 2A (annotated).

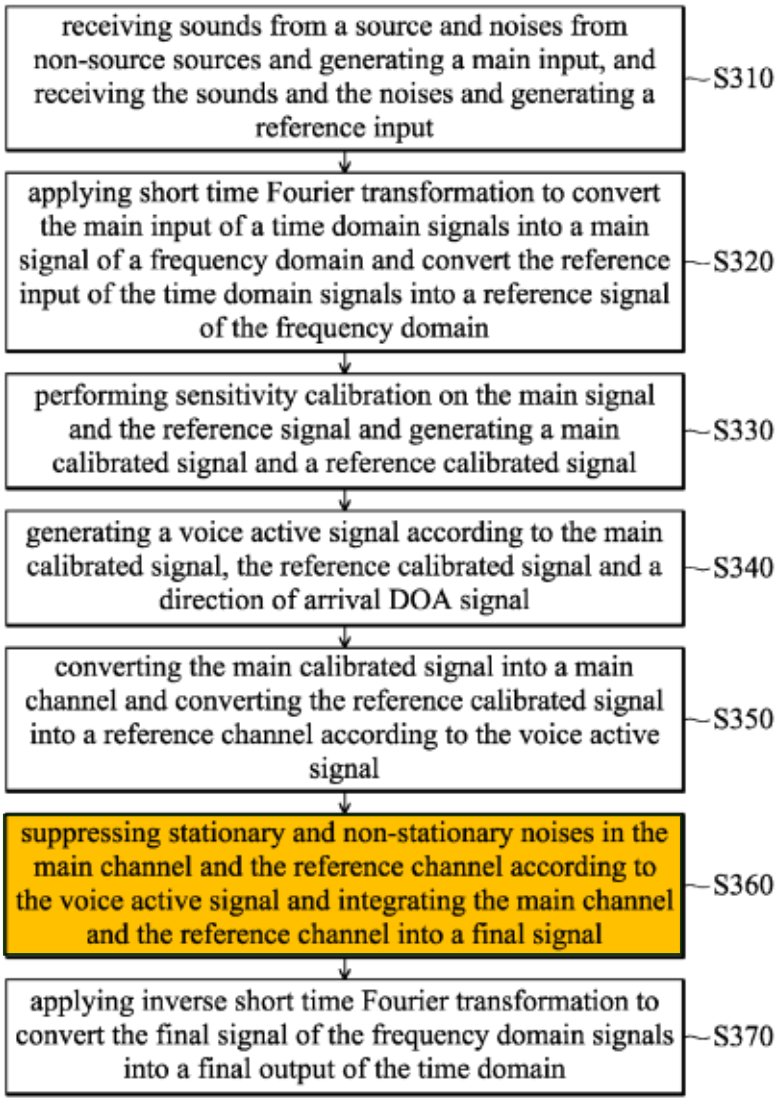


FIG. 3A

Li-730, FIG. 3A (annotated).

199. ***Li-730's adaptive filter performs noise reduction based on the voice incoming direction information N1/ N2.***

200. Li-730's adaptive filter performs noise reduction based on the voice incoming direction information because "[t]he noise suppressing unit 250 is used to further suppress stationary and non-stationary noises in the main channel N1 and the reference channel N2." Li-730, [0027], FIG. 2A, 3A. As explained in 1.c.ii above, main and reference channels (N1 and N2) include voice direction information because, they are generated using "steering vector signal V2" which "indicates [the] directional vector" of the source signal (voice signal) when the "voice active signal" is "turned on" (the power ratio of the source sound is greater than a predetermined threshold as discussed for [1a]). Li-730, [0026], [0031], claim 13, FIGS. 2A, 3C.

2. Claim 2: "The noise reduction apparatus according to claim 1, wherein the voice direction detector detects the voice incoming direction based on a phase difference between the first and second sound pick-up signals."

201. Li-730 discloses claim 2 because as explained for [1c], the beamformer (voice direction detector) uses a "steering vector" to determine the "directional vector at each frequency bin according to the voice active signal V1 which is provided by the VAD 230." Li-730, [0026], [0031], FIGS. 2A, 3D.

