IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re *Inter Partes* Review of:	)
U.S. Patent No. 9,226,090	)
Issued: Dec. 29, 2018	)
Application No.: 14/311,532	)
Filing Date: June 23, 2014	)

For: Sound Localization for an Electronic Call

FILED VIA P-TACTS

PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 9,226,090

TABLE OF CONTENTS

I. INTRODUCTION 1

II. MANDATORY NOTICES 2

A. Real Parties-in-Interest 2

B. Related Matters 2

C. Notice of Counsel and Service Information 3

D. Fee for IPR 4

E. Certification for Standing 4

III. CHALLENGES UNDER 35 U.S.C. § 103 5

IV. BACKGROUND 6

A. Technology Overview 6

B. ’090 Patent 10

C. The Challenged Claims 11

D. Prosecution History 11

V. PRIOR ART OVERVIEW 12

A. Bedingfield 12

B. Li 13

C. Andrea 14

D. Ballas 14

E. Sandgren 14

VI. LEVEL OF ORDINARY SKILL IN THE ART 15

VII. CLAIM CONSTRUCTION 15

VIII. GROUND 1: BEDINGFIELD IN VIEW OF ANDREA RENDERS OBVIOUS CLAIMS 1-3, 6, 15-17 16

A. Motivation to Combine 16

B. Reasonable Expectation of Success 18

C. Independent Claim 1 19

1. 1[pre] 19

2. 1[a]: Capture Binaural Sound 19

3. 1[b]: Designate SLP 22

4. 1[c]: Adjust Binaural Sound to Originate From SLP 26

5. 1[d]: Binaural Sound Occurs At SLP 26

D. Dependent Claims 2, 3, 6 27

1. Claim 2 27

2. Claim 3: Receive Location 29

3. Claim 6 30

E. Independent Claim 15 32

1. 15[pre] 32

2. 15[a]: First WED 32

3. 15[b]: Second WED Captures Binaural Voice 33

4. 15[c]: Memory 34

5. 15[d]: Processor 35

6. 15[d1]: Adjust Binaural Voice To SLP 35

7. 15[d2]: SLP Represents Origin 36

F. Dependent Claims 16-17 36

1. Claim 16: SLP At Empty Chair 37

2. Claim 17: Change ITD, ILD, HRTF 38

IX. GROUND 2: BEDINGFIELD IN VIEW OF ANDREA AND BALLAS RENDERS OBVIOUS CLAIM 18 39

A. Motivation to Combine 39

B. Reasonable Expectation of Success 40

C. Dependent Claim 18: Change ITD, ILD, HRTF To Maintain SLP 41

X. GROUND 3: BEDINGFIELD IN VIEW OF ANDREA AND SANDGREN RENDERS OBVIOUS CLAIM 19 43

A. Motivation to Combine 44

B. Reasonable Expectation of Success 45

C. Dependent Claim 19 46

1. 19[a]: Attenuation And Reverberation To Add Artificial Audial Cues 46

2. 19[b]: Virtual Objects 48

XI. GROUND 4: BEDINGFIELD IN VIEW OF BALLAS RENDERS OBVIOUS CLAIMS 8-10, 12, 14 48

A. Motivation to Combine 48

B. Reasonable Expectation of Success 49

C. Independent Claim 8 49

1. **8[pre]** 49

2. **8[a]:** Commence Call 50

3. **8[b]:** Designate SLP 50

4. **8[c]:** Voice Originates From SLP 51

5. **8[d]:** Maintain SLP But Alter ITD 51

D. Dependent Claims 9, 10, 12, 14 52

1. **Claim 9:** In Response To First Individual Moving 52

2. **Claim 10** 53

3. **Claim 12** 54

4. **Claim 14:** Receive Commands 55

XII. GROUND 5: LI IN VIEW OF ANDREA RENDERS OBVIOUS CLAIMS 1, 3, 6, 15-17 56

A. Motivation to Combine 56

B. Reasonable Expectation of Success 57

C. Independent Claim 1 58

1. **1[pre]** 58

2. **1[a]** 58

3. **1[b]** 60

4. **1[c]** 63

5. **1[d]** 65

D. Dependent Claims 3, 6 66

1. **Claim 3** 66

2. Claim 6 66

E. Independent Claim 15 68

1. 15[pre] 68

2. 15[a] 68

3. 15[b] 68

4. 15[c] 68

5. 15[d] 69

6. 15[d1] 69

7. 15[d2] 70

F. Dependent Claims 16-17 70

1. Claim 16 70

2. Claim 17 71

XIII. GROUND 6: LI IN VIEW OF ANDREA AND BALLAS RENDERS OBVIOUS CLAIMS 2, 18 73

A. Motivation to Combine 73

B. Reasonable Expectation of Success 74

C. Dependent Claim 2 76

1. 2[a] 76

2. 2[b] 76

D. Dependent Claim 18 76

XIV. GROUND 7: LI IN VIEW OF ANDREA AND SANDGREN RENDERS OBVIOUS CLAIM 19 77

A. Motivation to Combine 77

B. Reasonable Expectation of Success 79

C. Dependent Claim 19 80

1. Claim 19[a] 80

2. Claim 19[b] 80

XV. GROUND 8: LI IN VIEW OF BALLAS RENDERS OBVIOUS CLAIMS 8-10, 12, 14 80

A. Motivation to Combine 80

B. Reasonable Expectation of Success 81

C. Independent Claim 8 81

1. 8[pre] 81

2. 8[a] 81

3. 8[b] 82

4. 8[c] 82

5. 8[d] 83

D. Dependent Claims 9, 10, 12, 14 83

1. Claim 9 83

2. Claim 10 84

3. Claim 12 85

4. Claim 14 86

XVI. SECONDARY CONSIDERATIONS 86

XVII. THE BOARD SHOULD REACH THE MERITS OF THIS PETITION 86

XVIII. CONCLUSION 87

TABLE OF AUTHORITIES

Page(s)

Cases

Apple Inc. v. AGIS Software Development, LLC,
IPR2018-00819, Paper 9 (Nov. 7, 2018) 16

Cienna Corp. v. Oyster Optics, LLC,
IPR2021-00565, Paper 11 (Aug. 23, 2021) ..... *passim*

Eight KHZ, LLC v. Meta Platforms, Inc. et al.,
2:22-cv-000575 (W.D. Tex.) 3

Eight kHz, LLC v. Meta Platforms, Inc.,
No. 6:22-cv-00575-ADA-DTG (W.D. Tex.) 87

Western Digital Corp. v. SPEX Techs., Inc.,
IPR2018-0084, Paper 14 (April 25, 2018) 16

Statutes

35 U.S.C. §§ 102(a)(1) 5

35 U.S.C. § 102(a)(1)/(2) 5

35 U.S.C. § 102(a)(2) 5, 6

35 U.S.C. § 103 ..... 5, 8, 9, 10

35 U.S.C. § 314(a) 87

Other Authorities

37 C.F.R. § 42.10(b) 4

37 C.F.R. § 42.15(a) 4

37 C.F.R. § 42.24 94

37 C.F.R. § 42.24(a)(1) 94

37 C.F.R. § 42.104(a) 5

37 C.F.R. § § 42.24(a)(1)(i) 94

EXHIBIT LIST

Exhibit	Description
1001	U.S. Patent No. 9,226,090 ("'090 Patent")
1002	Prosecution History of U.S. Patent No. 9,226,090 ("'090 FH")
1003	Declaration of Dr. Kyriakakis
1004	*Curriculum Vitae* of Dr. Kyriakakis
1005	U.S. Patent No. 8,406,439 to Bedingfield et al. ("Bedingfield")
1006	U.S. Patent No. 9,131,305 to Li et al. ("Li")
1007	U.S. Patent No. 8,818,000 to Andrea ("Andrea")
1008	U.S. Patent Pub. No. 2003/0059070 to Ballas ("Ballas")
1009	U.S. Patent No. 7,606,375 to Bailey et al. ("Bailey")
1010	RESERVED
1011	RESERVED
1012	RESERVED
1013	U.S. Patent App. Pub. 2014/0064526 to Otto, et al. ("Otto")
1014	RESERVED
1015	Rumsey, Francis "Whose head is it anyway? Optimizing binaural audio." J. Audio Eng. Soc., Vol. 59, No. 9, 2011 ("Rumsey")
1016	RESERVED
1017	U.S. Patent App. Pub. 2008/0273708 to Sandgren, et al. ("Sandgren")
1018	RESERVED

Exhibit	Description
1019	RESERVED
1020	Scheinerman, Matt "Exploring Augmented Reality." Haverford College Computer Science (2009) ("Scheinerman")
1021	Comeau, Charles et al. "Headsight TV System." Electronics, 1961, pp. 86-90 ("Comeau")
1022	Sutherland, Ivan "The Ultimate Display." Proceedings of IFIP Congress, pp. 506-508 (1965) ("Sutherland1965")
1023	Sutherland, Ivan "A head-mounted three dimensional display." Fall Joint Computer Conference, pp. 757-764 (1968) ("Sutherland1968")
1024	Moller, Henrik, et al. "Head-Related Transfer Functions of Human Subjects." J. Audio Eng. Soc., Vol. 43, No. 5, 1995, pp. 300-321 ("Moller")
1025	Akeroyd, Michael. "The psychoacoustics of binaural hearing." International Journal of Audiology, 45:sup1, 25-33 (2006) ("Akeroyd")
1026	Middlebrooks, John, et al. "Sound Localization by Human Listeners." Annual Review of Psychology, 1991. 42:135-39 ("Middlebrooks")
1027	*Eight KHZ, LLC v. Meta Platforms, Inc. et al.*, 6:22-cv-575-ADA-DTG (W.D. Tex.), Plaintiff's Claim Construction Reply Brief (ECF 92) ("8KHZ Claim Construction Brief")
1028	RESERVED
1029	RESERVED
1030	RESERVED
1031	Cheng, Corey, et al. "Moving Sound Source Synthesis for Binaural Electroacoustic Music Using Interpolated Head-Related Transfer Functions (HRTFs)." ("Cheng")
1032	May 22, 2023 Deposition Transcript of Durand Begault, Ph.D. from *Eight KHZ, LLC v. Meta Platforms, Inc. et al.*, 6:22-cv-575-ADA-DTG (W.D. Tex.)
1033	U.S. Patent No. 257,453 to Ader("Ader")
1034	Olson, Harry. "Sound Reproducing Systems - Monaural, Binaural, Monophonic, and Stereophonic." Audio, 1958 ("Olson")

Exhibit	Description
1035	Kyriakakis, Chris. "Fundamental and Technological Limitations of Immersive Audio Systems." Proceedings of the IEEE, Vol. 86, No. 5, 1998 ("Kyriakakis 1998")
1036	Begault, Durand. "3-D Sound for Virtual Reality and Multimedia." 2000 ("Begault")
1037	Vertesi, Janet. "Seamful Space: Heterogeneous Infrastructures in Interaction." Science, Technology, & Human Values 2014, Vol. 39:2, 264-284 ("Vertesi")
1038	Torrejon, Alfonso, et al. "Improving Communication and Presence in Online Telepresence Systems." Workshop Proceedings of the 9th International Conference on Intelligent Environments, 2013 ("Torrejon")
1039	RESERVED
1040	RESERVED
1041	RESERVED
1042	RESERVED
1043	RESERVED
1044	U.S. Patent No. 2,955,156 to Heilig ("Heilig")

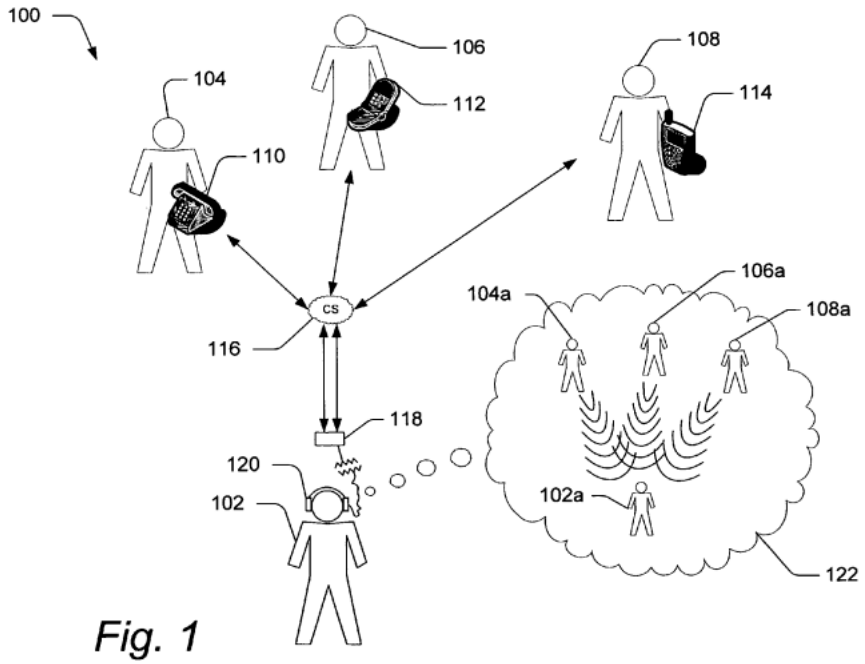
I. INTRODUCTION

Meta Platforms, Inc. ("Meta" or "Petitioner") respectfully requests *inter partes* review ("IPR") of claims 1-3, 6, 8-10, 12, 14-19 of U.S. Patent 9,226,090 (EX1001, "'090 Patent").

The foundational technology underlying virtual reality was developed in the 1960s with the Telesphere Mask, Headsight, and Ultimate Display. As virtual reality advanced, so did binaural sound simulation. Humans perceive sounds binaurally because both ears receive auditory stimuli, allowing the brain to compare the sound from each ear to determine the sound's source location. By 2014, capturing, processing, and generating binaural sound so a listener localizes its origin at a sound localization point, or SLP (i.e., point from where sound originates) were well-known, well-understood, and routinely used.

The '090 Patent simulates a face-to-face conversation during a call by positioning one participant's voice at an SLP near another participant. But rather than disclose some new innovation, the '090 Patent merely describes known wearable electronic devices (WEDs) applying known techniques during calls. Notably, nowhere does the '090 Patent describe the difference between the claimed subject matter and the prior art. Indeed, there is no patentable difference. For example, in 2007, AT&T filed for a patent titled "Synthetic Audio Placement" on a wearable device that simulates a conversation between remote users as if they were

"in the same location and speaking to each other directly."



EX1005, Figure 1

Accordingly, Petitioner respectfully requests that the Board institute review and find all challenged claims of the '090 Patent are unpatentable.

II. MANDATORY NOTICES

A. Real Parties-in-Interest

The real parties-in-interest are Meta Platforms, Inc., Meta Platforms Technologies, LLC, and Twisted Pixel Games, LLC.

B. Related Matters

U.S. Patent Office records indicate that the '090 Patent is assigned to Eight KHZ, LLC, ("PO"), which is currently asserting the '090 Patent in the following

concurrent litigation filed on June 3, 2022: *Eight KHZ, LLC v. Meta Platforms, Inc. et al.*, 2:22-cv-000575 (W.D. Tex.).

Petitioner has filed, at substantially the same time that this Petition was filed, petitions for *inter partes* review against related family members U.S. Patent Nos. 9,282,196 and 9,674,628, both of which are also asserted in the concurrent litigation. Petitioner has also filed, at substantially the same time that this Petition was filed, petitions for *inter partes* review against the following other patents asserted in the concurrent litigation: U.S. Patent No. 10,448,184; U.S. Patent No. 11,172,316; U.S. Patent No. 11,290,836; U.S. Patent No. 10,368,179; U.S. Patent No. 10,798,509; and U.S. Patent No. 10,917,737.

C. Notice of Counsel and Service Information

Lead Counsel	Back-Up Counsel
Lisa K. Nguyen (Reg. No. 58,018) lisa.nguyen@allenovery.com Postal & Hand-Delivery Address: Allen & Overy LLP 550 High Street Palo Alto, CA 94301 Telephone: (650) 388-1724	David M. Tennant (Reg. No. 48,362) david.tennant@allenovery.com Postal & Hand-Delivery Address: Allen & Overy LLP 1101 New York Ave NW Washington, DC 20005 Telephone: (202) 683-3891 Grace I. Wang (Reg. No. 69,892) grace.wang@allenovery.com Postal & Hand-Delivery Address: Allen & Overy LLP 1221 Ave of the Americas New York, NY 10020 Tel: (212) 756-1143

	David A. Hubbard (Reg. No. 73,621) david.hubbard@allenovery.com Postal & Hand-Delivery Address: Allen & Overy LLP 1221 Ave of the Americas New York, NY 10020 Tel: (212) 610-6357 Alan M. Billharz (Reg. No. 79,532) alan.billharz@allenovery.com Postal & Hand-Delivery Address: Allen & Overy LLP 1101 New York Avenue, NW Washington, DC 20005 Tel: (202) 683-3862 Eric E. Lancaster (*pro hac vice* to be filed) eric.lancaster@allenovery.com Postal & Hand-Delivery Address: Allen & Overy LLP 550 High Street Palo Alto, CA 94301 Telephone: (650) 388-1700

A Power of Attorney is being filed concurrently with this Petition in accordance with 37 C.F.R. § 42.10(b). Petitioner consents to electronic service by e-mail, using the email addresses identified above.

D. Fee for IPR

The Director is authorized to charge the fee(s) per 37 C.F.R. § 42.15(a) to Deposit Account 60-4184.

E. Certification for Standing

Petitioner certifies pursuant to 37 C.F.R. § 42.104(a) that the '090 Patent is available for IPR and that Petitioner is not barred or estopped from requesting an IPR challenging the claims on the grounds identified in herein.

III. CHALLENGES UNDER 35 U.S.C. § 103

Grounds	Claims	Prior Art
1	1-3, 6, 15-17	Bedingfield[1] in view of Andrea[2]
2	18	Bedingfield in view of Andrea and Ballas[3]
3	19	Bedingfield in view of Andrea and Sandgren[4]
4	8-10, 12, 14	Bedingfield in view of Ballas

[1] U.S. Patent 8,406,439 to Bedingfield et al. (EX1005, "Bedingfield") is prior art under at least 35 U.S.C. §§ 102(a)(1) and 102(a)(2).

[2] U.S. Patent 8,818,000 to Andrea (EX1007, "Andrea") is prior art under at least 35 U.S.C. § 102(a)(2).

[3] U.S. Patent Pub. 2003/0059070 to Ballas (EX1008, "Ballas") is prior art under at least 35 U.S.C. §§ 102(a)(1) and 102(a)(2).

[4] U.S. Patent Pub. 2008/0273708 to Sandgren et al. (EX1017, "Sandgren") is prior art under at least 35 U.S.C. § 102(a)(1)/(2).

