

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re *Inter Partes* Review of:)
U.S. Patent No. 9,282,196)
Issued: Mar. 8, 2016)
Application No.: 14/946,806)
Filing Date: Nov. 20, 2015)

**For: Moving a Sound Localization Point of a Computer Program During a
Voice Exchange**

FILED VIA P-TACTS

**PETITION FOR *INTER PARTES* REVIEW OF
U.S. PATENT NO. 9,282,196**

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	MANDATORY NOTICES	3
	A. Real Parties-in-Interest	3
	B. Related Matters.....	3
	C. Notice of Counsel and Service Information.....	3
	D. Fee for IPR	5
	E. 37 C.F.R. § 42.104(a) Certification.....	5
III.	CHALLENGES UNDER 35 U.S.C. § 103	5
IV.	BACKGROUND	6
	A. Technology Overview	6
	B. '196 Patent.....	10
	C. Challenged Claims	11
	D. Prosecution History	11
V.	PRIOR ART.....	11
	A. Li.....	11
	B. Bedingfield	12
	C. Ballas	13
	D. Sandgren.....	13
	E. Woo	13

VI.	LEVEL OF ORDINARY SKILL IN THE ART	14
VII.	CLAIM CONSTRUCTION	14
VIII.	GROUND 1: LI RENDERS OBVIOUS CLAIMS 1, 8, 15, 17.....	15
A.	Motivation to Modify	15
B.	Reasonable Expectation of Success	17
C.	Independent Claims 1 and 8:.....	18
1.	1[pre]/8[pre]: Move SLP of Voice of Computer Program.....	18
2.	1[a]/8[a]: Provide Through Earphone/WED Binaural Sound Such That Origin Occurs at SLP	24
3.	1[b]/8[b]: Receive Command to Move SLP	27
4.	1[c]/8[c]: Move SLP	28
5.	1[d]/8[d]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location	29
D.	Independent Claim 15:	29
1.	15[pre]: Move SLP of Voice of Computer Program	29
2.	15[a]: WED Provides Voice of Computer Program at SLP	29
3.	15[b]: Memory.....	30
4.	15[c]: Processor	30
5.	15[c1]: Receive Command to Move SLP.....	30
6.	15[c2]: Move SLP.....	30
7.	15[c3]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location	31

E.	Dependent Claim 17: Change ITD and ILD to Move SLP	31
IX.	GROUND 2: LI IN VIEW OF BALLAS RENDERS OBVIOUS CLAIMS 4, 6, 9-10, 14, 18	32
A.	Motivation to Combine	32
B.	Reasonable Expectation of Success	33
C.	Dependent Claim 4	35
1.	4[a]: Track Head	35
2.	4[b]: Adjust Voice To Continue To Appear To Originate From SLP	36
D.	Dependent Claim 6: Change Amplitude, ITD, and ILD to Continue to Originate from SLP as User Moves	37
E.	Dependent Claim 9: Maintain SLP But Alter Amplitude	38
F.	Dependent Claim 10	39
1.	10[a]: Monitor Movements	39
2.	10[b]: Adjust ITD and Amplitude	39
G.	Dependent Claim 14: Adjust ILD, ITD, and Amplitude So SLP Remains Fixed	39
H.	Dependent Claim 18: Change ITD and Amplitude to Maintain SLP	39
X.	GROUND 3: LI IN VIEW OF BALLAS AND SANDGREN RENDERS OBVIOUS CLAIM 5	40
A.	Motivation	40
B.	Reasonable Expectation of Success	41
C.	Dependent Claim 5: Adjust Amplitude to be Inversely Proportional to Square of Distance	42

XI.	GROUND 4: LI IN VIEW OF SANDGREN RENDERS OBVIOUS CLAIM 19	42
A.	Motivation to Combine	42
B.	Reasonable Expectation of Success	44
C.	Dependent Claim 19:.....	44
1.	19[a]: Change Attenuation and Reverberation.....	44
2.	19[b]: Virtual Objects.....	46
XII.	GROUND 5: LI AND WOO RENDERS OBVIOUS CLAIM 11.....	47
A.	Motivation to Combine	47
B.	Reasonable Expectation of Success	48
C.	Dependent Claim 11:.....	49
1.	11[a]: Physical Objects.....	49
2.	11[b]: Adjust Voice to Emulate Sound	49
XIII.	GROUND 6: BEDINGFIELD RENDERS OBVIOUS CLAIMS 1, 8, 15, 17	50
A.	Motivation to Modify	50
1.	<i>Voice of a computer program</i>	51
2.	<i>HPED</i>	51
B.	Reasonable Expectation of Success	52
1.	<i>Voice of a computer program</i>	52
2.	<i>HPED</i>	53
C.	Independent Claims 1 and 8	53

1.	1[pre]/8[pre]: Move SLP of Voice of Computer Program.....	53
2.	1[a]/8[a]: Provide Through Earphone/WED Binaural Sound Such That Origin Occurs at SLP	57
3.	1[b]/8[b]: Receive Command to Move SLP.....	60
4.	1[c]/8[c]: Move SLP	61
5.	1[d]/8[d]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location	62
D.	Independent Claim 15	63
1.	15[pre]: Move SLP of Voice of Computer Program.....	63
2.	15[a]: WED Provides Voice of Computer Program at SLP	63
3.	15[b]: Memory	63
4.	15[c]: Processor.....	63
5.	15[c1]: Receive Command to Move SLP	63
6.	15[c2]: Move SLP	64
7.	15[c3]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location.....	64
E.	Dependent Claim 17: Change ITD and ILD to Move SLP.....	64
XIV.	GROUND 7: BEDINGFIELD IN VIEW OF BALLAS RENDERS OBVIOUS CLAIM 4, 6, 9-10, 14, 18	65
A.	Motivation to Combine	65
B.	Reasonable Expectation of Success	67
C.	Dependent Claim 4.....	68
1.	4[a]: Track Head	68

2.	4[b]: Adjust Voice To Continue To Appear To Originate From SLP	68
D.	Dependent Claim 6: Change Amplitude, ITD, and ILD to Continue to Originate from SLP as User Moves	69
E.	Dependent Claim 9: Maintain SLP But Alter Amplitude	70
F.	Dependent Claim 10	70
1.	10[a]: Monitor Movements	70
2.	10[b]: Adjust ITD and Amplitude	70
G.	Dependent Claim 14: Adjust ILD, ITD, and Amplitude So SLP Remains Fixed	71
H.	Dependent Claim 18: Change ITD and Amplitude to Maintain SLP	71
XV.	GROUND 8: BEDINGFIELD IN VIEW OF BALLAS AND SANDGREN RENDERS OBVIOUS CLAIM 5	72
A.	Motivation to Combine	72
B.	Reasonable Expectation of Success	72
C.	Dependent Claim 5: Adjust Amplitude to be Inversely Proportional to Square of Distance	73
XVI.	GROUND 9: BEDINGFIELD IN VIEW OF SANDGREN RENDERS OBVIOUS CLAIM 19	74
A.	Motivation to Combine	74
B.	Reasonable Expectation of Success	75
C.	Dependent Claim 19	76
1.	19[a]: Change Attenuation and Reverberation	76
2.	19[b]: Virtual Objects	76

XVII. GROUND 10: BEDINGFIELD AND WOO RENDERS OBVIOUS CLAIM 11	76
A. Motivation to Combine	77
B. Reasonable Expectation of Success	78
C. Dependent Claim 11	79
1. 11[a]: Physical Objects	79
2. 11[b]: Adjust Voice to Emulate Sound.....	79
XVIII. SECONDARY CONSIDERATIONS	79
XIX. THE BOARD SHOULD REACH THE MERITS OF THIS PETITION.....	79
XX. CONCLUSION.....	80

TABLE OF AUTHORITIES

Page(s)

Cases

<i>Apple Inc. v. AGIS Software Development, LLC</i> , IPR2018-00819, Paper 9 (Nov. 7, 2018)	17
<i>Cienna Corp. v. Oyster Optics, LLC</i> , IPR2021-00565, Paper 11 (Aug. 23, 2021)	<i>passim</i>
<i>Eight KHZ, LLC v. Meta Platforms, Inc. et al.</i> , No. 2:22-cv-000575 (W.D. Tex.)	3, 92
<i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575-ADA-DTG, Dkt. Nos. 46, 48, 90 (W.D. Tex.)	83
<i>Western Digital Corp. v. SPEX Techs., Inc.</i> , IPR2018-0084, Paper 14 (April 25, 2018)	17

Statutes

35 U.S.C. § 102(a)(1)	6
35 U.S.C. § 102(a)(2)	5, 6
35 U.S.C. § 103	<i>passim</i>
35 U.S.C. § 314(a)	83

Other Authorities

37 C.F.R. § 42.10(b)	5
37 C.F.R. § 42.15(a)	5
37 C.F.R. § 42.24	90
37 C.F.R. § 42.24(a)(1)	90
37 C.F.R. § 42.24(a)(1)(i)	90
37 C.F.R. § 42.104(a)	5

37 C.F.R. § 42.104	5
--------------------------	---

EXHIBIT LIST

Exhibit	Description
1001	U.S. Patent No. 9,282,196 (“196 Patent”)
1002	Prosecution History of U.S. Patent No. 9,282,196 (“196 FH”)
1003	Declaration of Dr. Kyriakakis
1004	<i>Curriculum Vitae</i> 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	RESERVED
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	U.S. Patent Pub. 2002/0151996 to Wilcock (“Wilcock”)
1015	Rumsey, Francis “Whose head is it anyway? Optimizing binaural audio.” J. Audio Eng. Soc., Vol. 59, No. 9, 2011 (“Rumsey”)
1016	U.S. Patent Pub. No. 2009/0060235 to Woo (“Woo”)
1017	U.S. Patent App. Pub. 2008/0273708 to Sandgren, et al. (“Sandgren”)

Exhibit	Description
1018	RESERVED
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	<i>Eight KHZ, LLC v. Meta Platforms, Inc. et al.</i> , 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 <i>Eight KHZ, LLC v. Meta Platforms, Inc. et al.</i> , 6:22-cv-575-ADA-DTG (W.D. Tex.)
1033	U.S. Patent No. 257,453 to Ader (“Ader”)

Exhibit	Description
1034	Olson, Harry. “Sound Reproducing Systems - Monaural, Binaural, Monophonic, and Stereophonic.” Audio, 1958 (“Olson”)
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	Berkley, D.A., et al. “A Multimodal Teleconference System Using Hands-Free Voice Control.” 3rd International Conference on Spoken Language Processing, 1994 (“Berkley”)
1041	Fagel, Sascha, et al. “An Expandable Web-based Audiovisual Text-to-Speech Synthesis System.” 8th European Conference on Speech Communication and Technology, 2003 (“Fagel”)
1042	Taylor, Paul, Text-to-Speech Synthesis (Cambridge University Press 2009)
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, 4-6, 8-11, 14-15, and 17-19 of U.S. Patent 9,282,196 (EX1001, “’196 Patent”).

The foundational virtual reality technology began in the 1960s with the Telesphere Mask, Headsight, and Ultimate Display. Binaural sound simulation advanced with virtual reality advancements. 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 the sound originates) were well-known, well-understood, and routinely used.

The ’196 Patent describes a known system that provides binaural sound at an SLP and moves the SLP in response to a user command. But the ’196 Patent merely discloses known techniques. Notably, nowhere does the ’196 Patent describe the difference between the claimed subject matter and prior art. Indeed, there is no patentable difference. Three-dimensional sound systems that allowed users to move the SLP were prevalent before 2014. For example, in U.S. Patent 9,131,305 to Li (EX1006, “Li”), former Bell Labs scientists taught a three-dimensional sound system for “experiencing music, movies, video games” configured to allow for “movements of the sound sources” based on user input. *Id.*, 7:21-22, 15:8. In one application, the system simulates “a conference call, by placing different speaking

users at different spots to mimic a real conference room environment” through a GUI. *Id.*, 22:1-5.

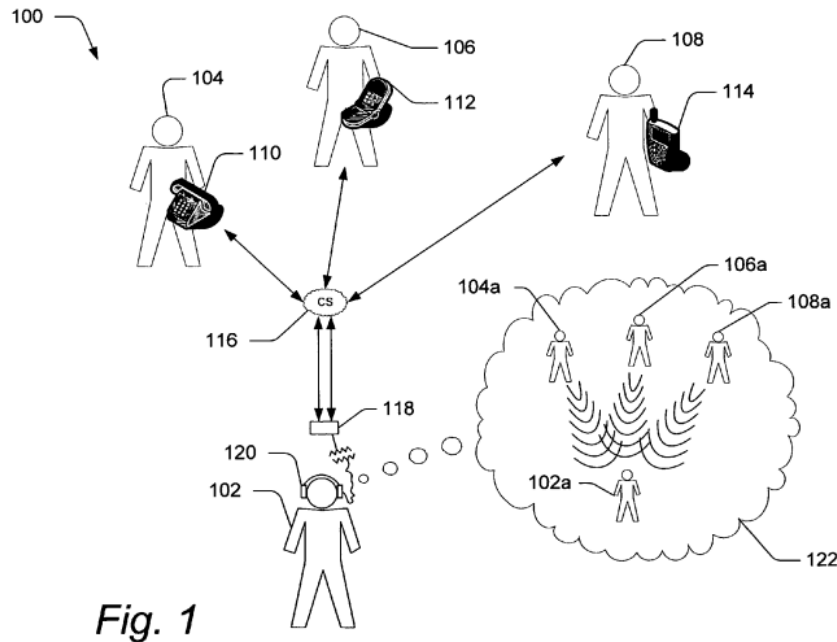


Fig. 1

EX1005, Fig. 1

In U.S. Patent 8,406,439 to Bedingfield (EX1005, “Bedingfield”), AT&T developed synthetic audio placement, creating a sound field around a listener during a conference call that allows the listener to hear sounds as if the participants “were in the same location and speaking to each other directly” and allows “the user [to] select[] a participant to move.” *Id.*, Abstract, 4:56-57, 17:23-24.

Petitioner requests that the Board institute review and find all challenged claims 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 '196 Patent is assigned to Eight KHZ, LLC ("PO"), which is currently asserting the '196 Patent in the following concurrent litigation filed on June 3, 2022: *Eight KHZ, LLC v. Meta Platforms, Inc. et al.*, No. 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. 9,226,090 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@allenoverly.com	David M. Tennant (Reg. No. 48,362) david.tennant@allenoverly.com

Postal & Hand-Delivery Address: Allen & Overy LLP 550 High Street Palo Alto, CA 94301 Telephone: (650) 388-1724	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@allenoverly.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@allenoverly.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@allenoverly.com Postal & Hand-Delivery Address: Allen & Overy LLP 1101 New York Avenue, NW Washington, DC 20005 Tel: (202) 683-3862 Eric E. Lancaster (<i>pro hac vice</i> to be filed) eric.lancaster@allenoverly.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 email, using the email addresses identified above.

D. Fee for IPR

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

E. 37 C.F.R. § 42.104(a) Certification

Petitioner certifies that the '196 Patent is available for IPR and that Petitioner is not barred or estopped from requesting IPR challenging the claims on the grounds identified herein.

III. CHALLENGES UNDER 35 U.S.C. § 103

Grounds	Claims	Prior Art
1	1, 8, 15, 17	Li ¹
2	4, 6, 9-10, 14, 18	Li in view of Ballas ²

¹ Li is at least 35 U.S.C. § 102(a)(2) prior art.

² U.S. Patent Pub. 2003/0059070 to Ballas (EX1008, “Ballas”) is at least 35 U.S.C. § 102(a)(1)/(2) prior art.