202. A POSITA would have understood that the steering vector represents the set of phase-delays (amount of delay) between the calibrated sound inputs C1 and C2 (first and second sound pick-up signals) originating from the first and

second microphones. Accordingly, a POSITA would have understood that the determined directional vector for each of the first and second calibrated signals (sound pick-up signals) is based on the difference between the phases (amount of delay) of the two signals.

3. Claim 3: "The noise reduction apparatus according to claim 2, wherein the adaptive filter performs the noise reduction process to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal

whereas the adaptive filter performs the noise reduction process to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

203. Li-730 discloses claim 3 because its noise suppression unit (adaptive filter) "suppress[es] stationary and non-stationary noises" (performs the noise reduction process) to reduce a noise component carried in "the main calibrated signal C1" or "reference calibrated signal C2" (the first or second sound pick-up signals) corresponding to main and reference channels N1 and N2. Li-730, [0026]-[0028].

204. ***Li-730 discloses reducing a noise component carried by the first sound pick-up signal (C1) using the second sound-pickup signal (C2)*. Li-730, [0027].**

205. As explained in [1d], Li-730's adaptive filter performs noise reduction using the first and second sound pick-up signals (C1, C2) because they are used to generate the main and reference channels N1 and N2 and "[t]he noise suppressing unit 250 is used to further suppress stationary and non-stationary *noises in the main channel N1* and the reference channel N2." Li-730, [0027], FIGS. 2A, 3A (emphasis added). The main channel N1 "correspond[s] to the sounds received from the source" i.e., the "speaker's voice" while reference signal N2 "correspond[s] to the noises received from non-source sources." Li-730, [0026], FIGS. 2A, 3D; *see also* [0020] (explaining that the "source" originates from "a speaker's mouth").

206. Thus, Li-730 discloses reducing a noise component carried by N1 corresponding to C1 (the speaker's voice) using N2 corresponding to C2 (the noises received from non-sources).

207. ***Li-730 discloses that C2 (the second sound pick-up signal) is used to reduce the noise component in C1 (the first sound pick-up signal) when C1 has an advanced phase because C1 corresponds to "source" sounds received from a front microphone that is positioned closer to the speaker's mouth.***

208. As explained in [1a], Li-730's noise reduction apparatus includes a "main microphone" disposed on a front panel of the apparatus. Li-730, [0020]. The

front microphone is typically[8] "closer to the source, e.g,. a speaker's mouth, than the reference microphone 204" and receives sound input M1 which is calibrated into signal C1. Li-730, [0020]-[0024]; *see also* [1a]. A POSITA would have understood that the input source sound received by the front microphone, and thus calibrated signal C1, likewise has an advanced phase relative to the non-source sounds received at the rear reference microphone, as was known in the art. Li-730, [0020]; Konchitsky, [0071] (explaining that sounds received from a "back microphone" are "delayed" relative to the "front microphone" and that the noise from the "delayed back microphone" is "subtracted" "from the front microphone signal.").

209. Further, because "[t]he noise suppressing unit 250 is used to further suppress stationary and non-stationary *noises in the main channel N1*" (generated using C1 as explained in [1d]) and to "integrate the main channel N1 and the reference channel N2 into a final signal"—that is, to suppress noises in main

[8] A POSITA would have understood that this is the case when the device is used in accordance with typical use, e.g., when the speaker is holding the apparatus and speaking into it; but, that there are alternative cases in which the rear microphone could pick up the speaker's voice first, e.g., where the speaker is not holding the device or the rear microphone is more closely aligned with the speaker's mouth.

channel N1 using N2 when N1 contains the source sound—a POSITA would have understood that the noise suppressing unit performs the noise reduction on N1/C1 when C1 has an advanced phase, as was known in the art. Konchitsky, [0057] ("At block 720, the signals are delayed by one sample. At block 730, the delayed rear microphone signal is subtracted from the front microphone signal.").

210. Thus, Li-730 discloses the noise reduction apparatus according to claim 2, wherein the noise suppression unit (adaptive filter) performs the noise reduction process to reduce a noise component carried by C1 (the first sound pick-up signal) using C2 (the second sound pick-up signal) when the first sound pick-up signal is received before the second pickup signal (has a more advanced phase than the second sound pick-up signal).