5	1, 3, 6, 15-17	Li[5] in view of Andrea
6	2, 18	Li in view of Andrea and Ballas
7	19	Li in view of Andrea and Sandgren
8	8-10, 12, 14	Li in view of Ballas

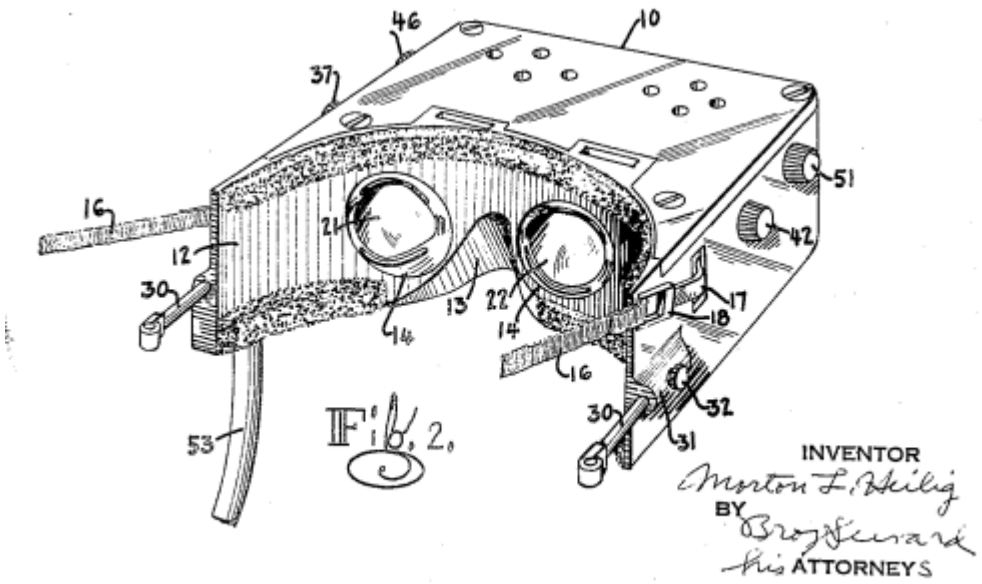
IV. BACKGROUND

A. Technology Overview

Wearable virtual reality dates back to the 1960s, when early pioneers invented head mounted displays ("HMDs") to provide immersive and realistic simulations. EX1003 ("Kyriakakis") ¶¶38-42. The earliest WEDs had displays synchronized with sound. More sophisticated devices adjusted the displays and sound in response to head tracking. Kyriakakis ¶38.

In 1960, Morton Heilig patented the first HMD having earphones and a screen for each eye. EX1044, Fig. 2 reproduced below; EX1020, 4; Kyriakakis ¶39.

[5] U.S. Patent 9,131,305 to Li et al. (EX1006, "Li") is prior art under at least 35 U.S.C. § 102(a)(2).



In 1961, Charles Comeau and James Bryan developed "Headsight," the first HMD to incorporate head tracking and provide images and sounds synchronized with the wearer's head movements. EX1020, 5; EX1021, 3-7; Kyriakakis ¶40. Later, Ivan Sutherland described an HMD that presented a three-dimensional image that "change[d] in exactly the way that the image of a real object would change for similar motions of the user's head." EX1023, 1; *see also* EX1022, 2; Kyriakakis ¶¶41-42.

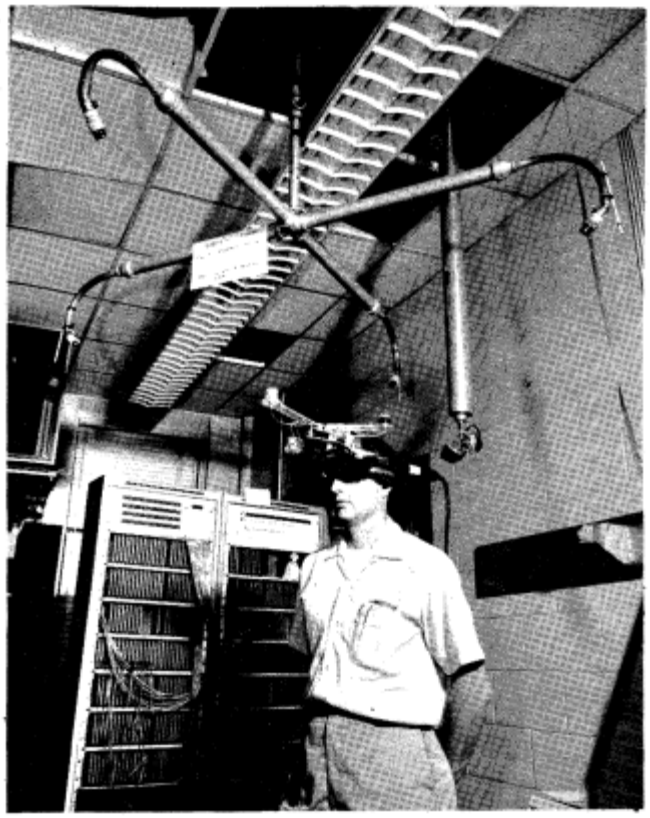
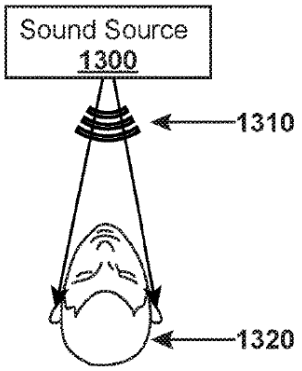


FIGURE 3—The mechanical head position sensor in use

As wearable HMD technology developed, so did binaural sound simulation. Kyriakakis ¶¶43-69. To determine sound source location, humans process and compare sound cues at each ear. *Id*. For a sound source directly in front of the human listener, the sound generally arrives at each ear at the same time. Kyriakakis ¶50.



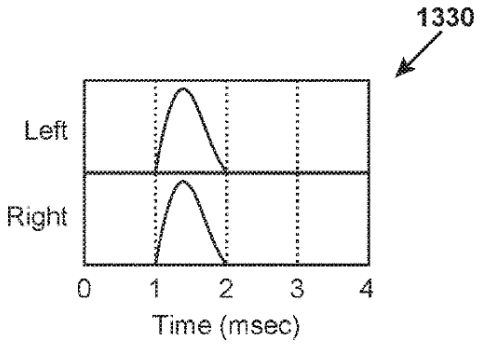
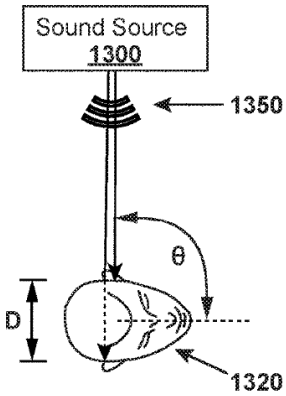


Figure 13A **Figure 13B**

EX1001, Figs. 13A-13B

But when the sound source is to the left, the sound arrives at the left ear slightly before the right ear, and the sound at the right ear will be less intense. *Id*. The difference in arrival time across the ears is the interaural time difference (ITD). *Id*.



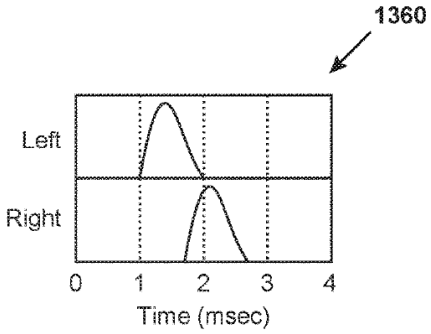


Figure 13C **Figure 13D**

EX1001, Figs. 13C-13D

Other factors impact how an ear receives sound: reverberation, reflection, refraction, and attenuation. Kyriakakis ¶52. Even the shape of the listener's ears, head, shoulders, and torso can change the sound. *Id*.

To capture these factors, for decades, a head related transfer function (HRTF) has been used to characterize how an ear receives a sound from a point in space. EX1024, 1; EX1025, 5; Kyriakakis ¶53. To create a "virtual auditory space," sounds are processed (convolved in the time domain) with a left HRTF and a right HRTF and played over headphones. EX1025, 5; Kryiakakis ¶¶54-55. Interpolation of HRTFs can also be used to mimic moving sound sources. EX1031. Indeed, by 2001, "sound spatialization and moving sound synthesis [were] both popular and widely accessible sound-sculpting tools." *Id.*

B. '090 Patent

The '090 Patent describes well-known concepts of providing sound "comparable in quality to sound that listeners would hear if they were communicating face-to-face with each other." EX1001, 1:14-16. The system captures a speaker's binaural voice, designates a SLP, adjusts the binaural voice to originate from the SLP, and provides the adjusted binaural voice to a listener.

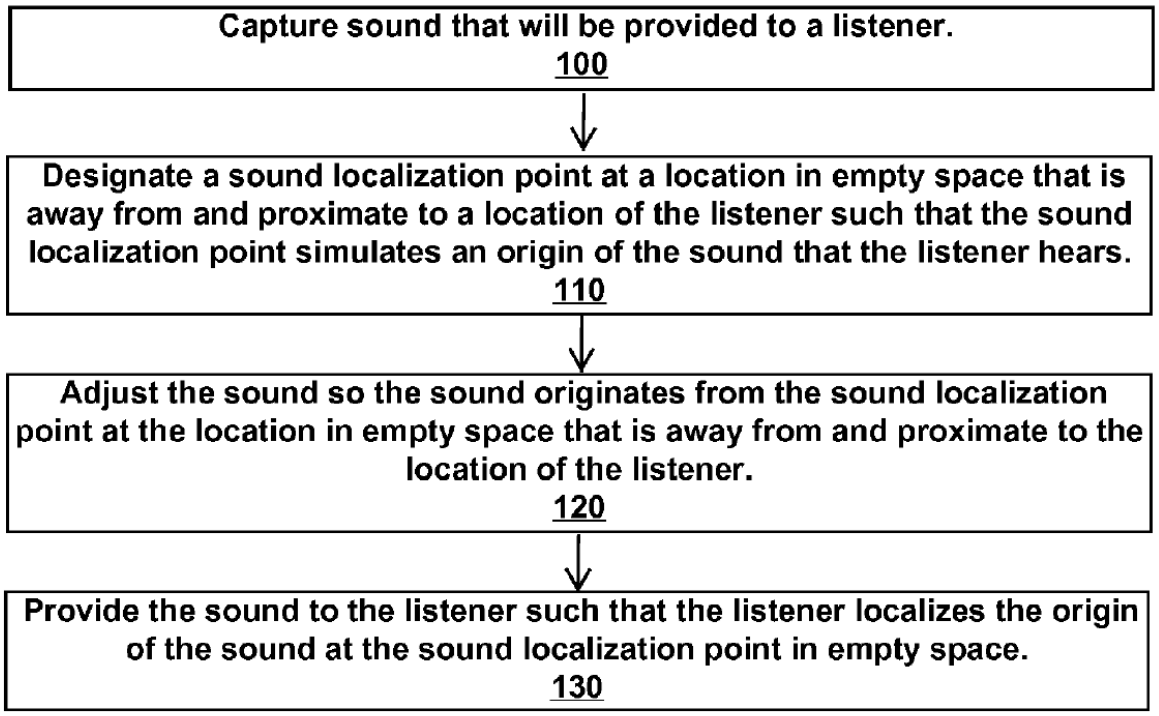


Figure 1

Notably, the '090 Patent does not describe positioning a party's voice at a SLP that is "away from and proximate to" another party as inventive, nor maintaining the origin of the voice at the SLP when the head moves.

C. The Challenged Claims

The effective filing date of the challenged claims is no earlier than June 23, 2014.

D. Prosecution History

The application for the '090 Patent was filed on June 23, 2014. The examiner issued no rejection. *See generally* EX1002.

On November 2, 2015, the Examiner allowed the claims. *Id.*, 121-127. The patent issued on December 29, 2015.

V. PRIOR ART OVERVIEW

A. Bedingfield

Bedingfield generally discloses all independent claim limitations, except for capturing binaural sound, and maintaining the SLP while the listener moves. Kyriakakis ¶84. Bedingfield was not cited or discussed during prosecution.

Bedingfield teaches synthetic audio placement for remote conference participants. EX1005, Abstract, 1:15-17, 1:50-2:16, 4:18-19; Kyriakakis ¶85.

Like the '090 Patent, Bedingfield provides the sound field over headphones (EX1005, 7:43-46, 13:34-43), which sense the listener's head movements and orientation (*id.*, 13:36-57), and the listener perceives the other participants as if they "were in the same location and speaking to each other directly." *Id.*, 4:31-5:3, Fig. 1; Kyriakakis ¶86.

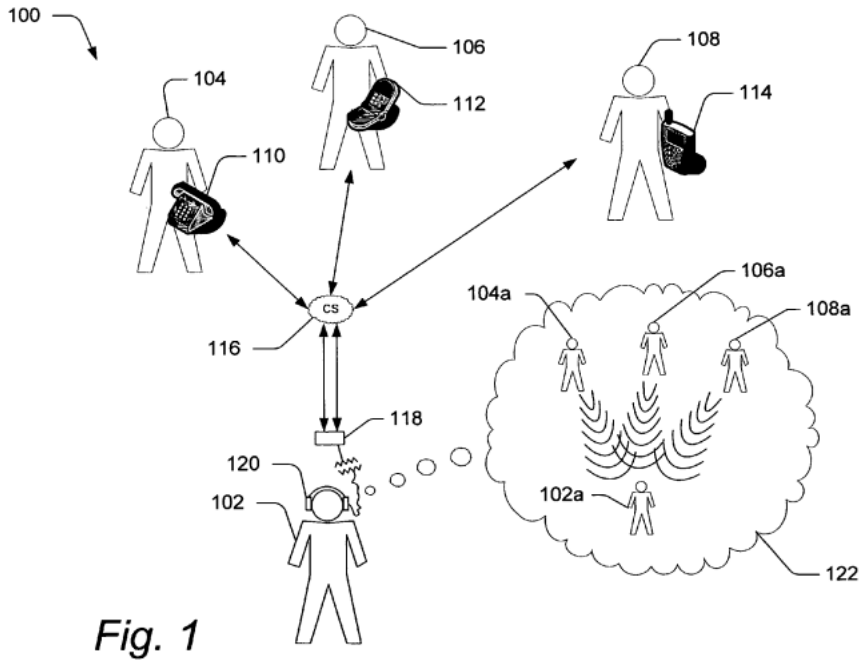


Fig. 1

B. Li

Li generally discloses all independent claim limitations, except for capturing binaural sound, and maintaining the SLP while the listener moves. Kyriakakis ¶87. Li was not cited or discussed during prosecution.

Li teaches a 3D sound system whereby users "experienc[e] music, movies, video games." EX1006, 7:21-22, 15:8, 16:4-7, 16:63-66, 19:39-20:52, Figs. 9, 16, 33; Kyriakakis ¶88.

Like the '090 Patent, Li's conferencing system "generate[s] configurable 3D sounds," by processing the speaker's voice to position "different speaking users at different spots." EX1006, 22:1-5, 33:14-18, 33:32-51, 33:62-66, 34:4-8, 34:20-51, 35:58-36:2. To generate a 3D sound, Li processes sound with HRTFs, which model

ILD and ITD cues for sound source localization. *Id.*, 1:65-2:21, 12:51-56, 22:49-52; Kyriakakis ¶89.

C. Andrea

Andrea discloses a WED that captures binaural sound. Kyriakakis ¶90. Andrea was not cited or discussed during prosecution.

Like the '090 Patent, Andrea manipulates the captured voice sound to improve its quality. EX1007, 1:44-56, 6:13-14, 11:6-27. Andrea's WED includes earbuds having a microphone at each ear (*id.,* 3:28-31), and the two microphones is beneficial for noise cancellation. *Id.*, 1:44-56, 14:55-58; Kyriakakis ¶91.

D. Ballas

Ballas discloses head tracking and maintaining an SLP even as the user's head moves. Kyriakakis ¶92. Ballas was not cited or discussed during prosecution.

Like the '090 Patent, Ballas "allow[s] the listener to change locations and head orientation without degradation of the audio signal." EX1008, ¶¶17-18, 25, cl. 1, 7, 13. Ballas's head tracking system provides the head location and orientation to a computer system that "processes the audio signals using a [HRTF] 106 thus producing spatialized audio." *Id.*, ¶¶17-18, 25-26. Ballas's HRTF filter "maintains the virtual location of the audio signals/sound" even as the user's head moves. *Id.*, Abstract, ¶17; Kyriakakis ¶93.

E. Sandgren

Sandgren discloses simulating virtual objects by changing the sound's reverberation and attenuation. EX1017, ¶¶6, 12, 77; Kyriakakis ¶94. Sandgren was not cited or discussed during prosecution.

Like the '090 Patent, Sandgren simulates sound-reflecting objects. *Id.*, Abstract, ¶¶2, 12, 16, 31-33, 69-70, Figs. 1, 5A-C; Kyriakakis ¶95.

VI. LEVEL OF ORDINARY SKILL IN THE ART

A person of ordinary skill in the art ("POSA") in 2014 would have had at least a bachelor's degree in computer science or engineering with a focus on audio signal processing and sound localization, with at least two years of relevant experience. Additional experience could substitute for less education, and additional education could likewise substitute for less experience. Kyriakakis ¶¶32-34.

This Petition does not turn on this precise definition, and the challenged claims would be unpatentable from the perspective of any reasonable person of ordinary skill in the art at the relevant time. Kyriakakis ¶35.

VII. CLAIM CONSTRUCTION

The prior art relied on in this Petition discloses the subject matter of the challenged claims under any reasonable construction, including their plain meaning.

Notably, PO contends in district court that "a processor" (i.e., 15[d]),[6] should be construed by its plain meaning and "[s]ection 112(f) does not apply." EX1027, 24-38. Solely for this IPR, Petitioner adopts PO's constructions of these terms for the challenged claims.[7] *See Western Digital Corp. v. SPEX Techs., Inc.*, IPR2018-0084, Paper 14, at 10-12 (April 25, 2018) (petitioner not required "to express its subjective belief regarding the correctness of its proposed constructions"); *Apple Inc. v. AGIS Software Development, LLC,* IPR2018-00819, Paper 9, at 9-11 (Nov. 7, 2018); *Cienna Corp. v. Oyster Optics, LLC*, IPR2021-00565, Paper 11, at 11-13 (Aug. 23, 2021). Thus, no express construction is required.

VIII. GROUND 1: BEDINGFIELD IN VIEW OF ANDREA RENDERS OBVIOUS CLAIMS 1-3, 6, 15-17

A. Motivation to Combine

It would have been obvious to use a WED that captures *binaural* sound, as taught in Andrea, as the communication device for Bedingfield's conference participants. Kyriakakis ¶¶96-106.

---

[6] *See* Claim Listing Appendix for claim language and labels.

[7] Petitioner reserves the right to argue alternative constructions in other proceedings, including that claims are indefinite.

A POSA would have recognized that capturing binaural sound (including binaural voice) and providing it to the listener would advance Bedingfield's stated goal of "allow[ing] a person to hear the speaker better" during an electronic call, such as the exemplified conference call. *See* EX1005, 1:44-46, Kyriakakis ¶98. Andrea discloses improving call quality through a noise-canceling solution for "directional listening," (EX1007, 1:44-51), requiring two microphones located at ears to capture binaural sound. EX1007, 6:13-14, 11:25-27, 11:6-13 (binaural sound capture allows for voice-noise separation), 15:18-31 (benefit of dual-channel earbuds is active noise reduction); Kyriakakis ¶¶99-101.