3	5	Li in view of Ballas and Sandgren ³
4	19	Li in view of Sandgren
5	11	Li in view of Woo ⁴
6	1, 8, 15, 17	Bedingfield ⁵
7	4, 6, 9-10, 14, 18	Bedingfield in view of Ballas
8	5	Bedingfield in view of Ballas and Sandgren
9	19	Bedingfield in view of Sandgren
10	11	Bedingfield in view of Woo

IV. BACKGROUND

A. Technology Overview

In the 1960s, early virtual reality pioneers invented head mounted displays (“HMDs”) to provide immersive and realistic simulated experiences. EX1003

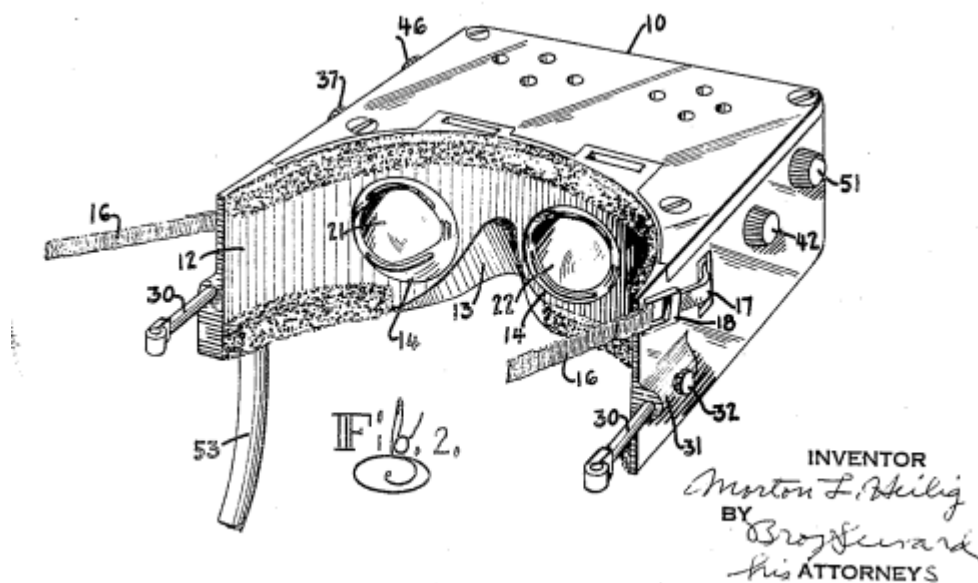
³ U.S. Patent Pub. 2008/0273708 to Sandgren (EX1017, “Sandgren”) is at least 35 U.S.C. § 102(a)(1)/(2) prior art.

⁴ U.S. Patent Pub. 2009/0060235 to Woo (EX1016, “Woo”) is at least 35 U.S.C. § 102(a)(1)/(2) prior art.

⁵ Bedingfield is at least 35 U.S.C. § 102(a)(1)/(2) prior art.

(“Kyriakakis”) ¶¶36-40. The earliest wearable electronic devices (“WEDs”) had displays synchronized with sound and some included head tracking systems that adjusted the displays and sound in response to head movements. *Id.*

In 1960, Heilig patented the first HMD having earphones and a screen for each eye. EX1044, Fig. 2 reproduced below; EX1020, 4; Kyriakakis ¶37.



In 1961, Comeau and Bryan developed “Headsight,” with head tracking to provide images and sounds synchronized with head movements. EX1020, 5; EX1021, 3-7; Kyriakakis ¶38. Later, 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; Kyriakasis ¶¶39-40.

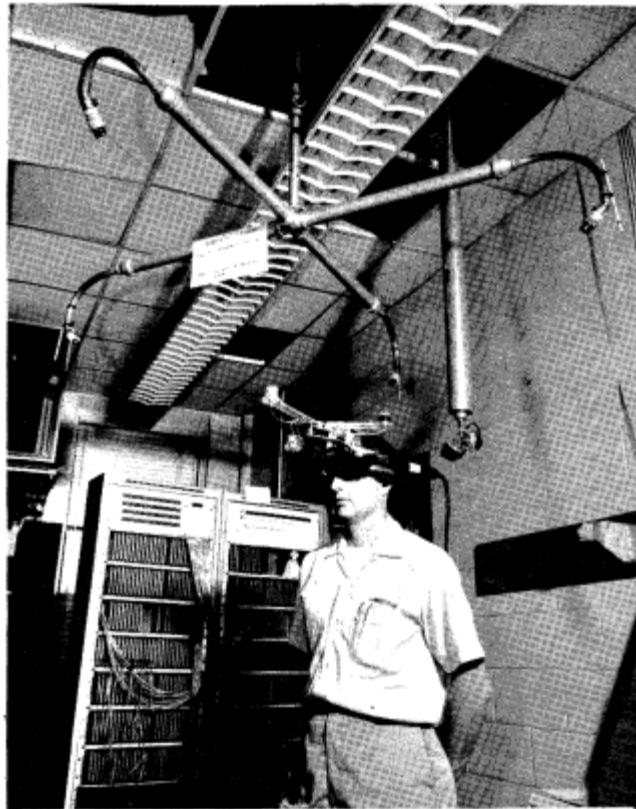


FIGURE 3—The mechanical head position sensor in use

As wearable HMD technology developed, so did binaural sound simulation. Kyriakakis ¶¶41-64. To determine a sound location, humans process and compare sound cues delivered to each ear. *Id.* ¶¶46-51. When a sound source is in front of a human, sound generally arrives at each ear at the same time. *Id.* ¶50.

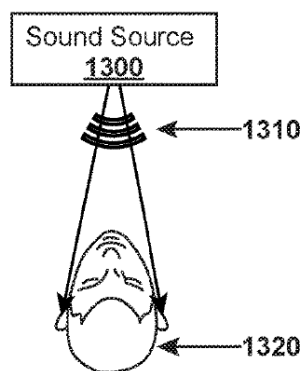


Figure 13A

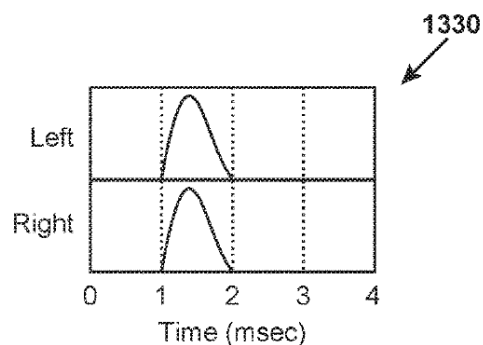


Figure 13B

EX1001, Figs. 13A-13B

When the sound source is to the left of the listener, the sound will arrive at the left ear slightly before the right ear, and the sound at the right ear will be less intense. *Id.* ¶48. The difference in arrival time across the ears is the interaural time difference (ITD). *Id.*

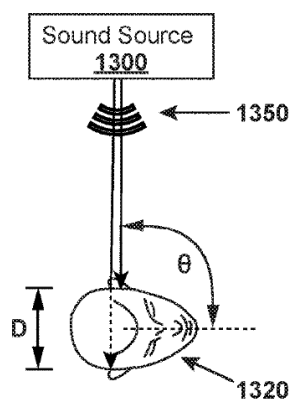


Figure 13C

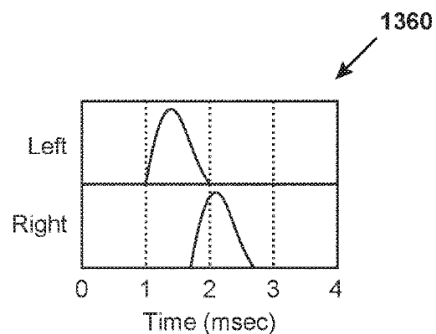


Figure 13D

EX1001, Figs. 13C-13D

Reverberation, reflection, refraction, and attenuation are other factors that impact how an ear receives sound from a point in space, and the shape of the listener's ears, head, shoulders, and torso can change the sound. *Id.* ¶50.

For decades, a head related transfer function (HRTF) has been used to capture these factors. EX1024, 1; EX1025, 5. 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. 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. '196 Patent

The '196 Patent is directed to well-known systems and methods of providing sound “comparable in quality to sound that the listeners would hear if they were communicating face-to-face with each other.” EX1001, 1:15-16. The '196 Patent describes “mov[ing] the SLPs” (*id.*, 36:12-17) and adjusting the sound “based on calculations made for sound attenuation and sound reverberation” (*id.*, 10:48-50). These may use Stoke's law for attenuation, Sabine's equation for reverberation, or other known equations. *Id.*, 19:46-59, 17:30-44.

Notably, the '196 Patent does not describe moving the SLP in response to a command as an inventive aspect.

C. Challenged Claims

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

D. Prosecution History

The application for the '196 Patent was filed on November 20, 2015. The examiner issued no rejection. *See generally* EX1002.

On January 27, 2016, the Examiner allowed the pending claims. *Id.*, 143-149. The '196 Patent issued on March 8, 2016.

V. PRIOR ART

A. Li

Li discloses all limitations of the independent claims in its conference call embodiment except *the voice of the computer program* which it discusses separately. Kyriakakis ¶76. Li was not cited or discussed during prosecution.

Li's 3D sound system allows users to “experienc[e] music, movies, video games” and to provide user inputs for “movements of the sound sources.” EX1006, 7:21-22, 15:8, 16:4-7, 16:63-66, 19:39-20:52, Figs. 9, 16, 33; Kyriakakis ¶78.

Like the '196 Patent, Li “generate[s] configurable 3D sounds” by processing the speaker's voice to position “different speaking users at different spots” during a

conference call. EX1006, 22:1-5, 33:14-18, 33:32-51, 33:62-66, 34:4-8, 34:20-51, 35:58-36:2; Kyriakakis ¶78. Li's "configurable sound field [] provides an interface to give the user the freedom of configuring the positions and movements of multiple sound tracks, in order to render a customized 3D binaural sound." EX1006, 20:26-36; Kyriakakis ¶78. By allowing the user to provide commands to "move the sound source up and down, etc., on the configurable sound field," Li moves the sound's SLP based on the user's command. EX1006, 20:20-36, 21:57-22:5, Kyriakakis ¶78.

B. Bedingfield

Bedingfield generally discloses all limitations of the independent challenged claims in its conference call system except *the voice of the computer program*. Kyriakakis ¶79. Bedingfield was not cited or discussed during prosecution.

Bedingfield teaches synthetic audio placement for different audio sources. EX1005, Abstract, 1:15-17, 1:50-2:16, 4:18-19. Bedingfield describes an "improved telephone or network-based communication system [that] may more closely simulate an in-person conference setting." *Id.*, 3:25-27; Kyriakakis ¶80.

Like the '196 Patent, Bedingfield provides the sound field over headphones so that the listener perceives the other remote conference participants as if they "were in the same location and speaking to each other directly." EX1005, 4:31-5:3, Fig. 1. Bedingfield's user makes selections including new locations for the other participants in the sound field, meaning their SLPs move within the sound field.

EX1005, 5:51-61, 6:43-45, Fig. 3; Kyriakakis ¶81.

C. Ballas

Ballas discloses head tracking and maintaining an SLP even as the user's head moves. Kyriakakis ¶82. Ballas was not cited or discussed during prosecution.

Like the '196 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 provides the head location and orientation to a computer system that "processes the audio signals using a [HRTF] filter 106 thus producing spatialized audio." *Id.*, ¶¶17-18, 25-26. The HRTF filter "maintains the virtual location of the audio signals/sound" even as the user's head moves. *Id.*, Abstract, ¶17; Kyriakakis ¶83.

D. Sandgren

Sandgren discloses an inverse-square relationship between sound intensity and distance, as well as simulating virtual objects in 3D audio. EX1017, ¶¶6, 12, 77; Kyriakakis ¶84. Sandgren was not cited or discussed during prosecution.

Like the '196 Patent, Sandgren simulates sound-reflecting objects and models sound wave absorption and reflection by objects within a room. *Id.*, Abstract, ¶¶2, 12, 16, 31-33, 69-70, Figs. 1, 5A-C; Kyriakakis ¶85.

E. Woo

Woo discloses detecting and photographing actual objects in the user's room and adjusting a sound output based on the object's position in the room. EX1016, ¶¶3, 53-54, Figs. 3, 5; Kyriakakis ¶86. Woo was not cited or discussed during prosecution.

Like the '196 Patent, Woo adjusts sound properties based on the presence of such objects. EX1016, ¶3, Abstract. Kyriakakis ¶87.

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 ¶¶30-36.

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 ¶33.

VII. CLAIM CONSTRUCTION

The prior art relied on herein 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., Elements 15[c] and 19[a]),⁶ should be construed by its plain meaning and “[s]ection 112(f) does not apply.” EX1027, 24-28. Solely for the purposes of this IPR, Petitioner adopts PO’s plain-meaning constructions for these terms for all challenged claims.⁷ *See Western Digital Corp. v. SPEX Techs., Inc.*, IPR2018-0084, Paper 14, at 10-12 (April 25, 2018) (petitioners 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: LI RENDERS OBVIOUS CLAIMS 1, 8, 15, 17

A. Motivation to Modify

A POSA would have found it obvious to modify Li’s conference call embodiment to provide *a voice of a computer program* as taught in other parts of Li. Kyriakakis ¶¶88-94.

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 the advantages of a computer program synthesizing a voice to provide a *voice of a computer program* as Li's sound source or binaural sound. Kyriakakis ¶89. For example, computerized text-to-speech synthesis would have allowed a speaking-impaired person to participate in Li's conference call system. Kyriakakis ¶¶89-93 (describing EX1042 explaining the development and application of text-to-speech, EX1040 describing a 1990s era teleconferencing system incorporating text-to-speech, and EX1041 which expanded text-to-speech from teleconferencing to videoconferencing); *see, generally*, EX1014. Computerized voice synthesis would have enabled Li's conference call system to control for voice quality, speed, and volume to improve clarity and minimize potential miscommunication, while obviating a human speaker recording every possible utterance, thereby saving time, cost, and storage space. Kyriakakis ¶93. Computer-synthesized voice also adds flexibility by removing the physical constraints of human speech and adapting to different languages, accents, emotions, and vocal ranges. *Id.*

The modification would have been a natural extension of Li's "3D sound...synthesis and reproduction" via sound generating devices or sound emitting devices. EX1006, 3:11-16, 3:39-44, 6:67-7:4; Kyriakakis, ¶94. Li already contemplates using its system in "a simple videogame" and "sophisticated virtual reality simulators" (EX1006, 1:32-35, 7:17-29, 10:3-11, 35:4-13), both of which the

POSA would have understood to use *a voice of a computer program* to enhance the user's experience. Kyriakakis ¶94.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying the conference call embodiment of Li to include *a voice of a computer program*. Kyriakakis ¶¶95-96. Li's portable device and wearable headphones provide binaural sound appearing to originate from an SLP. The "3D sound synthesis and reproduction to allow users to experience 3D sound...through a headset" in a conference call setting. EX1006, 7:1-4. Li uses its 3D sound system in conference calls, videogames, and virtual reality simulators (*id.*, 7:17-29, 10:3-11, 21:57-22:5, 35:4-13), and the generated or emitted sound includes "a person's voice [and] pre-recorded multiple sound tracks" (*id.*, 4:58-64, 14:40-46). Li already discloses sound generating devices and different sound sources, including musical instruments, a person's voice and pre-recorded sounds. EX1006, 4:58-64, 14:40-46; Kyriakakis, ¶96.

