

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re *Inter Partes* Review of:)
U.S. Patent No. 11,290,836)
Issued: March 29, 2022)
Application No.: 17/062,637)
Filing Date: Oct. 5, 2020)

For: **Providing Binaural Sound Behind an Image Being Displayed with an
Electronic Device**

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 11,290,836**

TABLE OF CONTENTS

I.	Introduction.....	1
II.	Identification of Challenges (37 C.F.R. §42.104(b))	2
III.	Background.....	2
	A. The '836 Patent	2
	B. POSAs	4
	C. Claim Construction.....	4
IV.	Ground 1: Claims 8-10 and 15-17 Are Unpatentable In View of Finnegan, Alone or in View of Begault.....	5
	A. Overview of Finnegan.....	5
	B. Overview of Begault	6
	C. Motivations to Modify and/or Combine	7
	D. Independent Claim 8	9
	1. 8[pre]: An Electronic Device.....	9
	2. 8[a]: Image Within A Near-Field Distance.....	9
	3. 8[b] and 8[d]: One or More Processors for Processing Sound for the Image, Including a DSP	11
	4. 8[c]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight.....	14
	5. 8[e]: DSP Processes Sound with Far-Field HRTFs While Electronic Device is at Near-Field Location.....	17
	E. Dependent Claims 9 and 10	19
	1. Claim 9: AR Image and Electronic Glasses.....	19
	2. Claim 10: VR Image and HMD	20
	F. Independent Claim 15	21

1.	15[pre]: An Electronic System.....	21
2.	15[a]: Image Within A Near-Field Distance.....	21
3.	15[b]: Processor for Processing Sound for the Image	21
4.	15[c]: HRTFs to Generate Binaural Sound.....	21
5.	15[d]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight.....	22
6.	15[e]: SLP Stays at Far-Field Distance While User’s Head Moves and Image Remains at Near-Field Distance	22
G.	Dependent Claims 16 and 17	24
1.	Claim 16: AR Display and Image.....	24
2.	Claim 17: VR Display and Image.....	25
V.	Ground 2(a): Claims 8-10 and 13-17 are Unpatentable Over Norris, Alone or in View of Begault and Algazi	25
A.	Overview of Norris (Ex.1007)	25
B.	Overview of Algazi (Ex.1008).....	26
C.	Motivations to Modify and/or Combine	27
D.	Independent Claim 8	28
1.	8[pre]: An Electronic Device	28
2.	8[a]: Image Within A Near-Field Distance.....	28
3.	8[b] and 8[d]: One or More Processors for Processing Sound for the Image, Including a DSP	29
4.	8[c]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight.....	31
5.	8[e]: DSP Processes Sound with Far-Field HRTFs While Electronic Device is at Near-Field Location.....	34
E.	Dependent Claims 9, 10, 13, and 14	36

1.	Claim 9: AR Image and Electronic Glasses.....	36
2.	Claim 10: VR Image and HMD	37
3.	Claim 13: AR Image of an IUA.....	37
4.	Claim 14: VR Image of an IUA.....	38
F.	Independent Claim 15	39
1.	15[pre]: An Electronic System.....	39
2.	15[a]: Image Within A Near-Field Distance.....	39
3.	15[b]: Processor for Processing Sound for the Image	39
4.	15[c]: HRTFs to Generate Binaural Sound.....	39
5.	15[d]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight.....	40
6.	15[e]: SLP Stays at Far-Field Distance While User’s Head Moves and Image Remains at Near-Field Distance	40
G.	Dependent Claims 16 and 17	42
1.	Claim 16: AR Display and Image.....	42
2.	Claim 17: VR Display and Image.....	42
VI.	Ground 2(b): Dependent Claims 11 and 12 are Unpatentable Over Norris, Alone or in View of Begault and Algazi, and Further in View of Sauk	43
A.	Overview of Sauk (Ex.1009).....	43
B.	Motivation to Combine Sauk	44
C.	Claim 11: Sensors Detect Rotation, and Binaural Sound Switches to Stereo	45
D.	Claim 12: Binaural Sound Switches to Stereo in Response to Change to Head Orientation Exceeding Threshold Azimuth Value.....	47

VII.	Ground 3: Claims 8-10 and 15-17 are Unpatentable Over Shivappa, Alone or in view of Zhou	48
A.	Overview of Shivappa (Ex.1010).....	48
B.	Overview of Zhou (Ex.1011)	50
C.	Motivation to Combine Shivappa with Zhou.....	51
D.	Claim 8	52
1.	8[pre]: An Electronic Device	52
2.	8[a]: Image Within A Near-Field Distance.....	52
3.	8[b] and 8[d]: One or More Processors for Processing Sound for the Image, Including a DSP	53
4.	8[c]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight.....	55
5.	8[e]: DSP Processes Sound with Far-Field HRTFs While Electronic Device is at Near-Field Location.....	58
E.	Dependent Claims 9 and 10	59
1.	Claim 9: AR Image and Electronic Glasses.....	59
2.	Claim 10: VR Image and HMD	59
F.	Independent Claim 15	60
1.	15[pre]: An Electronic System.....	60
2.	15[a]: Image Within A Near-Field Distance.....	60
3.	15[b]: Processor for Processing Sound for the Image	60
4.	15[c]: HRTFs to Generate Binaural Sound.....	60
5.	15[d]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight.....	61
6.	15[e]: SLP Stays at Far-Field Distance While User's Head Moves and Image Remains at Near-Field Distance	61

G.	Dependent Claims 16 and 17	63
1.	Claim 16: AR Display and Image.....	63
2.	Claim 17: VR Display and Image.....	64
VIII.	Secondary Considerations	64
IX.	The Board Should Reach the Merits of This Petition	64
X.	Mandatory Notices under 37 C.F.R. §42.8.....	65
A.	Real Parties-in-Interest.....	65
B.	Related Matters.....	65
C.	Grounds for Standing	65
D.	Lead and Backup Counsel and Service Information.....	65
E.	IPR Fee	67
XI.	Conclusion	67

TABLE OF ABBREVIATIONS

Term	Abbreviation
wearable electronic device	WED
head mounted display	HMD
portable electronic device	PED
handheld portable electronic device	HPED
three-dimensional	3D
sound localization point	SLP
digital signal processor	DSP
head-related transfer function	HRTF
virtual reality	VR
augmented reality	AR
intelligent user assistant	IUA

Exhibit List

Ex.	Description
1001	U.S. Patent No. 11,290,836 (“’836 patent”)
1002	U.S. Patent No. 11,290,836 File History (“’836 FH”)
1003	Declaration of Dr. Agnieszka Roginska in Support of Petition for <i>Inter Partes</i> Review of the ’836 Patent
1004	Curriculum Vitae of Dr. Agnieszka Roginska
1005	Daniel J. Finnegan et al., <i>Compensating for Distance Compression in Audiovisual Virtual Environments Using Incongruence</i> , Chi’ 16: Proc. of the 2016 CHI Conf. on Hum. Factors in Computing Sys. (2016) (“Finnegan”)
1006	Durand R. Begault, <i>3-D Sound for Virtual Reality and Multimedia</i> , NASA/TM—2000-209606 (Aug. 2000) (“Begault”)
1007	U.S. Patent No. 9,226,090 (“Norris”)
1008	V.R. Algazi et al., <i>The CIPIC HRTF Database</i> , IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (2001) (“Algazi”)
1009	U.S. Patent No 7,876,903 (“Sauk”)
1010	U.S. Patent No. 10,979,843 (“Shivappa”)
1011	Zhiying Zhou et al., <i>An Experimental Study on the Role of Software Synthesized 3D Sound in Augmented Reality Environments</i> , 16 Interacting with Computs. 989 (2004) (“Zhou”)
1012	Declaration of Ingrid Hsieh-Yee in Support of Petition for <i>Inter Partes</i> Review of the ’836 Patent (“Hsieh-Yee”)
1013	Curriculum Vitae of Ingrid Hsieh-Yee
1014	U.S. Patent No. 4,774,515 (“Gehring”)
1015	U.S. Patent No. 2,955,156 (“Heilig”)

Ex.	Description
1016	U.S. Patent No. 9,041,622 (“McCulloch”)
1017	U.S. Patent No. 10,509,466 (“Miller”)
1018	U.S. Patent No. 9,767,618 (“Najaf-Zadeh”)
1019	U.S. Patent No. 9,232,336 (“Nishio”)
1020	U.S. Patent No. 9,578,440 (“Otto”)
1021	U.S. Patent No. 1,624,486 (“Fletcher”)
1022	U.S. Patent No. 1,855,149 (“Jones”)
1023	U.S. Patent No. 8,873,761 (“Fukui”)
1024	U.S. Patent No. 9,787,846 (“Srinivasan”)
1025	U.S. Patent Application Publication No. US 2012/0183161 A1
1026	U.S. Patent No. 9,530,426 (“Wright”)
1027	U.S. Patent Application Publication No. US 2017/0115488 A1 (“Ambrus”)
1028	U.S. Patent No. 7,710,654 (“Ashkenazi”)
1029	Matthias Geier et al., <i>The SoundScape Renderer: A Unified Spatial Audio Reproduction Framework for Arbitrary Rendering Methods</i> (Convention Paper 7330), Audio Eng’g Soc’y 124th Convention (2008)
1030	V. Ralph Algazi & Richard O. Duda, <i>Headphone-Based Spatial Sound</i> , 28 IEEE Signal Processing Mag. 33 (2011) (“Algazi 2011”)
1031	S.S. Fisher et al., <i>Virtual Interface Environment Workstations</i> , Procs. of the Hum. Factors Soc’y, 32nd Annual Meeting 91 (1988)
1032	Wolfgang Fohl et al., <i>A System-On-Chip Platform for HRTF-Based Realtime Spatial Audio Rendering With an Improved Realtime Filter Interpolation</i> , 4 Int’l J. Advances Intelligent Sys. 309 (2011)
1033	Jason Jerald, <i>The VR Book – Perception and Interaction Design for Virtual Reality</i> (M. Tamer Özsu ed., 2015) (“The VR Book”)

Ex.	Description
1034	Masayuki Morimoto & Yoichi Ando, <i>On the Simulation of Sound Localization</i> , 1 J. of Acoustical Soc’y Japan (E) 167 (1980)
1035	C. Phillip Brown & Richard O. Duda, <i>A Structural Model for Binaural Sound Synthesis</i> , 6 IEEE Transactions on Speech and Audio Processing 476 (1998)
1036	Douglas S. Brungart & William M. Rabinowitz, <i>Auditory Localization of Nearby Sources. Head-related transfer Functions</i> , 106 J. Acoustical Soc’y of Am. 1465 (1999)
1037	Kai-Uwe Doerr et al., <i>Evaluation of a Low-Cost 3D Sound System for Immersive Virtual Reality Training Systems</i> , 13 IEEE Transactions on Visualization and Comput. Graphics 204 (2007)
1038	<i>Virtual Reality</i> , Cryptologic Quarterly 21 (Dec. 1, 2011)
1039	Stephan Paul, <i>Binaural Recording Technology: A Historical Review and Possible Future Developments</i> , 95 Acta Acustica United with Acustica 767 (2009)
1040	David Romblom & Bryan Cook, <i>Near-Field Compensation for HRTF Processing</i> (Convention Paper 7611), Audio Eng’g Soc’y 125th Convention (2008)
1041	Amy Turner et al., <i>Can the Perception of Depth in Stereoscopic Images Be Influenced by 3D Sound?</i> , 7863 Stereoscopic Displays & Applications (2011)
1042	Elizabeth M. Wenzel et al., <i>Development of a Three-Dimensional Auditory Display System</i> , 20 SIGCHI Bull. 52 (1988)
1043	Elizabeth M. Wenzel et al., <i>Localization with Non-Individualized Virtual Acoustic Display Cues</i> , Chi’ 91: Proceedings of the SIGCHI Conf. on Hum. Factors in Comput. Sys. 351 (1991)
1044	Frederic L. Wightman & Doris J. Kistler, <i>Headphone Simulation of Free-Field Listening. II: Psychophysical Validation</i> , 85 J. Acoustical Soc’y of Am. 868 (1989)
1045	Steven M. Kaplan, Wiley Electrical and Electronics Engineering Dictionary 335 (2004)

Ex.	Description
1046	Dmitry N. Zotkin et al., <i>Customizable Auditory Displays</i> , Proc. of the Inte'l Conf. on Auditory Display (July 2002)
1047	<i>Asserted U.S. Patent No. 11,290,836 "Providing Binaural Sound Behind an Image Being Displayed with an Electronic Device</i> , Plaintiff's Infringement Contentions Exhibit 27, <i>Eight KHZ, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. Mar. 3, 2023)
1048	<i>Asserted U.S. Patent No. 11,290,836 "Providing Binaural Sound Behind an Image Being Displayed with an Electronic Device</i> , Plaintiff's Infringement Contentions Exhibit 18, <i>Eight KHZ, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. Mar. 3, 2022)
1049	<i>Asserted U.S. Patent No. 11,290,836 "Providing Binaural Sound Behind an Image Being Displayed with an Electronic Device</i> , Plaintiff's Infringement Contentions Exhibit 36, <i>Eight KHZ, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. Mar. 3, 2023)
1050	Rasmus Konig Sorenson, <i>Review: Oculus Rift DK2 – Next-Level Virtual Reality</i> , flatpanelshd (Aug. 13, 2014), https://www.flatpanelshd.com/review.php?subaction=showfull&id=1407909173
1051	U.S. Patent No. 9,392,212 ("Ross")
1052	Mark Wherry, <i>Audio Add-ons For Macbook Pro, Apple Notes</i> , SoundonSound (Mar. 2007), https://www.soundonsound.com/techniques/audio-add-ons-macbook-pro
1053	Defendants Opening Claim Construction Brief, <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. Apr. 12, 2023), ECF No. 83
1054	Plaintiff's Claim Construction Reply Brief, <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. May 3, 2023), ECF No. 92
1055	Plaintiff's Claim Construction Reply Brief, Exhibit 7, <i>Eight kHz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. May 3, 2023), ECF No. 92-7