The POSA would have known that capturing binaural sound, as Andrea teaches, made it easier for the listener to distinguish between signal and noise, and improved call quality, providing spatialized sound to the listener and allowing for a more life-like conversation. Kyriakakis ¶102. Thus, a POSA would have recognized that applying Andrea's capturing *binaural* sound to Bedingfield would have enhanced the participants' experience on the electronic call. Kyriakakis ¶¶103-104.

The modification would have been a natural extension of Bedingfield's WED. Bedingfield already contemplates a "combined microphone input and speaker output" (i.e. headset). EX1005, 11:5-6, Fig. 11, 7:43-44; Kyriakakis ¶105. Bedingfield recognizes that "[a]n improved telephone or network-based

communication system may more closely simulate an in-person conference setting" by providing spatialized audio. EX1005, 3:25-27, 1:44-46. Capturing binaural voice to reduce noise would improve the simulated conference by improving sound quality and increasing SLP fidelity. Kyriakakis ¶106.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in applying Andrea's capturing binaural voice to Bedingfield. Kyriakakis ¶¶107-112.

Bedingfield and Andrea are analogous. Both describe capturing a user's voice using a WED and manipulating it to improve its communication. Kyriakakis ¶108. Both concern electronic communications, including VoIP and computer-based communications. EX1007, 1:49-51, EX1005, 7:47-50. Like Bedingfield, Andrea "relates to audio transmitting/receiving devices such as headsets with microphones." EX1007, 1:44-45; Kyriakakis ¶109.

Modifying Bedingfield would have been as simple as adding a second microphone to Bedingfield's WED for capturing binaural voice. Indeed, Andrea's devices "may be used in a multitude of different applications, including…communications such as [VOIP], PC to PC communications, PC to telephone communications." EX1007, 1:49-51; Kyriakakis ¶110. It would have been straightforward and beneficial to use a second microphone, which was well-known by 2014. Kyriakakis ¶111.

A POSA would have had a reasonable expectation of success, because the proposed modification involves nothing more than a combination of known prior art elements (microphones) according to known methods and known techniques (using multiple microphones to capture binaural voice) to yield predictable results (capture binaural sound), and involves use of a known technique to improve a similar device (WED) in the same way (adding a microphone). Kyriakakis ¶112.

C. Independent Claim 1

1. 1[pre]

To the extent limiting, Bedingfield discloses the preamble. EX1005, 1:15-17 (disclosing "*methods* and systems for providing synthetic audio placement"); Kyriakakis ¶¶113-115.

2. 1[a]: Capture Binaural Sound

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶116-125.

Bedingfield's Fig. 1 shows "an overall environment 100 for providing synthetic audio placement." EX1005, 4:18-19. This environment includes various participants, including listener 102 (*a listening person during a telephone call*[8]) and

[8] The Petition applies PO's interpretation that *telephone call* is "voice communication over an electronic network between at least two parties." *See* EX1027, 4.

participants 104, 106, and 108 (*talking person*). *Id.*, 4:19-23. The environment also includes a communication system 116 and a multi-channel processing device 118. *Id.*, 4:31-36. Kyriakakis ¶¶117-118.

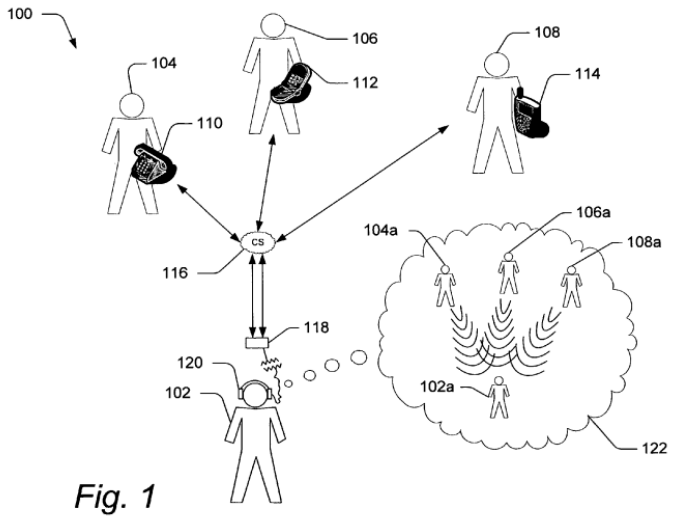


Fig. 1

Participants 104, 106, and 108 (*talking person*) have "communication devices 110, 112, and 114," which are devices such as "telephones, wireless phones, mobile phones." *Id.* 4:23-33. It was common in 2014 for *electronic earphones* such as earbuds or headphones to connect to such devices, and the POSA would have found it obvious that devices 110, 112, and 114 include an *electronic earphone located at a head of a talking person* because the *talking person* holds the device at the *head*. Kyriakakis ¶119. These devices also *captur[e] sound* because they include microphones that pick up audio signals representing the participant's voice. *Id.*

An example of participant 104's (a *talking person*) device is the "wearable speakers 1312" such as "stereo headphones," which include an *electronic earphone* that is *located at a head of a talking person*. EX1005, 7:43-46, 13:34-37; Figs. 14-15 (each participant utilizing wearable speakers), 14:11-23; Kyriakakis ¶120.

Bedingfield further discloses *sound will be provided to a listening person during a telephone call.* Kyriakakis ¶121. Bedingfield's multi-channel processing device 118 processes the captured sound from the participant 104 and outputs it (*sound will be provided*) over a multi-channel speaker system 120 that "enable[s] the listener 102 [*listening person*] to hear the speaking participant [*talking person*] in the sound field 122" as left and right sounds *during a telephone call*. EX1005, 4:18-19, 4:49-51. Because the "communication devices 110, 112, 114 may be *telephones*, wireless *phones*, mobile *phones*," Bedingfield describes a *telephone call*. *See id.*, 4:25-30 (emphasis added); Kyriakakis ¶121.[9]

Bedingfield discloses the *sound* as having two channels. EX1005, 4:45-47; Kyriakakis ¶122; *see* EX1032, 176:3-5 (PO's district court expert admitting providing binaural sound to listeners through earphones was known before 2014).

[9] The mapping for the phrase *during a telephone call* also applies to 1[b], 1[c], and 1[d], *infra*, and is not repeated in those sections.

While Bedingfield does not expressly teach *capturing* a *binaural sound*, this would have been obvious in view of Andrea. Kyriakakis ¶122; Sections VIII.A-B.

Andrea discloses earbuds (*electronic earphone*) worn by a user at left and right ears that "create a binaural recording." EX1007, 14:55-58; *see also id.*, 11:66–12:3, 3:28-31, 1:44-51, 14:55-61. Thus, Andrea's earbuds *captur[e] binaural sound*. Kyriakakis ¶¶123-124.

A POSA implementing Bedingfield would have been motivated to *capture binaural sound* taught by Andrea and would have had a reasonable expectation of success. Kyriakakis ¶125; Sections VIII.A-B.

3. 1[b]: Designate SLP

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶126-136.

Bedingfield discloses *a computer system* including at least a "multi-channel processing device 118 configured with the communication system 116," which together "creates the sound field 122 that is substantially similar… to an actual sound field where the listener 102 is present between all of participants 104, 106, 108." EX1005, 4:66-5:3. As Bedingfield further describes, the *computer system* simulates the sound field 122 so that the listener's perception of the call is as if the listener and participants "were in the same location and speaking to each other directly." *See* EX1005, 4:52-61. That is, a participant's *binaural sound* appears *to originate from the* location in the sound field (*SLP*). Kyriakakis ¶127.

Bedingfield discloses "method 300 of how synthetic audio placement may be organized" using multi-channel processing device 118 (part of the *computer system*). EX1005, 5:51-52, Fig. 3; Kyriakakis ¶128.

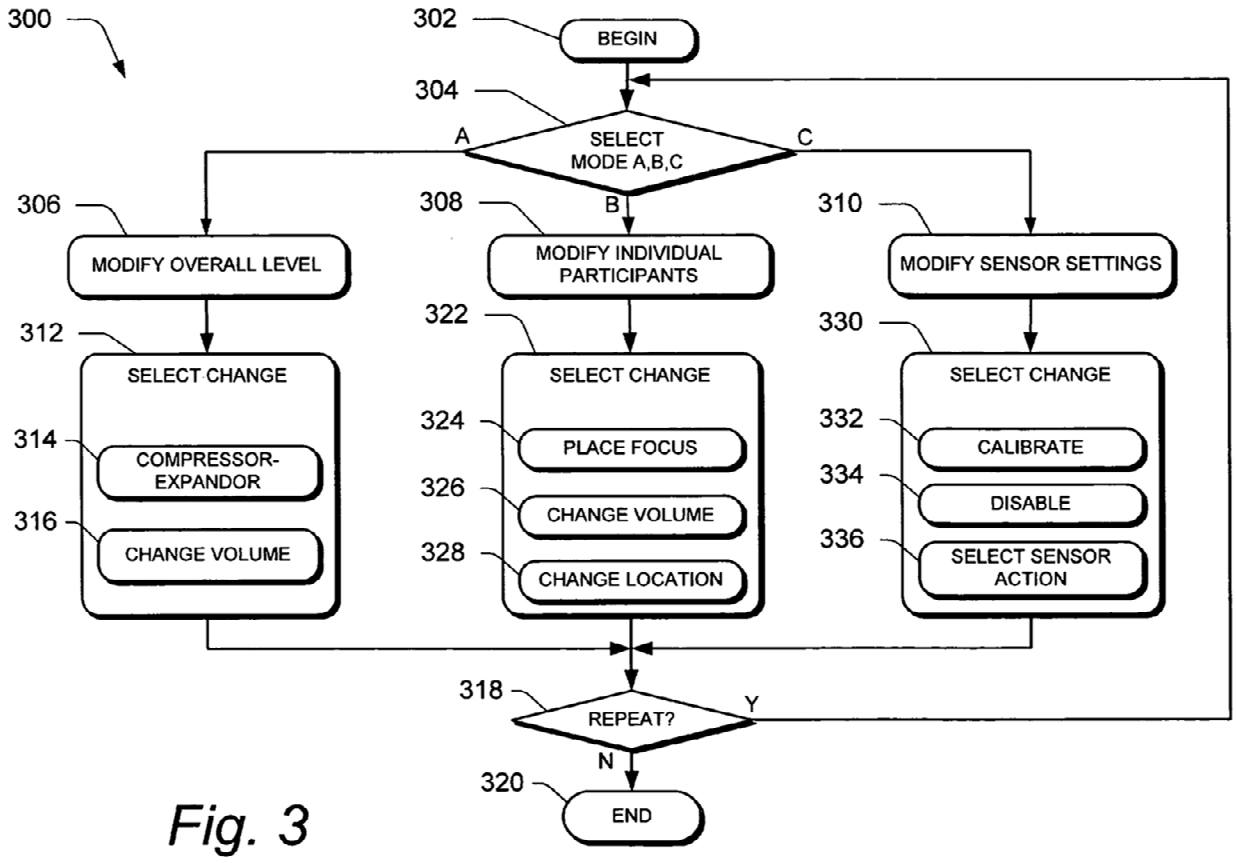


Fig. 3

In method 300, the listener (*listening person*) provides input to the multi-channel processing device 118 to "select the mode for modifying individual participants via route B at the block 308." EX1005, 6:27-29; Kyriakakis ¶129. For example, the listener selects to "change a participant's location," depicted as

locations 104a, 106a, and 108a in Fig. 1 (*designating an SLP*). EX1005, 6:29-33; *see also id.*, 6:43-45; Kyriakakis ¶129.

A POSA would have found it obvious to position the *SLP in empty space*, *i.e.*, at a location where no physical object exists. Kyriakakis ¶130; EX1005, 4:40-5:3, 9:60-64 (voice "occupy[ing] a 'vacant' location"). Bedingfield describes sound field configurations, including the circular configuration in Fig. 19a (reproduced below), and also placement of speaking participants as if they were "around a circular conference table." EX1005, 9:20-23, 16:36-48, Fig. 19a.

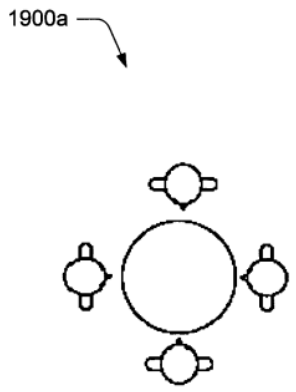


Fig. 19a

The POSA would have found it obvious to seat the listener (*listening person*) at a physical conference table, where an empty seat (*empty space*) is allocated for a virtual participant, as was known in the field. Kyriakakis ¶¶131-135 (describing EX1037's virtual participants at empty seats at the conference table, EX1038's

hybrid learning classroom where a virtual student takes "spatial ownership" of a physical empty seat, from which the virtual student's image is projected and the voice originates). This would have had the advantage of ensuring that the virtual teleconference environment achieves Bedingfield's goal to be as realistic as possible. Kyriakakis ¶135; EX1032, 195:12-196:10, 135:6-19, 236:2-24. Moreover, because each participant has its own sound field location that is separate from and located at "a distance from" the listener 102's location, each participant's *SLP* is *in empty space that is away from and proximate to the* listener 102 (*listening person*). Kyriakakis ¶135; EX1027, 8 (PO offering the construction "a distance from" for *away from and proximate to*).[10]

Bedingfield's simulated sound field 122 provides audio to the listener that is substantially similar, for example in "specific balance, volume, and timbre for each speaker" (EX1005, 4:52-5:3), "to an actual sound field where the listener 102 is present between all of participants 104, 106, 108" (*simulates an origin of the sound at the empty space that the listening person hears*). *Id.*; Kyriakakis ¶136.

---

[10] The mapping for the phrase *SLP in empty space that is away from and proximate to the listening person*, also applies to 1[c], 1[d], cl. 3, 6[b], 8[b], 8[c], 12[b], cl. 14, 15[d1], cl. 17, which recite the same phrase or a similar phrase, and the mapping is not repeated in those sections, *infra*.

4. 1[c]: Adjust Binaural Sound to Originate From SLP

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶137-138.

Bedingfield's multi-channel processing device 118 (part of *the computer system*) receives the audio signals (*sound*) and "process[es] the signals" (*adjust[s] the sound*) to "creat[e] a sound field 122" so that "the listener 102 [] hear[s] the speaking participant in the sound field 122." EX1005, 4:33-51; Kyriakakis ¶138; EX1027, 39-40 (PO admitting that the auditory adjustments for spatial audio location like ITD and ILD were "well-known"). As explained for 1[a], those audio signals (*sound*) are *captured at the earphone of the talking person*, and Bedingfield-Andrea renders obvious *binaural sound*. *See* Section VIII.C.2 (Ground 1, 1[a]); Kyriakakis ¶138.

5. 1[d]: Binaural Sound Occurs At SLP

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶139-140.

Bedingfield-Andrea's audio signals (*binaural sound*) are *provid[ed]* by the multi-channel speaker system 120 as a wearable headset or headphones (*electronic earphone located at a head of the listening person*). EX1005, Fig. 1; Kyriakakis ¶140.

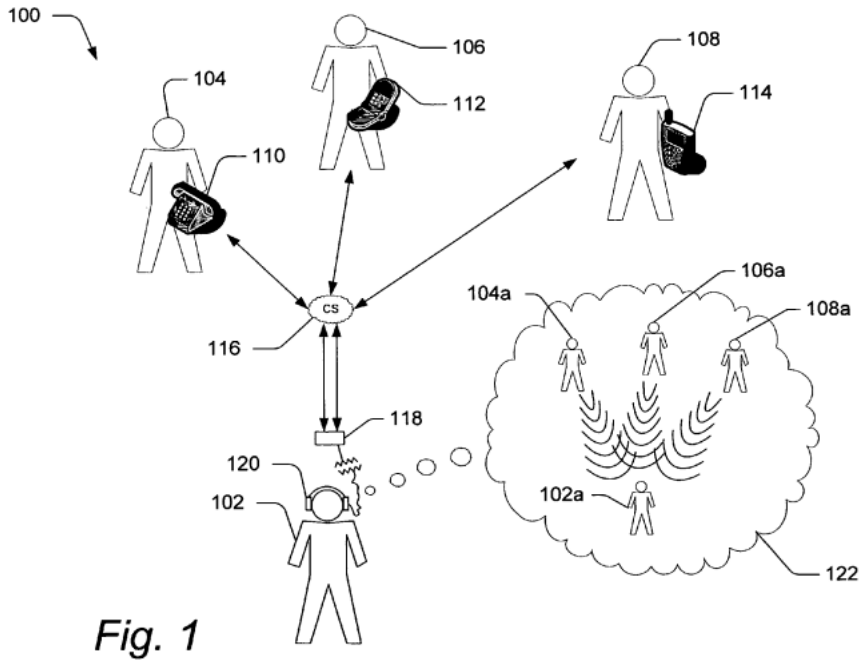


Fig. 1

D. Dependent Claims 2, 3, 6[11]

1. Claim 2

a) *2*[a]: Track Head Orientation

Bedingfield discloses this limitation. Kyriakakis ¶¶141-144.

Bedingfield discloses "sensors to detect [*track*] the movement of… listener 104." EX1005, 12:52-57. With reference to Fig. 13, Bedingfield's "wearable speakers 1312" (depicted as a headset) are worn on the *head* of the listener 104 (*the listening person*) and "detect [*track*] the *orientation* of [the] listener 104 [*the listening person*]." *Id.*, 13:38-39; Kyriakakis, ¶142.

[11] All dependent claims incorporate the analysis for the claims from which they depend.

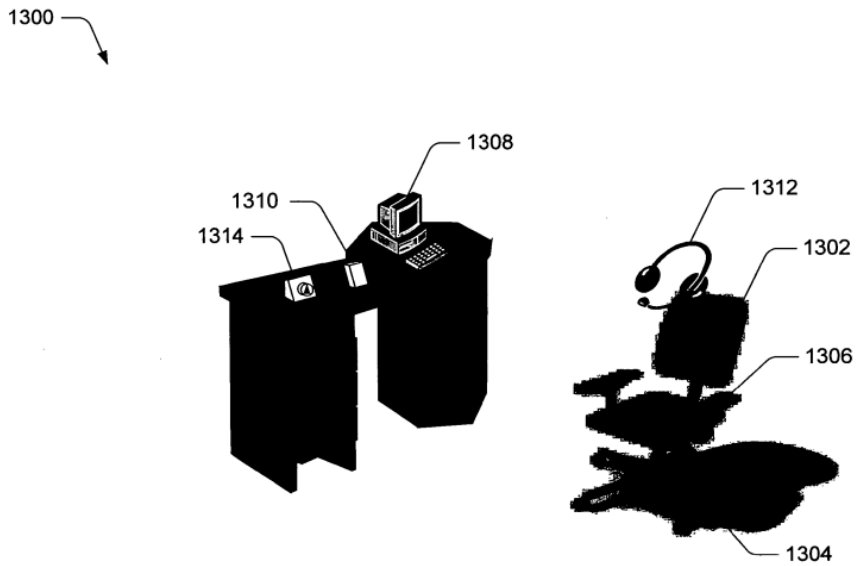


Fig. 13

Bedingfield's wearable speakers 1312 "sensemovements of the listener's 104 head" and the direction the listener is facing. *Id.*, 13:39-57 (describing motion detectors and sensors). By being worn on the head and detecting orientation and movement of the listener's head, Bedingfield *track[s] a head orientation of the listening person*. Kyriakakis ¶143.