A POSA would have recognized that although *a voice of a computer program* is not expressly part of Li's conference call embodiment, Li teaches using such voice in relation to other embodiments, and using such voice would have saved time and cost, added flexibility, and improved clarity. Kyriakakis ¶96. A POSA implementing Li's conference call embodiment would have understood to use a

computer-generated voice for Li's spatialized sound in the conference call. *Id.* This would have merely involved a combination of known prior art elements (spatialization and computer-generated voice) according to known methods and techniques (applying HRTF filters) to yield predictable results (spatialized sound), and involves use of known techniques (spatialization) to improve a similar device (wearable communication device) in the same way (filtering sounds to introduce spatial cues). *Id.*

C. Independent Claims 1 and 8:

1. 1[pre]/8[pre]: Move SLP of Voice of Computer Program

To the extent limiting, Li teaches the preambles. Kyriakakis ¶¶97-104.

Li discloses *one or more electronic devices in an electronic system* by describing a “configurable 3D sound system 900” that *execute[s] a method*. 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, 33; Kyriakakis ¶98.

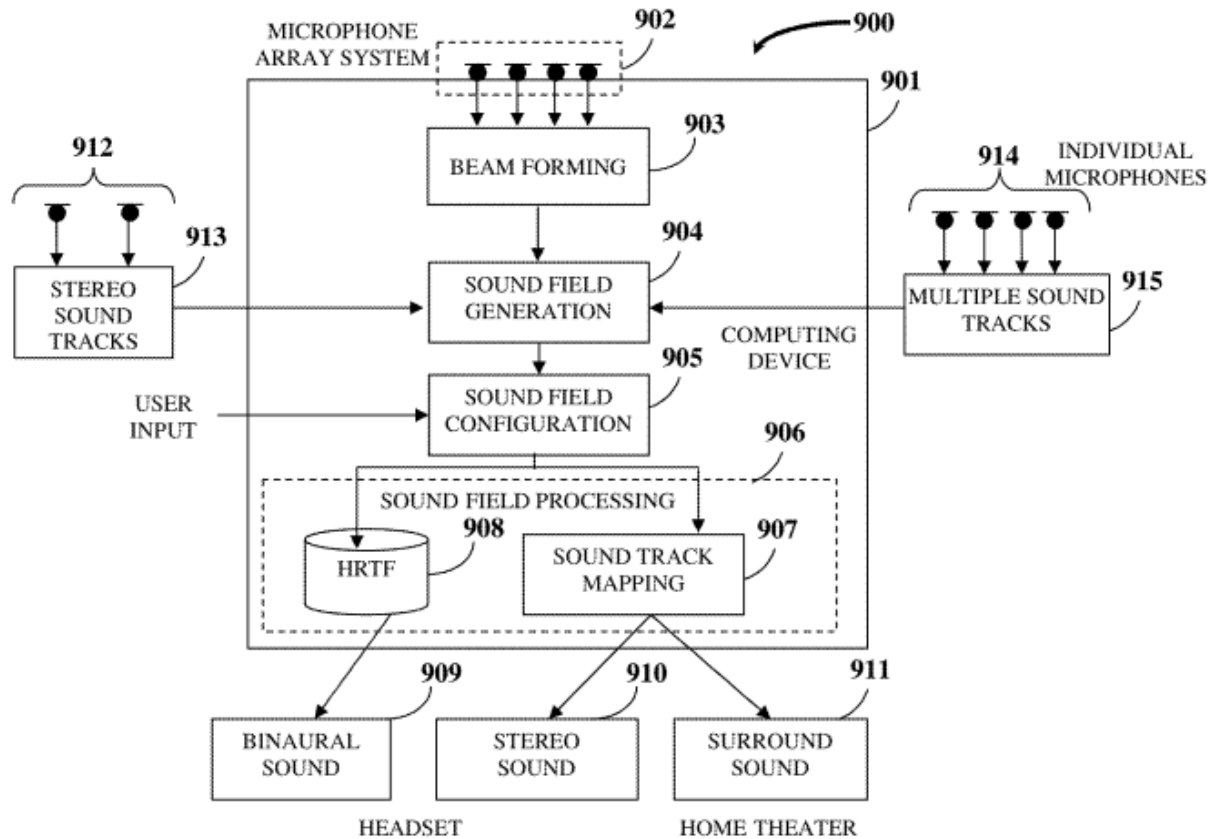


FIG. 9

In Li's Fig. 9 (reproduced above), the configurable 3D sound system 900 (*electronic system*) includes a computing device 901 (which the POSA would have known to include a processor and memory) and headphones or a headset (*one or more electronic devices*) that provide a "3D sound processing application" that "synthesizes and reproduces the user preferred sound field" and "performs sound track mapping 907 by convolving each of the sound tracks with corresponding HRTFs... to synthesis [sic] 3D binaural sound 909 for a headset user." EX1006, 17:1-4, 17:22-37; Kyriakakis, ¶98.

Specific to 8[pre], Li discloses *non-transitory computer readable storage medium storing instructions*. Kyriakakis ¶99. Li's configurable 3D sound system 900 includes computing device 901, which is an *electronic device*, and "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 (emphases added); *see also* 30:65-31:1, 35:24-50. Li's *electronic device* therefore includes a memory unit 3402 (*non-transitory computer readable storage medium*) and a processor 3401 that executes the *instructions*. Kyriakakis ¶99.

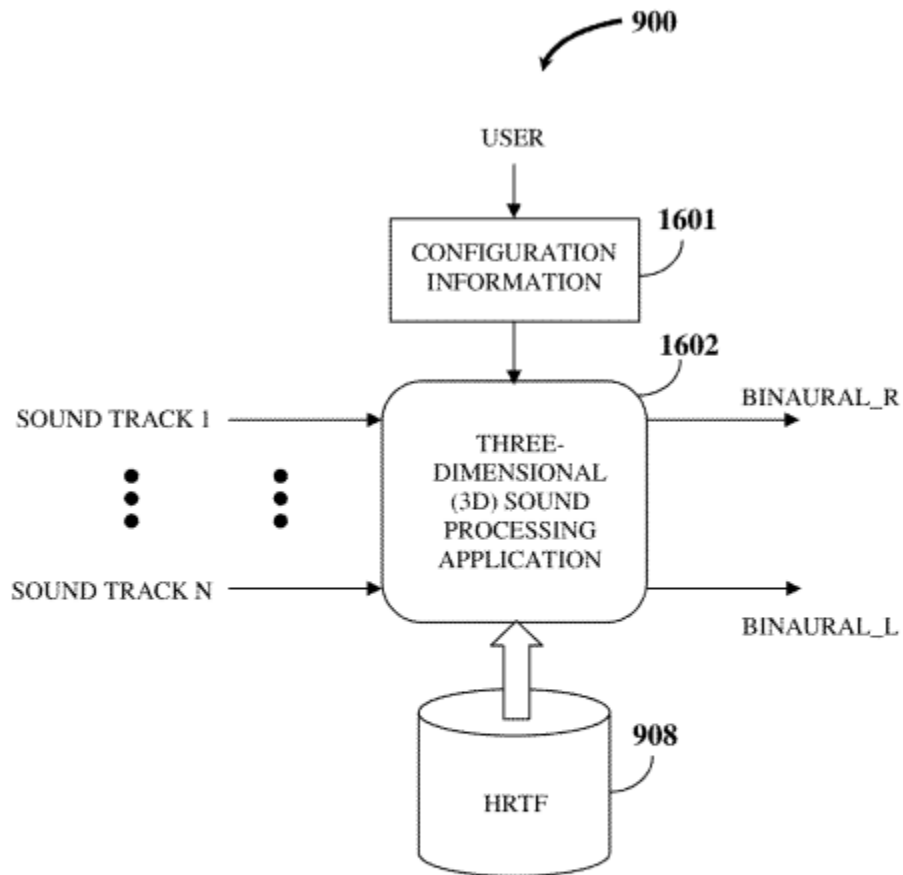


FIG. 16

Regarding *mov[ing] an SLP*, Li’s 3D sound processing application 1602 in Fig. 16 “generates a configurable sound field that provides an interface to give the user the freedom of configuring the positions and *movements* of multiple sound tracks, in order to render a customized 3D binaural sound.” *Id.*, 20:26-36 (emphasis added); Kyriakakis, ¶100.

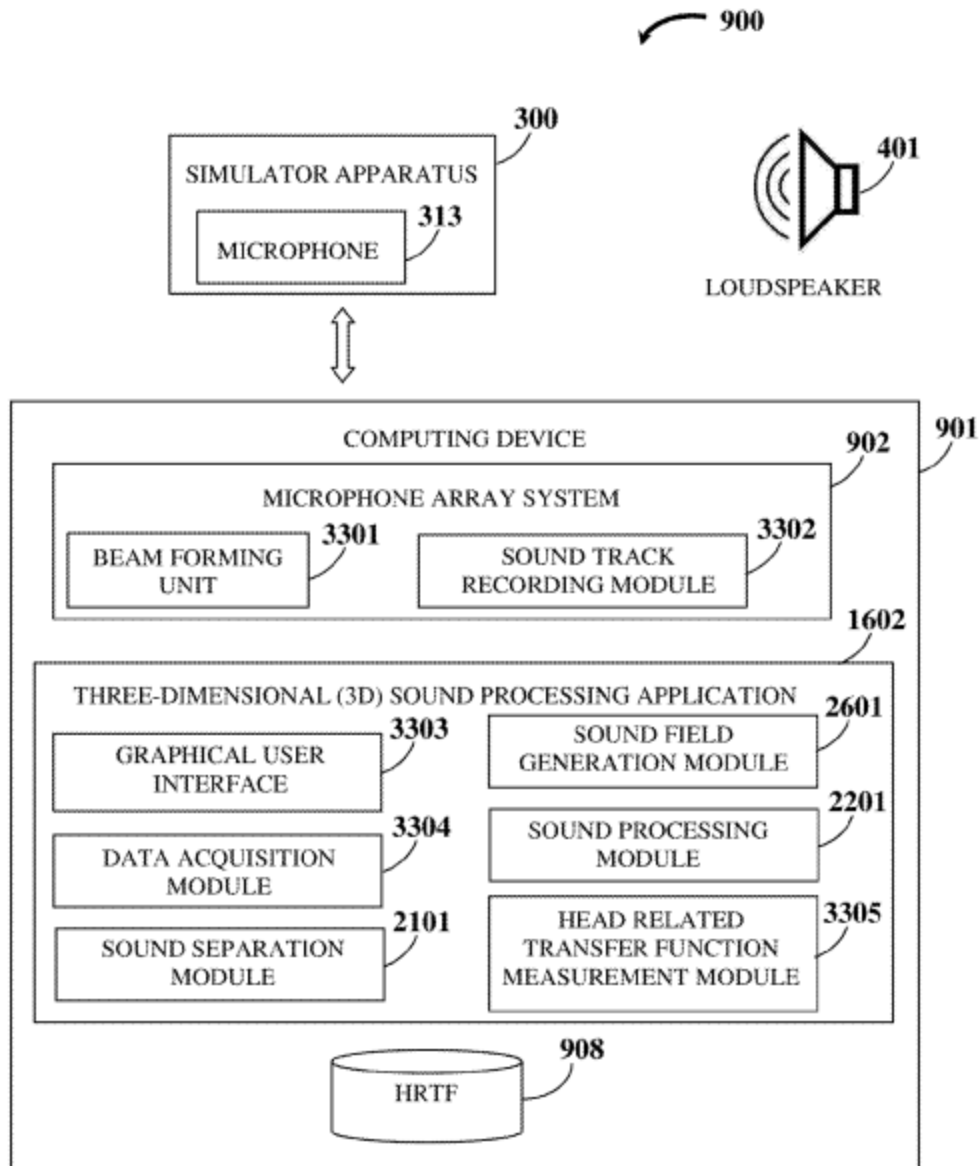


FIG. 33

As Li further describes in relation to Figs. 9 and 33, “the user configures the placements and *movements* of any available sound sources as the inputs in order to obtain a virtual reality scene” (*id.*, 22:53-67 (emphasis added)) and therefore

move[s] an SLP. See also id., 21:57-66 (describing system 900 as “allowing the user to move the sound source around the user 1703, move the sound source up and down, etc., on the configurable sound field 1700”). Kyriakakis, ¶101.

Although Li does not expressly disclose *a voice of a computer program* as part of its conference call embodiment, Li teaches using such voice in relation to other embodiments. Kyriakakis ¶102. Li’s sound source can be “sound generating devices or sound emitting devices,” “a person’s *voice*,” or “pre-recorded multiple sound tracks.” EX1006, 4:58-67, 14:40-48. Li also discloses a *computer program* to “generate configurable 3D sounds,” which involves processing the speaker’s voice to position “different speaking users at different spots.” *Id.*, 33:14-18, 33:32-51, 33:62-66, 34:4-8, 34:20-51, 35:58-36:2. As the POSA would have appreciated, implementing Li’s sound source as *a voice of a computer program* would have been advantageous because it would have saved time, cost, and storage space, added flexibility to vocal range, and improved clarity. Section VIII.A-B; Kyriakakis ¶103.⁸

In Li, moving the sound source (*mov[ing] a SLP*) takes place *during a voice exchange between the computer program and a user/person*. Kyriakakis ¶104.

⁸ The mapping for the phrase *voice of a computer program* also applies to 1[a]/8[a]-1[d]/8[d], 15[pre]-15[c3], which recite the same phrase or a similar phrase, and the mapping is not repeated in those sections, *infra*.

When Li's user "plac[es] different speaking users at different spots to mimic a real conference room environment" (EX1006, 21:66-22:5), this takes place *during* a "conference call," which involves *a voice exchange between the* processed voice of the speaker (*computer program*) *and a user/person*. Kyriakakis ¶104.⁹

2. 1[a]/8[a]: Provide Through Earphone/WED Binaural Sound Such That Origin Occurs at SLP

Li renders obvious this limitation. Kyriakakis ¶¶105-116.

Li's *electronic system* includes headphones or a headset, which includes *an electronic earphone on a person being located at a head of the user and being in communication with a smart phone (HPED)* (1[a]), or *a WED on the person* (8[a]). Specifically, Li depicts the configurable 3D sound system 900 (*electronic system*) in Fig. 9 as providing 3D binaural sound 909 (*binaural sound*) over a "headset." EX1006, 17:33-37; *see also* 20:44-48, 2:55-57, 3:6-10. Kyriakakis, ¶106.

Specific to 1[a], Li's headset *communicat[es] with an HPED*, as was common for *electronic earphones* such as earbuds or headphones. Kyriakakis ¶107. Li's "configurable 3D sound system 900 [*electronic system*]" can be used in portable

⁹ The mapping of *during a voice exchange between the computer program and a user/peron* also applies to 1[a]/8[a], 1[c], 1[d]/8[d], 15[pre], 15[c3], which have the same or similar phrase, and the mapping is not repeated, *infra*.

devices, for example, smart phones [*HPED*], tablet computing devices, etc.” for generating and playing a synthesized *binaural sound* 909. EX1006, 34:52-56, 34:60-35:13; Kyriakakis, ¶107. Thus, Li discloses the configurable 3D sound system 900 is included in a smart phone (*HPED*), which the POSA would have readily understood to connect to the headset (*electronic earphone*). Kyriakakis *Id.* Li therefore discloses that its headset (*electronic earphone*) is *in communication with* a smart phone (*HPED*). *Id.*; see also EX1006, 3:6-8 (“consumers often use headphones for audio rendering on a mobile device.”), 6:67-7:4, 7:33-37, 10:5-11, 14:14-15, 16:7, 17:35-37, 20:9-12, 20:44-48. Specific to 8[a], Li’s headset renders obvious *WED* under PO’s position. Kyriakakis ¶108. Not only were wireless headsets well-known and commonplace in 2014, but Li also discloses wireless connections. See EX1006, 30:38-43, 3:2-8, 36:33-37; Kyriakakis ¶108.

