

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.,
Petitioner,

v.

EIGHT KHZ, LLC,
Patent Owner.

Case IPR2023-01019
Patent 10,368,179 B1
and
Case IPR2023-01020
Patent 10,448,184 B1

**DECLARATION OF DR. AGNIESZKA ROGINSKA
IN SUPPORT OF PETITIONERS' REPLY**

Table of Contents

I.	Introduction and Qualifications	1
II.	Summary of Materials Reviewed and Considered	1
III.	Legal Understanding.....	7
IV.	Claim Construction.....	7
A.	Claims 7 and 14 of both patents: “environment” is not limited to physical environments.....	7
B.	Claim 1 of both patents: “3D virtual environment” is not a sound-only environment.....	12
C.	Claims 7 and 14 of the ’179 patent: “Determining, with the WED, when the user is leaving the environment” does not require independently determining	19
V.	The ’184 Patent.....	20
A.	’184 Patent - Ground 1: Claims 1, 3-4, 6-10, and 14-19 Are Unpatentable Over Clement, Osman, and Sauk.....	20
1.	Relevance in Sauk is not merely a function of the user’s decision.	20
2.	Sauk does not teach away from the combination	21
3.	’184 Patent - Independent Claims 1, 7, and 14.....	25
(1)	lower[ing]/reducing a volume of the binaural sound to warn (claims 1, 7, and 14).....	25
(2)	filtering to alter a spectral cue to warn (claim 1), degrading a spatial cue of the sound to alert (claim 7), or altering the processing of the binaural sound to alert (claim 14).....	26
4.	’184 Patent - Dependent Claims 15-19.....	31
(1)	Claim 15. The method of claim 14, wherein the DSP alters processing of the binaural sound to alert the user by changing a spatial cue that includes at least one of an interaural time difference (ITD), an interaural level difference (ILD), and the HRTFs.....	31

(2)	Claim 16. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by stopping, by the DSP, the processing of the sound with the HRTFs.	32
(3)	Claim 17. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by removing interaural time differences (ITDs) of the binaural sound being provided to the user.	32
(4)	Claim 18. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by eliminating one or more spatial cues in the binaural sound being provided to the user.	32
(5)	Claim 19. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by switching the binaural sound playing to the user to one of stereo sound and mono sound.....	33
5.	'184 Patent - Dependent Claims 8-10.....	33
(1)	Claim 8. The method of claim 7, wherein the spatial cue includes an interaural time difference (ITD).	33
(2)	Claim 9. The method of claim 7 further comprising: displaying, by the WED, the virtual image while the user plays a game that includes a character as the virtual image; and switching a voice of the character from the binaural sound to stereo sound when the WED no longer displays the virtual image during the game.	33
(3)	Claim 10. The method of claim 7 further comprising: alerting the user that the user is leaving the environment by switching the binaural sound to stereo sound.	34
6.	'184 Patent - Dependent Claims 3-4, and 6	34
(1)	Claim 3. The method of claim 1, wherein the warning occurs when the user is leaving a virtual area where the virtual image is located while remaining in the 3D virtual environment.	34

(2)	Claim 4. The method of claim 1 further comprising: tracking, with head tracking in the HMD, head movements of the user; and switching the sound from the binaural sound to stereo sound when a face of the user is directed away from the virtual image such that the virtual image is no longer visible to the user.....	35
(3)	Claim 6. The method of claim 1 further comprising: switching the sound from the binaural sound to stereo sound when the virtual image vanishes from the 3D virtual environment.	37
B.	'184 Patent - Ground 2: All Claims Are Unpatentable Over Lyren and Reif.....	37
1.	The '184 Patent's claims are not supported by its parent, the '140 Application	37
(1)	The '140 Application does not support warning or alerting the user.....	42
(2)	The '140 Application does not support warning the user that the user is leaving the 3D virtual environment (claim 1)	45
(3)	The '140 Application does not support alerting the user that the user is leaving "the environment" (claim 7) or leaving "a virtual image in the environment" (claim 14).....	50
2.	Combination of Lyren and Reif.....	54
VI.	The '179 Patent.....	55
A.	'179 Patent – Ground 1: Claims 1, 3-4, 6-10, and 14-19 are obvious over Clement, Osman, and Sauk.	55
1.	'179 Patent – Sauk does not teach away from the combination.....	55
2.	'179 Patent - Independent Claims 1, 7, and 14.....	55
(1)	lower[ing]/reducing a volume of the binaural sound to warn (claims 1 and 14).....	55
(2)	determining/sensing, with a WED, when a user is leaving an environment/3D virtual environment (claims 1, 7, 14)	55

(3)	degrading audio cues to prevent external localization of binaural sound to alert (claim 7)/filtering by the DSP, the binaural sound to warn (claim 1).....	58
3.	'179 Patent – Dependent claims 3-4, 6, 8-10, and 15-19	59
B.	'179 Patent – Ground 2: All claims are obvious over Lyren and Reif.	60
1.	The '179 Patent's claims are not supported by its parent, the '140 Application	60
(1)	The '140 Application does not support warning or alerting the user by modifying binaural sound and/or reducing the volume.	60
(2)	The '140 Application does not support warning the user that the user is leaving the 3D virtual environment (claim 1)	60
(3)	The '140 Application does not support alerting the user that the user is leaving “the environment” (claim 7) or leaving “a virtual image in the environment” (claim 14).....	60
(4)	The '140 Application does not support determining/sensing, with the WED/HMD, when the user is leaving the environment/3D virtual environment	61
2.	Combination of Lyren and Reif.....	62
VII.	Conclusion	62

I. Introduction and Qualifications

1. I, Agnieszka Roginska, have been retained by Meta Platforms, Inc., Meta Platforms Technologies, LLC, and Twisted Pixel Games, LLC (“Petitioners”) to provide my opinion concerning the validity of Claims 1-20 of U.S. Patent No. 10,368,179 (“the ’179 Patent”).

2. I previously provided a declaration dated June 6, 2023 with my opinions related to the level of skill of ordinary artisans, the state of the art described in the ’179 and ’184 patents at the time of the effective filing date of the two patents, and the disclosure of the prior art relative to the claims of the ’179 and ’184 patents. I understand that my original declaration was submitted as Exhibit 1003. I hereby incorporate Exhibit 1003 into this declaration by reference.

3. My experience and qualifications are set forth in Exhibit 1003 and in my curriculum vitae, which I understand was submitted as Exhibit 1004.

4. I have been asked to provide additional expert opinions regarding these matters in response to certain arguments in the Patent Owner Response and certain opinions of Patent Owner’s witness, Dr. Durand R. Begault.

II. Summary of Materials Reviewed and Considered

5. The opinions contained in this declaration are based on the documents I reviewed and my knowledge and professional judgment. In forming the opinions

Declaration for *Inter Partes* Review of USP 10,368,179 and 10,448,184

expressed in this declaration, while drawing on my experience, I reviewed the following documents:

Ex.	Description
1001-179	U.S. Patent No. 10,368,179 (“the ’179 Patent”)
1001-184	U.S. Patent No. 10,448,184 (“the ’184 Patent”)
1003	Declaration of Dr. Agnieszka Roginska
1004	<i>Curriculum Vitae</i> of Dr. Agnieszka Roginska
1005	U.S. Patent No. 9,690,374 (“Clement”)
1006	U.S. Patent No. 9,037,468 (“Osman”)
1007	U.S. Patent No. 7,876,903 (“Sauk”)
1008	U.S. Patent Application Publication No. 2017/0188168 (“Lyren”)
1009	U.S. Patent Application Publication No. 2018/0373412 (“Reif”)
1010	U.S. Patent Application Publication No. 2011/0206209 (“Ojala”)
1011	U.S. Patent No. 1,624,486 (“Fletcher”)
1012	U.S. Patent No. 1,855,149 (“Jones”)
1013	U.S. Patent No. 7,330,197 (“Kobayashi”)
1014	U.S. Patent No. 7,710,654 (“Ashkenazi”)
1015	U.S. Patent No. 8,520,872 (“Jang”)
1016	U.S. Patent No. 8,520,873 (“Mahabub”)
1017	U.S. Patent No. 9,041,622 (“McCulloch”)
1018	U.S. Patent No. 9,367,811 (“Ramachandran”)
1019	U.S. Patent No. 9,544,706 (“Hirst”)
1020	U.S. Patent No. 9,578,440 (“Otto”)
1021	U.S. Patent No. 9,674,507 (“Pirchheim”)
1023	U.S. Patent No. 9,767,618 (“Najaf-Zadeh”)
1025	U.S. Patent No. 10,070,239 (“Romigh”)
1027	U.S. Patent No. 10,095,031 (“Katz”)
1028	File History for U.S. Patent Application 16/106,140
1029	U.S. Patent No. 10,509,466 (“Miller”)
1030	U.S. Patent Application Publication No. 2012/0183161 (“Agevik”)
1031	U.S. Patent Application Publication No. 2016/0124502 (“Sawyer”)
1032	U.S. Patent Application Publication No. 2016/0292924

Declaration for *Inter Partes* Review of USP 10,368,179 and 10,448,184

	(“Balachandreswaran”)
1033	U.S. Patent Application Publication No. 2017/0105083 (“Nair”)
1034	U.S. Patent Application Publication No. 2017/0115488 (“Ambrus”)
1035	Korean Patent Application Publication No. KR 10-2011-0022252, including English Translation and Certification (“Moon”)
1036	Christoph Anthes et al., <i>State of the Art of Virtual Reality Technology</i> , 2016 IEEE AEROSPACE CONF. 1-19 (Mar. 2016) (“Anthes”)
1037	<i>Average Cost of RAM per Gigabyte Chart</i> , HumanProgress (“HumanProgress”)
1038	DURAND R. BEGAULT, 3-D SOUND FOR VIRTUAL REALITY AND MULTIMEDIA (NASA/TM-2000-209606) (Aug. 2000) (“Begault”)
1039	Tom Emrich, <i>Virtual Reality North: A Look at the Major VR Players in Canada</i> , betakit.com (July 15, 2015), betakit.com/virtual-reality-north-a-look-at-the-major-vr-players-in-canada/ (“Emrich”)
1040	Braxton Boren, <i>History of 3D Sound</i> (Am. Univ., Wash. D.C. 2018), https://dra.american.edu/islandora/object/auislandora%3A77787/datastream/PDF/view (“Boren”)
1041	Douglas S. Brungart et al., <i>The Interaction Between Head-Tracker Latency, Source Duration, and Response Time in the Localization of Virtual Sound Sources</i> , PROC. ICAD 04- 10TH MEETING INT’L CONF. AUDITORY DISPLAY 1-7 (July 6-9, 2004) (“Brungart 2004”)
1042	Douglas S. Brungart et al., <i>The Detectability of Headtracker Latency in Virtual Audio Displays</i> , PROC. ICAD 05-11TH MEETING INT’L CONF. AUDITORY DISPLAY 37-42 (July 6-9, 2005) (“Brungart 2005”)
1043	E. Colin Cherry, <i>Some Experiments on the Recognition of Speech, with One and with Two Ears</i> , 25 J. ACOUSTICAL SOC’Y AM. 975 (1953) (“Cherry”)
1044	Christof Faller, <i>Coding of Spatial Audio Compatible with Different Playback Formats</i> , AES 117th CONVENTION 1-12 (Oct. 28-31, 2004) (“Faller”)
1045	Mohammad Dashti et al., <i>Traffic Management: A Holistic Approach to Memory Placement on NUMA Systems</i> , ASPLOS’ 1-13 (Mar. 16-20, 2013) (“Dashti”)