Bedingfield's *tracking* of the *head orientation* is *with respect to the SLP*. Kyriakakis ¶144. As explained for 1[b], Bedingfield's sound field 122 has perceived locations (*SLP*), and listener 102 perceives the sound "as if the listener was present between all participants." EX1005, 4:55-66. The perceived locations (*SLP*) are fixed, meaning the listener's head orientation and movement is *with respect to the SLP*. *Id.*, 13:35-52 (describing the listener being oriented in different positions within the sound field by moving); *see also* 14:20-43, Figs. 14-15 (depicting the

sound field changing "to reflect the new sound field orientation" when the listener moves). Kryiakakis, ¶144

b) 2[b]: Adjust ITD

Bedingfield discloses this limitation. Kyriakakis ¶¶145-147.

As explained for 2[a], Bedingfield's listener (*listening person*) is oriented differently within the sound field 122 depending on (*in response to*) the listener's head turning from first to second positions (*the listening person changing the head orientation with respect to the SLP*). Kyriakakis ¶146.

Bedingfield also discloses *adjusting an ITD of the sound that the listening person hears* in order to orient the listener within the sound field 112. Kyriakakis ¶147. Bedingfield uses "various audio techniques" such as "delay" to place the different participants in different "configurations within the sound field." EX1005, 5:35-41. Bedingfield further describes applying "changes to delay" to enhance the "impression that a sound source has moved" relative to the listener. *Id.*, 3:56-59. The "delay" in this context is an *ITD*, commonly used to manipulate the perceived origin of a sound source in an audio scene. Kyriakakis ¶147; EX1027, 39-40. Thus, Bedingfield discloses *adjusting an ITD of the sound* delivered to the listener.

2. Claim 3: Receive Location

Bedingfield discloses this limitation. Kyriakakis ¶¶148-149.

As discussed for 1[b], the listener (*listening person*) provides input to "change a participant's location" 104a, 106a, 108a in the sound field 122. *See* EX1005, 6:29-45; Section VIII.C.3. Thus, Bedingfield's system *receiv[es], from the listening person, a location* for *where to position the SLP*. Kyriakakis ¶149. By tracking the *listening person's* orientation and relative location, and the *listening person* selecting a location for a participant, the *location specifies a distance from the listening person* within the sound field. *See* EX1005, 6:29-45; Kyriakakis ¶149.

3. Claim 6

a) 6[a]: Capture Binaural Sound

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶150-151.

As explained for 1[a], each of Bedingfield's participants 104, 106, and 108 is a *talking person*, and participant 104 is a first *talking person*. *See* Section VIII.C.2 (Ground 1, 1[a]); Kyriakakis ¶151. Participant 106 is a *second talking person*, and Bedingfield-Andrea renders obvious *captur[e] binaural sound* from participant 106 (*second talking person*). *Id.*

b) 6[b]: Designate Second SLP

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶152.

Bedingfield *designates* location 106a (*a second SLP*) that *simulates an origin of the sound from the* participant 106 (*second talking person*) *that the listener hears*. *See* Section VIII.C.3 (Ground 1, 1[b]); Kyriakakis ¶152.

c) 6[c]: Second SLP Adjacent To SLP

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶153-154.

Bedingfield depicts in Fig. 1 that location 106a (*second SLP*) is *position[ed] adjacent to* location 104a (*SLP*). Kyriakakis ¶154. As explained for 1[b], Bedingfield's sound field is simulated to provide the perception "as if the listener was present between all participants 104, 106, and 108." EX1005, 4:52-66. That is, the *binaural sound captured from* participant 104 (*talking person*) *appears to originate from the* location 104a (*SLP*) *and the binaural sound captured from* participant 106 (*second talking person*) *appears to originate from the* location 106a (*second SLP*). *See* Section VIII.C.3 (Ground 1, 1[b]); *see, also,* Section VIII.C.3 (Ground 1, 1[b]) Kyriakakis ¶154.

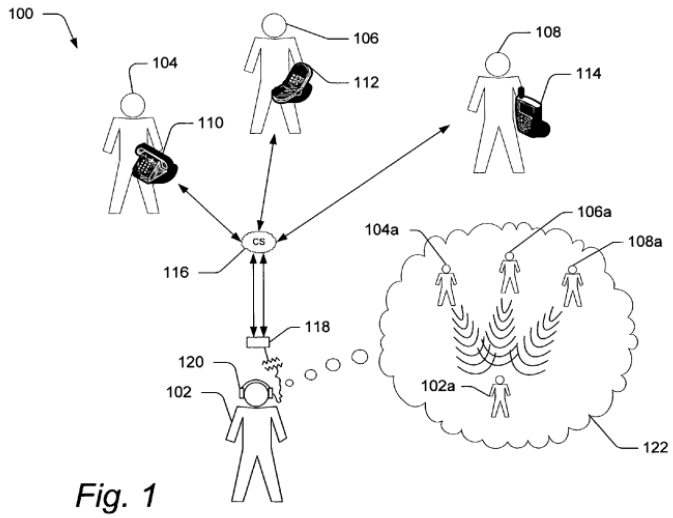


Fig. 1

E. Independent Claim 15

1. 15[pre]

To the extent limiting, Bedingfield discloses the preamble. Kyriakakis ¶¶155-156.

Bedingfield discloses *a computer system* as including at least a communication system 116, a multi-channel processing device 118, communication devices 110, 112, and 114, and multi-channel speaker system 120. EX1005, 4:31-51, Fig. 1; *see also* Section VIII.C.3 (Ground 1, 1[b]); Kyriakakis ¶156.

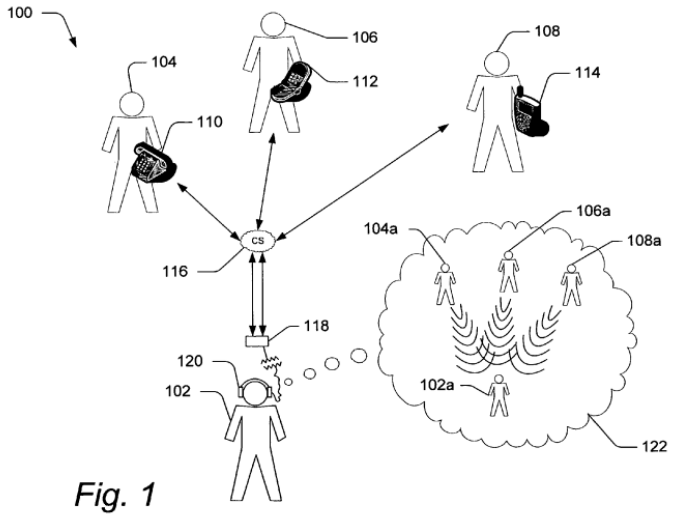


Fig. 1

2. 15[a]: First WED

Bedingfield discloses this limitation. Kyriakakis ¶¶157-158.

Bedingfield's *computer system* includes multi-channel speaker system 120, which as explained for 1[d], is a wearable headset or headphones (*first WED*) *that*

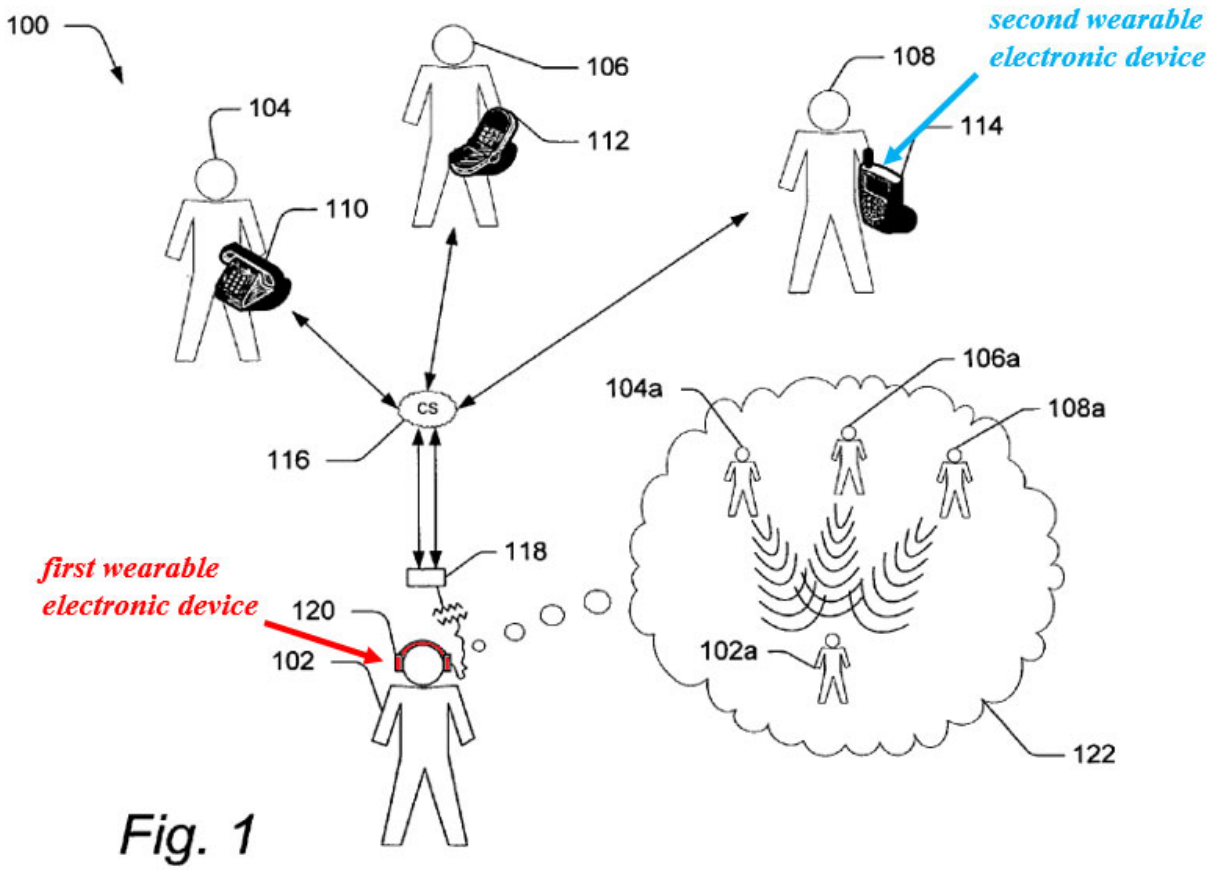
the listener 102 (*first person*) *wears.* EX1005, 4:18-19, 4:31-33, 4:49-51, 7:43-46, 13:34-37, 13:39-43 (describing wireless connections), Fig. 1; *see also* Section VIII.C.5 (Ground 1, 1[d]); Kyriakakis ¶158.

3. **15[b]:** Second WED Captures Binaural Voice[12]

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶159-160.

Bedingfield's *computer system* includes communication device 114 (*second WED*) *that* participant 108 (*a second person*) *wears*, *and that*, as explained for 1[a], *captures* a sound *of* participant 108 (*second person*) *during* a telephone call (*electronic phone call*) *with* the multi-channel speaker system 120 (*first WED*) *that the* listener 102 (*first person*) *wears*. *See* Section VIII.C.2 (Ground 1, 1[a]) (also explaining that Bedingfield-Andrea renders obvious *capture[] binaural* sound or *voice* from participant 108 (*second person*); Kyriakakis ¶160.

[12] The Petition applies PO's interpretation that *electronic phone call* is "voice communication over an electronic network between at least two parties." *See* EX1027, 8.



4. 15[c]: Memory

Bedingfield discloses this limitation. Kyriakakis ¶161.

Bedingfield states "[g]enerally, any of the functions described herein can be implemented using software," for which "program code can be *stored* in one or more computer readable *memory* devices." EX1005, 11:8-22 (emphasis added); *see also* 11:23-36, 12:5-44, Fig. 12. Such functionality includes synthetic audio placement within a sound field. EX1005, 1:15-18, 5:51-52, Fig. 3; Kyriakakis ¶161. Therefore, Bedingfield's components providing that functionality, including communication

system 116, the multi-channel processing device 118, communication devices 110, 112, and 114, and the multi-channel speaker system 120 (together, *the computer system*) included *a memory that stores instructions*. Kyriakakis ¶161.

5. 15[d]: Processor

Bedingfield discloses this limitation. Kyriakakis ¶162.

As explained for 15[c], "any of the functions" in Bedingfield, including synthetic audio placement, "can be implemented using software." EX1005, 11:8-22; *see also* 11:23-36, 12:5-44, Fig. 12; Kyriakakis ¶162. This includes "program code… executed on *processor(s)*." *Id.*, 11:13-22 (emphasis added); *see also* 11:60-12:16, 12:25-43. Therefore, the components that provide synthetic audio placement, namely Bedingfield's communication system 116, multi-channel processing device 118, communication devices 110, 112, and 114, and multi-channel speaker system 120 (together, *the computer system*), include *a processor that executes the instructions*. Kyriakakis ¶162; EX1032, 148:22-149:14 (PO's district court expert admitting the POSA would know how to implement 15[d], 15[d1], and 15[d2]).

6. 15[d1]: Adjust Binaural Voice To SLP

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶163.

Bedingfield processes (*adjust*) audio signals (*voice*) of the speaking participant (*second person*) *during* the conference call (*electronic phone call*) *in order to position a* perceived location 104a, 106a, 108a (*SLP*) *of the* audio signals

(*voice*) *to a location that is away from and proximate to the* listener (*first person*). *Id.* ¶163; Section VIII.C.4 (Ground 1, 1[c]); EX1027, 39-40. As explained for 15[b], that *voice* is a *binaural* voice. *See* Sections VIII.C.3 and VIII.E.3 (Ground 1, 1[b] and 15[b]); EX1032, 148:22-149:14.[13]

7. 15[d2]: SLP Represents Origin

Bedingfield-Andrea teaches this limitation. Kyriakakis ¶¶164-165.

As explained for 1[a], 1[d], 15[b], and 15[d1], Bedingfield's locations 104a, 106a, and 108a are each an *SLP* that *represents an origin of the binaural* sound *of the* participant (*second person*) *to the* listener (*first person*), and that *binaural* sound is the *binaural voice* from the participant. *See* Sections VIII.C.2, VIII.C.5, VIII.E.3, VIII.E.6 (Ground 1, 1[a], 1[d], 15[b], and 15[d1]); EX1032, 148:22-149:14; Kyriakakis ¶165.

F. **Dependent Claims 16-17**

[13] Even if limitation 15[d1] were construed as a means-plus-function term, it would have been obvious because PO admitted that the structural algorithm for performing adjust the binaural voice involves "well-known auditory adjustments... for spatial audio location like [ITD] and [ILD]" (EX1027), which are calculated using methods that the PO further admitted were known in the art.

1. **Claim 16**: SLP At Empty Chair

Bedingfield renders obvious this limitation. Kyriakakis ¶¶166-169; *see* Sections VIII.E.5-6 (Ground 1, 15[d]-15[d1).

The POSA would have known to position Bedingfield's *SLP* for speaking participants as if they were at a physical table, where an *empty* seat or *chair* is allocated for the virtual participant. Kyriakakis ¶167; *see* Section VIII.C.3 (Ground 1, 1[b]); *see also* EX1005, 9:20-23, 16:36-48, Fig. 19a (reproduced below).

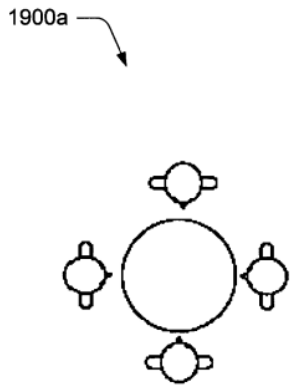


Fig. 19a

As Bedingfield explains, the sound field is provided to remote participants as if they were arranged "around a circular conference table." EX1005, 9:20-23, 16:36-48, Fig. 19a. Because conference participants are typically seated in chairs around conference tables, the POSA would have known to position the *SLP* for a remote

participant at an *empty chair*, and therefore *determine a location of an empty chair*, and *position the SLP at the empty chair*. Kyriakakis ¶168.

Moreover, when Bedingfield's user enters an input selecting a location for a participant's *SLP*, the *computer system* receives that input. EX1005, 6:43-45, 11:66–12:4; Kyriakakis ¶169. When the request is to position the *SLP* at an *empty chair*, Bedingfield's computer system *determine[s]* the *location of* the *empty chair* as the target for the *SLP* and *position*[s] the *SLP* at the *empty chair*. *See* EX1005. 1:62-63 ("*determin[e]* the synthetic position associated with the providing audio source"); Kyriakakis ¶169.

2. Claim 17: Change ITD, ILD, HRTF

Bedingfield renders obvious this limitation. Kyriakakis ¶¶170-174; *see* Sections VIII.E.5-7 (Ground 1, 15[d]-15[d2]).

Bedingfield discloses *ITD* is a cue that provides directionality and places a sound source in the sound field. *See* Section VIII.D.1 (Ground 1, 2). Another such cue is an *ILD*. To provide directionality in the sound field, Bedingfield's system applies a level difference between left and right ears, which is an *ILD*. EX1005, 9:48-57, Fig. 6; *see also id.*, 3:39-40, Table 1, 14:35-43; Kyriakakis ¶¶171-172.

Bedingfield further renders obvious *HRTF*. EX1005, 3:49-54 (describing the ears detecting position as a change in frequencies); Kyriakakis ¶173. A POSA would

have understood that Bedingfield's reference to using relative strength and attenuation of different sound frequencies for localization includes the well-known *HRTF*. Kyriakakis ¶173.

Thus, a POSA would have known that the *ILD*, *ITD*, and *HRTF* are all audio techniques that are *changed* in Bedingfield to *position the SLP* at the desired location. Kyriakakis ¶174.

IX. GROUND 2: BEDINGFIELD IN VIEW OF ANDREA AND BALLAS RENDERS OBVIOUS CLAIM 18

A. Motivation to Combine

It would have been obvious to modify Bedingfield-Andrea to track the user's head and adjust the sound to appear fixed even as the user's head moves, as Ballas taught. Kyriakakis ¶¶175-180. Bedingfield already discloses "detect[ing] the orientation of a listener 104" and "the sound field may correspondingly change to reflect the new sound field orientation." EX1005, 13:38-39, 14:11-43; *see also id.*, 1:56-59, 13:58-64, 18:49-52, Figs. 14, 15.

A POSA would have been motivated to combine the teachings of Bedingfield-Andrea and Ballas because they share a goal of providing a realistic experience. Kyriakakis ¶¶177-179. To achieve this, the POSA would have understood to track the user's head and provide a sound field accounting for the user's orientation and

movements, using Ballas's conventional head tracking and spatialization techniques. Kyriakakis ¶¶179-180.