Li also discloses the remaining elements of 1[a]/8[a]. In Li, the *SLP* is localized *at a first location* where no physical object exists that is at a distance from the user (*in empty space that is away from and proximate to the user/person*). Kyriakakis ¶109. 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 ¶110. For example, Li describes “placing different speaking users at different spots to mimic a real conference environment.” EX1006, 22:1-5. The POSA would have found it obvious to seat the *user/person* at a physical conference table, where an

empty seat at the table (*empty space*) is allocated for a virtual participant, as was known in the field. Kyriakakis ¶¶110-113 (describing EX1038 which provides empty chairs for remote teleconference participants; and describing EX1038 which discusses a hybrid learning classroom where a virtual student selects a physically unoccupied seat (an empty chair) to project the student's image and from which the student's voice will originate during the call). Such an implementation 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 stated goal to "match a vivid 3D vision field." EX1006, 3:11-16; Kyriakakis ¶110.

Furthermore, in relation to Fig. 1, Li discloses measuring HRTFs "in a quiet sound treated room free of impulsive background noise." EX1006, 10:39-54. These measured HRTFs "as disclosed in the detailed description of FIGS. 1-2" are "store[d]...in an HRTF database 908 exemplarily illustrated in FIG. 9." *Id.*, 19:62-67. The 3D sound processing application 1602 uses those same measured HRTFs from the database 908 to generate binaural sound 909. *Id.*, 17:27-37, 19:39-67. A POSA would understand that because the HRTFs are measured in a "quiet sound treated room" to generate the binaural sound, the *SLP* would be placed in a similarly object-free space. Kyriakakis ¶114. Moreover, because the sound sources are placed "at different spots to mimic a real conference room environment" (EX1006,

21:66-22:5), the *SLP at a first location is away from and proximate to the user/person* such that the sound source would be placed at an empty chair at a distance away from the user. Kyriakakis ¶115.¹⁰

3. 1[b]/8[b]: Receive Command to Move SLP

Li renders obvious this limitation. Kyriakakis ¶¶117-120.

Li's configurable 3D sound system 900 (*electronic system*) "acquires [receives] configuration information 1601 [command] from a user" including "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. Those commands "allow[] the user to move the sound source around the user 1703, move the sound source up and down, etc., on the configurable sound field 1700." *Id.*, 21:57-22:5; *see also* 15:29-32, 15:44-47. By allowing the user to move the sound source, Li discloses *receiving, from the user/person, configuration information (a command) to move the sound source for a speaker (SLP of the voice) from the first location to a second location that is next to (proximate) the first*

¹⁰ The mapping of *SLP at a first/second location in empty space that is away from and proximate to the user/person* also applies to 1[b]/8[b]-1[d]/8[d], 15[c1]-15[c3], which have the same or similar phrase, and the mapping is not repeated, *infra*.

location and at a distance from (away from and proximate to, see EX1027, 8) the user. Kyriakakis ¶118.

Specific to 1[b], Li's *command* is received *at the HPED* because the "user may configure the positions and movements of the sound sources on the generated configurable sound field 1700... by entering his/her preferences... via the GUI" "on the user's computing device 901," which is in a mobile phone (*HPED*). EX1006, 20:62-21:9, 21:47-60; *see also* 7:15-17 ("Users can control the configurations, for example, using a touch screen on their handheld devices."), 15:27-52, 15:61-16:8, 17:38-40, 18:1-21; Kyriakakis ¶119.

As explained for 1[pre]/8[pre], Li renders obvious the *voice of the computer program*. *See* Section VIII.C.1; Kyriakakis ¶120.

4. 1[c]/8[c]: Move SLP

Li renders obvious this limitation. Kyriakakis ¶121-122.

Li's configurable 3D sound system (*electronic system*) *moves, based on the configuration information (the command) the SLP around the user, or up and down, to different locations (from the first location to the second location) that are at a distance from the user to mimic a real conference room environment (proximate to the first location and away from and proximate to the user/person)*. EX1006, 21:57-22:5, 15:29-48; *see also* Section VIII.C.1-3 (Ground 1, 1[pre]/8[pre]-1[b]/8[b]); Kyriakakis ¶122.

5. 1[d]/8[d]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location

Li renders obvious this limitation. Kyriakakis ¶¶123-124.

Li's configurable 3D sound system (*electronic system*) provid[es], through the headset (*electronic earphone being located at a head of the user being in communication with the HPED (1[d]) or WED on the person (8[d])*), the 3D binaural sound to the user. See Section VIII.C.2 (Ground 1, 1[a]/8[a]). Moreover, in Li, after the user moves the sound source as described in 1[b]/8[b]-1[c]/8[c], the spots for the speakers have changed to new locations (*such that the origin of the voice of the computer program occurs at the SLP at the second location*). See Section VIII.C.1-4 (Ground 1, 1[pre]/8[pre]-1[c]/8[c]); Kyriakakis ¶124.

D. Independent Claim 15:

1. 15[pre]: Move SLP of Voice of Computer Program

To the extent limiting, Li renders obvious the preamble. See Section VIII.C.1 (Ground 1, 1[pre]/8[pre]); Kyriakakis ¶125.

2. 15[a]: WED Provides Voice of Computer Program at SLP

Li renders obvious this limitation, as Li's headphones or headset is *a WED that the user (person) wears*. See Section VIII.C.1-2 (Ground 1, 1[pre]/8[pre]-1[a]/8[a]); Kyriakakis ¶126.

3. 15[b]: Memory

Li renders obvious this limitation, as Li's configurable 3D sound system 900 employs and includes memory unit 3402 (*a memory*) *that stores programs (instructions)*. See Section VIII.C.1-2 (Ground 1, 1[pre]/8[pre]-1[a]/8[a]); Kyriakakis ¶127.

4. 15[c]: Processor

Li renders obvious this limitation, as Li's configurable 3D sound system 900 employs and includes processor 3401 (*processor*) *that executes the programs (instructions)*. See Section VIII.C.1-2 (Ground 1, 1[pre]/8[pre]-1[a]/8[a]); Kyriakakis ¶128.

5. 15[c1]: Receive Command to Move SLP

Li renders obvious this limitation. See Section VIII.C.3 (Ground 1, 1[b]/8[b]); Kyriakakis ¶129.

6. 15[c2]: Move SLP

Li renders obvious this limitation. See Section VIII.C.4 (Ground 1, 1[c]/8[c]); Kyriakakis ¶130.¹¹

¹¹ Assuming 15[c2] were construed as means-plus-function, it would have been obvious because PO admitted that the algorithm for moving an SLP uses “conventional techniques” (EX1027, 45-46).

7. 15[c3]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location

Li renders obvious this limitation. *See* Section VIII.C.5 (Ground 1, 1[d]/8[d]); Kyriakakis ¶131.¹²

E. Dependent Claim 17: Change ITD and ILD to Move SLP

Li renders obvious this limitation. Kyriakakis ¶¶132-137. As explained for 15[c2] and 1[pre]/8[pre]-1[c]/8[c], the POSA would have found it obvious for Li's processor 3401 to *execute[] the instructions to change HRTFs of the binaural sound in order to move the SLP of the voice of the computer program from the first location in empty space to the second location in empty space.* *See* Sections VIII.C.1-4, VIII.D.6. As a POSA would have recognized, Li discloses sound localization cues

¹² Assuming 15[c3] were construed as means-plus-function, it would have been obvious because Li describes “convolution” processes using HRTFs (*e.g.*, EX1006, 10:64–11:4, 16:65–17:4), and because PO's litigation expert opined that “[c]onvolution is the main algorithm required to perform the claimed processes,” while repeatedly conceding that convolution was “well-known.” EX1032 59:4-16 (admitting there were “many references”—“in the 1990s”—“discussing using convolution processes to process head related transfer functions with digital signal processors for binaural audio”).

including *an ITD and an ILD* that form the “combination of [] factors [that are] modeled by” HRTFs, such that both cues are included in the HRTFs that are used to *move the SLP*. EX1006, 1:65-2:21, Kyriakakis ¶132.

IX. GROUND 2: LI IN VIEW OF BALLAS RENDERS OBVIOUS CLAIMS 4, 6, 9-10, 14, 18

A. Motivation to Combine

It would have been obvious to modify Li to track the user’s head orientation and adjust the sound so that the sound appears fixed even as the user moves, as taught in Ballas. Kyriakakis ¶¶138-140. Li recognizes that to determine a sound source, humans rely on various cues including head motion. EX1006, 1:38-42. Both references have a shared goal of providing a realistic experience for the user and are structurally and functionally similar. Kyriakakis ¶¶138, 142. The POSA would have used Ballas’s conventional head tracking and spatialization techniques to track the user’s head and modify Li’s sound field accounting for the head orientation and movements. *Id.*

Li’s goal is to closely resemble “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; Kyriakakis ¶139. Ballas shares this goal of “a realistic experience” by providing “spatialized sound” for “teleconferencing and telepresence” and “virtual and augmented reality environments.” EX1008, ¶¶2, 9,

10. Li's method is for potentially lengthy stretches of time, such as for conferences, concerts, or radio dramas. EX1006, 21:57-22:5. For such applications, head tracking is key to make the experience realistic, as the user can freely move their head, and accounting for these movements in the sound field means the sound source origins appear fixed and therefore realistic. Kyriakakis ¶139. Modifying Li with Ballas would have also allowed the user to move while staying connected to the event, and maintained the sound field even as the user moves. *Id.*

The modification would have been a natural extension of Li's use of HRTFs to spatialize sound. Kyriakakis ¶140. Like Li, Ballas discloses an "[HRTF] filter to produce spatialized audio." EX1008, Abstract; *see also id.*, ¶¶16-17, 25. Ballas's head tracker provides head location and orientation measurements to a computer, which applies an "[HRTF] filter for maintaining virtual location of the audio signals." *Id.*, Abstract, ¶18; *see also id.*, ¶¶17, 19-20, 25-26, claims 1, 7. Using head tracking to adjust HRTFs would have allowed the user to move while maintaining fixed virtual sound locations, and would have advanced the common goals of both references. Kyriakakis ¶140.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Li with Ballas. Kyriakakis ¶¶141-143. Both use HRTFs to generate spatialized

binaural sound before providing it to a user over wearable headphones. *Id.*; see EX1006, Fig. 9.

Consistent with Li, Ballas's speakers 110 are attached to a "frame [that] is 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 is processed using HRTF filters 106 to produce spatialized audio. *Id.* Accordingly, a POSA would have recognized that Li and Ballas are structurally and functionally similar. Kyriakakis ¶142.

A POSA implementing Li would have recognized that Li's headset would attach to Ballas's frame, which also has head tracking system 104. Kyriakakis ¶143. Ballas's head tracking system 104 tracks the head's location and orientation and provides that information to a computer system for HRTF processing, producing spatialized audio that is fed to the speakers. EX1008, ¶25. 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 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 ¶143.

C. Dependent Claim 4

1. 4[a]: Track Head

Li in view of Ballas (“Li-Ballas”) teaches this limitation. Kyriakakis ¶¶144-146.

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’s “virtual reality head mounted visual display system” is an eyeglass frame or helmet. *Id.*, ¶25; *see also* Figs. 3, 4.

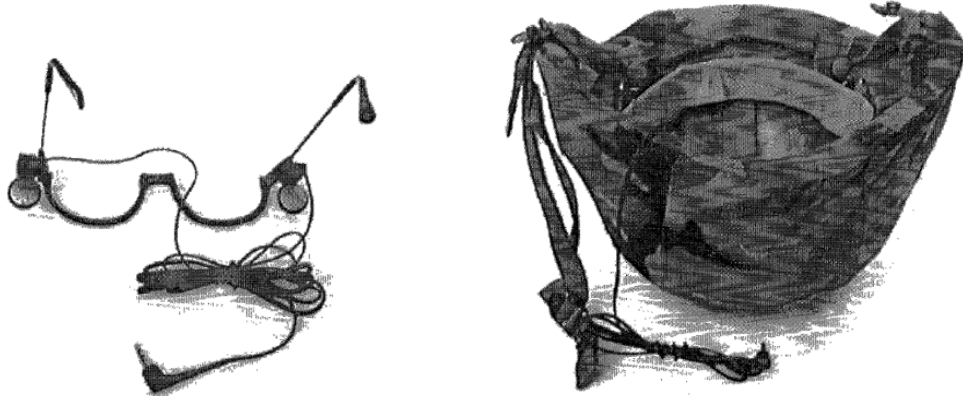


Figure 3. Eyeglass frame mount prototype Figure 4. Helmet mount prototype

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” (*tracking a head orientation of the user*). EX1008, ¶25; Kyriakakis ¶144.

Ballas discloses “a pair of [HRTFs]” are “for a particular location” (*the SLP at the second location*). *Id.*, ¶16. Specifically, 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 *as the user moves with respect to the SLP at the second location*. Kyriakakis ¶145.

A POSA would have been motivated to implement Li’s headphones as part of Ballas’s wearable frame having a head tracking system. Kyriakakis ¶146; *see also* Section IX.A (Motivation to Combine Li with Ballas).

2. 4[b]: Adjust Voice To Continue To Appear To Originate From SLP

Li-Ballas teaches this limitation. Kyriakakis ¶¶147-148.

As explained for 4[a], Ballas applies HRTF filters to (*adjusting*) the *sound* to maintain the sound’s virtual location (*so the sound... continues to appear to originate from the SLP at the second location*), which allows the user to move without degradation of the virtual location (*even as the user moves with respect to the second location*). *See* Section IX.C.1; Kyriakakis ¶144-148. In order for Ballas’s HRTF filters to “maintain[] [the] virtual location of the audio signals” even while the user moves, Ballas continually adjusts HRTFs because the *SLP* relative to the user’s position changes. Kyriakakis ¶148.

D. Dependent Claim 6: Change Amplitude, ITD, and ILD to Continue to Originate from SLP as User Moves

Li-Ballas teaches this limitation. Kyriakakis ¶¶149-152. As explained for 4[a]-4[b], Li-Ballas renders obvious adjusting (*changing*)... *the voice of the computer program... so a sound of the voice of the computer program continues to appear to originate from the SLP at the second location as the user moves. See* Section IX.C; Kyriakakis ¶¶144-148.

Li further explains that “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 (emphases added). Li defines the ITD as “the difference of arrival times of transmitted sound between the two ears” and the ILD as “the difference in level and/or intensity of the transmitted sound received between the two ears.” *Id.*, 1:65-2:4; *see also id.*, 12:54--57. Li also discloses *an amplitude* by describing “volume” as a “configurable parameter associated with the sound sources.” *Id.*, 5:18-22; Kyriakakis ¶¶150.

Li *chang[es]* the *ILD* and *ITD* for the voice as part of its bank of HRTF filters. Kyriakakis ¶¶151. In Li, *ILD* and *ITD* are modeled by two sets of HRTFs, which “are a bank by positions.” EX1006, 2:6-21; *see also id.*, 12:57-65. As the POSA would have understood, Li applies one HRTF (from the bank) to produce spatialized 3D sound for a specific position (*first location*), and applies a different HRTF to position

the 3D sound at a new position (*SLP* at the *second location*). *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 ¶151. Because each HRTF models a different combination of ILD and ITD cues, Li's selection of different HRTFs for different 3D sound positioning discloses *changing ILD and ITD for the voice*. Kyriakakis ¶151.

