UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM SERVICES LLC
Petitioner,

v.

SMART SPEAKER LLC
Patent Owner.

Patent No. 12,401,721

Inter Partes Review No. IPR2026-00276

DECLARATION OF HENRY H. HOUH

Table of Contents

I. INTRODUCTION 12
II. EXPERIENCE AND QUALIFICATIONS 12
III. MATERIALS CONSIDERED 17
IV. OVERVIEW OF THE '721 PATENT 18
A. Filing Date and Priority Dates 18
B. Specification 18
C. Claims 20
D. Prosecution History 26
V. RELEVANT LEGAL STANDARDS AND DEFINITIONS 31
A. Obviousness 31
B. Prior Art 33
VI. LEVEL OF ORDINARY SKILL IN THE ART FOR THE '721 PATENT 34
VII. CLAIM CONSTRUCTION 35
VIII. PRIOR ART 36
IX. ANALYSIS AND OPINIONS 37
A. Ground 1: Lim Renders Obvious Claims 1, 5-6, 9-15, 17, 21-22, 24-31, 33 37
1. Summary of Lim 37
2. Claim 1 45
3. Claim 5: "The first device according to claim 1, further comprising a second actuator that is configured for directly or indirectly affect, change, or produce, a second physical phenomenon." 115
4. Claim 6: "The first device according to claim 5, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current." 117
5. Claim 9: "The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone." 118

6. Claim 10: “The first device according to claim 1, wherein the IP address is stored in a volatile or non-volatile memory.” 119

7. Claim 11: ““The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type.” 120

8. Claim 12: “The first device according to claim 1, wherein the IP address is a static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address.” 121

9. Claim 13: “The first device according to claim 1, wherein the WLAN is according to, or base on, Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac.” 122

10. Claim 14: “The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.” 123

11. Claim 15: “The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution.” 124

12. Claim 17: “The first device according to claim 1, wherein the second device is integrated in, is part of, or is entirely included in, a household appliance having a primary function.” 125

13. Claim 21: “The first device according to claim 17, wherein the primary function of the appliance is associated with environmental control, and the appliance is part of a Heating, Ventilation and Air Conditioning (HVAC) system.” 126

14. Claim 22: “The first device according to claim 21, wherein the primary function of the appliance is associated with temperature control, and wherein the appliance is an air conditioner or a heater.” 126

15. Claim 24: “The first device according to claim 17, wherein the appliance is an answering machine, a telephone set, a home cinema system, a

High Fidelity (HiFi) system, a Compact Disc (CD) or Digital Video Disc (DVD) player, an electric furnace, a trash compactor, a smoke detector, a light fixture, or a dehumidifier." 127

16. Claim 25: "The first device according to claim 1, further configured to operate within a building, and wherein the server device is external to the building."....127

17. Claim 26: "The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays."....129

18. Claim 27: "The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode." 130

19. Claim 28: "The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice." 131

20. Claim 29: "The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker." 133

21. Claim 30: "The first device according to claim 1, wherein the sounder comprises an electric bell, a buzzer, a chime, a whistle, or a ringer."....134

22. Claim 31: "The first device according to claim 1, further comprising an enclosure that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors."....135

23. Claim 33: "The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves." 137

B. Ground 2: Lim and Wittenberg Render Obvious Claim 2.....138

1. Summary of Wittenberg 138

2. Claim 2: "The first device according to claim 1, wherein the first device or the server device is configured for processing the first and second human voice data using a voice recognition algorithm for identifying a voice of a specific person."....138

C. Ground 3: Lim and Gruber Render Obvious Claim 3-4, 11 144

1. Summary of Gruber 144

2. Claim 3: "The first device according to claim 1, further comprising a sensor that outputs sensor data that responds to a physical phenomenon, wherein the computer program instructions, when executed, further cause the first device to send to the server device, via the WLAN over the Internet, the sensor data, and wherein the first message is further responsive to the sent sensor data."....145

3. Claim 4: "The first device according to claim 3, wherein the sensor is a thermoelectric sensor that responds to a temperature or to a temperature gradient of an object using conduction, convection, or radiation, or wherein the sensor is a photoelectric sensor that responds to a visible or an invisible light or gamma rays."....155

4. Claim 11: "The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type." 156

D. Ground 4: Lim and Benesty Render Obvious Claims 7-8 158

1. Summary of Benesty....158

2. Claim 7: "The first device according to claim 1, further comprising an additional microphone for capturing human voice data, and wherein the capturing comprises capturing, by the additional microphone, the first and second human voice data." 160

3. Claim 8: "The first device according to claim 7, wherein the microphone and the additional microphone are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array."....163

E. Ground 5: Lim and Rossing Render Obvious Claims 9, 29, 33 164

1. Summary of Rossing....164

2. Claim 9: "The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or

wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone.” 167

3. Claim 29: “The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker.” 173

4. Claim 33: “The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves.” 176

F. Ground 6: Lim and Goldsmith Render Obvious Claims 13-15 177

1. Summary of Goldsmith.....177

2. Claim 13: “The first device according to claim 1, wherein the WLAN is according to, or base on, Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac.” 179

3. Claim 14: “The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.” 184

4. Claim 15: “The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution.” 185

G. Ground 7: Lim and Chatterjee Render Obvious Claim 18-20, 23 189

1. Summary of Chatterjee 189

2. Claim 18: “The first device according to claim 17, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation.” 190

3. Claim 19: “The first device according to claim 18, wherein the primary function of the appliance is heating food, and wherein the appliance is a microwave oven, an electric mixer, a stove, an oven, or an induction cooker.”

194

4. Claim 20: "The first device according to claim 18, wherein the appliance is a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker."....195

5. Claim 23: "The first device according to claim 17, wherein the primary function of the appliance is associated with cleaning, wherein the appliance primary function is associated with clothes cleaning and the appliance is a washing machine, or wherein the appliance is a vacuum cleaner." 196

H. Ground 8: Lim, and Rosenberger Render Obvious Claims 26-28....198

1. Summary of Rosenberger 198

2. Claim 26: "The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays."....200

3. Claim 27: "The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode." 204

4. Claim 28: "The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice." 205

I. Ground 9: Williams and Lim Render Obvious Claims 1, 5-6, 9-15, 17, 21-22, 24-29, 31, 33. 205

1. Summary of Williams....205

2. A POSITA would have been motivated to combine Williams and Lim 208

3. Claim 1....212

4. Claim 5: "The first device according to claim 1, further comprising a second actuator that is configured for directly or indirectly affect, change, or produce, a second physical phenomenon."....267

5. Claim 6: "The first device according to claim 5, wherein the second

physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current.” 270

6. Claim 9: “The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone.” 270

7. Claim 10: “The first device according to claim 1, wherein the IP address is stored in a volatile or non-volatile memory.” 271

8. Claim 11: “The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type.” 273

9. Claim 12: “The first device according to claim 1, wherein the IP address is a static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address.” 273

10. Claim 13: “The first device according to claim 1, wherein the WLAN is according to, or is compatible with, an Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard.” 274

11. Claim 14: “The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.” 275

12. Claim 15: “The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution.” 276

13. Claim 17: “The first device according to claim 1, wherein the second device is integrated in, is part of, or is entirely included in, a household appliance having a primary function.” 277

14. Claim 21: “The first device according to claim 17, wherein the primary

function of the appliance is associated with environmental control, and the appliance is part of a Heating, Ventilation and Air Conditioning (HVAC) system.” 277

15. Claim 22: “The first device according to claim 21, wherein the primary function of the appliance is associated with temperature control, and wherein the appliance is an air conditioner or a heater.” 278

16. Claim 24: “The first device according to claim 17, wherein the appliance is an answering machine, a telephone set, a home cinema system, a High Fidelity (HiFi) system, a Compact Disc (CD) or Digital Video Disc (DVD) player, an electric furnace, a trash compactor, a smoke detector, a light fixture, or a dehumidifier.” 278

17. Claim 25: “The first device according to claim 1, further configured to operate within a building, and wherein the server device is external to the building.” 279

18. Claim 26: “The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays.” 283

19. Claim 27: “The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode.” 283

20. Claim 28: “The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.” 283

21. Claim 29: “The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker. 284

22. Claim 31: “The first device according to claim 1, further comprising an enclosure that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors.” 285

23. Claim 33: “The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves.” 287

J. Ground 10: Williams, Lim and Wittenberg Render Obvious Claim 2.....287

1. Claim 2: “The first device according to claim 1, wherein the first device or the server device is configured for processing the first and second human voice data using a voice recognition algorithm for identifying a voice of a specific person.”.....288

K. Ground 11: Williams, Lim and Gruber Render Obvious Claim 3-4, 11 .293

1. Claim 3: “The first device according to claim 1, further comprising a sensor that outputs sensor data that responds to a physical phenomenon, wherein the computer program instructions, when executed, further cause the first device to send to the server device, via the WLAN over the Internet, the sensor data, and wherein the first message is further responsive to the sent sensor data.” 293

2. Claim 4: “The first device according to claim 3, wherein the sensor is a thermoelectric sensor that responds to a temperature or to a temperature gradient of an object using conduction, convection, or radiation, or wherein the sensor is a photoelectric sensor that responds to a visible or an invisible light or gamma rays.”.....298

3. Claim 11: “The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type.” 298

L. Ground 12: Williams, Lim, and Rosenberger Render Obvious Claims 5-6, 26-28 300

1. Claim 5: “The first device according to claim 1, further comprising a second actuator that is configured for directly or indirectly affect, change, or produce, a second physical phenomenon.”.....300

2. Claim 6: “The first device according to claim 5, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current.”.....303

3. Claim 26: “The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible

light is infrared, ultraviolet, X-rays, or gamma rays.”.....303

4. Claim 27: “The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode.” 304

5. Claim 28: “The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.” 304

M. Ground 13: Williams, Lim and Benesty Render Obvious Claims 7-8 305

1. Claim 7: “The first device according to claim 1, further comprising an additional microphone for capturing human voice data, and wherein the capturing comprises capturing, by the additional microphone, the first and second human voice data.” 305

2. Claim 8: “The first device according to claim 7, wherein the microphone and the additional microphone are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array.” 308

N. Ground 14: Williams, Lim and Rossing Render Obvious Claims 9, 29, 33 308

1. Claim 9: “The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone.”.....308

2. Claim 29: “The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker.” 310

3. Claim 33: “The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound

waves.” 312

O. Ground 15: Williams, Lim and Goldsmith Render Obvious Claims 13-15 314

1. Claim 13: “The first device according to claim 1, wherein the WLAN is according to, or is compatible with, an Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard..” 314

2. Claim 14: “The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.” 316

3. Claim 15: “The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution.” 317

P. Ground 16: Williams, Lim and Chatterjee Render Obvious Claim 18-20, 23 319

1. Claim 18: “The first device according to claim 17, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation.” 319

2. Claim 19: “The first device according to claim 18, wherein the primary function of the appliance is heating food, and wherein the appliance is a microwave oven, an electric mixer, a stove, an oven, or an induction cooker.” 321

3. Claim 20: “The first device according to claim 18, wherein the appliance is a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker.” 322

4. Claim 23: “The first device according to claim 17, wherein the primary function of the appliance is associated with cleaning, wherein the appliance primary function is associated with clothes cleaning and the appliance is a washing machine, or wherein the appliance is a vacuum cleaner.” 322

I, Henry H. Houh, make this declaration in connection with the proceeding identified above.

I. INTRODUCTION

1. I have been retained by counsel for Petitioner Amazon.com Services LLC ("Amazon") as a technical expert in connection with the proceeding identified above. I submit this declaration in support of Amazon's Petition for *Inter Partes* Review of United States Patent No. 12,401,721 ("the '721 patent").

2. I am being compensated at my normal rate, plus reimbursement for expenses, for my analysis. My compensation does not depend on the content of my opinions or the outcome of this proceeding.

II. EXPERIENCE AND QUALIFICATIONS

3. In formulating my opinions, I have relied upon my knowledge, training, and experience in the relevant art. My qualifications are stated more fully in my *curriculum vitae*, which has been provided as Appendix A. Here, I provide a brief summary of my qualifications.

4. I am over the age of 18 and am competent to make this declaration. I have personal knowledge or have developed knowledge of these technologies based upon education, training, or experience, of the matters set forth herein.

5. I received a Ph.D. in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology ("MIT") in 1998. Beforehand,

I received a Master of Science degree in Electrical Engineering and Computer Science in 1991, a Bachelor of Science degree in Electrical Engineering and Computer Science in 1989, and a Bachelor of Science degree in Physics in 1990, all from MIT.

6. I am currently self-employed as an independent technical consultant.

7. After I finished my undergraduate education, I became a research assistant in the Telemedia Network Systems ("TNS") group at the Laboratory for Computer Science. The TNS group built a high-speed gigabit network and created applications that ran over the network. Example video applications included ones for remote video capture, processing, storage, and display of video on computer terminals. In addition to working on the design of core network components, designing and building the high-speed links, and designing and writing the device drivers for the interface cards, I also set up the group's web server.

8. I also helped to build the web pages that initiated the above-mentioned video sessions via a web interface. Vice President Al Gore visited our group in 1996 and received a demonstration of—and remotely drove—a radio-controlled toy car with a wireless video camera mounted on it that was built by our group. This toy car device received commands transmitted over a

network from a remote computer, and video data from the toy car was transmitted wirelessly then over a computer network back to the user controller. On occasion, we allowed users visiting our web site to drive the toy car from their remote computer while they watched the video on their computer. The video stream was encoded by TNS-designed hardware, streamed over the TNS-designed network, and displayed using TNS-designed software.

9. I defended and submitted my Ph.D. thesis, titled "Designing Networks for Tomorrow's Traffic," in January 1998. As part of my thesis research, I analyzed local area and wide area flows to show a more efficient method for routing packets in a network, based on traffic patterns at the time.

10. From 1997 to 1999, I was a Senior Scientist and Engineer at NBX Corporation, a start-up that made business telephone systems for streaming packetized audio over data networks instead of using traditional telephone lines. Our system incorporated an auto-attendant, which was a basic interactive voice response (IVR) application to route calls to an extension or voice mailbox. For example, the auto-attendant could be programmed to provide a prompt such as "Press 1 for sales, 2 for support, 3 for directions, or dial an extension" and route the call accordingly based on the caller's touch tone input. Our voicemail system also provided voice prompts and menu selections in which a caller or user may select various options in the voicemail system.

11. As part of my work at NBX, I designed the core audio reconstruction algorithms for the telephones, as well as the packet transmission algorithms. I also designed and validated the core packet transport protocol used by the phone system. The protocol was used for all signaling in the phone system, including for the setup of conference calls. The NBX system also featured a computer interface for initiating phone calls, which could also initiate conference calls. The NBX system also supported the Telephony Application Programming Interface ("TAPI") that allowed other computer programs to integrate with our system telephony features. We obtained U.S. Patent No. 6,697,963, entitled "Telecommunication method for ensuring on-time delivery of packets containing time-sensitive data," as part of this work. NBX was later acquired by 3Com Corporation and the phone system is used by numerous businesses and remains available online.

12. From 1999-2004, I was employed by Empirix or its predecessor company, Teradyne. Empirix was a leader in test tools for telecommunications protocols and systems, providing functional testing tools as well as load testing tools. One of Empirix's core products was the Hammer, a telecommunications test platform that was used for functional and load testing. The Hammer was commonly used to test IVR systems, and Hammer also offered a testing service for IVRs. From my work with the testing service business, I received U.S.

Patent 7,590,542, titled "Method of generating test scripts using a voice-capable markup language." The '542 patent disclosed a method "for generating test scripts for testing an IVR from an IVR's voice-capable markup language application." (EX1032, Abstract.) From 2000-2001, I conceived of and built a test platform for testing Voice-over-IP (VoIP) based on highly parallel network processors. The first application on this new test platform was a cloud emulator for simulating the effects of transmitting VoIP over a busy network. The platform was based on a multi-core high performance network processor chip.

13. In 2005, as part of my role at BBN Technologies, I helped create BBN's Avoke call analytics business, which was a service for testing and monitoring the IVR systems within call centers. Avoke was used by companies and organizations such as Dell and the Center for Medicare and Medicaid Services (CMMS). Avoke analyzed millions of calls into the CMMS call center IVRs during the roll-out of the Affordable Care Act.

14. Also at BBN, I helped found PodZinger Inc., now known as RAMP Inc. PodZinger utilized BBN's speech recognition algorithms to search through the spoken words in audio and video segments. While I was Vice President of Operations and Technology, PodZinger followed its initial prototype with a full streaming audio and video search solution.

15. I have been awarded several United States patents, and I have several patent publications including the following examples:

- U.S. Patent No. 7,975,296, "Automated security threat testing of web pages";
- U.S. Patent No. 7,877,736, "Computer language interpretation and optimization for server testing";
- U.S. Patent No. 7,801,910, "Method and apparatus for timed tagging of media content";
- U.S. Patent No. 7,590,542, "Method of Generating Test Scripts Using a Voice-Capable Markup Language";
- U.S. Patent No. 6,967,963, "Telecommunication method for ensuring on-time delivery of packets containing time-sensitive data".

16. Additional details concerning my professional qualifications, experience, and publications are set forth in my curriculum vitae ("CV"), which is attached as Appendix A.

III. MATERIALS CONSIDERED

17. In preparing this declaration, I have reviewed, among other things, the '721 patent and file history. I have also considered each of the documents cited in this Declaration and the documents specifically identified below and in the

Exhibit List. In formulating my opinions, I have further relied upon my extensive experience in the relevant fields.

IV. OVERVIEW OF THE '721 PATENT

A. Filing Date and Priority Dates

18. The '721 patent was filed on May 8, 2025, and issued on August 26, 2025, less than 6 months from filing. The '721 patent claims earliest priority to a US provisional application filed on January 9, 2012, which I am informed that the Petitioner treats as the earliest effective filing date only for purposes of this proceeding.

B. Specification

19. The '721 patent has one independent claim and 32 dependent claims. The dependent claims sweep across conventional technologies in the field of electrical engineering. They span everything from generic sensors, to conventional microphone types, to standard networking protocols, to home appliances, and routine lighting and audio components. (EX1001, claims 2-32.)

20. Like the claims, the specification of US 12,401,721 is a vague aggregation of off-the-shelf technologies at the time of the invention. It recites long laundry lists of interchangeable components. "Actuators" could be speakers, "sounders," displays, coolers, heaters, lights, motors, pumps, etc.

(*E.g.*, EX1001, 94:3-21; 96:10-97:29; 99:5-21; 100:7-101:5; 104:8-29; 105:48-106:53; 109:42-110:25.) "Sensors" could be smoke sensors, blood pressure sensors, voltmeters, capacitance meters, magnetic sensors, strain gauge, angular sensors, etc. (EX1001, 17:35--45; 18:42-47; 18:65-66; 19:18-29; 19:30-35.) Various network technologies were described to be "known in the art." (*E.g.,* EX1001, 2:34-38; 6:42-43; 6:66-7:6; 9:35-38.) Each of these elements was well-known at the time, and the specification provides no technical details of how they are integrated. The result is a disjointed specification where components and processes are described vaguely and separately but never tied together in coherent and novel embodiments.

21. Although the claims of the '721 patent focus on a "captured human voice," the specification provides sparse and scattered references to voice. I note that a paragraph found in the Summary section describes voice control over a network, and that paragraph's description is at a high level of generality. It describes "a router in the building (or vehicle), connected to the one or more in-building (or in-vehicle) networks and to the external network, and may be operative to pass digital data between the in-building (or in-vehicle) and external networks." (EX1001, 32:25-29.) Further, "a first device in the building (or vehicle) . . . operative to transmit a digital data corresponding to the sensed voice to the router over the one or more in-building (or in vehicle) network" and "[a]

second device may be operative to execute actuator commands received from the router over the one or more in-building (or in-vehicle) networks." (EX1001, 32:61-33:3.) Further, a "control server may be operative to produce actuator commands in response to [a] voice processor output according to the control logic, and may be operative to transmit the actuator commands to the second device via the router, and the voice processor may be entirely or in part in the first device, the router, the control server, or any combination thereof." (EX1001, 33:7-33:13.)

22. The paragraph above, however, does not describe the claimed techniques involving multiple voice commands or the specific operation of multiple actuators across different devices. FIG. 19 of the '721 specification depicts a broad workflow of voice processing and is similarly lacking. (EX1001, FIG. 19, 134:33-135:25.)

C. Claims

23. I was asked to evaluate whether claims 1-33 of the '721 patent are patentable.

24. There is one independent claim. Claim 1 is directed to a first device for use with a Wireless Local Area Network (WLAN) in a building. The first device comprises: (1) a WLAN transceiver for communicating over the WLAN; (2) a microphone for capturing human voice data; (3) and a sounder for

converting an electrical energy to audible sound waves; (4) and one or more processors programmed with computer program instructions. The computer program instructions, when executed, cause the first device to (1) capture, using the microphone, first and second human voice data; (2) send, over the Internet via the WLAN, to a server device, the captured first and second human voice data; (3) receive, from the server device via the WLAN over the Internet, first and second messages, in response to the sending of the respectively captured first and second human voice data; (4) sound, using the sounder, digital audio content in response to the received first message; and (5) send, to a second device over the WLAN, a control message that comprises a control data for controlling a first actuator in the second device to directly or indirectly affect, change, or produce, a first physical phenomenon, in response to the received second message. And the each of the first and second device is addressable in the WLAN and in the Internet using a respective Internet Protocol (IP) address.

25. There are 32 dependent claims. Claim 2 requires the first device or server device to be configured for processing the first and second human voice data using a voice recognition algorithm for identifying the voice of a specific person.

26. Claim 3 requires the first device to further comprise a sensor that outputs sensor data that responds to a physical phenomenon, wherein the

computer program instructions, when executed, further cause the first device to send to the server device, via the WLAN over the Internet, the sensor data, and wherein the first message is received in response to the sent sensor data. Claim 4 requires that the sensor be a thermoelectric sensor that responds to a temperature or to a temperature gradient of an object using conduction, convection, or radiation, or wherein the sensor is a photoelectric sensor that responds to a visible or an invisible light or gamma rays.

27. Claim 5 requires that the first device to further comprise a second actuator that is configured for directly or indirectly affects, changes, produces, or creates a physical phenomenon. Claim 6 requires the physical phenomenon to comprise temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current.

28. Claim 7 requires the first device to comprise an additional microphone for capturing human voice data, and wherein the capturing comprises capturing, by the additional microphone, the first and second human voice data. Claim 8 requires that the microphone and additional microphone be arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array.

29. Claim 9 requires the microphone to be an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone.

30. Claim 10 requires the IP address be stored in a volatile or non-volatile memory.

31. Claim 11 requires the first device be further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type.

32. Claim 12 requires the IP address to be a static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address.

33. Claim 13 requires the WLAN be according to, or based on, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, Institute of Electrical and Electronics Engineers (IEEE) IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac. Claim 14 requires the WLAN to use a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.

34. Claim 15 requires the first device to be further configured or communicating over a cellular telephone network that is a Third Generation

(3G) network that uses Universal Mobile Telecommunications System (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global System for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution.

35. Claim 16 requires the first device be further configured for communicating over a cellular telephone network that is a Fourth Generation (4G) network that uses, or is compatible with, Evolved High Speed Packet Access (HSPA+), Mobile Worldwide Interoperability for Microwave Access (WiMAX), Long-Term Evolution (LTE), LTE-Advanced, Mobile Broadband Wireless Access (MBWA), or is based on IEEE 802.20-2008 standard.

36. Claim 17 requires the second device to be integrated in, part of, or entirely included in, a household appliance having a primary function. Claim 18 requires the primary functionality of the appliance to be associated with food storage, handling, or preparation. Claim 19 requires the primary function of the appliance to be heating food, and wherein the appliance is a microwave oven, an electric mixer, a stove, an oven, or an induction cooker. Claim 20 requires the appliance to be a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker. Claim 21

requires the primary function of the appliance to be associated with environmental control, and the appliance is part of an Heating, Ventilation and Air Conditioning (HVAC) system. Claim 22 requires the primary function of the appliance to be associated with temperature control, and wherein the appliance is an air conditioner or a heater. Claim 23 requires the primary function of the appliance to be associated with cleaning, wherein the appliance primary function is associated with clothes cleaning and the appliance is a washing machine, or wherein the appliance is a vacuum cleaner. Claim 24 requires the appliance to be an answering machine, a telephone set, a home cinema system, a High Fidelity (HiFi) system, a Compact Disc (CD) or Digital Video Disc (DVD) player, an electric furnace, a trash compactor, a smoke detector, a light fixture, or a dehumidifier.

37. Claim 25 requires the first device be further configured to operate within a building, and wherein the server device is external to the building.

38. Claim 26 requires the first device be further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays. Claim 27 requires the electric light source to comprise a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent

lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode.

39. Claim 28 requires the sounding to comprise playing digital audio content that is pre-recorded or synthesized, or simulating the voice of a human being or generating music, or sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice. Claim 29 requires the sounder to comprise an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker. Claim 30 requires the sounder to comprise an electric bell, a buzzer, a chime, a whistle, or a ringer.

40. Claim 31 requires the first device to further comprise an enclosure that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors. Claim 32 requires the enclosure be a wearable enclosure configured to be wearable on a human body.

41. Claim 33 requires the sounder be configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves.

D. Prosecution History

42. I understand that during prosecution of the ’721 patent, the Examiner did not issue any prior art-based Office Actions.

43. The application that matured into the '721 Patent was filed on May 8, 2025. (EX1003, 366-369.) The Patent Office granted the Applicant's petition to make special based on age on the basis of the inventor being over 65 years old. (EX1003, 364-365, 578-580.) Only a single Office Action was issued. And this Office Action contained mere double-patenting rejections with respect to three parent patents: U.S. Patent No. 11,128,710, U.S. Patent No. 12,316,706, and U.S. Patent No. 12,010,174. (EX1003, 203-228.) The Applicant submitted a Terminal Disclaimer in response. (EX1003, 191-196.) The Examiner then allowed the claims on July 3, 2025, within 2 months of filing, citing the entirety of claim 1 as the reason for allowance. (EX1003, 170-176.)

44. However, I note that the examination of the earliest patent, the '710 patent, was deficient. The application for the '710 patent underwent lengthy examination, with a total of ***six Office Actions***. The examiner repeatedly issued substantially identical rejections, which were ***affirmed by the Board***. Despite this, following only minor amendments, the Examiner abruptly allowed the application without offering any substantive explanation for the change in position. (*See* EX1004, 8-14.) The original claim 1 merely recited processing a single set of "human voice data." (EX1004, 1304.) After a non-final Office Action and a final Office Action, the Applicant appealed. (EX1004, 946-971,

861-887, 859.) The Examiner then reopened prosecution with a second non-final Office Action and a second final Office Action. (EX1004, 705-733, 635-670.) At this stage, the Applicant amended claim 1 for the first time, revising it so that the server transmits a message to a ***controlled device*** (rather than to a client device) to operate an actuator in the controlled device. (EX1004, 622-633.) The Board affirmed the Examiner's rejections of all claims, agreeing that the cited references disclosed the claimed concept of controlling a controlled device using voice captured by a client device (specifically, ***capturing human voice data by a client device, sending it via a wireless network to an Internet server for processing, receiving a message from the Internet server, and operating an actuator in a controlled device in response to the received message***). (EX1004, 389-421.)

45. After the PTAB decision, the Applicant further amended claim 1 of the '710 patent application to recite processing "first human voice data" and "second human voice data," both processed by the server to operate a first actuator and a second actuator in the controlled device. (EX1004, 344-346.) The Examiner issued two additional Office Actions in response. (EX1004, 229-248, 145-167.) The Applicant then amended claim 1 again, this time merely reciting that the first human voice data is processed to control a ***client device*** (instead of the controlled device). (EX1004, 121-123.) The only

difference between the claim the Board had already found obvious and the ultimately allowed claim in the '710 patent is that the ***client device can receive a voice command that causes operation of the client device itself***, in addition to receiving a voice command that causes operation of a controlled device (which the Board has found to be obvious). (EX1004, 121-123.) The claimed network architecture has been found to be obvious by the Board. (EX1004, 389-421.) Despite this minimal distinction, the Examiner allowed the application for the '710 patent, without providing any substantive justification for the abrupt shift from repeated rejections to allowance. (EX1004, 8-14.)

46. But, the Examiner erred in allowing the '710 patent. A device that receives a voice command and operates accordingly (*e.g.*, providing a visual or auditory output) was well-known at the time of the invention. The allowance was due to perfunctory searching that does not capture the new claim scope. After the Board decision, the Examiner should have had a laser focus on the new limitations and tailor the search strategy accordingly. However, instead of targeting the new limitations, the Examiner used completely identical patent classification search strings that he used before, and most of these are related to transportation and vehicles. (*See*, *e.g.*, EX1004, 25-74, 173-220.) I note that this is a critical error. Essentially, the new limitations were not properly searched for, let alone thoroughly considered and examined.

47. I note that the Examiner also performed a few cursory queries on Google Patents. (EX1004, 24.) Even those queries merely strung together two or three sweeping concepts such as "voice recognition," "external remote server," and "appliances." They were recycled concepts that the Board already rejected, not tailored search strings focusing on the new limitations. (EX1004, 24.)

48. The search strategy shows that the Examiner never appreciated how broad these new limitations were. For example, the Examiner searched for "actuator" when a broader search that includes speakers and displays would be more appropriate in light of the specification. (EX1004, 24.) Had he looked for devices that actuate displays or speakers, he would have uncovered references that cover the new limitations because they are conventional.

49. I note that the examination of the '706 patent and '174 patent were similarly flawed. Following the grant of the '710 patent, the '706 patent and '174 patent sailed through examination without any prior art based Office Action. For the '174 patent, there was only a single Office Action containing double patenting rejection with respect to parent patent U.S. Patent No. 11,128,710. (EX1052, 90-103.) After a Terminal Disclaimer was submitted, the Examiner allowed the claims, citing the entirety of claim 1 as the reason for allowance. (EX1052, 78-86, 38-40.) For the '706 patent, there was only a

single Office Action containing double patenting rejection with respect to two parent patents: U.S. Patent No. 11,128,710 and U.S. Patent No. 12,010,174. (EX1051, 45-64.) After a Terminal Disclaimer was submitted, the Examiner allowed the claims, citing the entirety of claim 1 as the reason for allowance. (EX1051, 12, 17-18.)

50. Because the Examiner relied on its similarity to the parent patents in allowing the '721 patent, the examination of the '721 patent is flawed at least for the same reasons above.

V. RELEVANT LEGAL STANDARDS AND DEFINITIONS

A. Obviousness

51. I have been informed that a claim is unpatentable if the differences between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious at the time the claimed subject matter was made to POSITA to which the subject matter pertains. I have also been informed that a patent claim may be obvious to a POSITA in view of the prior art teachings of a single reference or a combination of references.

52. To assess obviousness, I have been informed that I am to consider the scope and content of the prior art, the differences between the prior art and the claim, the level of ordinary skill in the art, and any secondary considerations to the extent that they exist.

53. I have also been informed that there are principles that may be used as further guidance in obviousness analysis (especially when considering combinations of references), which include considering whether:

- the claimed subject matter is simply a combination of prior art elements according to known methods to yield predictable results;
- the claimed subject matter is a simple substitution of one known element for another to obtain predictable results;
- the claimed subject matter uses known techniques to improve similar devices or methods in the same way;
- the claimed subject matter applies a known technique to a known device or method that is ready for improvement to yield predictable results;
- the claimed subject matter would have been "obvious to try" choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success;
- there is known work in one field of endeavor that may prompt variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations would have been predictable to a POSITA;
- there existed at the time of conception and reduction to practice a

known problem for which there was an obvious solution encompassed by the patent's claims; and

- there is some teaching, suggestion, or motivation in the prior art that would have led a POSITA to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed subject matter.

54. I have been informed that there are secondary considerations that may tend to show that a claimed invention would not have been obvious to a POSITA. These include, for example, commercial success, long-felt but unsolved needs, failure of others, and unexpected results. I have been further informed that there must be a relationship, or nexus, between any secondary considerations and the novelty of the claimed invention.

B. Prior Art

55. I am informed and understand that prior art includes patents and printed publications that existed before the effective filing date of the claimed invention. The effective filing date is the earliest claimed priority date or the earliest filing date of the application that describes the subject matter. A patent or published patent application is prior art if it was filed before the earliest filing date of the claimed invention, absent a few exceptions. I am informed that a patent or published patent application is entitled to have its effective filing date as the filing date of its provisional application if the disclosure of the provisional

application provides support for the claim(s) of the patent or published patent application.

VI. LEVEL OF ORDINARY SKILL IN THE ART FOR THE '721 PATENT

56. I am informed and understand that various factors can be considered in determining a person of ordinary skill in the art (POSITA). I am informed and understand that those factors include: (1) the education level of the inventors; (2) the type of problems encountered in the art; (3) prior art solutions to those problems; (4) the rapidity with which innovations are made; (5) sophistication of the technology; and (6) education level of active workers in the field.

57. A POSITA in the field of the '721 patent would have a bachelor's degree in electrical engineering, computer engineering, or equivalent field, and two years of experience working on remote device control and/or networked devices. An advanced graduate degree in the disciplines cited above may make up for the work experience indicated above.

58. I agree with this level of ordinary skill for the '721 patent. Based on my experience, as described above and in Appendix A, I meet the definition of the qualifications of a POSITA, now and as of the alleged priority date. For purposes of this declaration, in general, and unless otherwise noted, my

statements and opinions, such as those regarding my own experience, what a POSITA would have understood or known generally (and specifically related to the references I consulted herein), and how a POSITA would have understood and used the terminology found in the '721 patent, reflect the knowledge that existed in the relevant field at the time the '721 patent priority application was filed, and with the state of the art at that time.

VII. CLAIM CONSTRUCTION

59. I have been informed that claim construction requires consideration of the words of the claims themselves, the specification, the prosecution history, and extrinsic evidence concerning relevant scientific principles, the meaning of technical terms, and the state of the art. I have also been informed that the specification is usually definitive, as it was the single best guide to the meaning of a disputed term. I have been informed that claim terms are presumed to carry their ordinary and customary meaning as understood by a POSITA at the time of the purported invention. I have been informed that, in the context of this proceeding, which requires the Board to construe the terms only to the extent necessary to resolve the issue at hand—invalidity of the claims under 35 U.S.C. §103, Petitioner believes the Board need not construe any of the claims because, as set forth in the Declaration and Petition, the claims are invalid under

35 U.S.C. §103 under any reasonable interpretation, and thus, no claim construction is necessary to resolve any issue before the Board.

VIII. PRIOR ART

60. I understand that each of the references discussed herein qualifies as prior art against the '721 patent:

61. U.S. Publication No. 2013/0132094 ("Lim," EX1005) was filed on November 17, 2011, and published on May 23, 2013.

62. U.S. Publication No. 2012/0198339 ("Williams," EX1006) was filed on January 28, 2011, and published on August 2, 2012.

63. U.S. Patent No. 7,529,677 ("Wittenberg," EX1008) was filed on January 21, 2005, and published on May 5, 2009.

64. U.S. Publication No. 2012/0265528 ("Gruber," EX1009) was filed on September 30, 2011, and published on October 18, 2012.

65. "Springer Handbook of Speech Processing" ("Benesty," EX1010) was written by Jacob Benesty et al., and published by Springer-Verlag Berlin Heidelberg by at least December 11, 2007.

66. "Wireless Communications" ("Goldsmith," EX1040) was written by Andrea Goldsmith, and published by Cambridge University Press by at least January 20, 2006.

67. U.S. Publication No. 2013/0052946 ("Chatterjee," EX1012) was filed on August 23, 2011, and published on February 28, 2013.

68. U.S. Patent No. 8,340,975 ("Rosenberger," EX1013) was filed on October 4, 2011, and published on December 25, 2012.

69. "Springer Handbook of Acoustics" ("Rossing," EX1014) was edited by Thomas D. Rossing et al., and published by Springer Science+Business Media, LLC by at least July 23, 2007.

70. "Wireless Home Networking Simplified" ("Doherty," EX1007) was written by Jim Doherty et al., and published by Cisco Press by at least January 16, 2007.

IX. ANALYSIS AND OPINIONS

A. Ground 1: Lim Renders Obvious Claims 1, 5-6, 9-15, 17, 21-22, 24-31, 33

1. Summary of Lim

71. Lim was filed on November 17, 2011, and is prior art under §102(e). (EX1005, 1.) The Examiner never found, and Applicant never cited, Lim during prosecution. Lim relates to "controlling devices such as universal remote controls" and "appliances to be controlled." (EX1005, [0007].) Lim discloses that "a ***controlling device*** may be provisioned with means to receive voice input from a user, *e.g.*, **a *microphone***." (EX1005, [0007] (emphasis

added)[1]) Lim discloses that the controlling device would be a ***smart device 110*** in the form of a smart phone, as shown in FIGs. 1b and 5b. (EX1005, [0023].) The ***controlling device*** is "configured to control ***various controllable appliances, such as a television 102, a set top box ('STB') 104, a DVD player 106 and an AV receiver 108***." (EX1005, [0017].)

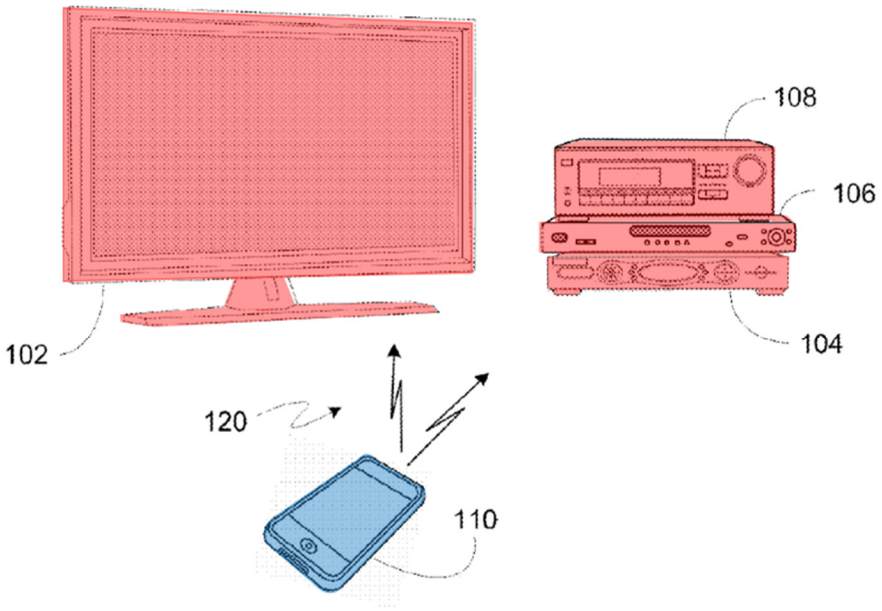


Figure 1b

72. Lim discloses that the ***controlling device*** (*e.g.*, ***first device***) receives speech commands and responds with an ***audio output*** from the device itself. For example, Lim discloses a "setup and configuration" stage for the

[1] All emphasis and annotations, such as highlighting and coloring, are added unless otherwise stated.

first/controlling device. (EX1005, [0025].) During this stage, the ***controlling device*** can "prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device." (EX1005, [0027].) As shown in FIG. 4 below, the ***first device*** responds to user's speech inputs with additional prompts, which can be auditory (*e.g.*, via a speaker which sounds digital audio contents). "This and subsequent prompts for voice input may take the form of an illuminated LED and/or 'speak' symbol, a displayed text or image, ***a synthesized spoken message*** (*e.g.*, where controlling device user feedback means 202 comprises ***speech synthesis means and a speaker***, or where a smart device app has access to speaker 330), etc." (EX1005, [0027].)

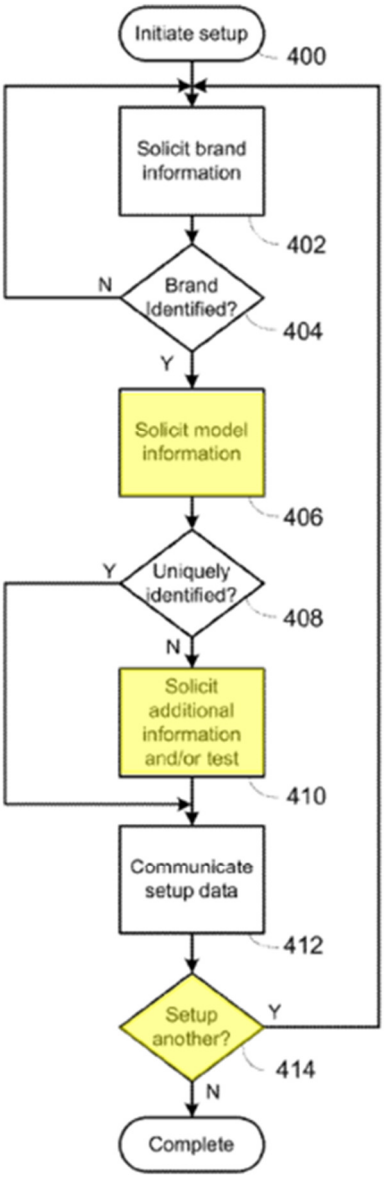


Figure 4

"Set up and configuration of controlling device" Lim, [0025].

73. Lim also discloses that the ***controlling device*** can be used to control appliances (*e.g.*, ***second devices***). Specifically, the ***controlling device*** receives speech commands and transmits ***control commands*** to ***second devices/appliances***. For example, "the ***controlling device 100*** may be capable

of ***transmitting commands*** to the ***appliances***, using any convenient IR, RF, Point-to-Point, or ***networked protocol 120***, to cause the appliances to perform operational functions. While illustrated in the context of a television 102, STB 104, DVD player 106 and AV receiver 108, it is to be understood that controllable appliances may include, but need not be limited to, ***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers***, etc." (EX1005, [0017].) As shown in FIG. 5b and FIG. 1b below, the ***controlling device 110*** transmits commands using the ***appliances' networked protocol 120***.

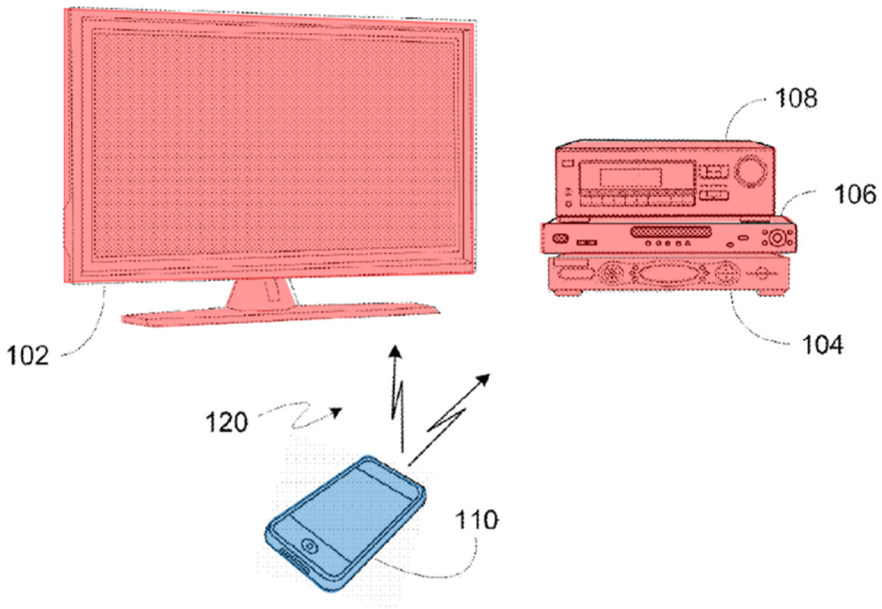


Figure 1b

74. FIG. 6 below depicts an exemplary process in which the ***first device*** receives speech commands and transmits commands to ***appliances*** in block 612 accordingly.

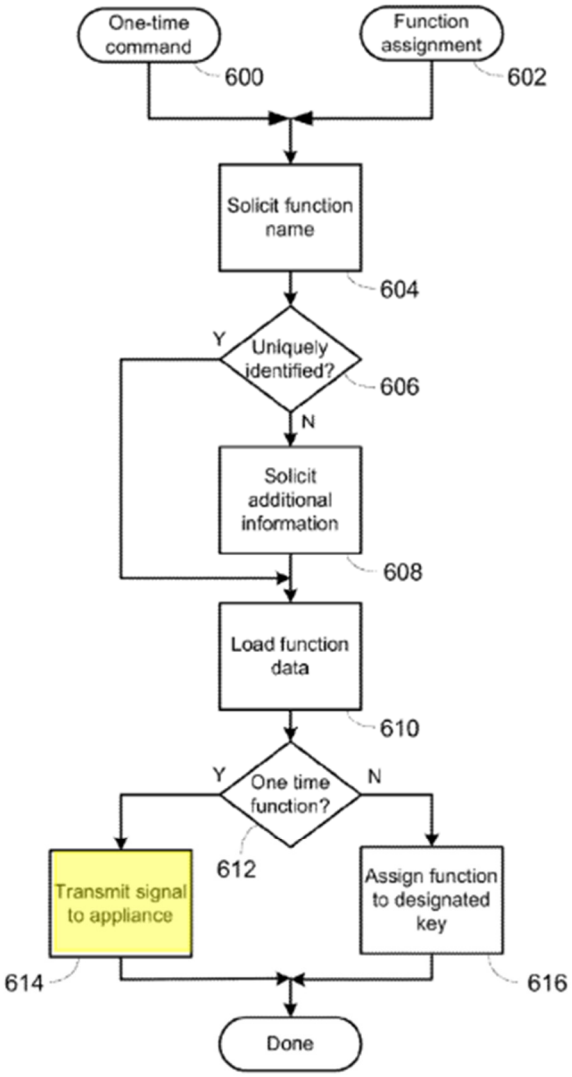


Figure 6

"Appliance command" Lim, [0035].

75. In Lim, the speech commands are transmitted to a ***server*** for processing. Further, Lim discloses (1) the ***server*** is accessible via the ***Internet***

and (2) the ***controlling device*** connects to the ***Internet*** via ***WiFi*** through a gateway device such as a router. For example, Lim discloses that "at step 402 the ***controlling device*** may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. . . . The resulting spoken syllables may be digitized and presented to a speech recognition means for processing . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027], [0034]-[0035].)

76. An exemplary network architecture is shown in FIG. 5b below. The "***remote server 500*** [is] accessible via a network 502 such as the ***Internet*** or PSTN." Lim discloses that "access to such network may be . . . via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].)

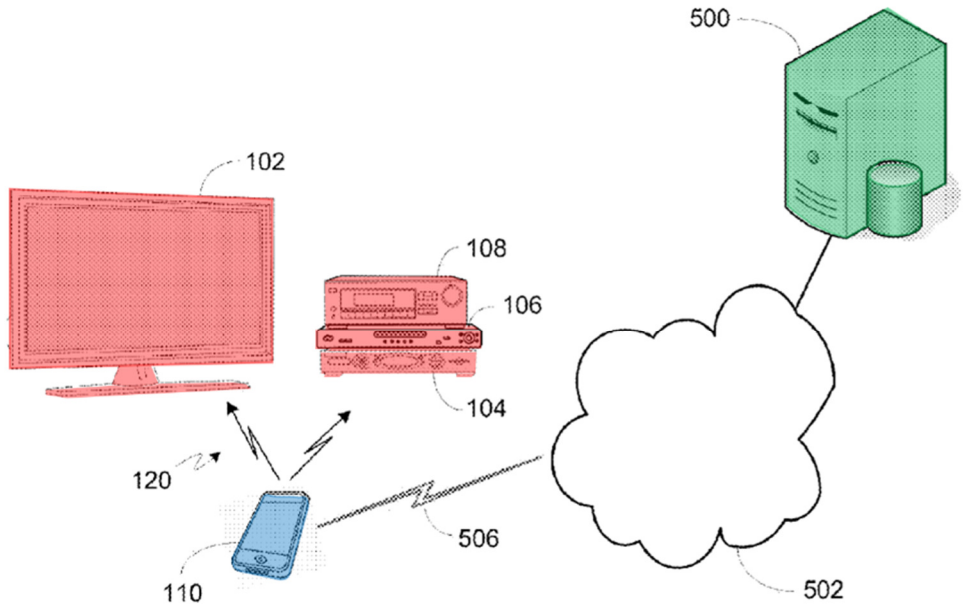


Figure 5b

77. Lim discloses that the ***controlling device*** has a microphone, a display, a speaker, various sensors, one or more processors, and WiFi/Bluetooth capabilities. "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . ***processing means 300*** which may comprise both an application processing section 302 and an RF/DSP processing section 304; an ***LCD display 306***; a keypad 308 which may comprise physical keys, touch keys overlaying LCD 306, or a combination thereof; a subscriber identification module (SIM) card 310; memory means 312 which may comprise ROM, RAM, Flash, or any combination thereof; ***WiFi and/or Bluetooth wireless interface(s) 314***; a wireless telephony interface 316; power management circuitry 318 with

associated battery 320; a USB interface 322 and connector 324; an audio management system 326 with associated ***microphone 328, speaker 330***, and headphone jack 332; built-in or add on IR communication means comprising a transmitter and/or receiver 336 with associated IR output diode(s) 334; and various optional accessory features 338 such as a digital camera, ***GPS, accelerometer***, etc." (EX1005, [0023].)

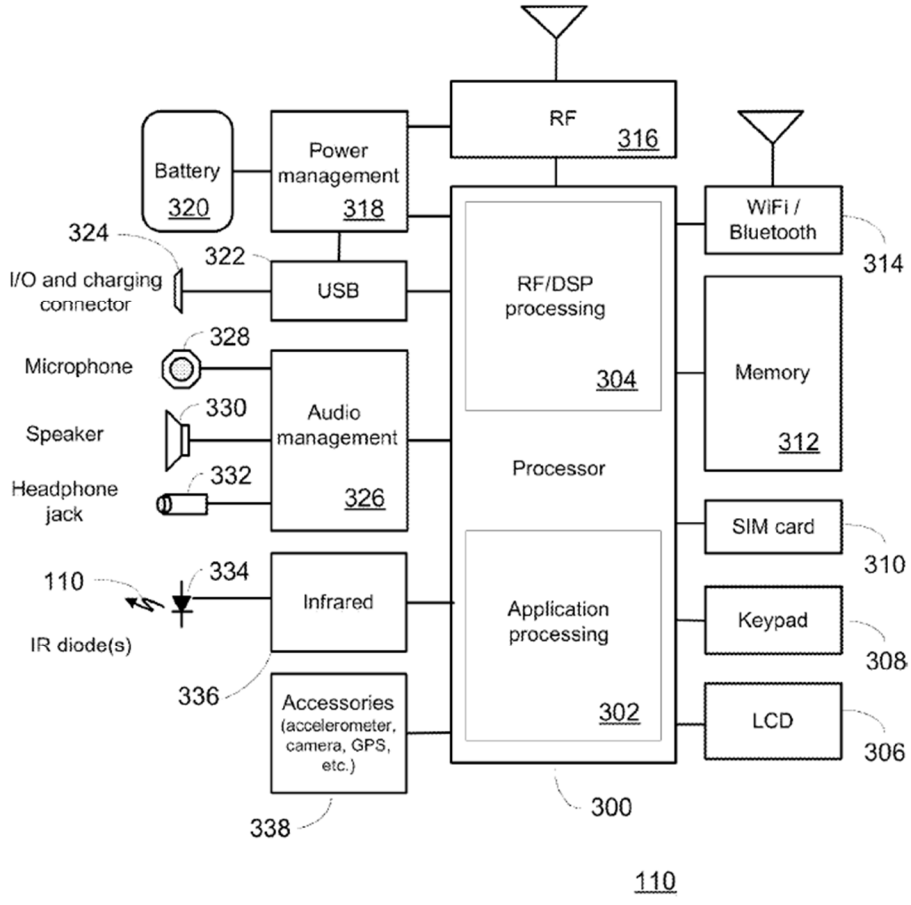


Figure 3

2. Claim 1

a. 1[PR]: "A first device for use with a Wireless Local Area Network (WLAN) in a building, the first device comprising:"

78. Lim discloses a ***first device*** and for use with a Wireless Local Area Network (WLAN) in a ***building***.

79. First, Lim's ***controlling device*** is a ***first device***. Lim discloses that "a ***controlling device*** may be provisioned with means to receive voice input from a user, *e.g.*, a microphone". (EX1005, [0007].) The "***controlling device 100*** is configured to control ***various controllable appliances, such as a television 102, a set top box ('STB') 104, a DVD player 106 and an AV receiver 108***." (EX1005, [0017].) Lim discloses that the ***controlling device*** may be "a ***smart device 110***, which device may comprise any of a smart phone, a tablet computer, a portable email device, an e-book, a hand-held games and/or game controller, a portable media player, etc." as shown in FIGs. 1b and 5b. (EX1005, [0018].)

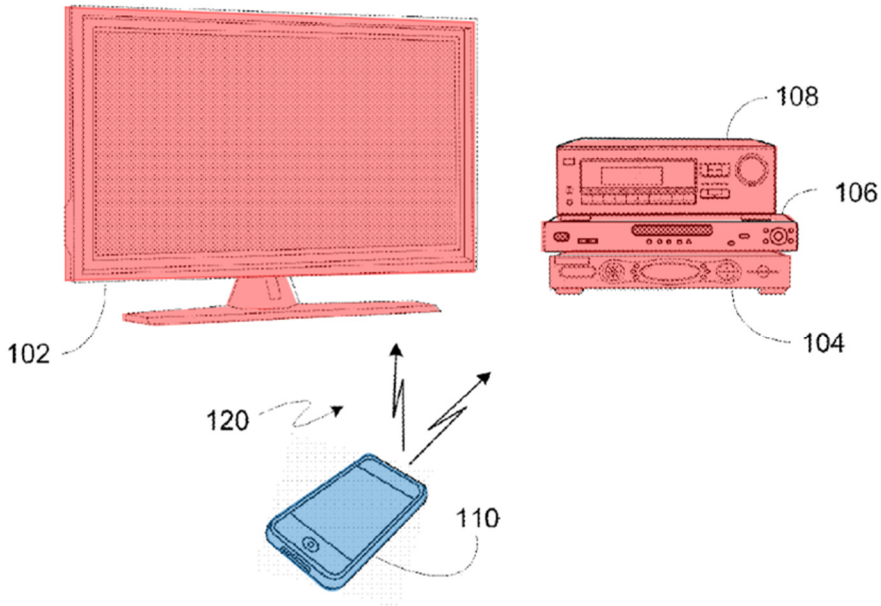


Figure 1b

80. Lim's ***first/controlling device*** is for use with a ***Wireless Local Area Network (WLAN)***. Lim's ***first device*** (*e.g.*, ***controlling device*** 100, 110) accesses ***remote server 500*** via WiFi through a router. Referring to FIG. 5b below, Lim discloses that "all or part of the said speech recognition processing [of human voice data captured by the ***first/controlling device***] may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the Internet" and "access to such network may be . . . via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].)

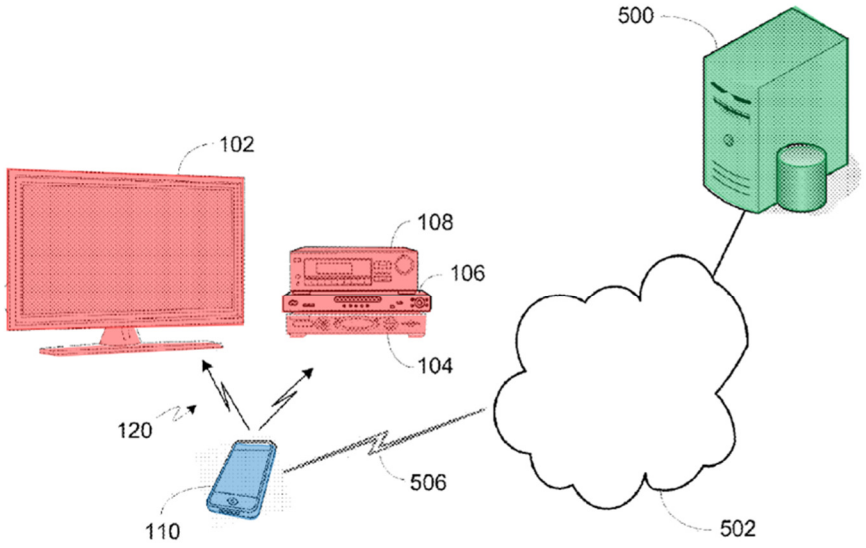


Figure 5b

81. A POSITA would have understood that what Lim describes is a WLAN network because it discloses a "local network connection" such as "WiFi" and it was common knowledge that WiFi implements a WLAN. (EX1005, [0027].) "Wireless local area networks (***WLANs***) are built around the family of IEEE 802.11 standards. . . . The 802.11a/b/g family of standards are collectively referred to as ***Wi-Fi***, for wireless fidelity." (EX1040, 600-601.) "***Wi-Fi*** is a registered trademark owned by the Wi-Fi Alliance, and it refers to ***wireless local area network (WLAN)*** products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards." (EX1013, 7:48-51.) This is consistent with the description of the '721 patent which states that "[i]n order to allow interoperability between equipment manufactured by different vendors, few ***WLAN standards*** have evolved, as part of the ***IEEE***

802.11 standard group, ***branded as WiFi*** (www.wi-fi.org). IEEE 802.11b describes a communication using the 2.4 GHz frequency band and supporting communication rate of 11 Mb/s, IEEE 802.11a uses the 5 GHz frequency band to carry 54 MB/s and IEEE 802.11g uses the 2.4 GHz band to support 54 Mb/s." (EX1001, 8:25-32.)

82. Lim discloses or certainly suggests that the ***first device*** is in a ***building***. Lim discloses that the ***smart device 110*** "may comprise any of a smart phone, a tablet computer, a portable email device, an e-book, a hand-held games and/or game controller, a portable media player." (EX1005, [0018].) It would have been obvious that these are devices that users can carry with them, including when the user is at ***home***, to conveniently control home appliances from home. For example, at the time of the invention, a "portable device [can be] readily carried by the operator" and it "allows an operator to easily and readily control ***home activities*** without having to directly interface with or use a PC, and which allows the operator to program and control the home gateway ***from essentially any area of the home*** without having to access a PC." (EX1021, 1:53-61, 2:19-26.)

b. 1[A]: "a WLAN transceiver for communicating over the WLAN;"

83. Lim discloses a ***first/controlling device*** with a ***WLAN transceiver*** for communicating over the WLAN.

84. Lim discloses that the "***controlling device*** 100 for use in commanding the functional operations of one or more appliances may include, . . . , Zigbee, ***WiFi*** , or Bluetooth ***transceiver***[.]" (EX1005, [0019], [0027].) As discussed, it was common knowledge that WiFi is a type of WLAN. (*Supra* § IX.A.2.a.; EX1040, 600-601; EX1013, 7:38-51; EX1001, 8:25-32.) Thus, a POSITA would have understood that a "WiFi [] transceiver" is a WLAN transceiver. While this statement refers to controlling device 100, it would have been obvious to a POSITA that the ***smart device 110*** similarly has a WLAN transceiver.