Bedingfield's goal is to provide a sound field that "closely simulate[s] an in-person conference setting." EX1005, 3:25-27. Ballas shares this goal. EX1008, ¶¶2, 9, 10. To simulate the in-person conference, tracking the listener's head and changing the sound field so that sound sources appear fixed are key. Kyriakakis ¶179.

The modification would have been a natural extension of Bedingfield's use of motion detectors, sensors, and devices to detect orientation. Kyriakakis ¶180; EX1005, 13:54-56. Modifying Bedingfield-Andrea with Ballas's teachings would have allowed the user to move while staying connected to the event, and achieved Bedingfield's express purpose to simulate a realistic in-person conference. Kyriakakis ¶180.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success modifying Bedingfield-Andrea with Ballas. Kyriakakis ¶¶[181-182. Bedingfield and Ballas both track the listener's movement and orientation and change the sound field accordingly. In Bedingfield, the user's "wearable speakers" are headphones. EX1005, 13:35-36. Ballas also discloses speakers attached to a wearable head

frame. EX1008, ¶25. A POSA would have recognized Bedingfield-Andrea and Ballas are structurally and functionally similar. Kyriakakis ¶181.

A POSA implementing Bedingfield-Andrea would have attached Bedingfield's headphones to Ballas's frame. Kyriakakis ¶182. This would have merely involved a combination of known prior art elements (headset and head tracking system) according to known methods and techniques (accounting for head movement) to yield predictable results (producing spatialized sound that appears fixed as the user moves), and involves use of known sound processing techniques (HRTF) to improve a similar device (wearable communication device) in the same way (filtering sounds to produce a sound field that appears fixed). *Id.*

C. **Dependent Claim 18:** Change ITD, ILD, HRTF To Maintain SLP

Bedingfield-Andrea in view of Ballas teaches this limitation. Kyriakakis ¶¶183-191; EX1032, 250:2-9 (PO's district court expert admitting it was known in the 1980s for "head-track virtual acoustic systems [to] keep a sound source fixed in a location in virtual space independent of head movement"); *see also* EX1036, 12, 99, 132-34, Figs. 4.37-4.38; *See* Sections VIII.E.5-7 and VIII.F.2 (Ground 1, 15[d]-15[d2] and 17).

Bedingfield also discloses adjusting the SLP as the user's head moves. EX1005, 13:34-52 (orientating the listener in a different position within the sound field when the listener moves), 13:58-64 ("adjust[ing] the focus" of the sound field

by changing its "orientation"); *see also id.*, 6:57-7:2, 4:7-12, Figs. 14-15 (reproduced below with annotations), 14:32-38; Kyriakakis ¶¶185-186.

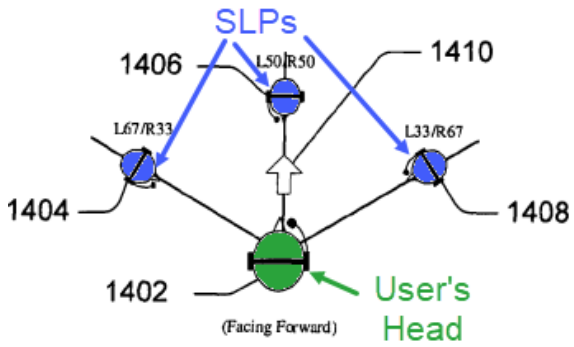


Fig. 14

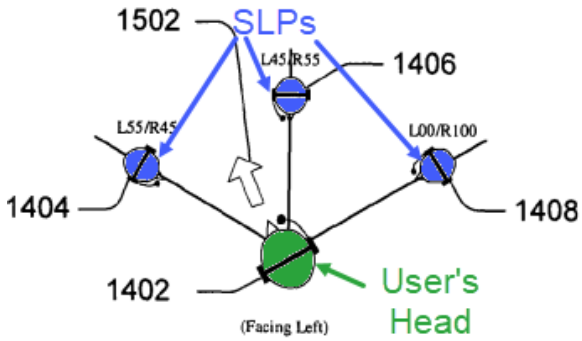


Fig. 15

Ballas "produc[es] virtual sound sources that are externally perceived and positioned at any orientation in azimuth and elevation from a listener." EX1008, Title, Abstract, ¶1. Ballas discloses "a virtual reality head mounted visual display system," such as an eyeglass frame or helmet. *Id.*, ¶25; *see also* Figs. 3, 4; Kyriakakis ¶187.

Figure 3. Eyeglass frame mount prototype

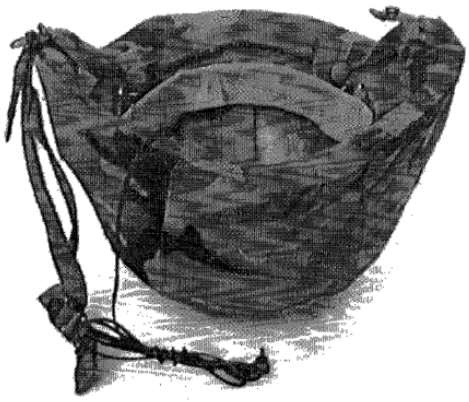


Figure 4. Helmet mount prototype

Mounted to the frame is a "head tracking system 104 [that] measures the location and orientation of a user's head and provides the measured information to the computer system 102." EX1008, ¶25; Kyriakakis ¶188.

Ballas discloses "a pair of [HRTFs]" that are "for a particular location" (*the SLP*). *Id.*, ¶16. Ballas's HRTF filters "maintain[] [the] virtual location of the audio signals thereby allowing the user to change location and head orientation without degradation of the virtual location of audio signals." *Id.*, Abstract, ¶¶17-20, claims 1, 7, 13. By allowing the user to move and tracking the user's head movements and orientation, while maintaining the virtual location of the audio signals, Ballas discloses *maintain[ing] the SLP fixed at the location during the electronic phone call while the first person physically moves away from the location and changes a head orientation with respect to the location*. Kyriakakis ¶189.

A POSA would have been motivated to apply Ballas's wearable frame with head tracking and HRTF processing to Bedingfield-Andrea's headphones, as this provides a more realistic virtual experience, allowing the user to move while fixing the perceived origin of the sound sources. Kyriakakis ¶¶190-191; *see also* Section IX.A-B (Motivation to Combine and Reasonable Expectation of Success).

X. GROUND 3: BEDINGFIELD IN VIEW OF ANDREA AND SANDGREN RENDERS OBVIOUS CLAIM 19

A. Motivation to Combine

It would have been obvious to modify Bedingfield-Andrea to simulate virtual objects, as taught in Sandgren. Kyriakakis ¶¶192-196. The references share a goal of providing a realistic and immersive experience for the user. Kyriakakis ¶¶192-193. To enhance the realism, the POSA would have been motivated to simulate sound-reflecting virtual objects, using Sandgren's conventional sound enhancement techniques. Kyriakakis ¶¶194-195.

Sandgren shares this goal of enhancing realism by rendering virtual 3D audio scenes and increasing externalization of sound through earphones. EX1017, ¶1, Abstract. To achieve this, Sandgren states it is "essential" to model the scatter sound waves of "objects in the room, e.g., tables and chairs." *Id.* ¶¶5-6, 19 (stating simulation of direct sound or main reflection from the floor, ceiling, and walls is "not enough").

According to Sandgren, a realistic sound simulator includes attenuation filters combined with reverberation generators. *Id.*, ¶¶62-63, Fig. 5C. Thus, a POSA would have been motivated to implement Bedingfield-Andrea to change attenuation and reverberation of the binaural sound to add artificial audial cues of virtual objects to provide a more realistic and immersive experience for the user, resulting in good externalization. Kyriakakis ¶195.

The modification would have been a natural extension of Bedingfield-Andrea's headphones providing a spatialized sound field. Kyriakakis ¶196. Like Bedingfield, Sandgren describes "simulating an audio scene or creating a virtual audio scene" over headphones. EX1017, ¶¶9-10; *see also* ¶¶11, 15-16, 19, 22, 46, 48-64, 70-100, Figs. 5A-C. Specifically, Sandgren describes "simulated sound-reflecting objects are simulated by processing a direct-sound signal with at least one [HRTF]…and generating simulated early-reflection signals." *Id.*, Abstract. A POSA would have recognized that the teachings of Sandgren would have enhanced Bedingfield's goal of an immersive experience by simulating virtual objects. Kyriakakis ¶196.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success modifying Bedingfield-Andrea with Sandgren. Kyriakakis ¶¶197-198. Bedingfield and Sandgren both simulate realistic auditory scenes having spatialized binaural sound and provide such sounds over headphones and are structurally and functionally similar. Kyriakakis ¶197.

A POSA implementing Bedingfield-Andrea would have recognized that sound spatialization would have been accompanied with Sandgren's attenuation filters and reverberation generators, to artificially simulate the presence of virtual objects. Kyriakakis ¶198. This would have merely involved a combination of

known prior art elements (localization cues with attenuation filters and reverberation generators) according to known methods and techniques (simulate sound-reflecting virtual objects) to yield predictable results (producing sound so that the user perceives the virtual objects), and involves use of a known technique (simulating sound reflections) to improve a similar device (wearable communication device) in the same way (filtering sounds to produce a realistic and immersive sound field). *Id.*

C. Dependent Claim 19

1. 19[a]: Attenuation And Reverberation To Add Artificial Audial Cues

Bedingfield-Sandgren teaches this limitation. Kyriakakis ¶¶199-203.

It would have been obvious to use Bedingfield's *processor to* perform this limitation in view of Sandgren. Kyriakakis ¶200.

Sandgren simulates sound-reflecting objects (*virtual objects*) to enhance the "[e]xternalization of the simulated sound source." EX1017, Abstract. In Sandgren, "[w]hen an object in a room produces sound, a sound wave expands outward from the source and impinges on walls, desks, chairs, and other objects that absorb and reflect different amounts of the sound energy." *Id.,* ¶2; *see also id.*, ¶¶12, 16, 70, Fig. 1; Kyriakakis ¶201.

Sandgren discloses *chang[ing] reverberation*. Kyriakakis ¶202. "Reflected sound energy reaching the listener is generally called *reverberation*." EX1017, ¶4,

47 (emphases added); *see also* ¶¶23-25, 48-66, 68-74, 100. Sandgren describes "[w]hen simulating a dynamic 3D audio scene…the parameters used by the [] early-reverberation generators 502, 502' must be updated continuously in order to simulate the reflection paths accurately."[14] *Id.*, ¶69, Figs. 5A-C; *see also id.*, ¶74. By updating parameters for a reverberation generator to simulate moving objects,

Sandgren also discloses *chang[ing] attenuation*. Kyriakakis ¶203. Sandgren's Figs. 5A-C depict audio simulators "having HRTF processors and an early-reflection generator," and depicts attenuation filters 504, 506, 520, and 522. EX1017, ¶¶31-33. These attenuation filters represent a gain or amplification applied to the direct sound (*e.g.*, 504, 506) or early reflections (*e.g.*, 520, 522). *Id.*, ¶¶49, 62, Figs. 5A-C; Kyriakakis ¶203. A POSA would have appreciated the attenuation applied to both direct sound and early reflections (including $a_k(n)$ representing the attenuation filter for the k-th reflection) *change* as Sandgren's *virtual objects* move and change location. EX1017, ¶¶74-75, 78-80, 89, Fig. 1; Kyriakakis ¶203.

[14] Sandgren refers to element 502 as an "early-reflection generator" as well as an "early-reverberation generator" and uses these names interchangeably. Petitioner therefore refers to Sandgren's element 502 as an "early-reverberation generator."

Sandgren contemplates the *sound* is a *binaural sound*. *Id.*; *see* EX1017, ¶11 (describing binaural impulse responses for a reverberant room), Figs. 5A-C (depicting left and right outputs); Kyriakakis ¶203.

Therefore, Sandgren discloses *change attenuation and reverberation of the… sound in order to add artificial audial cues of virtual objects near the second person*. EX1017, ¶¶4, 47, 82-84; Kyriakakis ¶¶202-203.

2. 19[b]: Virtual Objects

Bedingfield-Sandgren teaches this limitation. Kyriakakis ¶¶204-205.

As explained in 19[a], Sandgren's sound-reflecting objects (*virtual objects*) to enhance the "[e]xternalization of the simulated sound source." EX1017, Abstract. These objects are "simulated" and therefore *do not actually exist near the person but* are simulated in Sandgren's system, meaning they *are perceived to exist by the person from the artificial audial cues*. *Id.*, Abstract, ¶¶6, 23-25, 46-48, claims 1, 10, 19; Kyriakakis ¶205.

XI. GROUND 4: BEDINGFIELD IN VIEW OF BALLAS RENDERS OBVIOUS CLAIMS 8-10, 12, 14

A. Motivation to Combine

For the same reasons as explained for the Bedingfield-Andrea and Ballas combination, it would have been obvious to a POSA to implement the Bedingfield-only system to track the user's head orientation and adjust the sound so that the

sound appears fixed even as the user's head moves, as taught in Ballas. *See* Section IX.A (Ground 2, Motivation to Combine); Kyriakakis ¶206.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in making the Ballas modification to Bedingfield, for the same reasons as explained for the Bedingfield-Andrea and Ballas combination. *See* Section IX.B (Ground 2, Reasonable Expectation of Success); Kyriakakis ¶207.

C. Independent Claim 8

1. 8[pre]

To the extent limiting, Bedingfield discloses the preamble. Kyriakakis ¶¶208-209.

As explained for 15[pre], 15[c], and 15[d] Bedingfield discloses *a computer system* that has a memory (*a non-transitory computer readable storage medium*) *storing instructions* that cause a processor within the *computer system to execute a method*. *See* Section VIII.E.1, 4-5 (Ground 1, 15[pre, 15[c]-15[d]); Kyriakakis ¶209.

2. 8[a]: Commence Call[15]

Bedingfield discloses this limitation. Kyriakakis ¶¶210-211.

As explained for 1[a], Bedingfield discloses a telephone call (*electronic call*) that takes place (and is therefore *commence[d]* by Bedingfield's *computer system*) between listener 102 (*first individual*, depicted in Bedingfield's Fig. 1 as *wearing* headphones, or *an electronic earphone*) and participants 104, 106, and 108 (each a *second individual* who is *remote from* the listener 102). *See* Section VIII.C.2 (Ground 1, 1[a]); Kyriakakis ¶211.

3. 8[b]: Designate SLP

Bedingfield discloses this limitation. Kyriakakis ¶¶212-213.

As explained for 1[b], Bedingfield describes its *computer system* as *designat[ing]…an SLP…to the* listener 102 (*first individual*). *See* Section VIII.C.3 (Ground 1, 1[b]); Kyriakakis ¶213. In Bedingfield, that *designat[ion]* is *for the first individual* because the sound is delivered to the listener 102 (*first individual*) *through* the headphones (*electronic earphone). Id.* The *SLP simulates a location…from*

[15] The Petition applies PO's interpretation that *electronic call* is "voice communication over an electronic network between at least two parties." *See* EX1027, 4.

where an origin of a voice of any of participants 104, 106, 108 (*second individual*) *occurs for the* listener 102 (*first individual*). *Id.*

4. 8[c]: Voice Originates From SLP

Bedingfield alone or in view of Ballas teaches this limitation. Kyriakakis ¶¶214-215.

As explained for 1[d], Bedingfield's *computer system execute[s] the* telephone call (*electronic call*) *such that the origin of the voice of the* participant (*second individual*) *appears to the* listener (*first individual*) *to originate from… the SLP in empty space*. *See* Section VIII.C.5 (Ground 1, 1[d]). As explained for claim 18, the Bedingfield-Ballas teaches the *SLP* is a *fixed location…while the first individual moves and changes a location and a head orientation with respect to the SLP*. *See* Section IX.C (Ground 2, 18); Kyriakakis ¶215.

5. 8[d]: Maintain SLP But Alter ITD

Bedingfield alone or in view of Ballas teaches this limitation. Kyriakakis ¶¶216-218.

As explained for 8[c], Bedingfield and Ballas each discloses fixing (*maintaining*) *the origin of the voice of the second individual at the SLP in response to the first individual moving the head orientation with respect to the SLP*. *See* Section XI.C.4 (Ground 4, 8[c]); Kyriakakis ¶217.

As explained for 2[b] and claims 17 and 18, Bedingfield discloses *alter[ing] an ITD of the voice that the first individual hears*. *See* Sections VIII.D.2 and VIII.F.2 (Ground 1, 2[b] and 17) and Section IX.C (Ground 2, 18); Kyriakakis ¶218. A POSA would have appreciated it was commonplace to alter ITDs to maintain the origin of a sound while the user moves head orientation. Kyriakakis ¶218.

D. Dependent Claims 9, 10, 12, 14

1. **Claim 9:** In Response To First Individual Moving

Bedingfield alone or in view of Ballas teaches this limitation. Kyriakakis ¶¶219-221.

As explained for 8[c]-8[d], Bedingfield alone or in view of Ballas discloses *maintaining the origin of the voice of the second individual at the SLP… in response to the first individual moving closer to and away from the SLP*. *See* Sections XI.C.4-5 (Ground 4, 8[c]-8[d]); Kyriakakis ¶220.

As explained for claims 17 and 18, Bedingfield discloses *alter[ing] an ILD of the voice that the first individual hears*. *See* Section VIII.F.2 (Ground 1, 17) and Section IX.C (Ground 2, 18); Kyriakakis ¶221. A POSA would have appreciated it was commonplace to alter ILDs to maintain the origin of a sound while the user moves nearer and further away from the SLP. Kyriakakis ¶221.

2. Claim 10

a) 10[a]: Monitor Movements

Bedingfield alone or in view of Ballas teaches this limitation. Kyriakakis ¶¶222-223.

As explained for 2[a], as Bedingfield's listener's head moves and changes orientation, e.g., to focus on different participants (*in response to the first individual localizing the voice of the second individual during the electronic call*), Bedingfield senses and detects those movements (*monitor movements of the head orientation of the first individual*). *See* Section VIII.D.1 (Ground 1, 2); Kyriakakis ¶223. As explained for claim 18, Ballas discloses a head tracking system 104 to *monitor movements of the head orientation of the first individual in response to the first individual localizing the voice of the second individual during the electronic call. See* Section IX.C (Ground 2, 18); Kyriakakis ¶223.

b) 10[b]: Adjust HRTFs

Bedingfield-Ballas teaches this limitation. Kyriakakis ¶224-225.

As explained for claim 18 and 10[a], both Bedingfield and Ballas disclose changing (*adjust*) *HRTFs* to process audio signals and maintain a fixed location for a sound's SLP *in response to the movements of the head orientation of the first individual localizing the voice of the second individual during the electronic call*. *See* Section IX.C (Ground 2, 18) and Section XI.D.2 Ground 4, 10[a]); Kyriakakis

¶225. A POSA would have understood that changing HRTFs here involves *adjust[ing] HRTFs of the first individual to more accurately simulate real HRTFs of the first individual*, because head tracking and adjusting HRTFs would have provided a more realistic virtual experience, by allowing the user to move while fixing the perceived origin of sound sources. Kyriakakis ¶225.

3. Claim 12

a) 12[a]: Third Individual

Bedingfield discloses this limitation. Kyriakakis ¶¶226-227.

