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. 11,128,710

Inter Partes Review No. IPR2026-00145

DECLARATION OF HENRY H. HOUH

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. 11,128,710 ("the '710 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 programming 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."

(EX1031, 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 '710 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 '710 PATENT

A. Filing Date and Priority Dates

18. The ’710 patent was filed on July 23, 2017, and issued on September 21, 2021. The ’710 patent claims earliest priority to a US provisional application that was 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 specification of US 11,128,710 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, 94:1-19; 96:7-97:26; 99:4-19, 100:5-101:3; 104:8-29; 105:48-106:52; 109:41-110:24.) “Sensors” could be smoke sensors, blood pressure sensors, voltmeters, capacitance meters, magnetic sensors, strain gauge, angular sensors, etc. (*Id.*, 17:35-38; 17:39-45; 18:42-47; 18:65-66; 19:18-29; 19:30-35.) Various network technologies were described to be “known in the art.” (*Id.*, 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.

20. Although the claims of the ’710 patent focus on a “captured human voice,” the specification provides sparse and scattered references to voice. Only a

single paragraph—out of 171 columns of written description—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." 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, 32:57-33:13.)

21. 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 '710 specification

depicts a broad workflow of voice processing and is similarly lacking. (EX1001, FIG. 19, 134:33-135:25.)

22. Indeed, during prosecution, the Applicant was not able to point to a coherent embodiment including the recited combination of steps. Instead, the applicant described claim 1 as an amalgamation of various components and steps scattered across multiple figures (e.g., "item 61a, 61b in Figure 6", "item 23a, 23b or 23c in Figure 2", "item 60 in Figure 5", "item 13a, 13b, or 13c in Figure 1", "item 50 in Figure 5", "item 189 in Figure 2", "numerals 14a, 14b, or 22 in Figure 2 or Figure 24", "item 24 in Figure 2", "item 151 in Figure 15", "item 171 in Figure 17", "51a, 51b in Figure 5", "item 192 in Figure 19", "item 150a in Figure 15a", "item 170 in Figure 17", etc.). (*E.g.*, EX1003, 827-828.)

C. Claims

23. I was asked to evaluate whether claims 1-29 of the '710 patent are patentable.

24. There is one independent claim. Claim 1 is directed to a method for operating multiple actuators in response to captured human voice data. The method is for use with a client device and a controlled device in a building, each communicating over a wireless network and for use with an Internet-connected server device external to the building. The method comprises performing at least 14 steps: (1) capturing, by a microphone in the client device, a first human

voice data; (2) sending to the server, by the client device via the wireless network, the captured first human voice data; (3) receiving, by the server over the Internet, the captured first human voice data; (4) processing, by the server, the captured first human voice data; (5) responsive to the processing, sending a first message, by the server to the client device over the Internet; (6) receiving, by the client device via the wireless network, the first message; (7) operating a first actuator in the client device in response to the received first message; (8) capturing, by the microphone in the client device, a second human voice data; (9) sending to the server, by the client device via the wireless network, the captured second human voice data; (10 receiving, by the server over the Internet, the captured second human voice data; (11) processing, by the server, the captured second human voice data; (12) responsive to the processing, sending a second message, by the server to the controlled device over the Internet; (13) receiving, by the controlled device via the wireless network, the second message; and (14) operating a second actuator in the controlled device in response to the received second message.

25. There are 28 dependent claims. Claim 2 requires the processing to comprise performing a voice recognition algorithm for identifying the voice of a specific person.

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

further comprising sending to the server, by the client device via the wireless network, the sensor data, and wherein the first message is sent by the server in response to the 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 each of the first and second actuators 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 client device to comprise multiple microphones, and wherein the capturing comprises capturing, by the multiple microphones in the client device, the human voice data. Claim 8 requires that the multiple microphones 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.

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 client device or the controlled device to be addressable in the wireless network or the Internet using an address stored in a volatile or non-volatile memory of the respective device for uniquely identifying the respective device in the network. Claim 11 requires the address to be a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type or wherein the address is a layer 3 address and is static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address.

31. Claim 12 requires the wireless network to be a Wireless Personal Area Network (WPAN), that is according to, or based on, Bluetooth™ or Institute of Electrical and Electronics Engineers (IEEE) 802.15.1-2005 standards, or wherein the WPAN is a wireless control network that is according to, or based on, Zigbee™, IEEE 802.15.4-2003, or Z-Wave™ standards. Claim 13 requires the wireless network to be a Wireless Local Area Network (WLAN) that is 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 wireless network to use a wireless communication over a licensed or an unlicensed radio frequency band, that is an

Industrial, Scientific and Medical (ISM) radio band. Claim 15 requires the wireless network to be 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, or wherein the cellular telephone network is a Fourth Generation (4G) network that uses 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.