85. With respect to the ***smart device 110***, Lim discloses that "[a] representative smart device . . . in the form of a smart phone, may include . . . ***WiFi and/or Bluetooth wireless interface(s) 314***; a wireless telephony interface 316[.]" (EX1005, [0023].) A POSITA would have understood that a "WiFi wireless interface" refers to the hardware and software that enable WiFi communication, which necessarily include a WLAN transceiver. "The LAN/WAN network 100d may comprise suitable logic, circuitry, ***interfaces***, and/or code that may be ***operable to enable implementation*** of one or more wired and/or ***wireless LAN*** or WAN standards and/or protocols. . . . Exemplary

LAN technologies may comprise, for example, those based on IEEE 802.11 standards, including, for example, ***WiFi-based networks***." (EX1022, [0067].) Indeed, in or before 2012, mobile phones were commonly equipped with a WiFi transceiver to enable WiFi connectivity. "Each smartphone [smartphones that can communicate with existing WiFi infrastructure] nowadays is equipped with a WiFi transceiver . . . ." (EX1035, 3.) "[W]ireless mobile devices are being made with integrated WLAN transceivers[.]" (EX1027, 1:20-23.) Thus, a POSITA would have understood that Lim's ***controlling device 110*** has a WLAN transceiver.

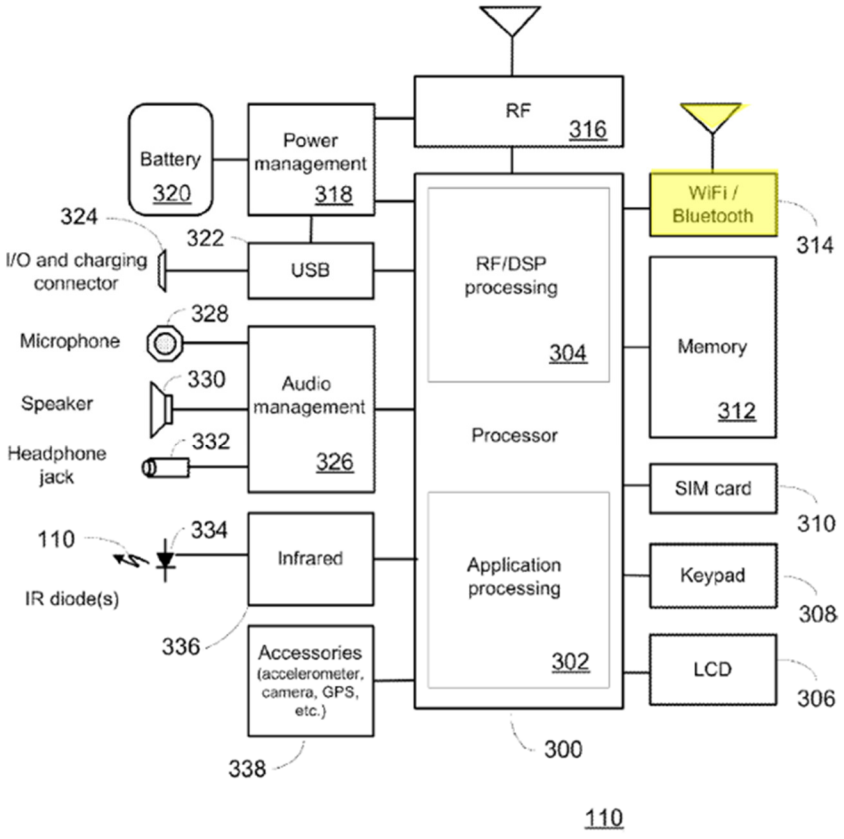
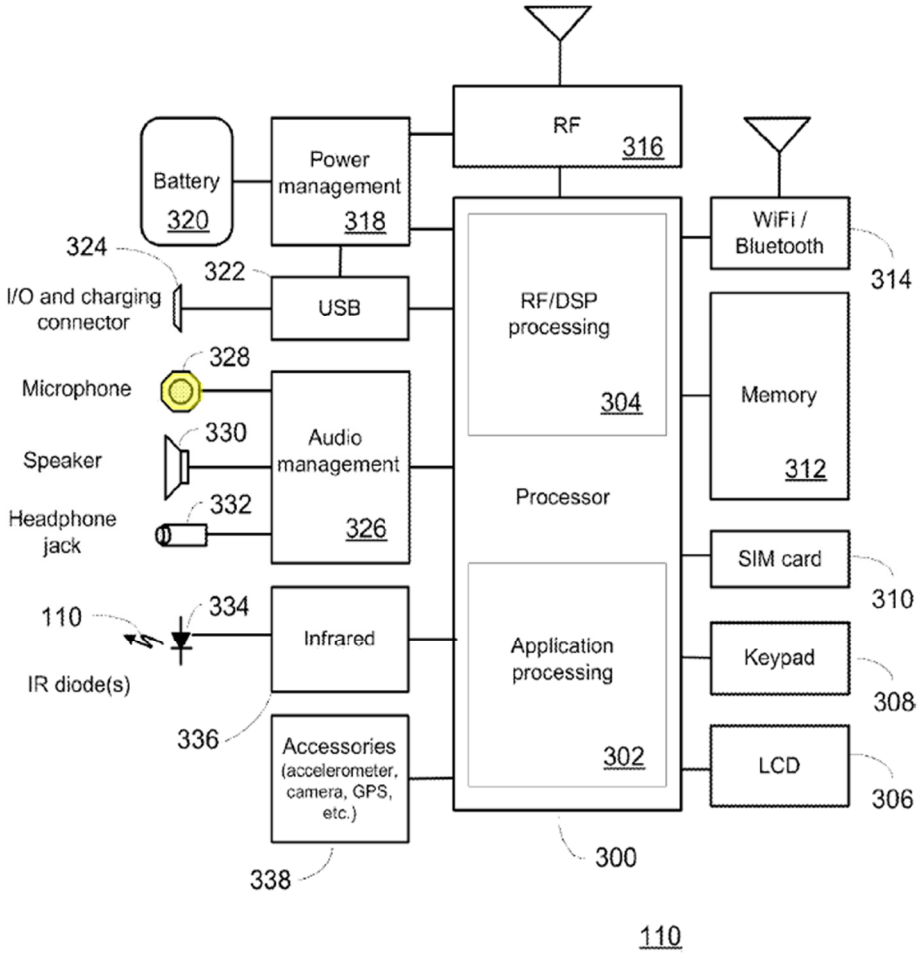


Figure 3

c. 1[B]: "a microphone for capturing human voice data;"

86. Lim discloses this element. First, Lim discloses that the ***first/controlling device*** comprises a ***microphone***. For example, "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include . . . speech input means comprising ***a microphone 224*** and associated circuitry 222 for example an analog to digital converter, speech digitizer, etc., as known in the art." (EX1005, [0019].) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated ***microphone 328***, speaker 330, and headphone jack 332 [.]" (EX1005, [0023].)

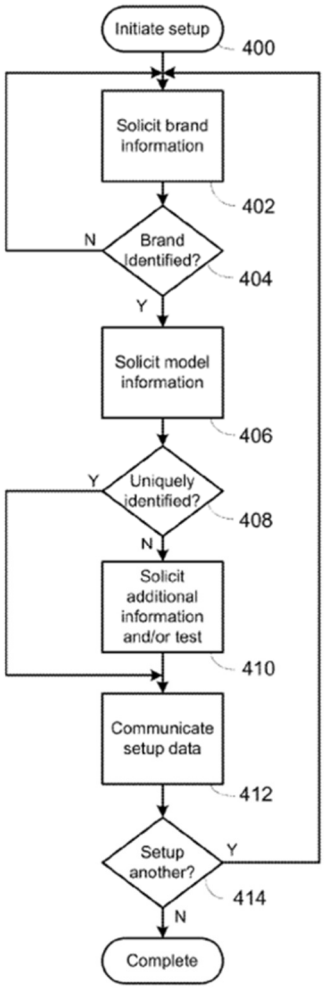


87. Lim's ***microphone*** in the ***first/controlling device*** captures ***human voice data***. Lim discloses many instances where the ***controlling device*** receives speech commands. For example, referring to FIG. 4 below, "during setup and configuration of the ***controlling device***, a user may provide appliance identification via ***spoken input***. . . ." (EX1005, [0025], [0013].)

88. FIG. 4 discloses numerous human voice data. "Once the setup mode has been initiated, at step 402 the controlling device may prompt the user to ***speak the brand name of the appliance*** to be added to the configuration of

the controlling device." (EX1005, [0027].) Further, "[a]t step 404, if the speech recognition means is unable to uniquely identify the desired brand, step 402 may be repeated in order to solicit additional identifying information from the user. For example, a user may be presented with a short list of possible brand identities from which to select; may be prompted to ***spell the first few characters of the brand name***; may be requested to ***speak the name of a geographic region or a service provider***; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].) "Once a brand has been successfully identified, at step 406 the user may be prompted to ***provide spoken appliance type and/or model number information*** for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, *e.g.*, to further narrow down the codeset choices." (EX1005, [0029].) "If the ***initially provided model information*** [*e.g.*, provided previously in step 406] is insufficient to uniquely identify the appliance, at step 410, additional user input may be solicited. For example, a user may be requested to ***indicate a device type*** if this was not already provided; to ***provide geographic region or service provider***; to select from a series of displayed images of appliances and/or OEM remote controls; etc." (EX1005, [0029].) "Thereafter, at step 414 the user may be prompted to ***indicate if they wish to configure additional appliances***. If so, the process may be repeated

starting at step 402 [where the controlling device prompts the user to speak the brand name of the appliance to be added to the configuration]." (EX1005, [0030].) "The resulting spoken syllables may be digitized and presented [by the first/controlling device] to a speech recognition means . . . ." (EX1005, [0027].) It would have been obvious to a POSITA that the ***controlling device***'s microphone captures these voice inputs.

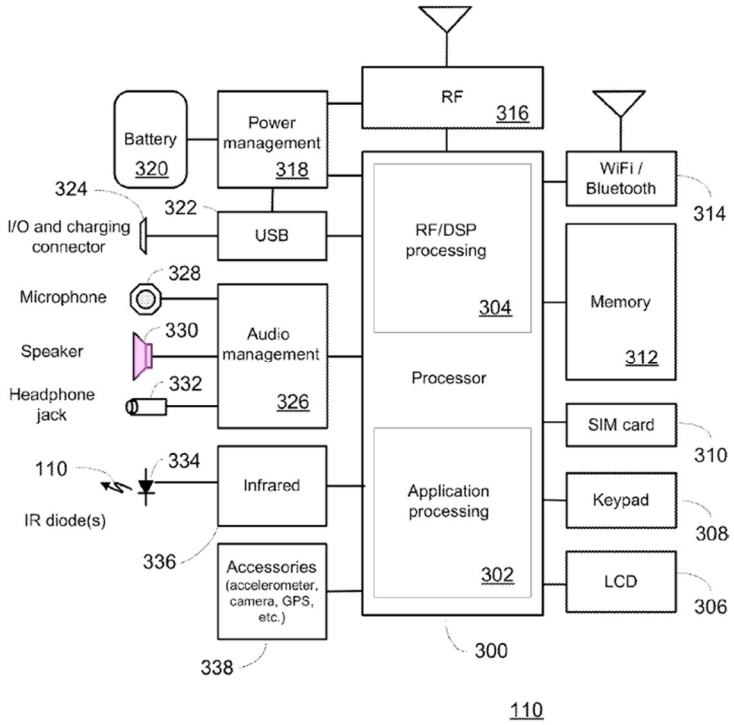


"Set up and configuration of controlling device" Lim, [0025].

Figure 4

d. 1[C]: "and a sounder for converting an electrical energy to audible sound waves;"

89. Lim's ***first/controlling device*** comprises a sounder for converting an electrical energy to audible sound waves. A POSITA would have understood that this element refers to a speaker. Lim's ***controlling device*** has a ***speaker***. For example, "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . a means 202 to provide feedback to the user (*e.g.*, one or more LEDs, illuminable keys, display, ***buzzer, voice synthesizer with associated speaker***, and/or the like)[.]" (EX1005, [0019].) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated microphone 328, ***speaker 330***[.]" (EX1005, [0023].)



90. The specification of the '721 patent confirms that a speaker is a sounder that converts electrical energy into audible sound waves. "A ***loudspeaker (a.k.a. speaker) is a sounder*** that produces sound in response to an electrical audio signal input, typically audible sound. The most common form of loudspeaker is the electromagnetic (or dynamic) type, uses a paper cone supporting a moving voice coil electromagnet acting on a permanent magnet. Where accurate reproduction of sound is required, multiple loudspeakers may be used, each reproducing a part of the audible frequency range. A loudspeaker is commonly optimized for middle frequencies; tweeters for high frequencies; and sometimes supertweeter is used which is optimized for the highest audible frequencies." (EX1001, 105:61-106:4.) "An actuator may be an

electroacoustic actuator, such as a sounder which converts electrical energy to sound waves transmitted through the air, an elastic solid material, or a liquid, usually by means of a vibrating or moving ribbon or diaphragm." (EX1001, 105:48-52.)

91. At the time of the invention, a POSITA would have understood that a speaker is a sounder for converting an electrical energy to sound waves. For example, the sounder can convert electrical energy into sound waves transmitted through the air through the movement of a diaphragm, often cone-shaped or dome-shaped, which produces "sound" or acoustical energy. "Most loudspeakers use the moving-coil principle (Figure 10.8). The coil is wound on a former at the apex of a stiff diaphragm which is held in flexible supports so that the coil is the intense field of a permanent magnet. A signal current through the coil will then ***vibrate the diaphragm, producing sound waves*** at the same frequency and with the same waveshape as the signal." (EX1018, 122, 123, Figure 10.8.)

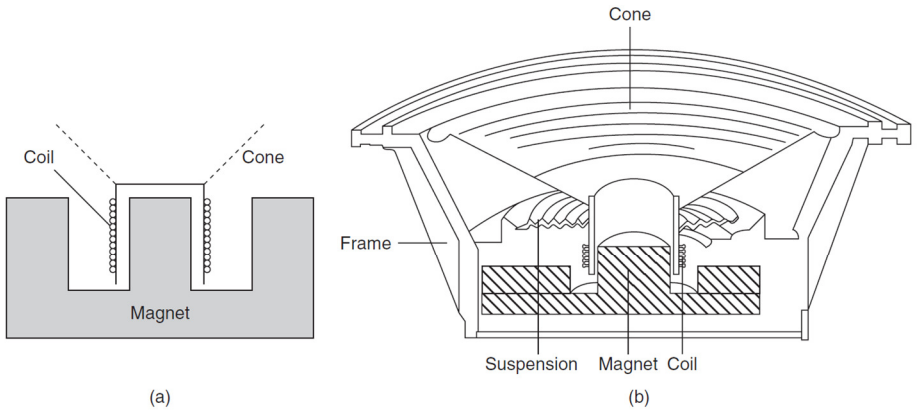


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

92. "[F]ew loudspeakers convert directly between electrical and acoustical energy. Instead, the electrical energy is converted to ***mechanical energy***, specifically the ***movement of a diaphragm*** . . . which then radiates acoustical energy." (EX1014, 763.)

93. As discussed above, a speaker is a sounder, thus a POSITA would have understood that Lim's ***speaker 330*** is a sounder for converting an electrical energy to audible sound waves.

94. Additionally, Lim's ***controlling device 100*** "may include . . . [a] ***buzzer*** [.]" (EX1005, [0019].) While this statement refers to controlling device 100, it would have been obvious that ***smart device 110*** is similarly equipped with a ***buzzer***. A buzzer is another example of a sounder, consistent with the description of the '721 patent. The '721 patent explains that, "***[a] sounder may be a buzzer (or beeper)***, a chime, a whistle or a ringer. Buzzers

may be either electromechanical or ceramic-based piezoelectric sounders which make a high-pitch noise, and may be used for alerting." (EX1001, 106: 56-62.)

e. 1[D]: "and one or more processors programmed with computer program instructions that, when executed, cause the first device to:"

95. Lim discloses this element. For example, Lim discloses that the controlling device 100 would include, "a ***processor 200*** coupled to a ROM memory 204, a RAM memory 205, . . . a non-volatile read/write memory 206" and "some or all of the memories 204, 205, 206 may include ***executable instructions*** (collectively, the controlling device program memory) that are intended to be ***executed by the processor 200 to control the operation of the controlling device 100***, as well as data which serves to define the necessary control protocols and command values for use in transmitting command signals to controllable appliances (collectively, the command data). In this manner, the processor 200 may be programmed to control the various electronic components within the controlling device 100, *e.g.*, to monitor the key matrix 216, to ***cause the transmission of signals***, etc." (EX1005, [0019], [0020].)

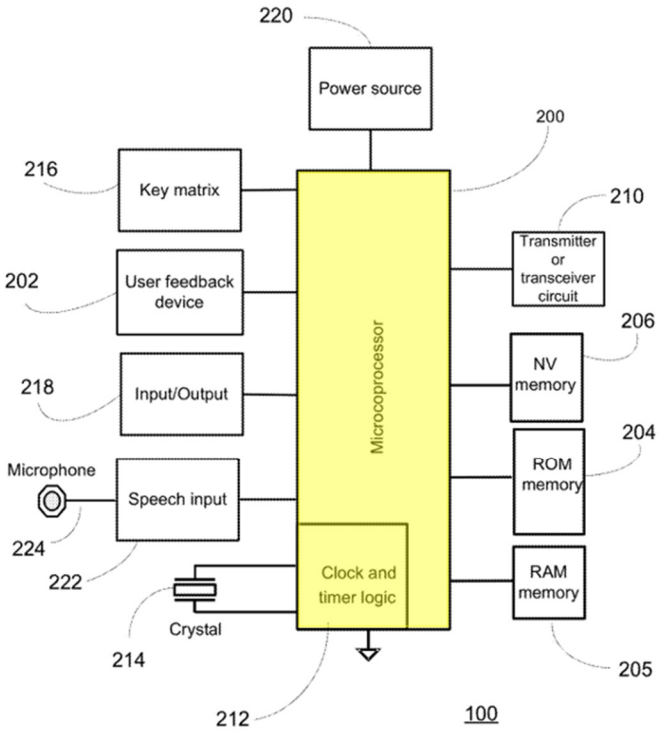


Figure 2

96. A POSITA would have understood that these statements apply to Lim's ***controlling device 110*** as well. As shown in FIG. 3 below, the smart device 110 includes "processing means 300 which may comprise both an application processing section 302 and an RF/DSP processing section 304; . . . memory means 312 which may comprise ROM, RAM, Flash, or any combination thereof[.]" (EX1005, [0023].) "It is known to offer an app for download and installation into memory 312 which app may serve to allow use of the smart phone to transmit commands suitable for operation of entertainment and other appliances." (EX1005, [0024].)

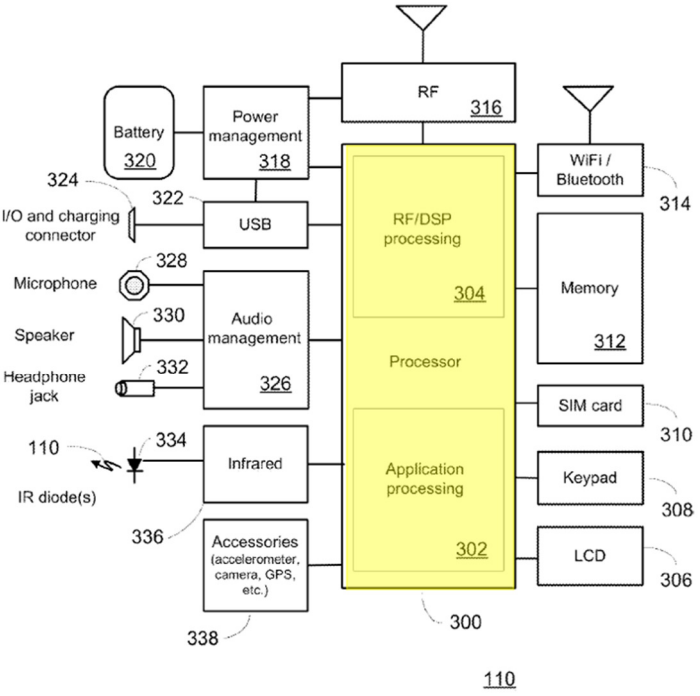


Figure 3

97. Execution of the computer program instructions in Lim's processors would cause the first/controlling device to perform the various operations disclosed in Lim. For example, "[t]o cause the controlling device 100 to perform an action, the controlling device 100 is adapted to be responsive to events, such as a sensed user interaction with the key matrix 216, ***sensed sound input***, etc. In response to an event, appropriate instructions within the program memory (hereafter the 'controlling device operating program') may be executed. . . . It will be appreciated that the controlling device operating program can be used not only to ***cause the transmission of commands and/or data to the appliances***, but also to perform local operations. While not limiting, local

operations that may be performed by the controlling device 100 may include ***displaying information/data***[.]" (EX1005, [0021]; *infra* §§ IX.A.2.f-IX.A.2.l.)

f. 1[E]: "capture, using the microphone, first and second human voice data;"

98. Lim discloses this element. As discussed, Lim discloses that the ***first/controlling device*** has a ***microphone***. (*Supra* § IX.A.2.c; EX1005, [0019], [0023].)

99. In Lim, the voice commands that trigger auditory responses by the first/controlling device are the recited "***first human voice data***"), and the voice commands that cause sending of a control message by the first/controlling device are the recited "***second human voice data***", as discussed below.

100. Lim discloses using a ***microphone*** in the ***first device*** to capture ***first human voice data***. Per the claim language of claim 1, "***first human voice data***" includes human voice data that causes the sounder to sound digital audio content in the recited manner. (*Infra* §§ IX.A.2.d, IX.A.2.i.) Thus, any of the human voice commands disclosed in Lim that causes a sounder (*e.g.*, ***speaker***) in the first device (*e.g.*, ***smart device 110***) to sound digital audio content in the recited manner discloses the recited "***first human voice data***."

101. Lim discloses many instances where the ***first device*** receives speech commands ("***first human voice data***") and provides a ***visual or audio response***

on the device itself. For example, Lim discloses a "setup and configuration" stage for the ***first device***. (EX1005, [0025], [0013].) "[D]uring setup and configuration of the controlling device, a user may provide appliance identification via ***spoken input***[.]" (EX1005, [0025].)

102. FIG. 4 discloses numerous examples of "***first human voice data***". Referring to FIG. 4 below, "[o]nce the setup mode has been initiated, at step 402 the controlling device may prompt the user to ***speak the brand name of the appliance*** to be added to the configuration of the controlling device." (EX1005, [0027].) Further, "[a]t step 404, if the speech recognition means is unable to uniquely identify the desired brand, step 402 may be repeated in order to solicit additional identifying information from the user. For example, a user may be presented with a short list of possible brand identities from which to select; may be prompted to ***spell the first few characters of the brand name***; may be requested to ***speak the name of a geographic region or a service provider***; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].) "Once a brand has been successfully identified, at step 406 the user may be prompted to ***provide spoken appliance type and/or model number information*** for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, *e.g.*, to further narrow down the codeset choices." (EX1005, [0028].) "If the ***initially***

provided model information [*e.g.*, provided previously in step 406] is insufficient to uniquely identify the appliance, at step 410, additional user input may be solicited. For example, a user may be requested to ***indicate a device type*** if this was not already provided; to ***provide geographic region or service provider***; to select from a series of displayed images of appliances and/or OEM remote controls; etc." (EX1005, [0029].) "Thereafter, at step 414 the user may be prompted to ***indicate if they wish to configure additional appliances***. If so, the process may be repeated starting at step 402 [where the controlling device prompts the user to speak the brand name of the appliance to be added to the configuration]." (EX1005, [0030].) It would have been obvious to a POSITA that the ***first/controlling device***'s microphone captures these voice inputs. Lim discloses that each such voice data is sent to the ***remote server 500*** for "speech recognition processing" and based on the processing, the ***first device*** provides a next prompt via "a displayed text or image", "a synthesized spoken message", etc. (EX1005, [0027]-[0030].) "[S]ubsequent prompts for voice input may take the form of . . . a ***synthesized spoken message*** (*e.g.*, where controlling device user feedback means 202 comprises speech synthesis means and a ***speaker***, or where a smart device app has access to ***speaker 330***), etc." (EX1005, [0027].) Thus, these voice commands cause the speaker of the

first/controlling device to provide an audio response. The user's voice inputs resulting in these prompts for voice input are all "***first human voice data***."

103. Lim discloses capturing, by the ***microphone*** in the ***first device***, a ***second human voice data***. To the extent that the recited "second human voice data" is different from the recited "first human voice data", Lim is clear that its system can receive and process ***multiple*** human voice commands. (EX1005, [0027]-[0029], [0035]-[0036], FIGs. 4 and 6.)

104. Further, Lim discloses human voice commands that can effectuate remote device control (*i.e.*, "***second human voice data***"). Per the claim language of claim 1, "***second human voice data***" includes human voice data that causes sending of a control message in the recited manner. (*Infra* § IX.A.2.j-IX.A.2.l.) Any of the human voice commands disclosed in Lim that causes sending of a control message to controllable appliances (*e.g.*, "***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers***") in the recited manner discloses the recited "***second human voice data***." (EX1005, [0017].)

105. FIG. 6 depicts an exemplary process in which the ***first device*** receives a speech command ("***second human voice data***") and transmits commands to ***appliances*** in block 612. For example, "at step 604 the user may

be prompted to ***speak the desired function name***," which "may be digitized and presented to a ***speech recognition means for processing***" "on a ***remote server 500***[.]" (EX1005, [0035]-[0036], [0027].) Based on the processing, the ***first/controlling device*** "transmit[s] the desired command to the target appliance" at block 614. (EX1005, [0036].) The user's voice input thus is a "***second human voice data***". It would have been obvious to a POSITA that the ***controlling device***'s microphone captures the voice input.

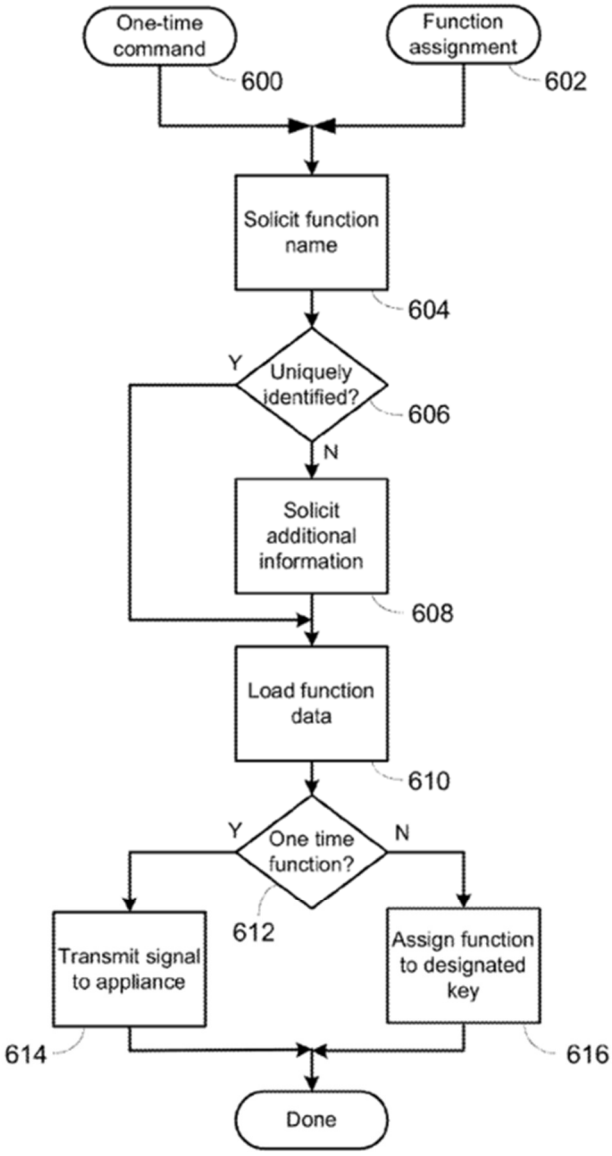


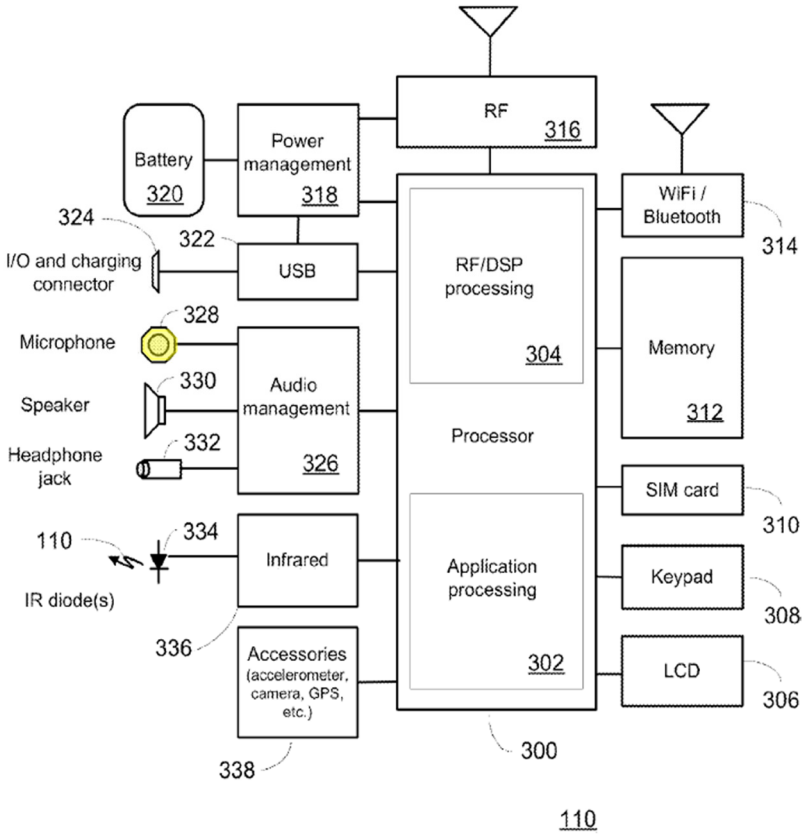
Figure 6

"Appliance command"
Lim, [0035].

106. Alternatively, to the extent that the first and second human voice data are captured/sent together (*e.g.*, as parts of one query) (which is not conceded), a POSITA would find it obvious that a user command to operate an appliance (*e.g.*, speaking the desired function name in FIG. 6) in Lim can

include multiple parts or phrases, such as a trigger phrase ("***first human voice data***") and the instruction ("***second human voice data***"), consistent with common implementations. (EX1023, 2:12-16, 34:63-67; EX1013, 8:20-32; EX1024, [0045], [0047].)

107. It would have been obvious to a POSITA that the first device's processor(s) are programmed with computer program instructions that, when executed, cause the first device to capture first and second human voice data using the microphone. As shown in Figure 3 (below), Lim's ***first device/controlling device***'s processor is connected to the ***microphone 328*** for operating the microphone to capture first and second human voice data. (*Supra* § IX.A.2.e; EX1005, [0021].)



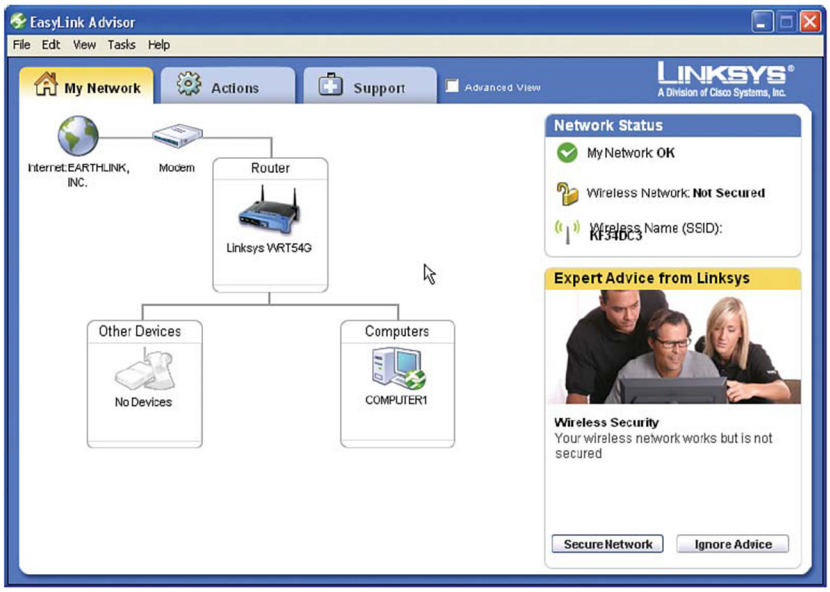
g. 1[F]: "send, over the Internet via the WLAN, to a server device, the captured first and second human voice data;"

108. Lim discloses this element. First, Lim discloses a ***server device***. Lim discloses that "all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027].) Lim states, "[a]s will become apparent, speech recognition and command data identification processing may be performed . . . ***remotely*** at a ***Web server***[.]" (EX1005, [0007].)

109. Further, Lim's ***first/controlling device*** communicates to the ***remote server 500*** over the Internet via the WLAN. Lim discloses that "access to such network [*i.e.*, 'a network 502 such as the Internet or PSTN'] may be . . . via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].) In other words, Lim's ***first device*** is connected to a WiFi network (*i.e.*, WLAN) via a router, which in turn accesses the ***remote server 500*** over the Internet. Indeed, this was the common configuration for a home network. As discussed in a textbook titled "Wireless Home Networking Simplified," "with a wireless LAN, the computers have to find a way to connect, or associate, themselves to the LAN, because they are not physically connected. The way they do this is by ***tuning themselves to the wireless router*** using the name of the router . . . . Once a computer has connected, or associated, itself to the access point, it becomes part of the LAN. . . . ***Now the computer can send and receive information*** to and from other computers on the LAN or ***to and from the Internet by transmitting over wireless to the wireless router (or access point)***. The wireless router acts as the central communication routing point . . . ." (EX1007, 11, 100.)

110. "Figure 7-28 shows what the initial ***network*** should look like if things are working properly, including ***connectivity to the Internet***." (EX1007, 96, Figure 7-28.)

Figure 7-28 EasyLink Graphical View of Your Network



111. Thus, Lim discloses a ***first device*** communicating with a ***server device*** over the ***Internet via the WLAN***.

112. Second, Lim discloses that the ***first/controlling device*** sends the captured human voice data (including the ***first human voice data*** and ***second human voice data***) to the ***server*** over the ***Internet via the WLAN***.

113. Regarding the ***first human voice data***, Lim discloses that "at step 402 the ***controlling device*** may prompt the user to speak the ***brand name of the appliance*** to be added to the configuration of the controlling device. . . . The resulting ***spoken syllables*** may be digitized and presented to a speech recognition means for processing and comparison to a phonetic database of brand names which have been cross referenced to appropriate codesets within the library of codesets. As appropriate for a particular embodiment, all or part of

said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the Internet or PSTN. In the latter case, access to such network . . . may be via a local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router (not illustrated)." (EX1005, [0027].) In other words, the "***first human voice data***" is sent by the ***controlling device*** over the Internet via the WLAN to a ***server device***.

114. It would have been obvious to a POSITA that other instances of "***first human voice data***" are similarly sent to the server for processing. Some of these are shown in FIG. 4. "At step 404, if the speech recognition means is unable to uniquely identify the desired brand, step 402 may be repeated in order to solicit additional identifying information from the user. For example, a user may be presented with a short list of ***possible brand identities*** from which to select; may be prompted to ***spell the first few characters of the brand name***; may be requested to speak the ***name of a geographic region or a service provider***; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].) Lim discloses that "[o]nce a brand has been successfully identified, at step 406 the user may be prompted to provide ***spoken appliance type and/or model number information*** for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if

necessary, *e.g.*, to further narrow down the codeset choices. . . . At step 408 the ***spoken syllables comprising the model identity*** may be processed and compared to a phonetic database of model numbers by the speech recognition means which have also been cross referenced to codesets within a library of codesets. If the initially provided model information is insufficient to uniquely identify the appliance, at step 410 additional user input may be solicited. For example, a user may be requested to ***indicate a device type*** if this was not already provided; to ***provide geographic region or service provider***; to select from a series of displayed images of appliances and/or OEM remote controls; etc." (EX1005, [0029].)

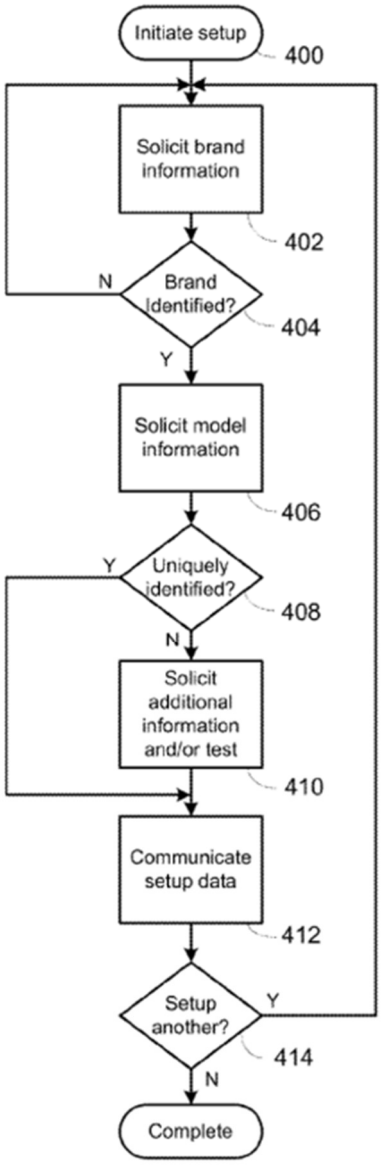


Figure 4 "Set up and configuration of controlling device" Lim, [0025].

115. The ***second human voice data*** is similarly sent. Regarding the ***second human voice data***, Lim discloses an exemplary process of using a configured ***controlling device*** to control home appliances using voice inputs. Referring to FIG. 6,"at step 604 the user may be prompted to speak the ***desired***

function name. As before this prompt may take the form of an illuminated LED and/or "speak" symbol, a displayed text or image, etc., as appropriate. The ***resulting spoken syllables*** may be digitized and presented to a speech recognition means for processing and comparison to a phonetic database of command function names which have been cross referenced to command data within a codeset that is applicable to the particular appliance model currently configured . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server*** accessible via a network such as the Internet or PSTN." (EX1005, [0035], [0034].) In other words, the "***second human voice data***" is sent by the ***controlling device*** over the Internet via the WLAN to a ***server device***.

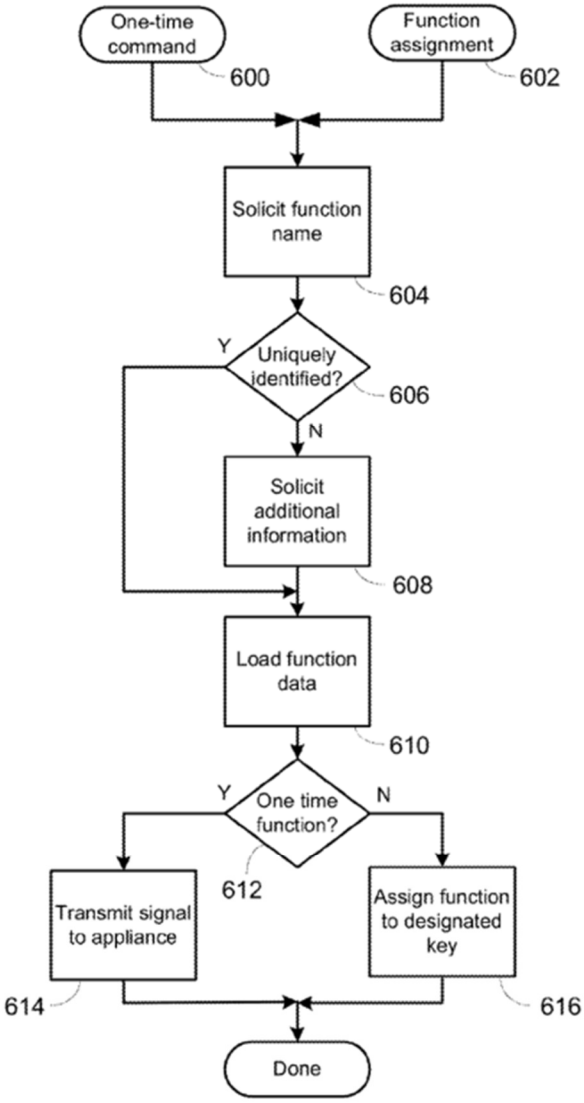


Figure 6

"Appliance command" Lim, [0035].

116. To the extent that the first and second human voice data are sent together (*e.g.*, as parts of one query) (which is not conceded), a POSITA would find it obvious that a user command to operate an appliance (*e.g.*, speaking the desired function name in FIG. 6) in Lim can include multiple parts, such as a

trigger phrase and the instruction, which would be sent together to the server. (EX1023, 2:12-16, 34:63-67; EX1013, 8:20-32; EX1024, [0045], [0047].)

117. It would have been obvious to a POSITA that the first device's processor(s) are programmed with computer program instructions that, when executed, cause the first device to send the captured first and second human voice data over the Internet via the WLAN to a server device. As shown in Figure 3 (below), Lim's ***first device/controlling device***'s processor is connected to the WiFi and/or Bluetooth wireless interface(s) 314 for operating the wireless interface to send the captured first and second human voice data to the server device over the Internet via the WLAN. (EX1005, [0023]; *supra* § IX.A.2.e; EX1005, [0021].)

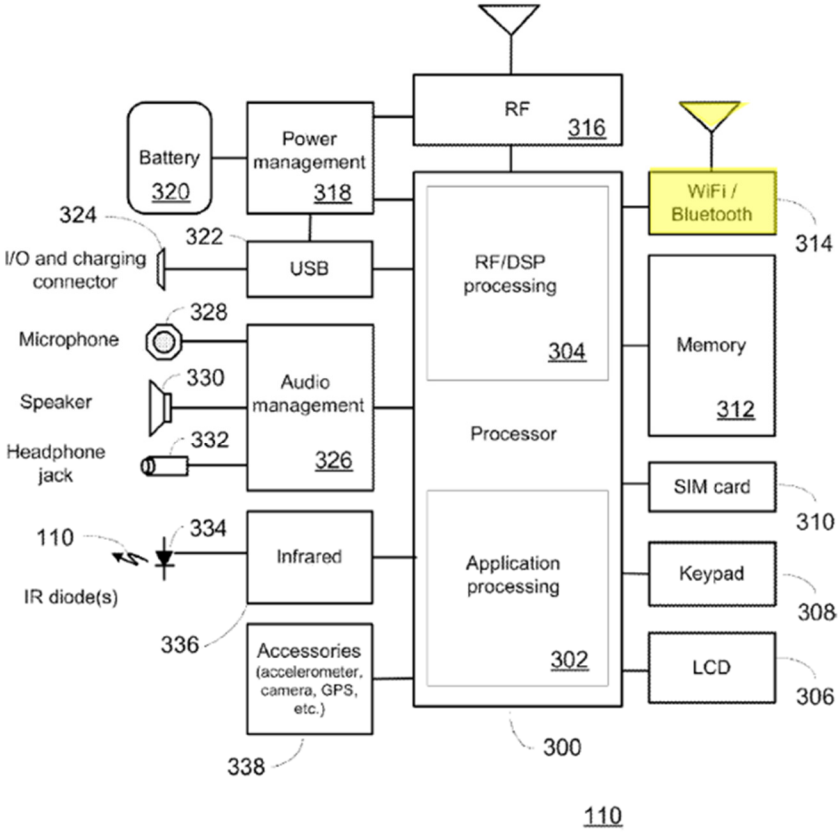


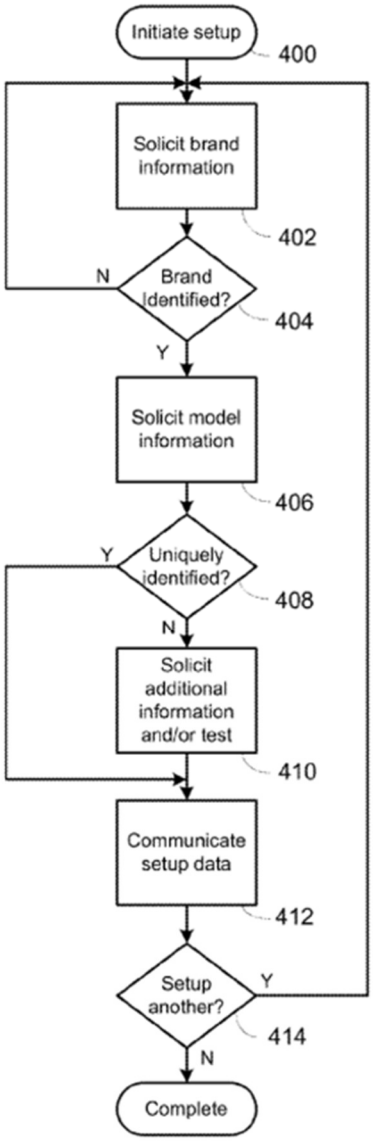
Figure 3

h. 1[G]: "receive, from the server device via the WLAN over the Internet, first and second messages, in response to the sending of the respectively captured first and second human voice data;"

118. Lim discloses or at least suggests the ***first/controlling device*** receives, from the ***server*** device via the WLAN over the Internet, ***first message*** and ***second message***, in response to the sending of the respectively captured ***first human voice data*** and ***second human voice data***.

119. First, Lim discloses or at least suggests that the ***first device*** receives a ***first message*** from the ***server*** in response to the sending of the captured ***first human voice data***. As discussed, Lim discloses the ***first/controlling device***

sends the captured ***first human voice data*** to the ***server*** for speech processing. (*Supra* § IX.A.2.g.) For example, in the setup and configuration stage in FIG. 4, the ***controlling device*** captures and sends ***first human voice data*** to the ***remote server*** for speech recognition processing. (*Supra* § IX.A.2.g; EX1005, [0027].)



"Set up and configuration of controlling device" Lim, [0025].

Figure 4

120. It would have been obvious to a POSITA that, after sending ***first human voice data*** to the ***remote server*** for speech recognition processing, the ***controlling device*** would receive some ***speech processing result(s)*** associated with the ***first human voice data*** from the ***remote server*** in order to provide a responsive next prompt. Indeed, Lim discloses providing responsive prompts (via a display or a speaker) following the user's voice inputs, and it would have been obvious that these responsive prompts are or derived from ***speech processing result(s)*** (*i.e.*, first message) received from the ***remote server***. A POSITA would have understood that the ***speech processing result(s)*** may be the recognized speech (*e.g.*, text strings), a result of processing the recognized speech (*e.g.*, semantic processing to determine meaning and adequacy of the speech), and/or information associated with the next prompt (*e.g.*, text or audio, information to display, information triggering a visual or auditory output such as flashing a light indicator, providing a buzz, or playing a sound). (*E.g.*, EX1009; EX1036, 307 ("The client side has to transmit the speech signal to the server. The server returns the recognition result for the client after processing."); EX1037, 2367 ("the server computer recognizes the speech and displays the recognized text back to the user").)

121. For example, "at step 402 the ***controlling device*** may prompt the user to ***speak the brand name of the appliance*** to be added to the configuration

of the controlling device. . . . The ***resulting spoken syllables*** may be digitized and presented to a speech recognition means for processing . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the Internet or PSTN." (EX1005, [0027].) A POSITA would have found it obvious that ***speech processing result(s)*** (*i.e.*, first message) is sent back to the ***first/controlling device*** such that "[a]t step 404, ***if the speech recognition means is unable to uniquely identify the desired brand*** [based on the ***speech processing result(s)***], step 402 may be repeated in order to solicit additional identifying information from the user. For example, a user may be presented with a short list of possible brand identities from which to select; may be prompted to spell the first few characters of the brand name; may be requested to speak the name of a geographic region or a service provider; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].) In this example, the ***speech processing result(s)*** can be the recognized speech (*e.g.*, text strings), a result of processing the recognized speech (*e.g.*, determining "***if the speech recognition means is unable to uniquely identify the desired brand***"), information associated with the responsive next prompt (*e.g.*, text or audio), and/or information to display, information triggering a visual or auditory output such as flashing a light indicator, providing a buzz, or playing a sound,

depending on the distribution of speech processing between the server and the controlling device.

122. In Lim, there are other instances of ***first messages*** received by the ***first/controlling device*** for it to determine the next prompt. For example, "at step 406 the user may be prompted to provide spoken appliance type and/or model number information for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, *e.g.*, to further narrow down the codeset choices. . . . At step 408 the ***spoken syllables comprising the model identity*** may be processed and compared to a phonetic database of model numbers by the speech recognition means which have also been cross referenced to codesets within a library of codesets. ***If the initially provided model information is insufficient to uniquely identify the appliance*** [based on the ***speech processing result(s)***], at step 410 additional user input may be solicited." (EX1005, [0029].) In this example, the ***speech processing result(s)*** can be the recognized speech (*e.g.*, text strings), a result of processing the recognized speech (*e.g.*, determining "[i]f the initially provided model information is insufficient to uniquely identify the appliance"), and/or information associated with the responsive next prompt (*e.g.*, text or audio, information to display, information triggering a visual or auditory output such as flashing a light indicator, providing a buzz, or playing a sound), depending on

the distribution of speech processing between the server and the controlling device.

123. "Thereafter, at step 414 the user may be prompted to ***indicate if they wish to configure additional appliances***. If so [based on the ***speech processing result(s)***], the process may be repeated starting at step 402." (EX1005, [0030].) In this example, the ***speech processing result(s)*** can be the recognized speech (*e.g.*, text strings), a result of processing the recognized speech (*e.g.*, determining "if the user indicates that they wish to configure additional appliances"), and/or information associated with the responsive next prompt (*e.g.*, text or audio, information to display, information triggering a visual or auditory output such as flashing a light indicator, providing a buzz, or playing a sound), depending on the distribution of speech processing between the server and the controlling device.

124. At the time of the invention, remote speech recognition servers routinely transmit results of speech recognition back to the requesting device. (*E.g.*, EX1009; EX1036, 307; EX1037, 2367.) Further, in light of Lim's disclosure of responsive prompts following the user's voice inputs, it would have been obvious to a POSITA that ***speech processing result(s)*** (*e.g.*, recognized speech, analysis of or information associated with the recognized speech, the next prompt, information to display, information triggering a visual

or auditory output such as flashing a light indicator, providing a buzz, or playing a sound) is sent from the ***remote server 500*** back to the ***first device*** for speech synthesis or display. (*Infra* § IX.A.2.i.) For example, "the application 120 [may] provide audio 222 for eventual transmission to the interface device 116, to be played on the speaker 206. The audio may be in the form of text, in which speech synthesis or text-to-speech conversion may be performed . . . by the operating logic 108 of the interface device 116." (EX1006, [0037].) "The client side has to transmit the speech signal to the server. The server returns the recognition result for the client after processing." (EX1036, 307.) "[T]he server computer recognizes the speech and displays the recognized text back to the user." (EX1037, 2367.) Further, it would have been obvious that such ***first message*** is received by the ***first device*** from the ***server*** "in response to the sending of the captured ***first human voice data***" because the recognized speech, analysis of or information associated with the recognized speech, the next prompt (*e.g.*, text or audio or information to display), and/or information triggering a visual or auditory output such as flashing a light indicator, providing a buzz, or playing a sound is generated based on the ***server***'s processing of ***first human voice data***.

125. Lim discloses or at least suggests its ***first/controlling device*** receives a ***second message*** from the ***server*** in response to the sending of the

captured ***second human voice data***. As discussed, Lim discloses the first device sending the captured ***second human voice data*** to the server. (*Supra* § IX.A.2.g.) It would have been obvious to a POSITA that, after sending ***second human voice data*** to the ***remote server*** for speech recognition processing, the ***controlling device*** would receive ***speech processing result(s)*** from the ***remote server***. The ***speech processing result(s)*** may be the recognized speech (*e.g.*, semantic processing to determine meaning and adequacy of the speech). (*E.g.*, EX1009; EX1036, 307; EX1037, 2367.)

126. For example, "at step 604 the user may be prompted to ***speak the desired function name*** . . . [t]he resulting spoken syllables may be digitized and presented to a speech processing means[.]" (EX1005, [0035].) A POSITA would have found it obvious that speech processing result(s) (*i.e.*, second message) is sent back to the first/controlling device such that "***[o]nce the desired function has been uniquely identified*** [based on the ***speech processing result(s)***], at step 610 the indicated command code data may be provisioned to the controlling device. As appropriate for the particular implementation, such provisioning may comprise downloading the data itself, supply of a pointer to an entry in a local database, etc." (EX1005, [0035]-[0036], [0030].) Thus, the second message is the ***speech processing result(s)*** from the server that allows the controlling device to determine whether the desired function has been

identified. In this example, the second message can be the recognized speech (*e.g.*, text strings), a result of processing the recognized speech (*e.g.*, determining whether the "desired function has been uniquely identified", or the identified function), etc., depending on the distribution of speech processing between the server and the controlling device. It would have been obvious that the ***second message*** is sent from the ***server*** to the first device in response to the sending of the captured ***second human voice data*** because the ***speech processing result(s)*** is transmitted based on the server's processing of the second human voice data.

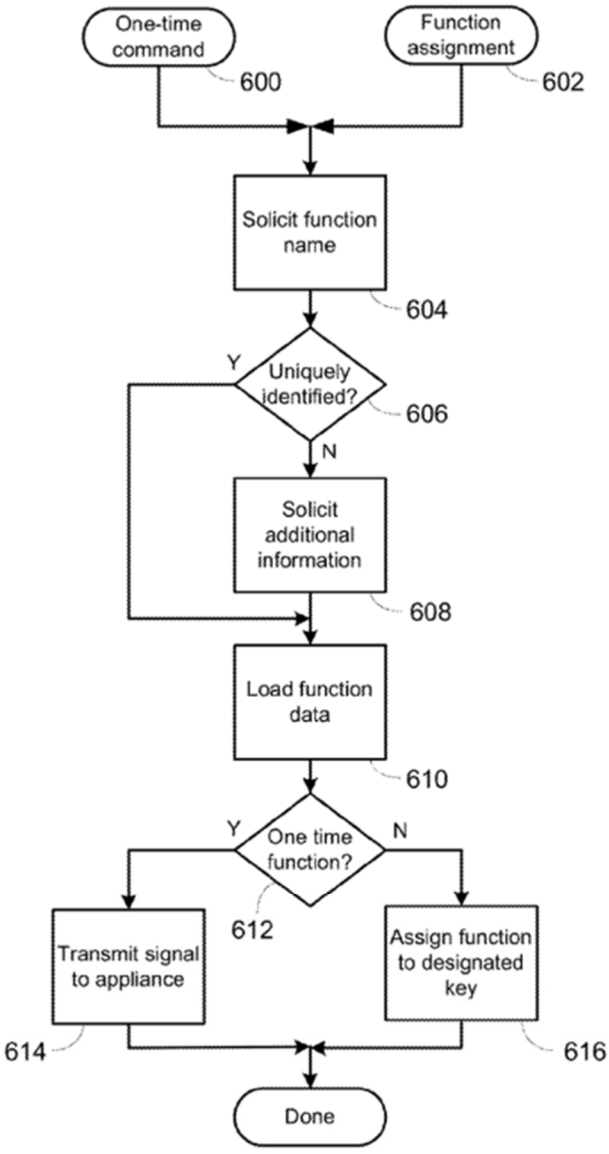


Figure 6

"Appliance command" Lim, [0035].

127. A second example of the second message is the "*command code data*" downloaded from the ***server 500*** in block 610 in response to the sending of the captured ***second human voice data***. Lim discloses that the command code data can be downloaded to the ***controlling device*** in block 610. "Once the

desired function has been uniquely identified [at the ***remote server device***], at step 610 the ***indicated command code data*** may be provisioned to the ***controlling device***. As appropriate for the particular implementation, such provisioning may comprise downloading the ***data*** itself, supply of a ***pointer*** to an entry in a local database, etc." (EX1005, [0036].) Lim states that the indicated command code data is downloaded from a library of codesets located externally at the same ***remote server 500*** accessible to the ***first/controlling device***. Lim discloses that "communication of the setup data may comprise downloading of a complete set of command code data into the controlling device from an off-platform database of codesets located for example . . . on a ***remote server 500***[.]" (EX1005, [0030].) "[A]n app may draw on a library of codesets suitable for commanding the functional operations of various types of appliances" and the library of codesets "may be located externally at a local PC or ***remote server device*** accessible to ***smart device 110*** for the download of only those code sets necessary for operation of the configured appliances[.]" (EX1005, [0024].) In other words, the second message is the downloaded command data/codesets from the ***server 500***. It would have been obvious to a POSITA that the ***second message*** is sent from the ***server*** to the ***first/controlling device*** in response to the sending of the captured ***second human voice data*** because the "***command code data***"/***codesets*** are identified and sent based on the

server's processing of the ***user command*** and are "necessary for operation of the configured appliances". (EX1005, [0030], [0024].)

128. Lim explains that command data can be or include "data which serves to define the necessary control protocols and command values for use in transmitting command signals to controllable appliances (collectively, the command data)." (EX1005, [0020].) Lim defines sets of command data or codesets to be "library comprising a multiplicity of command codes and protocols suitable for controlling various appliances - also referred to as sets of command data or codesets." (EX1005, [0022].)

129. Lim discloses that the first device receives the recited "***first message***" and "***second message***" from the server in response to the sending of the respectively captured ***first human voice data*** and ***second human voice data*** over the ***Internet via the WLAN***. As discussed above, Lim discloses a ***first/controlling device*** communicating with a ***server device*** ***via the WLAN*** over the ***Internet***. (*Supra* §§ IX.A.2.a, IX.A.2.g.) It thus follows that Lim's ***controlling device*** receives the recited "***first message***" and "***second message***" from the server in response to the sending of the respectively captured ***first human voice data*** and ***second human voice data*** over the ***Internet via the WLAN***.