1046	S.S. Fisher et al., <i>Virtual Interface Environment Workstations</i> , 32 PROC. HUM. FACTORS SOC’Y 32ND ANN. MEETING 1988 91-95 (Feb. 1988) (“Fisher”)
1047	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) (“Fohl”)
1048	MARK GRIMSHAW, GAME SOUND TECHNOLOGY AND PLAYER INTERACTION: CONCEPTS AND DEVELOPMENTS 22 (2011) (“Grimshaw”)
1049	Eric Grownney & Agnieszka Roginska, <i>Effect of Spatial Audio Processing on Attention Redirection and Speech Intelligibility in Multi-Talker Communication</i> , AES 152ND CONVENTION 1-10 (May 2022) (“Grownney”)
1050	Maxim Liadov, <i>A New Generation of Xtreme Fidelity Processors and Sound Cards</i> , iXBT Labs (Aug. 22, 2005) (“Liadov”)
1051	JESSE SCHELL, THE ART OF GAME DESIGN: A BOOK OF LENSES (2008) (“Schell”)
1052	Christoph Lameter, <i>Local and Remote Memory: Memory in a Linux/NUMA System</i> , SGI (June 20, 2006) (“Lameter 2006”)
1053	Christoph Lameter, <i>NUMA (Non-Uniform Memory Access): An Overview</i> , ACMqueue 1-12 (2013) (“Lameter 2013”)
1054	Matteo Marucci et al., <i>The Impact of Multisensory Integration and Perceptual Load in Virtual Reality Settings on Performance, Workload and Presence</i> , 11 NATURE SCI. REPS. 1 (2021) (“Marucci”)
1055	Joel D. Miller et al., <i>Latency Measurements of a Real-Time Virtual Acoustic Environment Rendering System</i> , PROC. 2003 INT’L CONF. AUDITORY DISPLAY 111-114 (July 6-9, 2003) (“Miller”)
1056	Francisco Cheng, <i>Mobile Processors 101: Why Smartphones are Smarter with an All-In-One Processor</i> , QUALCOMM ONQ BLOG (June 12, 2013) (“Cheng”)
1057	Masayuki Morimoto & Yoichi Ando, <i>On the Simulation of Sound Localization</i> , (E) 1 J. ACOUSTIC SOC’Y JAPAN 167 (1980) (“Morimoto”)

1058	RESERVED
1059	Stephen Paul, <i>Binaural Recording Technology: A Historical Review and Possible Future Developments</i> , 95 ACTA ACUSTICA UNITED WITH ACUSTICA 767 (2009) (“Paul”)
1060	Ben Kuchera, <i>Google Cardboard Adds ‘Inexpensive’ Spatial Audio to its SDK</i> , POLYGON (Jan. 13 2016, 12:00pm PST) (“Kuchera”)
1061	Paul Sauk & Duane Satorius, <i>Creating a Multi-Dimensional Communication Space to Improve the Effectiveness of 3-D Audio</i> , IEEE MILITARY COMM’CNS CONF. 1-7 (Oct. 29-31, 2007) (“Sauk 2007”)
1062	Mel Slater & Martin Usoh, <i>Presence in Immersive Virtual Environments</i> , IEEE VIRTUAL REALITY ANN. INT’L SYMP. 90-96 (Sept. 18-22, 1993) (“Slater”)
1063	JAKA SODNIK & SAŠO TOMAŽIČ, SPATIAL AUDITORY HUMAN-COMPUTER INTERFACES, SPRINGER BRIEFS COMPUT. SCI. (2015) (“Sodnik”)
1064	Paul James, <i>Sulon Cortex ‘Augmented Virtual Reality’ Headset, Interview with Don Henderson</i> , ROAD TO VR (Feb. 5, 2015) (“James”)
1065	Mona Lalwani, <i>Surrounded by Sound: How 3D Audio Hacks Your Brain</i> , THE VERGE (Feb. 12, 2015, 10:30am) https://www.theverge.com/2015/2/12/8021733/3d-audio-3dio-binaural-immersive-vr-sound-times-square-new-york (“Lalwani”)
1066	Dave Smith, <i>Three-Dimensional Audio Makes Virtual Reality so Much Better It’s Crazy</i> , BUSINESS INSIDER (Apr. 3, 2015, 1:06pm), https://www.businessinsider.com/virtual-reality-the-importance-of-3d-positional-audio-2015-4 (“Smith”)
1067	JASON JERALD, THE VR BOOK: PERCEPTION AND INTERACTION DESIGN FOR VIRTUAL REALITY (2015) (“VR Book”)
1068	Elizabeth M. Wenzel et al., <i>Development of a Three-Dimensional Auditory Display System</i> , 20 SIGCHI BULL. 52 (Oct. 1988) (“Wenzel 1988a”)
1069	Elizabeth M. Wenzel et al., <i>A Virtual Display System for Conveying</i>

	<i>Three-Dimensional Acoustic Information</i> , PROC. HUM. FACTORS SOC'Y 32ND ANN. MEETING 86-90 (1988) ("Wenzel 1988b")
1070	RESERVED
1071	Elizabeth M. Wenzel et al., <i>Localization with Non-Individualized Virtual Acoustic Display Cues</i> , PROC. SIGCHI CONF. HUM. FACTORS COMPUT. SYS. 351-359 (Apr. 1991) ("Wenzel 1991")
1072	Elizabeth M. Wenzel, <i>Localization in Virtual Acoustic Displays</i> , 1 PRESENCE TELEOPERATORS VIRTUAL ENV'TS 80 (Feb. 1992) ("Wenzel 1992")
1073	Qualcomm, <i>What's So Special About the Digital Signal Processor?</i> , QUALCOMM ONQ BLOG (Dec. 5, 2013) ("OnQ")
1074	Frederic L. Wightman & Doris J. Kistler, <i>Headphone Simulation of Free-Field Listening. II: Psychophysical Validation</i> , 85 J. ACOUSTIC SOC'Y AM. 868 (1989) ("Wightman")
1075	Jiann-Rong Wu et al., <i>Head Motion and Latency Compensation on Localization of 3D Sound in Virtual Reality</i> , PROC. ACM SYMP. VIRTUAL REALITY SOFTWARE TECH. 15-20 (Sept. 1997) ("Wu")
1076	Dmitry N. Zotkin et al., <i>Customizable Auditory Displays</i> , PROC. 2002 INT'L CONF. AUDITORY DISPLAY 1-10 (July 2-5, 2002) ("Zotkin")
1077	Gr. Brit. Provisional Patent Specification No. 394,325 ("Blumlein")
1078	Linley Gwennap, <i>Qualcomm Extends Hexagon DSP</i> , MICROPROCESSOR REP. 1-4 (Aug. 26, 2012) ("Gwennap")
1079	Stanley P. Lipshitz, <i>Stereo Microphone Techniques.....Are the Purists Wrong?</i> 34 J. AUDIO ENG'G SOC'Y 716 (1986) ("Lipshitz")
1080	U.S. Patent No. 9,851,786 ("Pedrotti")
1081	U.S. Patent No. 9,787,846 ("Srinivasan")
1082	U.S. Patent Application Publication No. 2015/0094142 ("Stafford")
1083	U.S. Patent Application Publication No. 2019/0043259 ("Wang")
1085	Oculus Connect 5 Spatial Audio for Oculus Quest and Beyond video clips (May 12, 2023)
1086	Oculus Connect 5 Spatial Audio for Oculus Quest and Beyond video clip transcriptions (May 12, 2023)
1087	Oculus Connect 5 Keynote Day 01 video clips (May 16, 2023)
1088	Oculus Connect 5 Keynote Day 01 video clip transcriptions (May 16,

	2023)
1093	U.S. Patent No. 5,173,944 (“Begault ‘944”)
1094	U.S. Patent No. 5,438,623 (“Begault ‘623”)
1147	Transcript of Deposition of Dr. Durand R. Begault
1148	Errata of Transcript of Deposition of Dr. Agnieszka Roginska
1149	D. R. Begault et al., “Headphone and head-mounted visual displays for virtual environments,” 15 th International Conference of the Audio Engineering Society. (1998)
1150	U.S. Patent Application Publication 2017/0148206 (“Donner”)
2006	Declaration of Dr. Durand R. Begault
2007	Curriculum Vitae of Dr. Durand R. Begault
2012	Supplemental Declaration of Dr. Durand R. Begault
2013	Updated CV of Dr. Durand R. Begault
2014	Transcript of Deposition of Dr. Agnieszka Roginska
2015	Excerpt from Merriam Webster’s Dictionary
Paper 10	Institution Decision (both ’184 and ’179)

III. Legal Understanding

6. I am not an attorney. However, I have been informed by counsel of certain legal standards and have applied those standards, described in Exhibit 1003, in formulating my opinions.

IV. Claim Construction

A. Claims 7 and 14 of both patents: “environment” is not limited to physical environments

7. Claim 7 of the ’184 patent recites, among other things, “alerting the user that the user is leaving *the environment*.” Claims 7 and 14 of the ’179 patent also recite this same language.

8. Claim 14 of the '184 patent recites “alerting the user that the user is leaving the virtual image in the environment.”

9. Dr. Begault states that “a PHOSITA would have understood the term ‘environment’ as used in claims 7 and 14 of the ‘179 and ‘184 Patents to mean ‘the physical surroundings in which a user operates.’” Ex.2012 ¶22. He further states that “[a] PHOSITA would also recognize that the concept of ‘leaving the environment’ to refer to the user moving from one physical environment condition to another physical environment condition.” Ex.2012 ¶ 25.

10. I disagree with Dr. Begault’s interpretation of “environment” and “leaving the environment” in claims 7 and 14 of the '179 and '184 patents. Specifically, the plain meaning of environment is not limited to the *physical* surroundings, but can also be virtual surroundings. In other words, the plain and ordinary meaning of environment encompasses at least both physical *and virtual* environments.

11. Nothing in the claims limits the “environment” to a physical environment. To the contrary, all of these claims recite “displaying, [with a wearable electronic device], a virtual image in an environment.” Thus, the claims explicitly require the environment to be able to support virtual images, encompassing both virtual and augmented realities. The inclusion of virtual and augmented realities is consistent with claim 12, which recites “the environment is a physical room where

the user is located, the virtual image is an augmented reality (AR) character that speaks to the user while the user plays a game.” I note that the specifications support that virtual images are displayed in augmented or virtual realities. *See* ’184 patent, 11:4-5 (“images that are part of an augmented or virtual reality”); 12:4-9 (“virtual image[] or images per an augmented reality”); 21:22-25 (image “appears in augmented reality or virtual reality”); 28:31-34 (“visually rendered character vanishing from an augmented or virtual environment”). At the very least, the recitation of “a virtual image” in these claims cannot in itself preclude “environment” from being a “virtual environment” because virtual images can be displayed in virtual environments. I note this is also shown by claim 1 of both patents with requires “displaying . . . a three-dimensional (3D) virtual environment that *includes a virtual image.*” And claim 1 further shows that an “environment” is not limited to physical environments because it explicitly recites a 3D “virtual environment.”

12. I also note that Dr. Begault does not rely on the surrounding claim language to justify his interpretation of environment.

13. The specification also supports that “environment” is not limited to a physical environment. For example, the specification repeatedly describes environments as being physical *or virtual*. ’184 patent, 28:16 (“a physical or a virtual environment”); 28:20 (“physical or virtual environment”); 28:24 (“physical or

virtual environment”); 28:28-29 (“physical or virtual environments”); 28:36-37 (“physical or virtual environment”). Further, the use of “3D virtual environment” in claim 1 of both patents supports that the term “environment” without modifiers must include such virtual environments and not be limited to only physical environments.

14. I note that Dr. Begault does not rely on any special definition of “environment” from the specifications of the patents to support his interpretation. Although Dr. Begault states that his interpretation is “consistent with the description in the specification” he does not cite to any particular passage of the specification. Ex.2012 ¶ 22.

15. Dr. Begault instead cites to three dictionary definitions. Ex.2012 ¶ 21. However, these definitions do not support limiting “environment” to a physical environment. Dr. Begault does not cite to any exhibits reproducing the definitions he cites, but instead provides links to the home page of three dictionaries. However, I understand that a definition from Merriam-Webster has been submitted as Exhibit 2013.

16. Nothing in the Merriam-Webster definition (“the circumstances, objects, or conditions by which one is surrounded”) appears to limit “environment” to a user’s physical surroundings (as opposed to virtual surroundings). That is, circumstances, objects, or conditions can be in a virtual environment. Dr. Begault appears to agree with me in this respect, as he states (after discussing the dictionary

definition) that “in the patents, one can be surrounded by circumstances, objects or conditions that are either real, virtual, or both simultaneously.” Ex.2012 ¶ 21. I agree, and the logical conclusion is that “environment” encompasses both physical and virtual environments.

17. I also note that Dr. Begault’s interpretation of “environment” contradicts other interpretations of “environment” he provides. For example, in his first declaration, Dr. Begault stated “a PHOSITA would recognize that the stated ‘environment’, a ‘3D virtual environment’, or a ‘virtual image in the environment’ in the subject claims *all to refer to the same type of immersive virtual world* context that the user interacts within.” Ex.2006 ¶ 13. As another example, in his second declaration, Dr. Begault stated that “[a] PHOSITA would have the understanding of the meaning of the word ‘environment’ both in its everyday meaning, as such *environments are presented in virtual and augmented reality*, and of the physical real environment in which a user would operate a virtual reality system.” Ex.2012 ¶ 64. Both of these interpretations appear to agree that “environment” is not limited to physical surroundings but can include virtual surroundings as well.