Li also *chang[es]* the *amplitude for the voice* as part of its bank of HRTF filters. Kyriakakis ¶152. In Li, “distance” and “volume” (*amplitude*) are each a “configurable parameter[] associated with the sound sources” and are therefore adjustable as *the user moves with respect to the SLP at the second location*, as Ballas contemplates with its head tracking system. EX1006, 5:18-22; *see also id.*, 15:36-40, 28:29-35, claims 8, 16, 19, 25, 35. Moreover, in a 3D sound field, a sound signal having a low volume is perceived to be further away than a louder sound. Kyriakakis ¶152. By having a bank of HRTFs for different “positions” (not just directions), Li applies different HRTFs for different distances, which involves *changing the amplitude for the voice* as the positions shift to change distance. *Id.*

E. Dependent Claim 9: Maintain SLP But Alter Amplitude

Li-Ballas teaches this limitation, as the volume (*amplitude*) is adjusted based on at least distance from the SLP. *See* Section IX.D (Ground 2, Claim 6); Kyriakakis ¶¶144-153.

F. Dependent Claim 10

1. 10[a]: Monitor Movements

Li-Ballas teaches this limitation. *See* Section IX.C.1 (Ground 2, 4[a]); Kyriakakis ¶¶144-146, 154.

2. 10[b]: Adjust ITD and Amplitude

Li-Ballas teaches this limitation. *See* Section VIII.E (Ground 1, 17) and Sections IX.C.1 and IX.D (Ground 2, 4[b] and 6); Kyriakakis ¶¶132-137, 144-152, 155.

G. Dependent Claim 14: Adjust ILD, ITD, and Amplitude So SLP Remains Fixed

Li-Ballas teaches this limitation. *See* Sections IX.C-D (Ground 2, cls. 4, 6); Kyriakakis ¶¶144-152, 156-157.

H. Dependent Claim 18: Change ITD and Amplitude to Maintain SLP

Li-Ballas teaches this limitation. Kyriakakis ¶¶158-160.

As explained in cl. 6, Li-Ballas discloses *chang[ing] an ITD and an amplitude of the binaural sound in order to maintain the SLP fixed at the second location in empty space... while the person... changes a head orientation with respect to the second location in empty space*. Kyriakakis ¶¶144-152. As explained in 15[c], Li's processor 3401 *executes instructions* to perform this processing, and as explained in 1[pre]/8[pre], this occurs *during the voice exchange*. Kyriakakis ¶128. Ballas explains "[t]he HRTF filter maintains virtual

location of the audio signals, thus allowing the listener to change locations and head orientation without degradation of the audio signal.” EX1008, ¶17; *see also id.*, Abstract, ¶18, claim 1. Ballas’s head tracking system 104 “measures the location and orientation of a user’s head.” *Id.*, ¶25; *see also id.*, Abstract, ¶¶17-20, 26, claims 1, 7, 13. As the POSA would have understood, Ballas’s head tracking system would have tracked *the person physically mov[ing] away from the second location in empty space*. Kyriakakis ¶160. To provide a realistic experience, Li-Ballas would have processed the sound *in order to maintain the SLP fixed at the second location* regardless of the user’s movements or head orientation. *Id.*

X. GROUND 3: LI IN VIEW OF BALLAS AND SANDGREN RENDERS OBVIOUS CLAIM 5

A. Motivation

A POSA would be motivated to combine Li and Ballas. *See* Section IX.A (Ground 2, Motivation to Combine); Kyriakakis ¶¶138-140. Li’s “volume” is a “configurable parameter[] associated with the sound sources” (EX1006, 5:18-22), and in Li’s 3D sound field, a sound signal having a low volume is perceived to be further away than a louder sound. Kyriakakis ¶161. A POSA would have looked to Sandgren to determine how to vary Li’s sound volume with distance, to mimic “what humans hear in the real world.” EX1006, 2:34-36, 22:1-5; *see also id.* 2:22-24, 13:57-60; Kyriakakis ¶162. Modifying Li with Sandgren to adjust a sound

amplitude according to an inverse-square law with distance (EX1017, ¶¶1, 77, Abstract) would have provided a more realistic and immersive experience for the user, and would have been a natural extension of Li's disclosure of a configurable volume. Kyriakakis ¶163.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Li with Ballas, and further modifying Li-Ballas with Sandgren. *See* Section IX.B (Ground 2, Reasonable Expectation of Success); Kyriakakis ¶¶141-143, 164-165. Li and Sandgren are structurally and functionally similar. Both use HRTFs to generate spatialized binaural sound and provide such sound over headphones. *Id.* Like Li, Sandgren creates virtual 3D audio scenes over headphones to externalize virtual sound sources, and the user perceives the sounds to originate from different SLPs. EX1017, ¶¶1, 46; Kyriakakis ¶164.

A POSA implementing Li would have recognized that Sandgren's teachings of the inverse square law and virtual object simulation would have been straightforward to implement in Li's spatialized sound processing. Kyriakakis ¶165. This would have merely involved a combination of known prior art elements (inverse square law and virtual objects) according to known methods and techniques (simulate distance and objects) to yield predictable results (realistic sound fields), and involves use of a known technique (sound adjustments) to improve a similar

device (wearable communication device) in the same way (adjusting sounds to produce a realistic sound field). *Id.*

C. Dependent Claim 5: Adjust Amplitude to be Inversely Proportional to Square of Distance

Li-Ballas and Sandgren teaches this limitation. Kyriakakis ¶¶166-168.

As explained for claim 6, Li-Ballas discloses *adjust[ing] an amplitude of the voice of the computer program... as the user moves with respect to the SLP at the second location in empty space. Id.* ¶¶144-152; See Section IX.D (Ground 2, 6).

Neither Li nor Ballas expressly states *adjust[ing] the amplitude... to be inversely proportional to a square of a distance between the user and the SLP*, but this would have been obvious in view of Sandgren. Kyriakakis ¶168. Sandgren discloses “[s]ound intensity generally follows an inverse-square law, meaning that for each doubling of distance, the intensity drops by 6 dB.” EX1017, ¶77. To provide a realistic sound field, it would have been obvious to *adjust an amplitude... to be inversely proportional to a square of a distance between the user and the SLP*. Kyriakakis ¶168.

XI. GROUND 4: LI IN VIEW OF SANDGREN RENDERS OBVIOUS CLAIM 19

A. Motivation to Combine

The POSA would have been motivated to modify Li to simulate virtual objects, as Sandgren taught. Kyriakakis ¶¶169-172. Li recognizes that to determine

a source of a sound, humans rely on various types of cues including reverberation. EX1006, 1:38-42. Sandgren teaches that reverberation and attenuation are important to realistically simulate 3D audio. *Id.*, ¶¶6, 12. Both references have a shared goal of providing a realistic and immersive experience for the user, and doing so would have had several benefits. Kyriakakis ¶¶169-170.

Li's primary goal is to closely simulate "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, 1:24-25 ("[s]ounds are a constant presence in everyday life and offer rich cues about the environment"), 1:46-47 (describing the need to accurately synthesize spatial sound to enhance immersiveness). Sandgren shares this goal. EX1017, ¶1 (describing virtual 3D audio scenes and increasing externalization); *see also id.*, Abstract. As Sandgren appreciates, "[a] room's reverberation depends on... the types of objects present in the room," and in order to "creat[e] a realistic 3D audio scene... it is not enough to concentrate on the direct sound." *Id.*, ¶¶5-6. Sandgren emphasizes that it is the reverberation from "objects in the room" that are "essential for good externalization," "not the main reflections from the floor, ceiling, and walls of a room." *Id.*, ¶19. Moreover, a realistic sound simulator includes attenuation filters combined with reverberation generators. *Id.*, ¶¶62-63, Fig. 5C. Thus, to achieve good externalization, Sandgren teaches that

reverberation and attenuation are essential to simulate, and the POSA would have been motivated to use both to simulate virtual objects. Kyriakakis ¶171.

The modification would have been a natural extension of Li's use of HRTFs to spatialize sound. Kyriakakis ¶172. Like Li, Sandgren uses HRTFs to "simulat[e] an audio scene or creat[e] a virtual audio scene" over headphones. EX1017, ¶9-10; *see also* ¶¶11, 15-16, 19, 22, 46, 48-64, 70-100, Figs. 5A-C. A POSA would have recognized that Sandgren's teachings would have enhanced Li's use of HRTFs for an immersive experience by introducing attenuation filters and reverberation generators indicative of virtual objects. Kyriakakis ¶172. This would have advanced the common goals of both references. *Id.*

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Li in accordance with the teachings of Sandgren. Kyriakakis ¶¶141-143, 164-165, 173; *See* Section X.B (Ground 3, Reasonable Expectation of Success).

C. Dependent Claim 19:

1. 19[a]: Change Attenuation and Reverberation

Li-Sandgren teaches this limitation. Kyriakakis ¶¶174-178.

Li discloses *reverberation* as a "source[] of localization cues" (EX1006, 1:39-42), but does not explicitly disclose that its *processor executes the instructions to*

perform this limitation, but this would have been obvious in view of Sandgren. Kyriakakis ¶175.

Sandgren simulates sound-reflecting objects (*virtual objects*) to enhance the “[e]xternalization of the simulated sound source.” EX1017, Abstract. “When 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 ¶176.

Sandgren discloses *chang[ing] reverberation*. Kyriakakis ¶177. “Reflected sound energy reaching the listener is generally called *reverberation*.” EX1017, ¶4, 47 (emphases added); *see also* ¶¶23-25, 48-66, 68-74, 100. As 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.”¹³ *Id.*, ¶69, Figs. 5A-C; *see also id.*, ¶74. By updating parameters for a reverberation generator to simulate moving objects,

¹³ 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 discloses *change... reverberation of the binaural sound in order to add artificial audial cues of virtual objects near the person*. Kyriakakis ¶177.

Sandgren also discloses *chang[ing] attenuation*. Kyriakakis ¶178. Sandgren's Figs. 5A-C depict audio simulators "having HRTF processors and an early-reflection generator," and each figure 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 ¶178. As the 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) will *change* as Sandgren's *virtual objects* move. EX1017, ¶¶74-75, 78-80, 89, Fig. 1; Kyriakakis ¶178. Sandgren further 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). Therefore, Sandgren discloses *change... attenuation of the binaural sound in order to add artificial audial cues of virtual objects near the person*. *Id.*

2. 19[b]: Virtual Objects

Li-Sandgren teaches this limitation. Kyriakakis ¶¶179-180.

As explained in 19[a], Sandgren's sound-reflecting objects (*virtual objects*) 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 sound 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 ¶¶174-180.¹⁴

XII. GROUND 5: LI AND WOO RENDERS OBVIOUS CLAIM 11

A. Motivation to Combine

The POSA would have found it obvious to modify Li to include artificial cues based on actual physical objects in the user’s room, as Woo teaches. Kyriakakis ¶¶181-182. Delivering sound perceived by the user to actually originate from the SLP in the user’s room would have enhanced Li’s conferencing system and achieved Li’s goal of providing a realistic and immersive experience. Kyriakakis ¶181. Woo photographs an actual object in the user’s room and adjusts the sound according to the object’s position in the room to provide better audio to the listener. *Id.* The POSA would have known to combine this teaching in Woo with Li’s system so that the sound is perceived by the user to actually originate from the SLP in the user’s room, and accounts for physical objects in the room. *Id.*

¹⁴ Assuming cl. 19 were construed as means-plus-function, it would still have been obvious over Li-Sandgren because Sandgren discloses specific algorithms for calculating attenuation and reverberation. *See, e.g.*, EX1017, Eqs. 1-24.

The combination would have been a natural extension of Li. Kyriakakis ¶182. For example, using Woo’s detection of actual physical objects and corresponding sound adjustment would have helped to achieve Li’s primary goal of simulating “what humans hear in the real world” in “a real conference room environment.” EX1006, 2:34-36, 22:1-5; *see also id.* 2:22-24, 13:57-60; Kyriakakis ¶182.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Li with Woo. Kyriakakis ¶¶183-184. Consistent with Li, Woo tracks a user’s position, adjusts the sound and optimizes the sound delivery. Kyriakakis ¶¶183. A POSA would have recognized that Li and Woo are structurally and functionally similar. Kyriakakis ¶183.

A POSA implementing Li would have recognized that the sound processing would have been accompanied with Woo’s sound adjustment based on actual physical objects in the user’s room, in order to artificially simulate their presence and enhance the user’s experience. Kyriakakis ¶184. Such an implementation would have merely involved a combination of known prior art elements (camera and sound processor) according to known methods and techniques (simulate spatialized sounds) to yield predictable results (producing sound so that the user perceives the sound as if it originates in the user’s actual room), and involves use of a known technique (generating realistic spatializations for sounds) 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 11:

1. 11[a]: Physical Objects

Li-Woo teaches this limitation. Kyriakakis ¶¶185-186.

While Li “mimic[s] a real conference room environment” (EX1006, 22:1-5), Li does not expressly disclose this limitation. But this would have been obvious in view of Woo. Kyriakakis ¶186. Woo discloses “photograph[ing] the object 200 and generat[ing] the plurality of images,” and “recogniz[ing] the position of the object 200” from the images. *Id.*, ¶¶53-54, Fig. 5; *see also id.*, ¶¶3, 40, 42-44, 47-49, 52, 55, Abstract. As the POSA would have appreciated, recognizing an object’s position means Woo *determine[s] physical objects located in a room with the person*. Kyriakakis ¶186.

2. 11[b]: Adjust Voice to Emulate Sound

Li-Woo teaches this limitation. Kyriakakis ¶¶187-188. As Woo explains, “[t]he sound processing apparatus 100 adjusts the property of the sound corresponding to the position of the object 200” that is determined from the photographic images. EX1016, ¶¶54-55, Fig. 5; *see also id.*, ¶¶3, 40, 45, 48-49, claims 1, 7, 13, 19. By adjusting the sound based on the object’s position in the images, Woo *adjust[s] the sound field based on the object 200 (physical objects*

located in the room). Kyriakakis ¶188. Such properties include “frequency, waveform, delay time, volume according to a frequency band and left/right balance” (EX1016, ¶45), which the POSA would have understood were properties to be accounted for in the sound field if there are physical objects that reflect and/or absorb sound. Kyriakakis ¶188. As the POSA would have understood, to meet Li’s goal of providing a realistic and immersive experience, it would have been obvious to, upon determination of physical objects in the same room, provide a realistic sound field accounting for the presence and movement of those physical objects, and provide *the voice of the computer program that the person hears to emulate a sound that the person would hear if the voice of the computer program actually originated from the SLP at the second location*. *Id.*; see Section VIII.C.5 (Ground 1, 1[d]/8[d]).

XIII. GROUND 6: BEDINGFIELD RENDERS OBVIOUS CLAIMS 1, 8, 15, 17

A. Motivation to Modify

It would have been obvious to a POSA to modify Bedingfield to (1) provide as binaural sound *a voice of a computer program during a voice exchange between the computer program and a user (or person)*, and (2) implement the multi-channel processing device as an *HPED*. Kyriakakis ¶¶189-193. A POSA would have been motivated to implement this modification for the following reasons.

1. *Voice of a computer program*

For the same reasons as discussed for Li in Ground 1, the POSA would have recognized certain advantages to provide a *voice of a computer program* as Bedingfield’s audio signal, as a computerized text-to-speech synthesis would have allowed a speaking-impaired person to participate in Bedingfield’s conference call system, improved Bedingfield’s clarity while saving time, cost, and storage space, and added flexibility. *See* Section VIII.A (Ground 1, Motivation to Modify). Kyriakakis ¶190.

The modification would have been a natural extension of Bedingfield’s “synthetic audio placement.” EX1005, Abstract, 1:15-17, 1:50-2:16, 4:18-19. Bedingfield already discloses “generating a synthetic audio signal” (EX1005, 1:63-64), which the POSA would have known to be *a voice of a computer program*. Kyriakakis ¶191.