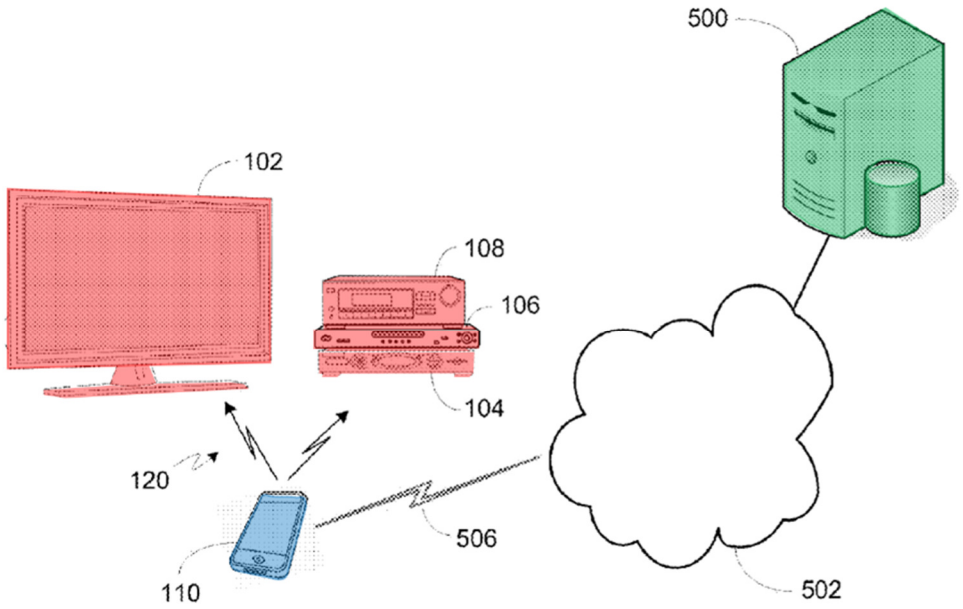


Figure 5b

130. Alternatively, as discussed, a POSITA would find it obvious that a user command to operate an appliance (*e.g.*, speaking the desired function name in FIG. 6) in Lim can include multiple parts, such as a trigger phrase ("***first human voice data***") and the instruction ("***second human voice data***"), consistent with common implementations. (EX1023, 2:12-16, 34:63-67; EX1013, 8:20-32; EX1024, [0045], [0047].) It follows that the processing of the first human voice data would trigger sending of ***corresponding speech processing results*** ("***first message***") to the controlling device to provide audio feedback as it does in FIG. 4. Further, the processing of the second human voice data would trigger sending of ***corresponding speech processing results*** of

the instruction and ***command data/codesets*** ("second message") as taught in FIG. 6. (*Infra* §§ IX.A.2.i-IX.A.2.l.)

131. It would have been obvious to a POSITA that the first device's processor(s) are programmed with computer program instructions that, when executed, cause the first device to receive first and second messages from the server device via the WLAN over the Internet in response to the sending of the respectively captured first and second human voice data. (*Supra* § IX.A.2.e; EX1005, [0021].)

i. 1[H]: "sound, using the sounder, digital audio content in response to the received first message;"

132. Lim discloses this element. As discussed, Lim discloses a ***speaker*** and a ***buzzer***, both of which are sounders, in the ***first/controlling device***. Lim discloses "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . a means 202 to provide feedback to the user (*e.g.*, one or more LEDs, illuminable keys, display, ***buzzer, voice synthesizer with associated speaker***, and/or the like)[.]" (*Supra* § IX.A.2.d; EX1005, [0019].) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated microphone 328, ***speaker 330***[.]" (EX1005,

[0023].) It would have been obvious that ***smart device 110*** is similarly equipped with a ***buzzer***.

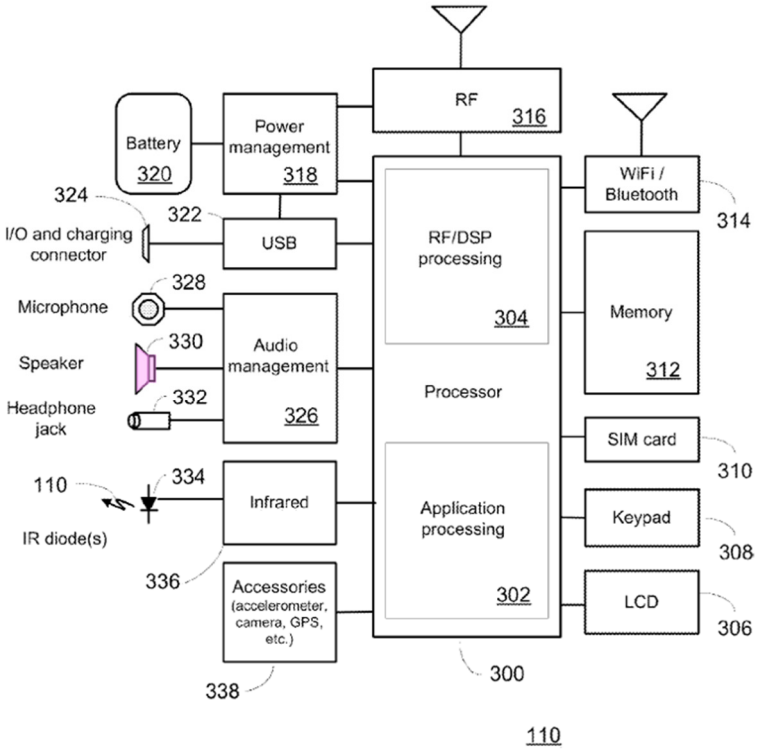
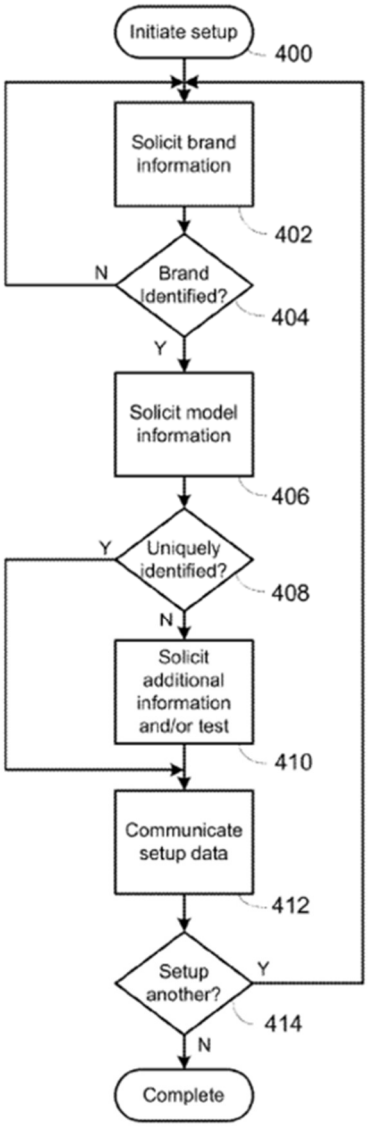


Figure 3

133. Lim discloses sounding digital audio content using a ***sounder*** in the ***first/controlling device*** in response to the received ***first message***. Specifically, Lim provides auditory prompts via a ***speaker*** in the ***controlling device***. The prompt is provided in response to a user voice input and, more specifically, in response to ***speech processing result(s)*** (*e.g.*, recognized speech, analysis of or information associated with the recognized speech, next prompts, and/or information triggering an auditory output such as playing a sound) received from

the speech recognition ***server 500***. For example, "at step 402 the ***controlling device*** may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. This and subsequent prompts for voice input may take the form of an illuminated LED and/or 'speak' symbol, a displayed text or image, a ***synthesized spoken message*** (*e.g.*, where controlling device ***user feedback means 202*** comprises speech synthesis means and ***a speaker***, or where a smart device app has access to ***speaker 330***), etc." (EX1005, [0027].) Lim suggests, and a POSITA would find it obvious, that the operation of the sounders (*i.e.*, prompts for voice input) are based on the ***speech processing result(s)***. For example, Lim discloses that after step 402, "[a]t step 404, ***if the speech recognition means is unable to uniquely identify the desired brand*** [based on ***speech processing result(s)***], step 402 may be repeated in order to ***solicit*** additional identifying information from the user. For example, a user may be ***presented*** with a short list of possible brand identities from which to select; may be ***prompted*** to spell the first few characters of the brand name; may be ***requested*** to speak the name of a geographic region or a service provider; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].) In addition, "***[o]nce a brand has been successfully identified*** [based on ***speech processing result(s)***], at step 406 the user may be ***prompted*** to provide spoken appliance type and/or model number information for processing and comparison

to a phonetic database of model number data which is applicable to the identified brand if necessary, e.g., to further narrow down the codeset choices." (EX1005, [0029].) A POSITA would have understood that these voice prompts are digital audio data sounded by the sounder (speaker) in the ***controlling device***.



"Set up and configuration of controlling device" Lim, [0025].

Figure 4

134. Additional prompts are provided in blocks 410 and 414. "***If the initially provided model information is insufficient to uniquely identify the appliance*** [based on ***speech processing result(s)***], at step 410 additional user input may be solicited. For example, a user may be ***requested*** to indicate a device type if this was not already provided; to provide geographic region or service provider; to select from a series of displayed images of appliances and/or OEM remote controls; etc." (EX1005, [0029].) And "at step 414 the user may be ***prompted*** to indicate if they wish to configure additional appliances. If so, the process may be repeated starting at step 402." (EX1005, [0030].) As another example, "a user may be ***requested*** to indicate a geographic region or the approximate age of an appliance; may be ***presented*** with a display of short list of function names from which to select; may be ***request*** to provide synonym(s) for the desired command function: etc." (EX1005, [0036].)

135. Additionally or alternatively, Lim's ***buzzer*** meets this element. Lim's ***first/controlling device*** "provide[s] feedback to the user (e.g., one or more LEDs, illuminable keys, display, ***buzzer***, voice synthesizer with associated speaker, and/or the like)." (EX1005, [0019].) Thus, Lim certainly suggests that providing the prompts discussed above would include sounding a buzzer, for example, to get the user's attention to start listening to the synthesized questions.

136. It would have been obvious that Lim sounds digital audio data via its ***speaker*** and/or ***buzzer*** in response to ***speech processing result(s)*** (*e.g.*, recognized speech, analysis of or information associated with the recognized speech, the next prompt, and/or information triggering an auditory output such as providing a buzz, or playing a sound) received from the server (*i.e.*, ***remote server 500***) because the ***speech processing result(s)*** is needed to determine if, when, and/or which next prompt should be sounded.

137. Further, it would have been obvious to a POSITA that the first device's processor(s) are programmed with computer program instructions that, when executed, cause the first device to sounds digital audio data via a ***sounder*** in response to the received ***first message***. (*Supra* § IX.A.2.e; EX1005, [0021].)

j. 1[I][1]: "and send, to a second device over the WLAN, a control message that comprises a control data for controlling a first actuator in the second device"

138. Lim discloses this element. First, Lim discloses sending to a ***second device*** over the ***WLAN***, a ***control message*** that comprises a control data. Lim's "***controlling device 100*** is configured to ***control various controllable appliances such as a television 102, a set top box ('STB') 104, a DVD player 106 and an AV receiver 108***." (EX1005, [0017].) Each of these controllable appliances is a ***second device***. Lim discloses that "[w]hile illustrated in the context of a television 102, STB 104, DVD player 106 and AV receiver 108, it is

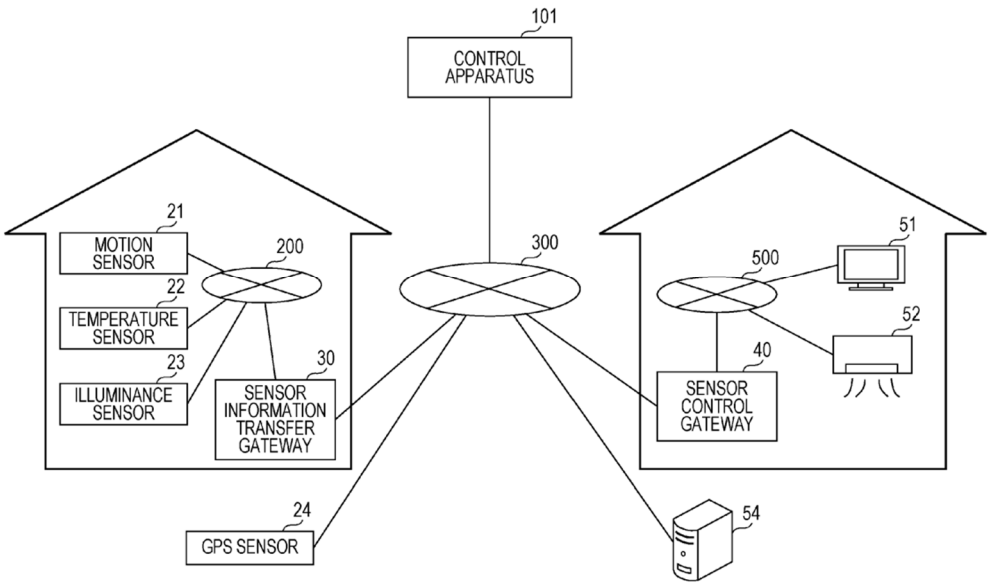
to be understood that controllable appliances may include, but need not be limited to, ***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers***, etc." (EX1005, [0017].)

139. Lim's ***second devices*** have actuators ("first actuator"). One example of the recited "first actuator" as disclosed in Lim's controllable appliances is an actuator providing visual information (*e.g.*, displays). Lim's controllable appliances such as "***a television 102, . . . DVD player 106***" have displays, which are actuators. (EX1005, [0017].) "Liquid crystal displays (LCDs) make use of a quite different principle. They dependent on a liquid crystal material that polarizes light when an electrical potential is applied . . . This is the principle now widely used for all types of displays, including computer monitors and television displays." (EX1018, 175.) The '721 patent itself explains that a display is an actuator. "An actuator may be a display for presentation of visual data or information, commonly on a screen, and may consist of an array (e.g., matrix) of light emitters or light reflectors, and may present text, graphics, image or video. A display may be a monochrome, gray-scale, or color type, and may be a video display." (EX1001, 25:33-38.) "The actuator may be a monochrome, grayscale or color display for visually

presenting information, and may consist of an array of light emitters or light reflectors. . . . The display may be a Liquid Crystal Display (LCD) display, a Thin Film Transistor (TFT), or an LED-backlit LCD display, and may be based on a passive or an active matrix." (EX1001, 40:50-41:6; *see also id*., 109:42-47, 110:61-67.)

140. This is consistent with prior findings of the Board in its Decision on Appeal affirming the Examiner's rejections of claims of the parent patent '710 that an image display device such as a TV is a controlled device (*i.e.*, second device) with actuators. "Nor does Appellant persuasively rebut the Examiner's mapping of Abe's control apparatus 101 to a control server (*see* Ans. 6)—a mapping that we find reasonable on this record given the control apparatus's external control capabilities for in-building networked sensor-based systems via Internet 300 in Figure 15. *See* Abe ¶¶ 93-96. As the Examiner indicates, Abe's in-building systems in Figure 15 include ***image display device 51*** and air conditioner 52 ***that are controlled devices with actuators***. *See* Ans. 6; Abe ¶ 95." (EX1004, 396.)

FIG. 15



141. Another example of the recited "first actuator" is an actuator providing sound (*e.g.*, speakers in the various media devices in Lim). As discussed above, a speaker is a sounder. (*Supra* § IX.A.2.d.) The specification of the '721 patent states that a sounder has or is an actuator. "An actuator may be an ***electroacoustic actuator, such as a sounder*** which converts electrical energy to sound waves transmitted through the air, an elastic solid material, or a liquid, usually by means of a vibrating or moving ribbon or diaphragm." (EX1001, 105:48-52.) A POSITA would have understood that Lim's media devices, such as ***a television 102, a set top box ('STB') 104, a DVD player 106 and an AV receiver 108***, comprise speakers to produce and control sound (*e.g.*, to playback a media item).

142. Another example of the "first actuator" is a motion actuator. Lim discloses that "controllable appliances may include, but need not be limited to, ***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers***, etc." (EX1005, [0017].) A POSITA would have understood that devices such as "***home lighting, drapery, fans, HVAC systems, thermostats***" have motion actuators that can be operated. Indeed, at the time of the invention, it was known that heating and cooling systems are integrated with motors and fans. "A unitary air conditioner is an apparatus consisting of one or more factory made assemblies that normally include an evaporator, a blower, a compressor, and a condenser, designed to be used together. The unit often performs heating as well as cooling functions." (EX1020, 401.) "Horizontal single package units (Fig. 13-3) are designed for both indoor and outdoor installation, with the majority of units incorporating air cooled or evaporative-cooled condensers. The most common design has supply and return air connections located in the same face with horizontal air flow." (EX1020, 401.)

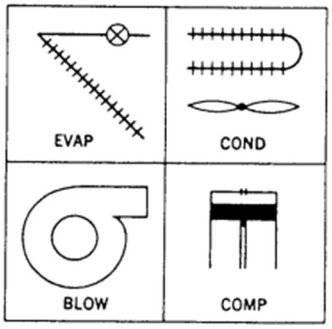


Fig 13-1. Self contained air conditioner

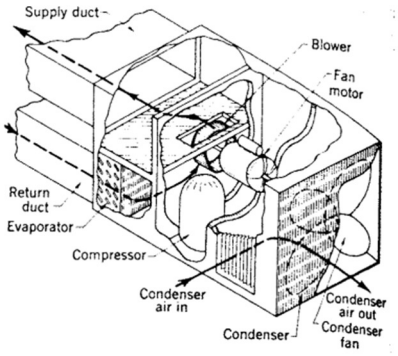


Fig 13-3. Horizontal self contained air conditioned unit

143. Further, as discussed, the Board agreed in its Decision on Appeal affirming the Examiner's rejections of claims 1-30 of the parent patent '710 that an air conditioner is a controlled device (*i.e.*, second device) with actuators. "As the Examiner indicates, Abe's in-building systems in Figure 15 include image display device 51 and ***air conditioner 52 that are controlled devices with actuators***. *See* Ans. 6; Abe ¶ 95." (EX1004, 396.)

144. In addition, a POSITA would have understood that Lim's VCRs contain actuators that spin the recording/reading head to read the tracks on the tape of the recording or playback. "To achieve the required speed of tape past the head, two or more heads are mounted on a spinning drum which is mounted to scan the head across the tape at a shallow angle. This allows the tape to move relatively slowly past the spinning drum. The helical scan system has been under development since the late 1950s, with the VHS system introduced in the early 1980s and largely unchanged today." (EX1018, 226.) And as shown in Figure 19.6 below, "[t]he analogue audio signal is recorded on a narrow strip

along one edge of the tape, while field sync. pulses are recorded on the opposite edge. The ***heads and tape travel*** in the same direction to minimize frictional losses and to extend the mechanical lifetime." (EX1018, 226-227.)

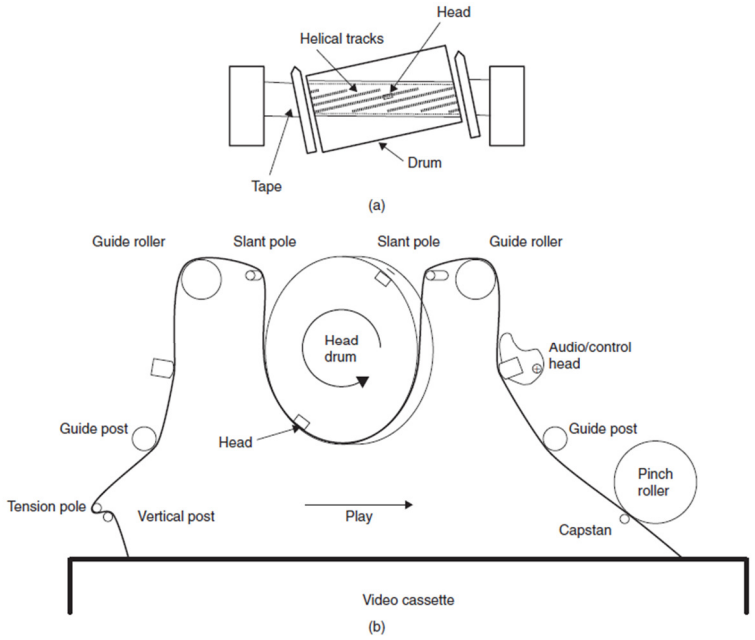


Figure 19.6 Tape transport of typical helical scan video recorder: (a) tracks on tape, and (b) tape path

145. As another example, CD and DVD drives use actuators for different parts of the spinning and reading process. The CD drive works in the following manner: "[t]he control signals set the speed of the drive motor, which is not constant; it runs faster when the inner tracks are being read and slower for the outer tracks, keeping the data speed constant. The beam position drives select the track and maintain focus, and the optical head reads the signals reflected from the track. These signals are then converted into conventional digital data in the interface." (EX1018, 279-280, Figure 23.7.)

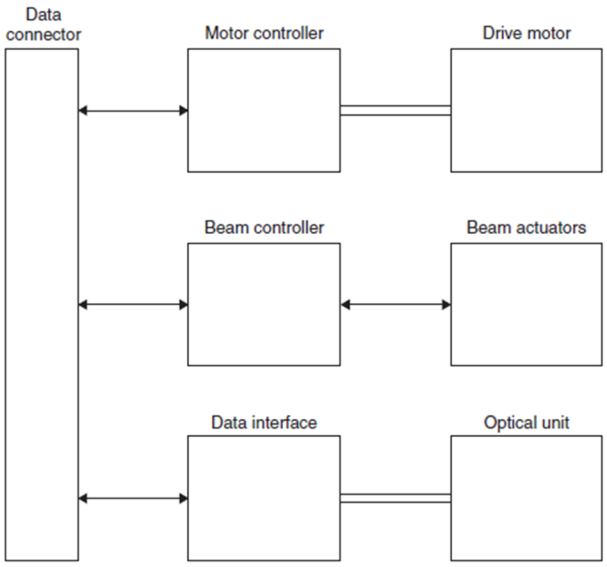


Figure 23.7 Block diagram, simplified, for a CD-ROM drive

146. Lim's ***first/controlling device*** sends a ***control message*** that comprises a control data for controlling a first actuator in the ***second device***. "As is known in the art, the ***controlling device 100*** may be capable of transmitting ***commands*** to the appliances, using any convenient IR, RF, Point-to-Point, or ***networked protocol 120***, to cause the appliances to perform operational functions." (EX1005, [0017].) A POSITA would have understood that this statement applies to Lim's ***controlling device 110*** as well. Similarly, an "app may serve to allow use of the ***smart phone*** to transmit ***commands*** suitable for operation of ***entertainment and other appliances***." (EX1005, [0024].) As shown in FIGs. 1b and 5b below, the ***controlling device 110*** transmits commands to the ***appliances*** via ***networked protocol 120***. It would have been obvious that Lim's ***commands*** are messages. As discussed,

transmission via "networked protocol 120" can include WiFi. (*Supra* §§ IX.A.2.a, IX.A.2.b.) Consistently, Lim's controlling device 110 has WiFi transceiver. (*Supra* §§ IX.A.2.a, IX.A.2.b.) A POSITA would have found it obvious that data transmitted over WiFi—such as Lim's commands—includes packets of information (*e.g.*, characters or bits) structured or arranged according to the WiFi protocol such that the receiving device can parse, validate, and interpret the content. Thus, Lim's commands are messages. This is also consistent with the '721 patent. (EX1001, 148:37-42 ("The term 'message' is used generically herein to describe at least an ordered series of characters or bits intended to convey a package of information (or a portion thereof), which may be transferred from one point to another, such as by using communication via one or more communication mechanisms or by transferring among processes.").)

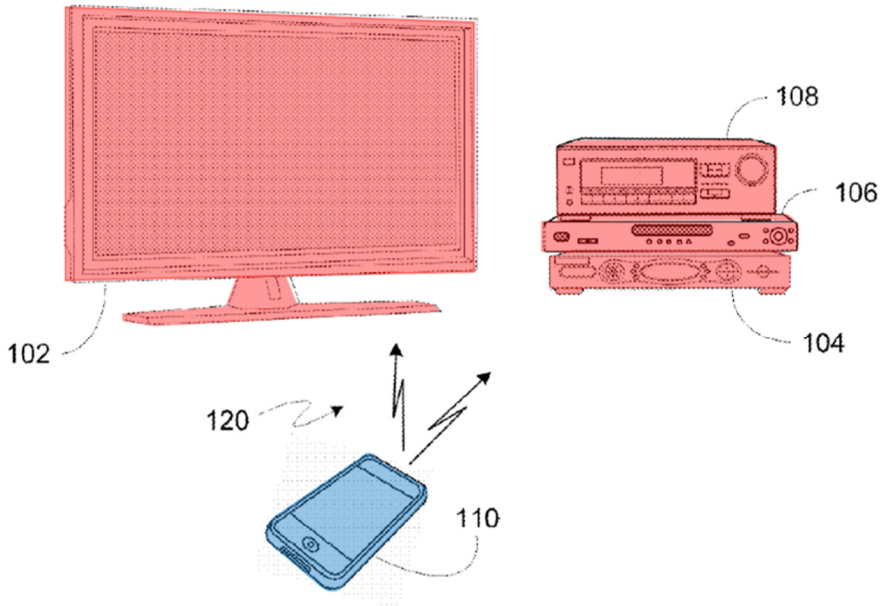


Figure 1b

147. Lim's ***control message*** is or comprises ***control data*** for controlling actuators in the ***second device***. For example, Lim's controlling device stores "data which serves to define the necessary control protocols and command values for use in transmitting ***command signals*** to controllable ***appliances*** (collectively, the ***command data***)." (EX1005, [0020].) As discussed above, such command/control data may be provisioned by the server . (*Supra* § IX.A.2.h.) "[T]he provisioned data may be utilized to transmit the desired ***command*** to the target ***appliance***." (EX1005, [0036].) Thus, Lim discloses that its "***command signals***" comprises control data (*e.g.*, Lim's "command values").

148. The actuators in Lim's ***second devices*** are controlled accordingly. For example, Lim discloses that such ***commands*** transmitted to the appliance, would for example, control a ***DVD player*** to, "[p]ower on," "[f]ast forward," "[p]ower off," "[f]ast reverse," "[p]ause," "[o]pen/close." (EX1005, [0031].) As another example, Lim discloses that "known synonyms for the command function [] may be used to cause a device such as a ***TV set or STB*** to tune from a currently selected channel to the immediately previous channel[.]" (EX1005, [0033].)

149. A POSITA would have understood that these operations involve controlling actuators such as displays, speakers, and motion actuators in these ***second devices*** such as ***TV, VCRs, DVRs, CD players, DVD players*** (*e.g.,* to playback a media item, control sound and/or provide visual information). (EX1018, 279-280, Figure 23.7.) For example, a POSITA would have understood that changing the channels of a TV or STB involve operating one or more displays and speakers (*i.e.*, actuators) in the TV or STB. It would have been obvious that the display and/or speaker and/or motors of these appliances are operated via the control message sent to the second devices. A POSITA would have understood that the motion actuators in Lim's "***fans, HVAC systems, thermostats***" are similarly controlled by the control message. (EX1020, 401, Fig. 13.1.)

150. Lim discloses that the ***control message*** is sent to the ***second device*** over the WLAN. Lim states "the ***controlling device 100*** may be capable of transmitting ***commands*** to the appliances, using any convenient IR, RF, Point-to-Point, or ***networked protocol 120***." (EX1005, [0017].) As shown in Figure 5b, the ***smart device 110*** communicates with the ***controllable appliances 102/104/106/108*** using ***networked protocol 120***. This "***networked protocol 120***"includes WiFi because Lim discloses that the first device (*e.g.*, ***smart device 110***) includes a ***WiFi transceiver/interface***. (EX1005, [0023], [0019].) Further, in or before 2012, home appliances that are WiFi-enabled were well-known. "[T]he mobile device may include a near field communication protocol that can be utilized for initiating and establishing connection with another wireless communication protocol (e.g., ***WiFi*** or Bluetooth) ***associated with the home appliance device***." (EX1012, [0018].) As discussed in "Wireless Home Networking Simplified", "[i]n the next wave, we see all types of devices connecting to the network, including some obvious devices such as ***TVs, stereos, and digital recording devices (DVRs)***, but also ***appliances***, ***home security systems, and electronics*** that can control home lighting, sprinkler systems, air conditioning, and answering machines." (EX1007, 321.) Thus, Lim certainly suggests, and it would have been obvious to a POSITA, that Lim's ***controlling device*** sends ***control messages*** via WiFi to the ***controllable***

appliances 102/104/106/108, which are WiFi-enabled and connected to the same home WLAN.

151. It would have been obvious to a POSITA that the first device's processor(s) are programmed with computer program instructions that, when executed, cause the first device to send a control message to a second device over the WLAN. (*Supra* § IX.A.2.e; EX1005, [0021].)

152. Lastly, with respect to claim 1, to the extent that the first and second human voice data are captured together (*e.g.*, as parts of one query) (which is not conceded), it would have been obvious that Lim, in response to receiving a user command related to operation of an appliance (*e.g.*, in FIG. 6), would provide a visual/audio feedback on the controlling device as it does in the configuration stage (FIG. 4) to provide confirmation and cause sending of a control message, thus meeting all elements of claim 1. Thus, a POSITA would have found it obvious, and Lim certainly suggests, that a query can result in both the first message and the second message, as Lim has the software and hardware configuration to do so.

k. 1[I][2]: "to directly or indirectly affect, change, or produce, a first physical phenomenon,"

153. As discussed above, Lim discloses controlling motors and actuators in the second devices (*e.g.*, ***televisions, VCRs, DVRs, DVD players, cable or***

satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers). Lim discloses that the first actuator in the second device is controlled to directly or indirectly affect, change, or produce, a first physical phenomenon.

154. The'721 patent states that a physical phenomenon comprises "temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current." (EX1001, 16:45-50, claims 5-6.) As discussed, the "first actuator" disclosed in Lim's second device include: speakers, displays, HVAC-related motors, VCR-related motors, etc. A speaker is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, audio or sound. A display is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, light. HVAC-related motors are configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, motion, temperature, or humidity. VCR-related motors are configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, motion of the drum head in a VCR. Thus, a POSITA would understand that Lim's actuator(s) (*e.g.*, speakers, displays, HVAC-related

motors, VCR-related motors) are controlled to directly or indirectly affect, change, or produce at least audio, light, motion, temperature, or humidity.

l. 1[I][3]: "in response to the received second message,"

155. Lim discloses this element.

156. As discussed, Lim discloses its ***first device*** receives a ***second message*** from the ***server***. The ***second message*** can be the ***speech processing result(s)*** that allows the ***first/controlling device*** to determine whether to proceed to block 610. (*Supra* § IX.A.2.h.) Responsive to the ***speech processing result(s)***, the ***first/controlling device*** obtains function data (block 610) and sends (block 614) the ***command*** to the ***appliance***.

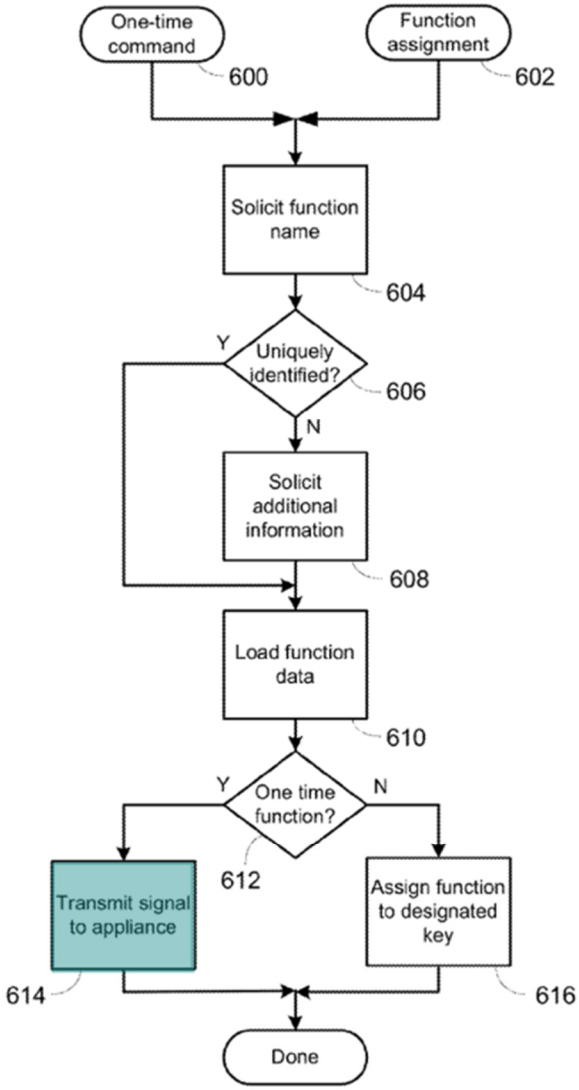


Figure 6

"Appliance command" Lim, [0035].

157. As discussed, a second example of the ***second message*** is the "***downloaded command code data***" from the ***remote server 500***. Lim discloses that an "app may serve to allow use of the ***smart phone*** to transmit ***commands*** suitable for operation of ***entertainment and other appliances***. Such an app may draw on a library of ***codesets*** suitable for ***commanding the functional***

operations of various types of appliances of multiple manufactures and/or models[.]" (EX1005, [0024].) "[T]his library . . . may be located externally at a local PC or ***remote server device*** accessible to ***smart device 110*** for the download of only those code sets necessary for operation of the configured ***appliances***[.]" (EX1005, [0024].) Lim discloses that the ***remote server device*** is the ***remote server 500***. (EX1005, [0030].) Then, after receiving the indicated command code data at block 610, "at step 614, the ***provisioned data*** may be utilized [for the ***controlling device***] to transmit the desired ***command*** to the ***target appliance***." (EX1005, [0036], [0017].)

158. In other words, at block 610, the ***smart device*** can download the ***second message*** (*e.g.*, ***indicated command data/codesets***) from the ***remote server device*** and, based on the downloaded data, transmit the ***command*** to the ***second device***. Thus, the ***command*** is sent by the ***first/controlling device*** to the ***second device*** in response to the received ***second message*** (*e.g.*, ***indicated command data/codesets***).

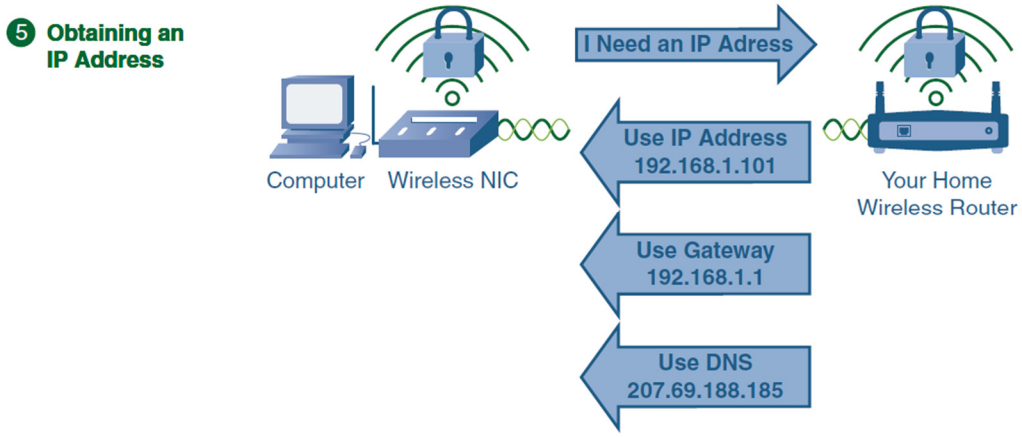
m. 1[J]: "wherein each of the first and second devices is addressable in the WLAN and in the Internet using a respective Internet Protocol (IP) address."

159. Lim discloses that each of the first and second device is addressable in the WLAN and in the Internet using an Internal Protocol (IP) address. As discussed, Lim discloses that each of its ***first device*** and ***second device*** is WiFi-

enabled and connects to the Internet via WLAN. (*Supra* §§ IX.A.2.a, IX.A.2.j; EX1012, [0018]; EX1007, 321.) A POSITA would have understood that this wireless network involves associating each device with a device identifier or address. Indeed, at the time of the invention, it was well-known that networked devices are associated with IP addresses. (EX1007, 99.) Indeed, Lim expressly confirms this. Lim discloses that the ***first/controlling device*** may request the user to "speak the name of a geographic region" and "used to filter search results when multiple possible matches were initially identified." (EX1005, [0028], [0038].) Lim discloses that "in embodiments which are ***Internet server based*** region information may be derivable from the ***IP address associated with the requesting device***[.]" (EX1005, [0038].) Thus, Lim expressly discloses that the requesting device (*e.g.*, the ***first/controlling device***) is addressable in the Internet using an IP address. (EX1005, [0038].)

160. A POSITA would have understood that the IP address may be an assigned IP address for the device (*e.g.*, by the router), the router's IP address, a combination of the two, or a globally unique IP address. "***Each computer or device that you want to join your wireless network needs a wireless network interface card (NIC)***." (EX1007, 99.) "If the wireless router grants access, it then typically ***assigns the NIC an IP address*** to use on the wireless network, at which point we can begin exchanging data," "***each device (computer, laptop,***

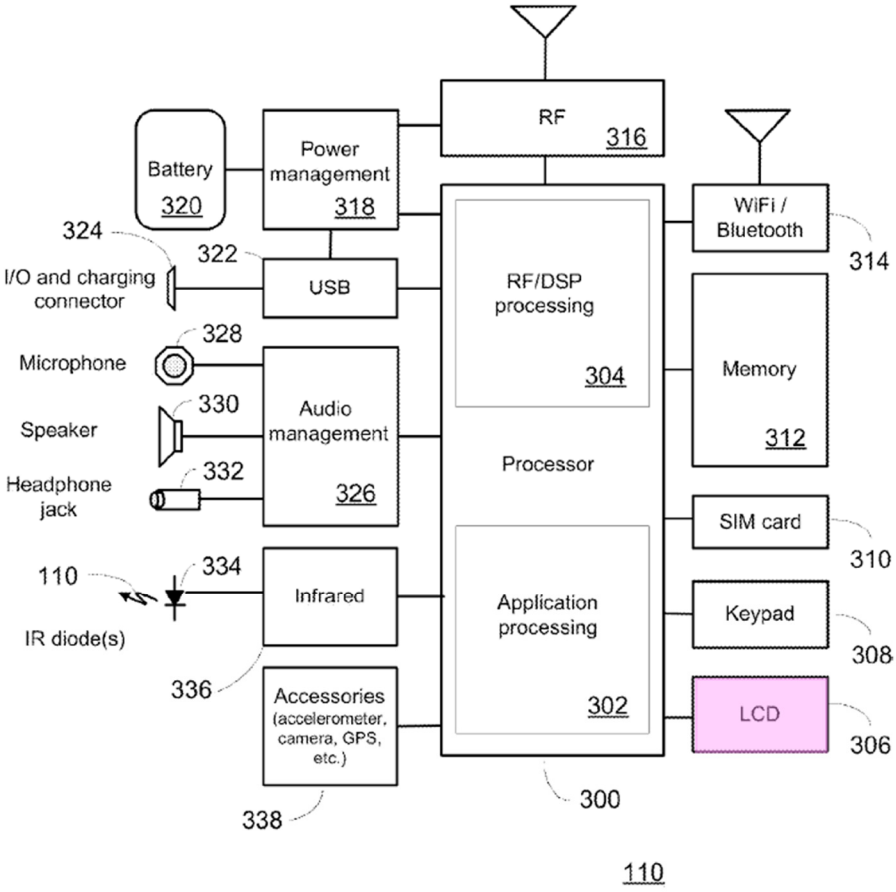
and so on) in a home network needs to have an IP address, which is how the device is known on the network. . . . Most home networks use IP addresses in the range of 192.168.1.100 to 192.168.1.149, and the ***router assigns each device an address*** from this pool when it boots." (EX1007, 108, 210, 175.)



3. Claim 5: "The first device according to claim 1, further comprising a second actuator that is configured for directly or indirectly affect, change, or produce, a second physical phenomenon."

161. Lim's ***first/controlling device*** further comprises a second actuator. For example, Lim discloses that the ***first/controlling device*** includes a display. Lim's "***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . a key matrix 216 (e.g., hard keys, soft keys such as a touch sensitive surface overlaid on a liquid crystal (LCD), . . . a means 202 to provide feedback to the user (e.g., one or more LEDs, illuminable keys, ***display***)[.]" (EX1005, [0019].) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an

LCD display 306; a keypad 308 which may comprise physical keys, touch keys overlaying LCD 306, or a combination thereof[.]" (EX1005, [0023].) As discussed above, a display is an actuator, consistent with the description of the '721 patent. (*Supra* §§ IX.A.2.j, IX.A.2.k.)



162. Moreover, Lim discloses that the first device has lights such as LEDs. The '721 patent explains an "actuator may be a light source that emits visible or non-visible light (infrared, ultraviolet, X-rays, or gamma rays) such as for illumination or indication. . . . The light source may be an electric light source for converting electrical energy into light . . . . The electric light source

may be based on Solid-State Lighting (SSL) such as a Light Emitting Diode (LED)[.]" (EX1001, 39:9-20.) Consistent with the '721 patent, Lim discloses that "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . , a key matrix 216 (*e.g.*, hard keys, soft keys such as a touch sensitive surface overlaid on a liquid crystal (LCD), . . . , a means 202 to provide feedback to the user (*e.g.*, ***one or more LEDs***, illuminable keys, display[.])" (EX1005, [0019].) It would have been obvious that ***smart device 110*** is similarly equipped with ***LEDs***.

163. Thus, the "second actuator" as recited and disclosed in Lim include displays and LED lights in the first device. A display is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, light. (*Supra* §§ IX.A.2.j, IX.A.2.k.) An LED light is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, light. (*Supra* §§ IX.A.2.j, IX.A.2.k.)

4. Claim 6: "The first device according to claim 5, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

164. As discussed above, Lim discloses a second actuator in Lim's first/controlling device that directly or indirectly affect, change, produce, or create at least light. (*Supra* § IX.A.3.)

5. Claim 9: "The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

165. A POSITA would have understood that this claim recites a standard microphone. (EX1001, 17:47-54.) Lim discloses such a standard microphone for capturing audio data. For example, Lim discloses that the "***controlling device 100*** for use in commanding the functional operations of one or more appliances may include . . . speech input means comprising ***a microphone 224*** and associated circuitry 222 for example an analog to digital converter, speech digitizer, etc., as known in the art." (EX1005, [0019].) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated ***microphone 328***, speaker 330, and headphone jack 332[.]" (EX1005, [0023].)

166. A POSITA would have understood that a standard microphone is an omnidirectional microphone based on sensing an incident sound-based motion of a diaphragm. (EX1014, 757, 760, 309.) Further, a POSITA would understand that dynamic microphone and piezoelectric microphone are standard and common implementations of microphone. (EX1014, 758-759.) Thus, a POSITA would understand (or it would be obvious) that Lim's microphone is a microphone as recited in claim 9.

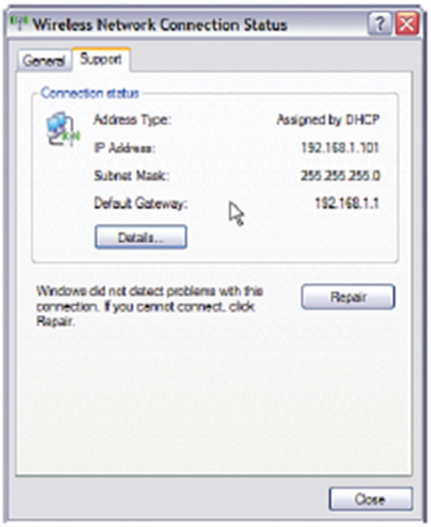
6. Claim 10: "The first device according to claim 1, wherein the IP address is stored in a volatile or non-volatile memory."

167. Lim discloses that the "***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, as needed for a particular application, a processor 200 coupled to a ***ROM memory 204, a RAM memory 205***, . . . a ***non-volatile read/write memory 206***[.]" (EX1005, [0019].) The ***smart device 110***, "illustrated in FIG. 3 in the form of a smart phone, may include . . . ***memory means 312 which may comprise ROM, RAM, Flash***, or any combination thereof[.]" (EX1005, [0023].)

168. As discussed above, Lim discloses that the controlling device is addressable in the Internet and WLAN with an IP address and that it communicates via WLAN and with remote server over the Internet via a wireless router. (*Supra* §§ IX.A.2.a, IX.A.2.j; EX1005, [0027], [0038], FIG. 5b.) A POSITA would have understood that the IP addresses would be locally stored in the disclosed memory in the ***controlling device 100*** and ***smart device 110***. For example, as shown in the figure below in "Wireless Home Networking Simplified", the networked devices can display the associated IP addresses, which means the IP addresses are locally stored in the device. (EX1007, 211, Figure 10-44, 265.) Thus, a POSITA would have understood

that the "ROM memory 204, RAM memory 205, non-volatile read/write memory 206, and memory means 312" as disclosed in Lim's first device (*e.g.*, ***smart device 110***) locally stores the IP addresses.

Figure 10-44 Verifying IP Address Assignment with Windows XP



7. Claim 11: ""The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type."

169. As discussed above, Lim discloses that the controlling device is addressable in the Internet and WLAN and that it communicates via WLAN and with the remote server over the Internet via a wireless router. (*Supra* §§ IX.A.2.a, IX.A.2.j; EX1005, [0027], [0038], FIG. 5b.) At the time of the invention, it was well-known that networked devices are associated with MAC addresses that are MAC-48, EUI-48 or EUI-64 type. (EX1007, 99, 166; EX1026, 2:50-59.) Thus, Lim suggests, and it would be obvious in view of

POSITA knowledge that Lim's networked devices are associated with MAC addresses. "***Each computer or device that you want to join your wireless network needs a wireless network interface card (NIC)***" and "***each wireless card has a unique identifier called a MAC address***." (EX1007, 99, 166.) "Most MAC addresses use one of the IEEE standardized protocols; ***MAC-48, EUI-48®, and EUI-64®***." (EX1026, 2:50-59.)

8. Claim 12: "The first device according to claim 1, wherein the IP address is a static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address."

170. As discussed above, Lim discloses that the controlling device is addressable in the Internet and WLAN with an IP address and that it communicates via WLAN and with the remote server over the Internet via a wireless router. (*Supra* §§ IX.A.2.a, IX.A.2.j, IX.A.6; EX1005, [0027], [0038], FIG. 5b; EX1007, 99, 108, 210, 175.)

171. At the time of the invention, it was well-known that Internet Protocol (IP) addresses are either static or dynamic, and are IPv4 addresses. "In the case of today's Internet, these addresses are Internet protocol version 4 (IPv4) addresses." (EX1029, 127.) "Internet protocol (IP) addresses are the unique numbers used to identify individual connections to the Internet." *Id.* "The broadband networks typically assign a *dynamic IP address* to your broadband service. *Dynamic* means the IP address can change[.]" (EX1007,

277.) Further, an IP address may be a "*static IP address*, which means . . . your address never changes." (EX1007, 277.)

9. Claim 13: "The first device according to claim 1, wherein the WLAN is according to, or base on, Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac."

172. Lim discloses or renders obvious this claim. As discussed, Lim discloses that the first device communicates with the ***remote server 500*** (*i.e.*, provisioning the speech recognition result and/or additional prompts for voice input and/or indicated command code data to the first device) over the ***Internet*** (*i.e.*, network 502), and ***via WiFi***, a form of WLAN, through a gateway device such as a ***router***. (*Supra* §§ IX.A.2.a, IX.A.2.g; EX1005, [0027], FIG. 5b.) Lim discloses that access to "a remote server 500 accessible via a network 502 such as the Internet" "may be via an intermediary appliance such as STB 104, illustrated at 506, or may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].)

173. A POSITA would have understood that WiFi implements a WLAN and is based on the IEEE 802.11 standards which include the 802.11a/b/g family. "Wireless local area networks (***WLANs***) are built around the family of IEEE 802.11 standards. . . . The 802.11a/b/g family of standards are collectively

referred to as ***Wi-Fi***, for wireless fidelity." (EX1040, 600-601.) "***Wi-Fi*** is a registered trademark owned by the Wi-Fi Alliance, and it refers to ***wireless local area network (WLAN)*** products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards." (EX1013, 7:48-51.) This is consistent with the description of the '721 patent which states that "[i]n order to allow interoperability between equipment manufactured by different vendors, few ***WLAN standards*** have evolved, as part of the ***IEEE 802.11 standard group***, ***branded as WiFi*** (www.wi-fi.org). IEEE 802.11b describes a communication using the 2.4 GHz frequency band and supporting communication rate of 11 Mb/s, IEEE 802.11a uses the 5 GHz frequency band to carry 54 MB/s and IEEE 802.11g uses the 2.4 GHz band to support 54 Mb/s." (EX1001, 8:25-32.) Thus, Lim discloses or suggests that its WLAN is according to the IEEE802.11 standards.

10. Claim 14: "The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

174. Lim discloses or renders obvious this claim. As discussed, Lim discloses that the first device communicates with the ***remote server 500*** (*i.e.*, provisioning the speech recognition result and/or additional prompts for voice input and/or indicated command code data to the first device) over the ***Internet*** (*i.e.*, network 502), and ***via WiFi***, a form of WLAN, through a gateway device

such as a ***router***. (*Supra* §§ IX.A.2.a, IX.A.2.g; EX1005, [0027], FIG. 5b.) Lim discloses that access to "a remote server 500 accessible via a network 502 such as the Internet" "may be via an intermediary appliance such as STB 104, illustrated at 506, or may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].)

175. A POSITA would understand that WiFi (*e.g.*, WLAN over 802.11b) is a form of ISM Band II frequency. (EX1040, 22-23, Table 1.2.) It was common knowledge at the time of the invention.

Table 1.2: Unlicensed U.S. spectrum allocations

Band	Frequency
ISM band I (cordless phones, 1G WLANs)	902–928 MHz
ISM band II (Bluetooth, 802.11b and 802.11g WLANs)	2.4–2.4835 GHz
ISM band III (wireless PBX)	5.725–5.85 GHz
U-NII band I (indoor systems, 802.11a WLANs)	5.15–5.25 GHz
U-NII band II (short-range outdoor systems, 80211a WLANs)	5.25–5.35 GHz
U-NII band III (long-range outdoor systems, 80211a WLANs)	5.725–5.825 GHz

176. Thus, Lim discloses or at least suggests that the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.

11. Claim 15: "The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal

Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution."

177. Lim discloses or renders obvious this claim. Lim's ***first/controlling device*** can be a wireless smartphone. (*Supra* § IX.A.2.a; EX1005, [0027].) Thus, Lim suggests that the first device is configured to communicate over a cellular telephone network. (EX1053.)

178. A POSITA would understand that 3G, the third-generation cellular network, GSM (Global System for Mobile Communications), and EDGE (Enhanced Data rates for GSM Evolution) are common and known cellular communications network in or before 2012. (EX1040, 12, 13, 599-600.) Further, wireless technology commonly included cellular connectivity in 2012. Thus, a POSITA would understand (and it would be obvious) that Lim's device is configured to communicate over the recited cellular telephone network.

12. Claim 17: "The first device according to claim 1, wherein the second device is integrated in, is part of, or is entirely included in, a household appliance having a primary function."

179. Lim discloses a second device that is integrated in, is part of, or is entirely included in, an appliance. As discussed, the second device in the Lim system includes "***various controllable appliances, such as a television 102, a set top box ('STB') 104, a DVD player 106 and an AV receiver 108***." (*Supra*

§ IX.A.2.j; EX1005, [0017].) "[C]ontrollable ***appliances*** may include, but need not be limited to, ***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers***, etc." (EX1005, [0017].) These devices are appliances having primary functions.

13. Claim 21: "The first device according to claim 17, wherein the primary function of the appliance is associated with environmental control, and the appliance is part of a Heating, Ventilation and Air Conditioning (HVAC) system."

180. Lim discloses controlling various appliances, such as "televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, ***fans, HVAC systems, thermostats***, personal computers, etc." (EX1005, [0017].) Thus, Lim discloses controlling an appliance that is part of a HVAC system and has a primary function associated with environmental control.

14. Claim 22: "The first device according to claim 21, wherein the primary function of the appliance is associated with temperature control, and wherein the appliance is an air conditioner or a heater."

181. As discussed, Lim discloses controlling various appliances, such as "televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, ***fans, HVAC systems, thermostats***, personal computers, etc." (EX1005, [0017].) A POSITA would have understood that fans, HVAC systems, and

thermostats are appliances associated with temperature control. Further, "HVAC" stands for heating, ventilation, and air conditioning and thus an HVAC system comprises an air conditioner or a heater.

15. Claim 24: "The first device according to claim 17, wherein the appliance is an answering machine, a telephone set, a home cinema system, a High Fidelity (HiFi) system, a Compact Disc (CD) or Digital Video Disc (DVD) player, an electric furnace, a trash compactor, a smoke detector, a light fixture, or a dehumidifier."

182. Lim discloses controlling various appliances, such as "***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs')***, amplifiers, ***CD players***, game consoles, ***home lighting***, drapery, fans, HVAC systems, thermostats, personal computers, etc." (*Supra* § IX.A.2.j; EX1005, [0017].)

16. Claim 25: "The first device according to claim 1, further configured to operate within a building, and wherein the server device is external to the building."

183. Lim discloses that the first device is configured to operate within a building. As discussed above, Lim discloses or certainly suggests that the ***first/controlling device*** (*e.g.*, a smart phone or other portable devices) is operated within a ***building*** such as a ***home***. Lim discloses that the ***smart device 110*** "may comprise any of a smart phone, a tablet computer, a portable email device, an e-book, a hand-held games and/or game controller, a portable media player, etc." (EX1005, [0018].) It would have been obvious that these are

devices that users can carry with them, including when the user is at ***home***, such that the user can conveniently control home appliances while they are at home. For example, at the time of the time of the invention, a "portable device [can be] readily carried by the operator" and it "allows an operator to easily and readily control ***home activities*** without having to directly interface with or use a PC, and which allows the operator to program and control the home gateway ***from essentially any area of the home*** without having to access a PC." (EX1021, 1:53-58, 2:19-26.)

184. Further, Lim discloses that the server device is external to the building. Lim's server device is a "***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027].) It would have been obvious to a POSITA that the ***remote server 500*** is external to the user's ***home*** because it is "remote." Lim further states, "[a]s will become apparent, speech recognition and command data identification processing may be performed . . . ***remotely*** at a ***Web server***[.]" (EX1005, [0007].) Further, Lim differentiates "remote server" from a "local PC". (EX1005, [0024].) At the time of the invention, the benefits for having a server in the cloud (*i.e.*, outside a user's home) perform resource-intensive tasks such as speech processing were well-known. "By performing complex speech and voice processing remotely, the local device can have fewer components, be smaller, use less power, and cost

less than a device that performs these operations locally. Further, upgrades to speech and voice recognition software would not require significant reconfiguration of the local device." (EX1008, 8:26-32; *see also* EX1006, [0022], Figure 2; EX1009, [0080], [0100]-[0102], [0110]-[0117].)

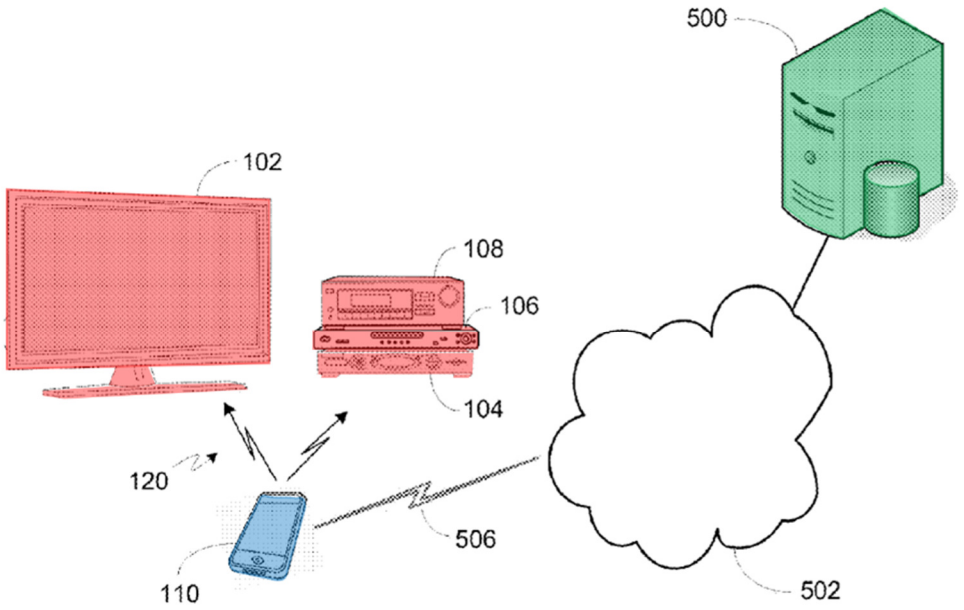


Figure 5b

17. Claim 26: "The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays."

185. As discussed, Lim's ***first device*** includes ***LEDs***. Lim discloses "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . , a means 202 to provide feedback to the user (*e.g.*, ***one or more LEDs***, illuminable keys, display, buzzer, voice synthesizer with associated speaker, and/or the like)[.]" (EX1005, [0019];

supra §§ IX.A.3, IX.A.4.) A POSITA would have understood that Lim's LEDs are electric light sources for converting an electrical energy into light that emits a visible light for illumination or indication. "Light-emitting diodes (LEDs) indicators are diodes that emit light when current passes in the forward direction." (EX1018, 174.) This is consistent with the '721 patent, which explains that "[t]he actuator may be a light source that emits visible or non-visible light (infrared, ultraviolet, X-rays, or gamma rays) such as for illumination or indication. . . . The light source may be an electric light source for converting electrical energy into light . . . . The electric light source may be based on Solid-State Lighting (SSL) such as a ***Light Emitting Diode (LED)***[.]" (EX1001, 39:9-20.)

18. Claim 27: "The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode."

186. As discussed, Lim's first actuators include ***LEDs*** in the ***first device***. Lim discloses "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . a means 202 to provide feedback to the user (e.g., ***one or more LEDs***, illuminable keys, display, buzzer, voice synthesizer with associated speaker, and/or the like)[.]" (EX1005, [0019]; *supra* § IX.A.3.)

19. Claim 28: "The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

187. Claim 28 depends on claim 27. Besides comprising an electric light source, Lim's device comprises a sounder that plays digital audio content as recited in claim 28.

188. For example, Lim discloses "a controlling device 100 for use in commanding the functional operations of one or more appliances may include, . . . a means 202 to provide feedback to the user (e.g., one or more LEDs), illuminable keys, display, ***buzzer, voice synthesizer with associated speaker***, and/or the like[.]" (EX1005, [0019].) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an LCD display 306; . . . an audio management system 326 with associated microphone 328, ***speaker 330***[.]" (EX1005, [0023]; *supra* §§ IX.A.2.d, IX.A.2.i.)