32. Claim 16 requires the controlled device to be integrated in, part of, or entirely included in, an appliance. Claim 17 requires the primary functionality of the appliance to be associated with food storage, handling, or preparation. Claim 18 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 19 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 20 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 21 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 22 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 23 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.

33. Claim 24 requires the first actuator to be 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 25 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.

34. Claim 26 requires the first actuator to be a motion actuator that causes linear or rotary motion.

35. Claim 27 requires the first actuator to be a sounder for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves. Claim 28 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, or wherein the sounder comprises an electric bell, a buzzer, a chime, a whistle, or a ringer. Claim 29 requires the operating of the first actuator 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.

D. Prosecution History

36. The application that matured into the ’710 Patent was filed on July 23, 2017. (EX1003, 1062.) The application 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* EX1003, 9-14.)

37. The original claim 1 merely recited processing a single set of “human voice data.” (EX1003, 1304.) After a non-final Office Action and a final Office

Action, the Applicant appealed. (EX1003, 946-971, 861-887, 859.) The Examiner then reopened prosecution with a second non-final Office Action and a second final Office Action. (EX1003, 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. (EX1003, 622-633.)

38. The Board affirmed the Examiner's rejections of all claims, agreeing that the cited references disclosed the claimed concept of capturing human voice data by a client device, processing it at a server, and operating an actuator in a controlled device. (EX1003, 389-421.) ***I note that the Board effectively found the second half of the issued claim 1 (limitations 1[H]-1[N]) to be obvious.*** (*Id*., 393-402.)

39. After the PTAB decision, the Applicant further amended claim 1 to recite processing "first voice data" and "second voice data," both processed by the server to operate a first actuator and a second actuator in the controlled device. (EX1003, 344-346.) The Examiner issued two additional Office Actions in response. (EX1003, 229-248, 145-167.)

40. 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). (EX1003, 121-130.) The only difference between the claim

the Board had already found obvious and the ultimately allowed claim is that the ***client device can receive a voice command that causes control of the client device itself***, in addition to receiving a voice command that causes control of a controlled device (which the Board has found to be obvious). (*Id*.) The claimed network architecture has been found to be obvious by the Board. (EX1003, 389-421.)

41. Despite this minimal distinction, the Examiner allowed the application, without providing any substantive justification for the abrupt shift from repeated rejections to allowance. (EX1003, 9-14.)

V. RELEVANT LEGAL STANDARDS AND DEFINITIONS

A. Obviousness

42. 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.

43. 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.

44. 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.

45. 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

46. 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 '710 PATENT

47. 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.

48. A POSITA in the field of the '710 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.

49. I agree with this level of ordinary skill for the '710 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 ’710 patent, reflect the knowledge that existed in the relevant field at the time the ’710 patent priority application was filed, and with the state of the art at that time.

VII. CLAIM CONSTRUCTION

50. 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

51. I understand that each of the references discussed herein qualifies as prior art against the ’710 patent:

52. U.S. Publication No. 2012/0198339 (“Williams,” EX1004) was filed on January 28, 2011, and published on August 2, 2012.

53. U.S. Patent No. 7,529,677 (“Wittenberg,” EX1007) was filed on January 21, 2005, and published on May 5, 2009.

54. U.S. Publication No. 2012/0265528 (“Gruber,” EX1008) was filed on September 30, 2011, and published on October 18, 2012.

55. “Springer Handbook of Speech Processing” (“Benesty,” EX1009) was written by Jacob Benesty et al., and published by Springer-Verlag Berlin Heidelberg by at least December 11, 2007.

56. “Wireless Communications” (“Goldsmith,” EX1035) was written by Andrea Goldsmith, and published by Cambridge University Press by at least January 20, 2006.

57. U.S. Publication No. 2013/0052946 (“Chatterjee,” EX1011) was filed on August 23, 2011, and published on February 28, 2013.

58. U.S. Patent No. 8,340,975 (“Rosenberger,” EX1012) was filed on October 4, 2011, and published on December 25, 2012.

59. "Springer Handbook of Acoustics" ("Rossing," EX1013) was edited by Thomas D. Rossing et al., and published by Springer Science+Business Media, LLC by at least July 23, 2007.