2. *HPED*

A POSA would have recognized the advantages of Bedingfield’s listener 102 (*user or person*) having an *HPED*. Kyriakakis ¶192. An *HPED* is easily transported without requiring any fixed location or physical wired connection. *Id.* Being *handheld* and *portable* would have meant that the listener 102 remains connected to Bedingfield’s conference call while moving. *Id.*

The modification would have been a natural extension of Bedingfield's disclosure of various *HPEDs* such as “telephones, wireless phones, [or] mobile phones” as connecting participants to the conference call system. *See* EX1005, 4:23-30, 11:41-51 (describing mobile or handheld computers, mobile telephone, and PDAs); Kyriakakis, ¶193.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Bedingfield with (1) *a voice of a computer program* and (2) implement the multi-channel processing device as an *HPED*. Kyriakakis ¶¶194-197.

Given that the proposed modification involves (1) a simple addition of another sound source and (2) using a handheld portable device, both of which have many advantages, it is nothing more than a combination of known prior art elements according to known methods and known techniques to yield predictable results, and involves use of known techniques to improve a similar device in the same way. Kyriakakis ¶195.

1. *Voice of a computer program*

Bedingfield provides audio to both ears that the listener perceives to originate from an SLP. Bedingfield's “sound field provid[es] synthetic audio placement of each of the plurality of remote audio signals.” EX1005, Abstract. Bedingfield contemplates using its sound field in a conference call setting. *Id.*, Abstract, 1:21-

46, 3:25-34, 5:4-15, Figs. 1-3). While Bedingfield does not expressly disclose *a voice of a computer program*, a POSA would have had a reasonable expectation of success using such a voice in Bedingfield, to save time and cost, add flexibility, and improve clarity. Kyriakakis ¶196. Modifying Bedingfield would have been as simple as using a computer-synthesized voice (instead of a recorded voice) as an audio signal, and would have been obvious especially in view of Bedingfield's teaching of "generating a synthetic audio signal." EX1005, 1:63-64; Kyriakakis ¶196.

2. *HPED*

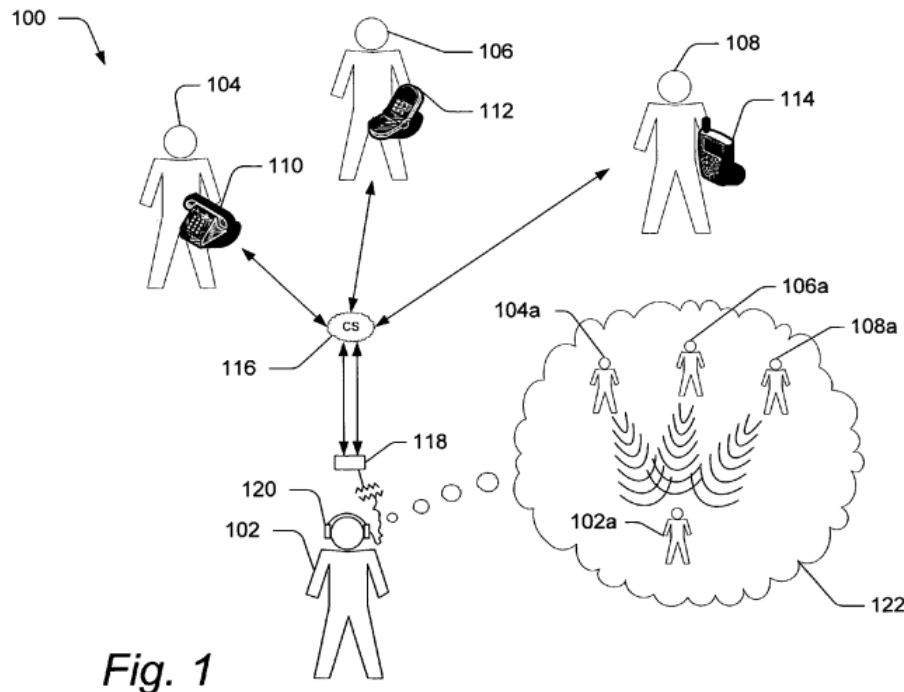
Bedingfield already describes providing sounds to the listener over headphones, which are an *HPED*. The POSA would have had a reasonable expectation of success implementing the multi-channel processing device 118 to be *handheld* and *portable* because such *HPEDs* were well-known and widely used well before the relevant time period. Kyriakakis ¶197. Bedingfield already contemplates various *HPEDs*, and the POSA would have found it straightforward to implement the multi-channel processing device 118 as an *HPED*. *Id.*

C. Independent Claims 1 and 8

1. 1[pre]/8[pre]: Move SLP of Voice of Computer Program

To the extent limiting, Bedingfield teaches the preambles. Kyriakakis ¶¶198-204.

Bedingfield discloses *one or more electronic devices in an electronic system* by describing a communication system 116 and a multi-channel processing device 118 that *execute a method*. EX1005, 4:31-51, Fig. 1, reproduced below).



As Bedingfield explains, the combination of the communication system 116, a multi-channel processing device 118, and multi-channel speaker system 120 (*electronic system*) has *one or more electronic devices* (which the POSA would have known to include a processor and memory). *Id.*, 4:31-51. Kyriakakis ¶199.

Specific to 8[pre], Bedingfield discloses a *non-transitory computer readable storage medium storing instructions*. Kyriakakis ¶200. Bedingfield states “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 (emphases added); *see also* 11:23-36, 12:5-44, Fig. 12. Bedingfield therefore discloses the communication system 116, the multi-channel processing device 118, and/or the multi-channel speaker system 120 (together, *the electronic system*) implement functions using software and would include *a memory that stores instructions*. Kyriakakis ¶200.

In Bedingfield, the communication system 116 and the multi-channel processing device 118 “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. In other words, Bedingfield’s *electronic system* simulates the sound field 122 as if the listener and participants “were in the same location and speaking to each other directly.” *Id.*, 4:52-61; Kyriakakis ¶201.

Regarding *mov[ing] an SLP*, Bedingfield depicts in Fig. 3 (reproduced below) “method 300 of how synthetic audio placement may be organized, including selections a listener may make,” including “on the multi-channel processing device 118” of Fig. 1. EX1005, 5:51-55; Kyriakakis ¶202.

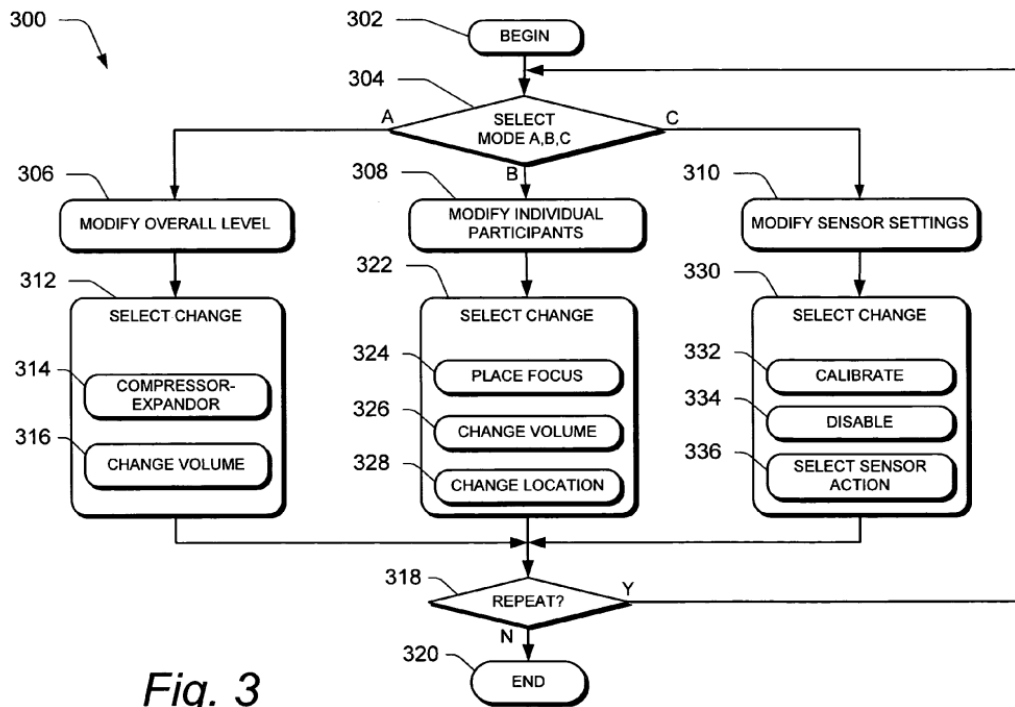


Fig. 3

“[A]t the block 328, the listener selects a participant and then selects a new location for the participant within the sound field” such as sound field 122. EX1005, 6:43-45; *see also* 6:29-32, 9:60-64, 14:7-10, 16:33-35. By using the multi-channel processing device 118 (part of the *electronic system*) to move a participant’s location within the sound field, Bedingfield’s *one or more electronic devices in an electronic system move an SLP of a voice*. Kyriakakis ¶202.

In Bedingfield, this takes place *during a voice exchange between the computer program and a user/person*. Kyriakakis ¶203. Bedingfield connects participants in a conference call (*see, e.g.,* EX1005, 5:4-10, Fig. 2) and user’s selections in Fig. 3 are available “when multiple participants are speaking during a conference” (*id.*, 6:36-

37), when a *voice exchange* amongst participants occurs (*id.*, 4:25-30, 4:36-61, 6:9-10, 9:60-64); Kyriakakis ¶203.¹⁵

Bedingfield's *voice exchange* includes a processed voice of a participant, or a *voice of a computer program*, as Bedingfield discloses “*computer program* products,” including “software modules” that “provide synthetic audio placement.” EX1005, 2:17-24, 12:9-15 (emphasis added); Kyriakakis ¶204. For the same reasons as explained above for Li, the POSA would have been motivated to implement Bedingfield's sound source as a *voice of a computer program*. *Id.*; Sections VIII.A-B.¹⁶

**2. 1[a]/8[a]: Provide Through Earphone/WED Binaural Sound
Such That Origin Occurs at SLP**

Bedingfield teaches this limitation. Kyriakakis ¶¶205-209.

¹⁵ The mapping for the phrase *during a voice exchange between the computer program and a user/peron* also applies to 1[a]/8[a], 1[c], 1[d]/8[d], 15[pre], 15[c3], which recite the same phrase or a similar phrase, and the mapping is not repeated in those sections, *infra*.

¹⁶ The mapping for the phrase *voice of a computer program* also applies to 1[a]/8[a]-1[d]/8[d], 15[pre]-15[c3], which recite the same phrase or a similar phrase, and the mapping is not repeated in those sections, *infra*.

Bedingfield's *electronic system* includes a multi-channel speaker system 120 depicted in Fig. 1 as headphones, which includes *an electronic earphone on a person being located at a head of the user and being in communication with a multi-channel processing device 118*, which would have been obvious to implement as an *HPED* (1[a]), or *a WED on the person* (8[a]). EX1005, 4:31-51, Fig. 1; *see also* 7:43-46, 13:34-36. It was common in 2014 for *electronic earphones* such as earbuds or headphones to connect to such processing devices. Kyriakakis ¶206. Specifically, Bedingfield's audio signals provided to the multi-channel speaker system 120 “enable the listener 102 to hear the speaking participant in the sound field 122,” representing the listener 102's perception of the participants 104, 106, and 108 as if they “were in the same location and speaking to each other directly.” EX1005, 4:47-5:3 (also stating sounds are provided to both ears). By rendering a sound field and providing synthetic audio placement using both ears, Bedingfield *provid[es] binaural sound to the user/person*. Kyriakakis ¶206.

Specific to 1[a], while Bedingfield does not expressly describe the multi-channel processing device 118 as an *HPED*, Bedingfield describes conference participants the conference call as having “communication devices 110, 112, 114 [which] may be telephones, wireless phones, [or] mobile phones” (EX1005, 4:23-30; *see also* 11:41-51 (“mobile, handheld, or laptop computer; a mobile telephone; a [PDA]”), any of which are *HPEDs*. Kyriakakis ¶207. The POSA would have

understood to implement the multi-channel processing device 118 as an *HPED*, allowing the listener 102 (*user/person*) to move with the multi-channel processing device 118 and remain connected to the conference. *Id.* Specific to 8[a], Bedingfield's headphones were wireless and therefore meet *WED* under PO's position. EX1005, 13:39-43 (describing wireless connection); Kyriakakis ¶207.

Bedingfield discloses the remaining elements of 1[a]/8[a]. The *SLP* is *at a first location in empty space that is away from and proximate to the user/person* because as Bedingfield depicts in Fig. 1, the sound field 122 “represents the perception resulting from the audio output” for the listener 102 and is similar to “an actual sound field where the listener 102 is present between all of participants 104, 106, 108.” EX1005, 4:40-5:3; *see also* 9:60-64 (describing a voice “occupy[ing] a ‘vacant’ location”). Because each participant has its own location in the sound field that is separate from and at a distance to the listener 102's location (*see analysis for 1[pre]/8[pre]*), each participant's *SLP* is *at a first location in empty space that is away from and proximate to the* listener 102 (*user/person*). Kyriakakis ¶¶208-209. Moreover, as explained in Ground 1 for Li, it would have been obvious to position the *SLP* in *empty space*, at a location where no physical object exists. Kyriakakis ¶¶105-120, 209. For example, Bedingfield's system “closely simulate[s] an in-person conference setting” (EX1005, 3:25-27, 4:31-5:3), such that the POSA would have found it obvious to seat the *user/person* at a physical conference table, where

an empty seat at the table (*empty space*) is allocated for a virtual participant, as was known in the field. Kyriakakis ¶¶110-115, 209.¹⁷

3. 1[b]/8[b]: Receive Command to Move SLP

Bedingfield teaches this limitation. Kyriakakis ¶¶210-212.

Bedingfield's *electronic system* includes the multi-channel processing device 118, or an *HPED* (specific to 1[b]). *See* Section XIII.C.2. Bedingfield's "listener may make selections on the multi-channel processing device 118 [*HPED*]." EX1005, 5:51-61, Fig. 3 (reproduced below); *see* Section XIII.C.1. Therefore, Bedingfield's multi-channel processing device 118 (*HPED*) *receives* selections *from the listener (user/person)*. Kyriakakis ¶211.

¹⁷ The mapping for *SLP at a first/second location in empty space that is away from and proximate to the user/person* also applies to 1[b]/8[b]-1[d]/8[d], 15[c1]-15[c3], which have the same or similar phrase, and the mapping is not repeated, *infra*.