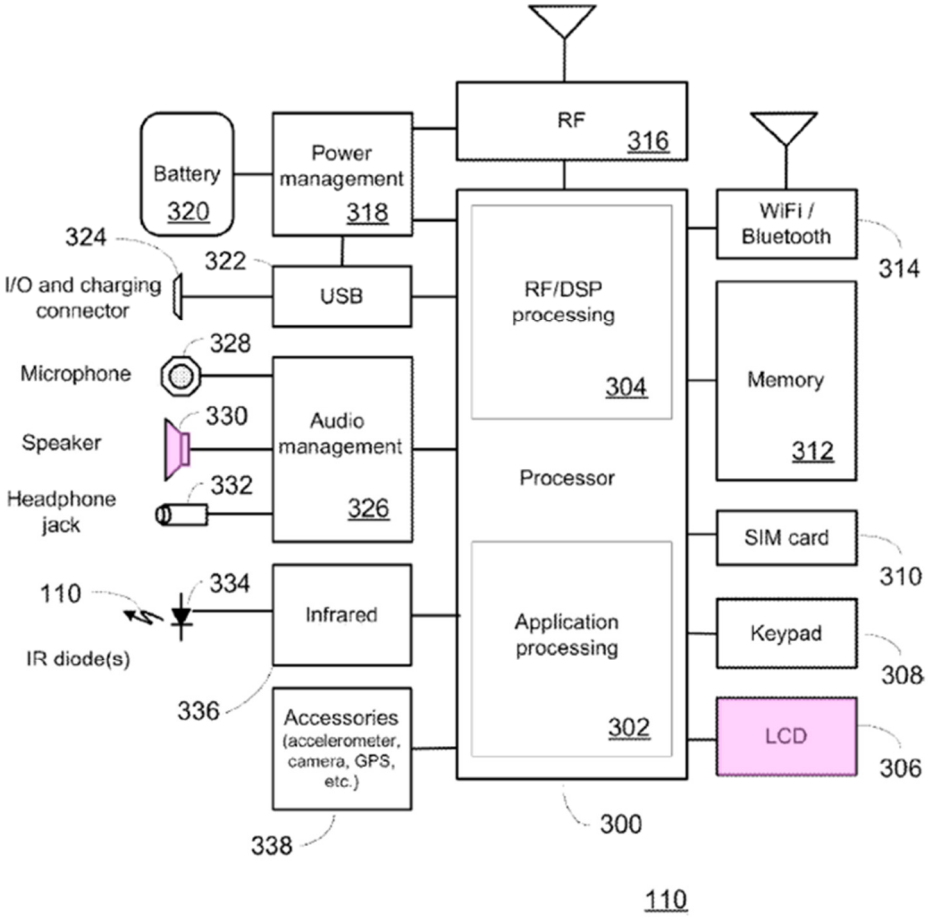


Figure 3

189. As discussed above, Lim discloses operations of its speaker in the first device (*e.g.*, controlling device 100 and smart device 110) by providing an auditory prompt. (*Supra* § IX.A.2.i.) The prompt is provided in response to a user voice input and, more specifically, in response to ***information*** (*e.g.*, recognized speech, analysis of or information associated with the recognized speech, and/or additional prompts) received from the speech recognition ***server***. For example, "at step 402 the controlling device may prompt the user to speak the brand name of the appliance to be added to the configuration of the

controlling device. This and subsequent prompts for voice input may take the form of an illuminated LED and/or 'speak' symbol, a displayed text or image, a ***synthesized spoken message*** (*e.g.*, where controlling device ***user feedback means 202*** comprises speech synthesis means and ***a speaker***, or where a smart device app has access to ***speaker 330***), etc." (EX1005, [0027].)

190. Thus, Lim discloses that the sounding comprises at least playing digital audio content that is synthesized and simulating the voice of a human being, and sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.

20. Claim 29: "The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

191. A POSITA would have understood that this claim recites standard speakers. (EX1001, 105:48-54, 105:61-106:4.) Lim discloses such a standard speaker for sounding digital audio content. As discussed, Lim discloses that its ***controlling device*** comprises a ***speaker***. For example, "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . a means 202 to provide feedback to the user (e.g., one or more LEDs, illuminable keys, display, ***buzzer, voice synthesizer with associated speaker***, and/or the like)[.]" (EX1005, [0019].) "A representative ***smart***

device, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated microphone 328, ***speaker 330***[.]" (EX1005, [0023].) A POSITA would have understood that an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker, a ribbon or planar magnetic loudspeaker are standard and common implementations of speakers. (EX1014, 321-322, 763.) Thus, a POSITA would have understood (or it would be obvious) that Lim's speaker is a loudspeaker as recited in claim 29.

21. Claim 30: "The first device according to claim 1, wherein the sounder comprises an electric bell, a buzzer, a chime, a whistle, or a ringer."

192. Lim discloses this claim. A POSITA would have understood that this claim refers to standard acoustic devices such as buzzers. (EX1001, 106:54-62.) As discussed, Lim discloses that its ***controlling device*** has a ***buzzer***. "A ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . a means 202 to provide feedback to the user (e.g., one or more LEDs, illuminable keys, display, ***buzzer,*** voice synthesizer with associated speaker, and/or the like)[.]" (EX1005, [0019].)

193. Thus, Lim's first device comprises a sounder which is or comprises a buzzer.

22. Claim 31: "The first device according to claim 1, further comprising an enclosure that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors."

194. Lim discloses an ***enclosure*** that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors for the ***first device***.

195. As discussed, Lim's first device comprises a WLAN transceiver, a sounder, a microphone, and one or more processors. (*Supra* §§ IX.A.2.b, IX.A.2.c, IX.A.2.d, IX.A.2.e.) For example, Lim's ***smart device 110***, "illustrated in FIG. 3 in the form of a smart phone, may include as needed for a particular application, ***processing means 300*** which may comprise both an application processing section 302 and an RF/DSP processing section 304; . . . ***WiFi and/or Bluetooth wireless interface(s) 314***; . . . an audio management system 326 with associated ***microphone 328, speaker 330***, and headphone jack 332[.]" (EX1005, [0023].)

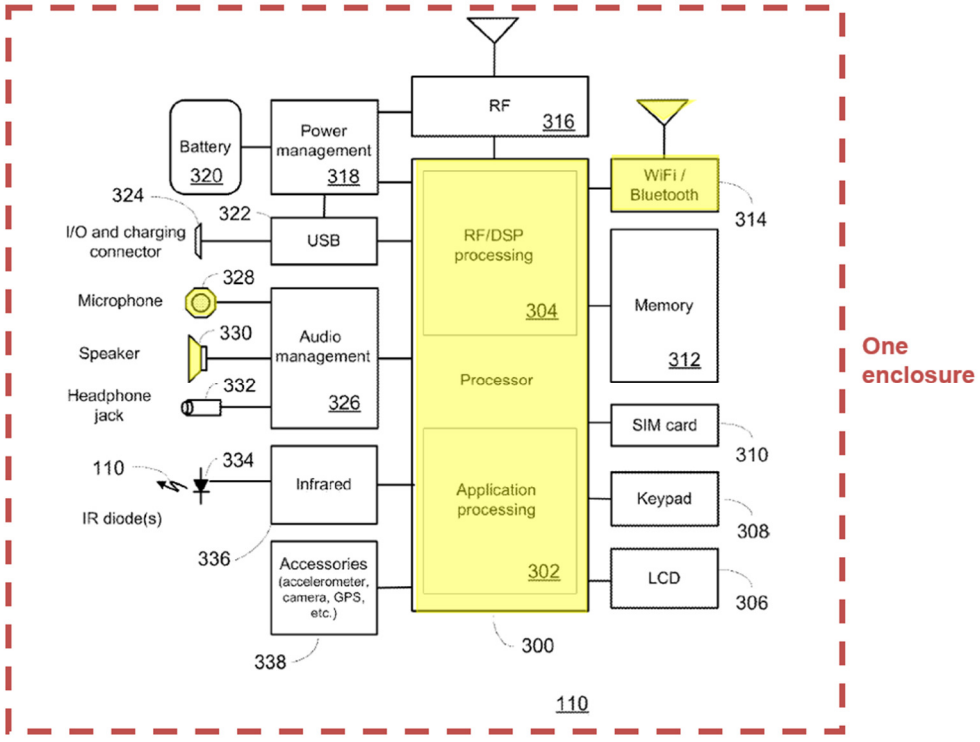


Figure 3

196. As shown in Figure 1b below, Lim's controlling device, ***smart device 110*** has a smart phone housing/enclosure.

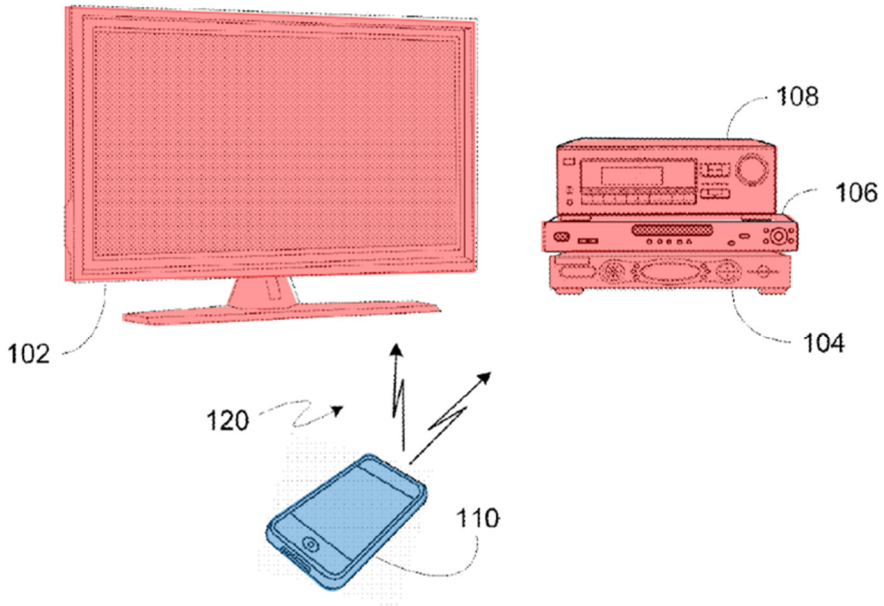


Figure 1b

197. Lim discloses that "FIG. 3 illustrates a block diagram of exemplary components of the controlling device of FIG. 1b[.]" (EX1005, [0012].) All the exemplary components of the smart device 110 which includes (the WLAN transceiver, the sounder, the microphone, and the one or more processors) are enclosed within an enclosure of the smart phone.

23. Claim 33: "The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves."

198. A POSITA would have understood that this claim recites a standard speaker. (EX1001, 105:61-106:4.) As discussed, Lim discloses a standard speaker for sounding digital audio content. (EX1005, [0019], [0023].) A POSITA would have understood that a standard speaker is a sounder for

converting an electrical energy to omnidirectional pattern emitted audible sound waves. (EX1018, 122; EX1014, 763, 765.) Thus, a POSITA would understand (or it would be obvious) that Lim's speaker is a sounder as recited in claim 33.

B. Ground 2: Lim and Wittenberg Render Obvious Claim 2

1. Summary of Wittenberg

199. Wittenberg was filed on January 21, 2005, published on May 5, 2009, and is prior art under § 102(a), (b), (e). (EX1008, 1.) The Examiner never found, and Applicant never cited, Wittenberg during prosecution. Wittenberg is related to "[a] technique for remotely processing a local audio command to control a local device[.]" (EX1008, Abstract.) Specifically, Wittenberg discloses "performing speech recognition processing on the audio signal at the remote site to determine whether the audio signal includes a command; performing voice recognition processing on the audio signal at the remote site to ***determine whether the audio signal has been supplied by an authorized user***; generating a command signal in response to the audio signal including a command and being supplied by an authorized user; and transmitting the command signal to a device at the local site to effect a change in a state of the local device." (EX1008, Abstract, Claim 1.)

2. Claim 2: "The first device according to claim 1, wherein the first device or the server device is configured for processing the

first and second human voice data using a voice recognition algorithm for identifying a voice of a specific person."

200. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2) Claim 2 is obvious in further view of Wittenberg. Wittenberg discloses, with reference to FIG. 2, that "[v]oice recognition processor analyzes audio signals received at the remote processor to determine whether the person issuing the voice command is authorized to issue the command (operation 250 in FIG. 2). The processing involves comparing the characteristics (e.g., frequency spectrum, speech patterns, etc.) of the audio signal to stored voice information to ***determine whether the audio signal voice matches that of an authorized user***. Algorithms and processing techniques for performing voice recognition are well known in the art and not described here in detail. Voice recognition processing provides an additional degree of security, which prevents an unauthorized person at the local site from affecting or manipulating the local controllable device in some manner." (EX1008, 5:61-6:6.)

201. At the time of the invention, it would have been obvious to combine the teachings of Wittenberg with Lim. Both Wittenberg and Lim relate to remotely processing audio information such as speech for controlling devices or providing services through devices. For example, Lim relates to "controlling devices such as universal remote controls" and "specific appliances to be

controlled." (EX1005, [0007].) Lim discloses that "the ***controlling device*** may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. . . . The resulting spoken syllables may be digitized and presented to a speech recognition means for processing . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027].) Wittenberg discloses "[a] method for remotely processing a locally-generated voice command to control a local device" and furthermore preventing unauthorized persons from controlling the local devices. (*See* EX1008, Claim 1.) Thus, Lim and Wittenberg are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same problem of controlling devices based on audio commands and ensuring security within the command/control system architecture.

202. A POSITA would have been motivated to combine Lim with Wittenberg to include the voice recognition of Wittenberg to achieve greater security and with a reasonable expectation of success. Wittenberg teaches additional processing techniques, for example, "the method comprising: (a) receiving at a local site an acoustic signal and generating an audio signal from the acoustic signal; (b) transmitting the audio signal to a remote site; (c) ***performing speech recognition processing*** on the audio signal at the remote site

to identify a voice command from the audio signal; (d) ***performing voice recognition processing*** on the audio signal at the remote site to determine whether the audio signal has been supplied by an authorized user[.]" (EX1008, Claim 1.) The voice recognition processing techniques can be combined with the teachings of Lim to yield a command/control system architecture that provides added security by determining whether an authorized person issued the voice command, with a reasonable expectation of success, as Lim already performs server-side speech recognition, making the addition a straightforward software modification.

203. Wittenberg describes a security issue that unauthorized users may be able to control devices via audio commands. "[F]or example, if an unauthorized person gains access to the portion of the system that receives voice commands, that person can instruct the system to perform unauthorized commands which may be harmful." (EX1008, 2:21-30.) Wittenberg expressly teaches the benefits for including voice recognition during processing to address the issue and provide enhanced security. "Voice recognition processing provides an ***additional degree of security***, which prevents an unauthorized person at the local site from affecting or manipulating the local controllable device in some manner." (EX1008, 5:61-6:6.)

204. Wittenberg recognizes the benefits of applying its techniques in numerous circumstances. "There are a numerous circumstances in the government/military, commercial, and private sectors where it would be advantageous to control a local device with a locally-generated voice command without placing the demands of speech processing on a local device and with the capability to provide a significant degree of security in the process of interpreting and carrying out voice commands." (EX1008, 2:25-31.) Specifically, Wittenberg recognizes the benefits of applying its techniques in home appliances. "For example, in the case where the local controllable device is a television, voice recognition processing can prevent an unauthorized person (e.g., a child) from accessing certain channels or ordering pay-per-view shows, whereas an authorized person (e.g., a parent) can be authorized to access all channels by issuing voice commands via a telephone or the like." (EX1008, 6:7-13.)

205. The Lim-Wittenberg combination results in a system where voice recognition technique as taught by Wittenberg is incorporated in Lim's server to analyze human audio commands in Lim. For example, the Lim-Wittenberg combination would determine if the person issuing the voice data is authorized to make the command (*e.g.*, homeowner) as a part of processing at the server (application engine 118) as taught by Wittenberg. This combination would

predictably enhance the security of the system because only authorized users would be able to control devices through their voice commands. The known benefits of Wittenberg would prompt the application of Wittenberg in the same field based on design incentives and market forces and yield the predictable result of improving the security of the control and command systems and prevent control by unauthorized users.

206. Because both Lim and Wittenberg include voice data processing and device control, a POSITA would be motivated to use the known algorithm and techniques of voice recognition processing of Wittenberg to improve the security of Lim in the same way. Further, it would have been obvious to try combining the teachings of Wittenberg and Lim. There are only a finite number of identified and predictable solutions for user authentication, such as authorization via biometrics like facial recognition, fingerprint, or making use of contextual data like device location and time to supplement the voice data processing. The combination is a straightforward software modification and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Wittenberg has successfully demonstrated that performing voice recognition can be implemented to improve security of the system and prevent unauthorized control, and thus, there would have been a reasonable expectation of success for

combining Lim and Wittenberg as Lim already performs server-side speech processing, making the addition a straightforward software modification. Moreover, at the time of the invention, algorithms and processing techniques for performing voice recognition were known. "Algorithms and processing techniques for performing voice recognition are well known . . . ." (EX1008, 6:1-3.)

C. Ground 3: Lim and Gruber Render Obvious Claim 3-4, 11

1. Summary of Gruber

207. Gruber was filed on September 30, 2011, and is prior art under § 102(e). (EX1009, 1.) The Examiner never found, and Applicant never cited, Gruber during prosecution. Gruber relates to "virtual assistants, and more specifically to mechanisms for improving interpretation and processing of commands provided to such an assistant." (EX1009, [0007].) Gruber discloses that "a virtual assistant uses context information (also referred to herein as 'context') to supplement natural language or gestural input from a user. This helps to clarify the user's intent and to reduce the number of candidate interpretations of the user's input, and reduces the need for the user to provide excessive clarification input. Context can include any available information that is usable by the assistant to supplement explicit user input to constrain an

information-processing problem and/or to personalize results." (EX1009, [0013].)

208. Gruber discloses that "a physical device running virtual assistant 1002 may have one or more sensors. Such sensors can provide sources of contextual information. Example of such information include, without limitation: the user's current location; the local time at the user's current location; the position, orientation, and motion of the device; the current light level, temperature and other environmental measures; the properties of the microphones and cameras in use; the current networks being used, and signatures of connected networks, including Ethernet, Wi-Fi and Bluetooth. Signatures include MAC addresses of network access points, IP addresses assigned, device identifiers such as Bluetooth names, frequency channels and other properties of wireless networks. Sensors can be of any type including for example: an accelerometer, compass, GPS unit, altitude detector, light sensor, thermometer, barometer, clock, network interface, battery test circuitry, and the like." (EX1009, [0218]-[0225].)

2. Claim 3: "The first device according to claim 1, further comprising a sensor that outputs sensor data that responds to a physical phenomenon, wherein the computer program instructions, when executed, further cause the first device to send to the server device, via the WLAN over the Internet, the sensor data, and wherein the first message is further responsive to the sent sensor data."

209. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2) Claims 3-4, 11 are obvious in further view of Gruber. First, Gruber discloses that the first device comprises a sensor that outputs sensor data that responds to a physical phenomenon. For example, Gruber discloses that "***a physical device*** running virtual assistant 1002 ***may have one or more sensors***. ***Such sensors can provide sources of contextual information***. Example of such information include, without limitation: the user's current location; the local time at the user's current location; the position, orientation, and motion of the device; the current light level, temperature and other environmental measures; the properties of the microphones and cameras in use; the current networks being used, and signatures of connected networks, including Ethernet, Wi-Fi and Bluetooth. Signatures include MAC addresses of network access points, IP addresses assigned, device identifiers such as Bluetooth names, frequency channels and other properties of wireless networks. Sensors can be of any type including for example: an accelerometer, compass, GPS unit, altitude detector, light sensor, thermometer, barometer, clock, network interface, battery test circuitry, and the like." (EX1009, [0218]-[0225].)

210. Gruber also discloses sending to the server, by the first device via the wireless network, the sensor data and that the first message is sent by the server in response to the sensor data. For example, Gruber discloses that the

contextual information provided by sensors can be transmitted to a speech processing server to process human voice data to clarify the user's intent, reduce the number of candidate interpretations, and constrain solutions during speech recognition, natural language processing, etc. For example, Gruber discloses that "[i]n addition to user input 2704 and background sources, virtual assistant 1002 can also draw on information from several sources of context, including for example ***device sensor data 1056***, application preferences and usage history 1072, dialog history and assistant memory 1052, personal databases 1058, personal acoustic context data 1080, current application context 1060, and event context 2706." (EX1009, [0179].) Gruber discloses that the "virtual . . . assistant 1002 also ***receives context information 1000***[.]" (EX1009, [0168].) "[D]ifferent types of virtual assistant features and/or functionalities may be implemented in a ***virtual assistant system*** which includes a client device (such as a personal digital assistant or smartphone running client software) and ***server system(s)*** . . .)[.]" (EX1009, [0080].)

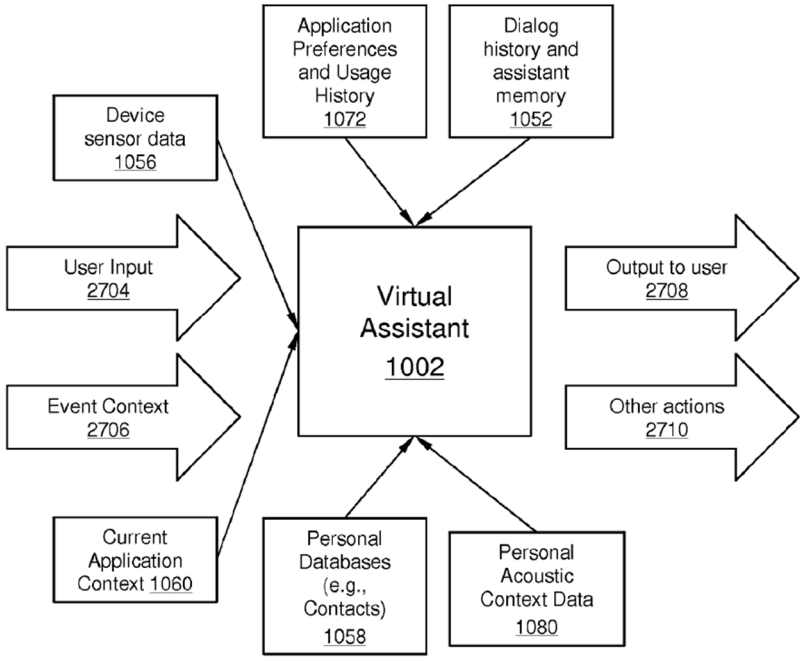


FIG. 1

211. "Client device 1304, which may be a mobile computing device or other device, can be the ***source of contextual information 1000*** such as ***device sensor data 1056*** . . . ***context 1000 can be communicated among distributed components such as client 1304 and server 1340***." (EX1009, [0366]-[0367].) "The following components are located at server 1340: complete vocabulary 2758b; complete library of language pattern recognizers 2760b;" and "[a]dditional components may be implemented as part of server 1340, including for example: language interpreter 2770; dialog flow processor 2780; output

processor 2790; . . . services orchestration 2782[.]" (EX1009, [0100] – [0102], [0110]-[0117].)

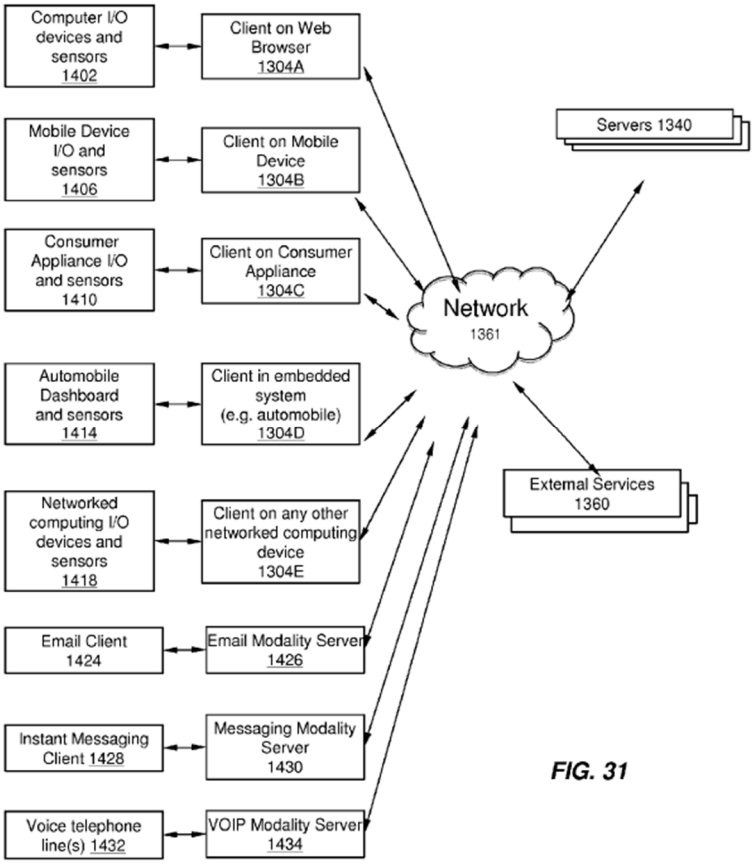


FIG. 31

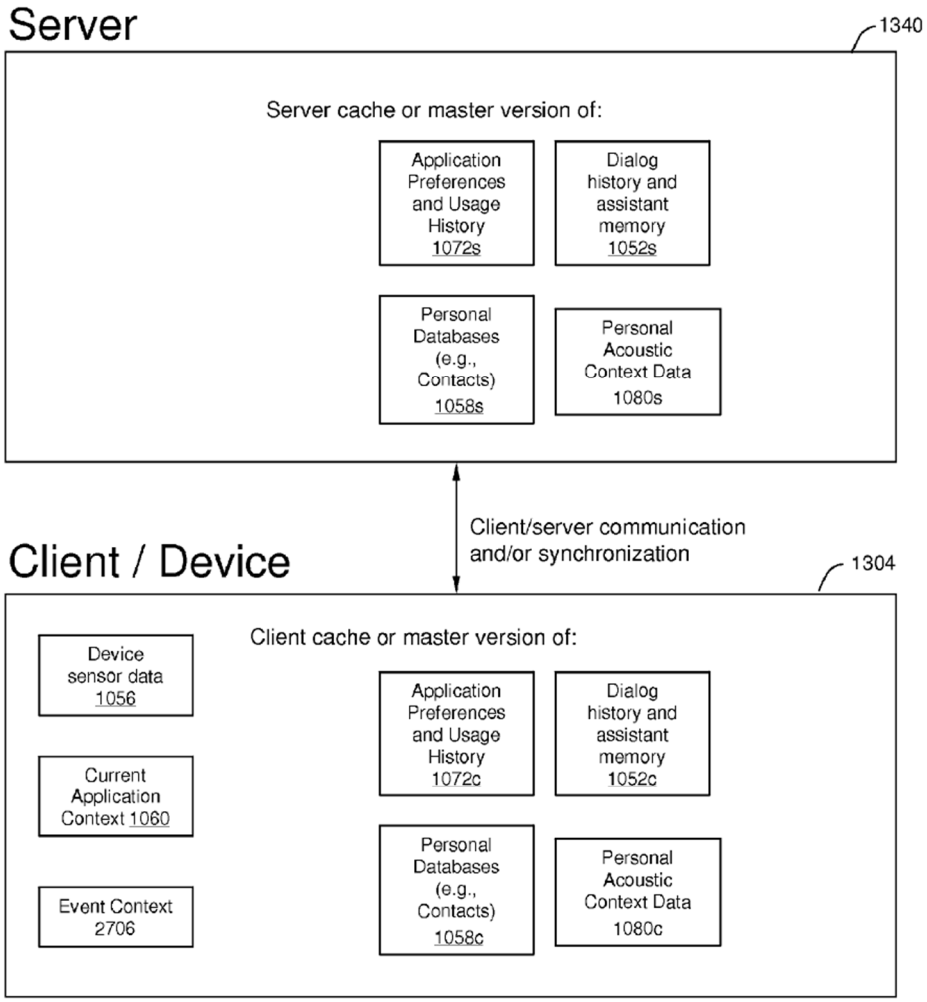


FIG. 6

212. Gruber further discloses that the contextual information provided by sensors (*e.g.*, environmental conditions) can be transmitted to a server (*e.g.*, server 1340) hoisting the vocabulary and language interpreter for "interpreting the received user input." (EX1009, [0100]-[0104], [0110]-[0111], Claims 43, 49.) For example, environmental conditions may serve as "current values" for

task parameters that may be implicit in a user's request. (EX1009, [0354].) If a user asks, "what is the weather like?" the sensor data can be used to provide an answer without the user needing to specify the location/environmental parameter. (EX1009, [0354].)

213. It would have been obvious to a POSITA to combine the teachings of Lim and Gruber. Lim and Gruber both relate to processing and responding to human speech commands. For example, Lim relates to "controlling devices such as universal remote controls" and "specific appliances to be controlled." (EX1005, [0007].) Lim discloses that "the ***controlling device*** may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. . . . The resulting spoken syllables may be digitized and presented to a speech recognition means for processing . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027].) Gruber discloses that the "invention relates to virtual assistants, and more specifically to mechanisms for improving interpretation and processing of commands provided to such an assistant." (EX1009, [0007].) Thus, Lim and Gruber are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problems of

processing and responding to human speech commands (*e.g.*, to provide services).

214. A POSITA would have been motivated to combine Lim with Gruber because Gruber teaches sending contextual information (*e.g.*, temperature, location), along with speech commands, to the server to improve server-side speech processing. Lim's controlling device already (1) has built-in sensors (*e.g.*, GPS, accelerometer) and (2) sends voice commands to the server for processing. Gruber teaches additional speech processing techniques of using contextual information obtained from device sensors to clarify the user's intent, reduce numbers of candidate interpretations, and/or constrain solutions during speech recognition and natural language processing. Thus, adding more sensors and sending sensor data to the server, as taught by Gruber, would be a straightforward, predictable improvement with a reasonable expectation of success. (EX1009, Abstract; EX1005, [0023].)

215. Further, Lim itself discloses that the first device would be a smart phone that includes sensors. Lim discloses "[a] representative smart device, illustrated in FIG. 3 in the form of a smart phone, may include . . . accessory features 338 such as a digital camera, GPS, accelerometer, etc." (EX1005, [0023].) Hence adding additional sensors in Lim to supply contextual

information to assist speech processing would be a straightforward, predictable improvement with a reasonable expectation of success.

216. Specifically, the Lim-Gruber combination would result in a system where sensors (as taught by Gruber and/or Lim) are incorporated in Lim's first device to capture sensor data (*e.g.*, temperature, location) to be transmitted to Lim's server to supplement or improve audio processing as taught by Lim and/or Gruber. Lim discloses a system that can capture sensor data (*e.g.*, audio, location information) using sensors (microphones, GPS) and transmit the captured sensor data to the server. Gruber discloses the benefits of supplementing speech processing with captured sensor data and/or related information. "This helps to clarify the user's intent and to reduce the number of candidate interpretations of the user's input, and reduces the need for the user to provide excessive clarification input." (EX1009, [0013].) For example, the Lim-Gruber combination would comprise sensors for capturing sensor data in the first device to be transmitted to Lim's server via the WLAN over the Internet to supplement processing. Lim's server (*e.g.*, remote server 500) would then use the captured sensor data as contextual information to supplement speech processing (*e.g.*, in FIGs. 4 and 6 of Lim) as taught by Gruber, and transmit the speech processing result(s) (*e.g.*, first message) back to the first device.

Combining the teachings of Lim and Gruber as such would yield predictable results of improving audio processing.

217. Incorporating Gruber's speech processing techniques into Lim yields the predictable result of improving accuracy of human voice data processing and improving the usability of the control system, because the system is more intelligent and the users are not required to provide excessive clarification input. Because both Lim and Gruber relate to processing of human voice data to provide services, a POSITA would be motivated to use the known techniques of Gruber to improve the speech processing capabilities of Lim in the same way. Further, it would have been obvious to try combining the teachings of Lim and Gruber. There are only a finite number of identified and predictable solutions for processing voice data and improving accuracy of the processing, such as capturing higher-fidelity audio data or drawing from non-audio cues. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Gruber has successfully demonstrated that device sensors can gather contextual information that would be useful in supplementing speech processing, and thus, there would have been a reasonable expectation of success for combining Lim and Gruber.

218. Thus, it would have been obvious to a POSITA to make use of device sensor data to process human speech without requiring excessive clarifications from users. The known benefits of Gruber would prompt the application of Gruber in the same field based on design incentives and other market forces.

3. Claim 4: "The first device according to claim 3, wherein the sensor is a thermoelectric sensor that responds to a temperature or to a temperature gradient of an object using conduction, convection, or radiation, or wherein the sensor is a photoelectric sensor that responds to a visible or an invisible light or gamma rays."

219. Claim 4 depends on claim 3. As discussed above, Gruber discloses various sensors including "an accelerometer, compass, GPS unit, altitude detector, light sensor, thermometer, barometer, clock, network interface, battery test circuitry, and the like." (EX1009, [0225].) The contextual information from these sensors can include "the current light level, temperature[,] and other environmental measures." (EX1009, [0222].) Thus, Gruber expressly discloses thermoelectric sensor in the device that responds to temperature as well as a photoelectric sensor that responds to light. A POSITA would have understood that photoelectric sensors are common light sensors and thermoelectric sensors are common temperature sensors. Light sensors (*i.e.*, light detectors) conducts optical detection: "[t]he process of optical detection involves the direct conversion of optical energy (in the form of photons) into an

electrical signal (in the form of electrons). (EX1030, 466.) "[T]hermometer[] . . . contacts the object or receives its electromagnetic radiation, then produces a physical response, or signal. That signal becomes a measure of the object's temperature." (EX1030, 116.)

220. It would have been obvious to a POSITA to combine the teachings of Lim and Gruber. (*Supra* § IX.C.2.)

4. Claim 11: "The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type."

221. Gruber discloses that its devices provide "signatures of connected networks, including Ethernet, Wi-Fi and Bluetooth. ***Signatures include MAC addresses of network access points, IP addresses assigned, device identifiers such as Bluetooth names, frequency channels and other properties of wireless networks.***" (EX1009, [0218]-[0225].) "Most MAC addresses use one of the IEEE standardized protocols; ***MAC-48, EUI-48®, and EUI-64®***." (EX1026, 2:50-59.)

222. It would have been obvious to a POSITA to combine the teachings of Lim and Gruber and look to Gruber for implementation details for Lim's networked device and addressing schemes (*e.g.*, MAC addresses) with a reasonable expectation of success. (*Supra* § IX.C.2.) Such combination would be obvious as Lim itself discloses that the controlling device is

addressable in the Internet and WLAN and that it communicates via WLAN and with remote server over the Internet via a wireless router and thus would suggest that Lim's networked devices can be associated with MAC addresses. (*Supra* §§ IX.A.2.a, IX.A.2.j; EX1005, [0027], [0038], FIG. 5b.) Incorporating a NIC with a unique MAC address into Lim's device would have been a routine and predictable design choice, because MAC addressing is a fundamental and ubiquitous feature of IEEE 802 networking and is required for ordinary device participation in standard WLANs.

223. A POSITA also would have been motivated by strong market incentives and product design requirements to include a NIC with a unique MAC address in Lim's device to ensure compatibility with the vast installed base of existing home routers, access points, and network infrastructure. Wireless consumer devices are expected to interoperate with standard Wi-Fi networks and MAC addressing is a universally adopted mechanism for network identification, association, filtering, and access control in such environments. Incorporating a unique MAC address would therefore predictably improve interoperability, simplify deployment, and allow the Lim device to function seamlessly with existing networks. These design incentives and market forces would have prompted a POSITA to apply the conventional MAC address

teachings to Lim to meet consumer expectations for wireless connectivity and broad network compatibility.

D. Ground 4: Lim and Benesty Render Obvious Claims 7-8

1. Summary of Benesty

224. Benesty was published by at least December 11, 2007, and is prior art under § 102(a), (b). (EX1039, 1.) The Examiner never found, and Applicant never cited, Benesty during prosecution. Benesty is a textbook titled "Springer Handbook of Speech Processing." (EX1010, Title.) Benesty provides an overview of well-known speech processing techniques, including speech recognition, speaker recognition, speech enhancement via noise reduction, and microphone arrays.

225. Benesty discloses that "[i]t is shown that microphone arrays have the ability to offer directional gains that can significantly improve the quality of signal pickup in reverberant and noisy environments." (EX1010, 1021.)

226. Benesty discloses that a microphone array can include multiple microphones arranged in an array. The schematic diagram FIG. 50.1 of Benesty illustrated a uniformly spaced linear microphone array of ***microphones*** spaced by a distance *d*.

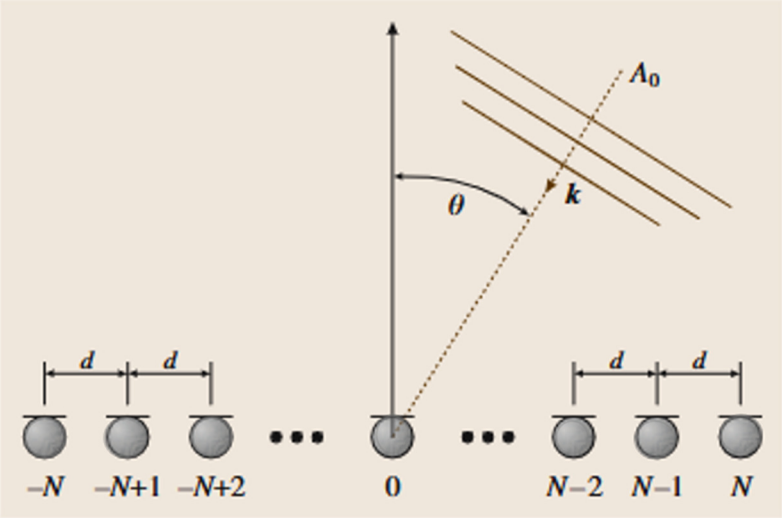


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

227. Benesty discloses that "for speech source localization, microphone array beam scanning and time delay estimation (TDE)-based localization methods are the two most widely used approaches. A beamformer is a spatial filter that operates on the microphone outputs in order to enhance the signal coming from one direction while suppressing noise and interference from other directions. Therefore, steering a microphone array beamformer and scanning across a room for the highest-energy output leads to an ***estimate of the direction*** of an active speech source. Using two arrays and by intersecting the two corresponding direction estimates, the source location is determined." (EX1010, 1044.)

228. Benesty is a textbook published by Springer and was well-known among POSITAs. A POSITA would have looked to Benesty for techniques of

applying multiple microphones arranged in a directional array in a device to, for example, improve the audio signal capturing and processing.

2. Claim 7: "The first device according to claim 1, further comprising an additional microphone for capturing human voice data, and wherein the capturing comprises capturing, by the additional microphone, the first and second human voice data."

229. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2) Claims 7-8 are obvious in further view of Benesty. For example, "[i]t is shown that ***microphone arrays*** have the ability to offer directional gains that can ***significantly improve the quality of signal pickup in reverberant and noisy environments***." (EX1010, 1021.) "One solution is to use beamforming microphone arrays, which can offer significant directional gain so as to result in similar audio performance to that of closely placed microphones." (EX1010, 1021.)

230. A microphone array includes multiple microphones arranged in an array, as shown below. The schematic diagram FIG. 50.1 of Benesty illustrates a uniformly spaced linear microphone array by ***microphones*** spaced by a distance *d*.

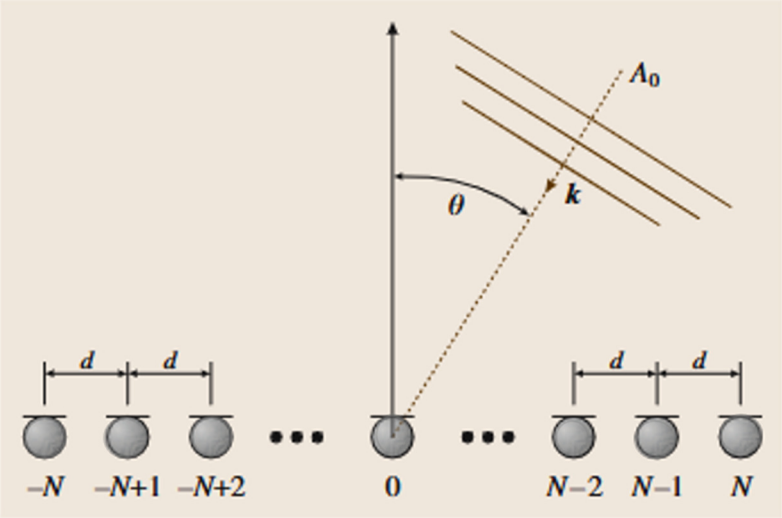


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

231. It would have been obvious to a POSITA to apply Benesty's techniques of including additional microphones to Lim's first device for capturing the first and second human voice data. Both Lim and Benesty are related to capturing and processing human audio data. Lim is directed to hands-free audio communication of a user with a first device. For example, Lim discloses that "a ***controlling device*** may be provisioned with means to receive voice input from a user, e.g., **a *microphone***[.]" (EX1005, [0007].) Benesty discloses hands-free audio communication and speech processing techniques. Benesty states that "[h]ands-free audio communication is now a major feature in mobile communication systems as well as audio and video conferencing systems. One problem that becomes evident to users of these systems is the decrease in communication quality due to the pickup of room

reverberation and background noise." (EX1010, 1021.) Thus, Lim and Benesty are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problem of providing hands-free audio communication between users and devices.

232. The Lim-Benesty combination would result in a system where multiple microphones as taught by Benesty are incorporated in Lim's first device to capture human voice data to improve the quality of captured audio. For example, Benesty's microphone array(s) would be included in Lim's controlling device 100 or smart device 110 to capture the first and second human voice data. Because Lim and Benesty relates to audio capturing and processing, a POSITA would be motivated to use the known techniques of Benesty to improve Lim's ability to capture and process audio. Even in use cases where background noise is desired, a microphone array can capture richer audio information, enabling more efficient and accurate processing, for example, separating or balancing speech and background noise when desired.

233. Combining the teachings of Lim and Benesty as such would yield predictable results of improved audio capturing and processing. Further, it would have been obvious to try combining the technical features of Benesty with the teachings of Lim, as there is a finite number of identified, predictable resolutions, without undue effort and with a reasonable expectation of success.

The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of potential solutions to improve audio signal pickup, such as additional software processing or incorporating additional hardware components. The known benefits of Benesty would prompt the application of Benesty in the same field based on design incentives and other market forces to produce a system architecture that responds to captured human voice data accurately, having improved audio communication signal pickup.

3. Claim 8: "The first device according to claim 7, wherein the microphone and the additional microphone are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array."

234. Benesty discloses that "for speech source localization, microphone array beam scanning and time delay estimation (TDE)-based localization methods are the two most widely used approaches. A beamformer is a spatial filter that operates on the microphone outputs in order to enhance the signal coming from one direction while suppressing noise and interference from other directions. Therefore, steering a microphone array beamformer and scanning across a room for the highest-energy output leads to an ***estimate of the direction*** of an active speech source. Using two arrays and by intersecting the two

corresponding direction estimates, the source location is determined." (EX1010, 1044.)

235. Benesty also discloses that "[a]lternatively, a TDE-based source localization algorithm involves a two-step procedure. In the first step, a set of relative time differences of arrival (TDOAs) for different microphone pairs are calculated. In the second step, the acoustic source location is estimated from the TDOAs and using the a priori knowledge about the locations of the microphones. TDE-based source localization algorithms have many merits: they cope well with both narrowband and broadband source signals; their localization resolution can be flexibly adjusted by varying the sampling rate and microphone array size; the effect of room reverberation is only of concern in the first step TDE and several recently developed robust TDE algorithms are promising for future practical use; and in general they are computationally efficient." (EX1010, 1044.)

236. It would have been obvious to a POSITA to combine the teachings of Lim and Benesty. (*Supra* § IX.D.2.)

E. Ground 5: Lim and Rossing Render Obvious Claims 9, 29, 33

1. Summary of Rossing

237. Rossing was published by at least July 23, 2007, and is prior art under § 102(a), (b). (EX1041, 1.) The Examiner never found, and Applicant

never cited, Rossing during prosecution. Rossing is a textbook titled "Springer Handbook of Acoustics." (EX1014, Title.) Rossing discloses how conventional microphones and speakers operate.

238. Rossing discloses how a conventional microphone operates. Rossing states that "[c]ommon to virtually all microphones ever made is a diaphragm. Most microphones do not directly sense the instantaneous pressure deviations of an airborne acoustic wave. Instead, the ***sound waves impinging on a diaphragm cause movement of the diaphragm***, and it is ***this movement that is sensed and converted to an electrical audio signal***." (EX1014, 757.)

239. Rossing also discloses different types of microphones such as a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone. "Two types of microphones avoid any mechanical attachment to the diaphragm: the ***ribbon microphone*** and the ***condenser microphone***. A popular variant of the ***dynamic microphone***, the ribbon microphone, uses a thin metal ribbon diaphragm suspended in a magnetic field." (EX1014, 758.) "A popular modern variant is the electret condenser microphone, perfected in 1962 by Sessler and West of Bell Laboratories." (EX1014, 759.)

240. Rossing discloses additionally, the piezoelectric microphones and carbon microphones. (EX1014, 758.)

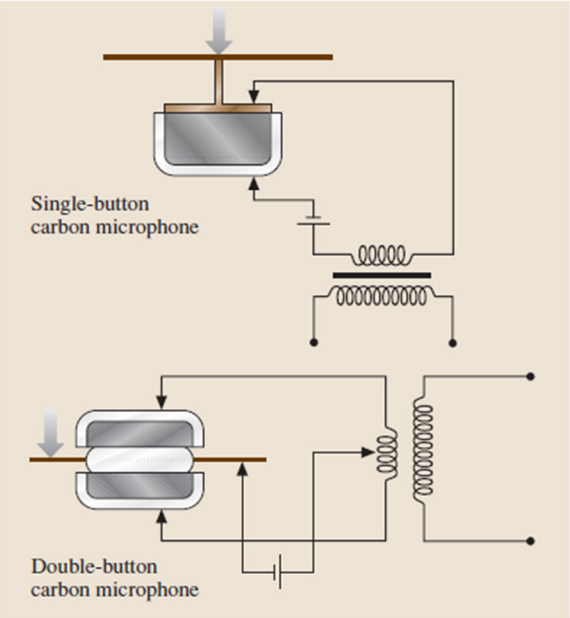


Fig. 18.11 A carbon microphone (After [18.25])

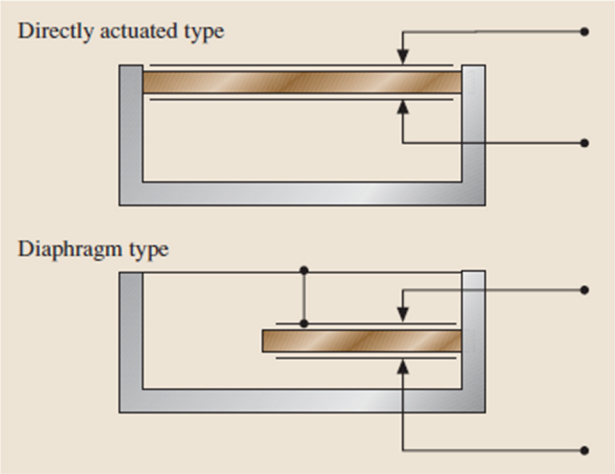


Fig. 18.12 Piezoelectric microphones (After [18.29])

241. Rossing also discloses how a conventional speaker operates. Rossing discloses that "[i]n a manner reminiscent of microphones, few loudspeakers convert directly between electrical and acoustical energy. Instead, the ***electrical energy is converted to mechanical energy, specifically the movement of a diaphragm***, often cone-shaped or dome-shaped, which then radiates acoustical energy, sometimes aided by an acoustical coupling horn or lens." (EX1014, 763.) "In the low-frequency range, the ***linear motion of an electrodynamical loudspeaker***, for example, can be described by a single-degree-of-freedom oscillator." (EX1014, 903.)

242. Rossing discloses that the soundwaves outputted by speakers are omnidirectional. "***Sound radiates from a loudspeaker in all directions***, depending on angle and frequency," and that "[i]t is rare for a speaker to have

truly constant directionality across its entire pass-band, in part from that fact that most are at least somewhat directional at mid and high frequencies, and, because of the long wavelengths involved, almost unavoidably ***omni-directional at low frequencies***." (EX1014, 765.)

243. Further, Rossing discloses that there can be different types of speakers including electromagnetic, piezoelectric, and electrostatic. "[M]any of the same transduction methods that have been used in microphones have at least been tried for loudspeakers including ***magnetic (also known as dynamic), piezoelectric***, and condenser." (EX1014, 763.)

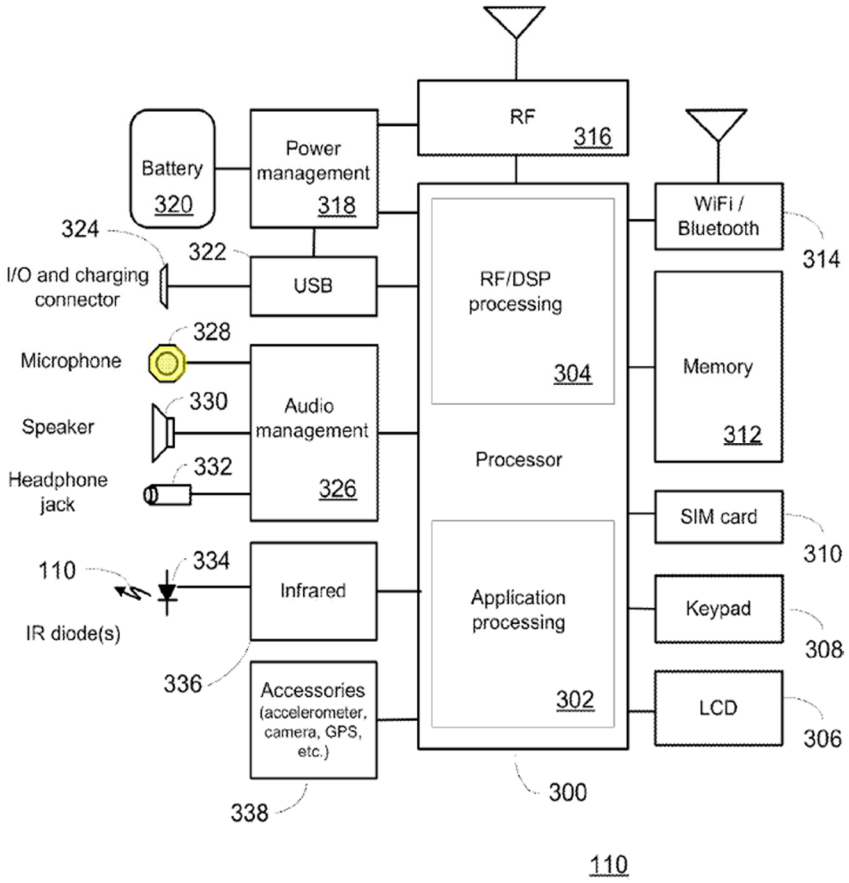
244. Rossing is a textbook published by Springer and was well-known among POSITAs. A POSITA would have looked to Rossing for detailed specifics and techniques of using microphones and speakers in the system disclosed in Lim.

2. Claim 9: "The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

245. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2) Claims 9, 29, 33 are obvious in further view of Rossing.

246. A POSITA would understand that the claim refers to a standard microphone. (*Supra* § IX.A.5; EX1001, 17:47-54.) Lim discloses that the

controlling device has a ***microphone***. For example, "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include . . . speech input means comprising ***a microphone 224*** and associated circuitry 222 for example an analog to digital converter, speech digitizer, etc., as known in the art." (EX1005, [0019].) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated ***microphone 328***, speaker 330, and headphone jack 332[.]" (EX1005, [0023].)



247. Rossing discloses that a microphone includes a diaphragm and operates by sensing movement of the diaphragm. Rossing states that "[c]ommon to virtually all microphones ever made is a diaphragm. Most microphones do not directly sense the instantaneous pressure deviations of an airborne acoustic wave. Instead, the sound waves impinging on a diaphragm cause movement of the diaphragm, and it is this movement that is sensed and converted to an electrical audio signal." (EX1014, 757.)

248. Rossing also discloses other variety of microphones such as the ribbon microphone or the condenser microphone. "Two types of microphones avoid any mechanical attachment to the diaphragm: the ribbon microphone and the condenser microphone. A popular variant of the dynamic microphone, the ribbon microphone, uses a thin metal ribbon diaphragm suspended in a magnetic field. Movement of the diaphragm induces a current in the diaphragm, which can then be amplified. In effect, the diaphragm acts as a single-turn moving coil." (EX1014, 758.) "The condenser microphone is based on the notion of electrical capacitance. A metal, or other conducting, diaphragm is placed close to a second, fixed conducting plate, forming a capacitor, and a polarizing voltage is applied between the plates. Movement of the diaphragm causes alterations in the capacitance, in turn causing corresponding alteration of the charge on the

plates, which is converted to a dynamic voltage by a resistor, and amplified as necessary." (EX1014, 758.)

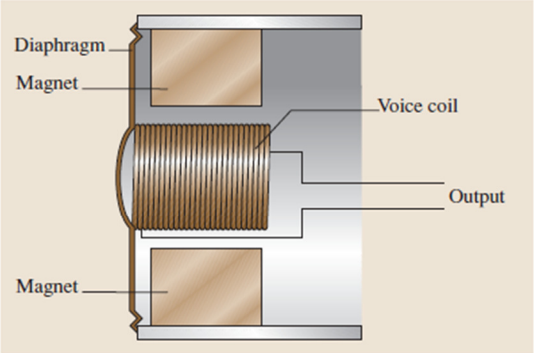


Fig. 18.10 A dynamic microphone uses a coil of wire attached to a diaphragm, suspended in a magnetic field (After [18.24])

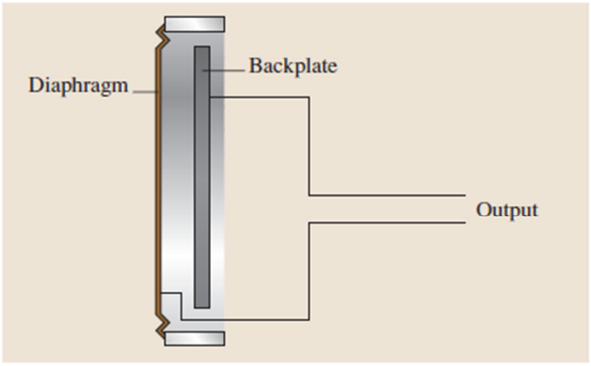


Fig. 18.9 A condenser microphone consists of a very thin diaphragm suspended parallel to a backplate (After [18.24])

249. Rossing discloses piezoelectric microphones and carbon microphones. "Many early telephone transmitters used carbon in one form or another, typically carbon granules coupled to the diaphragm such that vibration of the diaphragm would dynamically squeeze the granules, causing a variation in resistance that could be electrically sensed by an applied voltage. This arrangement could be scaled up in size and voltage enough to produce a sizeable signal, of particular importance prior to the advent of electrical amplification." (EX1014, 758.) "Another type of self-generating microphone is the crystal microphone, based on the piezoelectric effect discovered by the Curies in 1880. Here, the diaphragm constricts a thin crystal made of Rochelle salt, ceramic, or similar material, directly producing a voltage. The low compliance of the

piezoelectric element typically limits the performance of such a microphone, especially the extent of its frequency response." (EX1014, 758.)

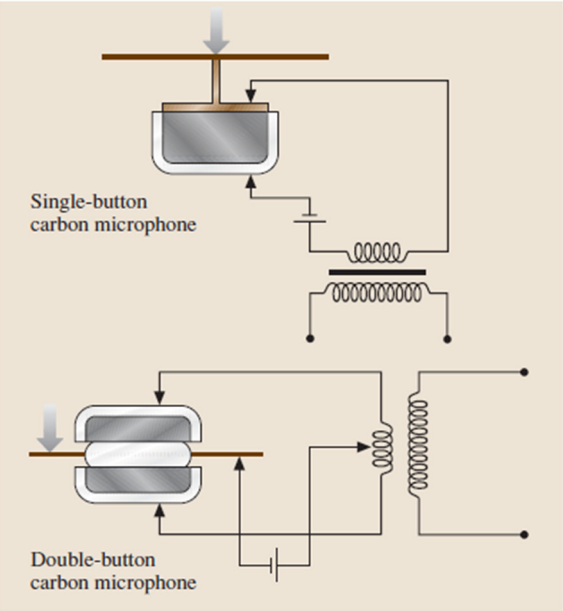


Fig. 18.11 A carbon microphone (After [18.25])

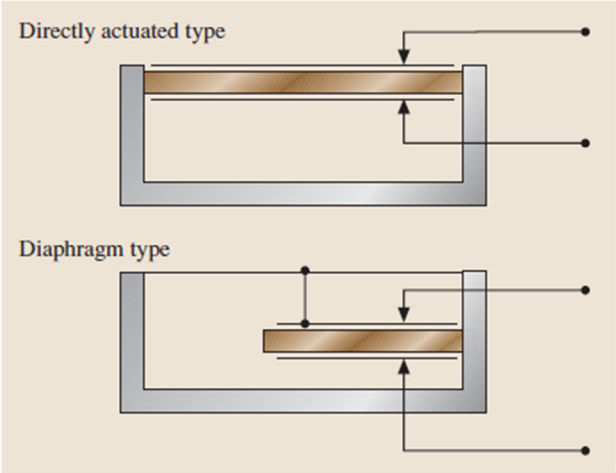


Fig. 18.12 Piezoelectric microphones (After [18.29])

250. Further, Rossing discloses an electret microphone. "A popular modern variant is the electret condenser microphone, perfected in 1962 by Sessler and West of Bell Laboratories. An electret is an object possessing a permanent electrostatic charge, making it rather like the electrical analog of the permanent magnet. Use of an electret to polarize the capacitor in a condenser microphone eliminates the need to supply a polarizing voltage for the capacitor. Since electret condenser microphones are used almost universally in cell phones, as well as other applications, they are probably the most popular type in current use. The diaphragm of electret condenser microphones is usually made of a thin sheet of flexible plastic which has had a metal coating deposited on it, a notably

lightweight arrangement that permits good sensitivity and frequency response. Such microphones also usually exhibit good sample-to-sample consistency, making them practical for multi-microphone arrays." (EX1014, 759.)

251. Rossing also discloses that omnidirectional microphones are conventional. Rossing discloses that "[f]or measurement of the lateral energy fraction (LEF), the early part (up to 80 ms) of this lateral sound energy is compared with the energy of the direct sound plus all early reflections picked up by an ordinary omnidirectional microphone where h_1 is the impulse response pressure recorded with a figure-of-eight microphone, whereas h is captured through the (usual) omnidirectional microphone." (EX1014, 309.)

252. Both Lim and Rossing are related to operating microphone(s) and speaker(s). (EX1005, [0017], [0019], [0023], [0027]-[0029], [0035]-[0036].) As discussed, Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1014, 757-758, 763, 765, 903.) Thus, Lim and Rossing are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problems of operating microphone(s) and speaker(s). Thus, a POSITA would have been motivated to look to Rossing for techniques and details regarding these components for implementing the system described in Lim.

253. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Lim, as various microphone features can improve Lim's microphone in the same way and yield predictable results. The known benefits of various features would prompt the application of Rossing in the same field based on design incentives and market forces. Further, it would have been obvious to try combining Rossing with the teachings of Lim, for example by using an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone in the first device to capture the first and second human data, without undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of types of microphones, such as a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone, that were commonly known in the industry and easy to manufacture and integrate into Lim.

3. Claim 29: "The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

254. A POSITA would have understood that this claim refers to a speaker. (*Supra* § IX.A.20; EX1001, 105:48-54, 105:61-106:4.) As discussed, Lim teaches operating an actuator (*e.g.*, ***speaker or speaker 303***) consistent with the descriptions of the '721 patent in the first device (*e.g.*, ***controlling device 100 or smart device 110***) in response to the received first message (*e.g.*, ***information*** such as recognized speech, analysis of or information associated with the recognized speech, and/or additional prompts). (*Supra* §§ IX.A.2.d, IX.A.2.i; EX1001, 105:48-52, 105:61-106:4; EX1005, [0027]-[0029].)

255. Rossing discloses that there are different types of speaker at the time of the invention, such as an electromagnetic speaker, a piezoelectric speaker, an electrostatic loudspeaker, or a magnetic loudspeaker. "[M]any of the same transduction methods that have been used in microphones have at least been tried for loudspeakers including ***magnetic (also known as dynamic), piezoelectric, and condenser***." (EX1014, 763.) "In 1929, E. W. Kellogg filed a patent for an ***electrostatic loudspeaker***, consisting of panels not unlike large condenser microphones." (EX1014, 763.) Rossing also discloses using "***small loudspeaker units*** (***piezoelectric or electro magnetic***) and the microphones 1/4" or even 1/8" condenser types." (EX1014, 321-322.)

256. As discussed, Lim and Rossing are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problems of operating microphone(s) and speaker(s). (*Supra* § IX.E.2.) Thus, a POSITA would have been motivated to look to Rossing for techniques and details related to speakers for implementing Lim. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional speakers to Lim' speaker, as various speaker features can improve Lim's speaker in the same way and yield predictable results. The known benefits of various features would prompt the application of Rossing in the same field based on design incentives and other market forces.

257. Further, it would have been obvious to try combining the technical features of Rossing with the teachings of Lim, for example, by using an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker in the first device to output audio reply, without undue effort and with a reasonable expectation of success. (*Supra* § IX.E.2.) There was a finite number of types of speakers, such as an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, that were commonly seen in the industry and easy to manufacture and integrate into the Lim system. (EX1014, 321-322, 763.)

4. Claim 33: "The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves."

258. A POSITA would have understood that this claim refers to a speaker. (*Supra* § IX.A.23; EX1001, 105:58-106:4.) As discussed above, Lim teaches operating a ***speaker or speaker 303*** in the first device. (*Supra* §§ IX.A.2.d, IX.A.2.i; EX1005, [0037].) As discussed above, at the time of the invention, a POSITA would have understood that a speaker is a sounder for converting an electrical energy to sound waves. (*Supra* § IX.A.2.d; EX1018, 122; EX1014, 763.)

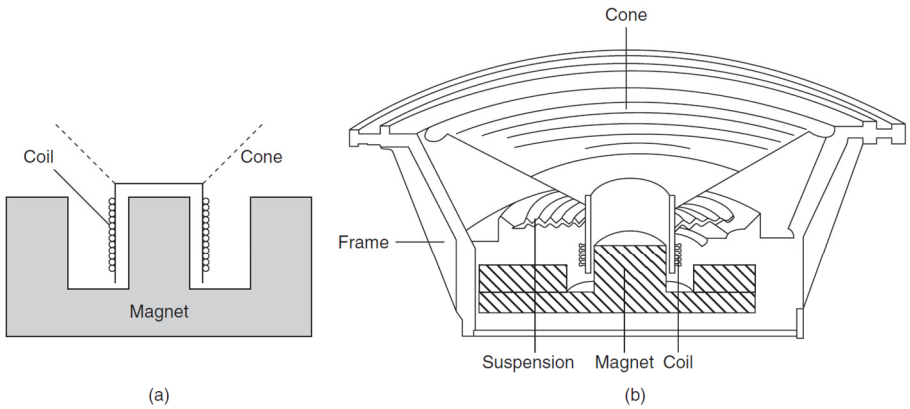


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

259. Further, in light of Rossing, a POSITA would have understood that the speaker operates by emitting audible sound waves that are of an omnidirectional, unidirectional, or bidirectional pattern. For example, "***[s]ound radiates from a loudspeaker in all directions***, depending on angle and

frequency. . . . It is rare for a speaker to have truly constant directionality across its entire pass-band, in part from that fact that most are at least somewhat directional at mid and high frequencies, and, because of the long wavelengths involved, almost unavoidably ***omni-directional at low frequencies***." (EX1014, 765.) It would have been obvious to combine Lim and Rossing. (*Supra* § IX.E.2.)

F. Ground 6: Lim and Goldsmith Render Obvious Claims 13-15

1. Summary of Goldsmith

260. Goldsmith was published by at least January 20, 2006, and is prior art under § 102(a), (b). (EX1040, 2.) The Examiner never found, and Applicant never cited, Goldsmith during prosecution. Goldsmith is a textbook titled "Wireless Communications." (EX1040, Title.) Goldsmith provides an overview of wireless communication techniques, including the wireless systems available at the time of publication in 2005.

261. Goldsmith discloses that the wireless systems available in 2005 includes, cellular telephone systems, cordless phones, wireless LANs, wide area wireless data services, broadband wireless access, paging systems, satellite networks, low-cost low-power radios such as Bluetooth and Zigbee, as well as ultrawideband radios. (EX1040, 8-21.)

262. Goldsmith discloses Bluetooth and Zigbee as two radio specifications based on the IEEE 802.15.4 standard that emerged as "it becomes feasible to embed them in more types of electronic devices, which can be used to create smart homes." (EX1040, 19-20.) Goldsmith discloses the wireless LAN standard, IEEE 802.11b standard, the standard which many companies developed their products on, and that "[t]wo additional standards in the 802.11 family" including "***IEEE 802.11a wireless LAN*** standard" and "***802.11g***." (EX1040, 15-16.)

263. Goldsmith discloses that both Bluetooth, and 802.11b WLAN are ISM Band II frequencies. (EX1040, 22-23, Table 1.2.)

Table 1.2: Unlicensed U.S. spectrum allocations

Band	Frequency
ISM band I (cordless phones, 1G WLANs)	902–928 MHz
ISM band II (Bluetooth, 802.11b and 802.11g WLANs)	2.4–2.4835 GHz
ISM band III (wireless PBX)	5.725–5.85 GHz
U-NII band I (indoor systems, 802.11a WLANs)	5.15–5.25 GHz
U-NII band II (short-range outdoor systems, 80211a WLANs)	5.25–5.35 GHz
U-NII band III (long-range outdoor systems, 80211a WLANs)	5.725–5.825 GHz

264. Further, Goldsmith discloses the 3G cellular system, GSM (Global Systems for Mobile Communications), and EDGE (Enhanced Data Services for GSM Evolution). "***The GSM standard*** uses a combination of TDMA and slow frequency hopping with frequency-shift keying for the voice modulation."

(EX1040, 12.) "The ***third generation (3G) cellular systems are based on a wideband CDMA standard*** . . . ." (EX1040, 13.) "***Enhanced Data Services for GSM Evolution (EDGE)***, further increases data rates . . . using a high-level modulation format combined with [FEC] coding." (EX1040, 13.) "***W-CDMA*** is the primary competing 3G standard to cdma2000. It has been selected as the ***3G successor to GSM***, and in this context is referred to as the ***Universal Mobile Telecommunications System (UMTS)***." (EX1040, 599-600, Table D.4.)

Table D.4: 3G digital cellular phone standards

	cdma2000				W-CDMA		
Parameter	1X	1XEV-DO	1XEV-DV	3X	UMTS	FOMA	J-phone
Channel bandwidth (MHz)	1.25	1.25	1.25	3.75	5	5	5
Chip rate (Mchips/s)	1.2288	1.2288	1.2288	3.6864	3.84	3.84	3.84
Peak data rate (Mbps)	.144	2.4	4.8	5–8	2.4[a]	2.4[a]	2.4[a]
Modulation	QPSK/MPSK (DL), BPSK/QPSK (UL)				QPSK (DL), BPSK (UL)		
Coding	Convolutional (low rate), Turbo (high rate)						
Power control (Hz)	800	800	800	800	1500	1500	1500

[a] 8–10 Mbps with HSDPA.

265. Goldsmith is a textbook published by Cambridge University Press and authored by Andrea Goldsmith, a professor of Electrical Engineering at the Stanford University. Goldsmith was well-known among POSITAs. A POSITA would have looked to Goldsmith for wireless communication systems, protocols, and techniques of applying them.

2. Claim 13: "The first device according to claim 1, wherein the WLAN is according to, or base on, Institute of Electrical and

Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac."

266. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2) Claims 13-15 are obvious in further view of Goldsmith.

267. As discussed, Lim discloses or suggests that its WLAN is according to the IEEE802.11 standards. (*Supra* § IX.A.9.)

268. Further, Goldsmith introduces various IEEE802.11 standards. Goldsmith discloses Wireless Local Area Network (WLAN) based on IEEE 802.11b as one of the common wireless network communication protocols well before the priority date .

269. Goldsmith discloses that wireless LAN was well-known and widely adopted. "Wireless LANs provide high-speed data transmissions within a small region (e.g., a campus or small building) as users move from place to place." (EX1040, 15.) "The second-generation of wireless LANs in the United States operate with 83.5 MHz of spectrum in the 2.4 GHz ISM band. A wireless LAN standard for this frequency band, the ***IEEE 802.11b standard***, was developed to avoid some of the problems with the proprietary first-generation systems. The standard specifies direct-sequence spread spectrum with data rates of around 1.6 Mbps (raw data rates of 11 Mbps) and a range of approximately 100m. The network architecture can be either star or peer-to-peer,

although the peer-to-peer feature is rarely used. ***Many companies developed products based on the 802.11b standard, and after slow initial growth the popularity of 802.11b wireless LANs has expanded considerably***. ***Many laptops come with integrated 802.11b wireless LAN cards. Companies and universities have installed 802.11b base stations throughout their locations, and many coffee houses, airports, and hotels offer wireless access***, often for free, to increase their appeal." (EX1040, 16.)

270. Goldsmith discloses a WLAN based on IEEE 802.11a and 802.11g. "Two additional standards in the 802.11 family were developed to provide higher data rates than 802.11b. The ***IEEE 802.11a wireless LAN*** standard operates with 300 MHz of spectrum in the 5 GHz U-NII band. The 802.11a standard is based on multicarrier modulation and provides 54-Mbps data rates . . . . Because 802.11a has much more bandwidth and consequently many more channels than 802.11b, it can support more users at higher data rates. There was some initial concern that 802.11a systems would be significantly more expensive than 802.11b systems, but in fact they quickly became quite competitive in price." (EX1040, 16.) In addition, Goldsmith also discloses "[t]he other standard, ***802.11g***" and that "[m]any wireless LAN cards and access points support all three standards to avoid incompatibilities." (EX1040, 16.)

271. It would have been obvious for a POSITA to apply Goldsmith's disclosure of various types of wireless networking techniques and protocols to Lim. Both Lim and Goldsmith are related to wireless communication technologies. Lim itself suggests a motivation to combine. As discussed, Lim discloses that the first device communicates with the ***remote server 500*** (*i.e.*, provisioning the speech recognition result and/or additional prompts for voice input and/or indicated command code data to the first device) over the ***Internet*** (*i.e.*, network 502), and ***via WiFi***, a form of WLAN, through a gateway device such as a ***router***. (*Supra* §§ IX.A.2.a, IX.A.2.g, IX.A.2.h; EX1005, [0027], FIG. 5b.) Lim discloses that access to "a remote server 500 accessible via a network 502 such as the Internet" "may be via an intermediary appliance such as STB 104, illustrated at 506, or may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].) Goldsmith discloses various existing wireless connection systems in a chapter titled "1.4 Current Wireless Systems" (EX1040, 8-21.) Thus, Lim and Goldsmith are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problem of providing wireless communications between devices.

272. A POSITA would be motivated to use the known wireless networking techniques of WLAN over IEEE 802.11b, 802.11a, and/or 802.11g

to improve the usability of Lim in the same way. Combining the teachings of Lim and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

273. A POSITA would have looked to Goldsmith's disclosure of WLAN based on IEEE 802.11b, IEEE 802.11a, and IEEE 802.11g as conventional wireless networking techniques and protocols that can be applied to Lim because Lim expressly discloses usage of WiFi, a form of WLAN. Further, Goldsmith explains the popularity of 802.11b wireless LANs and the wide coverage of locations with 802.11b base stations installed. (EX1040, 16.) Goldsmith discloses that 802.11a can support more users at higher data rates, and 802.11g can be used in either the 2.4 GHz and 5 GHz bands. Goldsmith discloses that "[m]any wireless LAN cards and access points support all three standards to avoid incompatibilities." (EX1040, 16.)

274. Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (*e.g.*, WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g) with the teachings of Lim, without undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of potential wireless network communications, such as infrared, BlueTooth,

Zigbee, 3G, 4G, UWB, and GSM, that were commonly known in the industry and were easy to apply to Lim. The known benefits and feasibility of including a well-known wireless networking technique such as WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g as taught by Goldsmith to connect in-home devices with a remote server would prompt the application of Goldsmith in the same field based on design incentives and other market forces.

3. Claim 14: "The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

275. Goldsmith summarizes in the table below the unlicensed radio frequency spectrum allocations in the U.S., which expressly shows that 802.11b WLANs are ISM Band II frequencies. (EX1040, 22-23, Table 1.2.)

Table 1.2: Unlicensed U.S. spectrum allocations

Band	Frequency
ISM band I (cordless phones, 1G WLANs)	902–928 MHz
ISM band II (Bluetooth, 802.11b and 802.11g WLANs)	2.4–2.4835 GHz
ISM band III (wireless PBX)	5.725–5.85 GHz
U-NII band I (indoor systems, 802.11a WLANs)	5.15–5.25 GHz
U-NII band II (short-range outdoor systems, 80211a WLANs)	5.25–5.35 GHz
U-NII band III (long-range outdoor systems, 80211a WLANs)	5.725–5.825 GHz

276. As discussed above, a POSITA would be motivated to combine the teachings of Goldsmith and Lim and apply the well-known wireless

communication techniques such as WLAN over 802.11b as taught in Goldsmith to Lim. (*Supra* §§ IX.A.2.a, IX.F.2.)

4. Claim 15: "The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution."

277. Goldsmith discloses 3G, the third-generation cellular network, GSM (Global System for Mobile Communications), and EDGE (Enhanced Data rates for GSM Evolution) as conventional wireless network communication protocols at the time of the invention.

278. Goldsmith discloses that "[t]he ***third-generation (3G) cellular systems are based on a wideband CDMA standard*** developed under the auspices of the International Telecommunications Union (ITU). The standard, initially called International Mobile Telecommunications 2000 (IMT-2000), provides different data rates depending on mobility and location, from 384 kbps for pedestrian use to 144 kbps for vehicular use to 2 Mbps for indoor office use. The 3G standard is incompatible with 2G systems, so service providers must invest in a new infrastructure before they can provide 3G service." (EX1040, 13.)

279. Further, Goldsmith discloses the details of GSM (Global Systems for Mobile Communications). "Many of the first-generation cellular systems in Europe were incompatible, and the Europeans quickly converged on a uniform standard for second-generation (2G) digital systems called GSM. ***The GSM standard*** uses a combination of TDMA and slow frequency hopping with frequency-shift keying for the voice modulation." (EX1040, 12.)

280. Goldsmith discloses an enhancement to GSM called EDGE, which stands for "Enhanced Data Services for GSM Evolution (EDGE)." (EX1040, 12-13.)

281. In addition, Goldsmith also discloses usage of W-CDMA as a 3G standard and that it is referred to as the Universal Mobile Telecommunications System (UMTS). "***W-CDMA*** is the primary competing 3G standard to cdma2000. It has been selected as the ***3G successor to GSM***, and in this context is referred to as the ***Universal Mobile Telecommunications System (UMTS)***. W-CDMA is also used in the Japanese FOMA and J-Phone 3G systems." (EX1040, 599-600, Table D.4.)

Table D.4: 3G digital cellular phone standards

	cdma2000				W-CDMA		
Parameter	1X	1XEV-DO	1XEV-DV	3X	UMTS	FOMA	J-phone
Channel bandwidth (MHz)	1.25	1.25	1.25	3.75	5	5	5
Chip rate (Mchips/s)	1.2288	1.2288	1.2288	3.6864	3.84	3.84	3.84
Peak data rate (Mbps)	.144	2.4	4.8	5–8	2.4[a]	2.4[a]	2.4[a]
Modulation	QPSK/MPSK (DL), BPSK/QPSK (UL)				QPSK (DL), BPSK (UL)		
Coding	Convolutional (low rate), Turbo (high rate)						
Power control (Hz)	800	800	800	800	1500	1500	1500

[a] 8–10 Mbps with HSDPA.

282. At the time of the invention, home appliances with cellular connectivity were known. For example, a patent application filed on March 8, 2006 discloses "a receiver configured to receive a signal from a wide-area wireless communication system comprising a plurality of transmitters configured to redundantly broadcast one or more wireless signals," "the receiver and said processor are embedded within the appliance," "the appliance is an air conditioner or heater regulated by a thermostat, a clothes dryer, a refrigerator, a freezer, or a dishwasher," and "the one or more wireless signals is any of an ***analog cellular signal, a digital cellular signal***, a general packet radio service signal, an ***enhanced data rate for GSM evolution (EDGE) service signal***." (EX1031, Claims 14-16, 18.)

283. It would have been obvious for a POSITA to combine Lim and Goldsmith. For example, a POSITA would be motivated to look to Goldsmith for available cellular protocols such as 3G using W-CDMA (UMTS), GSM, and

EDGE to implement Lim which discloses a smart phone suggesting cellular network communications. As discussed above, Lim and Goldsmith are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problem of providing wireless communications between devices. As discussed, Lim's first/controlling device can be a wireless smartphone. (*Supra* §§IX.A.2.a, IX.A.11; EX1005, [0027].) A POSITA would have understood that smartphones have cellular connectivity for wireless communication, which includes 3G as taught by Goldsmith as it discloses conventional cellular data technologies. (EX1040, 12-13.) A POSITA would be motivated to use the conventional cellular data technologies and wireless network communications techniques such as 3G using W-CDMA (UMTS), GSM, and EDGE as taught by Goldsmith to improve the usability of Lim in the same way. (*Supra* § IX.F.2.) Combining the teachings of Lim and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system. Further, wireless technology commonly included cellular connectivity in 2012, and Goldsmith teaches that 3G, GSM, are types of cellular technologies. (EX1040, 12-13.)

284. Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (*e.g.*, 3G over W-CDMA (UMTS), GSM, and EDGE) with the teachings of Lim, without undue effort and with a

reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of potential wireless network communications, such as infrared, BlueTooth, and Wi-Fi, that were commonly known in the industry and were easy to apply to Lim. The known benefits and feasibility of including a known cellular data technology and wireless networking technique such as 3G over W-CDMA (UMTS), GSM, and EDGE as taught by Goldsmith to connect in-home devices with a remote server would prompt the application of Goldsmith in the same field based on design incentives and other market forces.

G. Ground 7: Lim and Chatterjee Render Obvious Claim 18-20, 23

1. Summary of Chatterjee

285. Chatterjee was filed on August 23, 2011, and is prior art under §102(e). (EX1012, 1.) The Examiner never found, and Applicant never cited, Chatterjee during prosecution. Chatterjee generally relates to "a system and method for providing home automation using a mobile device." (EX1012, Abstract.)

286. Chatterjee describes that "home automation systems provide centralized control for lighting, HVAC (heating, ventilation and air

conditioning), appliances (such as refrigerators and dishwashers), and similar systems." (EX1012, [0001].)

287. Chatterjee discloses that "each home appliance device 120a-120c includes a wireless transceiver 125α-125c and device controller 125a-125c for facilitating communication with the mobile device. As will be used herein, home appliance devices represent any electrical or mechanical appliance, system, or machine used to perform a household function (*e.g.*, ***oven for cooking, dishwasher for cleaning, light fixtures for illumination, refrigerator for storage, and HVAC for cooling/heating the home environment***, etc.)." (EX1012, [0012].)

288. Specifically, Chatterjee discloses that the appliances can be controlled via a device controller integrated into the appliance devices. "Device controller 125a-125c may be a device or module electronically coupled to the home appliance device 120α-120c or an integrated unit of the home appliance device 120α-120, and is configured to execute programming instructions ***and control operations on the home appliance device such as controlling the temperature of a HVAC or toggling the light system on and off***." (EX1012, [0012].)

2. Claim 18: "The first device according to claim 17, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation."

289. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2) Claims 18-20, 23 are obvious in further view of Chatterjee.

290. Lim discloses controlling various home appliances, such as "***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers***, etc." (EX1005, [0017].)

291. Chatterjee discloses that "***home automation systems*** provide centralized control for ***lighting, HVAC (heating, ventilation and air conditioning), appliances (such as refrigerators and dishwashers)***, and similar systems." (EX1012, [0001].) Chatterjee discloses that "[a]s will be used herein, home appliance devices represent any electrical or mechanical appliance, system, or machine used to perform a household function (*e.g.*, ***oven for cooking, dishwasher for cleaning, light fixtures for illumination, refrigerator for storage, and HVAC for cooling/heating the home environment***, etc.)." (EX1012, [0012].) Chatterjee discloses that the appliances can be controlled via a device controller integrated into the appliance devices. "Device controller 125a-125c may be a device or module electronically coupled to the home appliance device 120α-120c or an integrated unit of the home appliance device 120α-120, and is configured to execute programming instructions ***and control***

operations on the home appliance device such as controlling the temperature of a HVAC or toggling the light system on and off." *Id.*

292. Further, Chatterjee discloses, with reference to FIG. 2, that the controlled home appliance includes a refrigerator.

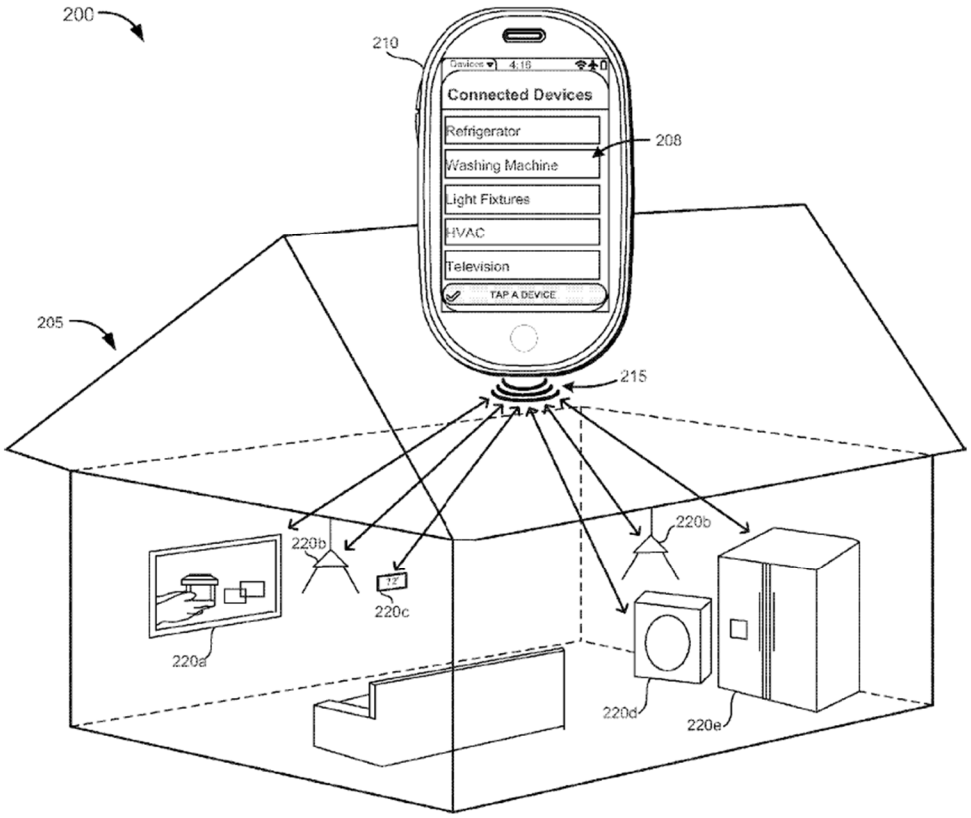


FIG. 2

293. It would have been obvious to a POSITA to combine the teachings of Lim and Chatterjee. Both Lim and Chatterjee are analogous art to the '721 patent. (EX1001, Claim 1, 135:59-63; EX1005, [0017], [0007].) All relate to a home automation system for remotely controlling home appliances. A

POSITA would understand Lim suggests other home appliances could be similarly controlled. Chatterjee discloses "a system and method for providing home automation using a mobile device." (EX1012, Abstract.) Thus, Lim and Chatterjee are from the same field of endeavor and reasonably pertinent to the same technical problems of remote control of home appliances. Combining Chatterjee with Lim would improve and expand the capabilities of Lim in the same way. Combining the teachings of Lim and Chatterjee as such would yield predictable results of enabling home automation.

294. Because Lim explicitly provides that its system can be used to control ***home automation devices***, ***HVAC, home lighting and thermostat***, respectively, and Chatterjee enumerates exemplary home automation devices and control of HVAC and home lighting systems, it would have been obvious to look to Chatterjee for additional home automation devices to be controlled. For example, the Lim-Chatterjee combination would result in a system where Lim's first device controls appliances such as oven, refrigerator, HVAC, and washing machine as taught by Chatterjee, thereby expanding the capabilities of Lim with a reasonable expectation of success.

295. Prior art such as Chatterjee has successfully demonstrated that the controlled smart appliances can be an oven, a dishwasher, light fixtures, and a refrigerator, and thus, there would have been a reasonable expectation of success

for combining Lim and Chatterjee. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Further, there are a finite number of identified and predictable home appliances to try controlling remotely. Thus, it would have been obvious to a POSITA to combine the teachings of Chatterjee and Lim to use Lim's techniques to control home appliances such as an oven, a dishwasher, light fixtures, and a refrigerator as disclosed by Chatterjee. The known benefits and feasibility of remote controlling home appliances such as oven, a dishwasher, light fixtures, and a refrigerator, as disclosed by Chatterjee would prompt the application of Chatterjee in the same field based on design incentives and other market forces.

3. Claim 19: "The first device according to claim 18, wherein the primary function of the appliance is heating food, and wherein the appliance is a microwave oven, an electric mixer, a stove, an oven, or an induction cooker."

296. As discussed, Chatterjee discloses that "***home automation systems*** provide centralized control for ***lighting, HVAC (heating, ventilation and air conditioning), appliances (such as refrigerators and dishwashers)***, and similar systems." (EX1012, [0001].)

297. Chatterjee specifically discloses that "[a]s will be used herein, home appliance devices represent any electrical or mechanical appliance,

system, or machine used to perform a household function (*e.g.*, ***oven for cooking***, dishwasher for cleaning, light fixtures for illumination, refrigerator for storage, and HVAC for cooling/heating the home environment, etc.)." (EX1012, [0012].)

298. It would have been obvious to a POSITA to combine the teachings of Lim and Chatterjee. (*Supra* § IX.G.2.)

4. Claim 20: "The first device according to claim 18, wherein the appliance is a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker."

299. As discussed, Chatterjee discloses that "***home automation systems*** provide centralized control for ***lighting, HVAC (heating, ventilation and air conditioning), appliances (such as refrigerators and dishwashers)***, and similar systems." (EX1012, [0001].)

300. Chatterjee specifically discloses that the controlled home appliance includes a refrigerator. Chatterjee discloses that "[a]s will be used herein, home appliance devices represent any electrical or mechanical appliance, system, or machine used to perform a household function (*e.g.*, oven for cooking, dishwasher for cleaning, light fixtures for illumination, ***refrigerator for storage***, and HVAC for cooling/heating the home environment, etc.)." (EX1012, [0012].)

301. Further, Chatterjee discloses, with reference to FIG. 2, that the controlled home appliance includes a refrigerator. FIG. 2 depicts, "a smart dish washer or washing machine 220d, and a ***smart refrigerator 220e***." (EX1012, [0014].)

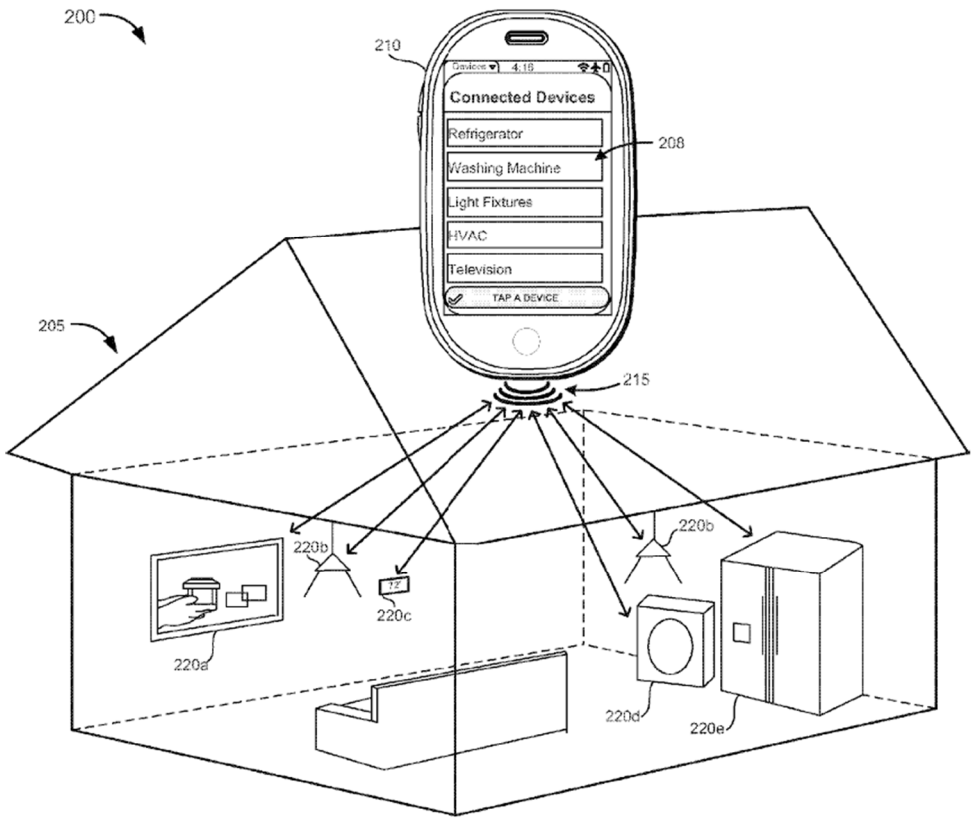


FIG. 2

302. It would have been obvious to a POSITA to combine the teachings of Lim and Chatterjee. (*Supra* § IX.G.2.)

5. Claim 23: "The first device according to claim 17, wherein the primary function of the appliance is associated with cleaning, wherein the appliance primary function is associated with clothes cleaning

and the appliance is a washing machine, or wherein the appliance is a vacuum cleaner."

303. As discussed, Chatterjee discloses that "***home automation systems*** provide ***centralized control for*** lighting, HVAC (heating, ventilation and air conditioning), ***appliances (such as refrigerators and dishwashers)***, and similar systems." (EX1012, [0001].)

304. Chatterjee specifically discloses, with reference to FIG. 2, that the controlled home appliance includes a washing machine. "In the present example of FIG. 2, ***home appliance devices*** include a television 220a, light fixtures/system 2026, HVAC 120c, a smart dish washer or ***washing machine 220d***, and a smart refrigerator 220e." (EX1012, [0014].)

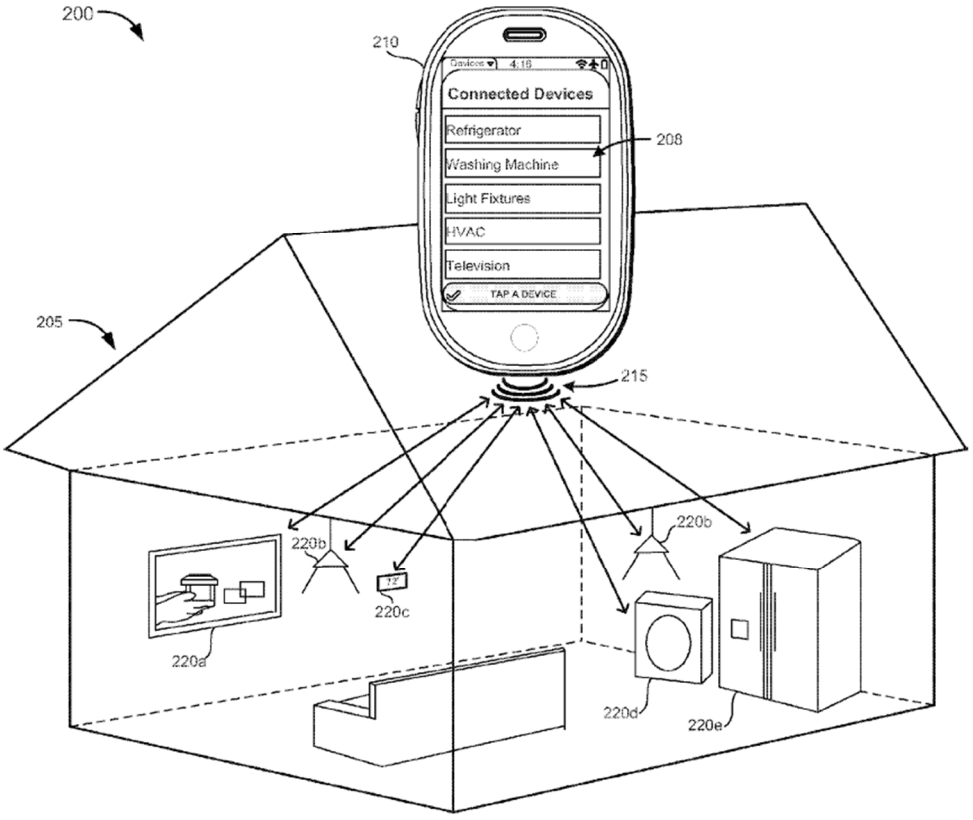


FIG. 2

305. It would have been obvious to a POSITA to combine the teachings of Lim and Chatterjee. (*Supra* § IX.G.2.)

H. Ground 8: Lim, and Rosenberger Render Obvious Claims 26-28

1. Summary of Rosenberger

306. Rosenberger was filed on October 4, 2011, and is prior art under §102(e). (EX1013, 1.) The Examiner never found, and Applicant never cited, Rosenberger during prosecution. Rosenberger relates to "improvements in

speech-recognizing control devices that respond to oral commands from a system user to produce a desired effect or result." (EX1013, 1:7-13.)

307. Rosenberger discloses that a device may include a multi-colored LED status light to visually indicate various device states and modes. . "Preferably, device D also includes … a ***multi-colored LED status light 34 for visually indicating various device states and modes***." (EX1013, 8:8-16.)

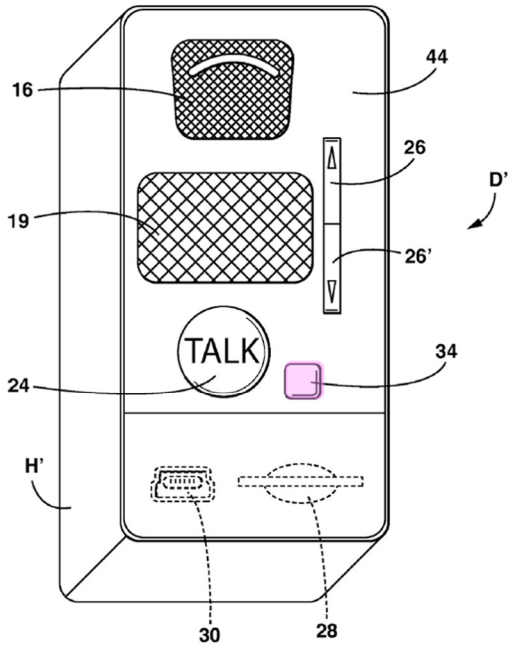


FIG. 2B

308. Rosenberger discloses that the multi-colored LED status light would serve to prompt users to provide speech input after receiving and recognizing a speech trigger phrase. "***Upon receiving and recognizing a speech trigger phrase*** and determining that it is in a better position to handle

subsequent user interaction than any other device that simultaneously recognized the same speech trigger phrase (see device coordination discussion below), or when the user depresses the 'Push to Talk' button 24 on the device, or in the event the device is instructed to wake up and expect a subsequent speech command by a system controller, the device beeps, ***changes its status light 34*** or plays a prerecorded or synthesized audio message (*e.g.* 'How may I help you?') through speaker 19 and/or 20 ***to prompt the user to say one of the speech commands from a known vocabulary or grammar stored in the speech pattern memory 14***." (EX1013, 8:20-32.)

2. Claim 26: "The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays."

309. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2.) Claims 26-27 are obvious in view of Lim or in further view of Rosenberger. (*Supra* §§IX.A.17, IX.A.18.) As discussed, Lim's ***first device*** includes ***LEDs***. (EX1005, [0019]; *supra* §§IX.A.3, IX.A.4.)

310. Further, Rosenberger discloses, with reference to FIG. 2B, a multi-colored LED status light in the device that visually indicate various device states and modes. "Preferably, device D also includes … a ***multi-colored LED status***

light 34 ***for visually indicating various device states and modes***." (EX1013, 8:8-16.)

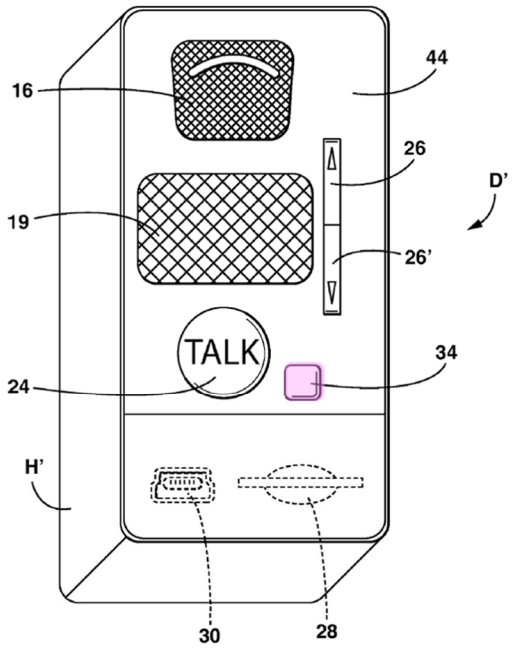


FIG. 2B

311. Specifically, Rosenberger discloses that "***[u]pon receiving and recognizing a speech trigger phrase*** and determining that it is in a better position to handle subsequent user interaction than any other device that simultaneously recognized the same speech trigger phrase (see device coordination discussion below), or when the user depresses the 'Push to Talk' button 24 on the device, or in the event the device is instructed to wake up and expect a subsequent speech command by a system controller, the device beeps, ***changes its*** *status light 34* or plays a prerecorded or synthesized audio message

(*e.g.* 'How may I help you?') through speaker 19 and/or 20 ***to prompt the user to say one of the speech commands from a known vocabulary or grammar stored in the speech pattern memory 14***." (EX1013, 8:20-32.)

312. At the time of the invention, it would have been obvious to combine the teachings of Rosenberger with Lim. Both Rosenberger and Lim relate to processing audio information such as human speech for controlling devices or providing services through devices. (*See* EX1005, [0007], [0027].) Rosenberger relates to "improvements in speech-recognizing control devices that respond to oral commands from a system user to produce a desired effect or result." (EX1013, 1:7-13.) Thus, Rosenberger and Lim are analogous to the '721 patent, from the same field of endeavor, and all relate to the same goal of processing and responding to audio commands.

313. Rosenberger discloses incorporating a multi-colored LED light in the device that responds to a speech trigger phrase received to "visually indicat[e] various device states and modes", for example by changing colors. Further, Lim itself discloses, "[t]his and subsequent prompts for voice input may take the form of an ***illuminated LED***." (EX1005, [0027]; *supra* §§IX.A.2.g, IX.A.2.h.) The technique of incorporating a light in the device that turns on/off and changes color to indicate the device status and prompt user to provide speech input can be combined with the teachings of Lim to improve the usability

of Lim' interface device, with a reasonable expectation of success as Lim itself already discloses the usage of LEDs.

314. Specifically, the Lim-Rosenberger combination results in a system that allows control of devices in response to processing of human voice data while incorporating the advantage of Rosenberger to signal or prompt users to provide voice input via its status light. For example, the Lim-Rosenberger combination would include a status light in the audio interface device, which turns on/off and/or changes color to indicate different device states and modes, as taught by Lim/Rosenberger.

315. The technique of incorporating an indicator light in the audio interface device disclosed in the Lim-Rosenberger combination yield the predictable result of improving usability of the device as the changes of the light can indicate to a user the current state or mode the device is operating in, and serves as an additional way to prompt users to provide the "first human voice data" and "second human voice data" to control the device, particularly as Lim explicitly discloses that the prompts for audio input may "take the form of an ***illuminated LED***". (EX1005, [0027].) Because both Lim and Rosenberger relates to capturing human voice data for controlling devices or providing services, a POSITA would be motivated to use the known techniques of Rosenberger to improve the usability of Lim's interface device in the same way

by enabling users to understand the shifts in device status and modes through the changes in the LED light, and to understand when the device requires additional voice input.

316. Further, it would have been obvious to try combining the teachings of Lim and Rosenberger. There are only a finite number of identified and predictable solutions for indicating to a user the device state or mode or prompt user to provide speech input, such as visual cues, auditory cues, and tactile cues. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Rosenberger has successfully demonstrated that devices may include a light for indicating device states and modes and prompting users to provide speech input, and thus, there would have been a reasonable expectation of success for combining Lim and Rosenberger. The known benefits of Rosenberger would prompt its application in the same field based on design incentives and other market forces.

3. Claim 27: "The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode."

317. As discussed, Rosenberger discloses a multi-colored LED status light in the device that visually indicates various device states and modes.

Rosenberger states that "[p]referably, device D also includes … a ***multi-colored LED status light 34*** ***for visually indicating various device states and modes***." (EX1013, 8:8-16.)

318. It would have been obvious to a POSITA to combine the teachings of Lim and Rosenberger. (*Supra* § IX.H.2.)

4. Claim 28: "The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

319. Claim 28 depends from claim 27. Besides comprising an electric light source, Lim-Rosenberger's device plays digital audio content as recited in claim 28 at least because Lim discloses claim 28 as discussed above. (*Supra* § IX.A.19.)

I. Ground 9: Williams and Lim Render Obvious Claims 1, 5-6, 9-15, 17, 21-22, 24-29, 31, 33.

1. Summary of Williams

320. Williams was filed on January 28, 2011, and is prior art under §102(e). (EX1006, 1.) The Examiner never found, and Applicant never cited, Williams during prosecution. Williams relates to "utilizing audio information that may already exist within a home or other location." (EX1006, [0007].) For example, "[a] user may ask a question, such as 'What's the recipe for deviled

eggs?'" and the system "may use the speech synthesis functionality of the system to provide audio answers to spoken questions." (EX1006, [0048].) Further, Williams' system "may accept spoken commands from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections." (EX1006, [0051]; *see also id.*, [0001], [0008], [0009], [0024].)

321. Various components of Williams are depicted in FIG. 1, which illustrates "an operational environment in which applications may utilize an application engine, and may respond to ***audio from user premises*** to provide various services to users." (EX1006, [0003].) Williams discloses that "[a]n individual ***interface device 116*** may include a microphone 202 that is configured to detect ambient noise, sounds, and ***speech***." (EX1006, [0023].) Further, "[w]ithin the ***home 102*** is a user 104 and several devices … [t]he illustrated devices include a ***media device 106***, representing any one or more of various types of devices such as ***televisions, audio players, video players.***" (*Id.*, [0011].) For network connection, "the devices within the ***home 102*** are illustrated as connecting to the network 114 via the in-home router 112." (EX1006, [0013].)

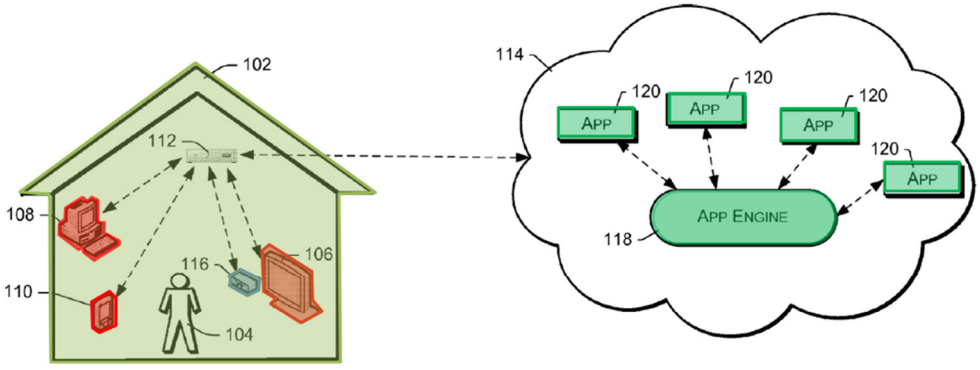


FIG. 1

322. Williams discloses that the network 114 includes "servers" including "a ***server*** 400 that may be used to implement the functionality of the application engine 118 and/or any of the applications 120." (EX1006, [0052].)

323. Williams discloses that the ***interface device 116*** receives ***speech commands***, which are processed by the ***server*** and cause actuation of a speaker of ***the interface device 116*** and/or control ***on-premise devices***. "In operation, the application engine 118 [of the ***server 400***] receives ***audio-based information*** from the ***interface devices 116***. The audio information may include audio streams, ***user commands or notifications derived from vocal utterances***, environmental information derived from on-premises audio, and so forth. The application engine 118 [of the ***server 400***] processes the audio-based information to determine various data about user activities, status, environment, commands, etc. This data can then be obtained by authorized applications 120,

which can ***act on the information to provide or enhance services to the user 104***." (EX1006, [0022].) For example, the server can "provide audio 222 for eventual transmission to the ***interface device 116***, to be played on the speaker 206." (EX1006, [0037].) Further, Williams discloses "allow[ing] a user to ***control various on-premises devices and equipment, such as media equipment and home-automation devices***. For example, such an application may accept spoken commands from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections." (EX1006, [0051].)

2. A POSITA would have been motivated to combine Williams and Lim

324. At the time of the invention, it would have been obvious for a POSITA to combine Williams and Lim. Williams and Lim both relate to processing and responding to human speech commands and controlling in-home devices accordingly. Williams discloses that "the application engine 118 is configured to receive the audio-based information 210. Depending on the type or format of the received information, the application engine 118 processes it to identify and create event, status, and/or command data relating to in-home activities." (EX1006, [0027].) "The application 120 may use the information from the event notifications 216 to initiate or customize services 218 on behalf

of users associated with the home 102." (EX1006, [0036].) Lim discloses that "a controlling device may be provisioned with means to receive voice input from a user, *e.g.*, a microphone, such that when placed in a setup mode the controlling device may solicit spoken appliance descriptive information, for example brand and model number, to be processed by a speech recognition system in order to identify the appliance to be controlled and the command data to which it is responsive." (EX1005, [0007].) Thus, both Williams and Lim are analogous art, from the same field of endeavor, and all relate to the same technical problems of processing and responding to human speech commands (*e.g.*, to provide services and/or to control appliances).

325. Lim teaches a technique for communicating information for controlling ***home appliances*** from a remote ***server***. Lim discloses that the ***server*** transmits the information to the ***first/controlling device*** (rather than the ***second devices***/appliance), which then transmits a command to the ***second device***. For example, "[o]nce the desired function has been uniquely identified, at step 610 the indicated command code data may be provisioned to the ***controlling device***. As appropriate for the particular implementation, such provisioning may comprise downloading the data itself, supply of a pointer to an entry in a local database, etc." (EX1005, [0036].) Then "the ***controlling device 100*** may be capable of transmitting ***commands*** to the appliances, using

any convenient IR, RF, Point-to-Point, or ***networked protocol 120***, to cause the appliances to perform operational functions." (EX1005, [0017].)

326. The Williams-Lim combination would result in a system where the ***server 400*** sends information for controlling ***second devices*** to Williams' ***interface device 116*** (rather than to Williams' ***second devices*** or ***appliances***). Williams' ***first device*** then in turn forwards such information (or command derived from the information) to the ***second devices*** or ***appliances***. This combination would predictably enhance the reliability and scalability of the system by having a centralized, local hub (*i.e.*, the controlling device) that receives control information from the remote server and transmits commands to the home appliances and other second devices. Moreover, the combination would yield predictable results of reducing the processing and network burden on the remote server, improving overall system responsiveness and efficiency. By adopting Lim's communication technique, the cloud server in Williams would no longer need to maintain continuous Internet connections with every appliance or second device in the home. Instead, the interface device could serve as a hub communication node, receiving control data from the server and issuing local commands to the appropriate appliance. The server only needs to establish and maintain connection with the hub, not with numerous individual second devices or appliances. This predictably reduces bandwidth usage,

improves latency, simplifies network management, and makes the system easier to scale, troubleshoot, and manage.

327. Further, the combination provides improved security, because a single hub communicating with the cloud minimizes the number of Internet-exposed endpoints and allows for one secured connection and centralized authentication/encryption handling. The hub/first device's connection with the second devices can be more reliable and faster than Internet connections, and thus can be better suited than the server for handling coordination among the second devices. Furthermore, in addition to providing improved security, the combination also simplifies management of security rules, as the appliances need not have any firewall rules configured to allow external devices (*e.g.*, the server device) to directly communicate with them. In addition, since the hub/first initiates a communication session with the server (*e.g.*, by sending the first and second human voice data to the server), and firewalls are configured to allow responses to client initiated requests, no special firewall rules are needed for the hub/first device either. Also, to the extent the server system comprises a cluster of servers, where each server in the cluster has its own different IP address, the combination does not require any special firewall rules allowing for any of the servers in the cluster to initiate a connection to either the hub/first device or any of the second devices.

328. The above discussed benefits would prompt the application of Lim in the same field based on design incentives and market forces and yield the predictable result of enhancing the simplicity and reliability of the system. Because both Williams and Lim include voice data processing and device control, a POSITA would be motivated to use the known communication techniques of Lim to improve the system of Williams in the same way. Further, it would have been obvious to try combining the teachings of Lim and Williams. There are only a finite number of identified and predictable solutions for transmitting messages and control signals from a remote server to home appliances, direct communications or via a hub device. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Lim has successfully demonstrated that the hub approach provides improved speed, reliability, security, and scalability, and thus there would have been a reasonable expectation of success for combining Williams and Lim.

3. Claim 1

a. 1[PR]: "A first device for use with a Wireless Local Area Network (WLAN) in a building, the first device comprising:"

329. Both Williams and Lim disclose this. (*Supra* § IX.A.2.a.)

330. For example, Williams discloses a ***first device*** and for use with a ***Wireless Local Area Network (WLAN)*** in a ***building***.

331. First, Williams discloses a ***first device*** in a ***building***. For example, Williams states that "the ***user's home 102*** may be equipped with one or more on-premises ***audio monitoring devices 116*** . . . having a microphone and a network interface". (EX1006, [0016].) "The ***interface device*** may also have a speaker or other form of audio output." *Id.* The ***device 116*** is in "network communications with various entities . . . via the network 114." (*Id.*, [0023].)

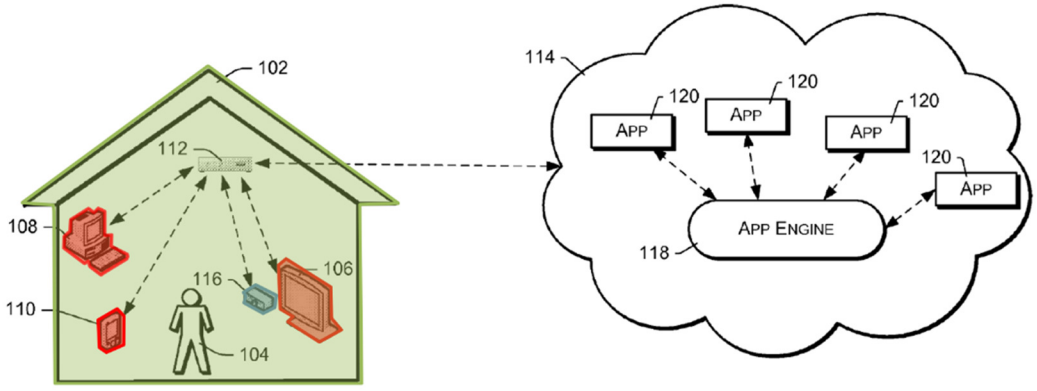


FIG. 1

332. FIG. 2 provides additional device on Williams' first device. FIG. 2 "is a block diagram illustrating additional details regarding the system of FIG. 1." (EX1006, [0004].) With reference to FIG. 2, Williams discloses that "[a]n individual ***interface device 116*** may include a microphone 202 that is configured to detect ambient noise, sounds, and speech. The ***interface device*** 102 [*sic*] may also include a network interface 204 that facilitates network

communications with various entities, including the application engine 118 and one or more applications 120, via the network 114. The ***interface device*** may also include a speaker 206, and operating logic 208." (EX1006, [0023].)

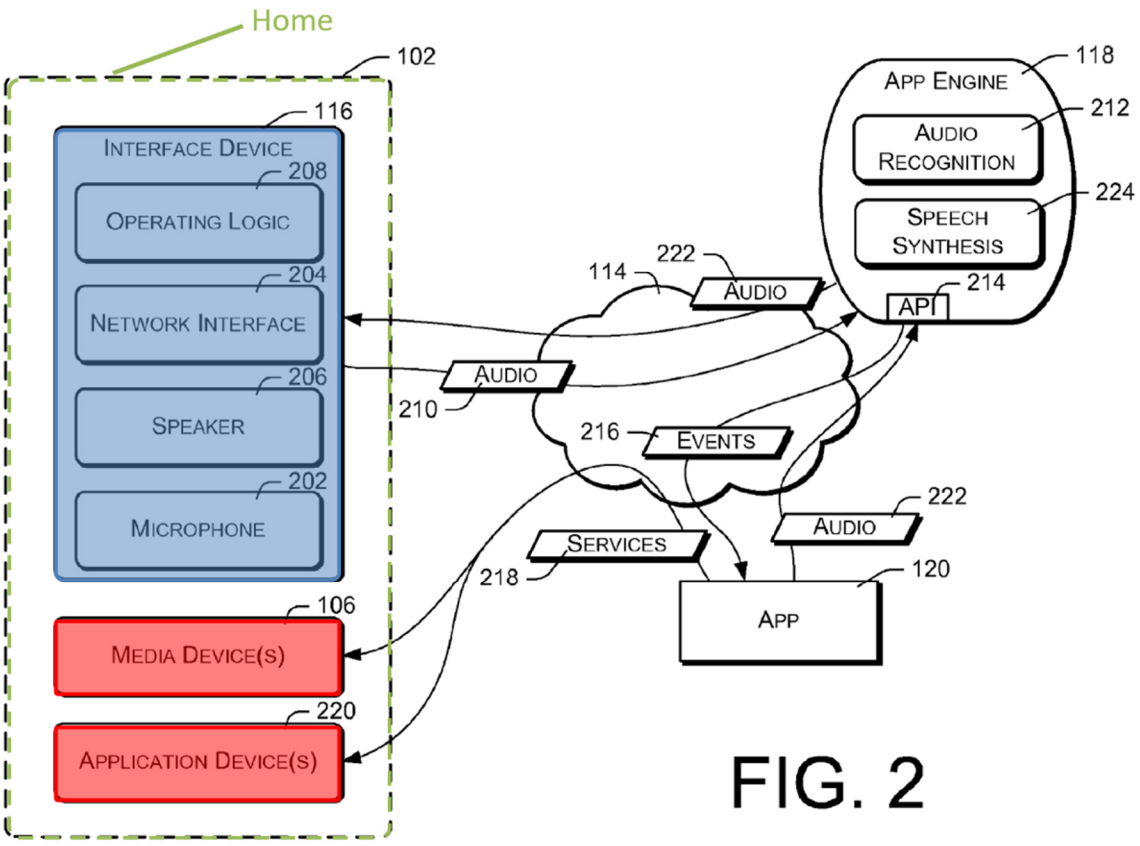


FIG. 2

333. Williams discloses a ***Wireless Local Area Network (WLAN)***. Wiliams discloses that the first device (*i.e.*, ***audio monitoring devices 116***) and the ***second device*** can connect to a local ***router or access point*** that in turn connects to the Internet. With reference to FIG. 1, Williams discloses "the ***devices shown within the home 102*** are merely examples of a wide variety of devices that may be present within a user premises or associated with a user.

Many such devices may have some sort of ***network connectivity***. In this case, the ***in-home devices*** connect to ***a router or access point 112***, which is in turn connected to ***a public network 114. Connection between in-home devices and the router 112 may be through wired or wireless technologies***." (EX1006, [0012].) In other words, Williams expressly discloses that each of the ***interface device 116*** and the ***media devices 106/108/110*** communicates over a wireless network through wireless router/access point 112. "[T]he ***interface devices 116*** communicate with and transmit environmental information to a cloud-based application service or engine 118, ***via the router 112*** and the network 114." (EX1006, [0020], [0023].)

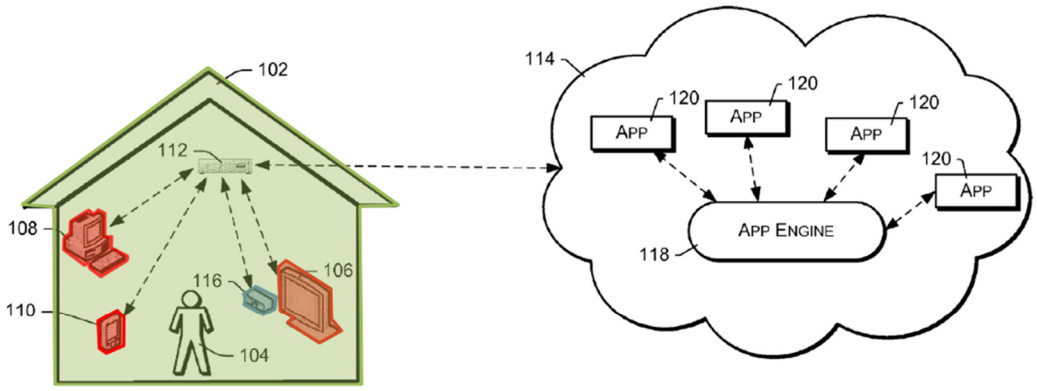


FIG. 1

334. A POSITA would have understood that the in-home wireless network in Williams is a WLAN network. A WLAN network is characterized

by a local access point (*e.g.*, router) that broadcast wireless signals to wireless-enabled devices. It was common knowledge that "with a ***wireless LAN***, the computers have to find a way to connect, or associate, themselves to the LAN, because they are not physically connected. The way they do this is by ***tuning themselves to the wireless router*** using the name of the router . . . Once a computer has connected, or associated, itself to the access point, it becomes part of the LAN . . . ***Now the computer can send and receive information*** to and from other computers on the LAN or ***to and from the Internet by transmitting over wireless to the wireless router (or access point)***. The wireless router acts as the central communication routing point." (EX1007, 11, 100.) Such a local wireless network through a local router is exactly what is shown in FIG. 1 of Williams. A POSITA would have recognized that Williams discloses a WLAN network in the ***home 102***.