60. U.S. Patent No. 7,062,339 ("Howard," EX1005) was filed on May 9, 2002, and published on June 13, 2006.

61. "Wireless Home Networking Simplified" ("Doherty," EX1006) 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: Williams Renders Obvious Claims 1, 5-6, 16, 23, and 29

1. Summary of Williams

62. Williams was filed on January 28, 2011, before the earliest provisional application filing date of the '710 patent, and thus qualifies as prior art under §102(e). (EX1004, 1.) The Examiner never found, and Applicant never cited, Williams during prosecution. Williams is generally related to "utilizing audio information that may already exist within a home or other location." (EX1004, [0007].) For example, "[a] user may ask a question, such as 'What's the 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." (EX1004, [0048].) As another example, 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." (EX1004, [0051]; *see also id*., [0001], [0008], [0009], [0024].)

63. For example, FIG. 1 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." (EX1004, [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***." (EX1004, [0023] (emphasis added)[1].) 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.***" (EX1004, [0011].)

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

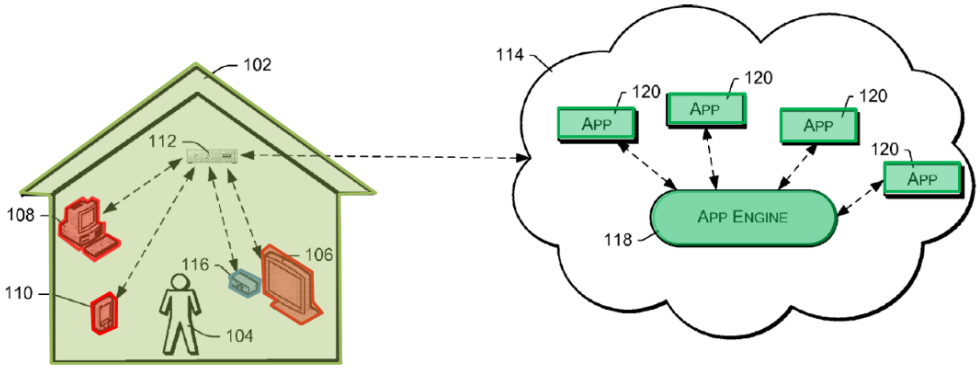


FIG. 1

64. As shown in FIG. 1, "the devices within the home 102 are illustrated as connecting to the network 114 via the in-home router 112." (EX1004, [0013].) 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." (EX1004, [0052].)

65. "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 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***." (EX1004,

[0022].) For example, the server can "provide audio 222 for eventual transmission to the ***interface device 116***, to be played on the speaker 206." (EX1004, [0037].) As another example, 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." (EX1004, [0051].)

66. Williams discloses a wide variety of wireless networking technologies that can be used. For example, with reference to FIG. 1, "[a]lthough the devices within the home 102 are illustrated as connecting to the network 114 via the in-home router 112, network connectivity might be implemented in various other ways. For example, 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." (EX1004, [0013].) "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." (EX1004, [0015].)

2. Claim 1

a. 1[PR][1]: "A method for operating multiple actuators in response to captured human voice data,"[2]

67. Williams discloses a method for operating multiple actuators in response to captured ***human voice data***. For example, FIG. 1 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." (EX1004, [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***." (EX1004, [0023].) A POSITA would have understood that Williams captures human voice data.

[2] I show how each claim preamble is satisfied, but do not concede that any preamble is limiting for any ground in the proceeding.

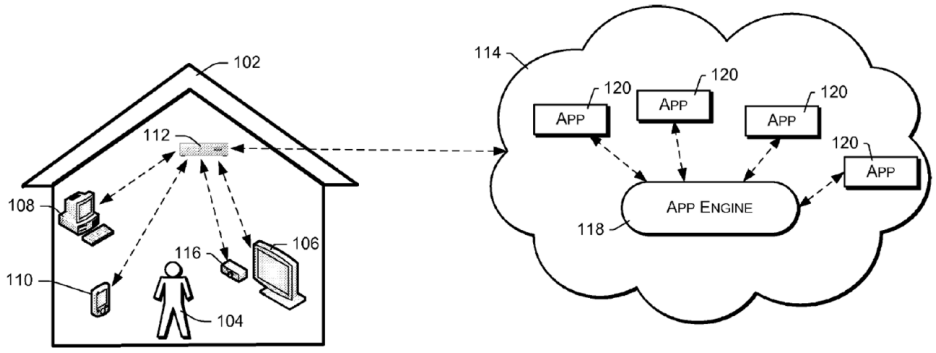


FIG. 1

68. Williams discloses that captured human speech data can be used to operate multiple actuators. Williams states, "[i]n operation, the application engine 118 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 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***." (EX1004, [0022].) For example, 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 a***ccept spoken commands*** from a user regarding ***on-premises devices***, and may respond by

interacting with such devices through available network connections." (EX1004, [0051].) As explained herein, each of the "on-premises devices" and the interface device contains actuators operable in response to captured human voice data. (*Infra* §§IX.A.2.k, IX.A.2.r.)

b. 1[PR][2]: "for use with a client device and a controlled device in a building,"

69. Williams discloses a ***client device*** and a ***controlled device*** in a ***building***.

70. First, Williams discloses a ***client 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". (EX1004, [0016].) A POSITA would have understood that the user's home 102 is or is part of a building. "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].) A POSITA would have understood that the device 116 communicates with and receives information from one or more server devices, such as server 400, and is thus a client device. (*Id.*, [0052]-[0057].)