18. I also disagree with Dr. Begault that “leaving the environment” refers to moving from one physical environment *condition* to another. *See* Ex.2012 ¶ 25. Dr. Begault does not explain his interpretation nor rely on anything from the ’179 or ’184 patents to justify it. I understand that in his deposition. Dr. Begault asserted

that any change in the physical environment constitutes leaving the environment. Ex.1147 209:3-9 (“Q: And your testimony here today is that any change in the circumstances or conditions or objects around you is leaving one environment and getting into a new environment; is that correct? . . . A: That is the way that we are operatively speaking about that, yes. . . .”). Under Dr. Begault’s understanding, moving a coffee mug constitutes leaving the environment. Ex.1147 208:9-209:19. I disagree that a mere change in the environment constitutes the *user leaving* the environment. As I discuss below in relation to the “3D virtual environment,” the “leaving” is something the *user* does, not the environment itself.

B. Claim 1 of both patents: “3D virtual environment” is not a sound-only environment

19. Claim 1 of both patents recites, among other things, “warning the user that the user is leaving the 3D virtual environment.”

20. Dr. Begault asserts that “a PHOSITA would have understood the term ‘3D virtual environment’ to mean more specifically ‘virtual surroundings in which a user operates that incorporate 3D virtual images and/or 3D sound.’” Ex.2012 ¶ 23 (emphasis added).

21. I disagree with Dr. Begault’s interpretation of “3D virtual environment” in claim 1 of both patents. Specifically, I disagree that the claimed “3D virtual environment” can be a sound-only virtual environment without incorporating 3D virtual images.

22. The claim language precludes a sound-only virtual environment in claim 1. In both patents, claim 1 explicitly requires that the 3D virtual environment “includes a virtual image.” Binaural sound is recited elsewhere in the claim, and “externally localizes in empty space to the virtual image in the 3D virtual environment.” Thus, the virtual image has a location in space (i.e., it is visual) and the binaural sound localizes to the same point that the image is displayed. The 3D virtual environment therefore cannot be sound only but must include a virtual image. I also note that claim 4 of the ’184 patent refers to the “virtual image [being] no longer *visible* to the user,” again reinforcing that the virtual image is a visual image that the user can see.

23. The claims also recite “displaying” the 3D virtual environment “with a display of a head mounted display (HMD) worn on a head of a user.” A POSA would understand that, by requiring that the 3D virtual environment be displayed, the claims further require a virtual image, not just 3D sound. In the asserted patents, images are *displayed* and sounds are *played* or *provided*. In claim 1 of the ’184 patent, for example, binaural sound “plays to the user” or is “playing to the user.” Claim 1 of the ’179 patent specifies that the binaural sound is “provided to the user.” The same is true in the specification, where *images* are displayed and sounds are played. *See, e.g.*, ’184 patent, 36:6-7 (“The display 1757 displays an image on or over the SLP so the wearer sees the position of this SLP.”), 6:20-22 (“An image of

the friend appears on the display of a laptop while the SLP 126 appears above the laptop.”); 6:65-67 (“As such, the sound localization points 154-157 coincide with visual displays or images of people with whom the user 150 talks.”); 7:3-7 (“[T]he OHMD 152 removes the accompanying visual image of the person from being displayed to the user.”), 11:10-12 (“The electronic glasses or head-mounted display provides or displays an image of Bob that coincides with the SLP in empty space.”), 11:39-41 (“Since Alice is not wearing any electronic glasses and cannot see a display, Bob’s voice localizes to the SLP without an accompanying image.”), 20:53-55 (“Soon Alice turns off her display and is left with only the binaural audio experience of their talk (i.e., Bob’s SLP no longer has an accompanying visual image).”). For further example, Figure 17 illustrates that a display of a system (e.g., 1729, 1757) is distinct from speakers or headphones that play sound (e.g., 1730, 1755). I note that Dr. Begault’s prior art research papers discuss virtual environments as requiring (visual) virtual images. Ex.1149 3 (“At a minimum, virtual reality systems usually involve visual interaction on a variable level of complexity. . . . Note that the level of visual information contained within the display drives the hierarchy; audio only appears at the ‘virtual environment’ level.”); *see also* Figure 2 (illustrating virtual image within virtual environments on left with “similar hierarchy of levels [that] can be applied to spatial auditory stimuli” on right), reproduced below.

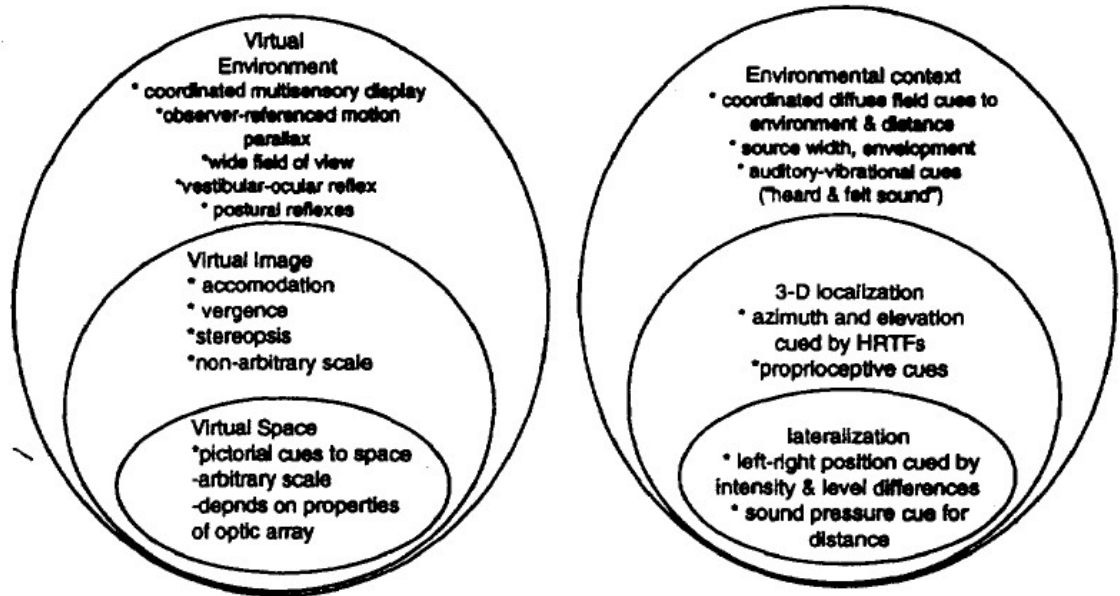


Figure 2. Hierarchies of visual and auditory fidelity levels compared (left diagram from [1]).

Ex.1149 Figure 2

24. Although the specification mentions a Virtual Audial Display in one passage ('184 patent, 30:35), a POSA would not understand the “displaying” of the “3D virtual environment” in claim 1 to refer to a Virtual Audial Display. As I mentioned above, the claims recite displaying the environment including a virtual image. Further, the binaural audio is recited separately from the displayed 3D virtual environment and the virtual image in claim 1. And a POSA would have understood that a Virtual Audial Display uses speakers, headphones, or earphones, and not the recited display of an HMD.

25. Dr. Begault asserts that the specifications are consistent with the fact that “the ‘3D virtual environment’ *includes* 3D audio.” Ex.2012 ¶ 27 (emphasis

added). But Dr. Begault's construction does not merely include 3D audio. Rather, Dr. Begault's interpretation would allow for the 3D audio to itself be the 3D virtual environment. And that possibility is precluded by the claim language as I discussed above. Specifically, the "3D sound environment" that Dr. Begault quotes from the specification is not encompassed by the claims which require the 3D virtual environment to include a virtual image and require that the 3D virtual environment to be displayed on a display of an HMD.

26. Although not addressed by Dr. Begault, the Patent Owner Response asserts that I "acknowledg[ed] that a 3D soundscape could be a 3D virtual environment." '184 POR at 8 (citing Ex.2014 176:3-6). What I actually said was that a 3D soundscape "could be *in*—sure, it could be *in* a virtual environment." Ex.2014 176:3-6. The Patent Owner Response omits that I repeatedly stated before this passage that a virtual environment would have both a visual and a sound modality—not either/or. Ex.2014 173:14-17 ("What I said was that a virtual environment or 3D virtual environment contains both modalities, remember? I talked about the visual modality and the sound modality."); 174:1-6 ("I think the clarification that I made is that a virtual environment would have the visual and the sound modality. I could, of course, close my eyes which is what I believe I said, but the environment, the virtual environment still contains those two modalities."); 174:15-17 ("I mean the environment still has the images as well as the sound. The images are still there.").

Thus, I never stated that a 3D virtual environment could be only a 3D soundscape without virtual images. Nevertheless, the relevant testimony about 3D soundscapes was not in the context of the claims, which plainly preclude an audio-only 3D virtual environment.

27. The Patent Owner Response also cites the disclosure of a “virtual source of binaural sound [localized] to a point” in the environment as supporting an audio-only 3D virtual environment. ’184 POR at 8 (quoting ’184 patent 3:36-39); *see also* POR at 8-9 (citing various other “examples of virtual localized sound sources localized to virtual points in a 3D environment”). But these passages merely describe 3D audio *in* an environment—not that 3D audio *is* a 3D virtual environment. As I said in my deposition, 3D sound can be *in* the environment, but it is not a 3D virtual environment in and of itself, and especially not in the context of the claims for the reasons I described above.

28. Dr. Begault further asserts that “[a] PHOSITA would recognize that the concept of leaving a ‘3D virtual environment’, refers to an act of a transition of user perception of a change in location between virtual environments, or a change in the location or existence of SLPs.” Ex.2012 ¶ 26. In other words, a *change* in the 3D virtual environment constitutes *leaving* the 3D virtual environment.

29. I disagree with Dr. Begault that a change in the environment constitutes *leaving* the environment. In all of the claims, “leaving” the environment is something

the *user* is doing. An environment may change around a user, but those changes do not constitute the *user leaving* the environment. Otherwise, under the claims, the user would need to be warned about every single minor change in the environment—a moved coffee cup—with an audio alert. Ex.1147 208:9-209:10. This absurd outcome shows that the 3D *environment changing* does not constitute the *user leaving* the 3D virtual environment. Dr. Begault’s quotations from his prior art book do not counsel otherwise. Ex.2012 ¶ 28 (quoting Ex.1038 011).

30. Further, Dr. Begault’s book describes that “[V]irtual reality systems would be better described as virtual environment systems” and that such systems usually include a “reality engine” that creates a “virtual world” including “visual, auditory, and haptic illusions.” Ex.1038 017-018; *see also* Figure 1.7 (showing “[c]omponents of a virtual reality or multimedia system” including “visual graphic representation” and “visual display” inputs). I also note that, in his first declaration, Dr. Begault stated “a PHOSITA would recognize that the stated ‘environment’, a ‘3D virtual environment’, or a ‘virtual image in the environment’ in the subject claims *all to refer to the same type of immersive virtual world* context that the user interacts within” and thus did not interpret the term as including a sound-only virtual environment. Ex.2006 ¶ 13.

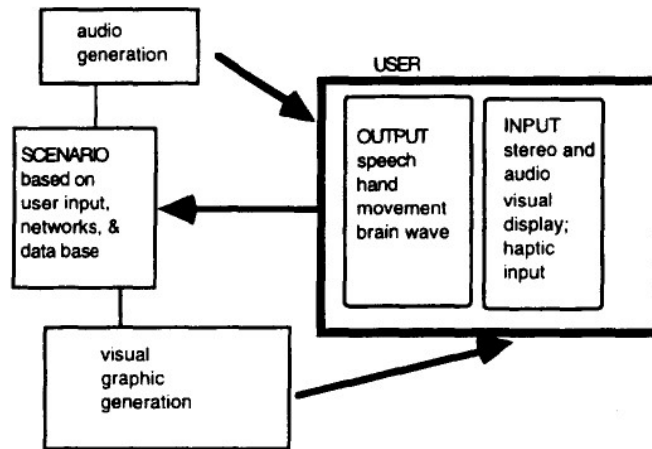


Figure 1.7. Components of a virtual reality or multimedia system. The scenario software is termed a “reality engine” within virtual reality; in a multimedia system, it is the software program that includes “authoring software.” The user receives **effector** data, primarily in the form of visual and audio input, and interacts with the scenario by outputting data from various **sensors**.

C. Claims 7 and 14 of the ’179 patent: “Determining, with the WED, when the user is leaving the environment” does not require independently determining

31. Claims 7 and 14 of the ’179 patent recite, among other things, “determining, with the WED, when the user is leaving the environment.”