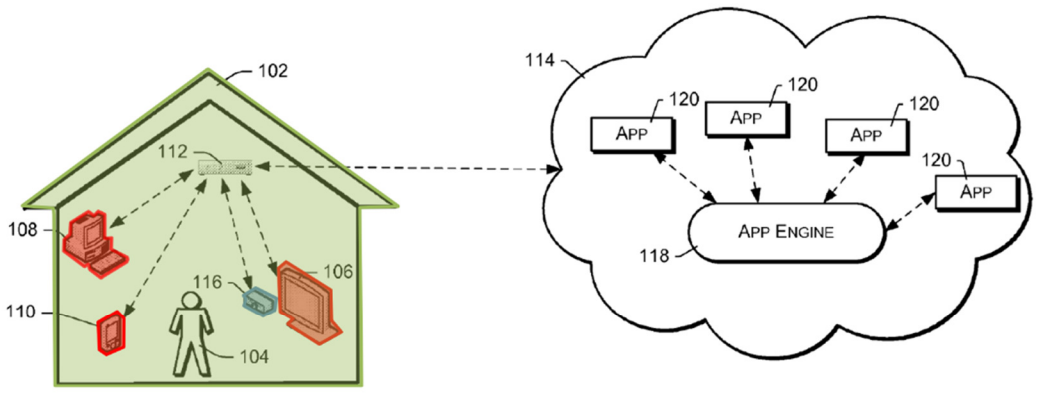


FIG. 1

335. Williams' WLAN network is consistent with the description of the '721 patent. The '721 patent explains that a device can "communicate (such as send sensor info or receive actuator commands) with the router (or gateway) or the control server using the same or different ***WANs used by the router***[.]" (EX1001, 28:7-11.) The '721 patent also explains that "[t]he external network or one or more of the in-building (or in-vehicle) networks may be a Local Area Network (LAN), and the router, the first device, or the second device may comprise a LAN interface, and the LAN interface may include a LAN port and a LAN transceiver." (EX1001, 46:1-5.)

336. Alternatively, I note that Williams also discloses network 114 may be a WLAN. (EX1006, [0015].) Williams states, "[t]he network 114 may thus include various types of communication networks, including ***local-area***

networks, wide-area networks, the Internet, wireless networks, cable television networks, telephone networks, cellular communications networks, etc. The techniques discussed herein may also be practiced in smaller networks, including ***localized*** and/or private networks." (EX1006, [0015].)

337. Williams' ***first device*** is for use with a ***Wireless Local Area Network (WLAN)***. As discussed above, Williams' ***first device*** connects to a WLAN network and thus is for use with a WLAN. (EX1006, [0012].) Further, Williams' ***audio interface device 116*** has the necessary configuration (*e.g.*, a wireless network interface) for use with a ***Wireless Local Area Network (WLAN)***. With reference to FIG. 2, Williams discloses that "[t]he ***interface device [116]*** may also include a ***network interface 204*** that facilitates network communications with various entities, including the application engine 118 and one or more applications 120, via ***the network 114***." (EX1006, [0023], [0016].) "To collect the audio information, ***network-enabled*** microphones can be placed at different activity centers within a home or user premises. The microphones may be incorporated in small, self-contained units, with ***wireless networking capabilities***, and configured to transmit audio-based information to a cloud-based application engine or service." (EX1006, [0007].)

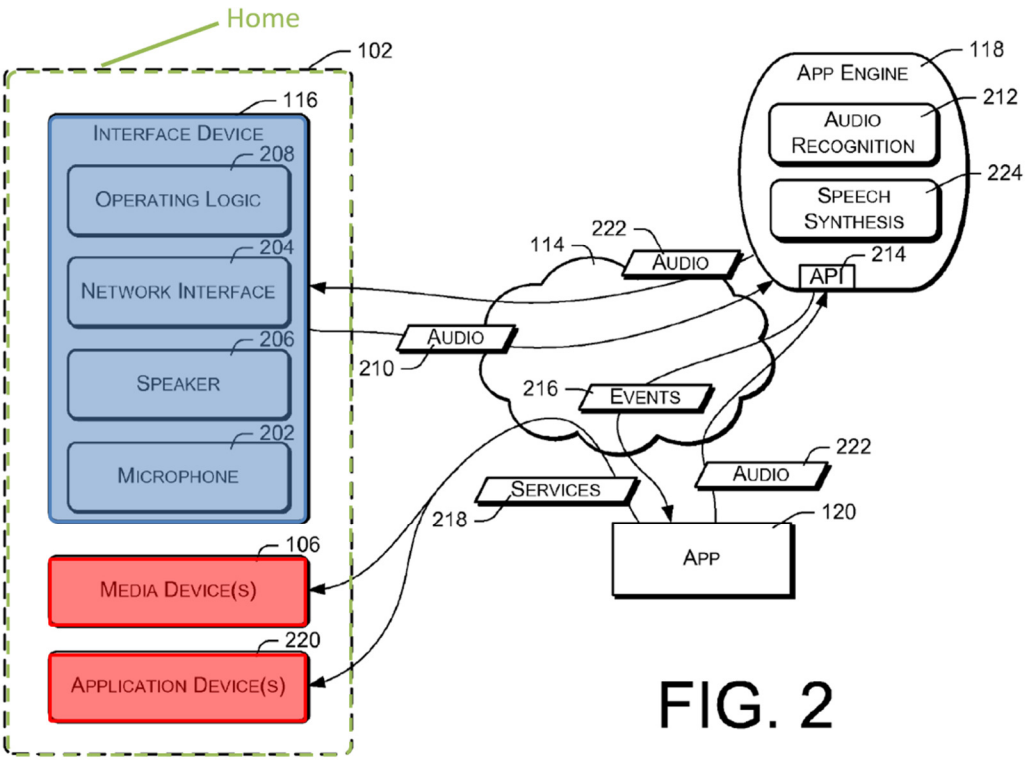


FIG. 2

338. As discussed above, Williams discloses that the ***first device*** is in a ***building***. Williams discloses that "[t]he user's ***home 102*** may be equipped with one or more on-premises ***audio monitoring devices 116***, referred to herein as ***in-home audio interface devices***." (EX1006, [0016], [0023].) The "***audio interface devices 116*** can be placed within the ***home 102*** so that their microphones detect ambient sound or noise within the ***home 102***." (EX1006, [0019].)

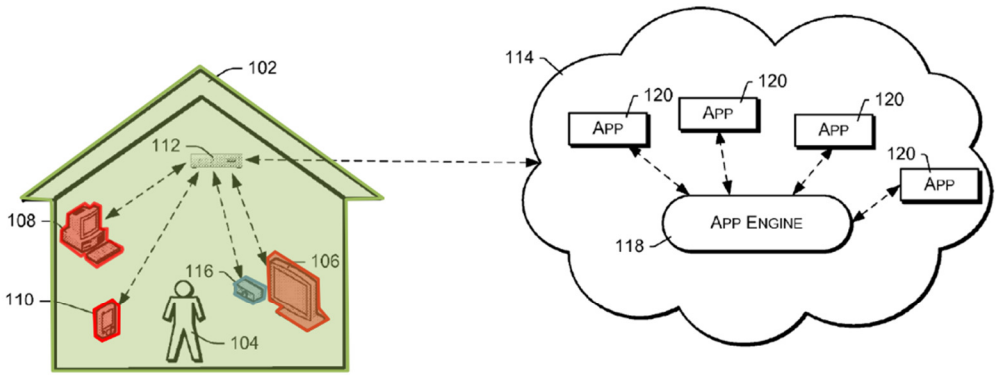


FIG. 1

339. Similarly, with reference to FIG. 2, Williams discloses that "***in-home audio interface devices 116*** may be located within the ***home 102***." (EX1006, [0023].)

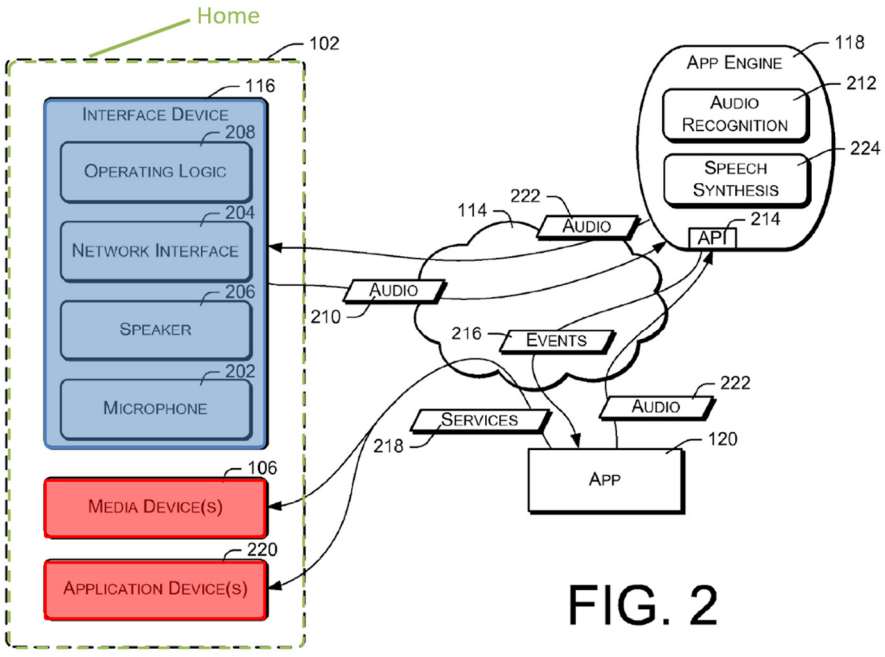


FIG. 2

b. 1[A]: "a WLAN transceiver for communicating over the WLAN;"

340. Both Williams and Lim disclose this element. (*Supra* § IX.A.2.b.)

341. For example, Williams discloses a ***first device*** having a WLAN transceiver for communicating over the WLAN. As discussed, Williams discloses all devices in the home (including ***audio interface device 116***) are part of a WLAN network and wirelessly communicate with the public network 114 via WLAN (*e.g.*, via a router or access point 112). In light of this disclosure, a POSITA would have understood that Williams' ***first device*** has a WLAN transceiver. A WLAN transceiver is the hardware component that receives wireless signals and is necessary for WLAN communication. "The mobile WLAN device 12a includes a WLAN transceiver 20 operatively coupled to antenna 22 and processing circuitry 24 operatively coupled to the WLAN transceiver to send and receive suitable information via a WLAN protocol." (EX1027, 3:17-20.) "The mobile WLAN devices 12a-12n send messages 34a-34n wirelessly using their respective WLAN transceivers, ***as known in the art***." (EX1027, 3:38-40.) "'[W]ireless terminal' represents any wireless device capable of wireless communication . . . Wireless terminal 100 includes a user interface 102, memory 104, antenna 106, transceiver 110 . . . Antenna 106, coupled to transceiver 110, receives and transmits wireless signals[.]"

(EX1028, [0013]-[0015].) Thus, it would have been obvious to a POSITA that Williams' ***first device*** would have the necessary hardware and software components, including a WLAN transceiver, to communicate in a WLAN network as disclosed.

342. Indeed, Williams' ***first device*** has a "network interface", which requires a transceiver. Williams discloses, with reference to Figure 2 below, that "[t]he ***interface device 102*** may also include a ***network interface 204*** that facilities network communications with various entities, including the application engine 118 and one or more applications 120, via the network 114." (EX1006, [0023], [0007], [0012], [0016], [0020].) The '721 patent explains that "a network interface may include a transceiver or modem and a communication port for coupling to the network[.]" (EX1001, 28:62-64.) Considering this disclosure, a POSITA would have understood that the "network interface 204" contains a WLAN transceiver for coupling the ***interface device*** to the WLAN network *e.g.*, via a WLAN interface card. (*Id.*, EX1007, 99-100.) "Each computer or device that you want to join your wireless network needs a ***wireless network interface card (NIC)***. A wireless NIC provides a number of functions, including: [e]nabling a computer (or device) to connect to your home network through a wireless signal; [s]ecuring the communications path between the ***wireless router and your computers and devices*** connected over wireless[;]

Providing a way for you to connect your computer to other wireless networks outside your home." (EX1007, 99-100).

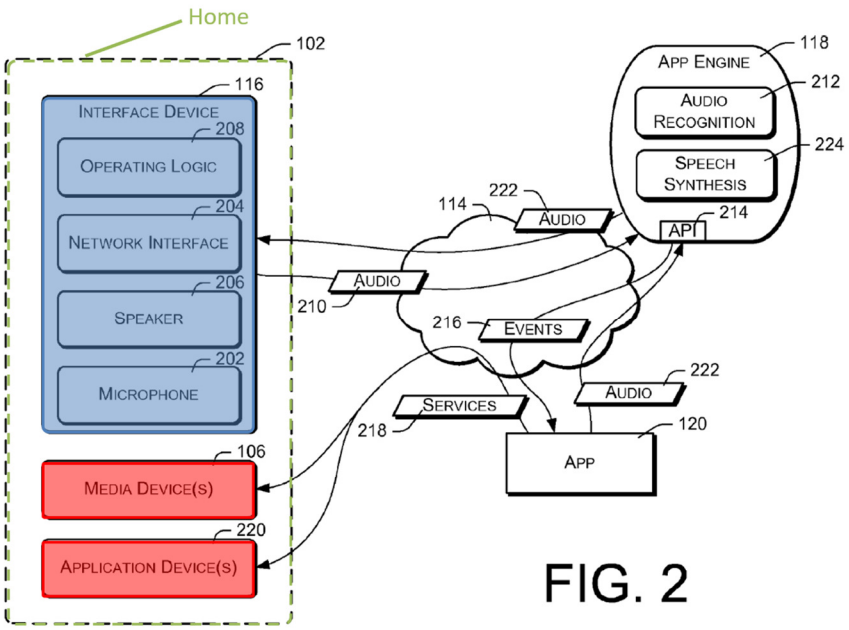


FIG. 2

c. 1[B]: "a microphone for capturing human voice data;"

343. Both Williams and Lim disclose this element. (*Supra* § IX.A.2.c.)

344. For example, Williams discloses a ***microphone*** in the ***first device***. With reference to FIG. 1, Williams discloses "[a]n ***audio interface device 116*** … having a ***microphone*** and a network interface." (EX1006, [0016]; *see also*, *id.*, [0007] ("network-enabled ***microphones*** can be placed at different activity centers within a home" "to collect the audio information" and "microphones may be incorporated in small, self-contained units"); EX1006, [0019] ("***audio interface devices 116*** can be placed within the home 102 so that their

microphones detect ambient sound or noise within the home 102"; *id*., claims 1-3 and 21.).

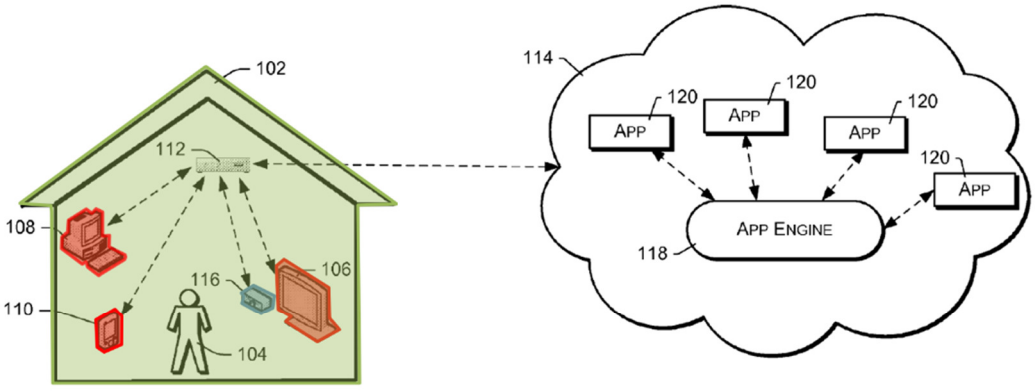


FIG. 1

345. Williams discloses capturing of ***human voice data*** by the microphone in the ***first device***. Williams discloses that the "***in-home interface device*** [] compris[es]: a ***microphone*** configured to receive audio from a nearby user[.]" (EX1006, Claim 21; *see also id.*, claims 1-3.) Williams also discloses a system comprising "a cloud-based service that is accessible by ***on-premises audio monitoring devices*** to receive ***audio-based information***", where "the ***audio-based information*** comprises ***audio streams***" and/or "***user commands derived from vocal utterances of the users.***" (EX1006, claims 1-3.) Williams further discloses that "[a]udio and audio-based information may be formatted as files, ***audio streams,*** data objects, notifications, text streams, and

so forth", and that "[t]he audio-based information may comprise [of] an ***audio stream*** or ***data*** derived from an ***audio stream.***" (EX1006, [0038]–[0039].) Williams teaches that "one or more ***in-home audio interface devices 116*** may be located within the home 102. An individual ***interface device 116*** may include a microphone 202 that is configured to detect ambient noise, sounds, and ***speech***." (EX1006, [0023], *see also, id*., [0016].) Williams further notes that "[i]t may also be convenient to place the ***interface devices*** near activity centers, such as in the kitchen, on desks, and so forth, so that ***user speech*** can be clearly received." (EX1006, [0019].) Thus, Williams' microphone captures human voice data.

d. 1[C]: "and a sounder for converting an electrical energy to audible sound waves;"

346. Both Williams and Lim disclose this element. (*Supra* § IX.A.2.d.)

347. For example, Williams discloses a ***speaker*** in the ***first device***. Williams discloses that "the ***interface device*** may also have a ***speaker*** or ***other form of audio output***." (EX1006, [0016].) With reference to FIG. 2, William discloses that "[t]he ***interface device*** may also include a ***speaker 206***, and operating logic 208." (EX1006, [0023].)

348. As discussed, a speaker is a sounder which converts electrical energy to audible sound waves transmitted through the air through the movement of a diaphragm. (*Supra* § IX.A.2.d; EX1001, 105:61-106:4;

EX1018, 122.) Thus, a POSITA would have understood that Williams discloses a sounder (*e.g.*, ***speaker 206***) in the first device (*e.g.*, ***interface device 116***) for converting an electrical energy to audible sound waves.

e. 1[D]: "and one or more processors programmed with computer program instructions that, when executed, cause the first device to:"

349. Both Williams and Lim disclose this element. (*Supra* § IX.A.2.e.)

350. For example, Williams discloses one or more processors in the first device programmed with computer program instructions. Williams discloses that "[t]he interface device may also include a speaker 206, and operating logic 208. The operating logic 208 may be implemented as ***one or more programmable processing units***, associated memory, and appropriate instructions stored in the memory and executed by the processing unit. Other types of logic may also be used to configure the interface device 116 to ***perform the functions described here***." (EX1006, [0023].) It would have been obvious to a POSITA that Williams' first device comprises one or more processors programmed with computer program instructions that, when executed, cause the first device to perform all the recited operations. (*Infra* §§IX.I.3.f-IX.I.3.l.)

f. 1[E]: "capture, using the microphone, first and second human voice data;"

351. Both Williams and Lim disclose this element. (*Supra* § IX.A.2.f.)

352. For example, Williams discloses capturing first and second human voice data using the microphone in the first device. As discussed, Williams' first device comprises a microphone. (*Supra* § IX.I.3.c.)

353. In Williams, the voice commands that trigger playing of audio 222 by the interface device are the recited "***first human voice data***", and the voice commands that cause control of media equipment and home-automation devices by the interface device are the recited "***second human voice data***," as discussed below. They can be parts of the same command.

354. Williams discloses using the ***microphone*** in the ***first device*** for capturing ***first human voice data***. Any of the human voice commands disclosed in Williams that cause a sounder (*e.g.*, ***speaker 206***) in the first device (*e.g.*, ***interface device 116***) to sound digital audio content in the recited manner discloses the recited "***first human voice data***." Referring to FIG. 2, "***audio-based information 210***" discloses the first human voice data as discussed below. (EX1006, [0022], [0024]-[0025]; *infra* § IX.I.3.i.)

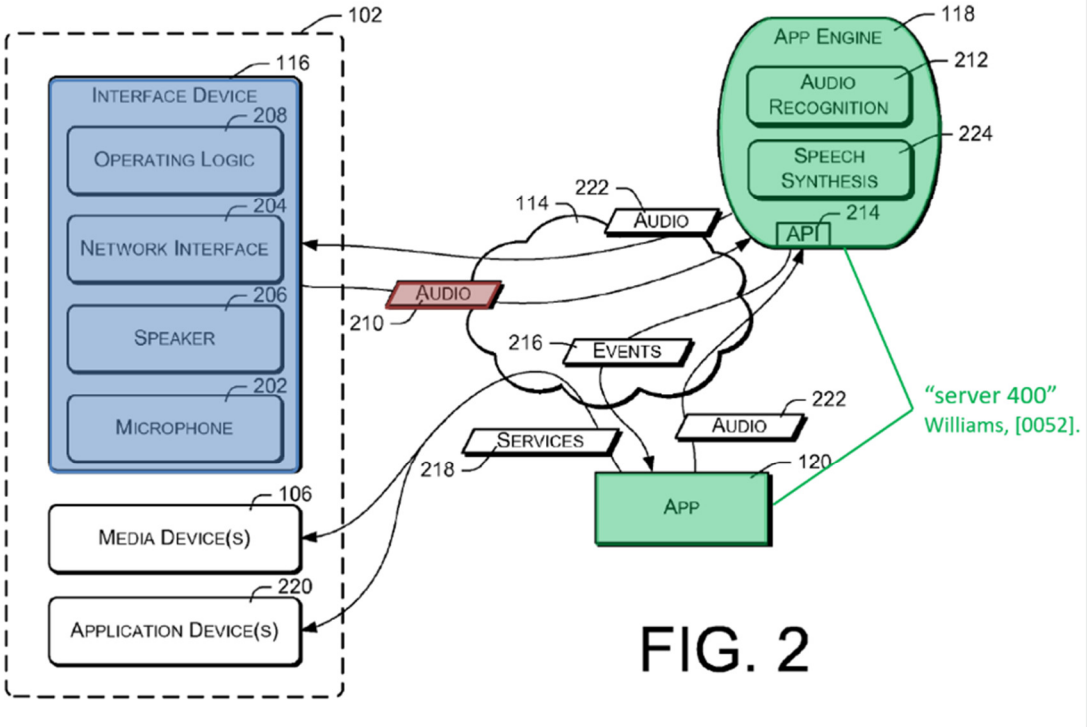


FIG. 2

355. Williams discloses other examples of ***first human voice data***. For example, Williams discloses that "[a]n application may provide general information in response to ***users' uttered questions***. ***A user may ask a question, such as 'What's the recipe for deviled eggs?***'" and "the application may use the speech synthesis functionality of the system to ***provide audio answers*** to spoken questions." (EX1006, [0048].)

356. As another example, Williams states that "a cloud-based application may allow users to order tickets without needing a computer. ***A user may tell the application what movie they want to see***. The application may use pre-configured information about the user (such as location) to determine which

theater is near the user. The application may respond to a spoken command to order movie tickets by using the speech synthesis functionality of the system to ***query the user regarding movie selections and times***." (EX1006, [0045].) "As another example, such an application may allow the user to ***inquire about a flight or other event***. The system might respond by sending information to the user via a device associated with the user, or by ***audibly responding using speech synthesis***." (EX1006, [0050].) It would have been obvious that these voice commands are captured using the microphone.

357. Williams discloses capturing, by the ***microphone*** in the ***first device***, a ***second human voice data***. To the extent that the recited "second human voice data" is different from the recited "first human voice data", Williams is clear that its system can receive and process ***multiple*** human voice commands. (EX1006, [0044], [0036], [0021].)

358. Further, Williams discloses human voice commands that can effectuate remote device control (*i.e.*, "***second human voice data***"). Any of the human voice commands disclosed in Williams that causes sending of a control message in the recited manner discloses the recited "***second human voice data***."

359. For example, "[a]n application might also be designed to allow a user to ***control various on-premises devices and equipment, such as media equipment and home-automation devices***. For example, such an application

may accept ***spoken commands*** from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections." (EX1006, [0051].) Such ***spoken commands*** are "second human voice data." "Cloud-based or network-based applications such as the applications 120 may communicate and operate in conjunction with a wide variety of network-connected devices, including ***entertainment devices, game boxes, media devices, home automation and security devices, communications devices, and so forth***." (EX1006, [0021]; *see also id.*, [0036].)

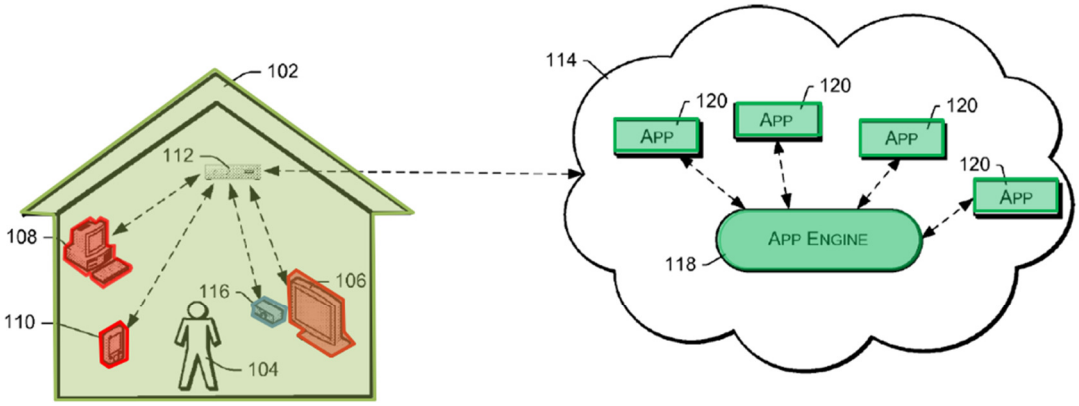


FIG. 1

360. A POSITA would have understood that the ***second human voice data*** (to the extent different from the first human voice data) is also captured using the microphone. To the extent that the first and second human voice data are captured/sent together (*e.g.*, as parts of one query) (which is not conceded), a

POSITA would find it obvious that "audio 210" (FIG. 2) can include multiple parts, such as a trigger phrase ("***first human voice data***") and the instruction ("***second human voice data***"), consistent with common interpretations. At the time of the invention, it was known to a POSITA that a speech command (*e.g.*, a natural-language query) to a virtual assistant may include at last two parts: a trigger/wake phrase and a specific command (*e.g.*, "turn on the TV"). (EX1023, 2:12-16, 34:63-67.) Further, trigger/wake phrases were known at the time of the invention. "***Upon receiving and recognizing a speech trigger phrase*** and determining that it is in a better position to handle subsequent user interaction than any other device that simultaneously recognized the same speech trigger phrase (see device coordination discussion below), . . . , or in the event the device is instructed to wake up and expect a subsequent speech command by a system controller, the device beeps, changes its status light 34 or plays a prerecorded or synthesized audio message (*e.g.* 'How may I help you?') through speaker 19 and/or 20." (EX1013, 8:20-32.). "The smartphone 208 may be programmed to recognize certain words that are stated into its microphone, as being ***words to trigger a search query*** that involves passing sound data up to the central server system 204 through the network 202, such as the [I]nternet." (EX1024, [0045], *see also* [0047], ("use a ***trigger word*** before the query, either to start the device listening for the query, or to define a context

for the query[.]").) Thus, the "audio 210" can equally be mapped to include or be both the first and second human voice data and is captured/sent together.

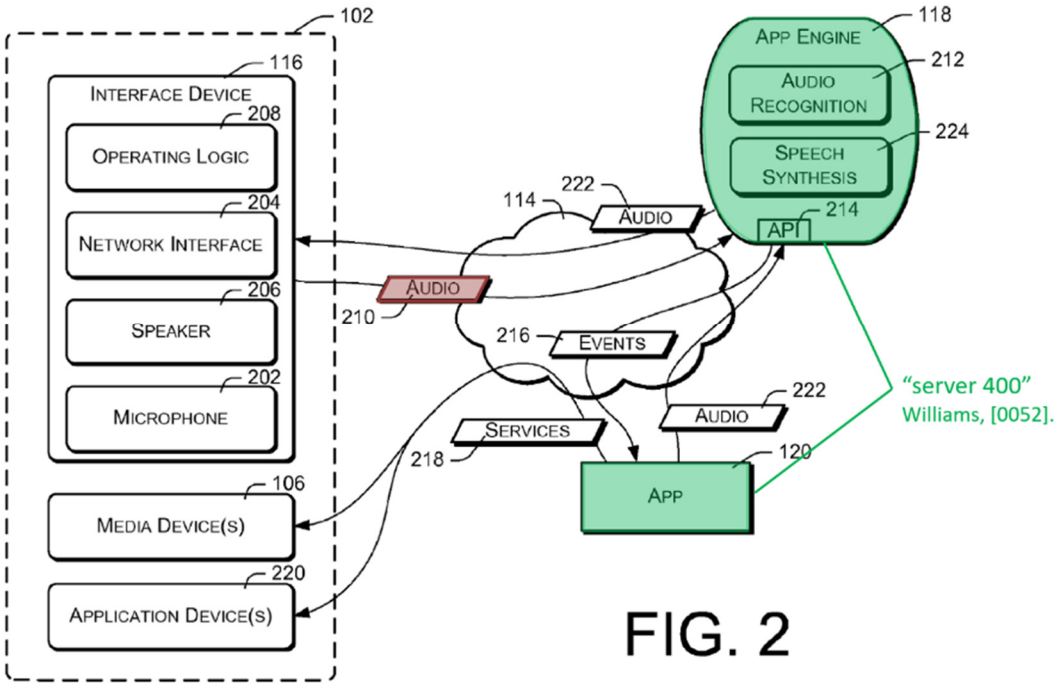


FIG. 2

g. 1[F]: "send, over the Internet via the WLAN, to a server device, the captured first and second human voice data;"

361. Both Williams and Lim disclose a ***first device*** sending the captured ***first human voice data*** and ***second human voice data*** to a ***server device*** over the ***Internet via the WLAN***. (*Supra* § IX.A.2.g.)

362. For example, Williams discloses a ***server device***. Williams discloses "a ***server 400*** that may be used to implement the functionality of the application engine 118 and/or any of the applications 120." (EX1006, [0052].)

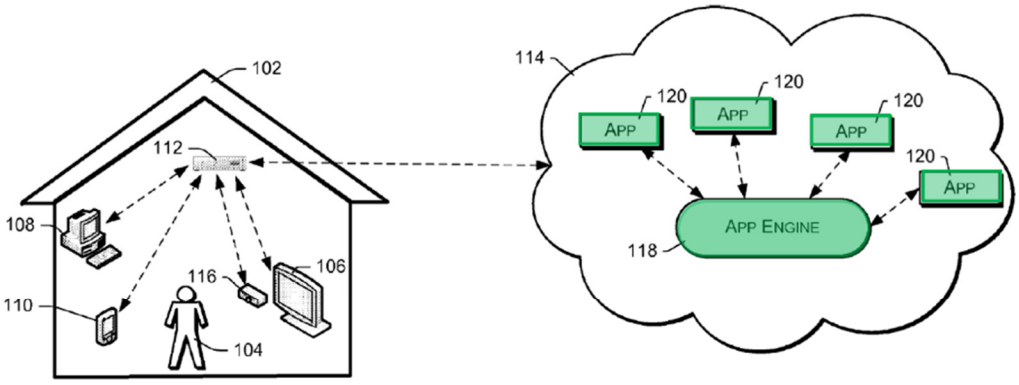


FIG. 1

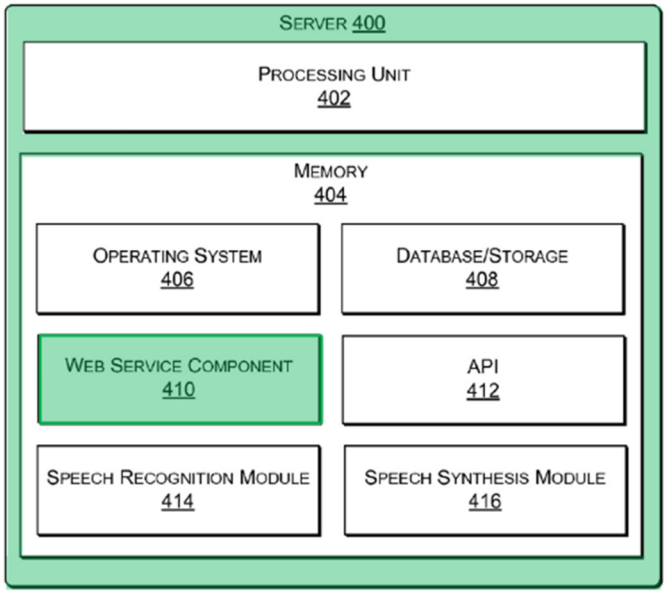


FIG. 4

363. As shown in FIG. 1, Williams discloses the in-home network in connected to cloud network 114. (EX1006, FIG. 1.) A POSITA would have understood that the server 400 implementing the application engine 118 and

applications 120 is a cloud server, in other words, ***Internet-connected***. For example, Williams discloses that "the ***server 400*** may also comprise a ***web service component 410*** that interacts with remote devices such as computers, media consumption devices, and so forth." (EX1006, [0055].) Williams discloses "a ***cloud-based*** application engine receives audio information from the interfaces and provides information to ***cloud-based*** applications based on the audio within the user premises. The other applications utilize the information to provide or enhance services to the users." (EX1006, Abstract; *see also id*., [0007]-[0009].) Williams uses "cloud" and "Internet" interchangeably. (*See, e.g.,* EX1006, [0001] ("Homes and other user premises are increasingly equipped with always-on ***Internet or 'cloud'*** connectivity.").)

364. Given FIG. 1, a POSITA would have understood that ***server 400*** (running 118 and 120) is an Internet-connected ***server device*** in the cloud, which is connected to the in-home WLAN (*e.g.*, via router 112). Williams discloses that "the network 114 represents what has come to be known as ***the 'cloud,' which may include a vast array of network-based entities, servers, clients, applications, and services.*** Blocks within the network 114 represent cloud-based applications and/or services that are relevant to this discussion." (EX1006, [0014].) "The network 114 may be a conglomeration of various discrete networks and systems, interconnected in various ways so that

for functional purposes they appear more or less as a unified network. The network 114 may thus include various types of communication networks, including local-area networks, wide-area networks, ***the Internet***, wireless networks, cable television networks, telephone networks, cellular communications networks, etc. The techniques discussed herein may also be practiced in smaller networks, including localized and/or private networks." (EX1006, [0015].) Further, Williams discloses that the "applications 120 may be ***web-based applications*** that are accessible to users via a network browser, operating in conjunction with the computer 108 or another Internet terminal or access device." (EX1006, [0021].) A POSITA would have understood that the in-home wireless network (*e.g.*, via router 112) is connected to the cloud-based ***server 400*** via the Internet. This is consistent with the descriptions of the '721 patent. The '721 patent explains that "[t]he system may comprise a router in the building, . . . an Internet-connected control server (referred herein also as 'cloud server' and 'gateway server') external to the building (or vehicle) storing the control logic, and communicatively coupled to the router over the Internet via the external network." (EX1001, 32:25-42.)

365. To the extent that the recited "server device" refers to a single server device or multiple server devices, Williams discloses both. Williams discloses that "FIG. 4 illustrates relevant components of ***a server 400*** that may

be used to implement the functionality of the application engine 118 and/or any of the applications 120. Generally, any of these entities may be implemented by ***one or more servers***, with the various functionality described above distributed in various ways across the different servers." (EX1006, [0052].) In other words, Williams discloses that the application engine 118 and/or any of the applications 120 can be implemented by one server device or multiple server devices.

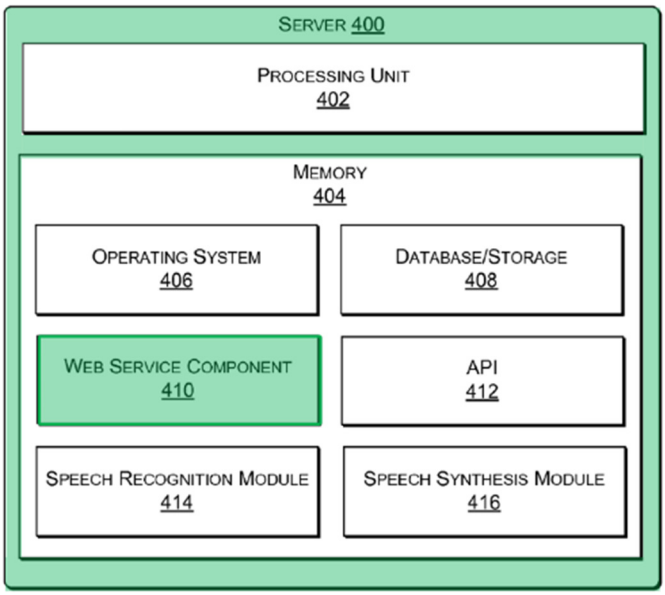


FIG. 4

366. I note that Williams mentions "[s]ome of the applications 120 may serve as resources for programs that execute on local devices such as the computer 108 or the personal device 110. Such local programs may operate in conjunction with the cloud-based applications 120 to provide various different

services and functionality, ranging from email services to games." (EX1006, [0021].) A POSITA would have understood that this does not mean that the server itself is local; rather, the server works with local devices. Williams refers to a local program that would provide a graphical user interface (*e.g.*, a graphical user interface for an email inbox or a game), which can update based on messages received from an application 120 that is hosted on a remote server device. Indeed, "most Internet-based email accounts use POP3 (Post Office Protocol v.3). POP3 allows an email client to ***collect the email on a remote server (POP server)*** . . . When an email client wishes to make use of the service, it establishes a TCP connection with the server host. When the connection is established, the POP3 server sends a greeting. The client and POP3 server then exchange commands and responses." (EX1017, 664.)

367. Second, Williams discloses that the ***first device*** sends the captured human voice data (including the ***first human voice data*** and ***second human voice data***) to the ***server***. As discussed above, application engine 118 and/or any of the applications 120 are implemented on a server device, ***server 400***. Williams discloses that "[i]n operation, the application engine 118 [of ***server 400***] receives ***audio-based information*** from the ***interface devices 116***. The audio information may include ***audio streams, user commands or notifications***

derived from vocal utterances, environmental information derived from on-premises audio, and so forth." (EX1006, [0022], *see also id.*, [0020].)

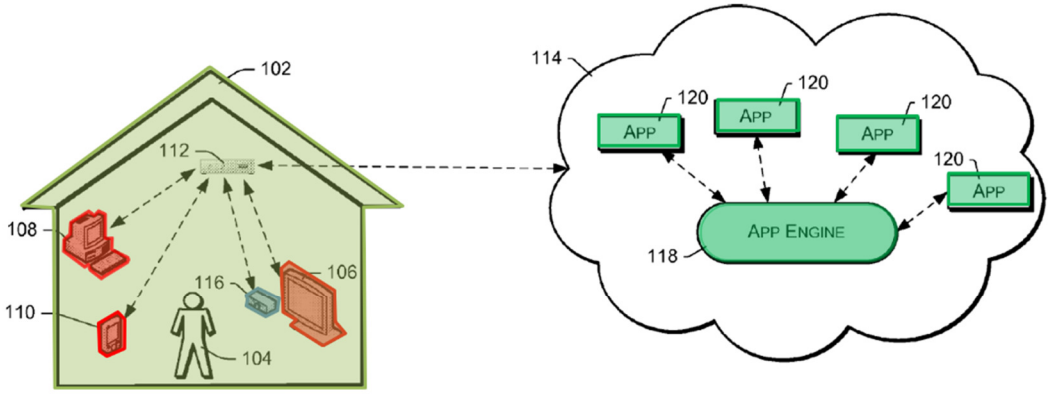


FIG. 1

368. Further, "FIG. 2 illustrates additional details regarding interactions between the described devices and services. As mentioned above, one or more in-home audio interface devices 116 may be located within the home 102. An individual ***interface device 116*** may include a microphone 202 that is configured to detect ambient ***noise, sounds, and speech***. The ***interface device 102*** [sic] may also include a network interface 204 that facilitates network communications with various entities, including the application engine 118 and one or more applications 120 [of ***server 400***], via the network 114." (EX1006, [0023].)

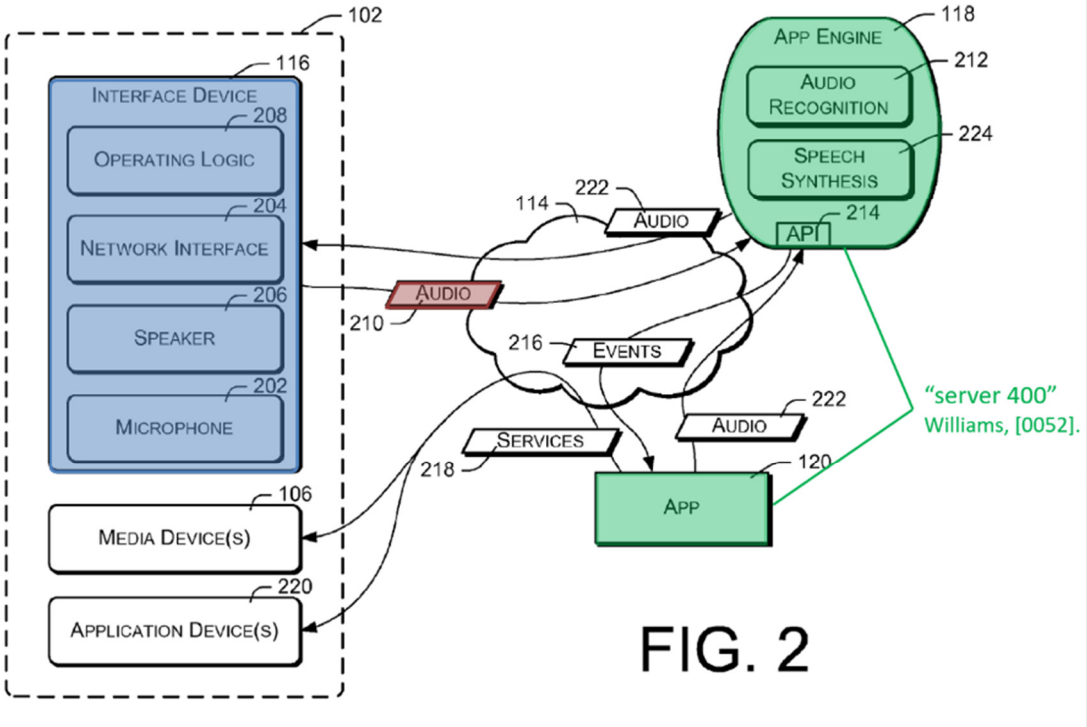


FIG. 2

369. Williams also discloses that "the ***interface device 116*** is configured to send ***audio-based information 210*** to the application engine 118 [of ***server 400***], based on ***ambient audio*** detected by the microphone 202 of the ***interface device 116***. The ***audio-based information 210*** may comprise a continuous audio stream, or may comprise separate audio streams corresponding to ***different periods of activity or noise within the home 102*** and nearby the ***interface device 116***. For example, an audio stream may be transmitted whenever ambient noise levels exceed a minimum threshold." (EX1006, [0024].) Williams further discloses that "the application engine 118 [of ***server 400***] is configured to receive the ***audio-based information 210***." Further, "[***a***]***udio and audio-based***

information may be formatted as files, ***audio streams, data objects***, notifications, text streams, and so forth" when communicated. (EX1006, [0038].) As shown in FIG. 2, ***audio 210*** is sent to the ***server*** by the ***interface device 116***.

370. As discussed, any of the human voice commands disclosed in Williams that cause a sounder (*e.g.*, ***speaker 206***) in the first device (*e.g.*, ***interface device 116***) to sound digital audio data in the recited manner discloses the recited "***first human voice data***." (*Supra* §IX.I.3.f; *infra* § IX.I.3.i.) Referring to FIG. 2, "***audio-based information 210***" discloses the first human voice data . (EX1006, [0022], [0024]-[0025].)

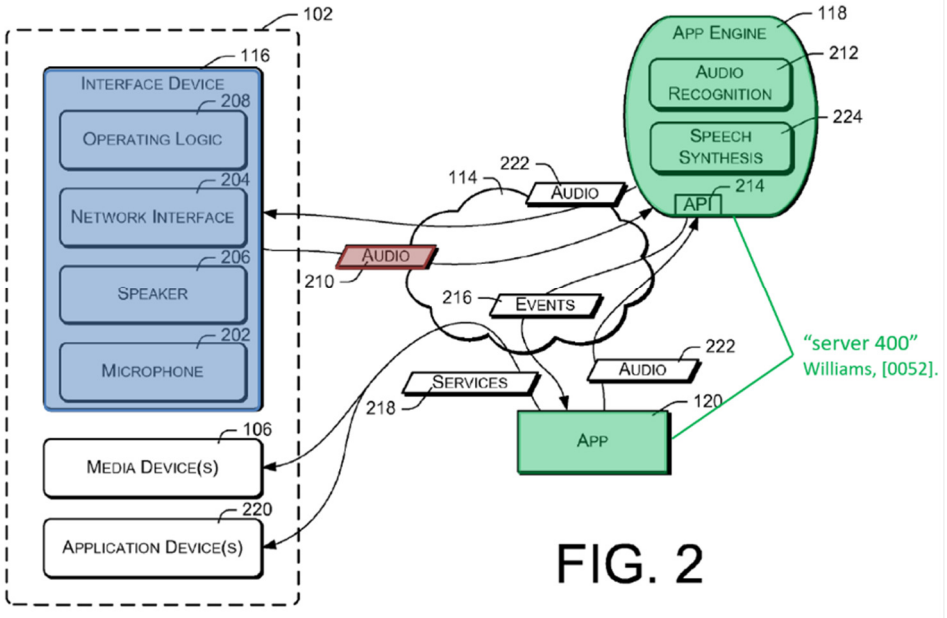


FIG. 2

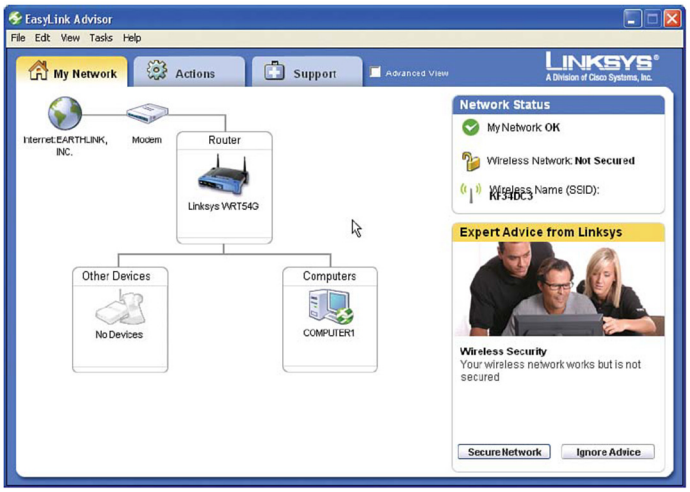
371. And any of the human voice commands disclosed in Williams that cause sending of a control message in the recited manner discloses the recited "***second human voice data***." (*Infra* § IX.I.3.j; *supra* § IX.I.3.f.) For example, "[a]n application might also be designed to allow a user to ***control various on-premises devices and equipment, such as media equipment and home-automation devices***. For example, such an application may accept ***spoken commands*** from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections." (EX1006, [0051].) Such ***spoken commands*** are "second human voice data." It would have been obvious to a POSITA that the ***second human voice data*** is similarly sent to the ***server***.

372. Williams discloses that the ***first device*** sends the captured human voice data (including the ***first human voice data*** and ***second human voice data***) to a ***server*** device over the ***Internet via the WLAN***. For example, with reference to FIG. 1, the audio data is sent by ***interface device 116*** to the router 112 via the in-home wireless network (*i.e.*, the WLAN) and to the ***server*** via the Internet (*e.g.*, network 114). Williams discloses "the ***interface devices 116*** communicate with and transmit environmental information to a cloud-based application service or engine 118 [of ***server 400***], ***via***

the router 112 and the network 114." (EX1006, [0020]; *see also supra* § IX.I.3.a.)

373. As discussed, a WLAN network is characterized by a local access point (*e.g.*, router) that broadcast wireless signals to wireless-enabled devices. (*Supra* §§ IX.A.2.a, IX.I.3.a; EX1007, 11, 100.) Thus, a POSITA would have recognized that Williams discloses a WLAN network in the ***home 102***. In other words, Williams' ***first device*** and controlled devices (*e.g.*, second devices) are connected to a WLAN network via a router. As discussed above, a router accesses a server over the Internet. (*Supra* §§ IX.I.3.a, IX.A.2.g.) Indeed, this was the common configuration for a home network. (*Id.*; EX1007, 11, see also, *Id.*, 100.) For example, "[f]igure 7-28 shows what the initial ***network*** should look like if things are working properly, including ***connectivity to the Internet***." (EX1007, 96; *see also id.*, Figure 7-28.)

Figure 7-28 EasyLink Graphical View of Your Network



374. Thus, Williams' network architecture of sending data (*e.g.*, ***first*** and ***second human voice data***) from the first device to the server and receiving messages from the server to the first device via the in-home wireless network (WLAN) which routes data to and from the ***server*** via the Internet (network 114) discloses the recited architecture of "over the Internet via the WLAN."

375. I note that network 114 by itself also discloses both the WLAN and the Internet. "***The network 114*** may thus include various types of communication networks, including ***local-area networks***, wide-area networks, ***the Internet***, ***wireless networks***, cable television networks, telephone networks, cellular communications networks, etc." (EX1006, [0015]; s*upra* § IX.I.3.a.)

376. In FIG. 2, ***audio 210*** is sent via the wireless network 114. Williams discloses that the ***interface device 116***, the application engine 118, and the application 120 [of ***server 400***] all communicates over the same wireless network (*e.g.*, network 114). (*Supra* § IX.I.3.a.) Here, I map "audio 210" as the first human voice data, however, as discussed above, a POSITA would have understood that the "audio 210" can equally be mapped to include both the first and second human voice data, or two instances of "audio 210" can be mapped to the first and second human voice data. (*Supra* §IX.I.3.f.)

h. 1[G]: "receive, from the server device via the WLAN over the Internet, first and second messages, in response to the sending of the respectively captured first and second human voice data;"

377. Both Williams and Lim disclose the ***first device*** receiving, from the ***server*** device via the WLAN over the Internet, ***first*** and ***second message***, in response to the sending of the respectively captured ***first*** and ***second human voice data***. (*Supra* § IX.A.2.h.)

378. First, Williams discloses that the first device receives a ***first message*** from the ***server*** in response to the sending of the captured ***first human voice data***. William discloses, with reference to FIG. 2, "API 214 of the application engine 118 [run on ***server 400***] may also allow the application 120 [run on ***server 400***] to provide ***audio 222*** for eventual transmission to the ***interface device 116***, to be played on the speaker 206." (EX1006, [0037].) "The ***audio*** may be in the form of text, in which speech synthesis or text-to-speech conversion may be performed by the application engine 118 or by the operating logic 108 of the interface device 116. In FIG. 2, the application engine 118 is illustrated as having a speech synthesis module 224. Alternatively, the ***audio 222*** may be provided from the application 120 as a digital audio stream." (EX1006, [0037].)

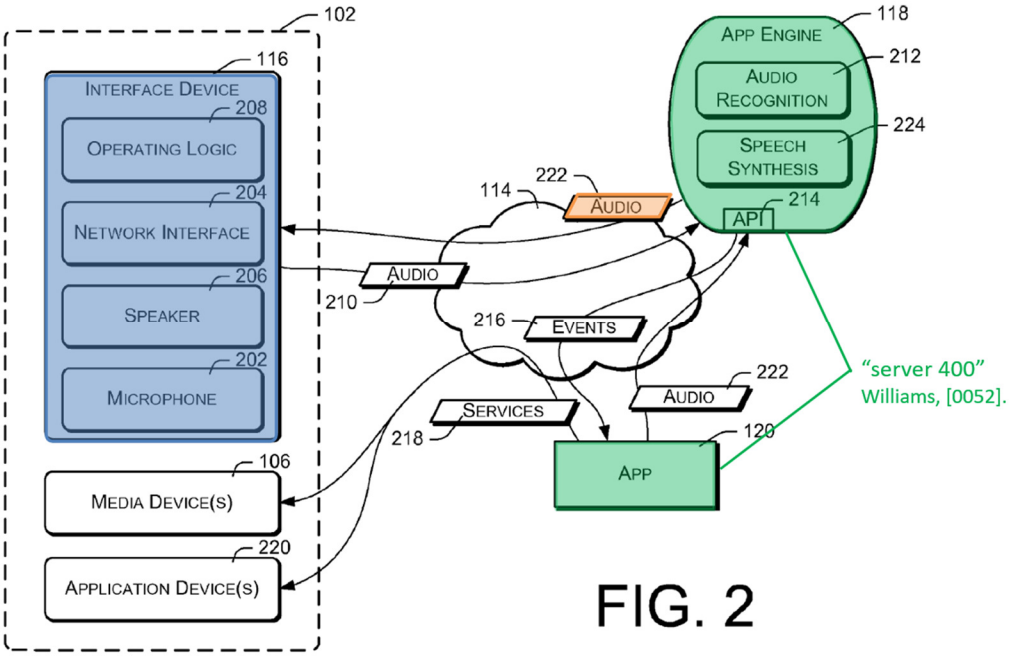


FIG. 2

379. Williams teaches that the ***first message*** is received by the ***first device*** from the ***server*** "in response to the sending of the captured ***first human voice data***" as recited. As discussed, Williams discloses the first device sending the captured ***first*** and ***second human voice data*** to the ***server***. (*Supra* § IX.I.3.g.) Williams teaches that the server device, ***server 400***, processes the captured human voice data and the ***first message*** is sent to and received by the first device in response. "[T]he application engine ***118*** [run on ***server 400***] is configured to receive the ***audio-based information 210*** … the application engine 118 [run on ***server 400***] ***processes*** it to identify and create event, status, and/or command data relating to in-home activities". (EX1006, [0027].) "The application engine 118 [run on ***server 400***] may keep a record of data

objects or event descriptors … Event descriptor objects may be created to identify or indicate the following types of information: ***spoken commands uttered by users*** [.]" (EX1006, [0027] - [0028].) "The application engine 118 [run on ***server 400***] ***processes*** the audio-based information to determine various data about user activities, status, environment, commands, etc. This data can then be obtained by authorized applications 120, which can ***act on the information to provide or enhance services to the user 104***." (EX1006, [0022].) For example, "the application engine may include audio recognition logic 212, such as speech-to-text conversion logic and other sound recognition functionality to recognize various types of sounds and media . . . [T]he application engine 118 exposes an application interface 214 (also referred to as an application programming interface or API) to the cloud-based applications 120. The API 214 allows an application 120 to obtain event descriptors that have been created by the application engine 118." (EX1006, [0034]-[0035].) "The API 214 of the application engine 118 may also allow the application 120 to provide ***audio 222*** for eventual transmission to the ***interface device 116***, to be played on the speaker 206." (*Id.*, [0037].) Further, Williams discloses that "the cloud-based service [running on the ***cloud server***] is responsive via the one or more application interfaces to provide ***audio*** to the ***on-premises audio monitoring devices*** at least partly in response to the cloud-based applications,"

or "at least partly in response to the network-based applications." (EX1006, Claims 9 and 16.) It would have been obvious that the ***first message*** is received "in response to the sending of the captured ***first*** . . . ***human voice data***" as recited because the ***first message*** is generated as the result of speech processing of the first human voice data.

380. Williams discloses several examples of providing to the user an ***audio response*** (*i.e.*, "first message") generated by the ***cloud server*** in response to the sending of the captured first human voice data. For example, Williams discloses that "[a]n application may provide general information in response to users' uttered questions. A user may ask a question, such as 'What's the recipe for deviled eggs?' The application may respond by emailing or otherwise providing a recipe for deviled eggs. In some cases, the application may use the speech synthesis functionality of the system to provide ***audio answers to spoken questions***." (EX1006, [0048].) As another example, "such an application may allow the user to inquire about a flight or other event. The system might respond by sending information to the user via a device associated with the user, or by ***audibly responding*** using speech synthesis." (EX1006, [0050].) As another example, Williams discloses "a cloud-based application may allow users to order tickets without needing a computer. A user may tell the application what movie they want to see. The application may use pre-configured information about the

user (such as location) to determine which theater is near the user. The application may respond to a spoken command to order movie tickets by using the speech synthesis functionality of the system to ***query the user regarding movie selections and times***." (EX1006, [0045].) A POSITA would have understood that the audio response or information associated with the audio response sent to the first device disclose the "first message".

381. Further, as discussed, the "audio 210" can equally be mapped to include or be both the first and second human voice data because "audio 210" results in both first and second messages. An example of providing to the user an ***audio response*** such as ***audio 222*** (*i.e.*, "first message") generated by the ***cloud server*** in response to the sending of the captured first human voice data is a confirmation of receipt of human voice data input.

382. Williams discloses that the ***first message*** is received by the ***first device*** from the ***server*** device via the WLAN over the Internet in response to the sending of the captured ***first human voice data***. With reference to FIG. 1, Williams discloses "the in-home devices connect to ***a router or access point 112***, which is in turn connected to ***a public network 114***. Connection between in-home devices and the router 112 may be through wired or ***wireless technologies***." (EX1006, [0012].) Thus, it would have been obvious that the ***audio 222*** is received by the "***router or access point 112***" from the ***cloud server***

via ***the Internet*** or "***public network 114***" and then wirelessly transmitted to the ***interface device*** via the home wireless network (*i.e.*, the WLAN). (EX1006, [0012]; *supra* § IX.I.3.g.)