As explained for 1[a], Bedingfield's *electronic call* includes the listener 102 (*first individual*) and three additional participants 104, 106, 108 (where any two participants are the *second and third individual*, for example where participant 104 is a *second individual* and participant 106 is a *third individual*). *See* Section VIII.C.2 (Ground 1, 1[a]); Kyriakakis ¶227; *see also* EX1005, Figs. 4-9, 11, 14-15, 17-19, 4:19-23 ("any number of participants may be in connection with the listener."). By disclosing an *electronic call* with at least three individuals on the call, Bedingfield discloses *add[ing] a third individual to the electronic call*. Kyriakakis ¶227.

b) 12[b]: Designate SLP For Third Individual

Bedingfield discloses this limitation. Kyriakakis ¶¶228-229.

As explained for 1[b], 6[b], and 8[b], Bedingfield teaches *designat[ing], for the first individual, an SLP that simulates a* location 106a (*location*) *in empty space*

from where an origin of a voice of the third individual occurs for the first individual through the electronic earphone. *See* Sections VIII.C.3 and VIII.D.5 (Ground 1, 1[b] and 6[b]) and Section XI.C.3 (Ground 4, 8[b]). As Bedingfield depicts in Fig. 1, Bedingfield's location 106a (*SLP of the third individual*) *is adjacent to* the location 104a (*SLP of the second individual*). Kyriakakis ¶229.

4. Claim 14: Receive Commands

Bedingfield discloses this limitation. Kyriakakis ¶¶230-232.

Bedingfield describes the listener (*first individual*) as "mak[ing] selections on the multi-channel processing device 118." EX1005, 5:51-61, Fig. 3.

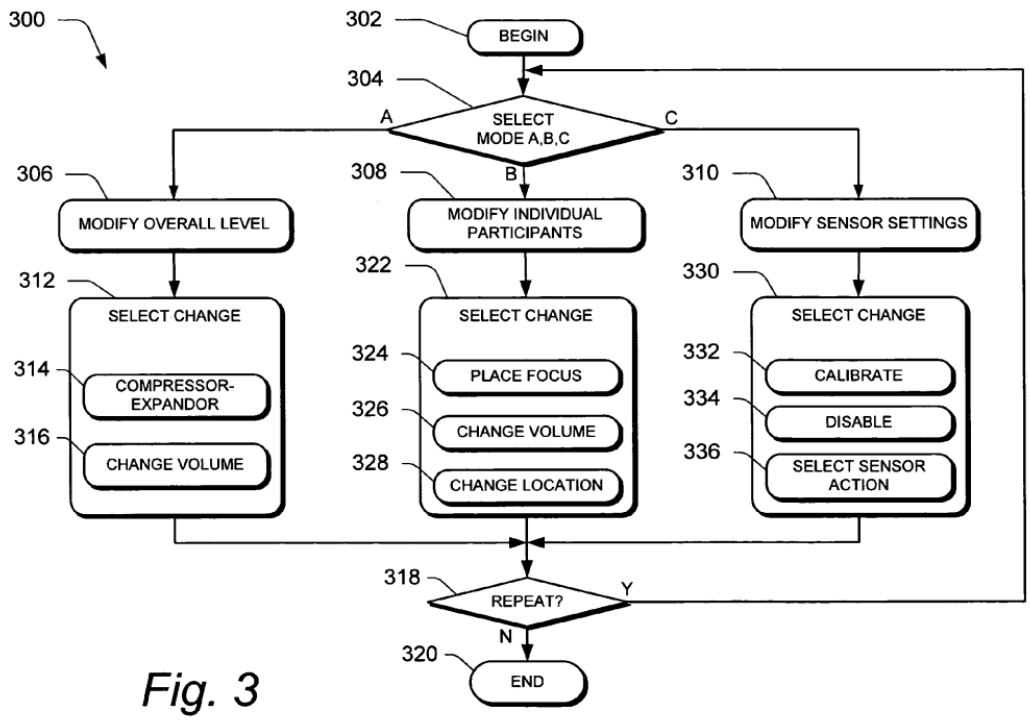


Fig. 3

These selections include "at the block 328… a new location for the participant within the sound field." EX1005, 6:43-45; *see also id.*, 6:29-32, 9:60-64, 14:7-10, 16:33-

35. By allowing the listener (*first individual*) to select a participant of the conference call and select a new location for that participant, Bedingfield discloses *receive, from the first individual, commands that move the SLP from a first point…to a second point.* Kyriakakis ¶¶231-232.

XII. GROUND 5: LI IN VIEW OF ANDREA RENDERS OBVIOUS CLAIMS 1, 3, 6, 15-17

A. Motivation to Combine

It would have been obvious to a POSA to use a WED that *captures binaural sound* as disclosed in Andrea in Li's 3D sound system, for the same reasons as explained for Ground 1. Kyriakakis ¶¶233-235.

A POSA would have recognized the advantages of applying Andrea's teachings of *capturing binaural voice* of a participant in an electronic phone call. Kyriakakis ¶234. For the same reasons as explained in Ground 1 (Section VIII.A), the POSA would have recognized that applying Andrea's teachings to Li's 3D sound system would have reduced noise, improved sound quality, and enhanced the experience of those participating in the call. Kyriakakis ¶234.

The modification would have been a natural extension of Li's disclosure of "[b]inaural recordings" that "sound closer to what humans hear in the real world." Kyriakakis ¶235; EX1006, 2:22-45; *see also id.*, 19:56-59. Li also contemplates that two audio recordings (e.g., one at each ear) would have included localization cues,

such as ITD and ILD. EX1006, 1:65-2:3. Capturing a participant's binaural voice to eliminate noise would have improved a simulated conference by improving the sound quality and create SLPs with increased fidelity. Kyriakakis ¶235.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in applying Andrea's capture of a binaural voice to Li's 3D sound system. Kyriakakis ¶¶236-238.

Li and Andrea are analogous. Both describe capturing a user's voice using a WED and manipulating that sound to improve its communication. Kyriakakis ¶237. Like Li, Andrea "relates to audio transmitting/receiving devices such as headsets with microphones." EX1007, 1:44-45. Further like Li, Andrea is directed to applications including "communications such as voice over internet protocol ("VoIP"), PC to PC communications, PC to telephone communications." EX1007, 1:49-51. A POSA would have had a reasonable expectation of success to improve sound quality and to produce more reliable localization cues. Kyriakakis ¶237.

Moreover, Li's 3D sound system is compatible with receiving input from a multi-microphone array, including an array of two microphones. EX1006, Fig. 9, 17:60-63. Thus, it would have been straightforward and beneficial to include a second microphone on the WED, and the modification to Li would have been a combination of known prior art elements according to known methods and known

techniques to yield predictable results, and involves use of a known technique to improve a similar device in the same way. Kyriakakis ¶238.

C. Independent Claim 1

1. 1[pre]

To the extent limiting, Li discloses the preamble. Kyriakakis ¶¶239-240.

Li discloses a *method* and system that generates a configurable 3D sound. EX1006, 3:39-44; Kyriakakis ¶240.

2. 1[a]

Li-Andrea teaches this limitation. Kyriakakis ¶¶241-248.

As Li explains as to Fig. 9, its configurable 3D sound system includes headphones or a headset, a microphone array system, and individual microphones, which together enable an electronic phone call (*a telephone call*, *see* EX1027, 4) between users. EX1006, 16:58-17:14, 17:22-37, 6:63-66; 7:41-8; 10:5-7, 19:39-20:52, 21:60-22:5 ("[T]he configurable 3D sound system 900 provides a method of communication among multiple people [located remotely from one another], for example, in a conference call, by placing different speaking users at different spots to mimic a real conference room environment."), 22:53-67, 23:21-43, 28:9-13, Figs. 9, 16, 33; Kyriakakis ¶¶242-243.

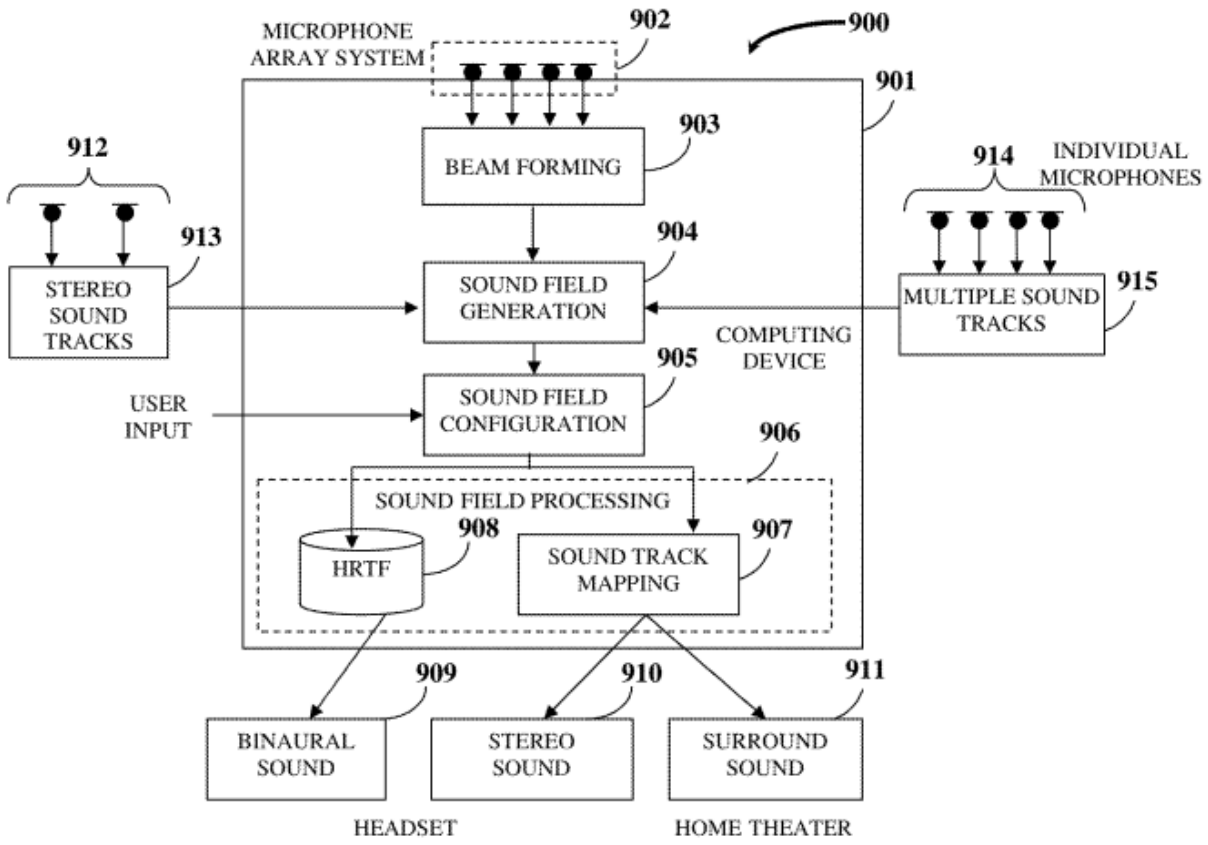


FIG. 9

Li's call is a *telephone call* under PO's interpretation (EX1027, 4) and allows a first user (*listening person*) and a second user (*a talking person*) to communicate using, *e.g.*, headphones and headsets. *See also* EX1006, 3:6-8 ("consumers often use headphones for audio rendering on a mobile device."), 6:68-7:4, 7:33-37, 10:5-11, 14:14-15, 16:7, 17:35-37, 20:9-12, 20:44-48; Kyriakakis ¶243.

The POSA would have readily understood that Li's headset in Fig. 9 includes a microphone and headphone, such that the two users (*listening person* and *talking person*) would have had their own headset or headphones. Kyriakakis ¶244. To the

extent this is not explicit in Li, Andrea discloses a headset, which would have been used in Li. EX1007, 14:55-58; *see also id.*, 12:1-2, 3:28-31, 1:44-51, 14:55-58; Kyriakakis ¶¶244-247.

The headset (*electronic earphone*) *located at* the *head of* the second (speaking) user (*talking person*) *captures sound*, including the user's voice, which *will be provided* to the first user (*listening person*). EX1006, 4:58-63 ("'Sound sources' refers to… a person's voice."), 18:16-21 ("recorded sound can be processed and reproduced to 3D sound."); Kyriakakis ¶245. Li's system also enables simultaneous recording of binaural sound using a microphone array. EX1006, 34:52-61; *see also id.*, 7:33-37 (describing individuals recording 3D sound), 2:22-32, 4:6-7, 34:52-61; EX1032:176:3-5.

To the extent Li does not expressly disclose capturing *binaural sound*, a POSA would have found it obvious to implement Li's *electronic earphone* to capture *binaural sound* taught by Andrea. Kyriakakis ¶248; Section XII.A-B. Andrea discloses a headset that includes two microphones to capture *binaural sound*. *See* Section VIII.C.2 (Ground 1, 1[b]). A POSA implementing Li would have therefore been motivated to *capture binaural sound* taught by Andrea. Kyriakakis ¶¶246-248.

3. 1[b]

Li-Andrea teaches this limitation. Kyriakakis ¶¶249-260.

Li's *computer system* is "configurable 3D sound system 900" used for communication among multiple users, including the *listening person* on a *telephone call* with *the talking person*. EX1006, 16:58-17:14, 19:39-20:52, 21:60-22:5, 22:53-67, 23:21-43, 28:9-13, Figs. 9, 16; Kyriakakis ¶250.

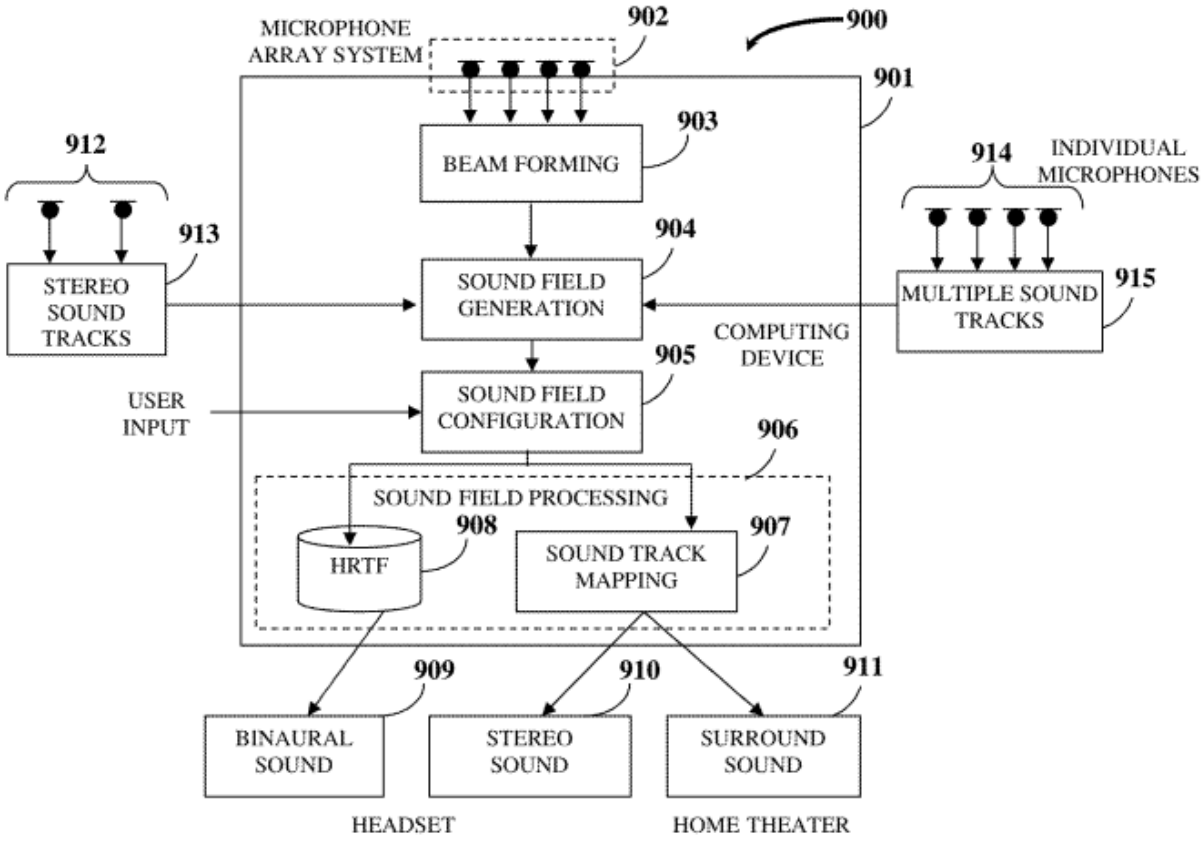


FIG. 9

Li's *computer system designat[es] an SLP* based on inputs received by a user. Kyriakakis ¶250. "[T]he user configures the placements and movements of any available sound sources as the inputs in order to obtain a virtual reality scene," and the resulting 3D sound is provided to a headset. EX1006, 22:53-67; *see also id.*,

21:57-66 (system 900 allows user to move the sound source), 20:37-40 (system 900 allows the user to set sound source locations), 22:39-42; *see also* Abstract, 2:42-45, 5:7-30, 15:40-42,15:61–16:8, Fig. 9; Kyriakakis ¶¶251-252.

A POSA implementing Li would have found it obvious to position the *SLP* in *empty space*, or at a location where no physical object exists. Kyriakakis ¶253. Li describes "placing different speaking users at difference spots to mimic a real conference environment." EX1006, 22:1-5. The POSA would have found it obvious to seat the listener (*listening person*) at a physical conference table, where an empty seat at the table (*empty space*) is allocated for a virtual participant, as was commonly done in the field. Kyriakakis ¶¶253-257 (describing EX1037 and EX1038). This would have positioned the virtual participant's *SLP* at the empty seat at the physical table (*empty space*), with the advantage of ensuring that the virtual teleconference environment achieves Li's goal to "match a vivid 3D vision field." EX1006, 3:11-16; Kyriakakis ¶257.

To the extent that *empty space* refers to a virtual environment with no virtual objects, Li describes as to Fig. 1, measuring HRTFs "in a quiet sound treated room free of impulsive background noise" (EX1006, 10:39-54) and using those HRTFs to generate binaural sound 909 (*id.*, 19:62-67, 17:27-37, 19:27-36). A POSA would understand that because HRTFs are measured in a "quiet sound treated room," the *SLP* would be placed in a similarly object-free space. Kyriakakis ¶258. Because the

sound sources are placed "at different spots to mimic a real conference room environment," such as seats around a conference table located "a distance from" the listener (EX1006, 21:66-22:5, 20:9-14; EX1027, 8), the *SLP at a first location* is *away from and proximate to the user*. Kyriakakis ¶258.[16]

In addition, Li's first user (*listening person*) perceives (*hears*) the binaural sound to originate at the SLP *during the telephone call* because Li's 3D sound system 900 is used to provide "a conference call" that "plac[es] different speaking users at different spots to mimic a real conference room environment." EX1006, 22:1-5; Kyriakakis ¶259.

To the extent Li does not expressly disclose *binaural sound*, a POSA would have found it obvious to implement Li's system to adjust *binaural sound* as taught by Andrea. *See* Section XII.C.2 (Ground 5, 1[a]); Kyriakakis ¶260.