32. This phrase does not require the WED to make the determination independently of a stationary tracking device. The claim language does not support such a construction. For example, the claims are open-ended by reciting that the method “compris[es]” the recited steps. In other words, the claims do not preclude the possibility of other unclaimed steps in a method, including steps performed by a stationary tracking device. Nor does the phrase “with the WED” mean that the WED cannot use information from another device such as a stationary tracking device. To the contrary, the claim language simply means the WED performs the claimed

determination without regard for *how* it does it, including whether the determination is performed with the help of any other device.

33. The specification also does not limit the claimed determining step. Although a WED includes various components that can potentially be used independently from a stationary tracking device to make the claimed determination (*see, e.g.*, Figure 20), nothing in the specification defines the “determining” step to be performed independently of any other device. *See* ’179 patent, 40:7-15.

34. It appears that Dr. Begault does not offer this claim construction in his declaration. *Cf.* Ex.2012 ¶¶ 20-29. Additionally, Dr. Begault opined in deposition that the claims do not preclude the use of external sensors and external reference points. Ex.1147 157:9-158:1. He also opined that the WED can perform the determining step with “an external tracking device” as well as an external camera. 159:19-160:12. I agree with these portions of Dr. Begault’s testimony at least as far as the determining step does not require the WED to make the determination independently of external devices.

V. The ’184 Patent

A. ’184 Patent - Ground 1: Claims 1, 3-4, 6-10, and 14-19 Are Unpatentable Over Clement, Osman, and Sauk

1. Relevance in Sauk is not merely a function of the user’s decision.

35. Dr. Begault states that in Sauk, “‘Relevancy’ or ‘importance’ is a function of the user’s selective decision regarding multiple streams of audio

information they are monitoring, as opposed to a transition of the user between the real-world and the virtual world.” Ex.2012 ¶ 85.

36. I disagree. Sauk discloses that “humans have an innate ability to select sounds which are of interest.” Sauk 10:35-37. “This natural ability helps humans reduce or eliminate those sounds that are not needed.” Sauk 10:37-39. Sauk discloses that the dimensionality of the audio makes the user *perceive* the sounds differently. Specifically, in cases where the audio is “presented in 1-D monophonic format” as opposed to “a 3-D binaural format,” the “audio is not presented with the same audio effect, [and] it is *perceived differently by a user*.” Sauk, 11:64-12:1. That is, “[t]he user can use this [dimensional] distinction to selectively focus on the voices of members of their own group.” Sauk, 12:1-3. Thus, although a user can choose to have the system change the dimensionality of an audio stream, the system’s changing of the dimensionality of the stream signals to the user’s “natural” and “innate ability to select sounds which are of interest” that a stereo or mono stream is less relevant and less important than streams presented in 3D. *See* Ex.1003, ¶ 112.

2. Sauk does not teach away from the combination

37. As I described in my first declaration, “[a] POSA would have combined Clement and Osman with Sauk to provide additional teachings regarding binaural sound, including signaling the importance of things or events to provide flexibility and additional options for different conditions, that enhance the system’s ability to

alert or warn the user when the user approaches the boundary or when transitioning from a virtual environment to the real world.” Ex.1003 ¶ 111. For example, Clement discloses an “audible indicator” when transitioning to the real world, and a POSA would have known that different audible indications may be preferable in different situations. Ex.1003 ¶ 111. For example, Clement teaches warning the user (e.g., with a buzz) when the user is moving towards an obstacle or other danger. Ex.1003, ¶ 175 (citing Clement 7:28-8:35). And a POSA would have sought Sauk’s teachings to introduce multiple layers of warnings, to ensure user safety as the user leaves a virtual environment. Ex.1003, ¶ 111.

38. POSAs would have understood from Sauk that binaural alerts can be used to alert/warn the user of immediate hazards, and that less relevant audio can be transitioned to mono/stereo to alert that user that the audio is becoming less important/relevant. Ex.1003, ¶¶ 83, 113, 173. POSAs would have understood that these alerts/warnings could be used together to introduce multiple layers of warnings: transitioning less relevant sound to mono/stereo would encourage a user to focus on the remaining binaural warning indicating an immediate hazard. A POSA would have combined Sauk with Clement and Osman to use a binaural alert when the user is *approaching* a hazard (as taught by Sauk) such as a wall (as taught by Clement). *See, e.g.*, Ex.1003 ¶ 113 (describing additional audible warnings for safety and to prevent injury based on Sauk’s disclosure of audio warnings for safety

hazards). The spatial aspect of this alert allows the user to know where the hazard is and in which direction to move to avoid the hazard. Sauk 12:46-51.

39. The POSA further would have combined Sauk with Clement and Osman such that the sound from the environment, which is becoming less important or relevant, would switch to mono or stereo from binaural. Ex.1003 ¶ 112. This transition would also allow the user to focus on the binaural alert of the physical hazard “to ensure that the warning would be heeded by the user.” Ex.1003 ¶ 112. Thus, on top of the binaural alert that the user is heading toward an obstacle, a POSA would have been motivated “to provide an *additional audible indication*” by way of “degrading the ‘binaural [audio]’” from the environment as it becomes less relevant. Ex.1003 ¶ 175.

40. In the Institution Decision, the Board stated that “Petitioner provides no persuasive reason why one of ordinary skill in the art would have ignored Sauk’s disclosure of adding binaural sound for alerts and instead look to Sauk’s non-alert context for removing binaural sound.” Institution Decision at 17.

41. But a POSA does not need to ignore Sauk’s disclosure of binaural alerts when approaching hazards. To the contrary a POSA would have been motivated to apply that teaching to Clement and Osman—binaural alerts when the user is approaching a hazard such as an obstacle. Ex.1003 ¶¶ 111-13, 175. Though the fact that the user is *leaving* the environment is relevant, a POSA would have understood

that the *audio* from the environment—or from the image in the environment—becomes less important or relevant as the user transitions out of the environment and/or approaches a hazard. Therefore, the POSA would have been motivated to apply Sauk’s teachings about relevance to that audio and degrade it to a mono or stereo dimensionality. It would make no sense to add directionality to the sound from the image in the environment as the user exits the environment and approaches a hazard that could be from a direction that’s different than the image, as there is no danger coming from the image. In other words, the user is leaving the image, not approaching the image.

42. A POSA would have been motivated to transition out of binaural audio to alert the user that it is leaving the environment whether or not the user is approaching an immediate hazard (which generates its own binaural alert). By transitioning the environment’s audio to mono/stereo, users can “selectively focus” on more relevant sounds, Sauk, 11:64-12:3, e.g., the environment the user is entering. Ex.2014 270:4-271:1. Degrading of the environment’s audio and the binaural hazard alert are complementary alerts, helping the user to focus on and avoid an immediate hazard. Roginska ¶¶ 112-113, 175. But even in scenarios without the immediate hazard, transitioning the environment’s audio to mono/stereo helps the user with a smooth transition out of the environment and helps them focus on the physical environment which has become more relevant.

3. '184 Patent - Independent Claims 1, 7, and 14

(1) lower[ing]/reducing a volume of the binaural sound to warn (claims 1, 7, and 14)

43. In my first declaration, I showed that this limitation was disclosed or rendered obvious by “the Clement combination” (Ex.1003 ¶ 317) including Clement and Osman. I noted that Clement with Osman discloses a 3D virtual environment that includes binaural audio that localizes from an image. Ex.1003 ¶ 164. I further noted that Clement described a transition out of an environment with a “serial fade out of the virtual audio” *Id.* ¶ 165 (quoting Clement 12:47-65). I described that in the combination of Clement and Osman, the audio that fades is binaural sound. *Id.* ¶ 168.

44. Dr. Begault does not seem to address the combination of Clement and Osman, but rather attacks the references in isolation. For example, Dr. Begault asserts that Clement does not disclose binaural audio. Ex.2012 ¶ 97. But he does not seem to dispute that Clement combined with Osman discloses binaural audio. Ex.1003 ¶ 168.

45. Dr. Begault further asserts that Sauk alone doesn’t teach “lowering the volume of the **warning**.” Ex.2012 ¶ 97. I note that the claims do not recite lowering the volume of a warning, but rather lowering/reducing volume of binaural **sound**, which the combination discloses as I described above and in my first declaration. Sauk does not teach away from “alerting by reducing the volume of binaural audio”

for the same reasons I described above in Section V.A.2. In fact, Sauk is consistent and compatible with Clement’s teaching of lowering the volume of less relevant sound as the user transitions out of the virtual environment. *See* Sauk 11:26-41.

46. In addition, my first declaration does not require combination with Sauk¹ for claim 14 of the ’179 patent. Ex.1003 ¶¶ 116-70.

**(2) filtering to alter a spectral cue to warn (claim 1),
degrading a spatial cue of the sound to alert (claim 7),
or altering the processing of the binaural sound to alert
(claim 14)**

47. Dr. Begault asserts that “Sauk’s disclosure of an idea that allows the users to pre-select the type of spatial audio rendering for voice communication in Selective Relevance Mode is [] a completely different concept than the binaurally-processed warning sounds used in his Alerts and Warnings mode.” Ex.2012 ¶ 87. I disagree, and Dr. Begault seems to ignore the combination of Clement, Osman, and Sauk, which would employ “multiple options” for audible indications “in different situations” as I explained in my first declaration and further discussed above in Section V.A.2. Ex.1003 ¶ 111. Further this combination would not base a change in

¹ Although I cite to Sauk at, e.g., ¶¶ 119, 140, 167, each these citations are in addition to disclosures from Clement and Sauk that render the claim obvious on their own, as I explained in my declaration. *See, e.g.*, Ex.1003 ¶¶ 117-18, 139, 164-66.

audio on user selection. Rather, the combination relies on Clement’s disclosure of detecting when a user is leaving an environment and alerting the user to that effect. Sauk teaches the POSA about binaural alerts and how users perceive audio with different dimensionality to have different relevance or importance. Ex.1003 ¶¶ 112-13. Sauk also teaches that how to switch from binaural to monaural/stereo—i.e., through the manipulation/removal of the binaural metadata. Sauk 11:41-12:13. And a POSA would have already known that the exact type of switching described in Sauk—switching from binaural to stereo/mono—could be used as an alert. *See* Ex.2012 ¶ 29 (“A PHOSITA would recognize that, prior to the advent of virtual reality, cueing a transition between visual and auditory environments through the use of sound is a familiar concept as old as the history of scene changes in dramatic stage arts” including in virtual auditory environments using “first monaural, then binaural audio...”); *see also id.* ¶ 76 (“A PHOSITA would recognize common and well-known methods of spatial processing that are relatively degraded compared to 3D sound techniques that use HRTF filtering.”). But a POSA would not rigidly apply Sauk to Clement such that a user would have to select the dimensionality of the audio, and therefore in the Clement-Osman-Sauk combination, Sauk’s disclosure of user selection does not appear to be relevant to the patentability of the claims.

48. Dr. Begault also asserts that “Sauk generates *new reproductions* of enunciated data based on the user switching between binaural, stereophonic, and

monophonic audio reproduction.” Ex.2006 ¶ 35. I understand that the Board disagreed with this argument in the Institution Decision, and Dr. Begault does not seem to have repeated it in his second declaration. Institution Decision at 17-18. I also disagree. As I noted in my first declaration, Sauk discloses that “a user could change the less relevant audio” to a lower dimensionality, and that “[t]his process also removes any localization cues for that audio.” Sauk 11:34-40; Ex.1003, ¶¶ 83, 173, 178, 181-82, 187, 321, 328. I further described that a POSA would have understood this includes altering the binaural audio’s spectral cues, degrading spatial cues, degrading audio cues, filtering the binaural sound, and altering processing of binaural sound. Ex.1003, ¶¶ 179-80. Dr. Begault does not seem to address or dispute these issues. *See* Institution Decision at 18.

49. Nevertheless, even replacing the binaural sound with stereo/mono would still meet the claims, because the binaural sound would still be filtered to alter a spectral cue (cl. 1), have a spatial cue degraded (cl. 7), and altered (cl.14) because the audio provided to the user is filtered, degraded, and altered such that the spatial cues (e.g., ILD, ITD) are not present in the audio from the environment played to the user, where the previously had been present in the audio from the environment played to the user. Institution Decision at 24-25; Sauk, 11:34-40; Ex.1003 ¶¶ 178-82. The same is true for the “filtering . . . to alter audio cues” (’179 claim 1) and “altering processing . . . to degrade audio cues” (’179 claim 7) recited in the ’179

patent. In fact, the '179 patent discloses that one “way of changing a sound” from binaural to non-binaural is “to render the original source with different spatial parameters or to re-render a sound source with different spatial parameters in order to adjust, degrade, or eliminate ... audio cues.” Ex.1001 5:33-62.