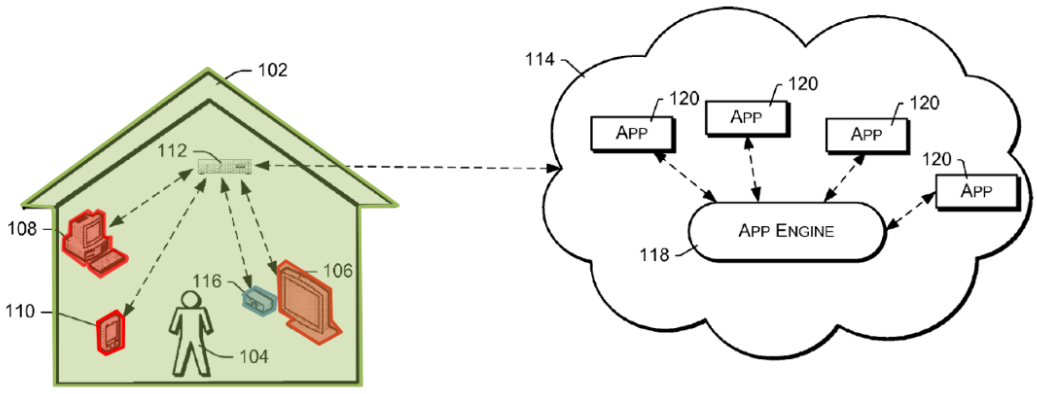


FIG. 1

71. FIG. 2 "is a block diagram illustrating additional details regarding the system of FIG. 1." (EX1004, [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." (EX1004, [0023].)