Ex.	Description
1056	U.S. Patent No. 5,173,944 (“Begault ’944”)
1057	U.S. Patent No. 5,438, 623 (“Begault ’623”)
1058	Deposition transcript of Durand Begault, Ph.D., <i>Eight Khz, LLC v. Meta Platforms, Inc.</i> , No. 6:22-cv-00575 (W.D. Tex. May 22, 2023)
1059	Areti Andreopoulou & Agnieszka Roginska, <i>Towards the Creation of a Standardized HRTF Repository</i> , Audio Eng’g Soc’y Convention Paper 8571, Oct. 20-23, 2011
1060	Chan Jun Chun et al., <i>Deep Neural Network Based HRTF Personalization Using Anthropometric Measurements</i> , Audio Eng’g Soc’y Convention Paper 9860, Oct. 18-21, 2017
1061	Karen Collins, <i>Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design</i> (2008)
1062	Sungmok Hwang & Youngjin Park, <i>HRIR Customization in the Median Plane Via Principal Components Analysis</i> , AES 31st Int’l Conf., London, UK (2007)
1063	Donald H. Merson & John N. Bowers, <i>Absolute and Relative Cues for the Auditory Perception of Egocentric Distance</i> , 8 Perception 311-22 (1979)
1064	Matthew Fam Kai Liang, <i>The AweSOME 5: High-End Sound Cards Review</i> (Oct. 11, 1998), https://assets.hardwarezone.com/2009/reviews/sound/roundup/soundcards.htm
1065	Matthew Fam Kai Liang, <i>The AweSOME: 5(Part 2): Explore: Technologies Behind ...</i> (Sept. 27, 1998), https://assets.hardwarezone.com/2009/reviews/sound/roundup/sound2.htm
1066	Pavel Zahorik, <i>Assessing Auditory Distance Perception Using Virtual Acoustics</i> , 111 J. Acoustical Soc’y. Am. 1832-46 (2002)
1067	Korean Patent Application Publication No. KR 10-2011-0022252, including English Translation and Certification (“Moon”)

Ex.	Description
1068	<i>Oculus: Complete Guide — History, Products, Founding, and More</i> , History-Computer.com (updated May 13, 2013), https://history-computer.com/oculus-history/
1069	Pat Lawler, <i>What’s So Special About the Digital Signal Processor?</i> , Qualcomm: OnQ Blog (Dec. 5, 2013), https://www.qualcomm.com/news/onq/2013/12/whats-so-special-about-digital-signal-processor (“Qualcomm blog”)
1070	John D. Sutter, <i>Microsoft’s Talking Paper Clip is Back</i> , CNN.com (Apr. 29, 2011), http://www.cnn.com/2011/TECH/web/04/29/microsoft.paper.clip/index.html#:~:text=Well%2C%20like%20it%20or%20not,2%3A%20Clippy's%20Second%20Chance.%22
1071	<i>Super Mario 64 Characters</i> , Mario.fandom.com, https://mario.fandom.com/wiki/Category:Characters_in_Super_Mario_Galaxy (last visited June 5, 2023)
1072	Mona Lalwani, <i>Surrounded By Sound: How 3D Audio Hacks Your Brain</i> , Theverge.com (Feb. 12, 2015), https://www.theverge.com/2015/2/12/8021733/3d-audio-3dio-binaural-immersive-vr-sound-times-square-new-york
1073	<i>Aureal Semiconductor</i> , Dosdays.co.uk, https://dosdays.co.uk/topics/Manufacturers/aureal.php (last visited June 5, 2013)
1074	Mike Andrawes, <i>Audio Report – Comdex ’99</i> , Anandtech.com (Nov. 30, 1999), https://www.anandtech.com/show/417
1075	Elizabeth M. Wenzel et al., <i>A Virtual Display System for Conveying Three-Dimensional Acoustic Information</i> , 32nd Proc. of the Hum. Factors Soc’y Annual Meeting 86 (1988) (“Wenzel 1988b”)
1076	Jean-Marc Jot et al., <i>Digital Signal Processing Issues in the Context of Binaural and Transaural Stereophony</i> , 98th Proc. of the Convention of the Audio Eng’g Soc’y (1995) (“Jot”)
1077	Henrik Moller et al., <i>Head-Related Transfer Functions of Human Subjects</i> , 43 J. Audio Eng’g Soc’y 300 (1995) (“Moller”)

Ex.	Description
1078	David A. Burgess, <i>Techniques for Low Cost Spatial Audio</i> , ACM Symp. on User Interface Software and Tech. 53 (1992) (“Burgess”)
1079	Navarun Gupta et al., <i>HRTF database at FIU DSP Lab</i> , IEEE Int’l Conf on Acoustics, Speech, and Signal Processing (2010) (“Gupta”)
1080	John C. Middlebrooks, <i>Individual Differences in External-Ear Transfer Functions Reduced by Scaling in Frequency</i> , 106 J. Acoustical Soc’y Am. 1480 (1999) (“Middlebrooks”)
1081	Elizabeth M. Wenzel, <i>Localization in Virtual Acoustic Displays</i> , 1 Presence Teleoperators & Virtual Env’ts 80 (1992) (“Wenzel 1992”)
1082	U.S. Patent No. 9,037,468 (“Osman”)
1083	Durant R. Begault et al., <i>Design and Verification of HeadZap, A Semi-Automated HRIR Measurement System</i> , 120th Proc. of the Audio Eng’g Soc’y Convention (2006) (“HeadZap”)
1084	Bill Gardner & Keith Martin, <i>HRTF Measurements of a KEMAR Dummy-Head Microphone</i> , MIT Media Lab Perceptual Computing – Tech. Rep. 280 (1994)
1085	David M. Howard & Jamie A.S. Angus, <i>Acoustics and Psychoacoustics</i> (3d ed. 2006)

I. Introduction

Meta Platforms, Inc., Meta Platforms Technologies, LLC, and Twisted Pixel Games, LLC (collectively, “Petitioners”) request IPR of claims 8-17 of U.S. Patent No. 11,290,836 (“’836 patent,” Ex.1001).

The ’836 patent relates to systems that display images and process related sounds with digital signal processors (“DSPs”) and head-related transfer functions (“HRTFs”), including “far-field” HRTFs, so sounds are positioned behind, but perceived to originate at, “near-field” images. But DSPs have been used for sound processing, including in virtual environments and with HRTFs, for decades. And before the ’836 patent’s earliest alleged priority date in May 2017, it was known that visual and audio stimuli could be positioned in virtual environments with sound behind images, for example, to compensate for distance compression, “a phenomenon where humans tend to underestimate the distance between themselves and target objects” in “virtual worlds rendered through head mounted displays” (“HMDs”). Finnegan 1, 11. Furthermore, “far-field” HRTFs—i.e., HRTFs measured at 1 m or more—were widely used before May 2017, as multiple publicly-available HRTF databases were built using such HRTFs. Algazi, 99. Accordingly, the challenged claims’ allegedly inventive configuration of near-field images with far-field sound was taught and/or rendered obvious by multiple prior art references disclosing HMDs that (i) display virtual images on their lenses (i.e., at near-field

distances within 1 m from the user) and (ii) render 3D sound with commonplace far-field HRTFs.

Petitioners ask the Board to institute review and find the challenged claims unpatentable.

II. Identification of Challenges (37 C.F.R. §42.104(b))

Ground 1: Claims 8-10 and 15-17 are unpatentable under § 103 over Finnegan (Ex.1005), alone or in view of Begault (Ex.1006).

Ground 2: Challenged claims 8-17 are unpatentable under § 103 over Norris (Ex.1007), alone or in view of other prior art. Specifically:

- Ground 2(a): Claims 8-10 and 13-17 would have been obvious over Norris, alone or in view of Begault and Algazi (Ex.1008); and
- Ground 2(b): Claims 11 and 12 would have been obvious over Norris, alone or in view of Begault and Algazi, and further in view of Sauk (Ex.1009).

Ground 3: Challenged claims 8-10 and 15-17 are unpatentable under § 103 over Shivappa (Ex.1010), alone or in view of Zhou (Ex.1011).

III. Background

A. The '836 Patent

The '836 patent purports to improve user experience in virtual environments by processing “sound for [an] image into binaural sound” so the sound “has a sound localization point (SLP)” “behind [an] image while the image is located a near-field

distance from a head of the user.” 836 patent Title, Abstract, 2:65-3:4.

According to the '836 patent, “[o]ne problem with electronically generated binaural sound or three-dimensional (3D) sound rendering is that listeners have a difficult time determining a location where the sound originates when this location is close to the listener,” for example, “less than one meter from the person (considered to be ‘near-field’).” *Id.* 2:11-20. However, “a listener may have more success in external localization when the binaural sound is processed to originate from about one meter or more from the person (considered to be ‘far-field’).” *Id.* 2:20-23; *see also id.* 27:23-25 (defining “near-field” as “less than one meter” and “far-field” as “greater than or equal to one meter”). The '836 patent further provides that because “modeling near-field virtual auditory space with HRTFs can be quite complex,” when a device attempts to process near-field sound, “[t]he listener may perceive the sound as originating from the wrong location.” *Id.* 2:24-38.

The '836 patent aims to solve these problems by “trick[ing] the listener into believing that the sound originates from [a] sound source when in fact the sound is convolved to originate from a location that is behind the sound source.” *Id.* 4:65-5:11, 5:64-67, 6:26-48. This can be done, the '836 patent states, by “process[ing] sound with HRTFs having coordinates of a far-field location even though the location of the source of the sound (e.g., an image being displayed) is a near-field distance from the listener.” *Id.* 6:49-67; Roginska (Ex.1003) ¶¶26-44.

B. POSAs

A person of ordinary skill in the art (“POSA”) as of the ’836 patent’s earliest alleged priority date (May 24, 2017) would have had a Bachelor’s degree in computer science/engineering, or a related field, with coursework in computer programming, digital signal processing, spatial audio, or virtual/augmented reality, and at least two years of industry experience, or academic research experience, in virtual/augmented reality, audio engineering, or an equivalent. Additional education can compensate for less experience, and vice-versa. Roginska ¶¶45-51.

C. Claim Construction

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

Additionally, Patentee contends in district court that limitations reciting “processors that process sound...” and “a digital signal processor (DSP) that processes sound...” (e.g., limitations 8[b] and 8[d]¹) should be construed by their plain meanings and “Section 112(f) does not apply.” Ex.1054, 24-38. Solely for

¹ The “Claims Appendix” attached to this Petition reproduces challenged claims with bracketed cross-referencing identifiers.

this IPR, Petitioner adopts Patentee’s plain-meaning positions for such limitations.²

IV. Ground 1: Claims 8-10 and 15-17 Are Unpatentable In View of Finnegan, Alone or in View of Begault

Finnegan and Begault teach or suggest every limitation of challenged claims 8-10 and 15-17. Finnegan disclosed processing sound and images incongruently, with sounds positioned behind associated images, to improve user perception. And Begault confirms that POSAs reading Finnegan would have known that a Finnegan-type system could and would use a DSP and HRTFs for sound processing.

A. Overview of Finnegan

Finnegan, published in May 2016, is § 102(a) prior art and was not considered during prosecution. Ex.1012 ¶¶66-76.

Finnegan described techniques for compensating for distance compression in virtual environments using incongruent stimuli—images and sounds presented at distinct distances. As Finnegan explained, “a key requirement for a sense of presence in Virtual Environments (VEs) is for a user to perceive space as naturally as possible.” Finnegan at Abstract. It was known that “[c]ompression”—“a phenomenon where humans tend to underestimate the distance between themselves and target objects”—occurs in virtual environments. *Id.* Finnegan thus described

² Petitioners reserve the right to argue alternative constructions in other proceedings, including that claims are indefinite—and do so in district court.

an experiment testing whether rendering audio at “an incongruent position further from the observer than the visual stimulus (IV)”—i.e., positioning sound behind images—“result[ed] in more accurate distance perception (DV) compared to conditions where both stimuli are at the same position (congruent conditions).” *Id.* 6.

Finnegan’s experiments used a VR HMD, specifically an Oculus Rift Development Kit 2, audio rendered using a custom plugin, and binaural headphones, to present users with perceived visual stimuli and congruent and incongruent sound. *Id.* 6-7. Finnegan concluded that “rendering the auditory and visual components of objects incongruently to one another could improve distance perception in a virtual environment.” *Id.* 11; Roginska ¶¶123-132.

B. Overview of Begault

Begault, published in 2000, is § 102(a) prior art and was not considered during prosecution. Hsieh-Yee ¶¶30-65; Begault Dep. (Ex.1058) 222:21-223:17.

Begault described processes for producing “3-D sound,” “spatialized sound,” and “binaural audio,” explaining these terms refer to techniques that use digital filters to position sounds in “virtual space.” Begault, x. Begault taught that sound should have an “interactive, spatial character” because one can “work more effectively within a VR environment if actions [a]re accompanied by appropriate sounds that seemingly emit from their proper locations.” *Id.* 12-13.

Begault described in detail HRTFs and “localization with HRTF cues,” noting the “inclusion of the binaural HRTF is fundamental for a 3-D sound system.” *Id.* 49, 99, Chpt. 2. Begault disclosed using digital signal processors (“DSPs”) for processing binaural sound, explaining that “[t]he 3-D audio component of an integrated audio subsystem ... can contain anywhere from one to as many DSPs as can be accommodated,” and that “[w]ithin a virtual reality or multimedia application, 3-D sound processing is almost always best accomplished digitally, using audio DSP chips.” Begault 100-101; Roginska ¶¶ 133-145.

C. Motivations to Modify and/or Combine

POSAs would have known and been motivated to modify Finnegan’s system to use a DSP for sound processing and/or to implement Finnegan’s system with a DSP as taught by Begault. Finnegan disclosed that “[i]n binaural environments,” like its tested environment, “digital signal processing provides abilities to alter the intensity, frequency and reverberation present in the audio signal,” Finnegan 3, and Begault taught that 3D sound processing is “almost always best accomplished” with DSPs. Begault 100-101.