50. Dr. Begault also states that “Sauk does not use the change in spatial audio rendering in any of the three modes to function as an alert.” But as I noted previously, a POSA would have known to use a degradation of spatial cues as an alert/warning. *See* Ex.1003, ¶ 112; *see also* Ex.2012 ¶ 29 (“A PHOSITA would recognize that, prior to the advent of virtual reality, cueing a transition between visual and auditory environments through the use of sound is a familiar concept as old as the history of scene changes in dramatic stage arts” including in virtual auditory environments using “first monaural, then binaural audio....”); *see also id.* ¶ 76 (“A PHOSITA would recognize common and well-known methods of spatial processing that are relatively degraded compared to 3D sound techniques that use HRTF filtering.”). Additionally, in the context of claim 14 of the '184 patent, Dr. Begault asserts that

“[a] PHOSITA would understand that that altering processing of the binaural sound by the DSP and reducing a volume of the binaural sound that localizes to the user at the virtual image ('184, claim 14) to mean changing some aspect of the position of the SLP (for example from binaural to stereo, or from binaural to monaural) in conjunction

with a reduction in amplitude (volume), so as to effect an alert to the listener (user) regarding a transition.” Ex.2012 ¶ 80.

In the context of claim 7 of the ’184 patent, he further asserts that a

“PHOSITA would reader ‘degrade a spatial cue and to lower a volume of the sound playing to the user’ (’184, claim 7) to mean using spatial auditory means of processing other than HRTFs to process a target sound (such as stereo or monaural processing), in conjunction with lowering the gain of the sound as described for claims 1 and 14 of the ’179 Patent above.” Ex.2012 ¶ 76.

Dr. Begault has not explained why he interprets the claims this way, but nevertheless the combination of Clement, Osman, and Sauk teaches these limitations even under Dr. Begault’s interpretations for the reasons discussed above and in my first declaration, for example by changing from binaural to mono/stereo in conjunction with lowering the volume.

51. Further, Dr. Begault’s Ground 2 arguments, discussed below, contradict this Ground 1 argument. There, Dr. Begault asserts that the mere switching from binaural to mono/stereo sound “implicitly” alerts the user that it is leaving the environment. Ex.2012 ¶ 59 (“adjustments in binaural sounds . . . would implicitly alert or warn the user....”). Dr. Begault’s deposition testimony also contradicts Patent Owner’s position on the prior art. *See* Ex.1147 237:12-19 (“a very salient change in auditory cuing would occur, which is, effectively, an alert”),

239:19-240:15 (“a noticeable change acts as an alert”), 265:17-266:20 (similar), 279:14-18 (similar). These statements by Dr. Begault support that the alert is obvious to a POSA. For example, a POSA would have known that they could use the change from binaural to mono/stereo as an alert and would have had reason to do so in the combination (to provide additional layers of warning to promote and ensure user safety).

4. '184 Patent - Dependent Claims 15-19

- (1) Claim 15. The method of claim 14, wherein the DSP alters processing of the binaural sound to alert the user by changing a spatial cue that includes at least one of an interaural time difference (ITD), an interaural level difference (ILD), and the HRTFs.**

52. In my first declaration I showed that converting binaural sound to stereo/mono sound changes or removes the inserted ITDs, ILDs, and HRTFs in the binaural sound. Ex.1003 ¶¶ 319-21, 171-85. I also noted that Sauk discloses its “process also removes any localization cues for that audio.” Ex.1003 ¶ 321 (quoting Sauk 11:35-40). Although Dr. Begault asserts that Sauk alone or “in combination with Osman and Clement disclose a method of alerting/warning the user by removing/degrading the binaural sound received by the user (claims 8, 10, 15-19 of the '184 Patent)” (Ex.2012 ¶ 100), he does not address the disclosures and arguments from my first declaration. *See* Ex.2012 ¶¶ 100-02.

- (2) Claim 16. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by stopping, by the DSP, the processing of the sound with the HRTFs.**

53. The combination of Clement, Osman, and Sauk renders claim 16 obvious as I discussed in my first declaration, along with the additional discussion I provided above for claim 15.

- (3) Claim 17. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by removing interaural time differences (ITDs) of the binaural sound being provided to the user.**

54. The combination of Clement, Osman, and Sauk renders claim 17 obvious as I discussed in my first declaration, along with the additional discussion I provided above for claim 15.

- (4) Claim 18. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by eliminating one or more spatial cues in the binaural sound being provided to the user.**

55. The combination of Clement, Osman, and Sauk renders claim 18 obvious as I discussed in my first declaration, along with the additional discussion I provided above for claim 15.

- (5) Claim 19. The method of claim 14 further comprising: alerting the user that the user is leaving the environment by switching the binaural sound playing to the user to one of stereo sound and mono sound.**

56. The combination of Clement, Osman, and Sauk renders claim 19 obvious as I discussed in my first declaration, along with the additional discussion I provided above for claim 15.

5. '184 Patent - Dependent Claims 8-10

- (1) Claim 8. The method of claim 7, wherein the spatial cue includes an interaural time difference (ITD).**

57. The combination of Clement, Osman, and Sauk renders claim 8 obvious as I discussed in my first declaration, along with the additional discussion I provided above for claim 15.

- (2) Claim 9. The method of claim 7 further comprising: displaying, by the WED, the virtual image while the user plays a game that includes a character as the virtual image; and switching a voice of the character from the binaural sound to stereo sound when the WED no longer displays the virtual image during the game.**

58. The combination of Clement, Osman, and Sauk renders claim 9 obvious as I discussed in my first declaration, along with the additional discussion I provided below for claim 4.

- (3) Claim 10. The method of claim 7 further comprising: alerting the user that the user is leaving the environment by switching the binaural sound to stereo sound.**

59. The combination of Clement, Osman, and Sauk renders claim 10 obvious as I discussed in my first declaration, along with the additional discussion I provided above for claim 15.

6. '184 Patent - Dependent Claims 3-4, and 6

- (1) Claim 3. The method of claim 1, wherein the warning occurs when the user is leaving a virtual area where the virtual image is located while remaining in the 3D virtual environment.**

60. In my first declaration, I showed that a POSA would have been motivated to combine Clement with Osman and Sauk to render claim 3 obvious. Ex.1003 ¶¶ 206-09, 348. I showed that “a POSA understood that transitions between virtual rooms were well-known in virtual reality environments.” Ex.1003 ¶ 207. I noted that Jang (Ex.1015) describes the well-known game “Second Life” where a user shifts between virtual spaces and noted a POSA would have been familiar with such games. *Id.* I further described Jang’s description that, in these well-known games, volume from a first space is gradually decreased while volume from a second space is gradually increased as part of transitioning a to a new room. *Id.* ¶ 208.

61. I also described that a POSA would have been motivated to apply the well-understood transitions between virtual rooms to the system of Clement and

Osman and would have had a reasonable expectation of success in doing so. Ex.1003 ¶ 209.

62. Dr. Begault focuses solely on Jang and asserts that “Petitioner does not rely on Clement, Osman, or Sauk to teach” the warning of claim 3. Ex.2012 ¶ 94. This is contradicted by my description above, where I applied a POSA’s knowledge to Clement and Osman to render the claim obvious. Dr. Begault does not seem to dispute any of the POSA knowledge I described nor does he seem to dispute the motivation to combine and reasonable expectation of success. Ex.2012 ¶ 94.

(2) Claim 4. The method of claim 1 further comprising: tracking, with head tracking in the HMD, head movements of the user; and switching the sound from the binaural sound to stereo sound when a face of the user is directed away from the virtual image such that the virtual image is no longer visible to the user.

63. I showed in my first declaration that the combination of Clement, Osman, and Sauk renders claim 4 obvious. For example, I discussed how Clement tracks the user’s head (Ex.1003 ¶ 211) and how a POSA knew about head tracking to implement 3D/binaural sound using HRTFs (*id.* ¶ 212). I also discussed how sound emanating from a source behind the user’s head has lower priority or relevance compared to sounds that the user is facing. *Id.* ¶ 213; *see also* ¶ 197.

64. While Sauk teaches using binaural sound when members of a group “lose sight of each other,” POR 35 (quoting Sauk 10:52-60), this disclosure does not relate to a “virtual image” that is no longer visible to the user. I previously cited

various passages from Osman and Jang that inform a POSA that when a virtual image is not displayed, its audio is often less relevant and is processed differently.

¶ 197. In contrast, the members of Sauk's group remain relevant because they continue to exist; their lack of visibility is not because the user has turned away from them. But virtual images, as in games, often become much less relevant, and eventually entirely irrelevant, when they disappear from the display. Ex.1003 ¶¶ 197, 213. This diminished relevancy is typical in the gaming context, and therefore gives POSAs a reason to apply Sauk's switching to 1-D or 2-D audio.

65. Dr. Begault does not appear to dispute the knowledge of a POSA. Ex.2012 ¶¶ 92-93. Instead he asserts that Osman describes that spatial awareness is more important when objects are not in view. *Id.* ¶ 93 (quoting Osman 7:31-42). Instead, a POSA would understand that this passage simply teaches what POSAs understood: when images are not in view, a good 3D sound model is important, assuming the sound is still relevant. But this passage does not contradict the general knowledge of a POSA that virtual images that are not in view are generally less relevant than those that are in view and it does not suggest that images that are not in view need 3D sound or that they are generally relevant. *See, e.g.,* Osman (Ex. 1006), 10:33-43 ("In one embodiment, the acoustic signals for the localized sounds are delivered with a higher volume when the object or person originating the sound is shown on the display of the portable device, or in the display 412 connected to

game console 422. . . . When the sound origin is not on the display, the volume of the sound is lower.”)

(3) Claim 6. The method of claim 1 further comprising: switching the sound from the binaural sound to stereo sound when the virtual image vanishes from the 3D virtual environment.

66. As I addressed above for claim 4, the combination of Clement, Osman, and Sauk renders claim 6 obvious. *See* Section V.A.6(2).

B. '184 Patent - Ground 2: All Claims Are Unpatentable Over Lyren and Reif

1. The '184 Patent's claims are not supported by its parent, the '140 Application

67. In my first declaration, I described that the '140 Application does not provide written description support for the claims of the '184 patent. Ex.1003 ¶¶ 352-55; *see also* ¶¶ 220-29. Specifically, the '140 Application does not provide support for alerting or warning the user that the user is leaving the [3D virtual environment/environment/virtual image in the environment] by reducing the sound's volume and/or otherwise altering it. Ex.1003 ¶¶ 354-55; ¶¶ 225-28. As a result, the effective filing date of all of the claims of the '184 patent is no earlier than the actual filing date of the '184 patent itself—a result that makes Lyren prior art to the '184 patent. Ex.1003 ¶¶ 223, 352.

68. Dr. Begault provides two tables that he states “give examples from the patent specifications for what a PHOSITA would recognize as physical environment

or virtual environment transitions.” Ex.2012 ¶ 31. According to Dr. Begault, “Table I gives examples of transitions in physical environments, and auditory cues used to inform the user of the transition. Table II gives examples of transitions in 3D virtual environments, and auditory cues used to inform the user of the transition.” Ex.2012 ¶ 31. However, none of these examples provide written description support for the claims, and Dr. Begault does not even assert that they do. Ex.2012 at 20-24.

69. Dr. Begault cites the following passage (Ex.2012 at 21):

Consider an example in which a user listens to an audio drama that was recorded in binaural sound but is played in mono sound through car speakers while the user drives the car. Upon arriving at a destination, the person wants to continue listening to the audio drama, steps out of the car, and places headphones on his head. The system continues streaming the audio drama to the person uninterrupted by sending the stream to the headphones rather than to the car. At this time, the system knows the audio drama is a binaural signal and switches the audio drama to binaural sound as it transmits to and plays through the headphones of the person.

’140 Application 109:29-110:4. Dr. Begault states that “[t]his example shows a transition of a user from one physical environment to another that cues a change in spatial audio rendering.” Ex.2012 at 21. But this passage discloses switching from mono to binaural sound; does not disclose lowering the volume; does not disclose

any alerts or warnings; and does not disclose a virtual image in an environment or any WED/HMD.

70. Dr. Begault cites the following passage (Ex.2012 at 21-22):

Consider an example in which an electronic device is set to provide a SLP of a voice of an intelligent personal assistant three feet in front of a listener. The electronic device includes a sensor (such as a camera or other type of sensor) to determine a distance from the electronic device and/or the listener to an object. When the object is within a predetermined distance (such as being within three feet of the listener), then the electronic device takes an action with regard to the SLP, such as moving the SLP, removing the SLP, switching or changing to stereo or mono sound, etc. This action prevents the voice from appearing to originate or to emanate from the object when such is not the desire or intention of the listener.