211. Conversely, and as explained below, a POSITA would have understood that when the rear facing microphone receives the source sound first, i.e., the second sound pick-up signal has the advanced phase, and the front microphones signal is delayed—the noise suppression unit reduces noise contained in C2 using the delayed non-source noise in C1.

212. ***Li-730 discloses reducing a noise component carried by the second sound pick-up signal (C2) using the first sound-pickup signal (C1).***

213. As discussed above, "[t]he noise suppressing unit 250 is used to further suppress stationary and non-stationary noises in the main channel N1 *and*

the reference channel N2." Li-730, [0027], FIG. 2A, 3A (emphasis added). A POSITA would have understood that suppressing noise in the reference signal (N2 corresponding to sound pick-up signal C2) includes using the first sound pick-up signal corresponding to C1/N1. Indeed, because only two sound signals are used in Li-730's noise reduction process (main and reference), a POSITA would have understood that the first signal is the only signal that could be used to suppress noise in N2.

214. Thus, Li-730 discloses reducing a noise component carried by N2 corresponding to C2 (the second sound pick-up signal) using N1 corresponding to C1 (the first sound pick-up signal).

215. ***Li-730 further discloses or renders obvious that C1 (the first sound pick-up signal) is used to reduce the noise component in C2 (the second sound pick-up signal) when C2 has an advanced phase because a POSITA would have understood that C2 corresponds to "source" sounds when the speaker's mouth is closer to the second microphone***.

216. As explained above, although the front microphone is typically "closer to the source, e.g. a speaker's mouth, than the reference microphone 204" (and receives sound input M1 which is calibrated into signal C1), a POSITA would have understood that there are alternative cases in which the rear microphone could

pick-up the speaker's voice first, e.g., where the speaker is not holding the device or the rear microphone is more closely aligned with the speaker's mouth.

217. In such situations, the source (speaker) sound M2 corresponding to C2 would have an advanced phase, and noise components in the C2/N2 signal would be suppressed using C1/N1 with the objective of enhancing the sound of the voice in the C2/N2 signal. Li-730, [0027]. That is, when the rear facing microphone receives the source sound first, i.e., the second sound pick-up signal has the advanced phase, and the front microphones signal is delayed—the noise suppression unit reduces noise contained in C2 using the delayed non-source noise in C1, as was known in the art. *See e.g.,* Konchitsky, [0057] ("The delayed front microphone signal is subtracted from the rear microphone signal.").

4. Independent Claim 12

a. [12preamble]: "An audio input apparatus comprising:"

218. Li-730 discloses [12preamble] because it discloses an "audio processing apparatus," such as a cell phone, for receiving audio inputs (M1, M2) via "main" and "reference" microphones 202, 204. Li-730, [0019], [0020].

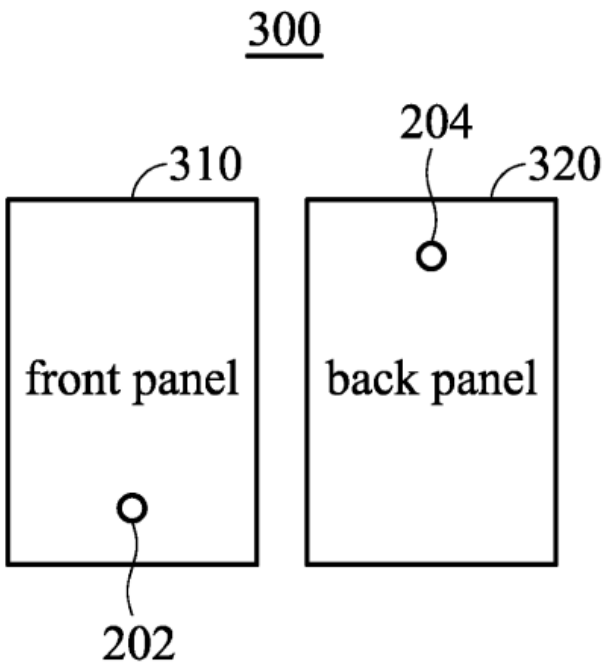


FIG. 2B

Li-730, FIG. 2B.