Notably, Dr. Durand Begault—the Begault reference author and Patentee’s expert in the parallel district court action—conceded that he disclosed in his own patents, from before 2014, “generating binaural sound through digital signal processors processing HRTFs.” Begault Dep. 135:6-25; *see also id.* 42:3-8, 251:20-

253:5 (NASA’s Convolvotron, implemented by 1988, “was using [] DSPs for sound processing”), Begault, 165-68 (describing “Convolvotron” and its “DSPs”); Exs.1056-57, U.S. Pat. Nos. 5,173,944 and 5,438,623.

Additionally, Finnegan taught that its virtual environment, including its visual and audio stimuli, was simulated with a “Macbook Pro” (Finnegan, 6), and, as Dr. Begault admitted, POSAs knew “you could use DSPs with Macs to process spatialized audio.” Begault Dep. 253:16-19, 254:23-255:2; Roginska¶¶193; Ex.1052 (discussing external DSP compatible with Macs).

Moreover, the Begault reference is just one of multiple prior art references demonstrating that POSAs would have known to implement a system like Finnegan’s with a DSP for processing sound. Roginska¶¶119-121 (collecting references). DSPs are and were used for such processing because they are fast, excel at signal processing tasks like HRTF processing, are reconfigurable and flexible, and are appropriately sized for use in systems like Finnegan’s. *Id.*¶¶188-191.

Accordingly, POSAs would have known and been motivated to implement, and would have reasonably expected to succeed in implementing, Finnegan’s system with a DSP for processing sound, as Begault prescribes. To POSAs, this would have been obvious to try and well within common sense, and required no more than combining known elements to improve similar devices in known ways. *Id.*¶¶185-193.

D. Independent Claim 8

1. 8[pre]: An Electronic Device

Finnegan taught claim 8's preamble, which recites: "An electronic device, comprising." Finnegan's experiments used an "Oculus Rift Development Kit 2 HMD," a "Macbook Pro," and a pair of "Binaural Headphones." Finnegan 2, 6; Roginska ¶¶194-95.

2. 8[a]: Image Within A Near-Field Distance

Finnegan disclosed limitation 8[a], which recites: "a display that displays an image that occurs within a near-field distance from a head of a user."

Finnegan used a Rift DK2 VR HMD that displayed images to users. Finnegan 2, 6, Fig. 1 ("example screenshot of what the participants saw inside the headset"), 7 (stimuli viewed "through the HMD viewport"). And, as the following images of the Rift DK2 show, the lenses in that device were well within 1 m from a user's head:



Ex.1050; The VR Book (Ex.1033), Fig. 3.3 (another image of an Oculus Rift); Roginska ¶199. Accordingly, Finnegan taught limitation 8[a]’s display of an image at a “near-field distance from a head of a user.” Roginska ¶¶196-201; *see also* ’836 patent, cl. 8 (requiring (i) the “display displays the image to the user at *the* near-field distance” and (ii) “the electronic device is located at *the* near-field distance”—i.e., the electronic device and the displayed image are co-located); Ex.1047, 4 (Patentee contends that an image on the lenses of a Quest HMD satisfies limitation 8[a] because the lenses are the HMD’s “display” and the image occurs “within a few inches” of the user).

Alternatively, if the pertinent “near-field distance from a head of a user” for limitation 8[a] is the distance to the location of a *perceived* image, Finnegan taught that too. In Finnegan’s experiments, users wore HMDs that presented *perceived* visual stimuli ranging between 0.5 to 9.5 meters in front of them while associated sounds were presented at congruent and incongruent locations. *Id.* 7. Any distance to a visual stimulus greater than or equal to 0.5 m and less than 1 m is a “near-field” distance per the ’836 patent. ’836 patent, 27:23-25; Roginska ¶¶202-203.³

³ *See also* Roginska ¶225 (Finnegan’s teachings could be applied to hardware other than HMDs, e.g., handheld portable electronic devices, which could and would

3. 8[b] and 8[d]: One or More Processors for Processing Sound for the Image, Including a DSP

Limitations 8[b] and 8[d] recite, “one or more processors that process sound for the image,” and “wherein the one or more processors include a digital signal processor (DSP),” respectively.

Finnegan taught that its virtual environment, including its visual and audio stimuli, was simulated with a “Macbook Pro” with an “Intel i7 processor.” Finnegan 6. Finnegan further disclosed that “[i]n binaural environments,” like the environment it tested, “digital signal processing provides abilities to alter the intensity, frequency and reverberation present in the audio signal.” *Id.* 3.

Based on these disclosures, it would have been obvious (i) to use a processor to process sound as limitation 8[b] requires, e.g., as Finnegan did with its Intel i7 processor, and (ii) that Finnegan’s “digital signal processing” could and would be implemented with a “digital signal processor (DSP),” for example, by moving sound processing from Finnegan’s Intel i7 processor to a DSP or adding an external DSP to Finnegan’s system to handle sound processing for images. §IV.C. Either of these would have rendered limitations 8[b] and 8[d] obvious. Roginska¶¶205-207. And as discussed, POSAs would have known and been motivated to use a DSP in

display images at near-field distances, with SLPs for binaural sound positioned behind such images (*see* §IV.D.4)).

Finnegan's system for sound processing for images because DSPs were routinely used for such processing before May 2017, including with Mac computers and in virtual environments, and were known to be advantageous. §IV.C. Thus, limitations 8[b] and 8[d] would have been obvious based on Finnegan alone. Roginska ¶¶204-209, 226-232.

In fact, Patentee's expert, Dr. Begault, effectively conceded the obviousness of these limitations:

Q. Let's look at claim 8 of the '836 Patent, please, Dr. Begault. Let me know when you're there.

A. I am there.

Q. Okay. I'd like you to read the second limitation of claim 8. It's a relatively lengthy one. And then my question for you, which I'll give you now so you know, is does that limitation recite two of the processor terms that you addressed in your declaration?⁴

A. Okay. Stand by. Yes.

Q. The first is the term that begins, One or more processors that process sound for the image to generate binaural sound that has a sound localization point, SLP, with a coordinate location that occurs at a far field distance located behind the image and along a line of sight from the user, right?

A. I see that.

⁴ See Ex.1055, ¶15.

Q. Okay. And the second processor term in the second limitation of claim 8 that you addressed is wherein the one or more processors include a digital signal processor, DSP, that processes the sound for the image with far-field head-related transfer functions, HRTFs, right?

A. Yes, I see that.

Q. Do you believe that before 2014, persons of skill in the art would have known how to implement systems performing the functions recited in those terms by resorting to known audio DSP structures within the art and common industry programming techniques and software libraries?

[A]: *Yes*.

Begault Dep. 200:3-201:9 (emphasis added and objection omitted).

Additionally and alternatively, limitations 8[b] and 8[d] would have been obvious over Finnegan in view of the Begault reference—specifically, in view of Begault’s descriptions of DSPs and prescription to use DSPs for processing 3D sound in virtual environments. Begault, 100-101. POSAs would have known, including based on Begault, that DSPs were advantageous for such sound processing and thus would have been motivated to implement a Finnegan-type system with

DSPs. §IV.C; Roginska ¶¶234-236. The sound processing DSP in such a combination would have rendered limitations 8[b] and 8[d] obvious. *Id.* ¶¶233-236.⁵

4. 8[c]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight

Limitation 8[c] requires that the sound processing by the claimed one or more processors, e.g., the DSP, “generate[s] binaural sound” that “has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance.”

Limitation 8[c] would have been obvious over Finnegan, alone or in view of Begault. First, POSAs would have known and understood that Finnegan’s sound processing generated (or at a minimum obviously generated) “binaural sound.”

⁵ Even if limitations 8[b] and 8[d] were construed as means-plus-function terms, they would have been obvious over Begault because Begault disclosed “convolution” processes using HRTFs (*e.g.*, Begault 105-09), and because Dr. Begault has opined that “[c]onvolution is the main algorithm required to perform the claimed processes,” while conceding that convolution was “well-known.” Ex.1055, ¶16; Begault Dep. 59:4-16 (admitting “many references” “in the 1990s” “discuss[ed] using convolution processes to process head related transfer functions with digital signal processors for binaural audio”).

Finnegan produced “binaural audio” playing through “Binaural Headphones.” Finnegan 2, 6. Moreover, Finnegan disclosed that its system “integrate[d] the SoundScape Renderer (SSR), a GPL licensed software implementation for binaural audio [1].” *Id.* 6 (citing “[1],” “Jens Ahrens, Matthias Geier, and Sascha Spors. 2008. The Sound-Scape Renderer: A unified spatial audio reproduction framework for arbitrary rendering methods” (“Ahrens”) (Ex.1029)).

Second, any incongruent binaural sound that Finnegan positioned at distances of 1 m or greater—which would occur in numerous instances in Finnegan’s tests ranging from 0.5 m to 9.5 m—would be at locations where a listener localizes sound (SLPs with coordinate locations) at “far-field” distances. *Id.* 7. All SLPs that are at one meter or greater are far-field SLPs. Roginska ¶¶210-14.

Third, the sounds Finnegan positioned at “far-field” distances were “behind” images that were “concurrently” displayed on a Rift HMD’s lenses, well within 1 m (i.e., at “near-field” distances) from the user for reasons discussed for limitation 8[a]. §IV.D.2; Finnegan, 6 (“visual cube” and “noise burst” stimuli presented “concurrently”); Roginska ¶¶215.

Fourth, POSAs would have known that Finnegan’s images and sounds were presented along a user’s line-of-sight. POSAs would have understood that to test whether “incongruent” sounds would be “integrated with the[ir] visual sensory signal[s],” which was Finnegan’s aim, sounds would be presented both behind visual

stimuli *and* along an observer’s line-of-sight. Finnegan 3. In fact, Finnegan required users to use a “chin rest” to secure their heads’ positions “before beginning the experimental conditions,” which would have secured users’ lines-of-sight so they were looking straight at Finnegan’s visual and audio stimuli. *Id.* 7; *id.* 6 (stimuli “constrained to the frontal view”); Roginska ¶¶216-17.

Additionally and alternatively, if the *perceived* distance to an image is the pertinent “near-field” distance for limitation 8[c]’s displayed image, Finnegan disclosed far-field SLPs located behind such images as well. Finnegan rendered audio “at an incongruent position further from the observer than [its] visual stimulus.” Finnegan 6; *id.* (“If we place the audio component at a greater distance from the observer,” then “the auditory sensory signal should be integrated with the visual sensory signal to produce a combined distance estimate that is closer to the intended distance.”); Roginska ¶¶218. And Finnegan explained that, in testing incongruent audio for visual stimuli presented over a range of locations from “0.5 to 9.5 meters in front of the observer,” audio offsets were calculated using an “incongruent positioning function.” *Id.* 7. That function—which “systematically position[ed] the audio component of a virtual object incongruently to its visual component”—is Finnegan’s Equation 2:

$$\phi = \left(\frac{\hat{y}}{k} \right)^{\frac{1}{\alpha}} \quad (2)$$

Id. 5, 7.

Application of this function to Finnegan’s testing parameters demonstrates that Finnegan tested “near-field” images while concurrently presenting incongruent audio at “far-field” distances. Roginska ¶¶219-22 (table of Equation 2 results demonstrates that for multiple distances to images between 0.5 m and 1 m, Finnegan used incongruent sounds at far-field distances). For example, when Finnegan’s Equation 2 is applied to a 0.75 m distance to a visual stimulus, it requires a distance to an incongruent audio source of 1.098 m, i.e., a far-field distance. *Id.* ¶¶223-225. And for reasons discussed above, Finnegan’s incongruent near-field visual and far-field audio stimuli were along a user’s line-of-sight.

5. 8[e]: DSP Processes Sound with Far-Field HRTFs While Electronic Device is at Near-Field Location

Limitation 8[e], which requires that the DSP in limitation 8[d] “processes the sound for the image with far-field head-related transfer functions (HRTFs) while the electronic device is located at the near-field distance from the head of the user,” would have been obvious over Finnegan, alone or in view of Begault.

First, as discussed, Finnegan, alone or in view of Begault, rendered obvious using a DSP to process sound for images, including to generate binaural sound. §§IV.D.3-IV.D.4.

Second, POSAs would have known and understood that Finnegan used “far-field head-related transfer functions (HRTFs)” to generate “binaural audio” played

through “Binaural Headphones.” Finnegan 2, 6. The “SoundScape Renderer (SSR)” “for binaural audio” integrated into Finnegan’s system “for spatial audio” (*id.* 6) “use[d] Head Related Transfer Functions (HRTFs) to reproduce [a] soundfield at the listeners ears.” Ahrens at 3; Roginska¶¶237-240. POSAs thus would have understood that the HRTFs used by this SoundScape Renderer and Finnegan were (or at a minimum obviously included) far-field HRTFs—the sounds Finnegan generated for users included far-field binaural sounds. §IV.D.4 (Finnegan tested incongruently-positioned far-field audio behind visual stimuli); Roginska¶241.

Third, the Rift DK2 HMD that Finnegan used would have been worn on, and located a near-field distance from, users’ heads, including while sounds for displayed images were processed as claimed. Roginska¶242; §IV.D.2.

Additionally and alternatively, limitation 8[e] would have been obvious over Finnegan in view of Begault. Begault prescribed using DSPs to process HRTFs to spatialize sound, further explaining that “[t]he use of HRTF spectral shaping is typically featured as the key component of a 3-D sound system.” *Id.* Chpt. 2, 41, 99-101; *see also id.* 118-19 (discussing “KEMAR” database “HRTFs,” which “have been used many times in both 3-D sound applications and localization studies”), 205 (KEMAR HRTFs measured at “1.4 meters”). Thus, even if Finnegan alone did not render limitation 8[e] obvious (it did), limitation 8[e] would have been obvious over Finnegan and Begault. Roginska¶¶243-46.

E. Dependent Claims 9 and 10

Claims 9 and 10 are unpatentable for the same reasons as claim 8 and as further explained below.