'140 Application 117:28-118:4. Dr. Begault states that this “example shows a transition, or change, in the circumstances or conditions of things within the physical environment of the user.” Ex.2012 at 22. But this passage discloses switching to stereo to mono sound based on the user getting closer to an SLP, not leaving the environment or leaving a virtual image in the environment. It also does not disclose lowering the volume or any alerts or warnings.

71. Dr. Begault cites the following passage (Ex.2012 at 22):

Consider the example above in which Alice wears earphones and localizes a voice of Bob in front of her during a phone call. The car has

a sensor that determines Alice is on a binaural call and instructs an HPED of Alice to switch the call to mono. As another example, when Alice enters the car, a system in the car pairs with the HPED and automatically switches the call from binaural to mono. As another example, a GPS device or object recognition device (such as a camera with object recognition software) determine that Alice is entering or in the car and provides a signal to the HPED or other source of the call to switch the call from binaural to mono.

'140 Application 120:20-27. Dr. Begault states that this “example shows a transition of a user from one physical environment to another that cues a change in spatial audio rendering..” Ex.2012 at 22. But this passage discloses does not disclose lowering the volume; does not disclose any alerts or warnings; and does not disclose a virtual image in an environment or any WED/HMD.

72. Dr. Begault cites the following passage (Ex.2012 at 23):

Consider an example in which a user wears or has an activity tracker or motion sensor (such as a device that monitors, tracks, and/or measures fitness-related metrics like distance walked, calories burned, rate of walking or running, etc.). The activity tracker or motion sensor detects when a person commences to walk quickly or run. When this event occurs, the computer system or electronic device changes or switches binaural sound.

'140 Application 161:15-20. Dr. Begault states that this “example shows a transition in 3D virtual environments (“switches binaural sound”) that cues a

change in spatial audio rendering.” Ex.2012 at 23. But this passage does not disclose the user leaving an environment, degrading a binaural sound, or lowering the volume; does not disclose any alerts or warnings; and does not disclose a virtual image in an environment or any WED/HMD.

73. Dr. Begault cites the following passage (Ex.2012 at 24):

Consider an example in which Glen is on a phone call with Alice in which a voice of Alice appears to Glen as stereo sound through speakers in an HPED that he holds. Alice informs Glen that she wants to talk to him “face to face” and requests that they meet in a visually rendered chat room. Glen goes into a quiet area in his house and dons a heads-up display (HUD) that couples with his HPED and meets Alice in the chat room. This action of donning the heads-up-display automatically switches the voice of Alice from stereo to binaural.

’140 Application 120:29-121:3. Dr. Begault states that this “example shows a transition in 3D virtual environments (moving to a visually rendered chat room) that cues a change in spatial audio rendering.” Ex.2012 at 24. But this passage discloses switching from stereo to binaural sound; does not disclose lowering the volume; does not disclose leaving a 3D virtual environment; and does not disclose any alerts or warnings.

74. Dr. Begault also cites to “scenario 110” which I address separately below. Ex.2012 at 23-24.

75. Dr. Begault also cites various passages as supposed disclosure of “physical environment or virtual environment transitions” and “auditory cues used to inform the user of the transition. Ex.2012 ¶ 31. He also describes what he calls “Events that precipitate an alert as described in the specification.” Ex.2012 ¶¶ 34-47. Nevertheless, Dr. Begault does not map these supposed environment transitions or events precipitating alerts to the language of any particular claim. Ex.2012 ¶¶ 30-31, pp. 20-24. I address these passages further in the subsequent discussion.

(1) The '140 Application does not support warning or alerting the user.

76. The '140 Application does not disclose an alert or warning to the user in the form of modifying binaural sound (i.e., filtering to alter a spectral clue (claim 1), altering processing to degrade a spatial cue (claim 7), altering processing by the DSP (claim 14), or lowering the volume (all claims)).

77. Dr. Begault states that “adjustments in binaural sound, such as a change in sound format from binaural to mono or stereo or the reduction in volume as disclose in the '140 Application, would *implicitly* alert or warn the user that the user is leaving the physical environment . . . or 3D virtual environment.” Ex.2012 ¶ 59 (emphasis added). Dr. Begault does not argue that such alerts are expressly disclosed, nor does he assert that they are inherently (or necessarily) disclosed. I note that Dr. Begault said the opposite in his deposition, explaining that not every alteration of binaural sound is an alert (Ex.1147 86:22-87:5) and therefore it cannot

be inherent. Further, I have been informed by counsel that implicit disclosure does not satisfy the written description requirement of the claims.

78. Nevertheless, the '140 Application is devoid of any disclosure about such adjustments to binaural audio serving as an alert or warning to the user. The '140 Application discloses that altering binaural sound is an example “action” that a system can take, e.g., when a person moves into a restricted area ('140 Application 120:13-18; Figure 6); however, the specification never describes or even suggests that such action serves as an alert or warning to the user. There is no warning about leaving an environment at all in this scenario—the system simply acts to remove 3D audio in areas where 3D audio is restricted. For instance, the '140 Application does not describe that switching to mono/stereo when Alice enters a car is a warning that she has entered her car or left any environment. Ex.1028 120:20-27. To the contrary, Alice knows she has entered the car because she did so purposefully, and doesn't need a warning to that effect. Dr. Begault agreed in his deposition that there is not an alert that “she's leaving or entering a car.” *See* Ex.1147 266:5-14. In contrast, in Clement (and virtual realities in general) a user might not realize they are approaching a boundary of a virtual environment and should be alerted to that fact, as disclosed by Clement. Clement 7:28-8:35.

79. Notably, the '140 Application identifies various alerts, and none of them alter binaural sound or lower the volume of the sound. *See* '140 Application

136:3-10 (“alerts include, but are not limited to, one or more of a displaying a visual warning, providing an audible sound, displaying or transmitting a message, altering or adding or removing an image or indicia, providing a command or instruction or notice to a process or computer program, actuating a light (such as a light emitting diode or LED), displaying a visual or perceivable indication or warning, playing an announcement, playing a video, and providing another indication that notifies the user.”). Nevertheless, these alerts inform the user that “binaural sound is being or will be provided”—not that the user is leaving the 3D virtual environment/environment/virtual image in the environment. ’140 Application, 136:12-15; *see also* ’140 Application, 137:6-8 (“Before switching from stereo sound to binaural sound, the earphones provide Alice with an audio warning of a voice speaking: ‘Switching to binaural.’”); 137:15-18 (“When the system switches to binaural sound . . . a display in the WEGs provides a green colored icon . . . that indicates to the person that binaural sound is activated.”). Notably, after the alert is acknowledged, “the electronic device switches to binaural sound.” ’140 Application, 136:28-29. The ’140 Application includes an audible alert that, e.g., rice has finished cooking, or a printer is out of paper, or a child is behind a car—not that the user has left any kind of environment—with a *binaural* alert, not by altering binaural audio or reducing volume. ’140 Application 137:32-138:27; *see* Ex.2012 ¶ 60. Indeed, where a child is behind a car, it would make no sense in that context to change an

alert from binaural to monaural/stereo. This is in stark contrast to the combination of Clement, Osman, and Sauk, where the sound altered from binaural to monaural are the sounds from the virtual environment that are no longer relevant.

(2) The '140 Application does not support warning the user that the user is leaving the 3D virtual environment (claim 1)

80. To the extent that Dr. Begault relies on his interpretation of “3D virtual environment” and “leaving the 3D virtual environment,” I disagree for the reasons I stated above in Section IV(B).

81. However, Dr. Begault does not even identify written description support for this limitation under his own interpretation of the claim. For example, Dr. Begault’s interpretation of “3D virtual environment” includes visual immersive environments, which he does not appear to argue written description of in the context of claim 1 (i.e., leaving such an environment and alerts for the same). Disclosure of having 3D audio turned off in a physically restricted area is not representative of an alert that the user is leaving a visually immersive environment. For example, 3D virtual environments add several complications not addressed or relevant in physical environments. For instance, 3D virtual environments require users or the system to first define the area within which the 3D virtual environment may safely exist. *E.g.*, Ex.1080 15:13-24; Clement 4:26-46. POSAs would have also understood that these 3D virtual environments require that special consideration is given to users leaving

such environments to ensure their safety as they leave such environments, including to avoid motion sickness and safely transition back to the real world. Clement 5:6-12. They also require unique tracking consideration to detect when users leave the specific identified boundary. Ex.1080 4:49-5:24. None of these, or similar questions, are addressed or considered in the '140 Application, and therefore a POSA would not have understood the '140 Application to disclose warning the user that it is leaving an immersive virtual environment.

82. Dr. Begault seems to try to combine disparate embodiments to find disclosure of leaving a visually immersive environment. Ex.2012 ¶¶ 53-55. But the '140 Application never describes “a virtual space where 3D audio is not permitted” contrary to Dr. Begault’s attempt to combine embodiments to find such a teaching. Ex.2012 ¶ 55. Nor is there any reason why a POSA would have understood the '140 Application to be teaching modifying Figure 6 to apply to visually immersive environments.

83. Nor does Dr. Begault address the claim as a whole, which requires warning a user that they are leaving a 3D virtual environment that “includes a virtual image.” None of Dr. Begault’s examples of leaving a 3D virtual environment include a virtual image—instead Dr. Begault describes them as disclosing a “*change in environmental condition*.” Ex.2012 ¶ 51; see ¶¶ 51-60. For instance, Dr. Begault relies upon the following embodiment (Ex.2012 ¶ 52):

Consider the example above in which Alice wears earphones and localizes a voice of Bob in front of her during a phone call. The car has a sensor that determines Alice is on a binaural call and instructs an HPED of Alice to switch the call to mono. As another example, when Alice enters the car, a system in the car pairs with the HPED and automatically switches the call from binaural to mono. As another example, a GPS device or object recognition device (such as a camera with object recognition software) determine that Alice is entering or in the car and provides a signal to the HPED or other source of the call to switch the call from binaural to mono.

'140 Application, 120:20-27. But in this embodiment, the user (Alice) is not even wearing an HMD when she goes into the restricted areas, but rather is wearing earphones and is using a handheld portable electronic device (HPED). And a POSA would have understood that a WED/HMD could not possibly be used in this embodiment: it would be unsafe for a driver to drive using a WED/HMD.

84. Nor does the following passage of the '140 Application provide written description support of the claims:

Consider an example in which Alice and Bob are both in a 3D visually rendered space talking binaurally face-to-face in full-duplex with each person in a medial plane of the other. Because they can see each other's visual representation, they experience accurate sound localization of each other's voices. Soon Alice turns off her display and is left with only the binaural audial experience of their talk (i.e., Bob's SLP no longer has an accompanying visual image). Due to the lack of a visual

cue, Alice cannot accurately locate the SLP of Bob's voice. The system detects that her screen is off, knows Bob's SLP is in her medial plane, and knows she has no head tracking hardware. The system makes a verbal announcement to Alice ("Adjusting localization") and moves Bob's SLP to a predetermined position that Alice has chosen for all communications with no visual image. Alice is familiar with where this location is and looks to Bob's SLP as she continues the conversation with Bob.

'140 Application, 127:28-128:7. But in this example there is no warning to Alice that she is leaving a "3D visually rendered space." To the contrary, the movement of the SLP for Bob's voice occurs after Alice has already left the 3D visually rendered space.

85. Finally, Dr. Begault cites description of Figure 16 in relation to "a method to provide binaural sound to a person and to take action when an event occurs." Ex.2012 ¶ 57. Dr. Begault cites the following passage:

Among other things, events can be triggered by changes in a user's or a remote user's network conditions, system resources, hardware, software, operating system notifications, the passage of time, the granting or denial of various resource and/or file permissions, a change in the ability to detect motion and/or object location and/or orientation and/or position in a physical or a virtual environment, or a change in the ability to detect an environment, its shape, acoustic properties, or noise level.

'140 Application 140:31-141:4. None of these disclosures are examples of a user leaving a 3D virtual environment. A POSA would understand that a change in the ability to detect position in a physical environment implies that the user is still in the environment and not leaving it. *See also* Institution Decision at 25. This list of changes largely relate to technical conditions of the user's equipment rather than the user leaving an environment. Dr. Begault does not even appear to assert that these are examples of the user leaving an environment, but rather "events that can precipitate altering binaural sound." Ex.2012 ¶ 57.

86. Dr. Begault also cites the following passage of the '140 Application (Ex.2012 ¶ 57):

Events are also triggered according to one or more audio cues detected in a user's or a remote user's physical or virtual environment such as cues indicating a location, position, or orientation, or a change in them, cues indicating a reference frame of a user or a remote user, a lateral or vertical motion, a change in distance, *a change in a physical or a virtual environment* such as its shape, acoustic properties, noise level, or placement of objects or structures in the environment.