4. 1[c]

Li-Andrea teaches this limitation. Kyriakakis ¶¶261-263.

[16] The mapping for the phrase *SLP in empty space that is away from and proximate to the listening person*, also applies to 1[c], 1[d], 3, 6[b], 8[b], 8[c], 12[b], 14, 15[d1], 17, which recite the same or similar phrase, and the mapping is not repeated in those sections, *infra*.

As explained for 1[b], Li's configurable 3D sound system (*computer system*) designates an SLP and processes sound using HRTFs (*adjusting the binaural sound*) to generate localized sound to the first user (*listening person*). *See* Section XII.C.3 (Ground 5, 1[b]); Kyriakakis ¶262; EX1027, 39-40. Li's configurable 3D sound system 900 includes a computing device 901, on which a "3D sound processing application" uses HRTFs to reproduce 3D binaural sound 909. EX1006, 17:1-4, 17:22-37, 17:45-49, 18:11-13, 20:26-36, Fig. 16; *see also id.*, 5:22-28; Kyriakakis ¶262.

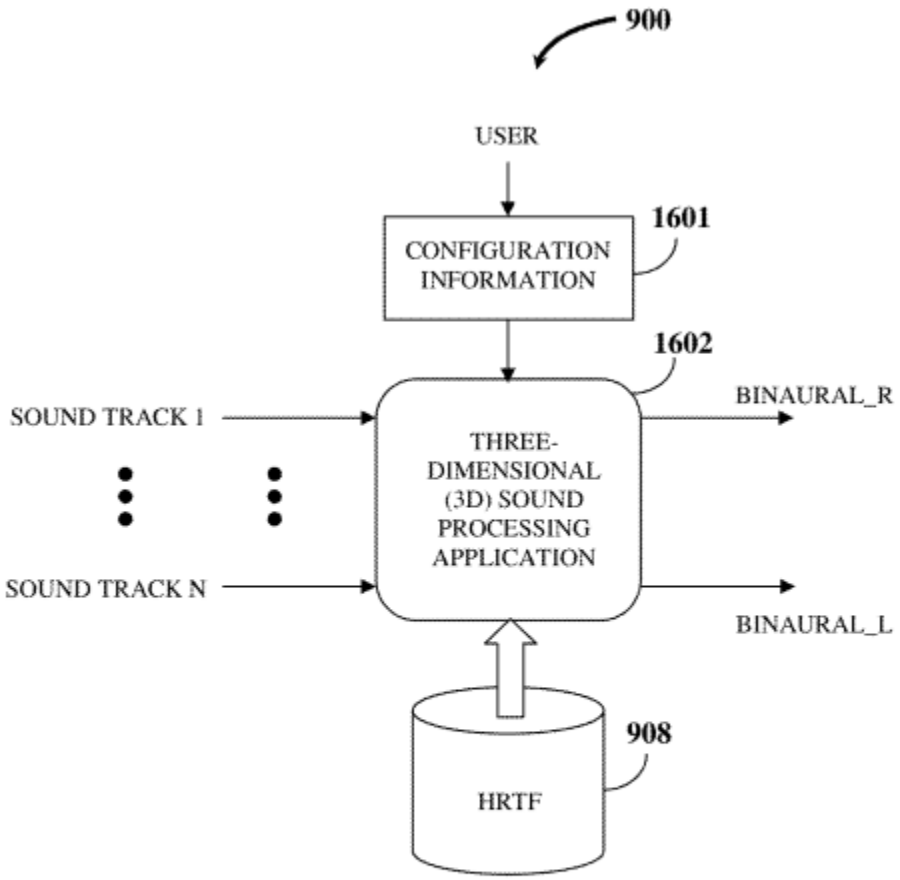


FIG. 16

As explained for 1[a], Li-Andrea teaches that *the binaural sound is captured*. *See* Section XII.C.2 (Ground 5, 1[a]); Kyriakakis ¶263.

5. 1[d]

Li teaches this limitation. Kyriakakis ¶¶264-265.

As explained for 1[a], Li contemplates that the first user (*listening person*) has a headset (*an electronic earphone located at a head of the listening person*). *See* Section XII.C.2 (Ground 5, 1[a]); Kyriakakis ¶265. As explained for 1[b]-1[c], Li discloses placing the sound sources at "different spots to mimic a real conference room environment" over headsets. *See* Sections XII.C.3-4 (Ground 5, 1[b]-1[c]); Kyriakakis ¶265.

D. Dependent Claims 3, 6

1. Claim 3

Li discloses this limitation. Kyriakakis ¶¶266-267.

As discussed for 1[b], Li's first user (*listening person*) "configures the placements" of the sound sources. *See* Section XII.C.3 (Ground 5, 1[b]); Kyriakakis ¶267. Thus, Li's system *receiv[es], from the listening person, a location* for *where to position the SLP*. Kyriakakis ¶267. By configuring placement and allowing the user to move the sound source within the sound field 1700, Li's *listening person* provides *a location that specifies a distance from the listening person*. *See* Section XII.C.3 (Ground 5, 1[b]); Kyriakakis ¶267.

2. Claim 6

a) 6[a]

Li-Andrea teaches this limitation. Kyriakakis ¶268-269.

As explained for 1[a], each of Li's conference call participants would have had a headset or headphones. *See* Section XII.C.2; Kyriakakis ¶269. As the POSA would have appreciated, each headset or headphone would have been modified by Andrea to *capture binaural sound* from each respective user. *See* Section XII.C.2 (Ground 5, 1[a]); Kyriakakis ¶269. Moreover, the POSA would have readily understood that Li's conference call is capable of including more than two participants, such that a third user (*second talking person*) participates in the conference call. Kyriakakis ¶269.

b) 6[b]

Li-Andrea teaches this limitation. Kyriakakis ¶270.

As explained for 1[b] and 6[a], Li *designates* a placement (*a second SLP*) that *simulates an origin of the sound from the* third user (*second talking person*) *that the listener hears*. *See* Sections XII.C.2 and XII.D.2.a) (Ground 5, 1[a] and 6[a]). A POSA would have modified Li to simulate *binaural sound* as taught in Andrea. Kyriakakis ¶270; *see* Section VII.A.

c) 6[c]

Li discloses this limitation. Kyriakakis ¶¶271-272.

As explained for 1[b], Li describes placing different speaking users at different spots. *See* Section XII.C.3 (Ground 5, 1[b]). A POSA would have readily appreciated that Li's goal to mimic a real conference room would have meant that

the different placements (*SLPs*) for the *talking persons* would have been positioned adjacent to each other. Kyriakakis ¶272.

E. Independent Claim 15

1. 15[pre]

To the extent limiting, Li discloses the preamble. *See* Section XII.C.3-4 (Ground 5, 1[b]-1[c]); Kyriakakis ¶273.

2. 15[a]

Li discloses this limitation, as each user would have had their own headset or headphones. Kyriakakis ¶274; *see* Section XII.C.2 (Ground 5, 1[a]).

3. 15[b]

Li-Andrea teaches this limitation. Kyriakakis ¶¶275-276.

As explained for 15[a], Li discloses a headset or headphones (*second WED*) *that* a second user (*second person*) *wears*, and the POSA would have found it obvious to apply Andrea's *capture of a binaural voice* to a *second WED*, as explained for 1[a]. *See* Section XII.E.2 (Ground 5, 15[a]), Section XII.C.2 (Ground 1, 1[a]); Kyriakakis ¶276.

4. 15[c]

Li discloses this limitation. Kyriakakis ¶277. Li's configurable 3D sound system is a *computer system* (*see* 15[pre]), which "employs the architecture of the computer system 3400" including "a *memory* unit 3402 for *storing* programs and

data" as well as "*instructions* [that are] fetched by the processor 3401." EX1006, 30:4-35, 31:18-35, Fig. 34; *see also* 30:65-31:1, 35:24-50. Li's *computer system* therefore includes a memory unit 3402 (*memory*) *that stores instructions*. Kyriakakis ¶277.

5. 15[d]

Li discloses this limitation. Kyriakakis ¶278. Li's configurable 3D sound system is a *computer system* (*see* 15[pre]), which "employs the architecture of the computer system 3400" including a *processor* 3401 that "*executes* computer programs" stored in the memory unit 3402. EX1006, 30:17-19, 30:26-29; *see also* 31:18-35, 32:32-33:31, 34:36-51, 35:17-50, 36:37-41. Li's *computer system* therefore includes a processor 3401 (*processor*) *that executes the instructions*. Kyriakakis ¶278.

6. 15[d1]

Li-Andrea teaches this limitation. Kyriakakis ¶¶279-280.

As explained for 1[c], Li processes sound by applying HRTFs to (*adjust*) the sound from the speaking user (*the binaural* sound *of the second person*) *during the electronic phone call in order to position* the localized sound at an *SLP that is away from and proximate to the* listening user (*first person*). *See* Section XII.C.4 (Ground

5, 1[c]); Kyriakakis ¶280; EX1027, 39-40. As explained for 15[b], that sound is a *binaural voice*, as taught by Andrea. *See* Section XII.E.3 (Ground 5, 15[b]).[17]

7. 15[d2]

Li-Andrea teaches this limitation. Kyriakakis ¶281. As explained for 1[a], 1[d], 15[b], and 15[d1], Li places speaking users at "different spots," each being an *SLP* that *represents an origin of the binaural* sound *of the* speaking user (*second person*) *to the* listening user (*first person*), and that *binaural* sound is the *binaural voice* from the speaking user. *See* Sections XII.C.2, XII.C.5, XII.E.3, and XII.E.6(Ground 5, 1[a], 1[d], 15[b], and 15[d1]); Kyriakakis ¶281; EX1032, 148:22-149:14.

F. Dependent Claims 16-17

1. Claim 16

Li teaches this limitation. Kyriakakis ¶¶282-285.

[17] Even if limitation 15[d1] were construed as a means-plus-function term, it would have been obvious over Li-Andrea because PO admitted that the structural algorithm for adjusting the binaural voice involves "well-known auditory adjustments…for spatial audio location like [ITD] and [ILD]" (EX1027) which are calculated using methods the PO further admitted were known.

As explained in 1[b], 1[c], 15[d]-15[d1], Li-Andrea teaches *a processor* that *executes instructions…to position the SLP* at a location *that is proximate to the first person*, and Li's placement of sound source locations are *SLPs*. *See* Sections VIII.C.2-3, VIII.E.5-6 (Ground 5, 1[b]-1[c], 15[d]-15[d1]). As explained for 1[b], the POSA would have known to position Li's *SLP* at an *empty chair* of a conference table, where the *empty chair* is allocated for the virtual participant. Kyriakakis ¶283; *see* Section XII.C.3 (Ground 5, 1[b]); *see also* EX1006, 22:1-5.

Because conference participants are typically seated in chairs around conference tables, the POSA would have known to position the *SLP* for a remote participant at an *empty chair*, and therefore *determine a location of an empty chair*, and *position the SLP at the empty chair*. Kyriakakis ¶284.

Moreover, Li explains that a user may enter a user input to select a location for a participant *SLP*. EX1006, 15:44-45, 27:16-22. In a teleconference setting mimicking an in-person, it would have been a common request to designate an *empty chair* for a participant. Kyriakakis ¶285. Upon receiving that user input, Li's computer system *determine[s]* the *location of* the *empty chair* as the target for the *SLP* and *position[s]* the SLP at the *empty chair*. EX1006, 11:58-62, 20:32-36; Kyriakakis ¶285.

2. Claim 17

Li teaches this limitation. Kyriakakis ¶¶286-290.

As explained in 15[d], 15[d1], 15[d2], Li's *processor executes instructions* that include applying *HRTFs to position an SLP of the binaural voice to a location that is away from and proximate to the first person*. *See* Section XII.E.5-7 (Ground 5, 15[d]-15[d2]); Kyriakakis ¶287.

Li explains "human hearing localizes sound sources in a three-dimensional (3D) spatial field, mainly by three cues: *an [ITD]* cue, *an [ILD]* cue, and a spectral cue." EX1006, 1:65-2:3; *see also id.*, 12:54-57; Kyriakakis ¶288.

Li also discloses *changing* the *ILD* and *ITD* as part of its bank of HRTF filters. EX1006, 2:6-21 (explaining how ILD and ITD "is modeled by" HRTFs); *see also id.*, 12:57-65; Kyriakakis ¶289. A POSA would have understood Li applies one HRTF (from the bank of HRTFs) to produce a spatialized 3D sound for a specific position, and applies a different HRTF (*change the HRTF*) to position the 3D sound at a new position (*SLP*). *See id.*; *see also id.*, 5:26-30, 10:55-11:9, 12:37-13:7, 15:53-16:4, 20:37-52, 22:44-52; Kyriakakis ¶289. Because each HRTF models a different combination of ILD and ITD cues, Li's selection of different HRTFs for different 3D sound positioning includes *changing ILD and ITD*. Kyriakakis ¶289.

Thus, based on Li's teachings and the knowledge in the art, a POSA would have known that the *ILD*, *ITD*, and *HRTF* are all sound processing techniques that would be *changed* in Li to *position the SLP* at the desired location. Kyriakakis ¶290.

XIII. GROUND 6: LI IN VIEW OF ANDREA AND BALLAS RENDERS OBVIOUS CLAIMS 2, 18

A. Motivation to Combine

It would have been obvious to a POSA to modify Li-Andrea to track the user's head orientation and adjust the sound so the sound appears fixed even as the user's head moves, as taught in Ballas. Kyriakakis ¶¶291-293. Li recognizes that to determine a sound source, humans rely on various cues including head motion. EX1006, 1:38-42. Both share a goal of providing a realistic experience for the user. Kyriakakis ¶¶292. To make the experience as realistic as possible, the POSA would have tracked the user's head orientation and movements and account for the tracking information in Li's sound field, using conventional head tracking and spatialization techniques taught by Ballas. *Id.*

Li's primary goal is to provide a sound field that closely resembles "what humans hear in the real world" and "mimic a real conference room environment." EX1006, 2:34-36, 22:1-5; *see also id.*, 2:22-24, 13:57-60. Ballas shares this goal of providing "a realistic experience" by providing "spatialized sound" for applications such as "teleconferencing and telepresence" and "virtual and augmented reality environments." EX1008, ¶¶2, 9, 10. For Li's applications such as conferences, concerts, or radio dramas (EX1006, 22:1-5), head tracking is key to make the experience as realistic as possible, because it allows the user to freely change

orientation and account for the user's head movements so that the origins of the sound sources appear fixed and therefore realistic to the user. Kyriakakis ¶292. Modifying Li-Andrea with Ballas's teachings would have allowed the user to move while staying connected to the event, and maintaining the sound field to appear fixed as the user moves and changes head orientation. *Id.*

The modification would have been a natural extension of Li's use of HRTFs to spatialize sound and generate a sound field. Kyriakakis ¶293. Like Li, Ballas discloses an "[HRTF] filter to produce spatialized audio." EX1008, Abstract; *see also id.*, ¶¶16-17, 25. Specifically, Ballas's "head tracking system determines the location and orientation of the listener's head and provides the measurements to a computer which processes audio signals" using an "[HRTF] filter for maintaining virtual location of the audio signals thereby allowing the user to change location and head orientation without degradation of the virtual location of audio signals." *Id.*, Abstract, ¶18; *see also id.*, ¶¶17, 19-20, 25-26, claims 1, 7. Thus, head tracking to adjust HRTFs would have allowed the user to change head location and orientation while maintaining fixed virtual sound locations. Kyriakakis ¶293. This would have advanced the common goals of both to provide a realistic environment in a teleconferencing application. *Id*.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Li-Andrea with Ballas. Kyriakakis ¶¶294-296. Li and Ballas both use HRTFs to generate spatialized binaural sound and provide such sound to a user over headphones on wearable devices. In Li-Andrea, the wearable device includes headphones worn on the user's head. EX1006, Fig. 9; Kyriakakis ¶294.

Consistent with Li-Andrea, Ballas's speakers 110 are on a "frame [] mounted on the user's head and moves as the head moves." EX1008, ¶25. The sound provided over the speakers to the user is "binaural in nature" and processed using HRTF filters to produce spatialized audio. *Id.* Accordingly, a POSA would have recognized that Li-Andrea and Ballas are structurally and functionally similar. Kyriakakis ¶295.

A POSA implementing the Li-Andrea system would have attached the Li-Andrea headset to Ballas's frame. Kyriakakis ¶296. Ballas's frame also includes a head tracking system, with measurements used to produce spatialized audio. EX1008, ¶25. This implementation would have merely involved a combination of known prior art elements (headset and head tracking system) according to known methods and techniques (accounting for head movement) to yield predictable results (producing spatialized sound that appears fixed as the user moves), and involves use of a known technique (HRTF) to improve a similar device (wearable communication

device) in the same way (filtering sounds to produce a sound field that appears fixed). Kyriakakis ¶296.

C. Dependent Claim 2

1. 2[a]

Li-Ballas teaches this limitation. Kyriakakis ¶¶297-298.

For the same reasons as explained for claim 18 in Ground 2 (Section IX.C.1), a POSA would have been motivated to mount the Li-Andrea headphones onto Ballas's wearable frame having head tracking, as this would have provided a more realistic virtual experience, allowing the user to move while fixing the perceived sound source origins. Kyriakakis ¶298; *see also* Section XIII.A (Motivation to Combine).

2. 2[b]

Li-Ballas teaches this limitation. Kyriakakis ¶¶299-300.

As explained for claim 18 in Ground 2 (Section IX.C.1), Ballas discloses *adjusting…the sound* as the *listening person chang[es] head orientation with respect to the SLP*. Kyriakakis ¶300. As explained for claim 17, Li explains that *ITD* is *adjust[ed]* to position a 3D sound, which the POSA would have understood to be performed to fix the sound's *SLP* while the user changes *head orientation*. Kyriakakis ¶300; EX1027, 39-40.

D. Dependent Claim 18

Li-Ballas teaches this limitation. Kyriakakis ¶¶301-303.

As explained in 15[d] and claim 17 in Ground 5, Li-Andrea teaches *a processor* that *executes instructions to change an ITD, an ILD, and an HRTF* to position *the SLP of the binaural voice to a location that is away from and proximate to the first person*, and as explained for claim 17, Li renders obvious *chang[ing] an ITD, an ILD, and an HRTF of the binaural sound*. *See* Section XIII.E.5-7, XIII.F.2 (Ground 5, 15[d]-15[d2], 17); Kyriakakis ¶302.

As explained for claim 18 in Ground 2, Ballas teaches *maintaining the SLP fixed at the location during the electronic phone call while the first person physically moves away from the location and changes a head orientation with respect to the location*. *See* Section IX.C (Ground 2, 18); Kyriakakis ¶303.

XIV. GROUND 7: LI IN VIEW OF ANDREA AND SANDGREN RENDERS OBVIOUS CLAIM 19

A. Motivation to Combine

It would have been obvious to modify Li-Andrea to adjust attenuation and reverberation of the sound to add artificial cues of virtual objects that do not actually exist but are perceived to exist, as taught in Sandgren, for the same reasons as discussed for Ground 3. Kyriakakis ¶¶304-309; Section X.A. Li recognizes that to determine a sound source, humans rely on various cues including reverberation. EX1006, 1:38-42. A POSA would have been motivated to combine these teachings

with Sandgren because the references have a shared goal of providing a realistic and immersive experience for the user. Kyriakakis ¶306. As discussed for Ground 3, Sandgren teaches that direct sound is not enough, and that reverberation and attenuation are important cues to fully simulate 3D audio. EX1017, ¶¶6, 12.