b. [12a]: "a first face and an opposite second face that is apart from the first face with a specific distance;"

219. Li-730 discloses [12a] because the audio processing apparatus includes a "front panel 310" (also called "front side") and an opposite "back panel 320" (also called "back side") separated by a "physical barrier" (a specific distance). Li-730, [0020], FIG. 2B.

c. [12b]: "a first microphone and a second microphone provided on the first face and the second face, respectively;"

220. Li-730 discloses [12b] because the "main microphone" (first microphone) 202 is on the front panel and the "reference microphone" (second microphone) 204 is on the rear panel. Li-730, [0020], FIG. 2B.

d. [12c]: "a speech segment determiner configured to determine whether or not a sound picked up by at least either the first microphone or the second microphone is a speech segment and"

221. [12c] is substantively similar to [1a]. Li-730 discloses [12c] for reasons discussed for [1a] above.

e. [12d]: "to output speech segment information when it is determined that the sound picked up by the first or the second microphone is the speech segment;"

222. [12d] is substantively similar to [1b]. Li-730 discloses [12d] for reasons discussed for [1b] above.

f. [12e]: "a voice direction detector configured, when receiving the speech segment information, to detect a voice incoming direction indicating from which direction a voice sound travels"

223. [12e] is substantively similar to [1c]. Li-730 discloses [12e] for reasons discussed for [1c] above.

g. [12e.i]: "based on a first sound pick-up signal obtained based on a sound picked up by the first microphone and a second sound pick-up signal obtained based on a sound picked up by the second microphone and"

224. [12e.i] is substantively similar to [1c.i]. Li-730 discloses [12e.i] for reasons discussed for [1c.i] above.

h. [12e.ii]: "to output voice incoming-direction information when the voice incoming direction is detected; and"

225. [12e.ii] is substantively similar to [1c.ii]. Li-730 discloses [12e.ii] for reasons discussed for [1c.ii] above.

i. [12f]: "an adaptive filter configured to perform a noise reduction process using the first and second sound pick-up signals based on the speech segment information and the voice incoming-direction information."

226. [12f] is substantively similar to [1d]. Li-730 discloses [12f] for reasons discussed for [1d] above.

5. Independent Claim 13

a. [13preamble]: "A noise reduction method comprising the steps of:"

227. Li-730 discloses [13preamble] because it discloses a noise reduction method. Li-730, [0028], Fig. 3A.

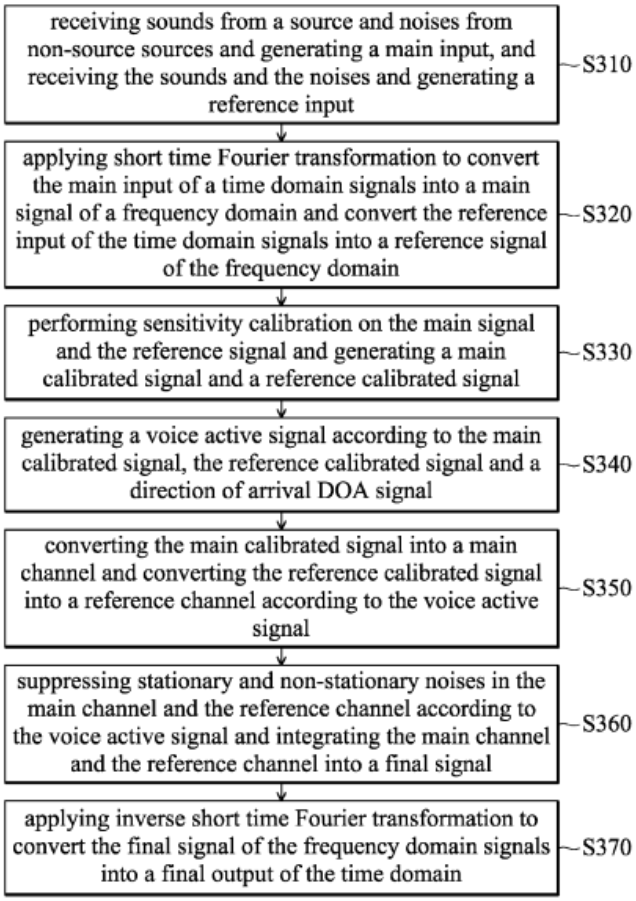


FIG. 3A

Li-730, FIG. 3A.