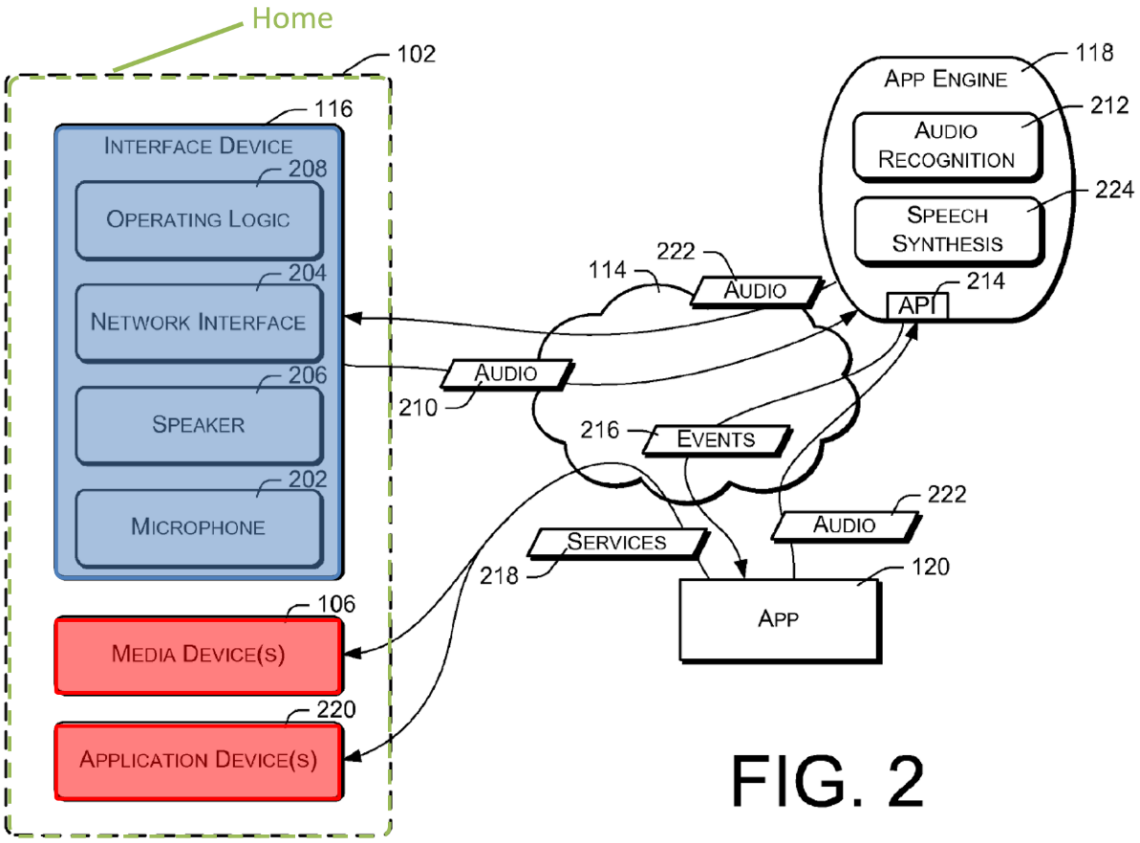


FIG. 2

72. Williams also discloses a ***controlled device*** in the building. With reference to FIG. 1, Williams discloses 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. (EX1004, [0011].) *See also* (EX1004, [0010].) Williams notes that "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." (EX1004, [0012].) For example, 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." (*Id.*, [0021], *see also id.*, [0007]).

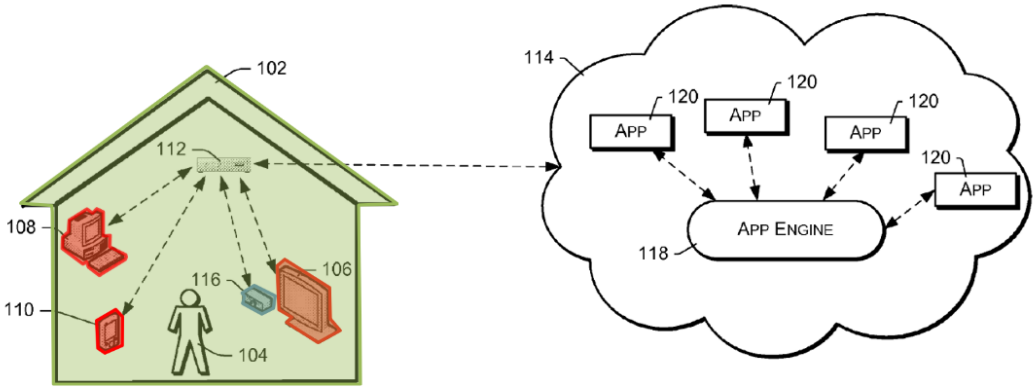


FIG. 1

73. The ***devices 106, 108, and 110*** are controlled devices. 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***." (EX1004, [0051].) Further, "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." (*Id.*) Thus, these devices can be controlled.

74. FIG. 2 "is a block diagram illustrating additional details regarding the system of FIG. 1." (EX1004, [0004].) With reference to FIG. 2, Williams discloses that "[t]he application 120 may use the information … to initiative or customize services 218 on behalf of users associated with the home 102 … [s]uch 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***." (EX1004, [0036].)