Moreover, the POSA would have understood to simulate sound-reflecting objects that do not actually exist but are perceived to exist in the virtual environment, using conventional sound enhancement techniques as taught by Sandgren. Kyriakakis ¶307. The primary goal of Li is to provide a sound field that closely resembles "what humans hear in the real world" and "mimic a real conference room environment." EX1006, 2:34-36, 22:1-5; *see also id.*, 2:22-24, 13:57-60. Li recognizes "[s]ounds are a constant presence in everyday life and offer rich cues about the environment." *Id.*, 1:24-25. "There is a need for accurately synthesizing such spatial sound to add to the immersiveness of a virtual environment." *Id.*, 1:46-47. Sandgren shares this goal, as discussed above for Ground 3. EX1017, ¶1; *see also id.*, Abstract.

Modifying Li-Andrea in accordance with Sandgren's teachings would have provided a more realistic and immersive experience for the user, for the same reasons explained above for Ground 3. Kyriakakis ¶308.

Moreover, the modification would have been a natural extension of the use of HRTFs in Li-Andrea to spatialize sound and generate a sound field, for the same

reasons explained above for Ground 3. Kyriakakis ¶309. A POSA would have recognized that the teachings of Sandgren would have enhanced the use of HRTFs in Li-Andrea for an immersive experience by introducing attenuation filters and reverberation generators indicative of virtual objects in the environment. Kyriakakis ¶309.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Li-Andrea with Sandgren, for the same reasons as explained for Ground 3. Kyriakakis ¶¶310-311; Section X.B. Specifically, both use HRTFs to generate spatialized binaural sound and provide such sound over headphones to a user, who perceives the sounds to originate from different SLPs. *See*, EX1017, ¶¶1, 46. Accordingly, a POSA would have recognized that Li-Andrea and Sandgren are structurally and functionally similar. Kyriakakis ¶310.

A POSA implementing the Li-Andrea system would have recognized that the HRTF processing that spatializes the sound would have been accompanied with Sandgren's attenuation filters and reverberation generators to artificially simulate the presence of virtual objects. Kyriakakis ¶¶311. Such implementation would have merely involved a combination of known prior art elements (HRTF and attenuation filters and reverberation generators) according to known methods and techniques (simulate sound-reflecting virtual objects) to yield predictable results (producing

sound so that the user perceives the virtual objects), and involves use of a known technique (simulating sound reflections) to improve a similar device (wearable communication device) in the same way (filtering sounds to produce a realistic and immersive sound field). *Id.*

C. Dependent Claim 19

1. Claim 19[a]

Li-Sandgren teaches this limitation. Kyriakakis ¶¶312-313.

Li discloses that *reverberation* is a "source[] of localization cues." EX1006, 1:39-42. It would have been obvious to implement Li's *processor* to perform this limitation, in view of Sandgren, for the same reasons discussed for Ground 3. Kyriakakis ¶313; Section X.C.1.

2. Claim 19[b]

Li-Sandgren teaches this limitation, for the same reasons discussed in Ground 3. *See* Section X.C.2 (Ground 3, 19[b]); Kyriakakis ¶314.

XV. GROUND 8: LI IN VIEW OF BALLAS RENDERS OBVIOUS CLAIMS 8-10, 12, 14

A. Motivation to Combine

For the same reasons as explained for the Li-Andrea combination in Ground 6, it would have been obvious to modify Li-alone to track the user's head orientation and adjust the sound so that the sound appears fixed even as the user's head moves,

as taught in Ballas. *See* Section XIII.A (Ground 6, Motivation to Combine); Kyriakakis ¶315.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in making the Ballas modification to Li-alone, for the same reasons as explained for Li-Andrea. *See* Section XIII.B (Ground 6, Reasonable Expectation of Success); Kyriakakis ¶316.

C. Independent Claim 8

1. 8[pre]

To the extent limiting, Li discloses the preamble. Kyriakakis ¶¶317-318.

As explained for 15[pre], 15[c], and 15[d] Li discloses *a computer system* that has a memory (*a non-transitory computer readable storage medium*) *storing instructions* that cause a processor within the *computer system to execute a method*. *See* Section XII.E.1, 4-5 (Ground 5, 15[pre], 15[c]-15[d]); Kyriakakis ¶318.

2. 8[a]

Li discloses this limitation. Kyriakakis ¶¶319-320.

As explained for 1[a], Li discloses an electronic phone call (*electronic call*) that takes place (and is therefore *commence[d]* by Li's *computer system*) between a first user (*first individual wearing an electronic earphone*) and a second user (*second*

individual who is *remote* from the first user). *See* Section XII.C.2 (Ground 5, 1[a]); Kyriakakis ¶320.

3. 8[b]

Li discloses this limitation. Kyriakakis ¶¶321-324.

As explained for 1[b], Li describes its *computer system* as *designat[ing]…an SLP* to the first user (*first individual*). *See* Section XII.C.3 (Ground 5, 1[b]); Kyriakakis ¶¶322-324. In Li, that *designat[ion]* is *for the first individual* because the designated sound is delivered to the first user (*first individual*) *through* the headphones (*electronic earphone*). *Id.* The *SLP simulates a location…from where an origin of a voice of* the second user (*second individual*) *occurs for the* first user (*first individual*). *Id.*

4. 8[c]

Li-Ballas teaches this limitation. Kyriakakis ¶¶325-327.

As explained for 1[d], Li's *computer system execute[s] the* electronic phone call (*electronic call*) *such that the origin of the voice of the* second user (*second individual*) *appears to the* first user (*first individual*) *to originate from…the SLP*. *See* Section XII.C.5 (Ground 5, 1[d]). As explained for claim 18, Li-Ballas teaches that the *SLP* is a *fixed location…while the first individual moves and changes a location and a head orientation with respect to the SLP*. *See* Section XIII.A-B, D (Ground 6, 18); Kyriakakis ¶¶326-327.

5. 8[d]

Li-Ballas teaches this limitation. Kyriakakis ¶¶328-332.

As explained for 8[c], Li-Ballas discloses fixing (*maintaining*) *the origin of the voice of the second individual at the SLP in response to the first individual moving the head orientation with respect to the SLP*. *See* Section XV.C.4 (Ground 8, 8[c]); Kyriakakis ¶329.

As explained for 2[b] and claims 17-18, Li discloses *alter[ing] an ITD of the voice that the first individual hears*. *See* Section XII.D.1 and XII.F.2 (Ground 1, 2[b] and 17) and Section XIII.C (Ground 2, 18); Kyriakakis ¶330. A POSA would have appreciated that it was commonplace to alter ITDs to maintain the origin of a sound while the user moves head orientation. Kyriakakis ¶¶331-332; *see also* Section IV.A (Technology Overview).

D. Dependent Claims 9, 10, 12, 14

1. Claim 9

Li-Ballas teaches this limitation. Kyriakakis ¶¶333-336.

As explained for 8[c] and 8[d], Li-Ballas teaches *maintaining the origin of the voice of the second individual at the SLP…in response to the first individual moving closer to and away from the SLP*. *See* Sections XV.C.4-5 (Ground 8, 8[c]-8[d]); Kyriakakis ¶334.

As explained for claims 17-18, Li discloses *alter[ing] an ILD of the voice that the first individual hears*. *See* Section XII.F.2 (Ground 1, 17), XIII.C (Ground 2, 18); Kyriakakis ¶335. A POSA would have appreciated that it was commonplace to alter ILDs to maintain the origin of a sound while the user moves nearer and further away from the SLP. EX1035, 944; Kyriakakis ¶336; *see also* Section IV.A (Technology Overview).

2. Claim 10

a) 10[a]

Li-Ballas teaches this limitation. Kyriakakis ¶¶337-339.

As explained for 2[a], Ballas discloses a tracking system for tracking the user's head orientation and location (*monitor movements of the head orientation of the first individual*). *See* Section XIII.C.1 (Ground 2, 2[a]); EX1008, ¶¶25-26; Kyriakakis ¶338. As explained for 1[b], Li's first user (*listening person*) "configures the placements" of the sound sources (*localizing the voice of the second individual during the electronic call*). *See* Section XII.C.3 (Ground 5, 1[b]); Kyriakakis ¶339.

b) 10[b]

Li-Ballas teaches this limitation. Kyriakakis ¶¶340-342.

As explained for claim 18 and 10[a], Li-Ballas renders obvious changing (*adjust*) *HRTFs* to process audio signals and maintain a fixed location for a sound's

SLP *in response to the movements of the head orientation of the first individual localizing the voice of the second individual during the electronic call*. *See* Section XIII.C (Ground 2, 18) and Section XV.D.2 (Ground 4, 10[a]); Kyriakakis ¶341. A POSA would have understood that changing HRTFs here involves *adjust[ing] HRTFs of the first individual to more accurately simulate real HRTFs of the first individual*, because head tracking and adjusting HRTFs would have provided a more realistic virtual experience for the user, by allowing the user to move while fixing the perceived origin of sound sources. EX1008, ¶¶2, 17; Kyriakakis ¶342.

3. Claim 12

a) 12[a]

Li discloses this limitation. Kyriakakis ¶¶343-344.

By disclosing an *electronic call* as a conference call, where the participants include at least a first user who "plac[es] different speaking users" in the sound field, Li discloses *add[ing] a third individual to the electronic call*. EX1006, 22:1-5, 10:5-6, 21:57-60; Kyriakakis ¶344.

b) 12[b]

Li discloses this limitation. Kyriakakis ¶¶345-347. For the same reasons as explained for 1[b], 6[b], and 8[b], Li teaches *designat[ing], for the first individual, an SLP that simulates a location in empty space from where an origin of a voice of the third individual occurs for the first individual through the electronic earphone.*

Sections XII.C.3 and XII.D.5 (Ground 5, 1[b] and 6[b]) and Section XV.C.3 (Ground 8, 8[b]). As Li "mimics a real conference room environment," the *SLP of the third individual is adjacent to SLP of the second individual*. EX1006, 22:1-5; Kyriakakis ¶¶346-347.

4. Claim 14

Li discloses this limitation. Kyriakakis ¶¶348-351.

Li's configurable 3D sound system 900 "acquires [*receives*] configuration information 1601 [*command*] from a user" that includes "user selections of configurable parameters, for example, an azimuth, an evaluation, a distance, a trace of movement, etc." and "give[s] the user the freedom of configuring the positions and movements of multiple sound tracks." EX1006, 20:20-36; *see also* 15:27-32, 21:57-22:5, 15:44-47. By allowing the user to provide input to move the sound source, Li discloses *receiving, from the first individual, configurable parameters and* configuration information (*commands*) *to move the* position (*SLP of the voice*) *from the first point to a second location that is* next to (*proximate*) *the first location and* a distance from (*away from and proximate to*) *the user*. Kyriakakis ¶¶349-351.

XVI. SECONDARY CONSIDERATIONS

There are no known secondary considerations. Should PO argue secondary considerations, Petitioner will seek leave to reply.

XVII. THE BOARD SHOULD REACH THE MERITS OF THIS PETITION

The Board should reach the merits of this petition and should not exercise its discretion to deny based on 35 U.S.C. § 314(a) or § 325(d). None of the prior art at issue in this petition was cited or relied upon during prosecution of the ’090 patent. As for the district court case, the court is currently deciding a motion to transfer. *See Eight kHz, LLC v. Meta Platforms, Inc.*, No. 6:22-cv-00575-ADA-DTG, Dkt. Nos. 46, 48, 90 (W.D. Tex.).

Petitioner reserves the right to address any discretionary issues raised by Patent Owner based on the facts and law as they stand at that time.

XVIII. CONCLUSION

For these reasons, Petitioner respectfully requests IPR of claims 1-3, 6, 8-10, 12, 14-19 of the ’090 Patent.

Date: June 9, 2023

Respectfully submitted,

By: /Lisa K. Nguyen/

Lisa K. Nguyen (Reg. No. 58,018)
lisa.nguyen@allenovery.com
Allen & Overy LLP
550 High Street
Palo Alto, CA 94301
Telephone: 650.388.1724

Counsel for Petitioner
Meta Platforms, Inc.

CLAIM LISTING APPENDIX

1[pre] A method, comprising:

[a] capturing, with an electronic earphone located at a head of a talking person, binaural sound that will be provided to a listening person during a telephone call;

[b] designating, with a computer system, a sound localization point in empty space that is away from and proximate to the listening person such that the sound localization point simulates an origin of the binaural sound at the empty space that the listening person hears during the telephone call;

[c] adjusting, with the computer system, the binaural sound captured at the earphone of the talking person so the binaural sound originates during the telephone call from the sound localization point in empty space that is away from and proximate to the listening person; and

[d] providing, with an electronic earphone located at a head of the listening person, the binaural sound to the listening person during the telephone call such that the origin of the binaural sound for the listening person occurs at the sound localization point in empty space that is away from and proximate to the listening person.

2[pre] The method of claim 1 further comprising:

[a] tracking a head orientation of the listening person with respect to the sound localization point;

[b] adjusting an interaural time difference of the sound that the listening person hears in response to the listening person changing the head orientation with respect to the sound localization point.

3. The method of claim 1 further comprising:

receiving, from the listening person, a location that specifies a distance from the listening person where to position the sound localization point in empty space that is proximate to and away from the listening person.

6[pre] The method of claim 1 further comprising:

[a] capturing binaural sound from a second talking person;

[b] designating a second sound localization point in empty space that is away from and proximate to the listening person such that the second sound localization point simulates an origin of the binaural sound from the second talking person that the listener hears;

[c] positioning the second sound localization point adjacent to the sound localization point such that the binaural sound captured from the talking person appears to originate from the sound localization point and the binaural sound captured from the second talking person appears to originate from the second sound localization point.

8[pre] A non-transitory computer readable storage medium storing instructions that cause a computer system to execute a method, comprising:

[a] commence an electronic call between a first individual wearing an electronic earphone and a second individual that are remote from each other;

[b] designate, for the first individual, a sound localization point that simulates a location in empty space from where an origin of a voice of the second individual occurs for the first individual through the electronic earphone such that the sound localization point in empty space is located away from and proximate to the first individual;

[c] execute the electronic call such that the origin of the voice of the second individual appears to the first individual to originate from a fixed location that is the sound localization point in empty space while the first individual moves and changes a location and a head orientation with respect to the sound localization point; and

[d] maintain the origin of the voice of the second individual at the sound localization point but alter an interaural time difference of the voice that the first individual hears in response to the first individual moving the head orientation with respect to the sound localization point.

9. The non-transitory computer readable storage medium storing instructions of claim 8 further to cause the computer system to execute the method comprising:

maintain the origin of the voice of the second individual at the sound localization point but alter an interaural level difference of the voice that the first individual hears in response to the first individual moving closer to and away from the sound localization point.

10[pre] The non-transitory computer readable storage medium storing instructions of claim 8 further to cause the computer system to execute the method comprising:

[a] monitor movements of the head orientation of the first individual in response to the first individual localizing the voice of the second individual during the electronic call;

[b] adjust head-related transfer functions (HRTFs) of the first individual to more accurately simulate real HRTFs of the first individual in response to the movements of the head orientation of the first individual localizing the voice of the second individual during the electronic call.

12[pre] The non-transitory computer readable storage medium storing instructions of claim 8 further to cause the computer system to execute the method comprising:

[a] add a third individual to the electronic call;

[b] designate, for the first individual, a sound localization point that simulates a location in empty space from where an origin of a voice of the third individual occurs for the first individual through the electronic earphone such that the sound localization point of the third individual is adjacent to the sound localization point of the second individual.

14. The non-transitory computer readable storage medium storing instructions of claim 8 further to cause the computer system to execute the method comprising:

receive, from the first individual, commands that move the sound localization point from a first point in empty space that is away from and proximate to the first individual to a second point in empty space that is away from and proximate to the first individual.

15[pre] A computer system, comprising:

[a] a first wearable electronic device that a first person wears;

[b] a second wearable electronic device that a second person wears and that captures a binaural voice of the second person during an electronic phone call with the first wearable electronic device that the first person wears;

[c] a memory that stores instructions; and

[d] a processor that executes the instructions to

[d1] adjust the binaural voice of the second person during the electronic phone call in order to position a sound localization point of the binaural voice to a location that is away from and proximate to the first person,

[d2] wherein the sound localization point represents an origin of the binaural voice of the second person to the first person.

16. The computer system of claim 15, wherein the processor further executes the instructions to determine a location of an empty chair that is proximate to the first person and position the sound localization point at the empty chair.

17. The computer system of claim 15, wherein the processor further executes the instructions to change an interaural time difference (ITD), an interaural level difference (ILD), and a head-related transfer function (HRTF) in order to position the sound localization point of the binaural voice to the location that is away from and proximate to the first person.

18. The computer system of claim 15, wherein the processor further executes the instructions to change an interaural time difference, an interaural level difference, and a head-related transfer function of the binaural sound in order to maintain the sound localization point fixed at the location during the electronic phone call while the first person physically moves away from the location and changes a head orientation with respect to the location.

19. The computer system of claim 15,

[a] wherein the processor further executes the instructions to change attenuation and reverberation of the binaural sound in order to add artificial audial cues of virtual objects near the second person,

[b]wherein the virtual objects do not actually exist near the second person but are perceived to exist by the first person from the artificial audial cues.

CERTIFICATE OF COMPLIANCE WITH 37 C.F.R. § 42.24

I hereby certify that this Petition complies with the word count limitation of 37 C.F.R. § § 42.24(a)(1)(i) because the Petition contains a total of 13,970 words, calculated by Microsoft Word's word-count feature. This total excludes the cover page, signature block, and the parts of the Petition exempted by Rule 37 C.F.R. § 42.24(a)(1).

Date: June 9, 2023

By: /Lisa K. Nguyen/

Lisa K. Nguyen (Reg. No. 58,018)
lisa.nguyen@allenovery.com
Allen & Overy LLP
550 High Street
Palo Alto, CA 94301
Telephone: 650.388.1724

Counsel for Petitioner Meta Platforms, Inc.

CERTIFICATE OF SERVICE

The undersigned certifies that a complete copy of this Petition for *Inter Partes* Review of U.S. Patent No. 9,226,090 and all Exhibits and other documents filed together with this Petition were served on the official correspondence address for the patent shown in PAIR and a courtesy copy to Eight kHz, LLC's current litigation counsel via FEDERAL EXPRESS next business day delivery on June 9, 2022:

Philip S. Lyren
Philip S. Lyren, PC
289 Woodland Avenue
Wadsworth, OH 44281

By: /Lisa K. Nguyen/

Lisa K. Nguyen (Reg. No. 58,018)
lisa.nguyen@allenovery.com
Allen & Overy LLP
550 High Street
Palo Alto, CA 94301
Telephone: 650.388.1724

Counsel for Petitioner Meta Platforms, Inc.