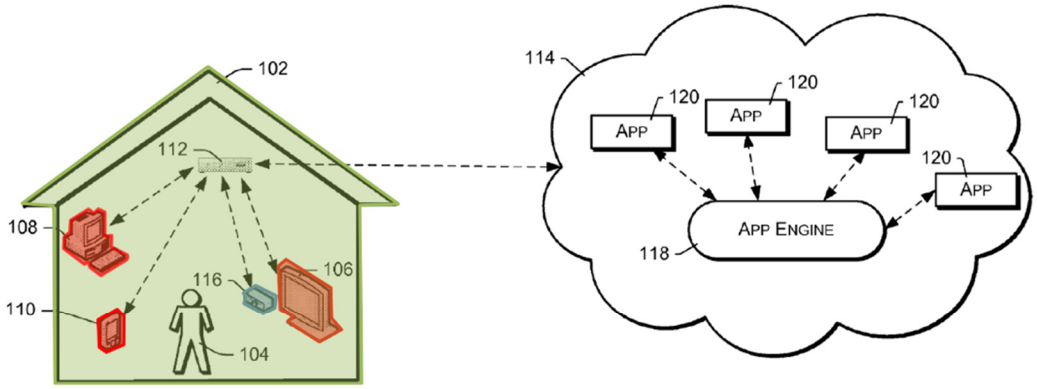


FIG. 1

383. Further, as shown in FIG. 2, ***audio 222*** is sent by the ***cloud server*** via the ***wireless network 114*** (*i.e.*, "the Internet") and the ***interface device 116*** receives the ***audio 222*** via the ***wireless network 114***, which by itself includes both WLAN and the Internet and thus discloses receiving data "via the WLAN over the Internet". (*See also supra* § IX.I.3.a.; EX1006, [0015].)

384. Williams discloses receiving, from the ***server device via the WLAN over the Internet***, a ***second message*** in response to the sending of the captured ***second human voice data***. The ***second human voice data*** (*e.g.*, "spoken commands from a user regarding on-premises devices") would be processed in

the same way as shown in FIG. 2. With reference to FIG. 2, Williams discloses that "[t]he application 120 [of the ***server***] may use the information from the event notifications 216 to initiate or customize ***services 218*** on behalf of users associated with the home 102. Such ***services*** may be provided through any of the ***media devices 106*** located in the home, as well as through other ***application devices 220, including computers 108, personal devices 110, and other devices available to users***." (EX1006, [0036], *see also*, *id.*, [0021], [0038], [0040].)

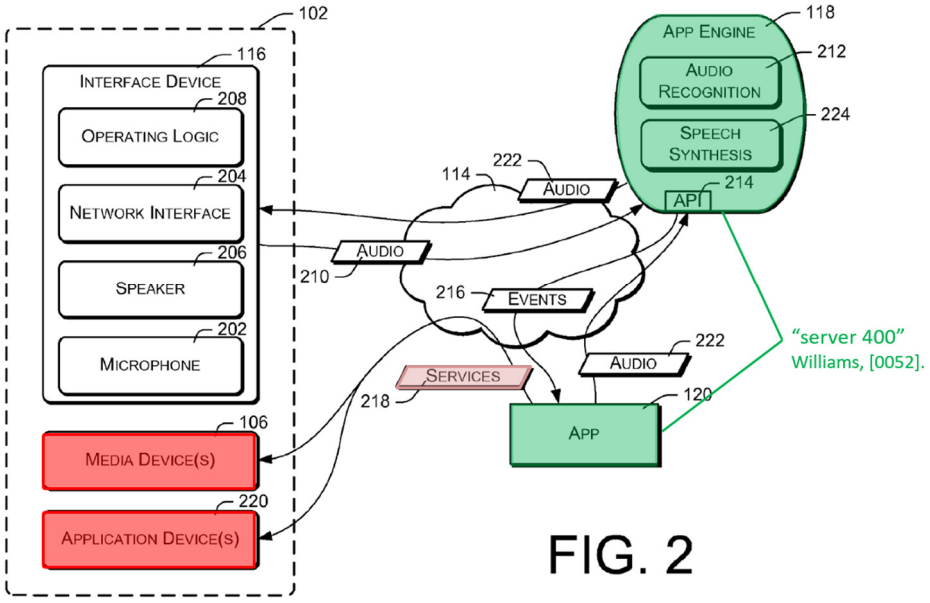


FIG. 2

385. Similarly, with reference to FIG. 3, "[a]t 310, the applications [of the ***server***] provide and/or enhance their ***services*** based on the notifications. For example, applications may act in response to received commands, or may provide ***certain information*** to a user that relates to the detected environment within the user premises." (EX1006, [0042].)

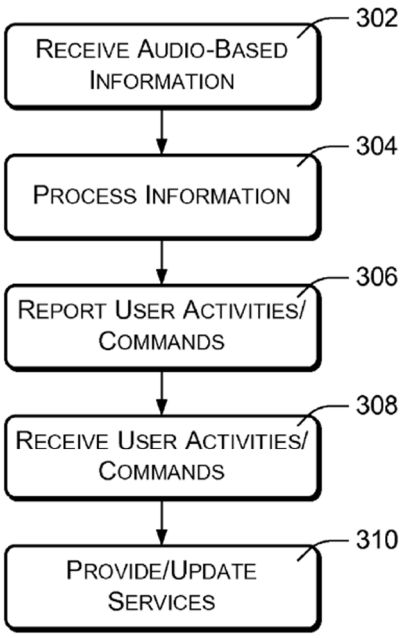


FIG. 3

386. A POSITA would have understood that providing *services 218* to a second device (*e.g.*, ***media devices 106 and/or application devices 220***) involves sending ***information*** related to *services 218* (*e.g.*, a control message) to the second device. Williams discloses that providing services includes "providing information." (EX1006, [0042].) Further, FIG. 2 expressly depicts that providing *services 218* involves transmitting information by the application 120 [of the ***server***] to the ***media devices 106 and/or application devices 220***. Further, Williams discloses that "*[s]ervices* may also be provided in conjunction with other

types of systems, including ***telephone systems, media players, automation systems and devices***, and so forth. Examples of applications 120 and associated *services 218* will be described below." (EX1006, [0036].) A POSITA would have understood that, in order to operate a second device, ***information*** related to the *service 218* (*e.g.*, a control message) is sent to and received by the second device. It was common knowledge that wirelessly controlling a device involves sending some information to the device wirelessly. For example, a control signal would be provided to a coffee maker appliance to turn off the appliance. "Similarly, the processing device 150 may provide a ***control signal to turn off the coffee maker appliance 110*** when the coffee maker appliance 110 reaches a 'warming an empty pot' state." (EX1019, 3:19-22.) As another example, a control signal would be provided to turn the appliance on. "In like manner, when the processing device 150 determines that the coffee maker appliance 110 has completed the 'brew state', it may ***communicates a control signal to turn on other appliances, such as the clock radio 140***." (EX1019, 6:23-27.) "A method for remotely processing a locally-generated voice command to control a local device [comprises] … generating a command signal . . . ***transmitting the command signal*** to the local device at the local site ***to effect a change in a state of the local device***." (EX1008, claim 1; *see also*, EX1012, [0012].)

387. As shown in FIGs. 1 and 2, William's ***media devices 106 and application devices 220*** receives ***services 218*** (the second message) from ***server 400*** over the wireless network 114 (*e.g.*, the Internet) and over the in-home WLAN (via router 112 in FIG. 1). (*See also* EX1006, [0012], [0013], [0052]-[0056].) As discussed, Williams' network architecture of sending data and receiving messages from the router 112 via the in-home wireless network (WLAN) which routes data/messages to and from the ***server*** via the Internet (network 114) discloses the recited architecture of "via the WLAN over the Internet."

388. To the extent that Williams does not disclose that the ***first device*** receives its ***second message***, Lim discloses its ***first device*** receives a ***first message*** and a ***second message*** from the ***server*** in response to the sending of the captured ***first human voice data*** and ***second human voice data***. (*Supra* § IX.A.2.h.) It would have been obvious to a POSITA that the Williams-Lim combination discloses this element. (*Supra* § IX.I.2.) In the Williams-Lim system, the ***server 400*** would send information (*e.g.*, related to ***services 218***) to the ***interface device 116*** (rather than the ***second devices***), which would then transmit commands to the ***second devices*** as taught by Lim, resulting in a more simplified, reliable, and secure system that is easier to scale, troubleshoot, and manage.

389. Again, to the extent that the first and second human voice data are captured/processed together (*e.g.*, as parts of one query) (which is not conceded), a POSITA would find it obvious that "audio 210" can include multiple parts, such as a trigger phrase ("***first human voice data***") and the instruction ("***second human voice data***"), consistent with common implementations. As shown in FIG. 2, processing of audio 210 triggers both first and second messages (***audio 222*** and ***services 218***). For example, the first human voice data triggers sending of "***audio 222***" ("***first message***") and the second human voice data triggers sending of information related to ***services 218*** ("***second message***"). (*Infra* §§IX.I.3.i-IX.I.3.l.)

i. 1[H]: "sound, using the sounder, digital audio content in response to the received first message;"

390. Both Williams and Lim disclose this element. (*Supra* § IX.A.2.i.)

391. Williams discloses sounding digital audio data using a ***sounder*** in the ***first device*** in response to the received ***first message***. First, Williams discloses a ***sounder***/***speaker*** in the ***first device***. As discussed, Williams discloses a ***speaker*** in the ***first device***. Williams discloses that "the ***interface device*** may also have a ***speaker*** or ***other form of audio output***." (EX1006, [0016].) With reference to FIG. 2, William discloses that "[t]he ***interface***

device may also include a ***speaker 206***, and operating logic 208." (EX1006, [0023].)

392. Williams discloses sounding digital audio data using a ***sounder/speaker*** in the ***first device*** in response to the received ***first message***. Specifically, Williams discloses that, "[t]he API 214 of the application engine 118 may also allow the application 120 to provide ***audio 222*** for eventual transmission to the ***interface device 116***, ***to be played on the speaker 206***. The audio may be in the form of text, in which speech synthesis or text-to-speech conversion may be performed by the application engine 118 or by the operating logic 108 of the interface device 116. In FIG. 2, the application engine 118 is illustrated as having a speech synthesis module 224. Alternatively, the ***audio 222*** may be provided from the application 120 as a digital audio stream." (EX1006, [0037].) It would have been obvious that digital audio data is sounded using the ***sounder*** in the ***first device*** in response to the ***audio 222*** (first message) received from the server 400.

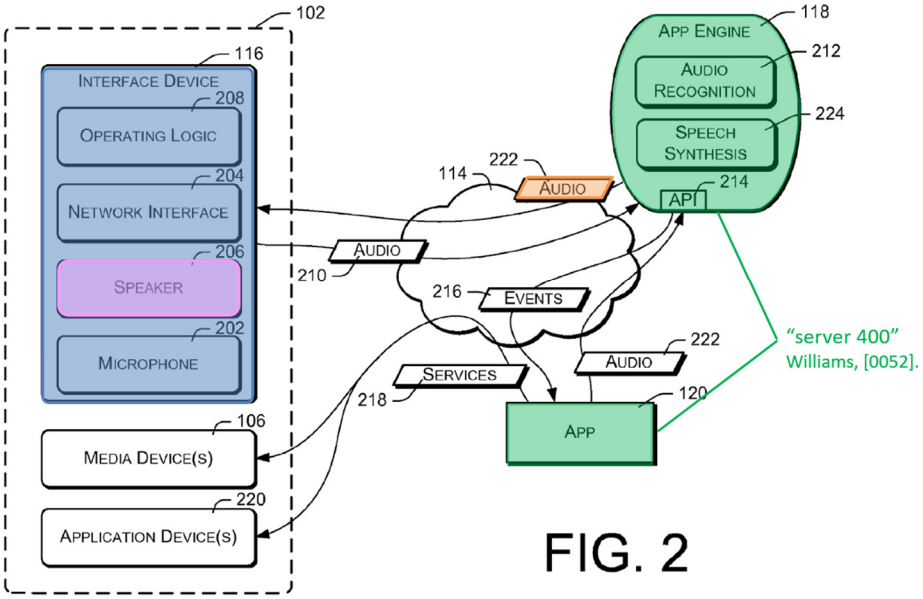


FIG. 2

393. As discussed above, a speaker is a sounder. (*Supra* §§ IX.A.2.d, IX.I.3.d; EX1001, 105:48-52, 105:61-106:4; EX1014, 763; EX1018, 122.) Thus, a POSITA would have understood that Williams discloses using a sounder (*e.g.*, ***speaker 206***) in the first device (*e.g.*, ***interface device 116***) to sound digital audio data in response to the received first message (*e.g.*, ***audio 222***) in the recited manner.

394. Williams discloses several exemplary use cases of sounding digital audio data using the ***speaker***. For example, "[t]he architecture above can be utilized to provide a wide variety of services and functionality, of which the following are simple examples." (EX1006, [0043].) For example, Williams discloses that "[a]n application may provide general information in response to users' uttered questions. A user may ask a question, such as 'What's the recipe

for deviled eggs?' The application may respond by emailing or otherwise providing a recipe for deviled eggs. In some cases, the application may use the speech synthesis functionality of the system to provide ***audio answers to spoken questions***." (EX1006, [0048].)

395. As another example, Williams states that "such an application may allow the user to inquire about a flight or other event. The system might respond by sending information to the user via a device associated with the user, or by ***audibly responding*** using speech synthesis." (EX1006, [0050].) As another example, Williams discloses "[a] user may tell the application what movie they want to see. The application may use pre-configured information about the user (such as location) to determine which theater is near the user. The application may respond to a spoken command to order movie tickets by using the speech synthesis functionality of the system to ***query the user regarding movie selections and times***." (EX1006, [0045].)

j. 1[I][1]: "and send, to a second device over the WLAN, a control message that comprises a control data for controlling a first actuator in the second device"

396. Williams-Lim discloses a ***first device*** sending, to the ***second device*** over the ***WLAN***, a ***control message*** that comprises a control data for controlling a first actuator in the second device. (*Supra* § IX.A.2.j.)

397. First, Williams discloses a ***second device***. Williams discloses "***control[ling]*** various ***on-premises devices and equipment***, such as ***media equipment*** and ***home-automation devices***." (EX1006, [0051]; *see also* claim 20.) Williams discloses that the server 400 "may accept spoken commands from a user regarding on-premises devices, and may ***respond*** by interacting with ***such devices*** through available network connections." (EX1006, [0051].) Williams discloses that "services may be provided through any of the ***media devices 106*** located in the home, as well as through other ***application devices 220, including computers 108, personal devices 110, and other devices available to users***". (EX1006, [0036]; *see also id.*, [0042].) "Cloud-based or network-based applications such as the applications 120 may communicate and ***operate*** in conjunction with a wide variety of network-connected devices, including ***entertainment devices, game boxes, media devices, home automation and security devices, communications devices***, and so forth." (EX1006, [0021].) These are examples of the ***second device***.

398. Williams' ***second devices*** have actuators. One example of the "first actuator" is an actuator providing visual information (*e.g.*, displays). As discussed, the '721 patent explains that a display is an actuator. (*Supra* § IX.A.2.j; EX1001, 40:51-41-7, *see also, id.*, 25:33-38, 109:41-46, 110:60-66, 144:64-145:1.) Williams teaches that "[w]ithin the home 102 is a user 104 and

several devices … [t]he illustrated devices include a ***media device 106***, representing any one or more of various types of devices such as ***televisions, audio players, video players***" within the home. (EX1006, [0011]; *see also id.*, [0009]-[0010], [0054].) Williams also discloses that "[t]he applications 120 may be web-based applications that are accessible to users via a ***network browser***, operating in conjunction with the ***computer 108*** or another Internet terminal or access device." (EX1006, [0021].) These controlled devices have displays. As discussed above, a display is or has an actuator. (*Supra* § IX.A.2.j; EX1004, 396; EX1025, FIG. 15.) Thus, a POSITA would have understood that Williams' ***media devices 106*** comprise actuators to provide visual information.

399. Another example of the recited "first actuator" is an actuator providing sound. As discussed, a speaker is an actuator. (*Supra* § IX.A.2.j; EX1001, 105:48-52, 105:61-106:4; EX1018, 122; EX1014, 763.) Williams' controlled devices include ***media device 106*** such as ***televisions, audio players, video players***. (EX1006, [0011]; *see also id.*, [0010].) Williams discloses "services may be provided through any of the ***media devices 106*** located in the home." (EX1006, [0036].) For example, Williams discloses providing "content items that are supplied to consuming devices such as ***electronic reader 104***." (EX1006, [0054].) A POSITA would have understood that

Williams' ***media devices 106*** comprise speakers to produce and control sound. Thus, Williams' ***media devices 106*** comprise actuators.

400. Another example of the "first actuator" is a motion actuator. Williams discloses that "[c]loud-based or network-based applications such as the applications 120 may communicate and ***operate*** in conjunction with a wide variety of network-connected devices, including entertainment devices, game boxes, media devices, ***home automation and security devices***, communications devices, and so forth." (EX1006, [0021]; *see also id.*, [0036] and [0051].) A POSITA would have understood that home automation and security devices (*e.g.,* automated doors or windows, automated lights, automated HVAC systems, automated cameras, automated locks, automated kitchen appliances) have motion actuators that can be operated. As discussed, at the time of the invention, it was known that heating and cooling systems are integrated with motors and fans. (*Supra* § IX.A.2.j; EX1020, 401, Figs. 13-1 and 13-3.)

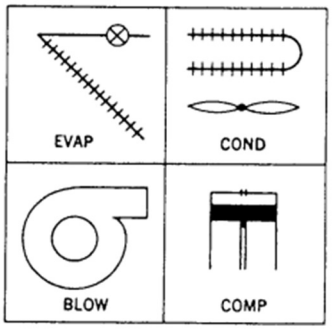


Fig 13-1. Self contained air conditioner

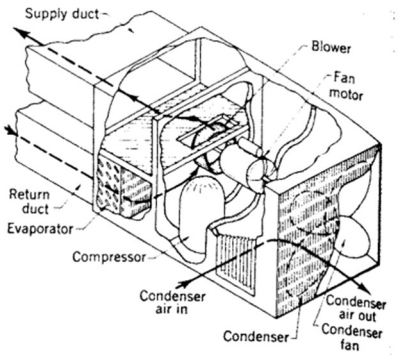


Fig 13-3. Horizontal self contained air conditioned unit

401. Further, consistent with the Board's Decision on Appeal affirming the Examiner's rejections of claims 1-30 of the parent patent '710, an air conditioner is a controlled device with actuators. "As the Examiner indicates, Abe's in-building systems in Figure 15 include image display device 51 and ***air conditioner 52 that are controlled devices with actuators***. *See* Ans. 6; Abe ¶ 95." (EX1004, 396; EX1025, FIG. 15.) Thus, Williams' ***home automation and security devices*** comprise actuators.

402. Williams discloses sending a ***control message*** that comprises a control data for controlling a first actuator to the ***second device***. As discussed, with reference to FIG. 2, Williams depicts sending information related to ***services 218*** by the application 120 of the ***server*** to the ***media devices 106 and/or application devices 220***. (EX1006, [0042].) Williams discloses that "[t]he application 120 [of the server] may use the information from the event notifications 216 to initiate or customize ***services 218*** on behalf of users associated with the home 102. Such ***services*** may be provided through any of the ***media devices 106*** located in the home, as well as through other ***application devices 220, including computers 108, personal devices 110, and other devices available to users***." (EX1006, [0036], *see also*, *id*., [0021], [0038], [0040].) Williams discloses that providing ***services*** includes "provid[ing] . . . information." (EX1006, [0042].) Further, FIG. 2 expressly depicts that providing ***services 218***

involves transmitting information by the application 120 of the ***server*** to the ***media devices 106 and/or application devices 220***.

403. As discussed, a POSITA would have understood that the information comprises a ***control message*** comprising control data to control the ***second devices***, consistent with common implementations. It was common knowledge that wirelessly controlling a device involves sending some information to the device wirelessly. For example, a control signal would be provided to a coffee maker appliance to turn off the appliance. "Similarly, the processing device 150 may provide a ***control signal to turn off the coffee maker appliance 110*** when the coffee maker appliance 110 reaches a 'warming an empty pot' state." (EX1019, 3:19-22.) As another example, a control signal would be provided to turn the appliance on. "In like manner, when the processing device 150 determines that the coffee maker appliance 110 has completed the 'brew state', it may ***communicate[] a control signal to turn on other appliances, such as the clock radio 140***." (EX1019, 6:23-27.) "A method for remotely processing a locally-generated voice command to control a local device [comprises] … generating a command signal . . . ***transmitting the command signal*** to the local device at the local site ***to effect a change in a state of the local device***." (EX1008, claim 1; *see also*, EX1012, [0012].) As discussed, Williams sends control message over the WLAN. A POSITA would have found it obvious that

data transmitted over WiFi—such as Williams' information—includes packets of information (*e.g.*, characters or bits) structured or arranged according to the WiFi protocol such that the receiving device can parse, validate, and interpret the content. Thus, Williams' information/command signals are messages. This is also consistent with the '721 patent. (EX1001, 148:37-42.)

404. Further, a POSITA would have understood that "***control[ling]*** various ***on-premises devices and equipment***, such as ***media equipment*** and ***home-automation devices***" involves controlling various "actuators" in media equipment and home-automation devices, for example, to playback a media item, or change the settings or turn on/off the home automation devices. As discussed, Williams discloses a variety of services that involve controlling actuators in ***second devices***. For example, "[s]ervices may also be provided in conjunction with other types of systems, including ***telephone systems, media players, automation systems and devices***, and so forth. Examples of applications 120 and associated services 218 will be described below." (EX1006, [0036].)

405. To the extent that Williams does not disclose that the ***control message*** is sent from the ***first device***, Lim discloses that the ***first device*** sends a ***control message*** to the ***second device*** in response to the received ***second message*** over the ***WLAN***. (*Supra* § IX.A.2.j.) It would have been obvious to

a POSITA that the Williams-Lim combination discloses this element. (*Supra* § IX.I.2.)

406. As discussed, Williams' ***first device*** and ***second devices*** are connected to a WLAN network via a router. (*Supra* §§ IX.I.3.a, IX.I.3.j.) As discussed in a textbook titled "Wireless Home Networking Simplified", "with a wireless LAN, the computers have to find a way to connect, or associate, themselves to the LAN, because they are not physically connected. The way they do this is by ***tuning themselves to the wireless router*** using the name of the router . . . ***Once a computer has connected, or associated, itself to the access point, it becomes part of the LAN . . . Now the computer can send and receive information to and from other computers on the LAN***[.]" (EX1007, 11, see also, *Id*., 100.) A POSITA would have understood that by connecting to the same WLAN network, the ***first device*** and ***second devices*** would communicate with each other. (EX1007, 11, 100.) Thus, in Williams-Lim, Williams' ***first device*** would send the ***control message*** to a ***second device*** over the ***WLAN***.

k. 1[I][2]: "to directly or indirectly affect, change, or produce, a first physical phenomenon,"

407. Williams-Lim discloses this. (*Supra* § IX.A.2.k.) As discussed above, Williams discloses controlling motors and actuators in the second device (*e.g.*, ***entertainment devices, game boxes, media devices, home automation and***

security devices, communications devices). A POSITA would have found it obvious that controlling Williams' ***speakers/home automation devices/media devices*** involves controlling actuators affecting, changing, producing, or creating a first physical phenomenon. (*Supra* §§IX.I.3.j, IX.A.2.j, IX.A.2.k.)

408. As discussed, a speaker is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, audio or sound. (*Supra* §§IX.I.3.j, IX.A.2.j, IX.A.2.k.) A display is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, light. *Id*. HVAC-related motors are configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, motion, temperature, or humidity. *Id*.

l. 1[I][3]: "in response to the received second message,"

409. In Williams-Lim, information related to ***services 218*** (second message) is sent to the first device. (*Supra* §IX.A.2.l.) Upon receiving the information, the first device then transmits the ***control message*** to the ***appliances*** or ***second device*** (*e.g.*, ***media devices 106 and/or application devices 220***) to operate the second device's actuators ("first actuators"). (*Supra* §IX.I.3.j.) A POSITA would find it obvious that the ***control message*** is or is derived from the information related to ***services 218*** (second message).

(EX1019, 3:19-22, 6:23-27; EX1008, Claim 1; EX1012, [0012].) Thus, it would have been obvious that the ***control message*** is sent in response to the received second message.

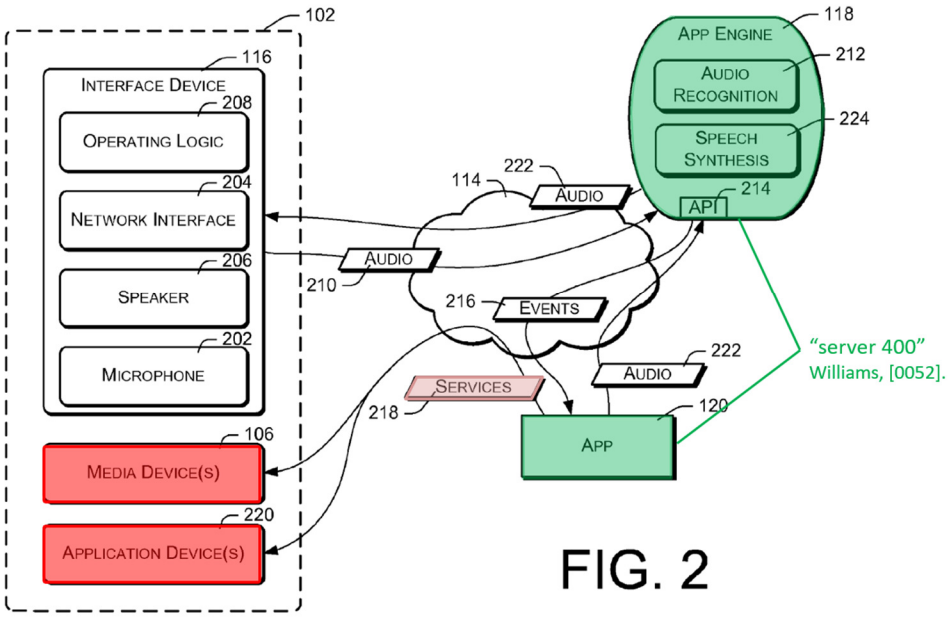


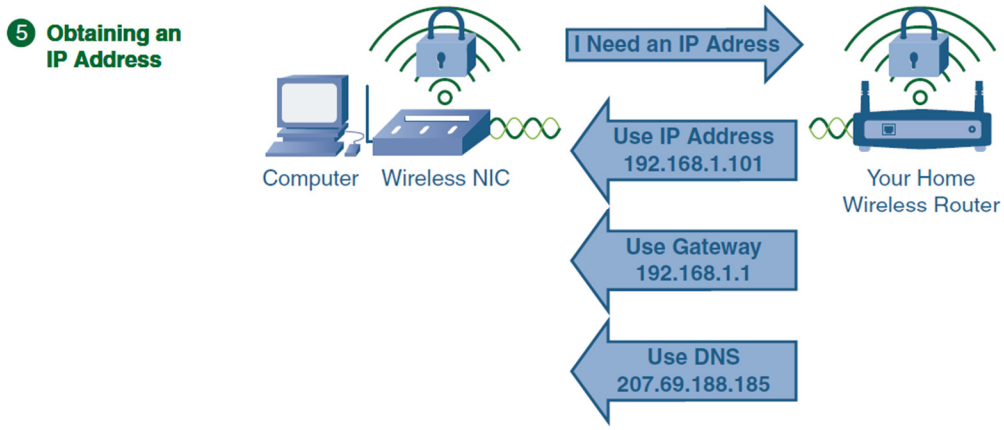
FIG. 2

m. 1[J]: "wherein each of the first and second devices is addressable in the WLAN and in the Internet using a respective Internet Protocol (IP) address."

410. Both Williams and Lim disclose this element. (*Supra* § IX.A.2.m.)

411. As discussed, Williams discloses that each of its first device and second device operates in a WLAN (*e.g.*, home network formed by router 112), which is connected to the Internet. (*Supra* §§ IX.I.3.a, IX.I.3.j) A POSITA would have understood that this wireless network involves associating each device with a device identifier or address (such as an IP address). Indeed, at

the time of the invention, it was well-known that networked devices are associated with IP addresses. The IP address may be an assigned IP address for the device (*e.g.*, by the router), the router's IP address, a combination of the two, or a globally unique IP address. "***Each computer or device that you want to join your wireless network needs a wireless network interface card (NIC)***." (EX1007, 99.) "If the wireless router grants access, it then typically ***assigns the NIC an IP address*** to use on the wireless network, at which point we can begin exchanging data," "***each device (computer, laptop, and so on) in a home network needs to have an IP address***, which is how the device is known on the network . . . Most home networks use IP addresses in the range of 192.168.1.100 to 192.168.1.149, and the ***router assigns each device an address*** from this pool when it boots." (EX1007, 108, 210, *see also*, *id.*, 175.)



4. Claim 5: "The first device according to claim 1, further comprising a second actuator that is configured for directly or indirectly affect, change, or produce, a second physical phenomenon."

412. As discussed, Lim itself discloses ***LEDs*** in the ***client device***. (*Supra* §§IX.A.3.) Lim discloses "a ***controlling device 100*** for use in commanding the functional operations of one or more appliances may include, . . . , a means 202 to provide feedback to the user (*e.g.*, ***one or more LEDs***)[.]" (*Id.*; EX1005, [0019].) It would have been obvious that the same description applies equally to ***smart device 110***.

413. The '721 patent states that LEDs are actuators. The '721 patent explains an "actuator may be a light source that emits visible or non-visible light (infrared, ultraviolet, X-rays, or gamma rays) such as for illumination or indication . . . The light source may be an electric light source for converting electrical energy into light . . . The electric light source may be based on Solid-State Lighting (SSL) such as a Light Emitting Diode (LED)[.]" (EX1001, 39:10-21.)

414. As discussed, both Williams and Lim are analogous art, from the same field of endeavor, and all relate to the same technical problems of processing and responding to human speech commands (*e.g.*, to provide services and/or to control appliances). Providing a light (*e.g.*, a second actuator) in the device that turns on/off and changes color to prompt user to provide speech input, as taught by Lim, can be combined with the teachings of Williams to improve the usability of Williams' interface device, with a reasonable

expectation of success. The technique of incorporating an indicator light in the audio interface device disclosed in the Williams-Lim combination yields the predictable result of improving usability of the device and serves as an additional way to prompt users to provide the "first human voice data" and "second human voice data" to control the device. Because both Williams and Lim relate to capturing human voice data for controlling devices or providing services, a POSITA would be motivated to use the known techniques of Lim to improve the usability of Williams' first device in the same way by enabling users to understand when the device requires additional voice input. Further, it would have been obvious to try combining the teachings of Williams and Lim. There are only a finite number of identified and predictable solutions for prompting the user to provide speech input, such as visual cues, auditory cues, and tactile cues. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Lim has successfully demonstrated that devices may include a light for prompting users to provide speech input, and thus, there would have been a reasonable expectation of success for combining Williams and Lim.

415. Thus, Williams-Lim discloses or at least suggests that its second actuator may be an LED indicator.

5. Claim 6: "The first device according to claim 5, wherein the second physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

416. As discussed, Williams-Lim's first device further comprises an LED light indicator (*e.g.*, a second actuator). The '721 patent confirms that LED lights are configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, light. (*Supra* §IX.A.4.)

417. Thus, Williams-Lim discloses this claim.

6. Claim 9: "The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

418. Both Williams and Lim disclose this claim. (*Supra* §IX.A.5.) A POSITA would have understood that these claims recite a standard microphone. (EX1001, 17:47-54.) For example, as discussed, Williams discloses a standard microphone for capturing audio data. (*Supra* § IX.I.3.c; EX1006, [0016].) As discussed, a POSITA would have understood that a standard microphone is an omnidirectional microphone based on sensing an incident sound-based motion of a diaphragm. (EX1014, 757, 309.) Further, a POSITA would understand that dynamic microphone and piezoelectric microphone are standard and common implementations of microphone. (EX1014, 758.) Thus, a POSITA

would understand (or it would be obvious) that Williams' microphone is the recited microphone in claim 9.

7. Claim 10: "The first device according to claim 1, wherein the IP address is stored in a volatile or non-volatile memory."

419. Both Williams and Lim disclose the IP address is stored in a volatile or non-volatile memory. (*Supra* § IX.A.6.) For example, Williams discloses that "[t]he ***interface device*** may also include a speaker 206, and operating logic 208. The operating logic 208 may be implemented as one or more ***programmable processing units, associated memory***, and appropriate instructions stored in the memory and executed by the progressing unit." (EX1006, [0023].)

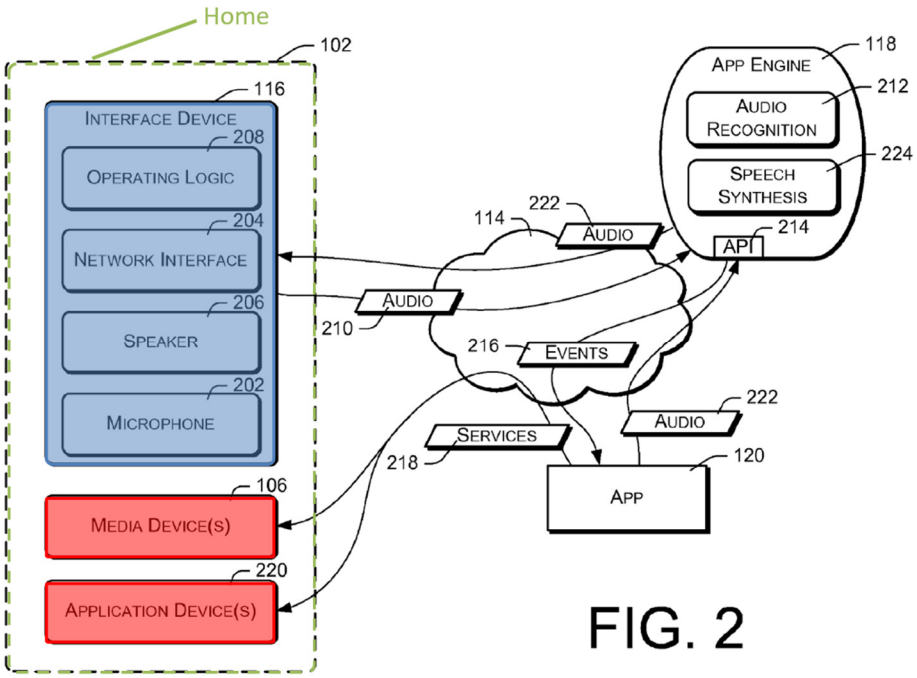
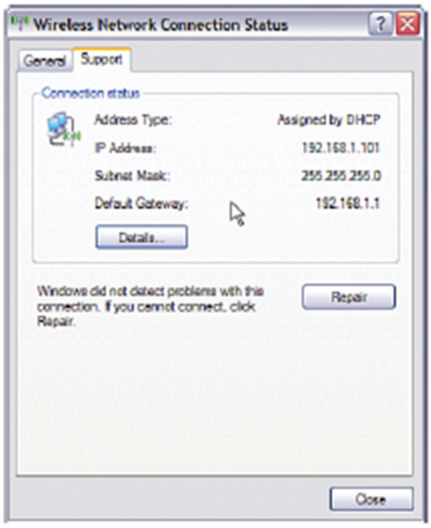


FIG. 2

420. As discussed above, Williams discloses that the first device is addressable in the Internet and WLAN with an IP address and that it communicates via WLAN and with remote server over the Internet via a wireless router. (*Supra* §§ IX.I.3.a, IX.I.3.j, IX.I.3.m; EX1006, [0012] – [0013], [0020], Figure 1; EX1005, [0027], [0038], FIG. 5b.)

421. A POSITA would have understood that the memory in the ***interface device*** as disclosed in Williams can be used to store the associated IP addresses. For example, as discussed above and shown below, networked devices can display the IP addresses, which means the IP addresses are locally stored in the device. (*Supra* § IX.A.6; EX1007, 211, Figure 10-44, *see also, Id.*, 265.)

Figure 10-44 Verifying IP Address Assignment with Windows XP



422. Thus, a POSITA would have understood that the "associated memory" as disclosed in Williams' first device (*e.g.*, ***interface device***) locally store the IP addresses.

8. Claim 11: "The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type."

423. Both Williams and Lim disclose or suggest this claim. (*Supra* § IX.A.7.) As discussed, Williams discloses that the first device is addressable in the Internet and that it communicates via WLAN and with remote server over the Internet via a wireless router. (*Supra* §§ IX.I.3.a, IX.I.3.g, IX.I.3.m, IX.A.2.a, IX.A.2.g, IX.A.2.m; EX1006, [0012] – [0013], [0020], Figure 1; EX1005, [0027], [0038], FIG. 5b.) Thus, Williams suggests, and it would be obvious in view of POSITA knowledge that Williams' networked device is associated with an MAC address that is of MAC-48, EUI-48, or EUI-64 type. "***Each computer or device that you want to join your wireless network needs a wireless network interface card (NIC)***" and "***each wireless card has a unique identifier called a MAC address***." (EX1007, 99, 166.) "Most MAC addresses use one of the IEEE standardized protocols; ***MAC-48, EUI-48®, and EUI-64®***." (EX1026, 2:50-59.)

9. Claim 12: "The first device according to claim 1, wherein the IP address is a static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address."

424. Both Williams and Lim disclose or suggest this claim. (*Supra* § IX.A.8.) For example, as discussed above, Williams discloses that the first/controlling device is addressable in the Internet with an IP address and that it communicates via WLAN and with remote server over the Internet via a wireless router. (*Supra* §§ IX.I.3.a, IX.I.3.g; EX1006, [0012] – [0013], [0020], Figure 1; EX1005, [0027], [0038], FIG. 5b; EX1007, 99, 166, 210.)

425. As discussed, at the time of the invention, it was well-known that Internet Protocol (IP) addresses are either static or dynamic, and are IPv4 addresses. (*Supra* § IX.A.8.) "In the case of today's Internet, these addresses are Internet protocol version 4 (IPv4) addresses." (EX1029, 127.) "Internet protocol (IP) addresses are the unique numbers used to identify individual connections to the Internet." *Id.* "The broadband networks typically assign a *dynamic IP address* to your broadband service. Dynamic means the IP address can change[.]" (EX1007, 277.) Further, IP address may be a "*static IP address*, which means . . . your address never changes." *Id.*

10. Claim 13: "The first device according to claim 1, wherein the WLAN is according to, or is compatible with, an Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard."

426. Both Williams and Lim disclose or renders obvious this claim. (*Supra* § IX.A.9.) Williams discloses WLAN and that the interface device 116

and home devices (media devices 106 and application devices 220) communicate with a wireless network (*e.g.*, via a router). (*Supra* § IX.I.3.a, IX.I.3.j, IX.I.3.m; EX1006, [0012] – [0013], [0020], Figure 1.) Williams expressly discloses usage of "other wireless data communications technologies" between interface device and application engine and "various network communications between the components of FIG. 2 [which includes the client device, the controlled device, and the server device] may be performed using existing networking protocols and techniques." (EX1006, [0020], [0012], [0013], [0038].)

427. A POSITA would understand that WLAN over IEEE 802.11a/b/g are common wireless technologies in or before 2012. (EX1040, 600-601; EX1013, 7:48-51.) Thus, a POSITA would understand that Williams discloses or suggests the WLAN is according to, or based on, Institute of Electrical and Electronics Engineers (IEEE) 802.11a, IEEE 802.11b, IEEE 802.11g.

11. Claim 14: "The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

428. Both Williams and Lim disclose or renders obvious this claim. (*Supra* § IX.A.10.) Williams discloses or suggests WLAN that is according to, or based on, Institute of Electrical and Electronics Engineers (IEEE) 802.11b

and a POSITA would understand that WLAN over 802.11b is a form of ISM Band II frequency. (*Supra* § IX.I.10; EX1040, 22-23, Table 1.2.)

12. Claim 15: "The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution."

429. Both Williams and Lim disclose or renders obvious this claim. (*Supra* § IX.A.11.) For example, Williams discloses that "***interface devices*** might use ***cellular*** or other wireless data ***communications technologies*** to communicate with the application engine 118 [of ***server 400***]," and "some in-home devices may communicate with the ***network 114 using cellular data technologies, or through other types of wired and wireless technologies that do not involve an in-home router or access point***." (EX1006, [0020], [0013].)

430. A POSITA would understand that 3G, the third-generation cellular network, GSM (Global System for Mobile Communications), and EDGE (Enhanced Data rates for GSM Evolution) are common and known cellular communications network in or before 2012. (EX1040, 12, 13, 599-600.) Further, wireless technology commonly included cellular connectivity in 2012.

Thus, a POSITA would understand (or it would be obvious that) Williams' device is configured to communicate over the recited cellular telephone network.

13. Claim 17: "The first device according to claim 1, wherein the second device is integrated in, is part of, or is entirely included in, a household appliance having a primary function."

431. Both Williams and Lim disclose a controlled device that is integrated in, is part of, or is entirely included in, an appliance. (*Supra* § IX.A.12.) As discussed, the controlled devices in the Williams system are "various ***on-premises devices and equipment***, such as ***media equipment*** and ***home-automation devices***" which can include devices such as "***televisions, audio players, video players***, and so forth." (*Supra* § IX.I.3.j; EX1006, [0051], [0011].) These devices are appliances.

14. Claim 21: "The first device according to claim 17, wherein the primary function of the appliance is associated with environmental control, and the appliance is part of a Heating, Ventilation and Air Conditioning (HVAC) system."

432. Williams-Lim discloses this claim. Williams discloses controlling "***home-automation devices***." (EX1006, [0051].) As discussed, Lim discloses controlling various appliances, such as "televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, ***fans, HVAC systems, thermostats***, personal computers, etc." (*Supra* § IX.A.2.j; EX1005, [0017].) These disclose

appliances whose primary functionality is associated with environmental control. It would have been obvious to combine Williams and Lim. (*Supra* § IX.I.2.)

15. Claim 22: "The first device according to claim 21, wherein the primary function of the appliance is associated with temperature control, and wherein the appliance is an air conditioner or a heater."

433. Williams-Lim discloses this claim. As discussed, Williams discloses controlling "***home-automation devices***" and Lim discloses controlling various appliances, such as "***fans, HVAC systems, thermostats***[.]" (*Supra* §§ IX.A.2.j, IX.I.3.j; EX1006, [0051]; EX1005, [0017].) A POSITA would have understood that fans, HVAC systems, and thermostats are appliances associated with temperature control. Further, "HVAC" stands for heating, ventilation, and air conditioning and thus an HVAC system comprises an air conditioner or a heater. It would have been obvious to a POSITA that the Williams-Lim combination discloses this claim. (*Supra* § IX.I.2.)

16. Claim 24: "The first device according to claim 17, wherein the appliance is an answering machine, a telephone set, a home cinema system, a High Fidelity (HiFi) system, a Compact Disc (CD) or Digital Video Disc (DVD) player, an electric furnace, a trash compactor, a smoke detector, a light fixture, or a dehumidifier."

434. Both Williams and Lim disclose this claim. (*Supra* §§IX.A.15, IX.A.2.a; EX1005, [0017].) For example, as discussed, the controlled device in the Williams system are devices such as "***televisions, audio players, video players***". (*Supra* § IX.I.3.j; EX1006, [0011].) These disclose a home cinema

system. Further, it would have been obvious to a POSITA that Williams' "video players" and "audio players" include DVD players and CD players. Indeed, CDs has been released commercially as early as 1982. (*See* EX1014, 746.)

17. Claim 25: "The first device according to claim 1, further configured to operate within a building, and wherein the server device is external to the building."

435. Both Williams and Lim disclose this claim. (*Supra* §§IX.A.16.) For example, Williams discloses that the first device is configured to operate within a building. As discussed above, Williams discloses that the ***first device*** is operated within a ***building***. Williams discloses that "[t]he user's ***home 102*** may be equipped with one or more on-premises ***audio monitoring devices 116***, referred to herein as ***in-home audio interface devices***." (EX1006, [0016], [0023].) The "***audio interface devices 116*** can be placed within the ***home 102*** so that their microphones detect ambient sound or noise within the ***home 102***." (EX1006, [0019].)

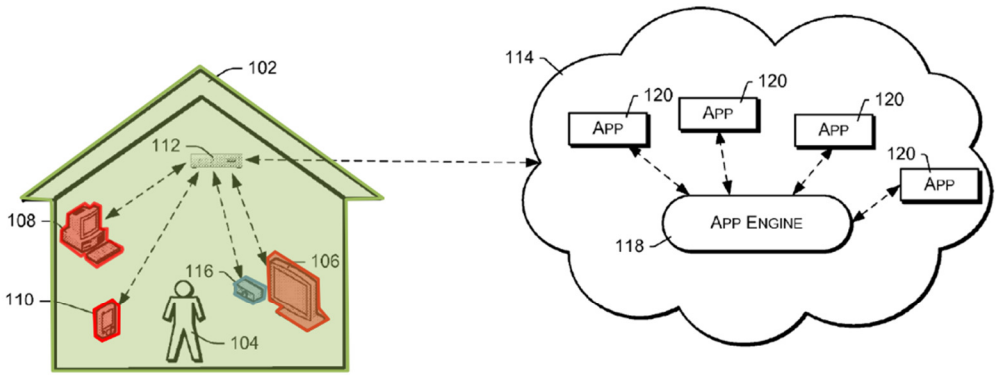


FIG. 1

436. Similarly, with reference to FIG. 2, Williams discloses that "***in-home audio interface devices 116*** may be located within the ***home 102***." (EX1006, [0023].)

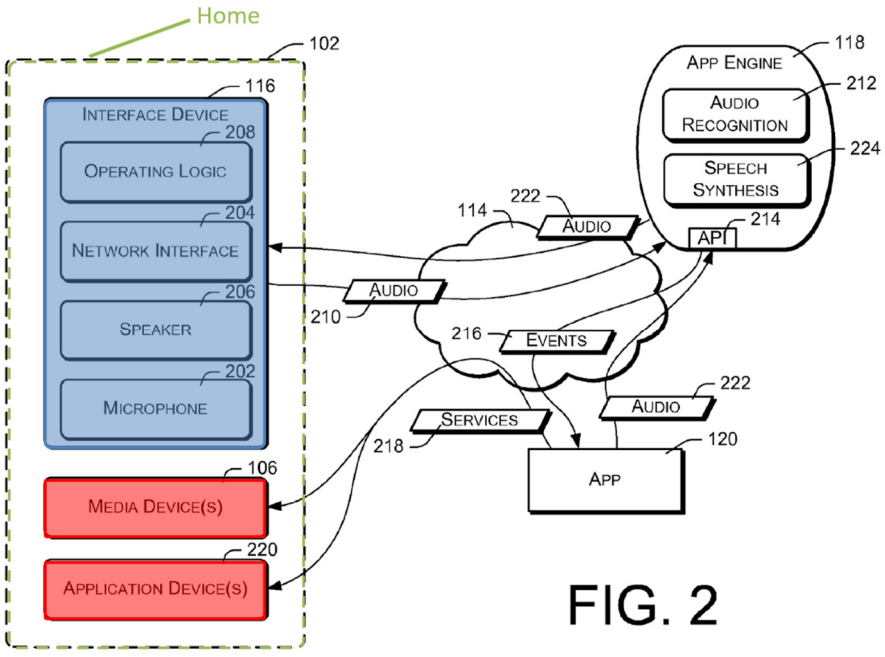


FIG. 2

437. Further, Williams discloses that the server device is external to the building. Williams discloses "a ***server 400*** that may be used to implement the functionality of the application engine 118 and/or any of the applications 120." (EX1006, [0052].)

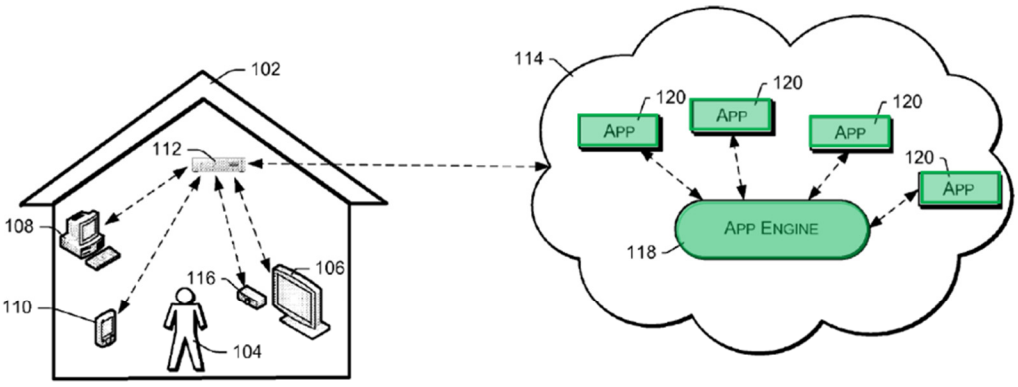


FIG. 1

438. Williams discloses that its ***server 400*** implementing the application engine 118 and applications 120 is external to the ***building*** (*e.g.*, the user premise 102 in FIG. 1). For example, Williams states that "[s]ervers may be located together or separately, and organized as virtual servers, ***server banks, and/or server farms***. The described functionality may be provided by the servers of a single entity or enterprise, or may utilize the servers and/or services of multiple entities or enterprises." (EX1006, [0052]; *see also id.*, [0020] – [0021].) At the time of the invention, a POSITA would have understood that

these "server banks" and the "server farms" disclosed in Williams are outside a residential and that they are connected to the home via the Internet. Indeed, at the time of the invention, server farms are understood to be located in remote facilities outside residential homes. "In recent years large investments have been made to build ***data centers***, or ***server farms***, purpose-built facilities providing storage and computing services[.]" (EX1015, 277.) "[T]he 10th busiest web site in the world, Wikipedia, is hosted on 350 servers having 856 cores, spread across three ***data centers***." (EX1015, 282.) "Data centers and server clusters represent a highly visible 'hot spot' with respect to energy consumption. In a ***server farm***, a larger number of high-performance server computers are collocated in one facility. ***Electricity is needed for both powering the server computers and for cooling the data center facility***." (EX1016, 1760.)

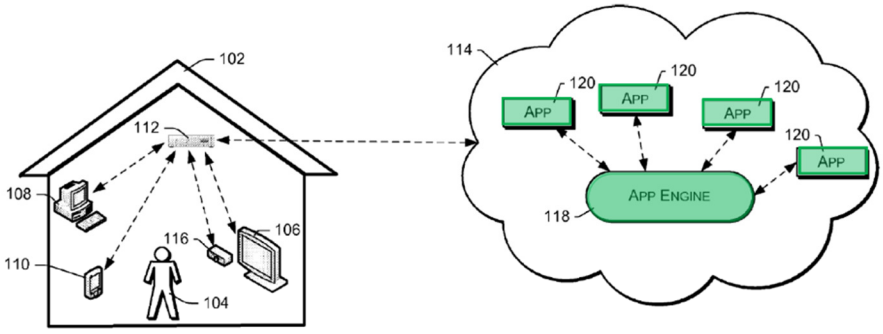


FIG. 1

18. Claim 26: "The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays."

439. Williams-Lim discloses this claim. As discussed, Lim discloses this claim. (*Supra* §IX.A.17.) It would have been obvious for a POSITA to combine Williams and Lim. (*Supra* §§ IX.I.2, IX.I.4.)

19. Claim 27: "The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode."

440. Williams-Lim discloses this claim. As discussed, Lim discloses this claim. (*Supra* § IX.A.18.) It would have been obvious for a POSITA to combine Williams and Lim. (*Supra* §§ IX.I.2, IX.I.4.)

20. Claim 28: "The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

441. Claim 28 depends on claim 27. As discussed, Lim discloses this claim. (*Supra* § IX.A.19.) Besides comprising an electric light source, Williams-Lim's device comprises a sounder that plays synthesized digital audio content as recited in claim 28. For example, as discussed above, Williams teaches "provid[ing] audio 222 for eventual transmission to the interface device

116, to be played on the ***speaker 206***." (*Supra* § IX.I.3.i, IX.I.3.h; EX1006, [0037].) Further, "[t]he audio [222] may be in the form of text, in which ***speech synthesis or text-to-speech conversion*** may be performed by the application engine 118 or by the operating logic 108 of the interface device 116." (EX1006, [0037], *see also id.*, [0045] and [0050].)

442. Further, Williams-Lim's sounder stimulates the voice of a human being and sounds a syllable, a word, a phrase, or a sentence using male or female voice. For example, Williams discloses that "the application may respond to a spoken command to order movie tickets by using the ***speech synthesis functionality*** of the system to ***query the user regarding movie selections and time***." (EX1006, [0046].) Thus, Williams-Lim discloses that the sounding of digital audio data by the sounder comprises at least playing digital audio content that is synthesized, simulating the voice of a human being, and sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.

21. Claim 29: "The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker.

443. Both Williams and Lim disclose or suggest this claim. (*Supra* §IX.A.20.) For example, Williams discloses a standard speaker for sounding

digital audio content. As discussed, Williams discloses that its interface device comprises a ***speaker***. (*Supra* §IX.A.2.d; EX1006, [0016], [0023].) As discussed, a POSITA would have understood that an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker, a ribbon or planar magnetic loudspeaker are standard and common implementations of speakers. (*Supra* §IX.A.20.) Thus, a POSITA would have understood (or it would be obvious) that Williams' speaker is the recited loudspeaker in claim 29.

22. Claim 31: "The first device according to claim 1, further comprising an enclosure that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors."

444. Both Williams and Lim disclose this claim. (*Supra* §IX.A.22.) For example, Williams discloses an ***enclosure*** that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors for the ***first device***.

445. As discussed, Williams's first device comprises a WLAN transceiver, a sounder, a microphone, and one or more processors. (*Supra* §§ IX.I.3.b, IX.I.3.c, IX.I.3.d, IX.I.3.e.) For example, with reference to FIG. 2, "[a]n ***individual interface device 116*** may include a ***microphone 202*** that is configured to detect ambient noise, sounds, and speech. The interface device 102 may also include a ***network interface 204*** that facilitates network communications with various entities, including the application engine 118 and

one or more applications 120, via the network 114. The interface device may also include a ***speaker 206***, and operating logic 208. The operating logic 208 may be implemented as one or more ***programmable processing units***, associated memory, and appropriate instructions stored in the memory and executed by the processing unit. Other types of logic may also be used to configure the interface device 116 to perform the functions described here." (EX1005, [0023].) As shown in FIGs. 1 and 2 below, Williams' ***interface device 102*** houses the WLAN transceiver, sounder, microphone, and processors within one enclosure.

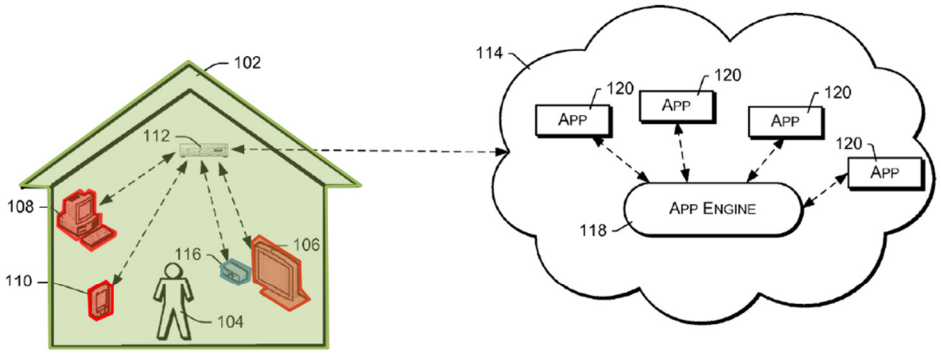


FIG. 1

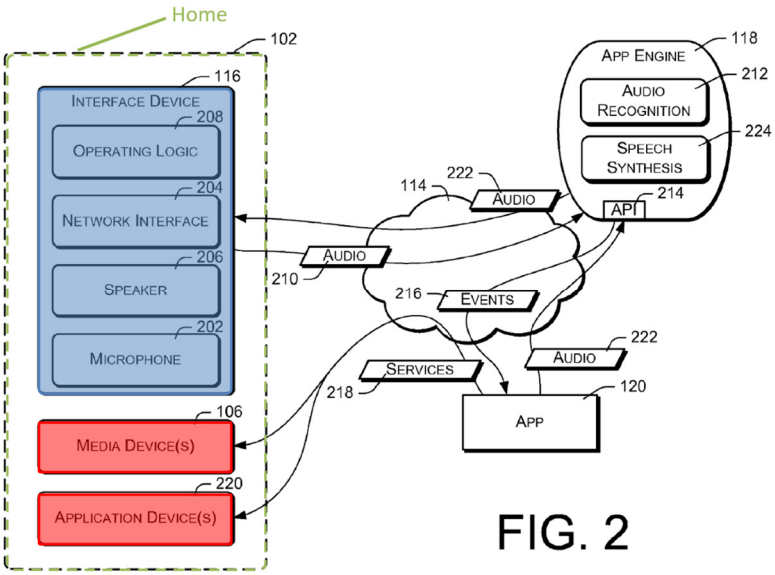


FIG. 2

23. Claim 33: "The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves."

446. Both Williams and Lim disclose or suggest this claim. (*Supra* §IX.A.23.) As discussed, Williams discloses a standard speaker for sounding digital audio content. (*Supra* §IX.A.2.d; EX1006, [0016], [0023].) As discussed, a POSITA would have understood that a standard speaker is a sounder for converting an electrical energy to omnidirectional pattern emitted audible sound waves. (EX1018, 122; EX1014, 763, 765.) Thus, a POSITA would understand (or it would be obvious) that Williams' speaker is a sounder as recited in claim 33.

J. Ground 10: Williams, Lim and Wittenberg Render Obvious Claim 2

1. Claim 2: "The first device according to claim 1, wherein the first device or the server device is configured for processing the first and second human voice data using a voice recognition algorithm for identifying a voice of a specific person."

447. Claim 1 is obvious in view of Williams-Lim. (*Supra* § IX.I.3.) Claim 2 is obvious in further view of Wittenberg.

448. Wittenberg discloses claim 2. (*Supra* § IX.B.2.)

449. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Wittenberg. Both Wittenberg and Williams-Lim relate to remotely processing audio information such as speech for controlling devices or providing services through devices. For example, Williams discloses that "[t]he application 120 may use the information from the event notifications 216 to initiate or customize services 218 on behalf of users associated with the home 102. Such services may be provided through any of the media devices 106 located in the home, as well as through other application devices 220, including computers 108, personal devices 110, and other devices available to users." (EX1006, [0036].) Lim relates to "controlling devices such as universal remote controls" and "specific appliances to be controlled." (EX1005, [0007].) Lim discloses that "the ***controlling device*** may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device . . . The resulting spoken syllables may be digitized and presented to a speech

recognition means for processing . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027].) Wittenberg discloses "a method for remotely processing a locally-generated voice command to control a local device" and furthermore preventing unauthorized persons from controlling the local devices. (*See* EX1008, Claim 1.) Thus, Williams-Lim and Wittenberg are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same problem of controlling devices based on audio commands and ensuring security within the command/control system architecture.