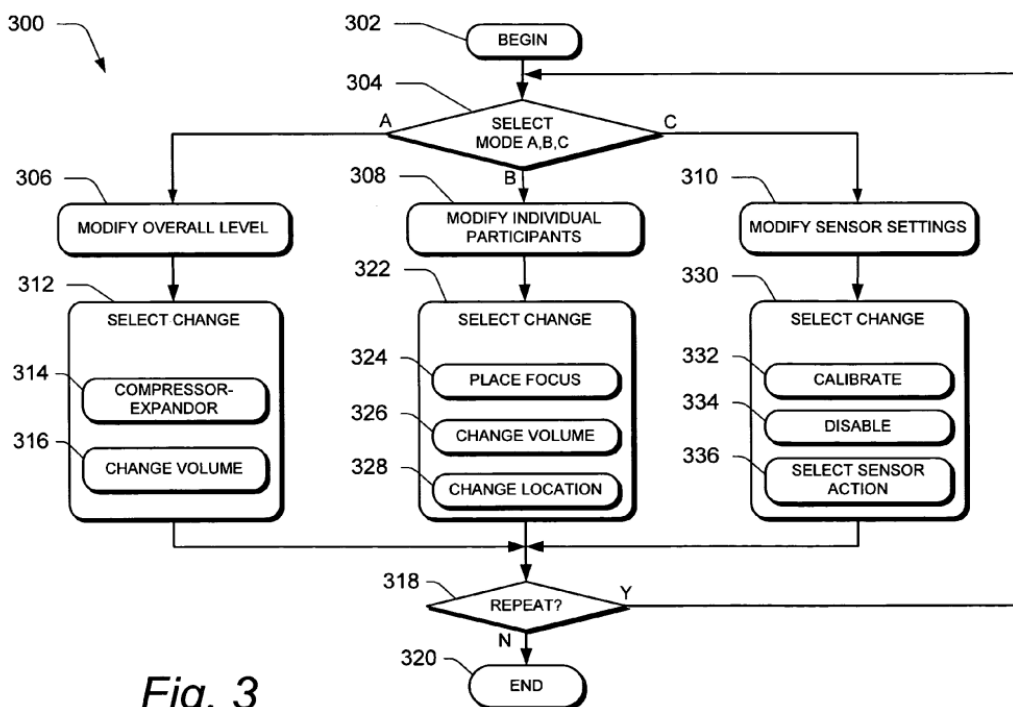


Fig. 3

At “block 328, the listener selects a participant and then selects 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 (*user/person*) to select a conference participant and their new location, Bedingfield discloses *receiving, from the user, a command to move the SLP of the voice... from the first location in empty space to a second location in empty space that is proximate to the first location and away from and proximate to the user.* Kyriakakis ¶211.

4. 1[c]/8[c]: Move SLP

Bedingfield teaches this limitation. Kyriakakis ¶¶213-214.

Bedingfield’s *electronic system moves, based on the command and, the location (SLP) of a speaker’s voice (voice of the computer program) from the*

previous position (*first location*) to the “new location” (*second location*), where the new location (*second location*) is within the sound field (*proximate to the first location*). EX1005, 4:18-5:3, 5:51-61, 6:29-32, 6:43-45, Figs. 1, 3; *see also* Sections XIII.C.1-3 (Ground 6, 1[pre]/8[pre]- 1[b]/8[b]); Kyriakakis ¶198-212, 214.

5. 1[d]/8[d]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location

Bedingfield teaches this limitation. Kyriakakis ¶¶215-216.

Bedingfield’s *electronic system provides, through headphones (electronic earphone being located at a head of the user being in communication with the HPED (1[d]) or WED on the person (8[d])), the audio rendering of the sound field to both ears (binaural sound) to the listener (user/person). See Section XIII.C.2 (Ground 6, 1[a]/8[a]); Kyriakakis ¶205-209, 216. Moreover, in Bedingfield, after the listener (user/person) selects at block 328 a new location for the participant within the sound field in 1[c]/8[c], the location for that participant is updated to the new location (such that the origin of the voice of the computer program occurs at the SLP at the second location). See Section XIII.C.1-4 (Ground 6, 1[pre]/8[pre]-1[c]/8[c]); Kyriakakis ¶198-214, 216.*

D. Independent Claim 15

1. 15[pre]: Move SLP of Voice of Computer Program

To the extent limiting, Bedingfield teaches the preamble. *See* Section XIII.C.1 (Ground 6, 1[pre]/8[pre]); Kyriakakis ¶198-204, 217.

2. 15[a]: WED Provides Voice of Computer Program at SLP

Bedingfield teaches this limitation, and Bedingfield's headphones is *a WED that the listener (person) wears*. *See* Section XIII.C.1-2 (Ground 6, 1[pre]/8[pre], 1[a]/8[a]); Kyriakakis ¶198-209, 218.

3. 15[b]: Memory

Bedingfield discloses this limitation. *See* Section XIII.C.1 (Ground 6, 1[pre]/8[pre]); Kyriakakis ¶198-204, 219.

4. 15[c]: Processor

Bedingfield discloses this limitation. *See* Section XIII.C.1 (Ground 6, 1[pre]/8[pre]); Kyriakakis ¶198-204, 220.

5. 15[c1]: Receive Command to Move SLP

Bedingfield teaches this limitation. *See* Section XIII.C.3 (Ground 6, 1[b]/8[b]); Kyriakakis ¶210-212, 221.

6. 15[c2]: Move SLP

Bedingfield teaches this limitation. *See* Section XIII.C.4 (Ground 6, 1[c]/8[c]); Kyriakakis ¶213-214, 222.¹⁸

7. 15[c3]: Provide Binaural Sound Such That Origin Occurs at SLP at Second Location

Bedingfield teaches this limitation. *See* Section XIII.C.5 (Ground 6, 1[d]/8[d]); Kyriakakis ¶215-216, 223.

E. Dependent Claim 17: Change ITD and ILD to Move SLP

Bedingfield teaches this limitation. Kyriakakis ¶¶224-225. Bedingfield discloses that *ITD* and *ILD* are directional cues that place a sound source in the sound field. Kyriakakis ¶224. Regarding *ILD*, Bedingfield explains that to provide directionality in the sound field, the 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. Regarding *ITD*, Bedingfield refers to applying a “delay” as an audio technique to place a sound source in the sound field, which the POSA would have known includes the well-known concept of an *ITD*. *Id.*, 5:34-43; Kyriakakis ¶224.

¹⁸ Even if limitation 15[c2] were construed as a means-plus-function term, it would have been obvious because PO admitted that the algorithm for moving an SLP uses “conventional techniques” (EX1032, 45-46).

As Bedingfield describes each of these cues as being used to set the direction or position of a sound source in the sound field, the POSA would have known to use Bedingfield's *processor to execute instructions to change an ITD and an ILD of the binaural sound in order to move the SLP*. Kyriakakis ¶225.

**XIV. GROUND 7: BEDINGFIELD IN VIEW OF BALLAS RENDERS
OBVIOUS CLAIM 4, 6, 9-10, 14, 18**

A. Motivation to Combine

The POSA would have found it obvious to apply Ballas's head tracking and fixing the perceived sound to Bedingfield. Kyriakakis ¶¶226-228. Bedingfield already discloses "detect[ing] the orientation of a listener 104" and that "the sound field may correspondingly change to reflect the new sound field orientation." EX1005, 13:38-39, 14:35-38. Bedingfield's listener also provides manual input to change the orientation, causing the sound field to adjust to appear fixed to the listener. *Id.*, 14:11-43; Figs. 14, 15. A POSA would have been motivated to combine Bedingfield and Ballas because both references have a shared goal of providing a realistic experience, and implementing Ballas's head tracking would relieve the user of having to input manual commands. Kyriakakis ¶226. To make this goal a reality, the POSA would have understood to track the user's head orientation and movements and use Ballas's conventional head tracking and spatialization techniques to account for the user's movements in synthesizing the

sound field. Kyriakakis ¶¶226. Thus, although Bedingfield does not explicitly describe head tracking and adjusting the sound field accordingly, this would have been obvious. Kyriakakis ¶226.

Bedingfield's primary goal is to provide a sound field that "closely simulate[s] an in-person conference setting." EX1005, 3:25-27. As explained for Ground 2, Ballas shares this goal. EX1008, ¶¶2, 9, 10. Bedingfield detects the listener's orientation and changes the sound field. EX1005, 1:56-59, 13:38-39, 13:58-64, 14:32-38, 18:49-52. To simulate the in-person conference setting, head tracking and changing the sound field so that the sound sources appear fixed to the listener are key. Kyriakakis ¶227. Modifying Bedingfield with Ballas would have allowed the user to move while staying connected, and achieved Bedingfield's express purpose to simulate an in-person conference setting. Kyriakakis ¶227.

The modification would have been a natural extension of Bedingfield's use of motion detectors, sensors, and devices to detect the listener's orientation. Kyriakakis ¶228; EX1005, 13:54-56. As explained for Ground 2, Ballas discloses head tracking and "maintaining virtual location of the audio signals." EX1008, Abstract, ¶18; *see also id.*, ¶¶17, 19-20, 25-26, claims 1, 7. Using Ballas's head tracking would have allowed the user to change head location and orientation while maintaining fixed virtual sound locations, simulating an in-person conference setting. Kyriakakis ¶228. This implementation would have advanced the common goals of both

Bedingfield and Ballas to provide a realistic environment in a teleconferencing application. Kyriakakis ¶228.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Bedingfield with Ballas. Kyriakakis ¶¶229-230. Both track the listener's movement and orientation and change the sound field accordingly. In Bedingfield, the user wears "wearable speakers" in the form of stereo headphones. EX1005, 13:35-36. As explained for Ground 2, Ballas also discloses speakers that are attached to a wearable head frame. EX1008, ¶25. Accordingly, a POSA would have recognized that Bedingfield and Ballas are structurally and functionally similar. Kyriakakis ¶229.

A POSA implementing the Bedingfield system would have recognized that Bedingfield's headphones could be attached to Ballas's frame. Kyriakakis ¶¶230. 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 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). *Id.*

C. Dependent Claim 4

1. 4[a]: Track Head

Bedingfield-Ballas teaches this limitation. Kyriakakis ¶¶231-233.

Bedingfield’s “device [] detect[s] the orientation of a listener 104” (EX1005, 13:38-39; *see also id.*, 6:54-58, 13:52-57), which is *tracking a[n] orientation of the user*. Kyriakakis ¶232. The listener’s orientation is relative to *the SLP*. *See* EX1005, 1:56-59, 18:49-52; Kyriakakis ¶232. Bedingfield does not explicitly disclose that what is *track[ed]* is the *orientation* of the user’s *head, as the user moves with respect to the SLP at the second location*. *Id.* However, this would have been obvious in view of Ballas, which discloses this limitation as explained for Ground 2. *See* Section IX.C.1 (Ground 2, 4[a]); Kyriakakis ¶¶144-146, 232.

The POSA would have been motivated to implement Bedingfield’s wearable speakers as part of Ballas’s wearable frame having a head tracking system. Kyriakakis ¶233; *see also* Section XIV.A (Motivation to Combine).

2. 4[b]: Adjust Voice To Continue To Appear To Originate From SLP

Bedingfield-Ballas teaches this limitation. Kyriakakis ¶¶234-236.

Bedingfield discloses “adjust[ing] the focus” of the sound field by “chang[ing] the orientation of the sound field 122.” EX1005, 13:58-64; *see also id.*, 6:57-7:2, 4:7-12. As Bedingfield describes for Figures 14 and 15 (reproduced below with

annotations), the listener provides a manual input to “change[] the focus [direction], [and] the sound field [] correspondingly change[s] to reflect the new sound field orientation.” *Id.*, 14:32-38; Kyriakakis ¶¶235-236.

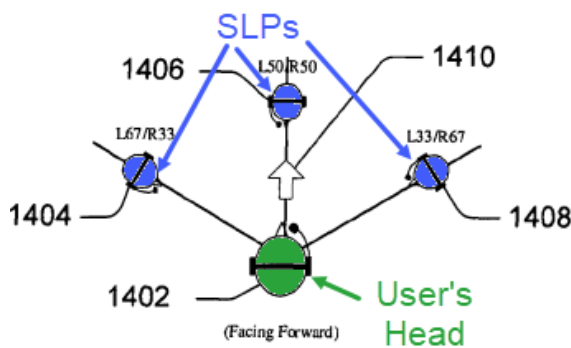


Fig. 14

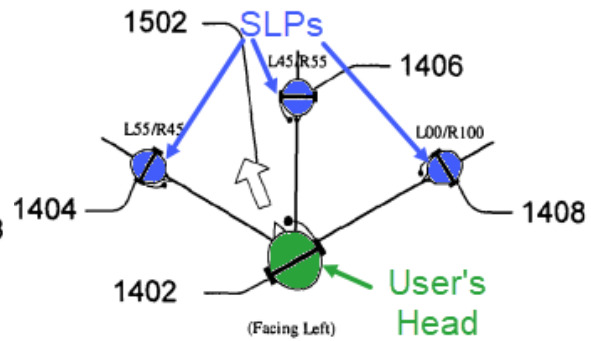


Fig. 15

Thus, Bedingfield discloses *adjusting a sound... that the user hears through the electronic earphone so the sound of the voice of the computer program continues to appear to originate from the SLP at the second location.* Kyriakakis ¶235.

While Bedingfield’s sound field is adjusted based on manual input, not according to the user’s movement, this would have been obvious in view of Ballas, which discloses this limitation as discussed for Ground 2. *See* Section IX.C.2 (Ground 2, 4[b]); Kyriakakis ¶¶147, 148, 236.

D. Dependent Claim 6: Change Amplitude, ITD, and ILD to Continue to Originate from SLP as User Moves

As explained in Ground 6 for claim 17, Bedingfield *chang[es] an ILD* and an *ITD*. Kyriakakis ¶¶224-225, 237. *See* Section XIII.E (Ground 6, 17).

Bedingfield refers to overall “volume” (*amplitude*) as providing relative distance information. EX1005, 3:45-47 (“A signal of lower volume sounds farther away than a signal of higher volume.”). As Bedingfield describes each of these cues as being used to set the direction or position of a sound source in the sound field, the POSA would have found it obvious to *chang[e] an amplitude, an ITD, and an ILD of the voice* to modify the sound field *so a sound of the voice... continues to appear to originate from the SLP... as the user moves*. Kyriakakis ¶237.

E. Dependent Claim 9: Maintain SLP But Alter Amplitude

Bedingfield-Ballas teaches this limitation. *See* Section XIV.D (Ground 7, Claim 6); Kyriakakis ¶¶224-225, 238.

F. Dependent Claim 10

1. 10[a]: Monitor Movements

Bedingfield-Ballas teaches this limitation. *See* Section XIV.C.1 (Ground 7, 4[a]); Kyriakakis ¶¶231-233, 239.

2. 10[b]: Adjust ITD and Amplitude

Bedingfield-Ballas teaches this limitation. *See* Section XIII.E (Ground 6, 17) and Sections XIV.C.1 and XIV.D (Ground 7, 4[b] and 6); Kyriakakis ¶¶224-225, 234-237, 240.

G. Dependent Claim 14: Adjust ILD, ITD, and Amplitude So SLP Remains Fixed

Bedingfield-Ballas teaches this limitation. Kyriakakis ¶¶241-242. *See* Sections XIV.C-D (Ground 7, cls. 4, 6); Kyriakakis ¶¶231-237.

H. Dependent Claim 18: Change ITD and Amplitude to Maintain SLP

Bedingfield-Ballas teaches this limitation. Kyriakakis ¶¶243-244. Bedingfield-Ballas renders obvious *chang[ing] an ITD and an amplitude of the binaural sound in order to maintain the SLP fixed at the second location in empty space... while the person... changes a head orientation with respect to the second location in empty space.* Section XIV.D (Ground 7, cl. 6); Kyriakakis ¶237. As explained in 15[c], Bedingfield's *processor executes instructions* to perform this processing, and as explained in 1[pre]/8[pre], this occurs *during the voice exchange.* Kyriakakis ¶¶198-204, 220. *See* Sections XIII.D.4, XIII.C.1 (Ground 6, 15[c] and 1[pre]/8[pre]).

As explained for Ground 2, Ballas further discloses *the person physically mov[ing] away from the second location in empty space.* Kyriakakis ¶¶158-160. *See* Section IX.H (Ground 2, cl. 18). To provide as realistic and as immersive of an experience as possible for the user, the Bedingfield-Ballas system would have processed sound *in order to maintain the SLP fixed at the second location* regardless of the user's movements or head orientation. Kyriakakis ¶244.