'140 Application 141:4-10. But as I previously discussed, a change in the environment does not constitute the user leaving the environment. *See* Sections IV.A-B. For example, even after a change in "shape, acoustic properties, noise level, or placement of objects or structures," it is still, according to the '140

Application, “*the environment*”—i.e., the same environment—and the user is still in it.

87. Additionally, Dr. Begault cites the following passage of the ’140 Application (Ex.2012 ¶ 57):

Events can be triggered by a change in a user's ability to associate visual cues or images with the associated audio cues, such as a visually rendered character vanishing from an augmented or virtual environment, the presence or absence of a physical object, a failure of a visual display system, a *degradation in the visibility of a user's physical or virtual environment*, or the impairment or failure of a user's physical eyesight.

’140 Application 141:15-20. Again, these examples either relate to equipment failure, failure of the user’s own physical eyesight, or a change in the environment—not the user leaving the environment. After all of these triggers, the user remains in the environment.

(3) The ’140 Application does not support alerting the user that the user is leaving “the environment” (claim 7) or leaving “a virtual image in the environment” (claim 14)

88. The ’140 Application also does not provide written description support for alerting the user that they are leaving the environment/leaving a virtual image in the environment in relation to claims 7 and 14. As above for claim 1, to the extent Dr. Begault relies on his incorrect interpretation of “environment,” I disagree with that interpretation for the reasons discussed above in Section IV(A). Nor do any

disclosures cites for leaving the “3D virtual environment” that I discussed above for claim 1 provide disclosure of leaving the “environment” under the correct interpretation of the claim. Dr. Begault fails to show written description support for the full scope of the claim which includes visually immersive environments, and even disclosure of the claims under a physical environment (which they have not shown) would not provide support for the full scope of the claim.

89. There is no written description support for warning a user that the user is leaving an environment or leaving an image in the environment in the context of claims 7 and 14. For example, in relation to Figure 3, the '184 patent discloses that 3D sound can be localized to “virtual images or images per an augmented reality” ('184 patent 11:65-12:15), but does not disclose alerting the user they are leaving an environment that includes a virtual image, or leaving a virtual image in that environment, by altering binaural sound and reducing/lowering volume, as claimed.

90. Dr. Begault cites the below passage about “scenario 110” from the '140 Application (Ex.2012 at 23-24; *see also* ¶ 35):

In scenario 110, a user 120 wears electronic earphones 122 while simultaneously localizing a voice of an intelligent personal assistant (IPA) to a first sound localization point 124 and a voice of a friend to a second sound localization point 126. As shown in box 128A, the user 120 localizes the voice of the IPA to his left and localizes the voice of his friend in front of himself and above his laptop computer that is situated on his desk. An image of the friend appears on a display of the

laptop while the SLP 126 appears above the laptop. As shown in the transition from box 128A to box 128B, the user 120 stops externally localizing a voice of his friend, and the sound localization point 126 disappears. When the call ends, the system changes the sound localization point 124 of the IPA and automatically moves it to be in front of the user 120. A voice of the friend switches to mono or stereo or gets localized internally to the user 120.

'140 Application 102:16-26. Dr. Begault describes this passage as “show[ing] a transition in 3D virtual environments (the location of a friend and the IPA) that cues a change in spatial audio rendering.” Ex.2012 at 24. This passage appears in a table he describes as examples of transitions in 3D virtual environments from “Binaural Sound #1 to Binaural Sound #2.” Ex.2012 at 23. But in scenario 110, the user is not wearing a WED that displays a virtual image as claimed, but rather the image is on a “display of the laptop.” '184 patent, 6:20-22. And this passage, which relates to Figure 1, discloses that 3D sound can be localized to images but do not disclose alerting the user they are leaving an environment that includes a virtual image, or leaving a virtual image in that environment, by altering binaural sound and reducing/lowering volume, as claimed.

91. Nor do other scenarios relating to Figure 1 provide such disclosure. For example, the '140 Application also recites:

In scenario 114, a user 150 wears an optical head mounted display (OHMD) 152 that simultaneously provides a plurality of sound

localization points 154, 155, 156, and 157 during a conference call with four individuals (each individual being represented with a visual image and accompanying SLP). As such, the sound localization points 154 - 157 coincide with visual displays or images of people with whom the user 150 talks. As shown in the transition from box 160A to box 160B, sound localization point 154 (appearing as a visual image of person) walks through a door and out of view of the user 150. When this occurs, the SLS providing sound to the OHMD 152 switches the voice of the corresponding person to mono, and the system providing video to the OHMD 152 removes the accompanying visual image of the person from being displayed to the user 150.

'140 Application 103:21-31. Like scenario 110, this scenario discloses that 3D sound can be localized to images but does not disclose alerting the user they are leaving an environment that includes a virtual image, or leaving a virtual image in that environment, by altering binaural sound and reducing/lowering volume, as claimed.

92. Dr. Begault asserts that “the example of the user moving from a non-restricted area to a restricted area is an example of the user leaving a 3D virtual environment.” Ex.2012 ¶ 56. As part of this conclusion, he combines various disclosures from disparate parts of the specification to conclude that “in combinations” these disclosures “demonstrate[] that the examples in Figure 6 regarding the restricted area take place in a combination of physical spaces and virtual spaces.” Ex.2012 ¶ 55. But to the extent Dr. Begault means to say that, for

example, the “restricted area” examples apply to a visually immersive environment, I disagree. The examples relating to Figure 6 relate to physical spaces where binaural audio would not be safe for the user to have, and such safety concerns would not be applicable in a virtual reality. For example, there is no reason why a user driving a virtual car would be restricted from hearing binaural audio while doing so. Nevertheless, the Figure 6 example does not have an environment that includes a virtual image, and Alice is not wearing an HMD. Nor does this example disclose alerting Alice by altering binaural sound or *reducing the volume*, which is required by the claims.

93. Further, as I discussed in the previous section, the ’140 Application’s disclosures relating to Figure 16 (e.g., change in ability to detect positions) does not correspond to the user leaving the environment. *See, e.g.*, Institution Decision at 25; Ex.2012 ¶ 57. But even if it was, there is no disclosure of alerting the user by altering binaural sound or *reducing the volume*, which is required by the claims.

2. Combination of Lyren and Reif

94. In my first declaration I showed that the combination of Lyren and Reif render all claims obvious. Dr. Begault does not appear to dispute any of these arguments, as he does not mention Lyren or Reif in either of his declarations.

VI. The '179 Patent

A. '179 Patent – Ground 1: Claims 1, 3-4, 6-10, and 14-19 are obvious over Clement, Osman, and Sauk.

95. In my first declaration I showed that the combination of Clement, Osman, and Sauk render claims 1, 3-4, 6-10, and 14-19 of the '179 patent obvious. Ex.1003 ¶¶ 95-216. That is still my opinion, and I add the below discussion to respond to Dr. Begault's testimony about these claims.

1. '179 Patent – Sauk does not teach away from the combination

96. Sauk does not teach away from the proposed combination with Clement and Osman for the same reasons I discussed for the '184 patent above in Section V.A.1-2.

2. '179 Patent - Independent Claims 1, 7, and 14

(1) lower[ing]/reducing a volume of the binaural sound to warn (claims 1 and 14)

97. I responded to Dr. Begault's testimony regarding these claim limitations above with respect to the '184 patent, and incorporate that discussion here for the corresponding limitations of the '179 patent. *See* Section V.A.3(1).

(2) determining/sensing, with a WED, when a user is leaving an environment/3D virtual environment (claims 1, 7, 14)

98. First, there does not seem to be any dispute that this claim limitation is disclosed or rendered obvious under a construction that does not require the WED

to make its determination independently of a stationary tracking device, which I showed above is how a POSA would have understood the claim. *See* Section IV.C.

99. Second, the Ground 1 combination discloses this limitation even under a construction requiring the determining step to be performed independently of a stationary tracking device. I described in my first declaration Clement's disclosure that "the HMD 100 may include a sensor that detects the user's proximity to a boundary of the tracking area 320, or that the user is within a threshold distance of a boundary of the tracking area 320." Ex.1003 ¶ 157 (quoting Clement, 7:4-16). I also described Clement's disclosure that "[t]he HMD 100 may also include a sensing system 160 including various sensing system devices 162-164 and a control system 170 including various control system devices 171-176 to facilitate manual user control and automated control of the HMD 100." Ex.1003 ¶ 157 (quoting Clement, 3:20-41). I also described in my deposition how the HMD could make such a determination. Ex.2014, 193:8-198:8, 202:15-205:16. For example I described that column 6, starting at line 23, described Figure 3C of Clement including "the HMD 100 and/or the tracking device 310 or other component of the system determines" the user position within the tracking area. Ex.2014, 197:7-198:8. When later asked "How would an HMD locate and track a user in the space?" I explained the various sensors on the HMD that would allow it to do so, including optical and audio sensors as well as gyroscopes. Ex.2014, 202:6-13. The crucial disclosure, however, is

Clement's teaching that the HMD is capable of locating and tracking the user. Ex.2014 205:12-16. For example, Clement describes that a sensor on the HMD such as a proximity sensor or "the camera 180 of the HMD 100 may detect" proximity of the user to the boundary of the tracking area. Clement 7:7-16. Indeed, POSAs knew before the time of the alleged inventions that HMDs could independently determine position and orientation information and when a user is leaving an environment using cameras, light-emitting trackers, inertial sensors (e.g., accelerometers and/or gyroscopes), and any type of tracker (e.g., magnetic, optical, electrical, sonar, radar, ultrasonic, or any combination thereof). *See* Ex.1080 (Pedrotti) 5:8-24; 4:59-61 ("Such trackers, and the processors which determine the position and orientation of such trackers, are well know to those of skill in the art."); *see also* Ex.1150 (Donner) [0013] (HMD based display system "determines the current pose (that is, one or both of position and rotational orientation)" of the HMD "relative to reference coordinate frame (that is, 'world space') and then generates imagery of a scene from a perspective associated with that pose (that is, 'screen space')."), [0027-0028] (hardware includes an inertial management unit (IMU) (including, for example a gyroscope, magnetometer, and accelerometer), image capture, image sensors, structured-light or time-of flight (ToF) depth sensor). From these teachings a POSA would have been able to use the sensors on Clement's HMD to determine that the user is leaving the environment independent of a stationary tracking device.

100. I also note that the '179 patent does not provide any express disclosure of a WED determining that the user is leaving an environment independently of a stationary tracking device. And the various sensors and trackers in the WED of the '179 patent were well-known in the art, including head trackers and location trackers. *See* '179 patent, Figure 20; *see also* Ex.1003 ¶ 45 (“‘head tracking’ is a ‘core element[] for VR experiences.’”) (citing Ex.1067 at 263).

(3) degrading audio cues to prevent external localization of binaural sound to alert (claim 7)/filtering by the DSP, the binaural sound to warn (claim 1)

101. Claim 1 of the '179 patent recites “warning the user that the user is leaving the 3D virtual environment by filtering, by the DSP, the binaural sound to alter audio cues to prevent the user from experiencing external localization of the sounds.” Claim 7 of the '179 patent recites “alerting the user that the user is leaving the environment by altering processing of the binaural sound to degrade audio cues in order to prevent the user from experiencing external localization of the binaural sound.”

102. Dr. Begault states that “[a] PHOSITA would recognize that the ‘prevent[ing] the user from experiencing external localization of the binaural sound’ means the user hears ‘inside the head localization.’” Ex.2012 ¶ 67. According to Dr. Begault, this occurs when the sound is “presented to both ears (monaural sound)” or “when the sound is presented in stereo.” Ex.2012 ¶ 67. Dr. Begault does not appear

to distinguish the claims from the prior art based on this interpretation of the “prevent[ing]” limitation. However, as I described above with respect to the “degrading” limitation of the ’184 patent, the combination of Clement, Osman, and Sauk renders obvious the switching from binaural to stereo/monaural sound. *See* Section V.A.3(2). And as I showed above, Sauk does not teach away from the limitation or “teach the opposite” of the claims. *See* Section VI.A.1; V.A.1-2. That is, the audio from the environment is not a physical hazard that the user would need alerting of its position and thus benefit from 3D audio as it leaves the environment—to the contrary, it becomes less relevant as the user leaves the environment, and the user needs to be more aware of their new environment rather than the environment they’re leaving. The degrading/filtering of the binaural audio warns/alerts the user in conjunction with an additional layer of warning when the user is heading towards danger, as I previously discussed. *See* Section V.A.1-2.