1. Claim 9: AR Image and Electronic Glasses

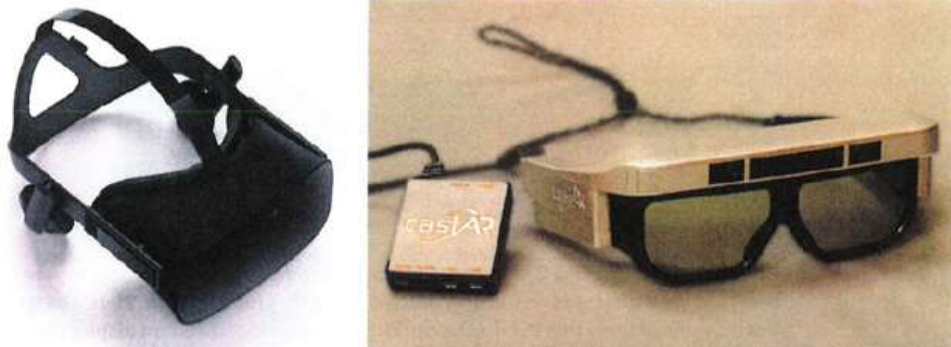
Claim 9 recites: “The electronic device of claim 8, wherein the image is an augmented reality (AR) image, and the display is included with electronic glasses worn on the head of the user.”

Claim 9 would have been obvious over Finnegan because POSAs would have known that Finnegan’s disclosures apply to displaying augmented reality (“AR”) images, including with “electronic glasses,” which are worn on users’ heads. Roginska ¶¶250-252. In describing the motivation for its experiments, Finnegan discusses how “[i]ncongruent multisensory environments,” like the one it aimed to design, are applicable in “augmented reality (AR) environments,” citing Zhou’s study of such environments. Finnegan 3-4, 13 n.46; §VII.B (Zhou overview). Finnegan further taught that its disclosures could be applied in “mixed reality” settings, e.g., environments that “incorporate movement and interaction with physical and tangible props” from the “real world.” Finnegan 11.

From these disclosures, POSAs would have known that Finnegan’s disclosures apply to AR settings and to displaying AR images. THE VR BOOK 29 (“mixed reality” environments include VR and AR environments), Fig. 3.1. Before

May 2017, POSAs knew that AR and VR applications have “many similarities,” and “the convergence of AR and VR [wa]s not new from a research perspective.” *Id.* 484. In fact, by the early 2000s, there were virtual games in which players “seamless[ly] transition[ed] from AR to VR.” Zhou 1004-09, Fig. 11; Roginska ¶¶253-54.

POSAs also would have known and been motivated to use “electronic glasses” for presenting AR images in an AR environment applying Finnegan’s teachings. Roginska ¶¶255-257. “Electronic glasses” were and remain standard AR hardware, and are worn on the head. The VR Book 32-33 (describing AR “HMDs [that] enable computer-generated cues to be overlaid onto the visual field”), Fig. 3.3 (excerpted below) (images of Oculus Rift and “CastAR” HMDs, i.e., AR “electronic glasses”):



Roginska ¶¶255-257.

2. Claim 10: VR Image and HMD

Finnegan disclosed claim 10, which recites: “The electronic device of claim 8, wherein the image is a virtual reality (VR) image, and the display is included with a head mounted display worn on the head of the user.” Finnegan displayed VR

images with an Oculus Rift DK2 HMD, i.e., an HMD worn on a user's head.

Finnegan Abstract, 2, 6; §IV.D.2; Roginska ¶¶258-259.

F. Independent Claim 15

1. 15[pre]: An Electronic System

The preamble of claim 15 recites: “An electronic system, comprising.”

Finnegan discloses an electronic system. §IV.D.1; Roginska ¶260.

2. 15[a]: Image Within A Near-Field Distance

Limitation 15[a], which recites, “a display that displays an image that is located within a near-field distance from a head of a user viewing the image,” would have been obvious for the same reasons as limitation 8[a]. §IV.D.2. Additionally, POSAs would have known that Finnegan presented images at near-field distances from users “viewing th[ose] image[s].” *Id.*; Roginska ¶¶261.

3. 15[b]: Processor for Processing Sound for the Image

Limitation 15[b], which recites, “one or more processors that process sound for the image,” would have been obvious for the same reasons as limitations 8[b] and 8[d]. §IV.D.3; Roginska ¶¶262.

4. 15[c]: HRTFs to Generate Binaural Sound

Limitation 15[c], which requires processing sound for an image “with head-related transfer functions (HRTFs) to generate binaural sound,” would have been obvious for the same reasons as limitation 8[c] and 8[e]. §§IV.D.4-IV.D.5. POSAs

would have known that the HRTFs are used to generate binaural sound. *Id.*; Roginska ¶¶263.

5. 15[d]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight

Limitation 15[d], which requires that the processed and generated binaural sound “has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance,” would have been obvious for the same reasons as limitation 8[c]. §IV.D.4; Roginska ¶¶264.

6. 15[e]: SLP Stays at Far-Field Distance While User’s Head Moves and Image Remains at Near-Field Distance

Limitation 15[e] requires that “the coordinate location of the binaural sound continues to occur at the far-field distance located behind the image when the head of the user moves away from viewing the image and the image remains within the near-field distance from the head of the user.”

A coordinate location of binaural sound occurring at a far-field distance located behind an image would have been obvious over Finnegan, alone or in view of Begault, for reasons discussed. §§IV.D.4, IV.F.5. The additional recitations of limitation 15[e] would have been obvious over Finnegan in view of Begault for the following reasons.

Finnegan explained that its experiments were designed to prompt “incongruent design in VE development,” and expressly identified a virtual environment in which “the observer can move freely about [a] scene” as “an example of where [its] methods could be applied.” Finnegan 11. Thus, POSAs would have known and been motivated to implement Finnegan’s incongruent positioning methods for addressing distance compression in virtual environments in which “observer[s] can move freely.” Finnegan 11. And, POSAs designing such environments would have known to control for head movement, including when “the head of [a] user moves away from viewing [an] image and the image remains within the near-field distance from the head of the user,” as claimed. Roginska ¶¶265-267. That type of 3D audio control was well known before May 2017, as Begault demonstrates. *Id.* ¶¶267.

Begault describes techniques to process binaural sound for virtual reality applications, including sound processing techniques using HRTFs and DSPs. In those descriptions, Begault explains that “[i]n virtual reality, the key to immersivity is directly tied to the success of the interface in providing realistic information.” Begault 165. Therefore, Begault continues, “head-tracked virtual acoustic systems represent the best approach to implementing sound, so that *sound sources can remain fixed in virtual space independent of head movement, as they are in natural space hearing.*” *Id.*; *id.* 133 (teaching that a 3D audio system should “give

the illusion of a sound that remains fixed in space independent of head movement” and discussing commercially available systems using this technique), Figs. 4.37-4.38; Begault Dep. 250:2-9 (admitting systems in the 1980s used this technique).

Consequently, POSAs would have known and been motivated to keep sound sources “fixed in virtual space independent of head movement” (e.g., when a user’s head moves away from an image), as Begault taught, including in a setting with incongruent audio at far-field distances and images at near-field distances, as Finnegan taught. Roginska ¶¶268-69.

Further, Finnegan’s near-field images would obviously remain within “near-field distance[s] from the head of the user”—on the HMD’s lenses—until, for example, the user’s head movement caused them to drop from the field of view. Roginska ¶¶270-71.

G. Dependent Claims 16 and 17

Claims 16 and 17 are unpatentable for the same reasons as claim 15 and as further explained below.

1. Claim 16: AR Display and Image

Claim 16, which recites, “The electronic system of claim 15, wherein the display is an augmented reality (AR) display worn on the head of the user viewing the image, and the image is an AR image,” would have been obvious for the same

reasons as claim 9. §IV.E. And POSAs would have known that users wearing AR displays, e.g., AR HMDs, would view AR images. *Id.*; Roginska¶¶272-75.

2. Claim 17: VR Display and Image

Claim 17, which recites, “The electronic system of claim 15, wherein the display is a virtual reality (VR) display worn on the head of the user viewing the image, and the image is an VR image,” would have been obvious for the same reasons as claim 10. §IV.E.2. And POSAs would have known that users wearing VR displays, e.g., VR HMDs, would view VR images. *Id.*; Roginska¶¶276-77.

V. Ground 2(a): Claims 8-10 and 13-17 are Unpatentable Over Norris, Alone or in View of Begault and Algazi

Norris, Begault, and Algazi teach or suggest every limitation of challenged claims 8-10 and 13-17. Norris disclosed processing and rendering virtual images and sounds with the claimed SLP-behind-image configuration. Begault confirms that Norris’s sound processing could and would be accomplished with a DSP. And Algazi discloses a database of far-field HRTFs that POSAs would have known to use with Norris’s systems, including because Norris prescribes such use.

A. Overview of Norris (Ex.1007)

Norris, issued in December 2015, and also assigned to Patentee, is § 102(a) prior art and was not considered during prosecution.

Norris is directed to systems that use SLPs to simulate locations in empty space from where voices are perceived to originate during calls. Norris, Abstract,

1:19-21, 2:50-52, 3:1-2, Fig. 1. Norris disclosed that an SLP “can exist in a three-dimensional (3D) position described by an azimuth or horizontal angle, elevation or vertical angle, and distance from the listener,” and that sound can be processed so it appears to “emanate from or originate at” the SLP, including with “HRTFs.” *Id.* 3:46-50, 4:21-29, 28:24-32. Norris further taught the use of HRTFs from a “database of common approximate profiles.” *Id.* 21:47-50, 23:49-52, 24:12-21.

Norris’s systems comprise “wearable electronic device[s]” (or “WEDs”), including HMDs and “electronic glasses” that display AR images on their lenses. *Id.* 41:48-59 (“augmented reality imagery is projected through or reflected off a surface of a lens”); *id.* 9:4-17, 30:3-43, 37:29-39, 40:59-41:9, 42:23-32, Fig. 25; Roginska ¶¶146-50.

B. Overview of Algazi (Ex.1008)

Algazi, published in 2001, is § 102(a) prior art and was not considered during prosecution. Ex.1012 ¶¶77-101.

Algazi described “a public-domain database of high-spatial-resolution” HRTFs “measured at the U.C. Davis CIPIC⁶ Interface Laboratory.” Algazi, Abstract. Algazi recognized the “central role” HRTFs play in “spatial audio systems” and disclosed that the CIPIC laboratory measured HRTFs with each “subject seated at

⁶ Center for Imaging Processing and Integrated Computing.

the center of a 1-m radius hoop” and “[s]ound source location[s]” “specified by the azimuth angle θ and elevation angle ϕ in interaural-polar coordinates.” *Id.*, 99; Roginska ¶¶151-54.

C. Motivations to Modify and/or Combine

POSAs would have known and been motivated to modify Norris’s system to use a DSP for sound processing and/or to implement Norris’s system with a DSP as taught by Begault. Norris taught that “digital signal processing (DSP)” provides sound to listeners (Norris 29:9-12, 42:23-32), and as noted, Begault taught that 3D sound processing is “almost always best accomplished” with DSPs. Begault 100-101. And for the same reasons POSAs would have known to modify a Finnegan system to use a DSP, or to combine a Finnegan system with a Begault DSP, as discussed in Ground 1, POSAs would have known to modify a Norris system to use a DSP, or to combine a Norris system with a Begault DSP. §IV.C.

Skilled artisans also would have known and been motivated to use HRTFs from Algazi’s database with Norris’s systems. Norris prescribes using HRTFs from a “database of common approximate profiles,” Norris, 21:47-50, and persons of ordinary skill in the art would have understood that using database HRTFs can be more efficient than creating HRTFs from scratch. Algazi discloses that its CIPIC database HRTFs provide better results than other public HRTF databases (Algazi, 99), and the CIPIC database is widely used in the audio community.

Roginska¶¶282-83. Thus, POSAs would have known and been motivated to implement, and would have reasonably expected to succeed in implementing, Norris’s systems with Algazi’s HRTFs. Indeed, combining the teachings of Norris with Begault, or with Begault and Algazi, would have been obvious to try and well within common sense for POSAs, and would have required no more than combining known elements according to known methods. Roginska¶¶279-85.

D. Independent Claim 8

1. 8[pre]: An Electronic Device

Norris taught claim 8’s preamble, which recites: “An electronic device, comprising.” Norris’s systems and methods comprise electronic devices. Norris, 2:8-11, 2:33-35, 2:40-46, 40:10-41:19, Figs. 17, 25-26; Roginska¶¶291-92.

2. 8[a]: Image Within A Near-Field Distance

Norris disclosed limitation 8[a], which recites: “a display that displays an image that occurs within a near-field distance from a head of a user.”

Norris disclosed wearable electronic devices (“WEDs”) that display images to users on a display, including “head-mounted displays” and “electronic glasses.” Norris, 41:48-59; 31:52-32:48, 37:29-39, 40:59-41:9, Figs. 17, 19-20, 25-26. And, as discussed in Ground 1, before May 2017, persons of ordinary skill in the art understood that HMDs, e.g., Norris’s “electronic glasses,” display images on their lenses—well within 1 m from the head of a user, i.e., “within a near-field distance.” Roginska¶¶293-296; The VR Book, Figs. 2.1 and 3.3 (showing HMDs and AR

electronic glasses); Norris, 30:40-43 (“listener 1710 sees through a display of the wearable electronic device 1720 a real-time image of the band from a front row seat at the venue”), Fig. 17, 41:47-59, 40:4-23. Further, as noted in Ground 1, Patentee alleges in district court that limitation 8[a] is satisfied by images displayed on HMD lenses inches in front of a user’s eyes, which is how Norris’s HMDs work. §IV.D.2.

3. 8[b] and 8[d]: One or More Processors for Processing Sound for the Image, Including a DSP

Limitations 8[b] and 8[d] recite, “one or more processors that process sound for the image,” and “wherein the one or more processors include a digital signal processor (DSP),” respectively.

Norris disclosed a “processor unit” comprising “a processor,” such as a “CPU” or “microprocessor” for “perform[ing] operations and tasks that implement one or more blocks of [Norris’s] flow diagrams.” Norris, 41:10-60. And Norris’s flow diagrams relate to processing sound for images. *Id.* Figs. 1-10. Norris also describes and claims processing sound for associated virtual images. *Id.* 31:65-32:48 (“sound localization point 2050 can appear to the listener 2020 as a virtual image of the speaker 2010,” and “listener 2020 can see the sound localization point 2050 as it moves in the room 2055”), Fig. 20, 40:59-41:9, 30:37-43, cls. 15, 17-19. Additionally, Norris taught that “digital signal processing” provides sound to listeners. *Id.* 29:9-12, 42:23-32.