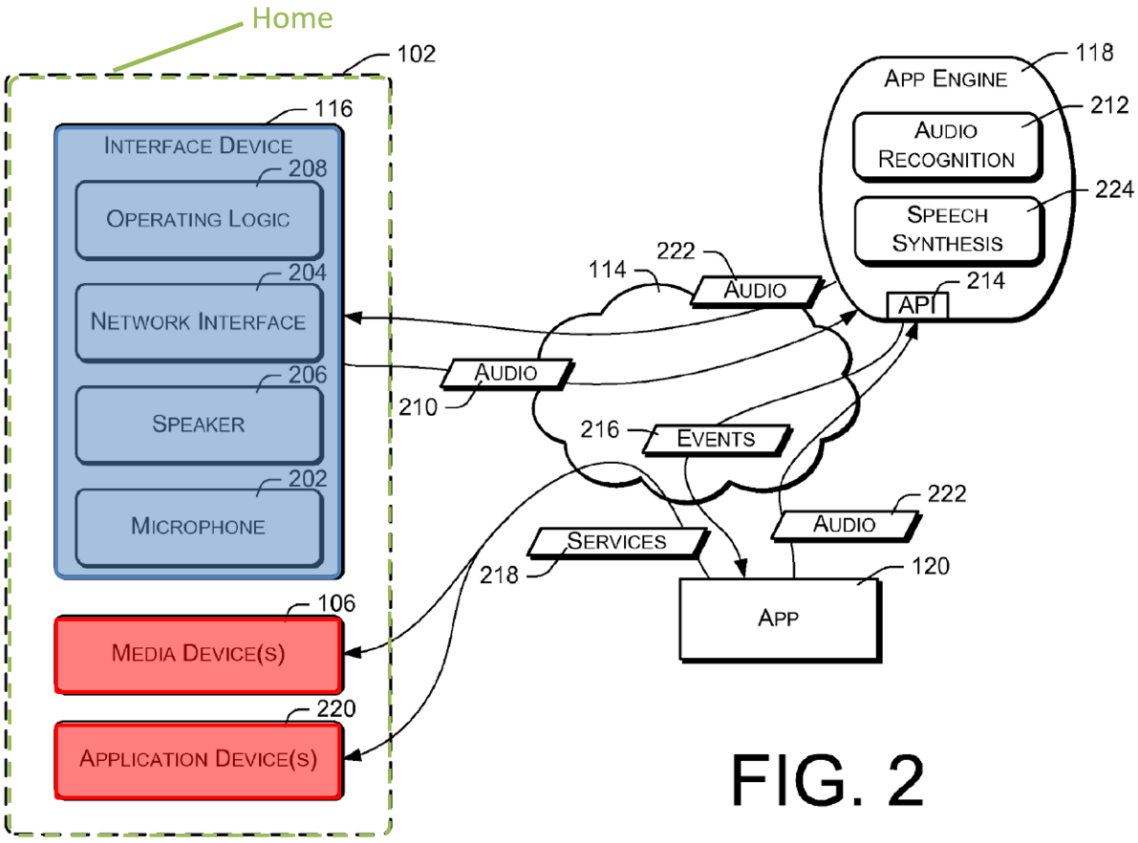


FIG. 2

75. A POSITA would have understood that Williams discloses that the ***client device*** and the ***controlled device*** are in the same ***building***. For example,

"FIG. 1 illustrates an environment 100 in which these techniques might be practiced. The environment includes ***a home or other user premises 102***. User premises may include ***houses***, offices, automobiles, and other spaces." (EX1004, [0010].) Further, as discussed above, Williams discloses 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.***" (EX1004, [0011].) Further, "[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***." (EX1004, [0016], [0023].) Specifically, 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***." (EX1004, [0019].)

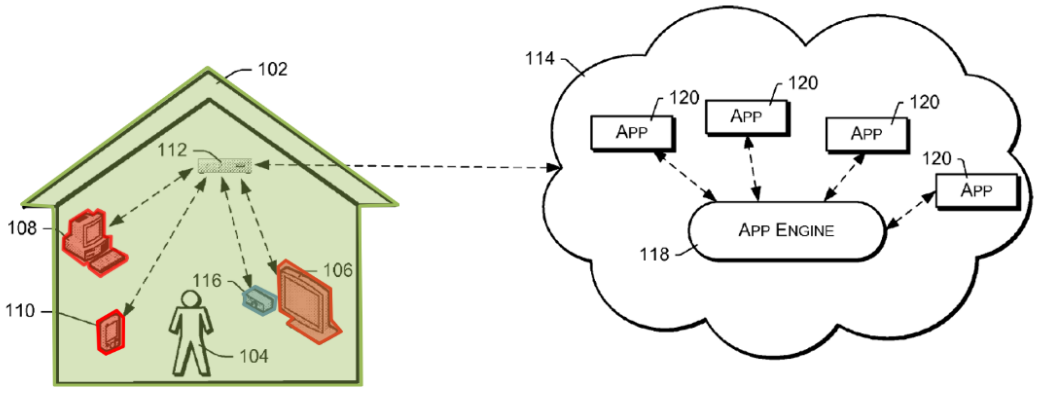


FIG. 1

76. Similarly, with reference to FIG. 2, Williams discloses that "***in-home audio interface devices 116*** may be located within the ***home 102***[,]" and 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***." (EX1004, [0023] and [0036].)