XV. GROUND 8: BEDINGFIELD IN VIEW OF BALLAS AND SANDGREN RENDERS OBVIOUS CLAIM 5

A. Motivation to Combine

A POSA would be motivated to combine the teachings of Bedingfield and Ballas. *See* Section XIV.A (Ground 7, Motivation to Combine); Kyriakakis ¶¶226-228, 245. Bedingfield recognizes that a sound’s “volume” (*amplitude*) provides distance information. EX1005, 3:45-47 (“A signal of lower volume sounds farther away than a signal of higher volume.”). A POSA would have been motivated to look to Sandgren to determine how to vary Bedingfield’s sound volume with distance, to provide a realistic conferencing experience. Kyriakakis ¶245. Modifying Bedingfield and Ballas in accordance with Sandgren’s teachings to adjust a sound amplitude according to an inverse-square law with distance (EX1017, ¶¶1, 77, Abstract) would have provided a realistic experience, and would have been a natural extension of Bedingfield’s volume setting. Kyriakakis ¶245. There is nothing in Ballas that would prevent a POSA from modifying Bedingfield with Sandgren’s teachings. Kyriakakis ¶245.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Bedingfield with Ballas. *See* Section XIV.B (Ground 7, Reasonable Expectation of Success). Kyriakakis ¶¶194-197, 246-247. A POSA would have had a reasonable

expectation of success in further modifying Bedingfield-Ballas with Sandgren. Kyriakakis ¶246. Bedingfield and Sandgren both simulate realistic auditory scenes having spatialized binaural sound over headphones. Kyriakakis ¶246. As explained in Ground 3, Sandgren creates virtual 3D audio scenes to externalize virtual sound sources. *See* Section X.B (Ground 3, Reasonable Expectation of Success). Accordingly, Bedingfield and Sandgren are structurally and functionally similar. Kyriakakis ¶246.

A POSA implementing Bedingfield would have recognized that Sandgren's teachings of the inverse square law and virtual object simulation would have been straightforward to implement in Bedingfield's spatialized audio. Kyriakakis ¶247. This would have merely involved a combination of known prior art elements (inverse square law and virtual objects) according to known methods and techniques (simulate distance and objects) to yield predictable results (realistic sound fields), and involves use of a known technique (sound adjustments) to improve a similar device (wearable communication device) in the same way (adjusting sounds to produce a realistic sound field). *Id.*

C. Dependent Claim 5: Adjust Amplitude to be Inversely Proportional to Square of Distance

Bedingfield-Ballas in view of Sandgren teaches this limitation. Kyriakakis ¶¶248-250.

As explained for claim 6, Bedingfield-Ballas teaches *adjust[ing] an amplitude of the voice of the computer program... as the user moves with respect to the SLP at the second location in empty space. See Section XIV.D (Ground 7, cl. 6); Kyriakakis ¶¶237, 249.*

As explained in Ground 3, Sandgren discloses an inverse-square law, and to provide a realistic sound field, it would have been obvious to *adjust an amplitude... to be inversely proportional to a square of a distance between the user and the SLP. See Section X.C (Ground 3, 5); Kyriakakis ¶¶166-168, 250.*

XVI. GROUND 9: BEDINGFIELD IN VIEW OF SANDGREN RENDERS OBVIOUS CLAIM 19

A. Motivation to Combine

It would have been obvious to modify Bedingfield 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. Kyriakakis ¶¶251-252. A POSA would have been motivated to combine the teachings of Bedingfield and Sandgren because a POSA would have recognized that both references have a shared goal of providing a realistic and immersive experience for the user. Kyriakakis ¶251. To make this goal a reality, 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 taught by

Sandgren. *Id.* Thus, although Bedingfield does not describe changing attenuation and reverberation of the binaural sound to add artificial audial cues of virtual objects, a POSA would have been motivated to perform such an implementation, as doing so has the benefits described for Ground 4. *Id.* Modifying Bedingfield in accordance with Sandgren’s teachings of adjusting attenuation and reverberation to simulate virtual objects would have provided a more realistic and immersive experience for the user. *Id.*

Moreover, the modification would have been a natural extension of Bedingfield’s providing a sound field of spatialized sounds over headphones. Kyriakakis ¶252. Like Bedingfield, Sandgren “simulat[es] an audio scene or creat[e] a virtual audio scene” over headphones. EX1017, ¶¶9-10; *see also* ¶¶11, 15-16, 19, 22, 46, 48-64, 70-100, Figs. 5A-C. For the same reasons as discussed for Ground 4, a POSA would have recognized that the teachings of Sandgren would have enhanced Bedingfield’s goal of an immersive experience by introducing attenuation filters and reverberation generators indicative of virtual objects in the environment. Kyriakakis ¶252. This implementation would have advanced the common goals of both Bedingfield and Sandgren to provide a realistic and immersive experience for the user. Kyriakakis ¶252.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Li in accordance with the teachings of Sandgren. *See* Section XV.B (Ground 8, Reasonable Expectation of Success). Kyriakakis ¶253.

C. Dependent Claim 19

1. 19[a]: Change Attenuation and Reverberation

Bedingfield-Sandgren teaches this limitation. Kyriakakis ¶254. Sandgren discloses reverberation generators and attenuation filters *to change attenuation and reverberation of binaural sound to add artificial audial cues of virtual objects near a person*. Kyriakakis ¶254. *See also* Section XI.C.1 (Ground 4, 19[a]).¹⁹

2. 19[b]: Virtual Objects

Bedingfield-Sandgren teaches this limitation. Kyriakakis ¶255. Sandgren discloses objects that *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.* *See also* Section XI.C.2 (Ground 4, 19[b]).

XVII. GROUND 10: BEDINGFIELD AND WOO RENDERS OBVIOUS CLAIM 11

¹⁹ Even if limitation 19[a] were construed as a means-plus-function term, it would still have been obvious over Bedingfield in view of Sandgren because Sandgren discloses structural algorithms for calculating attenuation and reverberation. *See, e.g.,* EX1017, Eqs. 1-24.

A. Motivation to Combine

It would have been obvious to a POSA to modify Bedingfield to adjust the voice to include artificial cues based on actual physical objects in the user's room, as taught in Woo. Kyriakakis ¶¶256-257. As the POSA would have recognized, an implementation that deliver a sound perceived by the user to actually originate from the SLP in the user's room, would have enhanced Bedingfield's conferencing system and achieved Bedingfield's goal of simulating a face-to-face conference as closely as possible. Kyriakakis ¶256. Woo photographs an actual object in the user's room and adjusts the sound according to the object's position in the room. Kyriakakis ¶256. The POSA would have known to combine this conventional teaching in Woo with Bedingfield's system so that the sound is perceived by the user to actually originate from the SLP in the user's room, and accounts for physical objects in the room. Kyriakakis ¶256.

The combination would have been a natural extension of Bedingfield's remote conferencing system. Kyriakakis ¶257. For example, using Woo's detection of actual physical objects in the room and corresponding adjustment of the sound would have helped to achieve Bedingfield's primary goal of providing an enhanced sound field that "closely simulate[s] an in-person conference setting." EX1005, 3:25-27; Kyriakakis ¶257. Moreover, the POSA would have understood that adjusting a

sound based on photographs of actual objects in the room would have enhanced Bedingfield's sound delivery and provided a more realistic conferencing experience. Kyriakakis ¶257.

B. Reasonable Expectation of Success

A POSA would have had a reasonable expectation of success in modifying Bedingfield in accordance with the teachings of Woo. Kyriakakis ¶¶258-259. Consistent with Bedingfield's system, Woo tracks a user's position, adjusts certain properties of a sound and optimizes the delivery of the sound for the user. Kyriakakis ¶¶258. Accordingly, a POSA would have recognized that Bedingfield and Woo are structurally and functionally similar. Kyriakakis ¶258.

A POSA implementing the Bedingfield system would have recognized that the processing of the sound would have been accompanied with Woo's adjustment of the sound based on actual physical objects in the user's room, in order to artificially simulate their presence and simulate an in-person conference. Kyriakakis ¶259. Such an implementation would have merely involved a combination of known prior art elements (camera and sound processor) according to known methods and techniques (simulate spatialized sounds) to yield predictable results (producing sound so that the user perceives the sound as if it originates in the user's actual room), and involves use of a known technique (generating realistic spatializations for sounds) to improve a similar device (wearable communication device) in the

same way (filtering sounds to produce a realistic and immersive sound field). *Id.* ¶259.

C. Dependent Claim 11

1. 11[a]: Physical Objects

Bedingfield-Woo teaches this limitation. Kyriakakis ¶260. Woo discloses *to determine physical objects located in room with person* by photographing objects to recognize their location in a room. Kyriakakis ¶260. *See* Section XII.C.1 (Ground 5, 11[a]).

2. 11[b]: Adjust Voice to Emulate Sound

Bedingfield-Woo teaches this limitation. Kyriakakis ¶261. Woo *adjust[s]* the sound field *based on the object 200 (physical objects located in the room)* by describing that the sound is adjusted based on the object's position in the images *to emulate sound that a person would hear if a voice of the computer program actually originated from SLP at the second location*. Kyriakakis ¶261. *See* Section XII.C.2 (Ground 5, 11[b]).

XVIII. SECONDARY CONSIDERATIONS

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

XIX. 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 herein was cited or relied upon during prosecution. 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.

XX. CONCLUSION

For these reasons, Petitioner respectfully requests IPR of claims 1, 4-6, 8-11, 14-15, and 17-19 of the '196 Patent.

Date: June 9, 2023

Respectfully submitted,

By: /Lisa K. Nguyen/

Lisa K. Nguyen (Reg. No. 58,018)

lisa.nguyen@allenoverly.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 executed by one or more electronic devices in an electronic system that move a sound localization point (SLP) of a voice of a computer program during a voice exchange between the computer program and a user, the method comprising:

[a] providing, through an electronic earphone being located at a head of the user and being in communication with a handheld portable electronic device (HPED), binaural sound to the user during the voice exchange between the computer program and the user such that an origin of the voice of the computer program occurs at the SLP at a first location in empty space that is away from and proximate to the user;

[b] receiving, from the user and at the HPED, a command to move the SLP of the voice of the computer program from the first location in empty space to a second location in empty space that is proximate to the first location and away from and proximate to the user;

[c] moving, based on the command and during the voice exchange between the computer program and the user, the SLP of the voice of the computer program from the first location in empty space to the second location in empty space that is proximate to the first location and away from and proximate to the user; and

[d] providing, through the electronic earphone being located at the head of the user and being in communication with the HPED, the binaural sound to the user during the voice exchange between the computer program and the user such that the origin of the voice of the computer program occurs at the SLP at the second location in empty space that is proximate to the first location and away from and proximate to the user.

4. The method of claim 1 further comprising:

[a] tracking a head orientation of the user as the user moves with respect to the SLP at the second location; and

[b] adjusting a sound of the voice of the computer program that the user hears through the electronic earphone so the sound of the voice of the computer program continues to appear to originate from the SLP at the second location even

as the user moves with respect to the second location.

5. The method of claim 1 further comprising:

adjust an amplitude of the voice of the computer program to be inversely proportional to a square of a distance between the user and the SLP at the second location in empty space as the user moves with respect to the SLP at the second location in empty space.

6. The method of claim 1 further comprising:

changing an amplitude, an interaural time difference, and an interaural level difference of the voice of the computer program as the user moves with respect to the SLP at the second location so a sound of the voice of the computer program continues to appear to originate from the SLP at the second location as the user moves.

8[pre] A non-transitory computer readable storage medium storing instructions that cause one or more electronic devices in an electronic system to execute a method that moves a sound localization point (SLP) of a voice of a computer program during a voice exchange that is between the computer program and a person, the method comprising:

[a] provide, through a wearable electronic device on the person, binaural sound to the person during the voice exchange that is between the computer program and the person such that an origin of the voice of the computer program occurs at the SLP at a first location in empty space that is away from and proximate to the person;

[b] receive, from the person, a command to move the SLP of the voice of the computer program from the first location in empty space to a second location in empty space that is proximate to the first location and away from and proximate to the person;

[c] move, in response to the command, the SLP of the voice of the computer program from the first location in empty space to the second location in empty space that is proximate to the first location and away from and proximate to the person; and

[d] provide, through the wearable electronic device on the person, the binaural sound to the person during the voice exchange that is between the computer program and the person such that the origin of the voice of the computer program occurs at the SLP at the second location in empty space that is proximate to the first location and away from and proximate to the person.

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

maintain the origin of the voice of the computer program at the SLP at the second location in empty space but alter an amplitude of the voice of the computer program that the person hears in response to the person moving closer to or away from the SLP at the second location.

10. The non-transitory computer readable storage medium storing instructions of claim 8 further to cause the one or more electronic devices in the electronic system to execute the method comprising:

[a] monitor movements of a head orientation of the person in response to the person moving with respect to the SLP at the second location in empty space; and

[b] adjust an interaural time difference (ITD) and an amplitude of the voice of the computer program in response to the movements of the head orientation of the person as the person moves with respect to the SLP at the second location.

11. The non-transitory computer readable storage medium storing instructions of claim 8 further to cause the one or more electronic devices in the electronic system to execute the method comprising:

[a] determine physical objects located in a room with the person; and

[b] adjust, based on the physical objects located in the room, the voice of the computer program that the person hears to emulate a sound that the person would hear if the voice of the computer program actually originated from the SLP at the second location.

14. The non-transitory computer readable storage medium storing instructions of claim 8 further to cause the one or more electronic devices in the electronic system

to execute the method comprising:

adjust an interaural level difference, an interaural time difference, and an amplitude for the voice that the person hears so the SLP of the voice of the computer program remains fixed at the second location in empty space while a head orientation of the person moves with respect to the second location.

15. An electronic system with one or more electronic devices that move a sound localization point (SLP) of a voice of a computer program during a voice exchange between the computer program and a person, the electronic system comprising:

[a] a wearable electronic device that the person wears and that provides the voice of the computer program to the person at the SLP at a first location in empty space that is away from and proximate to the person;

[b] a memory that stores instructions; and

[c] a processor that executes the instructions to

[c1] receive a command from the person to move the SLP of the voice of the computer program from the first location in empty space to a second location in empty space that is proximate to the first location and away from and proximate to the person,

[c2] move the SLP of the voice of the computer program in response to the command from the first location in empty space to the second location in empty space, and

[c3] provide binaural sound to the person through the wearable electronic device during the voice exchange between the computer program and the person such that an origin of the voice of the computer program occurs at the SLP at the second location in empty space that is proximate to the first location and away from and proximate to the person.

17. The electronic system of claim 15, wherein the processor further executes the instructions to change an interaural time difference (ITD) and an interaural level difference (ILD) of the binaural sound in order to move the SLP of the voice of the computer program from the first location in empty space to the second location in empty space.

18. The electronic system of claim 15, wherein the processor further executes the instructions to change an interaural time difference (ITD) and an amplitude of the binaural sound in order to maintain the SLP fixed at the second location in empty space during the voice exchange while the person physically moves away from the second location in empty space and changes a head orientation with respect to the second location in empty space.

19. The electronic system of claim 15, 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 person, wherein the virtual objects do not actually exist near the person but are perceived to exist by the 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 13,998 words, calculated by Microsoft Word's word-count feature. This total excludes the cover page, signature block, and the parts of the Petitions exempted by 37 C.F.R. § 42.24(a)(1).

Dated: June 9, 2023

By: /Lisa K. Nguyen/

Lisa K. Nguyen (Reg. No. 58,018)

lisa.nguyen@allenoverly.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 9,282,196 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, 2023:

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@allenoverly.com
Allen & Overy LLP
550 High Street
Palo Alto, CA 94301
Telephone: 650.388.1724

*Counsel for Petitioner
Meta Platforms, Inc.*