3. ’179 Patent – Dependent claims 3-4, 6, 8-10, and 15-19

103. I responded to Dr. Begault’s testimony regarding these dependent claims limitations above with respect to the ’184 patent, and incorporate that discussion here for the corresponding limitations of the ’179 patent. *See* Sections V.A.3-6.

B. '179 Patent – Ground 2: All claims are obvious over Lyren and Reif.

1. The '179 Patent's claims are not supported by its parent, the '140 Application

(1) The '140 Application does not support warning or alerting the user by modifying binaural sound and/or reducing the volume.

104. The '140 Application does not disclose an alert or warning to the user in the form of modifying binaural sound (i.e., filtering to alter a spectral clue (claim 1), altering processing to degrade a spatial cue (claim 7), altering processing by the DSP (claim 14), or lowering the volume (all claims)) for the reasons I described in my first declaration, along with the above additional commentary I provided for the '184 patent. *See* Section V.B.1(1); Ex.1003 ¶¶ 224-29.

(2) The '140 Application does not support warning the user that the user is leaving the 3D virtual environment (claim 1)

105. The '140 Application does not disclose this limitation for the reasons I described in my first declaration, along with the above additional commentary I provided for the '184 patent. *See* Section V.B.1(2); Ex.1003 ¶¶ 224-29.

(3) The '140 Application does not support alerting the user that the user is leaving “the environment” (claim 7) or leaving “a virtual image in the environment” (claim 14)

106. The '140 Application does not disclose this limitation for the reasons I described in my first declaration, along with the above additional commentary I provided for the '184 patent. *See* Section V.B.1(3); Ex.1003 ¶¶ 224-29.

(4) The '140 Application does not support determining/sensing, with the WED/HMD, when the user is leaving the environment/3D virtual environment

107. The '140 Application does not provide written description support for the “determining” limitations of claims 1, 7, and 14. Ex.1003 ¶¶ 224-29. For example, there is no support for the user leaving the claimed environment/3D virtual environment, as discussed above, and therefore no support for determining that the user has left it, either. Section VI.B.1(2)-(3). Additionally, Alice does not even wear a WED/HMD in the example of entering a restricted area such as a car, and therefore no determining could be made with a WED in those examples. '140 Application, 120:20-27. And boilerplate language that various embodiments may be implemented with an electronic device or WED does not help. '140 Application, 167:9-14. For example, a POSA would not have understood such language to be explicit disclosure that a WED is used in every embodiment—instead this catch-all language discloses that the methods/blocks can be performed in a variety of ways, and does not show written description support because the embodiments of, e.g., Alice entering a car, clearly do not and would not use a WED. For example, the Figure 6 scenarios of the '140 Application do not and would not include virtual images. Ex.1028 119:18-120:1. A POSA would not have had any reason to modify these embodiments to add a virtual image as Alice, e.g., walks around outside. But even if this embodiment

were modified to substitute the HPED for a WED/HMD, the act of Alice getting into the car would not constitute her leaving an environment that includes virtual images.

2. Combination of Lyren and Reif

108. In my first declaration I showed that the combination of Lyren and Reif render all claims obvious. Dr. Begault does not appear to dispute any of these arguments, as he does not mention Lyren or Reif in either of his declarations.

VII. Conclusion

109. I have provided these opinions based on the information currently available to me. I reserve the right to expand or modify my opinions as this proceeding continues and as my investigation and study continues. I may also supplement my opinions in response to any additional information that becomes available to me, including opinions provided by Patent Owner or its expert(s).

110. All statements made herein of my own knowledge are true, and all statements made on information and belief are believed to be true. Further, I am aware that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. § 1001. I declare under penalty of perjury that the foregoing is true and correct.

Declaration for *Inter Partes Review* of USP 10,368,179 and 10,448,184

Executed on June 26, 2024 in New York, NY.

A handwritten signature in black ink, appearing to read 'A. Roginska', written over a horizontal line.

Dr. Agnieszka Roginska, Ph.D.

APPENDIX

The following listing of the challenged claims adds letters to the claim limitations for clarity:

'179 Patent Claims

1[pre] A method comprising:

[a] storing head-related transfer functions (HRTFs) in memory of a head mounted display (HMD) worn on a head of a user;

[b] displaying, with a display of the HMD worn on the head of the user, a three-dimensional (3D) virtual environment that includes a virtual image;

[c] processing, by a digital signal processor (DSP) in the HMD, sounds with the HRTFs to generate binaural sound that is provided to the user and that externally localizes in empty space to the virtual image in the 3D virtual environment;

[d] sensing, with one or more sensors in the HMD, when the user leaves the 3D virtual environment that includes the virtual image; and

[e] warning the user that the user is leaving the 3D virtual environment by filtering, by the DSP, the binaural sound to alter audio cues to prevent the user from experiencing external localization of the sounds and by lowering a volume of the sounds being provided to the user.

2. The method of claim 1 further comprising:

warning the user that the user is leaving the 3D virtual environment by removing, by the DSP, background sounds that occur in the 3D virtual environment as part of a game the user is playing.

3. The method of claim 1, wherein the warning occurs when the user leaves a virtual room while remaining in the 3D virtual environment.

4. The method of claim 1 further comprising:

tracking, with head tracking in the HMD, head movements of the user;
and

switching the sounds from the binaural sound to stereo sound when a face of the user is directed away from the virtual image.

5. The method of claim 1 further comprising:

switching the sounds from the binaural sound to stereo sound when the user activates a physical switch located in the HMD.

6. The method of claim 1 further comprising:

switching the sounds from the binaural sound to stereo sound when the virtual image vanishes in the 3D virtual environment.

7[pre] A method executed by one or more electronic devices, the method comprising:

[a] storing head-related transfer functions (HRTFs) in memory;

[b] displaying, by a display of a wearable electronic device (WED) worn on a head of a user, a virtual image in an environment;

[c] processing, by a digital signal processor (DSP), sounds with the HRTFs to generate binaural sound in empty space that is provided to the user and that externally localizes to the empty space and to the virtual image in the environment; and

[d] determining, with the WED, when the user is leaving the environment; and

[e] alerting the user that the user is leaving the environment by altering processing of the binaural sound to degrade audio cues in order to prevent the user from experiencing external localization of the binaural sound.

8. The method of claim 7 further comprising:

alerting the user that the user is leaving the environment by reducing a volume of the binaural sound being provided to the user.

9. The method of claim 7 further comprising:

displaying, by the WED, the virtual image while the user plays a game that includes a character as the virtual image; and

switching a voice of the character from the binaural sound to stereo sound when the character vanishes from the environment during the game.

10. The method of claim 7 further comprising:

alerting the user that the user is leaving the environment by switching the binaural sound to stereo sound.

11. The method of claim 7 further comprising:

switching the binaural sound to stereo sound upon activation of a switch on the WED.

12. The method of claim 7 further comprising:

wherein the WED includes electronic glasses, the environment is a physical room where the user is located, the virtual image is an augmented reality (AR) character that speaks to the user while the user plays a game, the WED includes head tracking, and a digital signal processor (DSP) in the WED reduces a volume of the binaural sound to alert the user when a face of the user is not directed to the virtual image in the environment.

13. The method of claim 7 further comprising:

sensing, with the WED, when a physical object interferes with a location of the virtual image in the environment; and

switching the binaural sound to stereo sound in response to sensing the physical object interferes with the location of the virtual image in the environment.

14[pre] A method executed by one or more electronic devices, the method comprising:

[a] storing head-related transfer functions (HRTFs) in memory;

[b] displaying, with a wearable electronic device (WED) worn on a head of a user, a virtual image in an environment;

[c] processing, by a digital signal processor (DSP), sound with the HRTFs to generate binaural sound that externally localizes in empty space from the virtual image in the environment;

[d] determining, with the WED, when the user is leaving the environment; and

[e] alerting the user that the user is leaving the environment by reducing a volume of the binaural sound that localizes from the virtual image and that is being provided to the user.

15. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by degrading the binaural sound being provided to the user as the user is leaving the environment.

16. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by altering, by the DSP, processing of the binaural sound so the user perceives less external localization.

17. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by removing interaural time differences (ITDs) of the binaural sound being provided to the user.

18. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by eliminating audio cues in the binaural sound being provided to the user.

19. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by switching the binaural sound playing to the user to stereo sound.

20. The method of claim 14 further comprising:

detecting, with a microphone in the WED, a voice of a person in the environment of the user; and

switching the binaural sound to stereo sound in response to detecting the voice.

'184 Patent Claims

1[pre] A method comprising:

[a] storing head-related transfer functions (HRTFs);

[b] displaying, with a display of a head mounted display (HMD) worn on a head of a user, a three-dimensional (3D) virtual environment that includes a virtual image;

[c] processing, by a digital signal processor (DSP), sound with the HRTFs to generate binaural sound that plays to the user and that externally localizes in empty space at the virtual image in the 3D virtual environment; and

[d] warning the user that the user is leaving the 3D virtual environment by filtering, by the DSP, the binaural sound to alter a spectral cue and lower a volume of the sound playing to the user.

2. The method of claim 1 further comprising:

warning the user that the user is leaving the 3D virtual environment by removing, by the DSP, background sound that occurs in the 3D virtual environment as part of a game the user is playing.

3. The method of claim 1, wherein the warning occurs when the user is leaving a virtual area where the virtual image is located while remaining in the 3D virtual environment.

4. The method of claim 1 further comprising:

tracking, with head tracking in the HMD, head movements of the user; and

switching the sound from the binaural sound to stereo sound when a face of the user is directed away from the virtual image such that the virtual image is no longer visible to the user.

5. The method of claim 1 further comprising:

switching the sound from the binaural sound to stereo sound when the user activates a button located on the HMD.

6. The method of claim 1 further comprising:

switching the sound from the binaural sound to stereo sound when the virtual image vanishes from the 3D virtual environment.

7[pre] A method executed by one or more electronic devices, the method comprising:

[a] storing head-related transfer functions (HRTFs);

[b] displaying, by a display of a wearable electronic device (WED) worn on a head of a user, a virtual image in an environment;

[c] processing, by a digital signal processor (DSP), sound with the HRTFs to generate binaural sound that plays to the user and that externally localizes at the virtual image in the environment; and

[d] alerting the user that the user is leaving the environment by altering processing of the sound by the DSP to degrade a spatial cue and to lower a volume of the sound playing to the user.

8. The method of claim 7, wherein the spatial cue includes an interaural time difference (ITD).

9. The method of claim 7 further comprising:

displaying, by the WED, the virtual image while the user plays a game that includes a character as the virtual image; and

switching a voice of the character from the binaural sound to stereo sound when the WED no longer displays the virtual image during the game.

10. The method of claim 7 further comprising:

alerting the user that the user is leaving the environment by switching the binaural sound to stereo sound.

11. The method of claim 7 further comprising:

switching the binaural sound to stereo sound upon activation of a button on the WED.

12. The method of claim 7 further comprising:

wherein the WED includes electronic glasses, the environment is a physical room where the user is located, the virtual image is an augmented reality (AR) character that speaks to the user while the user plays a game, the WED includes head tracking, and a digital signal processor (DSP) in the WED reduces a volume of the binaural sound to alert the user when a face of the user is not directed to the virtual image in the environment.

13. The method of claim 7 further comprising:

sensing, with the WED, when a physical object interferes with a location of the virtual image in the environment; and

switching the binaural sound to stereo sound in response to sensing the physical object interferes with the location of the virtual image in the environment.

14. A method executed by one or more electronic devices, the method comprising:

[a] storing head-related transfer functions (HRTFs);

[b] displaying, with a wearable electronic device (WED) worn on a head of a user, a virtual image in an environment;

[c] processing, by a digital signal processor (DSP), sound with the HRTFs to generate binaural sound that externally localizes to the user at the virtual image in the environment; and

[d] alerting the user that the user is leaving the virtual image in the environment by altering processing of the binaural sound by the DSP and reducing a volume of the binaural sound that localizes to the user at the virtual image.

15. The method of claim 14, wherein the DSP alters processing of the binaural sound to alert the user by changing a spatial cue that includes at least one

of an interaural time difference (ITD), an interaural level difference (ILD), and the HRTFs.

16. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by stopping, by the DSP, the processing of the sound with the HRTFs.

17. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by removing interaural time differences (ITDs) of the binaural sound being provided to the user.

18. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by eliminating one or more spatial cues in the binaural sound being provided to the user.

19. The method of claim 14 further comprising:

alerting the user that the user is leaving the environment by switching the binaural sound playing to the user to one of stereo sound and mono sound.

20. The method of claim 14 further comprising:

detecting, with a microphone in the WED, a voice of a person in the environment of the user; and

switching the binaural sound to stereo sound in response to detecting the voice.