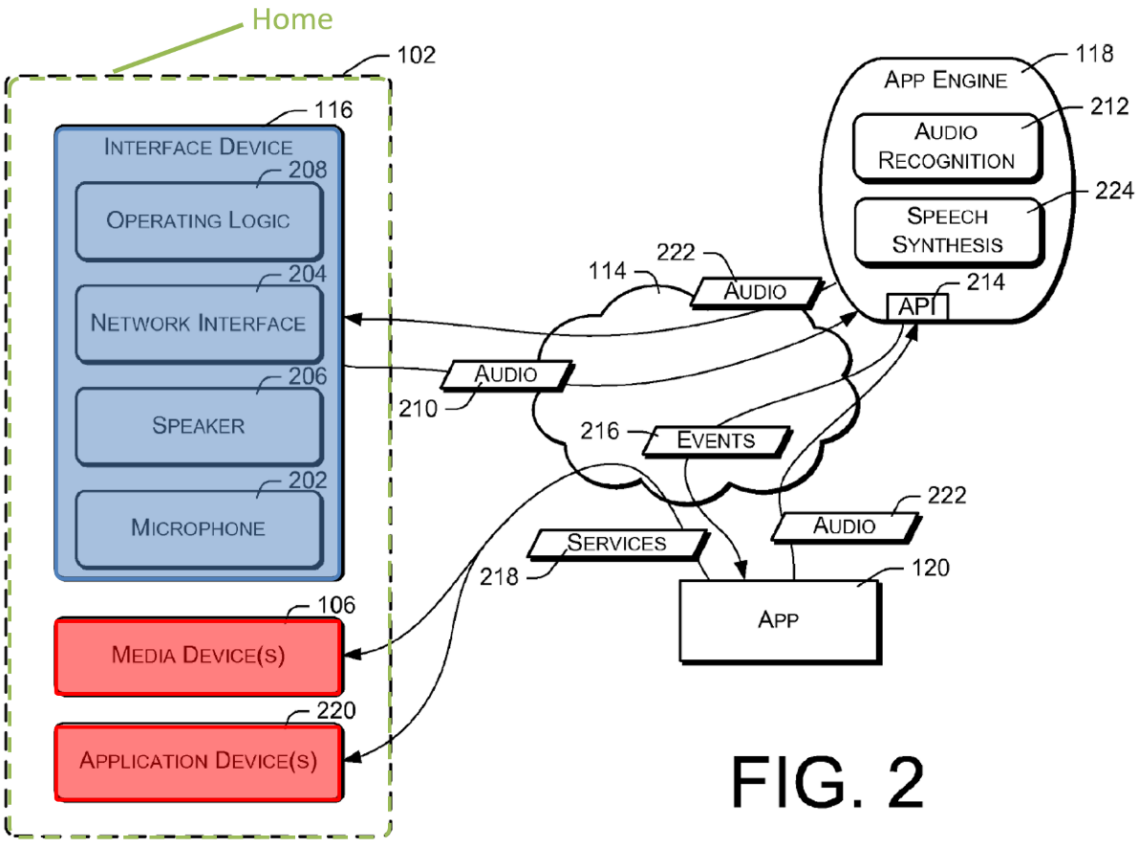


FIG. 2

77. Williams discloses that it is advantageous to place the ***client device*** and the ***controlled device*** in the same ***building***. "It may be convenient in many situations to place the ***audio interface devices*** near ***media devices*** within the ***home***

102, such as the ***media device 106***" and 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." (EX1004, [0019].)

78. Williams discloses its architecture may extend to multiple user premises. "FIG. 1 illustrates an environment 100 in which these techniques might be practiced. The environment includes a home or other user premises 102. User premises may include houses, offices, automobiles, and other spaces. Although a single home 102 is illustrated in FIG. 1, most implementations will include multiple premises and users." (EX1004, [0010].) In other words, there may be multiple user premises, each user premise having a respective interface device 116 (i.e., the recited "client device") and respective media devices 106 (i.e., the recited "controlled devices").

c. 1[PR][3]: "each communicating over a wireless network, and"

79. Williams discloses that each of the ***client device*** and the ***controlled device*** communicates over a wireless network.

80. Williams discloses the ***client device*** and the ***controlled device*** can communicate over a wireless network. For example, each of the ***interface device 116*** and the ***media devices 106/108/110*** communicates over a wireless network through wireless router/access point 112 consistent with the description of the '710 patent. In other words, the wireless network can be the in-home wireless network

formed by the wireless router/access point 112 and the in-home devices. 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***." (EX1004, [0012].) "[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." (EX1004, [0020].) This is consistent with the '710 patent, which 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 '710 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." (*Id*., 46:1-5.)