Based on these disclosures, it would have been obvious to POSAs reading Norris (i) to use a processor to process sound as limitation 8[b] requires (as Norris expressly teaches), and (ii) that the “digital signal processing” Norris discloses could and would be implemented with a digital signal processor (DSP), for example, by using a DSP as Norris’s processor or adding a DSP to Norris’s system to handle sound processing for images. Roginska ¶¶297-303. Either of these would have rendered obvious limitations 8[b] and 8[d]. *Id.* ¶¶ 297-303, 313-314. And, as explained in Ground 1, POSAs would have known and been motivated to use a DSP for sound processing for images because DSPs were routinely used and known to be advantageous for such processing, before May 2017, as Patentee’s expert, Dr. Begault, admitted. §§ IV.C, IV.D.3; Roginska ¶¶315-316. Again, DSPs are fast, excel at signal processing tasks like HRTF processing that can be used to generate 3D audio, are reconfigurable and flexible, and are appropriately sized for use in systems like Norris’s. Roginska ¶317.

Accordingly, limitations 8[b] and 8[d] would have been obvious based on Norris alone. *Id.* ¶¶317-18; *see also* Begault Dep. 200:3-201:9 (Dr. Begault effectively conceded the obviousness of limitations 8[b] and 8[d]) (reproduced in §IV.D.3).

Additionally and alternatively, limitations 8[b] and 8[d] would have been obvious over Norris in view of Begault—specifically, in view of Begault’s

descriptions of DSPs and prescription to use DSPs for sound processing in virtual environments. Begault, 100-101. As discussed above, POSAs would have known, including based on Begault, that DSPs were advantageous for sound processing for virtual images and thus would have been motivated to implement a Norris-type system with a DSP handling Norris's "digital sound processing." §V.C; Roginska¶¶319-22. The DSP in such a combination would have rendered obvious both limitation 8[b]'s "one or more processors that process sound for the image" and limitation 8[d]'s "the one or more processors include a digital signal processor (DSP)." Roginska¶¶321; *see also* note 5.

4. 8[c]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight

Limitation 8[c] requires that the sound processing by the claimed one or more processors, e.g., the DSP, "generate[s] binaural sound" that "has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance." Limitation 8[c] would have been obvious over Norris, alone or in view of Begault.

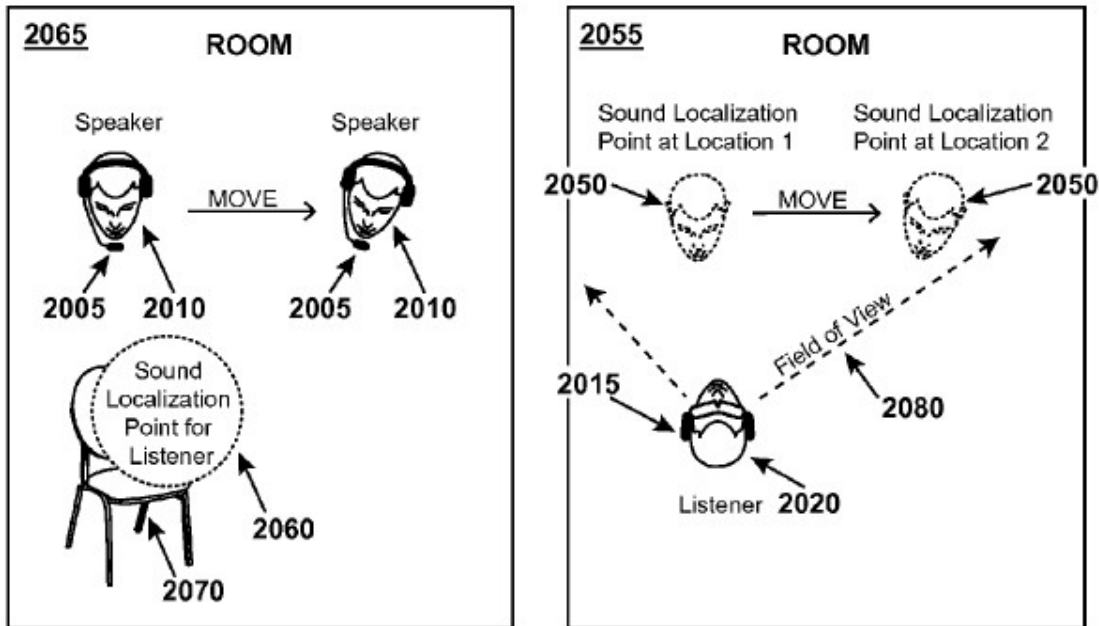
First, Norris disclosed "provid[ing] a sound localization point in empty space" by "generat[ing] binaural sound" "that localizes to a listener." Norris, 22:60-23:3.

Second, Norris disclosed SLPs with coordinate locations that occur at far-field distances, i.e., at distances of "greater than or equal to one meter" ('836 patent 27:23-

25). Norris, 20:51-21:6 (“voice... localized at a distance of ten feet”), 34:5-21 (“electronic device of John establishes a sound localization point for the voice of Paul to be five feet in front of John”), 4:67-5 (“in a Cartesian coordinate system,” “sound is adjusted to change one or more of the (X, Y, Z) coordinates. For instance, when a listener is located at (0, 0, 0), an origin of the sound or the sound localization point for the listener appears to originate from (10, 17, 25)”), 5:25-31 (“binaural sound is processed to change the sound localization point from being two feet away from the listener to being thirty feet away”), 13:55-14:9 (“voice... localizes... fifteen (15) feet away”), 15:4-7 (“calls... in the morning are localized five feet from the individual, and calls that occur in the afternoon or evening are localized eight feet from the individual”), 15:41-42, 15:52-54, 17:45-54, 18:51-54, 20:56-59, 22:62-65, 34:5-21; *id.* 42:23-32; Roginska ¶¶304-306.

Third, Norris’s far-field SLPs would all be behind their associated images when those images are displayed on the lenses (display) of a Norris HMD (i.e., at near-field distances). §V.D.2; Roginska ¶¶307-308.

Fourth, POSAs would have known that, in addition to being behind associated images, Norris’s far-field SLPs would be along a user’s line of sight to the images. Roginska ¶¶308-09. Norris’s Figure 20 shows that configuration:



Norris, Fig. 20. In Figure 20, a “first person speaker” (2010) in the left room (2065) wears a headset with a microphone (2005) and communicates with a “second person,” the “listener” (2020), in the right room (2055). *Id.* 31:66-32:4. The WED 2005 of speaker 2010 “captures binaural sound and transmits sound to the listener 2020 who localizes the sound at a sound localization point 2050 located in [his] room 2055.” *Id.* 32:4-7. Similarly, the WED 2015 of listener 2020 “captures binaural sound and transmits sound to the speaker 2010 who localizes the sound at a sound localization point 2060 located in [his] room 2065.” *Id.* 32:7-11.

In the right room (2055), the listener’s WED 2015 “is a pair of electronic glasses with a display that provides an image or video of the speaker 2010 in a field of view 2080 of the listener.” *Id.* 32:40-48. That virtual image of the speaker is displayed on the lenses (display) of the listener’s electronic glasses, with its

associated SLP (2050) occurring behind the image on the lenses. SLP 2050 is (or at a minimum obviously would occur) along listener 2020's line of sight to the virtual image of the speaker, in the listener's depicted "field of view." *Id.* Fig. 20; Roginska ¶¶310-11. Moreover, Figure 20 is just one of multiple examples in Norris that POSAs would have known use the claimed configuration of SLPs and images—with images on HMDs short distances in front of users' eyes and associated far-field SLPs behind, and along a line of sight to, the images. Roginska ¶312; Norris, 30:19-43, 20:51-21:6.

5. 8[e]: DSP Processes Sound with Far-Field HRTFs While Electronic Device is at Near-Field Location

Limitation 8[e] requires that the DSP recited in limitation 8[d] "processes the sound for the image with far-field head-related transfer functions (HRTFs) while the electronic device is located at the near-field distance from the head of the user." Limitation 8[e] would have been obvious over Norris, alone or in view of Begault and Algazi.

First, as discussed, Norris, alone or in view of Begault, rendered obvious using a DSP to process sound for images, including to generate binaural sound. §§V.D.3-V.D.4.

Second, POSAs would have known and understood that Norris processed binaural sounds with "far-field head-related transfer functions (HRTFs)." Norris disclosed processing sound for images with HRTFs, including for "far field distances

greater than about one meter,” where a “far field HRTF” can be used. Norris, 28:24-32; *id.* 5:25-31, 28:16-29:8, 26:9-27; Roginska ¶¶323-325. Furthermore, to POSAs, Norris disclosed (or at a minimum rendered obvious) processing sound with “far-field” HRTFs through its numerous descriptions of localizing sound to distances “greater than or equal to one meter” (’836 patent 27:23-24), e.g., those discussed in Section V.D.4. Roginska ¶¶323-25. Also, POSAs would have known and been motivated to use far-field HRTFs, including to process sounds for near-field images, because that would require fewer stored HRTFs, could use commonly available far-field HRTF databases, and could leverage other adjustments for producing near-field effects, such as adjusting intensity or volume. Roginska ¶¶326-28.

Third, the HMDs Norris discloses would have been worn on users’ heads, and those HMDs and their lenses would have been located at near-field distances from users’ heads, including while sound for displayed images was processed as claimed. Roginska ¶¶327; Norris Fig. 20; The VR Book Figs. 2.1, 3.3 (showing HMDs and AR electronic glasses).

Additionally and alternatively, limitation 8[e] would have been obvious over Norris, alone or in view of Begault, and further in view of Algazi. Norris taught that HRTFs from public databases can be used with its systems and methods (Norris 21:47-50), and Algazi described and disclosed one of the most well-known databases

for such HRTFs—the CIPIC database—which POSAs would have known and been motivated to use. §§V.B-V.C; Algazi 99; Roginska¶¶328-332.

Algazi taught that in the CIPIC database “[a]ll HRTFs were measured with the subject seated at the center of a *1-m radius* hoop.” Algazi, 99. Thus, POSAs would have known and understood that an implementation of Norris’s systems and methods using Algazi’s CIPIC database HRTFs, per Norris’s instruction, would use “far-field” HRTFs—HRTFs measured at distances of “greater than or equal to one meter” (’836 patent 27:23-25). Roginska¶¶329-330; §V.C.

E. Dependent Claims 9, 10, 13, and 14

Claims 9, 10, 13, and 14 are unpatentable for the same reasons as claim 8 and as further explained below.

1. Claim 9: AR Image and Electronic Glasses

Claim 9 recites: “The electronic device of claim 8, wherein the image is an augmented reality (AR) image, and the display is included with electronic glasses worn on the head of the user.”

Norris disclosed “electronic eyeglasses” worn on users’ heads “in which augmented reality imagery” (i.e., an “augmented reality (AR) image”) “is projected through or reflected off a surface of a lens” (i.e., displayed). Norris, 41:48-59; *id.* 32:42-48 (WED 2015 is “a pair of electronic glasses with a display that provides an image or video of the speaker 2010 in a field of view 2080 of the listener 2020”),

40:59-41:9, Fig. 20; Roginska¶¶333-37.

2. Claim 10: VR Image and HMD

Claim 10 recites: “The electronic device of claim 8, wherein the image is a virtual reality (VR) image, and the display is included with a head mounted display worn on the head of the user.”

Norris disclosed that its WEDs include “head-mounted displays,” which are worn on users’ heads, and an “imagery system” “such as an optical projection system, a *virtual image display system*, virtual augmented reality system, and/or a spatial augmented reality system” for use with those HMDs for displaying images. Norris, 41:48-59, 40:59-41:9. Norris further described electronic calls between “first and second individuals” in “separate virtual realities,” and taught location and head orientation tracking for “augmented” and “virtual reality” applications. *Id.* 6:29-45, 9:12-15. Accordingly, POSAs would have known and understood that Norris’s systems and methods use HMDs worn on users’ heads that display virtual reality images, as claimed. Roginska¶¶338-39.

3. Claim 13: AR Image of an IUA

Claim 13 recites: “The electronic device of claim 8, wherein the image is an augmented reality (AR) image of an intelligent user assistant (IUA) that is located less than one meter away from the head of the user, and a voice of the IUA occurs at a coordinate location behind the AR image of the IUA.”

Norris disclosed, or at a minimum rendered obvious, AR images located less than one meter from the head of a user, e.g., images on the lenses of its “electronic glasses” for AR (§§V.D.2 and V.E.1), and SLPs located behind such images (§V.D.4). Additionally, Norris disclosed that its systems and methods can utilize “an intelligent personal assistant.” Norris, 41:33-35; Roginska¶¶340-41. POSAs would have known and understood from Norris that the “intelligent personal assistant” (the claimed IUA) could and would (i) be presented as an AR image and (ii) vocalize from Norris’s far-field SLPs behind such an image. Roginska¶¶342-43; Ex.1051, Ross, 17:63-18:9 (“a character and/or avatar” “walks the user through a scene in virtual space and points things out,” similar to a “commentary track”).

4. Claim 14: VR Image of an IUA

Claim 14 recites: “The electronic device of claim 8, wherein the image is a virtual reality (VR) image of an intelligent user assistant (IUA) that is located less than one meter away from the head of the user, and a voice of the IUA occurs at a coordinate location behind the VR image of the IUA.”

Norris disclosed, or at a minimum rendered obvious, VR images located less than one meter from the head of a user, e.g., images on the lenses of its HMDs for VR (§§V.D.2 and V.E.2), and SLPs located behind such images (§V.D.4). Additionally, as noted, Norris disclosed using “an intelligent personal assistant.” Norris, 41:33-35. POSAs would have known and understood that Norris’s

“intelligent personal assistant” (the claimed IUA) could and would (i) be presented as a VR image and (ii) vocalize from Norris’s far-field SLPs behind such an image. Roginska ¶¶344-46; Ross, 17:63-18:9.