450. A POSITA would have been motivated to combine Williams-Lim with Wittenberg to include the voice recognition of Wittenberg to achieve greater security and with a reasonable expectation of success. Williams discloses authorization and authentication safeguards to address security concerns. "Authorization and authentication safeguards are employed to ensure that only entities authorized by users are able to access this information." (EX1006, [0008].) Wittenberg teaches additional processing techniques, for example, "the method comprising: (a) receiving at a local site an acoustic signal and generating an audio signal from the acoustic signal; (b) transmitting the audio signal to a remote site; (c) ***performing speech recognition processing*** on

the audio signal at the remote site to identify a voice command from the audio signal; (d) ***performing voice recognition processing*** on the audio signal at the remote site to determine whether the audio signal has been supplied by an authorized user[.]" (EX1008, Claim 1.) The voice recognition processing techniques can be combined with the teachings of Williams-Lim to yield a command/control system architecture that provides added security by determining whether an authorized person issued the voice command, with a reasonable expectation of success, as Lim and Williams already perform server-side speech recognition, making the addition a straightforward software modification.

451. Wittenberg describes a security issue that unauthorized users may be able to control devices via audio commands. "[F]or example, if an unauthorized person gains access to the portion of the system that receives voice commands, that person can instruct the system to perform unauthorized commands which may be harmful." (EX1008, 2:21-30.) Wittenberg expressly teaches the benefits for including voice recognition during processing to address the issue and provide enhanced security. "Voice recognition processing provides an ***additional degree of security***, which prevents an unauthorized person at the local site from affecting or manipulating the local controllable device in some manner." (EX1008, 5:61-6:6.)

452. Wittenberg recognizes the benefits of applying its techniques in numerous circumstances. "There are a numerous circumstances in the government/military, commercial, and private sectors where it would be advantageous to control a local device with a locally-generated voice command without placing the demands of speech processing on a local device and with the capability to provide a significant degree of security in the process of interpreting and carrying out voice commands." (*Id.*, 2:25-31.) Specifically, Wittenberg recognizes the benefits of applying its techniques in home appliances. "For example, in the case where the local controllable device is a television, voice recognition processing can prevent an unauthorized person (*e.g.*, a child) from accessing certain channels or ordering pay-per-view shows, whereas an authorized person (*e.g.*, a parent) can be authorized to access all channels by issuing voice commands via a telephone or the like." (EX1008, 6:7-13.)

453. The Williams-Lim-Wittenberg combination results in a system where voice recognition technique as taught by Wittenberg is incorporated in Williams-Lim's server to analyze human audio commands in Williams-Lim. For example, the Williams-Lim-Wittenberg combination would determine if the person issuing the voice data is authorized to make the command as a part of processing at the server (application engine 118) as taught by Wittenberg. This

combination would predictably enhance the security of the system because only authorized users would be able to control devices through their voice commands. The known benefits of Wittenberg would prompt the application of Wittenberg in the same field based on design incentives and market forces and yield the predictable result of improving the security of the control and command systems and prevent control by unauthorized users.

454. Because both Williams-Lim and Wittenberg includes voice data processing and device control, a POSITA would be motivated to use the known algorithm and techniques of voice recognition processing of Wittenberg to improve the security of Williams-Lim in the same way. Further, it would have been obvious to try combining the teachings of Wittenberg and Williams-Lim. There are only a finite number of identified and predictable solutions for user authentication, such as authorization via biometrics like facial recognition, fingerprint, or making use of contextual data like device location and time to supplement the voice data processing. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Wittenberg has successfully demonstrated that performing voice recognition can be implemented to improve security of the system and prevent unauthorized control, and thus, there would have been a reasonable expectation of success for

combining Williams-Lim and Wittenberg. Moreover, at the time of the invention, algorithms and processing techniques for performing voice recognition were known. "Algorithms and processing techniques for performing voice recognition are well known . . . ." (EX1008, 6:1-3.)

K. Ground 11: Williams, Lim and Gruber Render Obvious Claim 3-4, 11

1. Claim 3: "The first device according to claim 1, further comprising a sensor that outputs sensor data that responds to a physical phenomenon, wherein the computer program instructions, when executed, further cause the first device to send to the server device, via the WLAN over the Internet, the sensor data, and wherein the first message is further responsive to the sent sensor data."

455. Claim 1 is obvious in view of Williams-Lim. (*Supra* § IX.I.3.) Claims 3-4, 11 are obvious in further view of Gruber.

456. Gruber discloses claim 3. (*Supra* § IX.C.2.)

457. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Gruber. Williams-Lim and Gruber both relate to processing and responding to human speech commands. Williams discloses that "the application engine 118 is configured to receive the audio-based information 210. Depending on the type or format of the received information, the application engine 118 processes it to identify and create event, status, and/or command data relating to in-home activities." (EX1006, [0027].) "The application 120 may use the information from the event notifications 216 to

initiate or customize services 218 on behalf of users associated with the home 102." (EX1006, [0036].) Lim relates to "controlling devices such as universal remote controls" and "specific appliances to be controlled." (EX1005, [0007].) Lim discloses that "the ***controlling device*** may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device . . . The resulting spoken syllables may be digitized and presented to a speech recognition means for processing . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027].) Gruber discloses that the "invention relates to virtual assistants, and more specifically to mechanisms for improving interpretation and processing of commands provided to such an assistant." (EX1009, [0007].) Thus, Williams-Lim and Gruber are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problems of processing and responding to human speech commands (*e.g.*, to provide services).

458. A POSITA would have been motivated to combine Williams-Lim with Gruber because Gruber teaches sending contextual information (*e.g.*, temperature, location), along with speech commands, to the server to improve server-side speech processing. Williams-Lim's controlling device already (1)

has built-in sensors (*e.g.*, GPS, accelerometer) and (2) sends voice commands to the server for processing. Gruber teaches additional speech processing techniques of using contextual information obtained from device sensors to clarify the user's intent, reduce numbers of candidate interpretations, and/or constrain solutions during speech recognition and natural language processing. Thus, adding more sensors and sending sensor data to the server, as taught by Gruber, would be a straightforward, predictable improvement with a reasonable expectation of success. (EX1009, Abstract; EX1005, [0023].)

459. Further, Lim itself discloses that the client device, for example, would be a smart phone that includes sensors. Lim discloses "a representative smart device, illustrated in FIG. 3 in the form of a smart phone, may include . . . accessory features 338 such as a digital camera, GPS, accelerometer, etc." (EX1005, [0023].) Hence adding additional sensors in Williams-Lim to supply contextual information to assist speech processing would be a straightforward, predictable improvement with a reasonable expectation of success.

460. Specifically, the Williams-Lim-Gruber combination would result in a system where sensors (as taught by Gruber and contemplated by Lim) are incorporated in Lim's first device to capture sensor data (*e.g.*, temperature, location) to be transmitted to Williams-Lim' server to supplement audio processing as taught by Williams-Lim and/or Gruber. Williams-Lim discloses

a system that can capture sensor data (*e.g.*, audio, location information) using sensors (microphones, GPS) and transmit the captured sensor data to the server. Williams discloses that "the application engine 118 receives audio-based information from the interface devices 116," and "[a]n individual interface device 116 may include a microphone 202 that is configured to detect ambient noise, sounds, and speech." (*See* EX1006, [0022], [0023].) Gruber discloses the benefits of supplementing speech processing with captured sensor data and/or related information. "This helps to clarify the user's intent and to reduce the number of candidate interpretations of the user's input, and reduces the need for the user to provide excessive clarification input." (EX1009, [0013].) For example, the Williams-Lim-Gruber combination would comprise sensors for capturing sensor data in the audio interface device to be transmitted to Williams' server via the WLAN over the Internet to supplement processing, Williams' server (*e.g.*, server 400) can use the captured sensor data as contextual information to process human voice data as taught by Gruber to determine how to respond (*e.g.*, audio responses, services). to supplement speech processing. Combining the teachings of Williams-Lim and Gruber as such would yield predictable results of improving audio processing.

461. Incorporating Gruber's speech processing techniques into Williams-Lim yields the predictable result of improving accuracy of human

voice data processing and improving the usability of the control system, because the system is more intelligent and the users are not required to provide excessive clarification input. Because both Williams-Lim and Gruber relates to processing of human voice data to provide services, a POSITA would be motivated to use the known techniques of Gruber to improve the speech processing capabilities of Williams-Lim in the same way. Further, it would have been obvious to try combining the teachings of Williams-Lim and Gruber. There are only a finite number of identified and predictable solutions for processing voice data and improving accuracy of the processing, such as capturing higher-fidelity audio data or drawing from non-audio cues. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Gruber has successfully demonstrated that device sensors can gather contextual information that would be useful in supplementing speech processing, and thus, there would have been a reasonable expectation of success for combining Williams-Lim and Gruber.

462. Thus, it would have been obvious to a POSITA to make use of device sensor data to process human speech without requiring excessive clarifications from users. The known benefits of Gruber would prompt the

application of Gruber in the same field based on design incentives and other market forces.

2. Claim 4: "The first device according to claim 3, wherein the sensor is a thermoelectric sensor that responds to a temperature or to a temperature gradient of an object using conduction, convection, or radiation, or wherein the sensor is a photoelectric sensor that responds to a visible or an invisible light or gamma rays."

463. Claim 4 depends on claim 3. As discussed, Gruber discloses claim 4. (*Supra* § IX.C.3.)

464. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Gruber. (*Supra* § IX.K.1.)

3. Claim 11: "The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type."

465. As discussed, Gruber discloses claim 11. (*Supra* § IX.C.4.) Gruber discloses that a client device can have signatures such as "MAC addresses". (EX1009, [0218] - [0225].)

466. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Gruber and look to Gruber for implementation details for Williams-Lim's networked device and addressing schemes (*e.g.*, MAC addresses) with a reasonable expectation of success as Williams-Lim's first device is addressable in the Internet and that it communicates via WLAN and with remote server over the Internet via a wireless router. (*Supra* §§ IX.I.3.a,

IX.I.3.g, IX.I.3.m, IX.A.2.a, IX.A.2.g, IX.A.2.m, IX.K.1; EX1006, [0012] – [0013], [0020], Figure 1; EX1005, [0027], [0038], FIG. 5b.)

467. Incorporating a NIC with a unique MAC address into Williams-Lim's device would have been a routine and predictable design choice, because MAC addressing is a fundamental and ubiquitous feature of IEEE 802 networking and is required for ordinary device participation in standard WLANs.

468. A POSITA also would have been motivated by strong market incentives and product design requirements to include a NIC with a unique MAC address in Williams-Lim's device to ensure compatibility with the vast installed base of existing home routers, access points, and network infrastructure. Wireless consumer devices are expected to interoperate with standard Wi-Fi networks and MAC addressing is a universally adopted mechanism for network identification, association, filtering, and access control in such environments. Incorporating a unique MAC address would therefore predictably improve interoperability, simplify deployment, and allow the Williams-Lim device to function seamlessly with existing networks. These design incentives and market forces would have prompted a POSITA to apply the conventional MAC address teachings to Williams-Lim to meet consumer expectations for wireless connectivity and broad network compatibility.

L. Ground 12: Williams, Lim, and Rosenberger Render Obvious Claims 5-6, 26-28

1. Claim 5: "The first device according to claim 1, further comprising a second actuator that is configured for directly or indirectly affect, change, or produce, a second physical phenomenon."

469. Claim 1 is obvious in view of Lim. (*Supra* § IX.I.3.) Claims 5-6, 26-27 are obvious in view of Lim or in further view of Rosenberger.

470. As discussed, Lim discloses this claim. (*Supra* §§IX.A.3.)

471. Further, Rosenberger discloses its device comprises "a ***multi-colored LED status light 34*** ***for visually indicating various device states and modes***." (EX1013, 8:8-16.)

472. At the time of the invention, it would have been obvious to combine the teachings of Rosenberger with Williams-Lim. Both Rosenberger and Williams-Lim relate to processing and responding to audio information such as human speech for controlling devices or providing services through devices. (*See* EX1006, [0027] and [0036]; EX1005, [0007], [0027].) Rosenberger relates to "improvements in speech-recognizing control devices that respond to oral commands from a system user to produce a desired effect or result." (EX1013, 1:7-13.) Thus, Rosenberger and Williams-Lim are analogous to the '721 patent, from the same field of endeavor, and all relate to the same goal of processing and responding to audio commands.

473. Rosenberger discloses incorporating a multi-colored LED light in the device that responds to a speech trigger phrase received to "visually indicat[e] various device states and modes", for example by changing colors. (EX1013, 8:15-16.) Further, Lim itself discloses, "[t]his and subsequent prompts for voice input may take the form of an ***illuminated LED***." (EX1005, [0027]; *Supra* §§IX.A.2.h, IX.A.3.) The technique of incorporating a light (*e.g.*, a second actuator) in the device that turns on/off and changes color to indicate the device status and prompt user to provide speech input can be combined with the teachings of Williams-Lim to improve the usability of Williams-Lim' interface device, with a reasonable expectation of success.

474. Specifically, the Williams-Lim-Rosenberger combination results in a system that allows control of devices in response to processing of human voice data while incorporating the advantage of Rosenberger to signal or prompt users to provide voice input via its status light. For example, the Williams-Lim-Rosenberger combination would include Rosenberger's LED light indicator (*e.g.*, an additional actuator) in Williams-Lim's first device, which can be operated (*e.g.*, turns on/off and/or changes color) to indicate different device states and modes in response to the information received from the server, as taught by Lim/Rosenberger.

475. The technique of incorporating an indicator light in the audio interface device disclosed in the Williams-Lim-Rosenberger combination yield the predictable result of improving usability of the device as the changes of the light can indicate to a user the current state or mode the device is operating in, and particularly, serves as an additional way to prompt users to provide the "first human voice data" and "second human voice data" to control the device. Because both Williams-Lim and Rosenberger relates to capturing human voice data for controlling devices or providing services, a POSITA would be motivated to use the known techniques of Rosenberger to improve the usability of Williams-Lim' first device in the same way by enabling users to understand the shifts in device status and modes through the changes in the LED light, and to understand when the device requires additional voice input.

476. Further, it would have been obvious to try combining the teachings of Williams-Lim and Rosenberger. There are only a finite number of identified and predictable solutions for indicating to a user the device state or mode or prompt user to provide speech input, such as visual cues, auditory cues, and tactile cues. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Rosenberger has successfully demonstrated that devices may include a light for indicating device states and

modes and prompting users to provide speech input, and thus, there would have been a reasonable expectation of success for combining Williams-Lim and Rosenberger. The known benefits of Rosenberger would prompt its application in the same field based on design incentives and other market forces.

2. Claim 6: "The first device according to claim 5, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

477. As discussed, Williams-Lim's first device further comprises an LED light indicator (*e.g.*, a second actuator). The '721 patent confirms that LED lights are actuators configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, light. (*Supra* §IX.A.4.)

478. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Rosenberger. (*Supra* § IX.L.1.) For example, LED lights as taught by Lim/Rosenberger would be incorporated into Williams' first device to provide feedback to the user or indicate device states and modes in response to information received from the server.

3. Claim 26: "The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays."

479. Rosenberger discloses claim 26. (*Supra* § IX.H.2.)

480. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Rosenberger. (*Supra* § IX.L.1.)

4. Claim 27: "The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode."

481. Rosenberger discloses claim 27. (*Supra* § IX.H.3.)

482. Further, Lim itself discloses ***LEDs*** in the ***client device***. (EX1005, [0019]; *supra* §§IX.A.3.)

483. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Rosenberger. (*Supra* § IX.L.1.)

5. Claim 28: "The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

484. Claim 28 depends on claim 27. Both Williams and Lim disclose this claim. (*Supra* §§ IX.A.19, IX.A.2.i.)

485. Besides comprising an electric light source, Williams-Lim-Rosenberger's sounder plays a digital audio content that is pre-recorded or synthesized as recited in claim 28. For example, as discussed above, Williams teaches "provid[ing] audio 222 for eventual transmission to the interface device 116, to be played on the ***speaker 206***." (*Supra* § IX.I.3.i, IX.I.3.h; EX1006,

[0037].) Further, "[t]he audio [222] may be in the form of text, in which ***speech synthesis or text-to-speech conversion*** may be performed by the application engine 118 or by the operating logic 108 of the interface device 116." (EX1006, [0037], *see also id.*, [0045] and [0050].)

486. Further, Williams-Lim-Rosenberger's sounder stimulates the voice of a human being and sounds a syllable, a word, a phrase, or a sentence using male or female voice. For example, Williams discloses that "the application may respond to a spoken command to order movie tickets by using the ***speech synthesis functionality*** of the system to ***query the user regarding movie selections and time***." (EX1006, [0046].) Thus, Williams-Lim-Rosenberger discloses that the sounding of digital audio data by the sounder comprises at least playing digital audio content that is synthesized, simulating the voice of a human being, and sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.

M. Ground 13: Williams, Lim and Benesty Render Obvious Claims 7-8

1. Claim 7: "The first device according to claim 1, further comprising an additional microphone for capturing human voice data, and wherein the capturing comprises capturing, by the additional microphone, the first and second human voice data."

487. Claim 1 is obvious in view of Williams-Lim. (*Supra* § IX.I.3.) Claims 7-8 are obvious in further view of Benesty.

488. Benesty discloses claim 7. (*Supra* § IX.D.2.)

489. It would have been obvious to a POSITA to apply Benesty's techniques of including multiple microphones to Williams-Lim' first device. Both Williams-Lim and Benesty are related to capturing and processing human audio data. Williams-Lim is directed to hands-free audio communication of a user with a first device. For example, FIG. 1 of Williams illustrates "an operational environment in which applications may utilize an application engine, and may respond to ***audio from user premises*** to provide various services to users." (EX1006, [0003].) Lim discloses that "a ***controlling device*** may be provisioned with means to receive voice input from a user, *e.g.*, **a *microphone***". (EX1005, [0007].) Benesty discloses hands-free audio communication and speech processing techniques. Benesty states that "hands-free audio communication is now a major feature in mobile communication systems as well as audio and video conferencing systems. One problem that becomes evident to users of these systems is the decrease in communication quality due to the pickup of room reverberation and background noise." (EX1010, 1021.) Thus, Williams-Lim and Benesty are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problem of providing hands-free audio communication between users and devices.

490. The Williams-Lim-Benesty combination would result in a system where multiple microphones (additional microphones) as taught by Benesty are incorporated in Williams-Lim's first device to capture first and second human voice data to improve the quality of captured audio. For example, Benesty's microphone array(s) would be included in Williams' audio interface device to capture the first and second human voice data. Because Williams-Lim and Benesty relates to audio capturing and processing, a POSITA would be motivated to use the known techniques of Benesty to improve Williams-Lim' ability to capture and process audio, especially in a room where the microphone is not placed near the mouth of the user. Even in use cases where background noise is desired, a microphone array can capture richer audio information, enabling more efficient and accurate processing, for example, separating or balancing speech and background noise when desired.

491. Combining the teachings of Williams-Lim and Benesty as such would yield predictable results of effectively improve audio capturing and processing. Further, it would have been obvious to try combining the technical features of Benesty with the teachings of Williams-Lim, as there is a finite number of identified, predictable resolutions, without undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by

successful implementations in the field. There was a finite number of potential solutions to improve audio signal pickup, such as additional software processing or incorporating additional hardware components. The known benefits of Benesty would prompt the application of Benesty in the same field based on design incentives and other market forces to produce a system architecture that responds to captured human voice data accurately, having improved audio communication signal pickup.

2. Claim 8: "The first device according to claim 7, wherein the microphone and the additional microphone are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array."

492. Benesty discloses claim 8. (*Supra* §§ IX.D.2 - IX.D.3.)

493. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Benesty. (*Supra* § IX.M.1.)

N. Ground 14: Williams, Lim and Rossing Render Obvious Claims 9, 29, 33

1. Claim 9: "The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

494. Claim 1 is obvious in view of Lim. (*Supra* § IX.A.2) Claims 9, 29, 33 are obvious in further view of Rossing.

495. Rossing discloses claim 9. (*Supra* § IX.E.2)

496. A POSITA would understand that the claim refers to a standard microphone. (EX1001, 17:47-54.) As discussed, Williams discloses use of such a ***microphone*** in the ***first device***, consistent with the description of the '721 patent. (EX1001, Claim 9, 17:47-54.) With reference to FIG. 1, Williams discloses "[a]n ***audio interface device 116*** … having a ***microphone*** and a network interface." (EX1006, [0016]; *see also*, *id.*, [0007], [0019], claims 1-3 and 21). Further, Lim also discloses that the ***controlling device*** has a ***microphone***. "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated ***microphone 328***, speaker 330, and headphone jack 332 [.]" (EX1005, [0023], *see also*, [0019].)

497. Both Williams-Lim and Rossing are related to operating microphone(s) and speaker(s). (EX1006, [0016]; *see also*, *id.*, [0007], [0019], claims 1-3 and 21.) As discussed, Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1014, 757-758, 763, 765, 903.) Thus, Williams-Lim and Rossing are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problems of operating microphone(s) and speaker(s). Thus, a POSITA would have been motivated to look to Rossing for techniques and details regarding these components for implementing the system described in Williams-Lim.

498. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Williams-Lim, as various microphone features can improve William's microphone in the same way and yield predictable results. The known benefits of various features would prompt the application of Rossing in the same field based on design incentives and market forces. Further, it would have been obvious to try combining Rossing with the teachings of Williams-Lim, for example by using an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone in the client device to capture the first and second human data, without undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of types of microphones, such as a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone, that were commonly known in the industry and easy to manufacture and integrate into Williams-Lim.

2. Claim 29: "The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

499. Rossing discloses claim 29. (*Supra* § IX.E.3.)

500. A POSITA would have understood that this claim refers to a speaker.

501. As discussed, Rossing discloses that there are different types of speaker at the time of the invention, such as an electromagnetic speaker, a piezoelectric speaker, an electrostatic loudspeaker, or a magnetic loudspeaker. (*Supra* § IX.E.3, EX1014, 763, 321-322.)

502. As discussed, Williams and-Lim and Rossing are from the same field of endeavor and reasonably pertinent to the same technical problems of operating microphone(s) and speaker(s). Both Williams and Lim disclose sounding a sounder (*e.g.*, ***speaker 206, speaker 303***), consistent with the descriptions of the '721 patent, in the first device (*e.g.*, ***interface device 116, smart device 110***) in response to the received first message (*e.g.*, ***audio 222,*** recognized speech, analysis of or information associated with the recognized speech, and/or the next prompt). (*Supra* §§ IX.I.3.d, IX.I.3.i, IX.A.2.d, IX.A.2.i; EX1006, [0046], [0047], [0031], FIG. 4; EX1005, [0027]-[0029]; EX1001, 105:48-52, 105:61-106:4.) Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1014, 757-758, 763, 765, 903.) Thus, a POSITA would have been motivated to look to Rossing for techniques and details related to speakers for implementing

Williams-Lim, as various speaker features can improve William-Lim's speaker in the same way and yield predictable results. The known benefits of various features would prompt the application of Rossing in the same field based on design incentives and other market forces.

503. Further, it would have been obvious to try combining the technical features of Rossing with the teachings of Williams-Lim, for example, by using an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker in the client device to output audio reply, without undue effort and with a reasonable expectation of success. (*Supra* § IX.N.1.) There was a finite number of types of speakers, such as an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, that were commonly seen in the industry and easy to manufacture and integrate into the Williams-Lim system. (EX1014, 321-322, 763.)

3. Claim 33: "The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves."

504. Rossing discloses claim 33. (*Supra* § IX.E.4.)

505. A POSITA would have understood that this claim refers to a standard speaker. (EX1001, 105:58-106:4.) As discussed above, at the time of the invention, a POSITA would have understood that a speaker is a sounder for converting an electrical energy to sound waves. (EX1018, 122; EX1014, 763.) Further, as discussed, the '721 patent discloses the same. "An actuator may be an ***electroacoustic actuator, such as a sounder*** which converts electrical energy to sound waves . . . A ***loudspeaker (a.k.a. speaker) is a sounder*** that produces sound in response to an electrical audio signal input, typically audible sound." (EX1001, 105:48-52, 105:61-106:4.)

506. As discussed above, both Williams and Lim disclose sounding a sounder (*e.g.*, ***speaker***) in the first device. (*Supra* §§ IX.I.3.d, IX.I.3.i, IX.A.2.d, IX.A.2.i; EX1006, [0046], [0047], [0031], FIG. 4; EX1005, [0027]-[0029].)

507. Further, in light of Rossing, a POSITA would have understood that the speaker operates by emitting audible sound waves that are of an omnidirectional, unidirectional, or bidirectional pattern. For example, "***[s]ound radiates from a loudspeaker in all directions***, depending on angle and frequency … It is rare for a speaker to have truly constant directionality across its entire pass-band, in part from that fact that most are at least somewhat directional at mid and high frequencies, and, because of the long wavelengths

involved, almost unavoidably ***omni-directional at low frequencies***." (EX1014, 765.) It would have been obvious to combine Williams-Lim and Rossing. (*Supra* § IX.N.1.)

O. Ground 15: Williams, Lim and Goldsmith Render Obvious Claims 13-15

1. Claim 13: "The first device according to claim 1, wherein the WLAN is according to, or is compatible with, an Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard.."

508. Claim 1 is obvious in view of Williams-Lim. (*Supra* § IX.I.3.) Claims 13-15 are obvious in further view of Goldsmith. Goldsmith discloses claim 13. (*Supra* § IX.F.2.)

509. It would have been obvious for a POSITA to apply Goldsmith's disclosure of various types of wireless networking techniques and protocols to Williams-Lim. Both Williams-Lim and Goldsmith are related to wireless communication technologies. Williams expressly discloses usage of "other wireless data communications technologies" between interface device and application engine and "various network communications between the components of FIG. 2 [which includes the client device, the controlled device, and the server device] may be performed using existing networking protocols and techniques." (EX1006, [0020], [0012], [0013], [0038].) Lim discloses that access to "a remote server 500 accessible via a network 502 such as the

Internet" "may be via an intermediary appliance such as STB 104, illustrated at 506, or may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].) Goldsmith discloses various existing wireless connection systems in a chapter titled "1.4 Current Wireless Systems" (EX1040, 8-21.) Thus, Williams-Lim and Goldsmith are analogous art to the '721 patent, from the same field of endeavor, and all relate to the same technical problem of providing wireless communications between devices.

510. A POSITA would be motivated to use the known wireless networking techniques of WLAN over IEEE 802.11b, 802.11a, and/or 802.11g to improve the usability of Williams-Lim in the same way. Combining the teachings of Williams-Lim and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

511. A POSITA would have looked to Goldsmith's disclosure of WLAN based on IEEE 802.11b, IEEE 802.11a, and IEEE 802.11g as conventional wireless networking techniques and protocols that can be applied to Lim because Lim expressly discloses usage of WiFi, a form of WLAN. Further, Goldsmith explains the popularity of 802.11b wireless LANs and the wide coverage of locations with 802.11b base stations installed. (EX1040, 15-16.) 802.11a can support more users at higher data rates, and 802.11g can be used in either the 2.4

GHz and 5 GHz bands. Goldsmith discloses that "many wireless LAN cards and access points support all three standards to avoid incompatibilities." *Id*.

512. Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (*e.g.*, WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g) with the teachings of Williams-Lim, without undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of potential wireless network communications, such as infrared, BlueTooth, Zigbee, 3G, 4G, UWB, and GSM, that were commonly known in the industry and were easy to apply to Williams-Lim. The known benefits and feasibility of including a well-known wireless networking technique such as WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g as taught by Goldsmith to connect in-home devices with a remote server would prompt the application of Goldsmith in the same field based on design incentives and other market forces.

2. Claim 14: "The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

513. Goldsmith discloses claim 14. (*Supra* § IX.F.3.)

514. As discussed above, a POSITA would be motivated to combine the teachings of Goldsmith and Williams-Lim and apply the well-known wireless communication techniques such as WLAN over 802.11b as taught in Goldsmith to Williams-Lim. (*Supra* §§ IX.O.1, IX.F.2.)

3. Claim 15: "The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution."

515. Goldsmith discloses claim 15. (*Supra* § IX.F.4.)

516. Williams discloses that "***interface devices*** might use ***cellular*** or other wireless data ***communications technologies*** to communicate with the application engine 118 [of ***server 400***]," and "some in-home devices may communicate with the ***network 114 using cellular data technologies, or through other types of wired and wireless technologies that do not involve an in-home router or access point***." (EX1006, [0020], [0013].)

517. It would have been obvious for a POSITA to combine and apply Goldsmith's disclosure of conventional wireless communication protocols such as 3G using W-CDMA (UMTS), GSM, and EDGE to Williams-Lim. As discussed above, Williams-Lim and Goldsmith are analogous art to the '721

patent, from the same field of endeavor, and all relate to the same technical problem of providing wireless communications between devices. Williams discloses that the "***interface devices*** might use ***cellular*** or other wireless data ***communications technologies*** to communicate with the application engine 118 [of ***server 400***][.]" (EX1006, [0020].) Goldsmith discloses conventional cellular data technologies. (EX1040, 12-13.) A POSITA would be motivated to use the conventional cellular data technologies and wireless network communications techniques such as 3G using W-CDMA (UMTS), GSM, and EDGE as taught by Goldsmith to improve the usability of Williams-Lim in the same way. (*Supra* § IX.F.4.) Combining the teachings of Williams-Lim and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

518. Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (*e.g.*, 3G over W-CDMA (UMTS), GSM, and EDGE) with the teachings of Williams-Lim, without undue effort and with a reasonable expectation of success. (*Supra* §IX.O.1.) The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of potential wireless network communications, such as infrared, BlueTooth, and Wi-Fi, that were commonly known in the industry and were

easy to apply to Williams-Lim. In addition, wireless technology commonly included cellular connectivity in 2012, and Goldsmith teaches that 3G, GSM, are types of cellular technologies. (EX1040, 12-13.) The known benefits and feasibility of including a known cellular data technology and wireless networking technique such as 3G over W-CDMA (UMTS), GSM, and EDGE as taught by Goldsmith to connect in-home devices with a remote server would prompt the application of Goldsmith in the same field based on design incentives and other market forces.

P. Ground 16: Williams, Lim and Chatterjee Render Obvious Claim 18-20, 23

1. Claim 18: "The first device according to claim 17, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation."

519. Claim 1 is obvious in view of Williams-Lim. (*Supra* § IX.I.3.) Claims 18-20, 22 are obvious in further view of Chatterjee.

520. Chatterjee discloses claim 18. (*Supra* § IX.G.2.)

521. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Chatterjee. Both Williams-Lim and Chatterjee analogous art to the '721 patent. (EX1001, Claim 1, 135:59-63; EX1006, [0007], [0051].) All relate to a home automation system for remotely controlling home appliances. Williams discloses that "[a]n application might also be designed to allow a user to control various on-premises devices and equipment, such as

media equipment and ***home-automation devices***." (EX1006, [0051].) Lim discloses controlling various appliances, such as "***televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes ('STBs'), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers***, etc." (EX1005, [0017].) A POSITA would understand Williams-Lim suggests other home appliances could be similarly controlled. Chatterjee discloses "a system and method for providing home automation using a mobile device." (EX1012, Abstract.) Thus, Williams-Lim and Chatterjee are from the same field of endeavor and reasonably pertinent to the same technical problems of remote control of home appliances. Combining Chatterjee with Williams-Lim would improve and expand the capabilities of Williams-Lim in the same way. Combining the teachings of Williams-Lim and Chatterjee as such would yield predictable results of enabling home automation.

522. Because Williams and Lim explicitly provide that their systems can be used to control ***home automation devices***, ***HVAC, home lighting and thermostat***, respectively, and Chatterjee enumerates exemplary home automation devices and control of HVAC and home lighting systems, it would have been obvious to look to Chatterjee for additional home automation devices to be controlled. For example, the Williams-Lim-Chatterjee combination

would result in a system where Williams-Lim's first device controls appliances such as oven, refrigerator, HVAC, and washing machine as taught by Chatterjee, thereby improving and expanding the capabilities of Williams-Lim with a reasonable expectation of success.

523. Prior art such as Chatterjee has successfully demonstrated that the controlled smart appliances can be an oven, a dishwasher, light fixtures, and a refrigerator, and thus, there would have been a reasonable expectation of success for combining Williams-Lim and Chatterjee. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Further, there are a finite number of identified and predictable home appliances to try controlling remotely. Thus, it would have been obvious to a POSITA to combine the teachings of Chatterjee and Williams-Lim to use Williams-Lim' techniques to control home appliances such as an oven, a dishwasher, light fixtures, and a refrigerator as disclosed by Chatterjee. The known benefits and feasibility of remote controlling home appliances such as oven, a dishwasher, light fixtures, and a refrigerator as disclosed by Chatterjee would prompt the application of Chatterjee in the same field based on design incentives and other market forces.

2. Claim 19: "The first device according to claim 18, wherein the primary function of the appliance is heating food, and wherein

the appliance is a microwave oven, an electric mixer, a stove, an oven, or an induction cooker."

524. Chatterjee discloses claim 19. (*Supra* § IX.G.3.)

525. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Chatterjee. (*Supra* § IX.P.1.)

3. Claim 20: "The first device according to claim 18, wherein the appliance is a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker."

526. Chatterjee discloses claim 20. (*Supra* § IX.G.4.)

527. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Chatterjee. (*Supra* § IX.P.1.)

4. Claim 23: "The first device according to claim 17, wherein the primary function of the appliance is associated with cleaning, wherein the appliance primary function is associated with clothes cleaning and the appliance is a washing machine, or wherein the appliance is a vacuum cleaner."

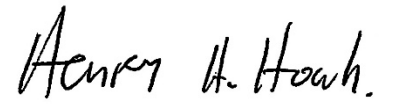
528. Chatterjee discloses claim 23. (*Supra* § IX.G.5.)

529. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Chatterjee. (*Supra* § IX.P.1.)

* * *

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and that these statements were made with knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code.

Dated: February 27, 2026



Henry H. Houh

Henry H. Houh

15 Ballard Ter.
Lexington, MA 02420
henry.houh@gmail.com
cell: 781-325-8273

Education

Massachusetts Institute of Technology, Cambridge, MA

- PhD in Electrical Engineering and Computer Science, February 1998. "Designing Networks for Tomorrow's Traffic," thesis supervised by Professor David Tennenhouse and Professor John Guttag. GPA 4.7/5.0
- Master of Science in Electrical Engineering and Computer Science, February 1991. "Demonstration of a laser repetition rate multiplier," thesis. GPA 4.5/5.0
- Bachelor of Science in Electrical Engineering and Computer Science, June 1989. "Boundary element analysis of arbitrarily shaped dielectric structures," thesis. GPA 4.7/5.0
- Bachelor of Science in Physics, February 1990. GPA 4.7/5.0

MITx, Cambridge, MA

- 2020 – 2021: Completed four courses in the Statistics and Data Science MicroMasters program, a sequence of MIT-offered graduate-level courses in statistics and machine learning.

Experience

Houh Consulting Inc. / Independent Consultant

- 2009 - present: Technical consultant specializing in Social Networking, Web 2.0, Web Site Development, Data Networking, Optical Networking, Telecommunications, Media Streaming and Voice Over IP. Clients include: Akin Gump Strauss Hauer & Feld LLP, Arnold & Porter Kaye Scholer LLP, BBN, Baker Botts LLP, Barnes & Thornburg LLP, BBN Technologies, Beus Gilbert McGroder PLLC, Bracewell LLP, Christensen O'Connor Johnson Kindness PLLC, Cooley LLP, Covington & Burling LLP, Cozen O'Connor P.C., DLA Piper, Davis Polk & Wardwell LLP, Desmarais LLP, Duane Morris LLP, Erise IP, P.A., Faegre Baker Daniels LLP, Finnegan, Henderson, Farabow, Garrett & Dunner, LLP, Fish & Richardson P.C., Goldman Ismail Tomaselli Brennan & Baum LLP, Haynes and Boone, LLP, Kaye Scholer, Kellogg Huber Hanson Todd Evans & Figel PLLC, Kirkland & Ellis LLP, Klarquist Sparkman, LLP, McCarthy Tetrault LLP, McGuireWoods LLP, Morgan, Lewis & Bockius LLP, Munger Tolles & Olson LLP, Orrick, Herrington & Sutfliffe LLP, O'Melveny & Myers LLP, Orrick, Herrington & Sutcliffe LLP, Paul Hastings LLP, Paul, Weiss, Rifkind, Wharton & Garrison LLP, Perkins Coie LLP, Quinn Emanuel Urquhart & Sullivan, LLP, Sidley Austin LLP, Smith Baluch LLP, Sterne, Kessler, Goldstein & Fox PLLC, Troutman Sanders LLP, Venable LLP, Weil, Gotshal & Manges LLP, Wiley Rein LLP, Wilmer Cutler Pickering Hale and Dorr LLP, and Winston & Strawn LLP.

BlocksCAD Inc. (H3XL spin-out)

- 2014 – 2025: Founder and member of Board of Directors. BlocksCAD is a visual Computer-Aided Design (CAD) program, designed to teach children mathematics and geometric visualization. BlocksCAD is being used in U.S. schools and worldwide in maker spaces, with 350,000+ users. BlocksCAD was selected to be part of the LearnLaunch 2017 Boost Accelerator, the MIT Play Labs 2018 Summer Accelerator, and the MassChallenge 2018 Boston Cohort. Recipient of DARPA DSO grant in 2014, and SBIR grant from the USDA in 2019. BlocksCAD was acquired by zSpace in 2025.

H3XL Inc. d/b/a Einstein's Workshop (formerly Lexington Robotics)

- 2009 - 2022: Founder and President. Started a science, technology, engineering and math enrichment program and creative/maker space, in 7,000 square feet of space. Serve 2,000+ kids and families annually. As of end of 2022, have delivered an estimated 250,000 student-instruction-hours of STEM courses. Principle Investigator for 2-year DARPA grant to improve 3D Computer Aided Design tool developed in-house for the purposes of teaching 8 year olds and up 3D CAD.

Eons

- 2008 - 2009: Chief Technology Officer. Created product that Eons acquired from BBN Technologies. Integrated BBN product with Eons social networking platform and significantly increased the Eons group creation rate. Helped evaluate advertisement platform offerings and rolled out the "Boom Network" advertising network. Eons raised $32 million from General Catalyst Partners, Charles River Ventures, Sequoia Capital, and Intel Capital.

BBN Technologies

- 2007 - 2008: Delta Division, Vice President of Technology. Grew "Boomerang" counter-sniper project engineering team and significantly de-risked $10 million worth of product deliveries. Identified sales team leading to $10 million dollars of new sales within 6 months and $100+ million in additional orders in the following two years. Boomerang was a significant asset leading to the acquisition of BBN by Raytheon in 2009. Created new business plan and grew team; launched new fully-featured social networking web site in 5 months. Served as lead expert witness in patent infringement lawsuit, resulting in $58 million jury award to client; verdict for patents I testified on were upheld on appeal which resulted in a $120 million settlement.
- 2004 - 2007: Delta Division, Director of Technology, responsible for commercializing IP and creating new businesses. Hired and grew division's initial engineering team. Wrote three business plans, two of which were funded. For call center business plan, acted as general manager, hiring and managing engineering team, inside sales team, and identifying and recruiting a new general manager. Identified and recruited other key employees to Delta Division, including senior members of team leading to successful internal sales growth and spin-outs of projects. Contracted by BBN to BBN spin-out PodZinger as VP of Operations and Technology. Additional areas studied include mesh networking, network security, satellite communications, and semantic web.

PodZinger Inc. (BBN spin-out, also known as EveryZing and now RAMP)

Amazon EX1002 Appendix A Page 2

- 2006: Vice President of Operations and Technology. Significantly upgraded capability of consumer-facing search site and redeployed web site from company to co-location facility. Identified key portions of infrastructure for upgrading and cost reduction. Helped write business plan, evaluating advertisement models of revenue. Hired in operations replacement and phased back to BBN.

Commonwealth Capital

- 2004: Entrepreneur-in-residence (informally), performed technical due diligence on business plans, brainstormed ideas for new businesses with venture partner. With venture partner, left for portfolio company BBN to form core of commercialization team.

Empirix, Inc./Teradyne, Inc.

- 2001 - 2004: Chief Technologist, Engineering Manager, Web Application Test Group. Researched potential new product areas; developed product plan and prototype. Responsible for three new and existing products. Managed off-shore development team. Chief architect for all web testing products. Re-architected core testing product, helped write javascript interpreter. Provided technical vision for core product.
- 2000 - 2001: Chief Technologist, Communication Infrastructure Test Group. Responsible for incorporating new technology internally, tracking new technologies, technical evaluation of partnerships and potential acquisitions. Helped develop division strategy. Developed plans which formed core capabilities for successful new products introduced in 2004-5.
- 2000 - 2001: Engineering Manager, Communications Infrastructure Test Group. Execution of new product plan developed in prior role. Grew team from four existing engineers to team of over 30 on immediate team and over 40 on project. Delivered new platform in one year. Platform and derivatives accounted for large portion of booked products for the division within 2 years and is currently (2008) a key portion of new product offerings.
- 2000: Empirix was formed as a spin-out of Teradyne in January 2000. Reported to CEO in carve-out of Empirix from Teradyne.
- 1999: (Teradyne) Director of Business Development, Software Test Units. Reported directly to Chairman of the Board/Founder and then to general manager of software test unit (6 divisions of Teradyne). Evaluated and researched acquisition and partnership candidates. Internally assessed technology position in market and gaps in product lines. Worked with senior division staff to develop new product strategies. Attended internal Teradyne divisional board meetings. Chairman served as my mentor.

3Com Corporation/NBX Corporation

- 1999: Software Engineer 5. Continued work after 3Com acquisition of NBX. Built cross-division relationships for new products and research directions. NBX was acquired by 3Com in March 1999.
- 1997 - 1999: (NBX) Senior Scientist and Engineer. Work in IP Telephony. Architected next-generation product. Protocol design and validation for core protocol now used tens of millions of times daily. Led team in integration of IP protocols into current product. Designed audio reconstruction algorithms. Developed applications for bug analysis and diagnosis of system problems. Implementation of network simulator. Work on collaborative projects with external

partners. Worked to identify gaps in product. Representative at numerous trade shows. Innovated novel methods of using product.

MIT Department of Electrical Engineering and Computer Science, Cambridge, MA

- 1991 - 1998: Research assistant, Telemedia Network Systems Group. Design, development, and implementation of Gigabit ATM network for distributed multimedia system. Studied host interface design issues. ATM network simulation.
- Spring 1989, Fall 1990, Spring 1995: Teaching assistant, Computation Structures digital systems course. (Spring 1995 Head TA)
- 1988 - 1989: Head laboratory teaching assistant for Computation Structures. Responsibilities included writing and revising lab assignments, and maintaining the lab.
- 1987: Laboratory teaching assistant for Computation Structures.
- 1987: Design, construction, and programming of 16-bit computer.

Agora Technology Group, Incorporated

- 1994 - 1996: Founder and CEO. Conceived and oversaw development of targeted advertising-supported Web sites. Responsible for company's vision and direction. Attended the first two WWW Conferences, presenting a workshop and paper at the first, and appearing on the "Commercialization and Economics of the Web" panel and chairing the "Where Commercial Services and the Web Are Headed" panel at the second. Sold company intact; is currently an operating stand-alone company.

AT&T Bell Laboratories, Holmdel, NJ

- 1989 - 1990: Implementation of cascadable all-optical fiber logic gate. Modelocking of all-fiber erbium laser. Construction of modelocked laser repetition rate booster. Strong optics laboratory and fiber optic experience.
- Summer 1988, 1987: Research in integrated optics. Analysis of rectangular waveguides using microwave modeling. Fabrication of integrated optical components.

Honors

- MIT Alumni Association Great Dome Award, 2010, Baker 60th Anniversary Reunion Co-Chair (highest group award given by MIT Alumni Association)
- MIT Alumni Association Presidential Citation Award (now known as Great Dome), 2008, Member of MIT Chairman's Salon committee
- MIT Alumni Association Bronze Beaver Award, 2007 (highest individual award given by MIT Alumni Association)
- MIT Alumni Association Volunteer Honor Roll, February 2004
- MIT Alumni Association Lobdell Award, 1999
- Boston Museum of Science Gold Pin for 1000 hours of Volunteer Service, April 1999
- MIT Alumni Association Presidential Citation Award (now known as Great Dome), 1997, Member of Alumni Online Communications Committee

Patents and Patent Publications

- U.S. Patent #9,697,231, H. Houh, J. Stern, R. Spina, M. Meteer, "Methods and apparatus for providing virtual media channels based on media search," July 4, 2017.
- U.S. Patent #9,697,230, H. Houh, J. Stern, "Methods and apparatus for dynamic presentation of advertising, factual, and informational content using enhanced metadata in search-driven media applications," July 4, 2017. See also WO2007056485.
- U.S. Patent #7,975,296, L. Apfelbaum, H. Houh, T. Mayberry and G. Friedman, "Automated security threat testing of web pages," July 5, 2011. See also US20030159063, WO2003067405.
- U.S. Patent #7,877,736, H. Houh and J. N. Stern, "Computer language interpretation and optimization for server testing," January 25, 2011. See also US20050138104, WO2005043300.
- U.S. Patent #7,801,910, H. Houh and J. N. Stern, "Method and apparatus for timed tagging of media content," September 21, 2010. See also US20070112837, US20090222442, WO2007056535.
- U.S. Patent #7,590,542, D. C. Williams, W. C. Hand, H. Houh, A. R. Seeley, "Method of Generating Test Scripts Using a Voice-Capable Markup Language," September 15, 2009. See also EP1530869, US20030212561, WO2003096663.
- U.S. Patent #6,967,963, H. H. Houh, P. Anderson, C. Gadda, "Telecommunication method for ensuring on-time delivery of packets containing time-sensitive data," November 22, 2005. See also EP1060400, WO2000033092, CA2318774.
- U.S. Patent #5,144,375, M. C. Gabriel, H. H. Houh, N. A. Whitaker, "Sagnac Optical Logic Gate," September 1, 1992. Also issued as European Patent # EP0456422, July 23, 1997, German Patent #DE69126913, August 28, 1997
- U.S. Patent Application Publication No. 20020015387, "Voice Traffic Packet Capture and Analysis Tool for a Data Network" (Abandoned in 2007)
- U.S. Patent Application Publication No. 20020016708, "Method and Apparatus for Utilizing a Network Processor as Part of a Test System" (Abandoned in 2007)
- U.S. Patent Application Publication No. 20020016937, "Method and Apparatus for Utilizing a Network Processor as Part of a Test System." (Abandoned in 2007) See also WO2002011413.
- U.S. Patent Application Publication No. 20070106646, "User-directed navigation of multimedia search results" (Abandoned in 2009)
- U.S. Patent Application Publication No. 20070106660, "Method and apparatus for using confidence scores of enhanced metadata in search-driven media applications" (Abandoned in 2009)
- U.S. Patent Application Publication No. 20070106685, "Method and apparatus for updating speech recognition databases and reindexing audio and video content using the same" (Abandoned in 2015)
- U.S. Patent Application Publication No. 20070118873, "Methods and apparatus for merging media content" (Abandoned in 2015)
- U.S. Patent Application Publication No. 20090222442, "User-directed navigation of multimedia search results" (Abandoned in 2016)
- U.S. Patent Application 11/395,732, "Search snippet creation for audio and video data" (Abandoned in 2009)
- U.S. Patent Application 11/774,931, "Methods and apparatus for managing a social networking web site"
- U.S. Patent Application 11/774,947, "Methods and apparatus for organizing media files"
- U.S. Patent Application 11/774,956, "Methods and apparatus for managing an online event"

- U.S. Provisional Patent Application 61/086,909, “Measuring and ranking relationship activity”
- U.S. Provisional Patent Application 61/086,914, “Detecting media object commonality”
- U.S. Provisional Patent Application 61/086,904, “Message categorization based on message characteristics”
- U.S. Provisional Patent Application 61/086,905, “Photo tagging to request action”

Publications

- “IP switching: server driven flow classification,” H. H. Houh, *Proceedings of the Washington University Workshop on Integration of IP and ATM* , November 1996.
- “Aurora at MIT,” D. D. Clark, H. H. Houh, and D. L. Tennenhouse, Editors, *MIT Laboratory for Computer Science Technical Report 673*, December 1995.
- “ViewStation Applications: Implications for Network Traffic,” C. J. Lindblad, D. Wetherall, W. Stasior, J. F. Adam, H. H. Houh, M. Ismert, D. Bacher, B. Phillips, and D. L. Tennenhouse, *IEEE Journal of Selected Areas in Communications*, 1995.
- “The VuNet Desk Area Network: Architecture, Implementation, and Experience,” H. H. Houh, J. F. Adam, M. Ismert, C. J. Lindblad, and D. L. Tennenhouse, *IEEE Journal of Selected Areas in Communications*, 13 (4), May, 1995.
- “Reducing the Complexity of ATM Host Interfaces,” H. H. Houh and D. L. Tennenhouse, Hot Interconnects II Symposium Proceedings, Stanford, August 11-12, 1994.
- “Media-intensive data communications in a `desk-area' network,” J. F. Adam, H. H. Houh, M. Ismert, and D. L. Tennenhouse, IEEE Communications, August 1994.
- “ViewStation Applications: Intelligent Video Processing Over A Broadband Local Area Network,” C. J. Lindblad, D. J. Wetherall, W. Stasior, B. Phillips, D. Bacher, J. Adam, H. Houh, M. Ismert, and D. L. Tennenhouse, Proceedings of the 1994 USENIX Symposium on High-Speed Networking, Oakland, CA, August 1994.
- “The Media Gateway: Live Video on the World Wide Web,” H. H. Houh, C. J. Lindblad, J. Soo, D. L. Tennenhouse, and D. J. Wetherall, Workshop at the 1994 World Wide Web Conference, Geneva, Switzerland, May 1994.
- “Active Pages: Intelligent Nodes on the World Wide Web ,” H. H. Houh, C. J. Lindblad, and D. J. Wetherall, Proceedings of the 1994 World Wide Web Conference, Geneva, Switzerland, May 1994.
- “Wavelength Division vs. Code Division Access Methods for Optical Networks,” H. H. Houh, Area Exam Paper, May 1993.
- “Experience with the VuNet: A Network Architecture for a Distributed Multimedia System,” J. F. Adam, H. H. Houh, D. L. Tennenhouse, Proceedings of the 18th Conference on Local Computer Networks, pp. 70-76, September 1993
- “The VudBoard: A Simple DMA Interface,” H. H. Houh, Proceedings of the 4th Gigabit Minijam, January 1994.
- “A Software-Oriented Approach to the Design of Media Processing Environments,” D. L. Tennenhouse, J. Adam, D. Carver, H. Houh, M. Ismert, C. Lindblad, W. Stasior, D. Weatherall, D. Bacher, and T. Chang., submitted to the International Conference on Multimedia Computing and Systems, May 1994.
- “A Network Architecture for Distributed Multimedia Systems,” J. F. Adam, H. H. Houh, M. Ismert, and D. L. Tennenhouse, submitted to the International Conference on Multimedia Computing and Systems, May 1994.

- "The Viewstation Collected Papers," D. L. Tennenhouse, J. Adam, C. Compton, A. Duda, D. Gifford, H. Houh, M. Ismert, C. Lindblad, W. Stasior, R. Weiss, D. Wetherall, D. Bacher, D. Carver, and T. Chang, MIT Laboratory for Computer Science Technical Report, MIT/LCS/TR-590, November 1993.
- "The Viewstation Collected Papers II," W. F. Stasior, D. L. Tennenhouse, J. F. Adam, D. R. Bacher, H. H. Houh, M. Ismert, C. J. Lindblad, B. M. Phillips, D. J. Wetherall, MIT Laboratory for Computer Science Technical Report, MIT/LCS/TR-696, April 1996.
- "A System's Perspective of the Sagnac Fiber Logic Gates and Their Possible Applications," A. Huang, N. Whitaker, C. Gabriel, H. Avramopoulos, P. M. W. French, H. H. Houh, and I. Chuang, Applied Optics, September 10, 1994
- "Complete Switching in a Three-Terminal Sagnac Switch," H. Avramopoulos, P. M. W. French, M. C. Gabriel, H. H. Houh, N. A. Whitaker, T. Morse, IEEE Phot. Tech. Lett. **3** (3), 235
- "Complete Switching in a Three-Terminal Sagnac Switch," H. Avramopoulos, P. M. W. French, M. C. Gabriel, H. H. Houh, N. A. Whitaker, IEEE/LEOS Annual Meeting, Paper PDP-13, November 1990
- "All-optical phase-locked oscillator," N. A.Whitaker, Jr., H. H. Houh, H. Avramopoulos, T. F. Morse, IEEE/LEOS Annual Meeting, Paper ELT2.4/MOO3, November 1990
- "Passive modelocking of an all-fiber erbium laser," H. Avramopoulos, H. H. Houh, N. A. Whitaker, M. C. Gabriel, T. F. Morse, IEEE/LEOS Conference on Optical Amplifiers and their Applications, Paper PDP-8, August 1990
- "Transverse modes, waveguide dispersion, and 30ps recovery in submicron GaAs/AlAs microresonators," J. L. Jewell, S. L. McCall, A. Scherer, H. H. Houh, N. A. Whitaker, A. C. Gossard, and J. H. English, Appl. Phys. Lett. **55** (1), July 3, 1989

Leadership, Activities and Interests

- Leadership
  - Reading Symphony Orchestra
    - Board of Directors (Treasurer), 2022 – present
  - Discovery Museum (Acton, MA)
    - Board of Directors, 2023 – present
    - Board of Directors, Development Committee, 2023 – present
    - Science and Technology Advisory Council, 2012 – present
    - The Discover Museum received the 2024 National Medal for Museum and Library Service, the nation's highest honor given to museums and libraries that make significant and exceptional contributions to their communities.
  - Tau Beta Pi National Engineering Honor Society
    - Executive Councilor and Secretary, 2024 – 2025, Treasurer 2026
    - Advisor, MA B Chapter at MIT, 2003 – present
    - District Director (National Officer), Tau Beta Pi, New England Area, 1991 – 2003
    - President, MA B Chapter at MIT, Fall 1988 – Spring 1989
    - Laureate award, 1989
  - MIT Alumni Association Board of Directors
    - MIT Annual Giving Board, 3 year term, 2024 – 2027
    - National Selection Committee, 3 year term, 2020 – 2023
    - K-12 STEM Initiatives Co-chair, 2013 – 2017
    - Awards Committee Chair, 2012 – 2014

        - Awards Committee, 3 year term, 2011 – 2014
        - Vice President, 2 year term, 2004 – 2006
        - Alumni Ad-Hoc Committee, Chair, 2005-2006
        - Board Member, 2 year term, 1997 – 1999
    - MIT Club of Boston
        - Board of Directors, 2006 – 2011
        - K-12 Initiatives Chair, 2009 – 2012
        - VP of Communications, MIT Club of Boston, 2003 – 2006
        - Past-President, MIT Club of Boston, 2002 – 2003
        - President, MIT Club of Boston, 2001 – 2002
        - President-Elect, MIT Club of Boston, 2000 – 2001
        - VP of Programs, MIT Club of Boston, 1999 – 2000
        - Activities Super-Chair, MIT Club of Boston, 1998 – 1999
    - MIT Enterprise Forum of Cambridge, Inc.
        - Past Chair, 2009 – 2011
        - In-NOW-vation Co-chair, 2010
        - Chair, 2007 – 2009
        - Vice Chair, 2005 – 2007
        - Executive Board Member, 2002 – 2011
        - Winter Workshop Co-Chair, February 2007 – conceived idea for conference, which sold-out and produced largest attendance numbers in recent memory
        - Spring Workshop Co-Chair, Spring 2004
        - Membership Committee Chair, Fall 2003 - 2006
        - 25th Anniversary Dinner Chair, Fall 2003
        - As Membership Chair and Board Member, started Special Interest Groups in 2004; a SIG won the MIT Presidential Citation award, the MIT Alumni Association's highest award for organizations, in 2006
    - Estabrook Elementary School PTA
        - Advisory committee to the superintendent on PCB issue, 2010-2011
        - 4th Grade after-school science program co-organizer, 2010-2012
        - 4th and 5th Grade before-school Math Olympiad co-organizer, 2009-2013
        - 5th Grade BBQ and Yearbook Committee, 2011, 2013
        - Family Math Night volunteer, 2008-2012
    - MIT Class of 1989
        - Class Agent, 2024 – present
        - 35-year Reunion Gift Committee, Co-chair, 2024
        - Secretary, five consecutive 5 year terms, 1989 – 2014
        - 30-year Reunion Committee, 2019
        - 25-year Reunion Committee, 2014
        - 20-year Reunion Committee and Gift Committee, 2009
        - 15-year Reunion Committee and Gift Committee, 2004
        - 10-year Reunion Committee and Gift Committee, 1999
        - 5-year Reunion Committee, 1994
        - Interim Treasurer, 1993 – 1994
        - Instituted annual senior class career fair, now raising over $100,000 annually for senior class activities, Fall 1988
    - Strong, consistent record of leadership dating to high school
- Acting

Amazon EX1002 Appendix A Page 8

  - ‘21’ (Sony Pictures), credited as “Chinatown Dealer,” 2007, Kevin Spacey's movie about the MIT Blackjack Team inspired by “Bringing Down the House” by Ben Mezrich, opened nationwide on March 28, 2008. 21 was the number one movie in US for two weeks and number one globally for one week. 21 also topped the DVD sales, Blu-ray sales and DVD rental charts.
  - Spring Lake Theater Company, first New York-area off-broadway production of “A Chorus Line,” played role of Mark, Summer 1990
- Former member of the MIT Blackjack team
- Producer for 10,000 Maniacs’ 2013 album “Music from the Motion Picture”
- Executive Producer for 10,000 Maniacs’ 2015 album “Twice Told Tales”
- Producer for 10,000 Maniacs’ 2016 DVD “10,000 Maniacs Live”
- Executive Producer for OK Go’s 2025 album “And The Adjacent Possible”
- 2024-25 production of “Romeo & Juliet” at the Circle in the Square Theater, nominated for 2025 Tony for Best Revival of a Play, Investor
- “The Interestings,” a new musical in progress by Sara Bareilles and Sarah Ruhl, Investor
- Violist, violinist, cellist, harpist, guitarist, singer, actor: played in many amateur/semi-professional groups including Firebird Pops Orchestra, Boston Civic Symphony Orchestra, Reading Symphony Orchestra, MIT Summer Philharmonic Orchestra, Carlisle Chamber Orchestra, Horizon Ensemble, Rivers Symphony Orchestra, Opera51 Orchestra, American Music Abroad Summer 2025 Tour, Horizon Ensemble, New England Film Orchestra, Freisinger Chamber Orchestra, Merrimack Valley Philharmonic, Longwood Symphony, MIT Symphony, and Somerville Community Chorus
- Violist, “Coming Home” album by Karen Phillips, 2014
- Google Scholar citations: 2224, h-index: 22, i10-index: 27 (since 2019: 539/12/12)
- Erdos-Bacon number of 7
- Erdos-Bacon-Sabbath number of 11