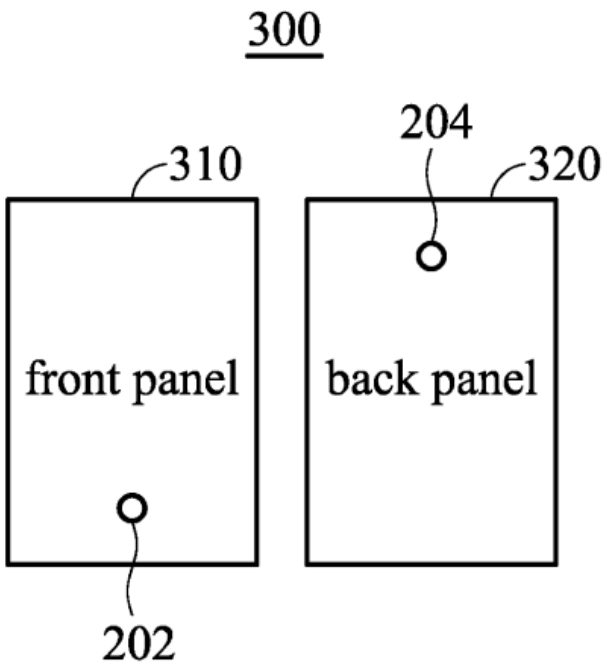


FIG. 2B

Li-730, FIG. 2B.

b. [13a]: "determining whether or not a sound picked up by at least either a first microphone or a second microphone is a speech segment;"

228. [13a] is substantively similar to [1a]. Li-730 discloses [13a] for reasons discussed for [1a] above.

c. [13b]: "detecting a voice incoming direction indicating from which direction a voice sound travels, based on a first sound pick-up signal obtained based on a sound picked up by the first microphone and a second sound pick-up signal obtained based on a sound picked up by the second microphone, when it is determined that the sound picked up by the first or the second microphone is the speech segment; and"

229. Li-730 discloses [13b] for reasons discussed for [1b], [1c], and [1c.i] above.

d. [13c]: "performing a noise reduction process using the first and second sound pick-up signals based on speech segment information indicating that the sound picked up by the first or the second microphone is the speech segment and voice incoming-direction information indicating the voice incoming direction."

230. Claim [13c] is substantively similar to [1d]. Li-730 discloses [13c] for reasons discussed for [1d] above.

6. Claim 14: "The noise reduction method according to claim 13, the voice incoming direction is detected based on a phase difference between the first and second sound pick-up signals."

231. Claim 14 is substantively similar to claim 2. Li-730 discloses claim 14 for reasons discussed for claim 2 above.

7. Claim 15: "The noise reduction method according to claim 14, wherein the noise reduction process is performed to reduce a noise component carried by the first sound pick-up signal using the second sound pick-up signal when the first sound pick-up signal has a more advanced phase than the second sound pick-up signal whereas the noise reduction process is performed to reduce a noise component carried by the second sound pick-up signal using the first sound pick-up signal when the second sound pick-up signal has a more advanced phase than the first sound pick-up signal."

232. Claim 15 is substantively similar to claim 3. Li-730 discloses claim 15 for reasons discussed for claim 3 above.

XI. CONCLUSION

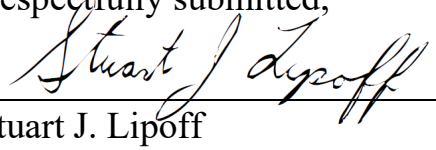
233. For the reasons set forth in Section X, it is my opinion that one skilled in the art would have found claims 1-20 of the '259 Patent anticipated or obvious.

234. 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 this 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.

235. 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 that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Date: February 7, 2025

Respectfully submitted,



Stuart J. Lipoff