F. Independent Claim 15

1. 15[pre]: An Electronic System

The preamble of claim 15 recites: “An electronic system, comprising.” Norris discloses an electronic system. §V.D.1; Roginska ¶¶347.

2. 15[a]: Image Within A Near-Field Distance

Limitation 15[a], which recites, “a display that displays an image that is located within a near-field distance from a head of a user viewing the image,” would have been obvious for the same reasons as limitation 8[a]. §V.D.2. Additionally, POSAs would have known that Norris presented images at near-field distances from users “viewing th[ose] image[s].” *Id.*; Roginska ¶¶348.

3. 15[b]: Processor for Processing Sound for the Image

Limitation 15[b], which recites, “one or more processors that process sound for the image,” would have been obvious for the same reasons as limitations 8[b] and 8[d]. §V.D.3; Roginska ¶¶349; *see also* note 5.

4. 15[c]: HRTFs to Generate Binaural Sound

Limitation 15[c], which requires processing process sound for an image “with head-related transfer functions (HRTFs) to generate binaural sound,” would have been obvious for the same reasons as limitation 8[c] and 8[e]. §§V.D.4-V.D.5.

POSAs would have known that the HRTFs are used to generate binaural sound. *Id.*; Roginska¶350.

5. 15[d]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight

Limitation 15[d], which requires that the processed and generated binaural sound “has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance,” would have been obvious for the same reasons as limitation 8[c]. §V.D.4; Roginska¶351.

6. 15[e]: SLP Stays at Far-Field Distance While User’s Head Moves and Image Remains at Near-Field Distance

Limitation 15[e] requires that “the coordinate location of the binaural sound continues to occur at the far-field distance located behind the image when the head of the user moves away from viewing the image and the image remains within the near-field distance from the head of the user.”

A coordinate location of binaural sound occurring at a far-field distance located behind an image would have been obvious over Norris, alone or in view of Begault and Algazi, for reasons discussed. §§V.D.4, V.F.5. The additional recitations of limitation 15[e] would have been obvious over Norris, alone or in view of Begault and Algazi, for the following reasons.

Norris disclosed a processor executing instructions to “change an interaural time difference, an interaural level difference, and a head-related transfer function of [] binaural sound in order to maintain [a] sound localization point fixed at [a] location” “while the first person physically moves away from the location and changes a head orientation with respect to the location.” Norris, cl. 18, *id.* cl. 8.

Norris further disclosed that “[m]ultiple sound localization points [can] exist in [a] room, and binaural synthesis of sound dynamically occurs in real-time as the listener moves his head and/or body.” *Id.* 8:50-53. Thus, “[a]s [a] listener walks around the room, the sound localization points remain fixed or unmoved to provide virtual sound sources located in the room.” *Id.* 8:53-55. That way, “[i]f the virtual sound source represents a fixed object, then the virtual sound source remains fixed in the room.” *Id.* 8:55-57. “For instance, sound that emanates from a virtual water fountain is a fixed sound source since the virtual water fountain does not move.” *Id.* 8:57-59; Roginska ¶¶352-353.

Consequently, POSAs would have understood that Norris’s systems and methods could and would maintain binaural sound at far-field distances located behind near-field images displayed on the lenses of a Norris AR or VR HMD, e.g., the “virtual water fountain,” even when a user’s head moves away from viewing that image, so the user could continue to properly localize such sound sources. Roginska ¶¶354-355. As noted in Ground 1, and as Begault confirms, this type of

sound processing—accounting for head movement while maintaining binaural sound source locations—was well known to POSAs before May 2017. §IV.F.6; Begault Dep. 250:2-9.

Further, Norris’s near-field images, e.g., the “virtual water fountain,” would obviously remain within “near-field distance[s] from the head of the user”—on Norris’s HMD’s lenses—until, for example, the user’s head movement caused them to drop from the field of view. Roginska¶¶354.

G. Dependent Claims 16 and 17

Claims 16 and 17 are unpatentable for the same reasons as claim 15 and as further explained below.

1. Claim 16: AR Display and Image

Claim 16, which recites, “The electronic system of claim 15, wherein the display is an augmented reality (AR) display worn on the head of the user viewing the image, and the image is an AR image,” would have been obvious for the same reasons as claim 9. §V.E.1. And POSAs would have known that users wearing AR displays, e.g., AR HMDs, would view AR images. *Id.*; Roginska¶¶356-59.

2. Claim 17: VR Display and Image

Claim 17, which recites, “The electronic system of claim 15, wherein the display is a virtual reality (VR) display worn on the head of the user viewing the image, and the image is an VR image,” would have been obvious for the same

reasons as claim 10. §V.E.2. And POSAs would have known that users wearing VR displays, e.g., VR HMDs, would view VR images. *Id.*; Roginska ¶¶360-61.

VI. Ground 2(b): Dependent Claims 11 and 12 are Unpatentable Over Norris, Alone or in View of Begault and Algazi, and Further in View of Sauk

Claims 11 and 12 are unpatentable for the same reasons as claim 8 (§V.D) and as further explained below.

Specifically, the additional limitations of claims 11 and 12, which relate to switching from binaural sound to stereo sound based on head rotation, would have been obvious over Sauk, which taught such switching.

A. Overview of Sauk (Ex.1009)

Sauk, issued January 25, 2011, is § 102(a) prior art and was not considered during prosecution.

Sauk taught a “communication space” where sound relevance can be signified by how sound is processed. Sauk, Title, 11:26-40. Sauk described a “Selective Relevance Mode” using “a plurality of audio dimensions such as 3-D (binaural), 2-D (stereophonic), and 1-D (monophonic).” *Id.* 11:26-40. Sauk explained that, because users can perceive differences when audio is provided in different dimensions, “changing an audio format for sounds from binaural to monophonic or stereophonic audio signifies a different level of relevancy or importance for such audio.” *Id.* Thus, as certain sounds become less relevant, the system can “change

the less relevant audio to a monophonic (1-D) or stereophonic (2-D) dimension.”

Id.; Roginska ¶¶155-162.

B. Motivation to Combine Sauk

POSAs would have known and been motivated to combine the teachings of Sauk with Norris, alone or with Begault and Algazi.⁷

Sauk, like Norris, aims to provide advantageous mechanisms for improving the provision of binaural audio, including in AR, VR, and mixed-reality settings. *E.g.*, Sauk, 2:45-47 (“The invention concerns a method and apparatus for producing, combining, and customizing virtual sound environments to provide the user with understandable information regarding their surroundings.”); §V.A (Norris overview). POSAs would have recognized before May 2017 that Sauk describes “innovative applications for actually putting [] binaural audio information to practical use” (Sauk, 2:34-41), and that Sauk’s method of switching from binaural to stereo sound to indicate relevance of audio could and would enhance user experience and perceptual accuracy while lowering processing requirements. Roginska ¶¶286-88.

Moreover, POSAs would have had a reasonable expectation of success in combining the teachings of Norris, alone or with Begault and Algazi, and Sauk. Like Norris, Begault, and Algazi, Sauk taught ways to enhance binaural sound, including

⁷ For overviews of Norris, Algazi, and Begault, see Sections V.A, V.B, and IV.B.

through the use of HRTFs and implementations on WEDs. And Norris and Sauk implement HRTF processing methods and head and position tracking in analogous ways. Sauk, 4:33-5:5 (position and head tracking), 7:24-8:33 (sound processing with HRTFs); Norris, 6:29-45 (position and head tracking), 4:20-29 (sound processing with HRTFs); Roginska ¶289; *see also* Begault Dep. 208:20-209:8 (testifying that sound systems “prior to 2014” could “use different types of spatial display” and could “switch sound from binaural audio to stereo audio or monaural audio,” e.g., “as a form of alert”).

C. Claim 11: Sensors Detect Rotation, and Binaural Sound Switches to Stereo

Claim 11 recites: “The electronic device of claim 8 further comprising: one or more sensors that detect rotation of the electronic device, wherein the binaural sound automatically switches to stereo sound in response to the one or more sensors detecting the rotation of the electronic device.” Claim 11 would have been obvious over Norris, alone or in view of Begault and Algazi, and further in view of Sauk.

Norris disclosed one or more sensors that detect rotation, including a “head orientation tracker” that detects rotation of a HMD by “measur[ing] one or more rotational head orientations of yaw (side-to-side or left and right), pitch (up and down), and roll (tilting) of the head.” Norris, 6:29-45, Fig. 26 (Sensors 2630), 40:4-58 (describing Figure 26 and sensors 2630); Roginska ¶¶363-64.

Norris further taught that a user can arrange SLPs “in any order,” including “mov[ing] any of them to the foreground,” e.g., “to make the sound louder,” “moving any of them to the background,” e.g., “to make the sound softer,” and “moving any of them to increase or decrease priority or access to resources.” Norris, 36:17-25. Thus, based on Norris, POSAs would have known and been motivated to prioritize different sounds (such as those directly in front of the user) and the processing of sounds based on SLP location, including by degrading sound where appropriate. And Sauk disclosed a useful way to do this, explaining that a system can “change [] less relevant audio to a monophonic (1-D) or stereophonic (2-D) dimension,” to “signif[y] a different level of relevancy or importance” and “remove any localization cues for” that audio. Sauk, 11:25-40; Roginska ¶365.

Thus, in view of Norris, alone or with Begault and Algazi, and further in view of Sauk, POSAs would have known and been motivated to degrade binaural audio, including by automatically switching it to stereo sound, as a WED is rotated. *Id.* ¶¶366-67; §VI.B. For example, per Sauk’s teachings, using a sensor to detect motion of the WED and automatically switching binaural sound to stereo sound for images away from which a user is rotating could and would indicate that such sound was (or was becoming) less relevant. Sauk, 11:26-41. POSAs would have known that when binaural sound is so switched to stereo sound, a user would still be able to hear the less relevant stereo sound but would focus on the binaural sound. Or, in a

scenario where the sound is rotated out of view to one side of the user, the user could still hear the out-of-view sound in stereo sound (which requires less processing resources) while still being able to recognize that the sound comes from the side. POSAs also would have understood that altering sound this way could and would conserve processing power and extend battery life, both powerful motivations for POSAs implementing binaural audio in AR and VR systems. Roginska ¶¶367-68.

D. Claim 12: Binaural Sound Switches to Stereo in Response to Change to Head Orientation Exceeding Threshold Azimuth Value

Claim 12 recites: “The electronic device of claim 8, wherein the binaural sound automatically switches to stereo sound in response to the electronic device detecting a change to a head orientation of the user that exceeds a threshold value in an azimuth direction.” Claim 12 would have been obvious over Norris, alone or in view of Begault and Algazi, and further in view of Sauk, for the same reasons as claim 11. §VI.C.

Just as POSAs would have known that automatically switching binaural sound to stereo sound for images away from which a user is rotating would helpfully indicate that such audio is (or is becoming) less relevant and could conserve resources (*see id.*), POSAs would have known and understood that doing that when a user rotates past a threshold value in an azimuth direction—i.e., when the user turns far enough that some images and sound sources rotate out of view—would have the same benefits. Roginska ¶¶369-71.

VII. Ground 3: Claims 8-10 and 15-17 are Unpatentable Over Shivappa, Alone or in view of Zhou

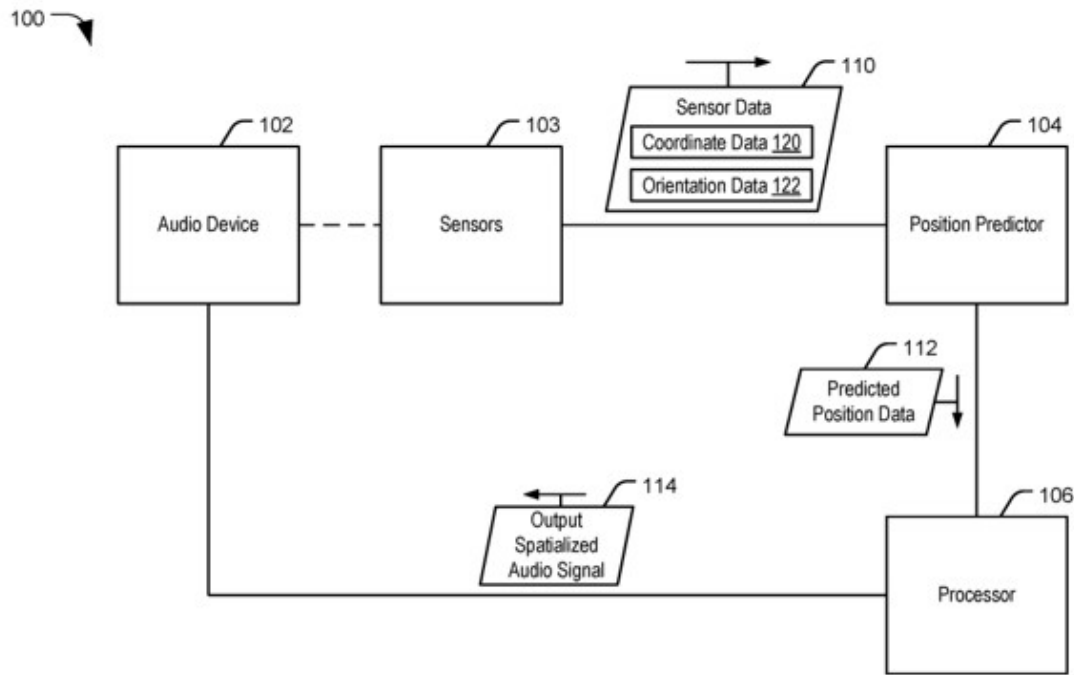
Challenged claims 8-10 and 15-17 would have been obvious over Shivappa, alone or in view of Zhou. Shivappa disclosed processing and rendering virtual images and sounds with the claimed SLP-behind-image configuration. And Zhou confirms that POSAs would have known Shivappa's sound processing could and would be accomplished with far-field HRTFs.

A. Overview of Shivappa (Ex.1010)

Shivappa, filed April 8, 2016 and issued in April 2021, is § 102(a)(2) prior art and was not considered during prosecution.

Shivappa disclosed using 3D sound rendering techniques to allow a user to perceive 3D sound (i.e., binaural sound), including the direction and distance of a sound source. Shivappa, 5:22-36, 24:49-53 (“audio processing device may be configured to process object-based 3D audio” “with 3D coordinates and directivity”). Shivappa's “3D sound rendering” techniques include “performing binauralization” with “head related transfer functions (HRTFs).” *Id.* 7:60-8:4, 23:29-52, 24:45-57; *see also id.*, References Cited (listing publications addressing binaural and 3D sound and HRTFs, including references discussing “HRTF Audio” in Oculus devices).

Shivappa disclosed an audio processing device 100 that generates spatialized audio signals using processor 106 based on predicted position data of the user or audio device 102, and sensors 103 for detecting position changes:



Id. Fig. 1A, 5:6-20. Shivappa used a “DSP” for this processing. *Id.* Fig. 7 (item 710), 27:63-28:24, 28:58-29:22. Shivappa further disclosed that its “audio device 102 includes (or is integrated within) a head mounted display (HMD) of a virtual reality (VR) system or an augmented reality (AR) system.” *Id.* 5:37-6:3, Fig. 2.

Additionally, Shivappa disclosed that sensors on its HMDs determine the location and orientation of the audio device and user, and then the processor 106, e.g., the “DSP,” generates spatialized audio signals for sound sources in a virtual environment accounting for a user’s head and other movement. *Id.* 6:51-55, 8:28-42. That way, for example, a user may initially perceive a flute playing to the left, but when the user turns left, the sound changes so the user perceives the flute to be playing from in front. *Id.* 6:7-11; Roginska ¶¶163-70.

B. Overview of Zhou (Ex.1011)

Zhou, published in 2004, is § 102(a) prior art and was not considered during prosecution. Ex.1012 ¶¶102-27.

Zhou “examined the effect of 3D sound in a larger AR environment,” which included an “AR environment” and a “VR environment” and “a seamless transition” between the two. Zhou 1004-07, Fig. 11; *id.* Abstract, §1.

Zhou recognized that it had been known for years that “primary benefits of auditory spatial information” include “relieving processing demands on the visual modality, enhancing spatial awareness, and directing attention to important spatial events.” *Id.* 1004-05. Zhou aimed to show additional benefits, including helping users localize and identify different spatial objects, and to assess “the impact of 3D sound on depth perception.” *Id.* Abstract, 1005.

Zhou recognized that HRTFs can be used “in virtual and augmented environments” to externalize sound sources and obtain “[e]xact spatial 3D sound” (i.e., binaural sound). *Id.* 991. To “generate 3D sound,” Zhou used “open source software development kits (SDKs)” incorporating “general HRTFs” from a database, which “update[d] the relative position of the virtual sound source[s] to [a user’s] head co-ordinate system.” *Id.*, 994 (“[W]e applied [] OpenAL”), 1013 (“3D sound generated by OpenAL ... applying HRTFs”), 994-95 (“3D sound co-ordinates” used to position sound), Fig. 1 (depicting same).

In a first experiment, Zhou placed sound sources “around 1.5 m away in the front of the listener[s].” *Id.* 993. In a second, Zhou used an “InterSense IS-900 system ... for precise motion tracking” and “the user’s head position, orientation, and the virtual sound sources’ 3D position as input parameters for [its] OpenAL software,” including while co-players collaborated in a game while “walk[ing] around in a three by three meter region.” *Id.* 997-98, 1006-07, Figs. 9-11.

From its experiments, Zhou concluded that “the addition of 3D sound can be a very effective way to complement the user experiences in AR environments” because “[i]t helps depth perception” and “enables a more realistic environment and more immersive feeling.” *Id.* 1014; Roginska ¶¶171-81.

C. Motivation to Combine Shivappa with Zhou

POSAs would have known and been motivated to combine Shivappa’s systems and methods, which process HRTFs to produce binaural sound, with Zhou’s teachings of using spatialized 3D audio in AR and VR settings. Roginska ¶372. POSAs would have known that Zhou’s 3D audio could complement user experiences, including in the AR and VR systems on which Shivappa focuses, because, as Zhou teaches, 3D audio helps with depth perception and enables a more realistic environment and immersive feeling. Zhou 1014; Shivappa 3:16-19; Roginska ¶373. Further, POSAs combining Shivappa’s systems and methods with Zhou’s teachings would have had a reasonable expectation of success because,

before May 2017, the methods and systems Zhou disclosed (in 2004) would have been well within a POSA's common sense and skill, and would have required no more than combining prior art elements (e.g., HMDs, HRTFs, and open source audio SDKs) according to known methods. Roginska¶374.

D. Claim 8

1. 8[pre]: An Electronic Device

Shivappa taught the preamble of claim 8, which recites: “An electronic device, comprising.” Shivappa, 5:37-6:3 (“audio device 102” “includes (or is integrated within) a head mounted display (HMD)”), 24:49-64, Figs. 1, 2, 7; Roginska¶¶379-80.

2. 8[a]: Image Within A Near-Field Distance

Shivappa disclosed limitation 8[a], which recites: “a display that displays an image that occurs within a near-field distance from a head of a user.”

Shivappa disclosed VR and AR HMDs with “display screen[s] for displaying visual information.” Shivappa, 5:37-53, Fig. 2, Fig. 7, 28:25-26 (“display 728”). And POSAs would have known that HMDs display images on their lenses, within “a distance less than one meter,” i.e., a “near-field distance,” from the head of a user ('836 patent 27:23-25). Shivappa, 5:42-45 (“The VR system (or the AR system) may be configured to display media content, such as movies, and to provide interactive content, such as video games, presentations, virtual meetings, etc.”). Further, as noted in Grounds 1 and 2, Patentee alleges in district court that limitation

8[a] is satisfied by images displayed on HMD lenses inches in front of a user's eyes, which is how Shivappa's HMD works. §IV.D.2; Roginska¶¶381-383.

3. 8[b] and 8[d]: One or More Processors for Processing Sound for the Image, Including a DSP

Limitations 8[b] and 8[d] recite, “one or more processors that process sound for the image,” and “wherein the one or more processors include a digital signal processor (DSP),” respectively. Shivappa disclosed, or at a minimum rendered obvious, both limitations.

Shivappa disclosed an “audio processing device” that “provide[s] spatialized audio signals to an audio device” that “may be integrated in a virtual reality (VR) system or augmented reality (AR) system.” Shivappa, 3:16-18, 3:60-4:19. Shivappa further disclosed a “processor” that “perform[s] 3D sound rendering to generate a 3D output audio signal,” including through “binauralization” and “head-related transfer functions (HRTFs),” which “enable a user to perceive a direction or a distance of sounds in a 3D audio space relative to a location of a user.” *Id.* 3:19-23, 5:37-6:19, 7:53-8:4, 23:29-52, 24:45-57, Fig. 1A (“processor” 106). And the binaural sounds processed by Shivappa's system are for virtual images. For example, a user can “play[] a game” where “a car,” i.e., a virtual image of a car, “drives to the left of the user in the game,” and the “audio output enables the user to perceive a sound of a car as coming from the user's left side.” *Id.* 3:23-26. Shivappa's claims also recite processors that, among other things, “adjust an

orientation of the 3D sound field” and “output the binaural audio signal ... for playback,” including where the audio is “associated with a virtual reality or augmented reality application” and a “display is configured to display visual information associated with the virtual reality or augmented reality application.” *Id.* cls. 1, 6; Roginska ¶¶384-87.

Additionally, Shivappa disclosed that its system’s processor, which processes sound for images as described above, is “a digital signal processor (DSP).” Shivappa, 28:1-3, 28:21-24 (“processor 710” can be “a digital signal processor (DSP)” and “configured to generate [an] output spatialized audio signal”), 28:4-5 (“The processor 710 may include or correspond to the processor 106 of FIGS. 1A-C, 2, and 3.”); 7:40-44 (“processor 106 may include ... a digital signal processor (DSP)”), 13:40-45 (“Each component illustrated in FIGS. 1A-C may be implemented using hardware,” e.g., “a DSP”), 21:34-39, 22:39-44, 25:11-16, Fig. 7 (“Processor (e.g., DSP)”); Roginska ¶¶397-398; *see also* Begault Dep. 200:3-201:9 (Dr. Begault effectively conceded the obviousness of limitations 8[b] and 8[d]) (reproduced in §IV.D.3).⁸

⁸ Even if the processor-related elements of limitations 8[b] and 8[d] were construed as means-plus-function terms, they would have been obvious over Shivappa and Zhou because Zhou disclosed using “HRTFs” to “produce 3D sound effects” from a

4. 8[c]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight

Limitation 8[c] requires that the sound processing by the claimed one or more processors, e.g., the DSP, “generate[s] binaural sound” that “has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance.” Limitation 8[c] would have been obvious over Shivappa, alone or in view of Zhou.

Shivappa disclosed virtual images and associated binaural sounds (e.g., binaural sounds generated for virtual doors, characters, instruments, conference speakers, and vehicles), and POSAs would have known these sounds could and would be located at SLPs with coordinate locations occurring at far-field distances (“greater than or equal to one meter,” ’836 patent 27:23-25) from a user. Shivappa, 5:30-36 (“a user may perceive the auditory sound as being in 3D, which may enable the user to perceive direction, distance, or both of one or more sound sources corresponding to the auditory sound. For example, a user may perceive a sound of

“convolution” process (*e.g.*, Zhou, 992), and because, as noted (*see* note 5), Dr. Begault has opined that “[c]onvolution is the main algorithm required to perform the claimed processes,” while conceding that convolution was “well-known.” Ex.1055, ¶16; Begault Dep. 59:4-16.

a door opening to their right ... for an auditory sound of an opening door”), 8:42-47 (“[A] sound source may be a character speaking in a movie, an instrument playing in a concert, a person speaking during a teleconference, a vehicle making noise in a virtual reality video game, or any other source of sound corresponding to the input spatialized audio signal.”); Roginska ¶¶388-90.

POSAs also would have understood that Shivappa’s binaural sounds would (or obviously would) occur (i) behind associated images presented at near-field distances on Shivappa’s HMDs’ lenses, as discussed in Section VII.D.2, and (ii) along a user’s line of sight because the user can look in the direction of the binaural source and its associated image. Shivappa, 3:14-31 (“The audio outputs may be three-dimensional (3D) audio outputs that enable a user to perceive a direction or a distance of sounds in a 3D audio space relative to a location of a user, either in a game (or other virtual reality environment) or in reality.”), 5:42-54 (“The visuals and audio output by the virtual reality system may change based on movement of the user (e.g., movement of the user’s head).”), 24:49-53; Roginska ¶¶390-91. Thus, Shivappa alone renders obvious limitation 8[c].

Additionally and alternatively, limitation 8[c] would have been obvious over Shivappa in view of Zhou. Zhou taught “generat[ing] 3D sound,” i.e., binaural sound, using open source SDKs that incorporated HRTFs from available databases, and Zhou disclosed positioning binaural sound at far-field distances while associated

images were located at near-field distances—namely, on the lenses of an HMD. For example, in its first experiment, Zhou’s SDK applied HRTFs to place binaural sounds “around 1.5 m away” from a listener, which is a far-field distance, along the listener’s line of sight and while images associated with the sound sources were presented to the listener on the lenses of Zhou’s HMD, i.e., at a near-field distance. Zhou 993, 1013. Similarly, in a second experiment, Zhou tracked co-players “walk[ing] around in a three by three meter region,” and placed binaural sounds at the locations of those co-players’ controllers (their “wands”). *Id.* 997-98, 1006-07. POSAs would have known that such co-players could and would be separated by distances of one meter or more, i.e., far-field distances, in a nine square meter play area. Roginska ¶¶92-396. Thus, in this scenario as well, binaural sounds (e.g., the “buzz” of a co-player’s airplane) would be positioned at far-field distances (on co-player wands separated by a meter or more) while images were positioned at near-field distances (e.g., the image of an airplane on the lenses of a first co-player’s HMD), including while co-players were in each other’s lines of sight. *Id.* Further, limitation 8[c]’s required configuration of an SLP behind an image along a user’s line of sight is disclosed by Zhou’s Figure 1. Zhou Fig. 1 (showing a user in an HMD looking through the screen that displays images, to coordinates for a virtual sound source behind that screen) (reproduced in §VII.B); Roginska ¶¶392-96

Accordingly, POSAs combining Zhou’s teachings with Shivappa’s—e.g., to improve depth perception and enable a more realistic environment and immersive feeling (§VII.C)—would have known and been motivated to implement sound processing using the configurations of far-field SLPs and near-field images required by limitation 8[c], which renders this limitation obvious. Roginska ¶¶392-96.

5. 8[e]: DSP Processes Sound with Far-Field HRTFs While Electronic Device is at Near-Field Location

Limitation 8[e], which requires that the DSP recited in limitation 8[d] “processes the sound for the image with far-field head-related transfer functions (HRTFs) while the electronic device is located at the near-field distance from the head of the user,” would have been obvious over Shivappa, alone or in view Zhou.

First, as discussed, Shivappa disclosed and rendered obvious using a DSP to process sound for images, including to generate binaural sound. §§V.D.3-V.D.4. Second, Shivappa, alone or in view of Zhou, rendered obvious a DSP processing sound for images at far-field locations, which POSA would have known could and would be rendered with far-field HRTFs. §VII.D.4. Third, Shivappa’s HMDs are located at near-field distances from users’ heads when they are used, including during the processing and rendering of virtual images and audio. §VII.D.2; Shivappa, Fig. 2; Roginska ¶¶399-405.

E. Dependent Claims 9 and 10

Claims 9 and 10 are unpatentable for the same reasons as claim 8 and as further explained below.

1. Claim 9: AR Image and Electronic Glasses

Claim 9 recites: “The electronic device of claim 8, wherein the image is an augmented reality (AR) image, and the display is included with electronic glasses worn on the head of the user.”

Shivappa disclosed that its audio processing system can be used in an HMD of an “augmented reality (AR) system” with “a display screen for displaying visual information,” i.e., AR images in an AR system. Shivappa, 5:37-42. Shivappa also disclosed that its HMD “may be a pair of virtual reality goggles.” *Id.* 13:64. These disclosures would have rendered claim 9 obvious, including claim 9’s recitation of “electronic glasses,” which like “goggles” were a well-known variant of VR and AR HMDs. Roginska ¶¶406-411; The VR Book, Fig. 2.1 (depicting “head-mounted displays and viewers over time”), Fig. 3.3 (presenting images of an Oculus Rift next to a “CastAR” HMD, i.e., AR “electronic glasses”) (reproduced in §IV.E.1).

2. Claim 10: VR Image and HMD

Claim 10 recites: “The electronic device of claim 8 , wherein the image is a virtual reality (VR) image, and the display is included with a head mounted display worn on the head of the user.”

Shivappa disclosed that its audio processing system can be used in an HMD of “a virtual reality (VR) system” with “a display screen for displaying visual information,” i.e., VR images in a VR system. Shivappa, 5:37-42; *id.* 3:14-23, 13:46-67, Fig. 2; Roginska ¶¶412-415.

F. Independent Claim 15

1. 15[pre]: An Electronic System

The preamble of claim 15 recites: “An electronic system, comprising.” Shivappa discloses an electronic system. §§VII.A, VII.D.1; Roginska ¶416.

2. 15[a]: Image Within A Near-Field Distance

Limitation 15[a], which recites, “a display that displays an image that is located within a near-field distance from a head of a user viewing the image,” would have been obvious for the same reasons as limitation 8[a]. §VII.D.2. Additionally, POSAs would have known that Shivappa presented images at near-field distances from users “viewing th[ose] image[s].” *Id.*; Roginska ¶417.

3. 15[b]: Processor for Processing Sound for the Image

Limitation 15[b], which recites, “one or more processors that process sound for the image,” would have been obvious for the same reasons as limitations 8[b] and 8[d]. §VII.D.3; Roginska ¶418; *see also* note 8.

4. 15[c]: HRTFs to Generate Binaural Sound

Limitation 15[c], which requires processing sound for an image “with head-related transfer functions (HRTFs) to generate binaural sound,” would have been

obvious for the same reasons as limitation 8[c] and 8[e]. §§VII.D.4-VII.D.5. POSAs would have known that the HRTFs are used to generate binaural sound. *Id.*; Roginska¶419.

5. 15[d]: Far-Field SLP Behind a Near-Field Image Along a Line of Sight

Limitation 15[d], which requires that the processed and generated binaural sound “has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance,” would have been obvious for the same reasons as limitation 8[c]. §VII.D.4; Roginska¶420.

6. 15[e]: SLP Stays at Far-Field Distance While User’s Head Moves and Image Remains at Near-Field Distance

Limitation 15[e] requires that “the coordinate location of the binaural sound continues to occur at the far-field distance located behind the image when the head of the user moves away from viewing the image and the image remains within the near-field distance from the head of the user.”

A coordinate location of binaural sound occurring at a far-field distance located behind an image would have been obvious over Shivappa, alone or in view of Zhou, for reasons discussed. §§VII.D.4, VII.F.5. The additional recitations of

limitation 15[e] would have been obvious over Shivappa, alone or in view of Zhou, for the following reasons.

Shivappa disclosed that its “visuals and audio output” “may change based on movement of the user (e.g., movement of the user’s head).” Shivappa, 5:45-48. “To illustrate, if the user’s head moves or rotates, or if the user moves or rotates in the virtual environment, the processor 106 modifies the input spatialized audio file such that a perceived direction or distance of one or more sound sources in a 3D space is modified based on the movement.” *Id.* 8:38-42, 6:37-47. Thus, “if the user is playing a game and a car drives to the left of the user in the game, the audio output enables the user to perceive a sound of a car as coming from the user’s left side. If the user turns to the right in the game such that the car is behind the user, the audio processing device processes the spatialized audio signal to cause the audio output to change such that the user experiences the sound of the car as coming from behind the user.” *Id.* 3:23-31. And Zhou’s 3D audio system operates similarly. For example, when a player in Zhou’s game “walks closer to the sound sources (the castle and the princess therein or the co-player), the sound becomes louder while it becomes softer when the player walks away.” Zhou 1006-07.

In view of these disclosures, POSAs would have understood that Shivappa’s systems, with or without Zhou’s 3D audio techniques, could and would maintain binaural sound occurring at far-field distances located behind near-field images

displayed on the lenses of Shivappa’s AR or VR HMDs, even when a user’s head moves away from viewing that image, so the user could continue to properly localize such sound sources. Roginska ¶¶421-25. As noted in Grounds 1 and 2, this type of sound processing—accounting for head movement while maintaining binaural sound source locations—was well known to POSAs before May 2017. §IV.F.6; Begault Dep. 250:2-9.

Further, Shivappa’s near-field images, e.g., a car, character, or conference speaker, would obviously remain within “near-field distance[s] from the head of the user”—on Shivappa’s HMD’s lenses—until, for example, the user’s head movement caused them to drop from the field of view. Roginska ¶¶423.

G. Dependent Claims 16 and 17

Claims 16 and 17 are unpatentable for the same reasons as claim 15 and as further explained below.

1. Claim 16: AR Display and Image

Claim 16, which recites, “The electronic system of claim 15, wherein the display is an augmented reality (AR) display worn on the head of the user viewing the image, and the image is an AR image,” would have been obvious for the same reasons as claim 9. §VII.E.1. And POSAs would have known that users wearing AR displays, e.g., AR HMDs, would view AR images. *Id.*; Roginska ¶¶425-28.

2. Claim 17: VR Display and Image

Claim 17, which recites, “The electronic system of claim 15, wherein the display is a virtual reality (VR) display worn on the head of the user viewing the image, and the image is an VR image,” would have been obvious for the same reasons as claim 10. §VII.E.2. And POSAs would have known that users wearing VR displays, e.g., VR HMDs, would view VR images. *Id.*; Roginska ¶¶429-30.

VIII. Secondary Considerations

Petitioners are unaware of any secondary considerations that affect—let alone overcome—the obviousness cases set out above. Should Patentee proffer any secondary considerations evidence in its preliminary response, Petitioners should be permitted to reply.

IX. The Board Should Reach the Merits of This Petition

The Board should reach the merits of this petition and should not exercise its discretion to deny based on 35 U.S.C. § 314(a) or § 325(d). None of the prior art at issue in this petition was cited or relied upon during prosecution of the ’737 patent. As for the district court case, the court is currently deciding a motion to transfer. *Eight kHz v. Meta Platforms*, No. 6:22-cv-00575, Dkt. 46 (W.D. Tex.).

Petitioners reserve the right to address any discretionary issues raised by Patentee based on the facts and law as they stand at that time.

X. Mandatory Notices under 37 C.F.R. § 42.8

A. Real Parties-in-Interest

Petitioners are the only real parties-in-interest.

B. Related Matters

Patent Center indicates that U.S. Patent Application No. 17/703,528, 16/246,501 (now U.S. Patent No. 10,219,095), 16/354,182 (now U.S. Patent No. 10,791,409), 17/062,637 (now U.S. Patent No. 10,805,758), and 17/703,528, which is currently pending, are related to the '836 patent, which issued from Application No. 17/062,637.

The '836 patent is at issue in *Eight KHZ LLC v. Meta Platforms, Inc.*, No. 6:22-cv-00575 (W.D. Tex.).

C. Grounds for Standing

Petitioners certify that the '836 patent is eligible for IPR and they are not barred from requesting same.

D. Lead and Backup Counsel and Service Information

Pursuant to 37 C.F.R. §§ 42.8(b)(3), 42.8(b)(4), and 42.10(a), Petitioners designate the following lead counsel:

- Inge A. Osman (Reg. No. 74,480), inge.osman@lw.com, Latham & Watkins LLP; 555 Eleventh Street, NW, Ste. 1000; Washington, DC 20004-1304; 202.637.3308 (Tel.); 202.637.2201 (Fax).

Petitioners also designate the following backup counsel:

Petition for *Inter Partes* Review of USP 11,290,836

- Christopher W. Henry (Reg. No. 60,907), christopher.henry@lw.com, Latham & Watkins LLP; 200 Clarendon Street; Boston, MA 02116; 617.880.4550 (Tel.); 617.948.6001 (Fax).
- Jonathan M. Strang (Reg. No. 61,724), jonathan.strang@lw.com, Latham & Watkins LLP; 555 Eleventh Street, NW, Ste. 1000; Washington, DC 20004-1304; 202.637.2362 (Tel.); 202.637.2201 (Fax).
- Linfong Tzeng (Reg. No. 65,140), linfong.tzeng@lw.com, Latham & Watkins LLP; 140 Scott Drive, Menlo Park, CA 94025; 650.463.3092 (Tel.); 650.463.2600 (Fax).
- Kimberly Q. Li (Reg. No. 74,713), kimberly.li@lw.com, Latham & Watkins LLP; 200 Clarendon Street; Boston, MA 02116; 617.880.4529 (Tel.); 617.948.6001 (Fax), and
- Razi Safi (Reg. No. 75,594), razi.safi@lw.com, Latham & Watkins LLP; 555 Eleventh Street, NW, Ste. 1000; Washington, DC 20004-1304; 202.637.2120 (Tel.); 202.637.2201 (Fax).

Pursuant to 37 C.F.R. § 42.10(b), a Power of Attorney from Petitioners is attached. Petitioners consent to electronic service.

E. IPR Fee

The Director is authorized to charge the requisite fee to Deposit Account No. 506269.

XI. Conclusion

IPR should be instituted, and each challenged claim should be found unpatentable.

Respectfully submitted,

Dated: June 8, 2023

By: / Inge A. Osman /

Inge A. Osman (Reg. No. 74,480)
inge.osman@lw.com
Latham & Watkins LLP
555 Eleventh Street, NW, Ste. 1000
Washington, DC 20004-1304
Telephone: 202.637.2200
Fax: 202.637.2201

*Counsel for Petitioners
Meta Platforms, Inc., Meta Platforms
Technologies, LLC, and Twisted Pixel
Games, LLC*

Claims Appendix

This “Claims Appendix” reproduces the challenged claims, with bracketed identifiers added for cross-referencing.

Claim 8:

[pre] An electronic device, comprising:

[a] a display that displays an image that occurs within a near-field distance from a head of a user;

[b] and one or more processors that process sound for the image

[c] to generate binaural sound; that has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance,

[d] wherein the one or more processors include a digital signal processor (DSP)

[e] that processes the sound for the image with far-field head-related transfer functions (HRTFs) while the electronic device is located at the near-field distance from the head of the user.

Claim 9:

The electronic device of claim 8, wherein the image is an augmented reality (AR) image, and the display is included with electronic glasses worn on the head of the user.

Claim 10:

The electronic device of claim 8, wherein the image is a virtual reality (VR) image, and the display is included with a head mounted display worn on the head of the user.

Claim 11:

The electronic device of claim 8 further comprising:

one or more sensors that detect rotation of the electronic device, wherein the binaural sound automatically switches to stereo sound in response to the one or more sensors detecting the rotation of the electronic device.

Claim 12:

The electronic device of claim 8, wherein the binaural sound automatically switches to stereo sound in response to the electronic device detecting a change to a head orientation of the user that exceeds a threshold value in an azimuth direction.

Claim 13:

The electronic device of claim 8, wherein the image is an augmented reality (AR) image of an intelligent user assistant (IUA) that is located less than one meter away from the head of the user, and a voice of the IUA occurs at a coordinate location behind the AR image of the IUA.

Claim 14:

The electronic device of claim 8, wherein the image is a virtual reality (VR) image of an intelligent user assistant (IUA) that is located less than one meter away from the head of the user, and a voice of the IUA occurs at a coordinate location behind the VR image of the IUA.

Claim 15:

[pre] An electronic system, comprising:

[a] a display that displays an image that is located within a near-field distance from a head of a user viewing the image;

[b] and one or more processors that process sound for the image

[c] with head-related transfer functions (HRTFs) to generate binaural sound,

[d] that has a sound localization point (SLP) with a coordinate location that occurs at a far-field distance located behind the image and along a line-of-sight from the user to the image while the display displays the image to the user at the near-field distance,

[e] wherein the coordinate location of the binaural sound continues to occur at the far-field distance located behind the image when the head of the user moves away from viewing the image and the image remains within the near-field distance from the head of the user.

Claim 16:

The electronic system of claim 15, wherein the display is an augmented reality (AR) display worn on the head of the user viewing the image, and the image is an AR image.

Claim 17:

The electronic system of claim 15, wherein the display is a virtual reality (VR) display worn on the head of the user viewing the image, and the image is an VR image.

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 13,830 words, which is the sum of 13,780 words calculated by Microsoft Word's word-count feature and 50 words hand-counted in the figures. This total excludes the cover page, signature block, and the parts of the Petition exempted by 37 C.F.R. § 42.24(a)(1).

By: / Inge A. Osman /

Inge A. Osman (Reg. No. 74,480)
inge.osman@lw.com
Latham & Watkins LLP
555 Eleventh Street, NW, Ste. 1000
Washington, DC 20004-1304
Telephone: 202.637.2200
Fax: 202.637.2201

*Counsel for Petitioners
Meta Platforms, Inc., Meta Platforms
Technologies, LLC, and Twisted Pixel
Games, LLC Meta Platforms, Inc.*

CERTIFICATE OF SERVICE

The undersigned certifies that a complete copy of this Petition for *Inter Partes* Review of U.S. Patent No. 11,290,836 and all Exhibits and other documents filed together with this Petition were served on the official correspondence address for the patent shown in Patent Center:

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

via FEDERAL EXPRESS next business day delivery, on June 8, 2023.

By: / Inge A. Osman /

Inge A. Osman (Reg. No. 74,480)
inge.osman@lw.com
Latham & Watkins LLP
555 Eleventh Street, NW, Ste. 1000
Washington, DC 20004-1304
Telephone: 202.637.2200
Fax: 202.637.2201

Counsel for Petitioners
Meta Platforms, Inc., Meta Platforms
Technologies, LLC, and Twisted Pixel
Games, LLC