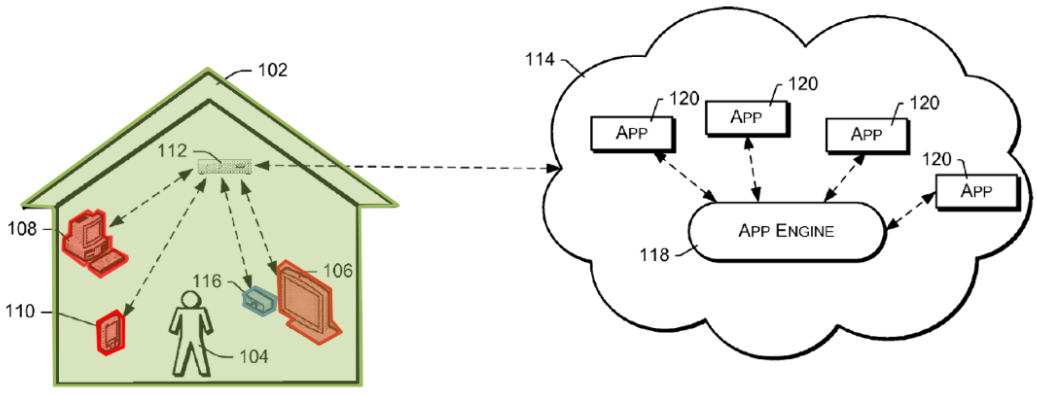


FIG. 1

81. I note that Williams discloses other types of wireless networks meeting the limitation. For example, with reference to FIG. 1, "[a]lthough the devices within the home 102 are illustrated as connecting to the network 114 via the in-home router 112, network connectivity might be implemented in various other ways. For example, 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." (EX1004, [0013].) "[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***. As mentioned above, such cloud-based communication might be implemented in many ways, and is not limited to the particular router-based configuration shown in FIG. 1. As an example, the

interface devices might use ***cellular or other wireless data communications technologies*** to communicate with the application engine 118." (EX1004, [0020], [0023].)

82. Further, some or all of the network 114 can be or be part of the recited "wireless network." "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." (EX1004, [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." (EX1004, [0015].) Alternatively, some or all of the network 114, in combination with the in-home network, also disclose the recited "wireless network".

83. FIG. 2 shows that the ***interface device 116***, the ***media device(s) 106***, and the ***application device(s) 220*** all communicates over a wireless network (e.g.,

network 114). "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, via ***the network 114***." (EX1004, [0023].) Williams further discloses that the "various network communications between the components of FIG. 2 may be performed using existing networking protocols and techniques." (EX1004, [0038].)

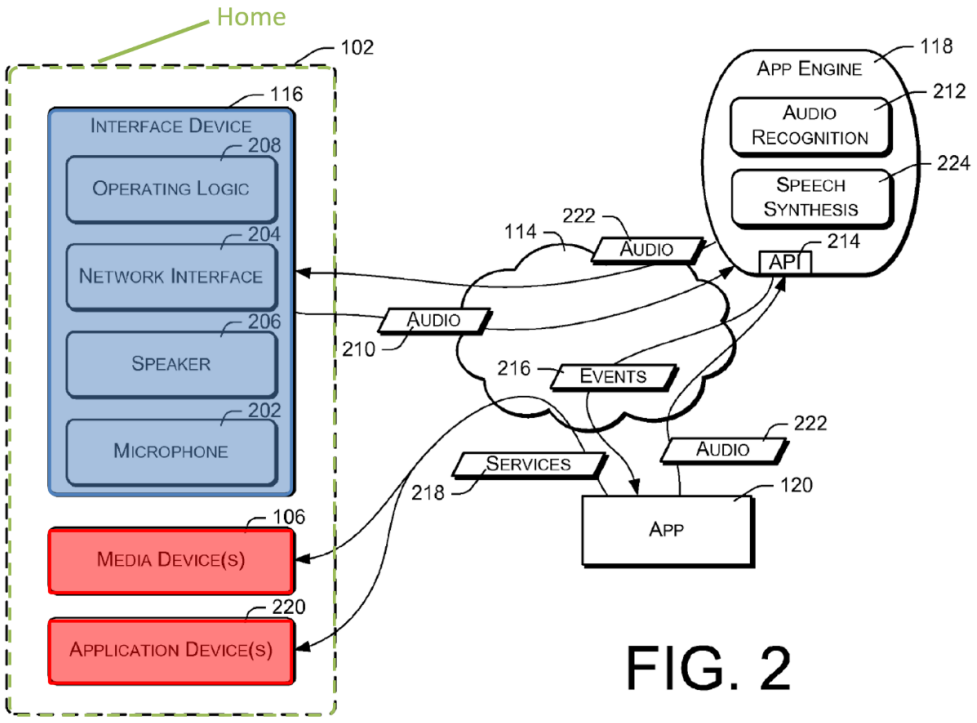


FIG. 2

84. I note that Williams discloses that the "***audio interface device 116*** may in some embodiments comprise a device having a microphone and

a ***network interface***." (EX1004, [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." (EX1004, [0007].) As shown, Williams discloses a device (i.e., ***audio interface device 116***) incorporated with microphones that comprises a network interface and with wireless networking capabilities, thus, it would have been obvious to a POSITA that the ***audio interface device 116*** would communicate over a wireless network.

d. 1[PR][4]: "for use with an Internet-connected server device external to the building, the method comprising:"

85. Williams discloses the method is for use with an ***Internet-connected server device*** 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." (EX1004, [0052].)

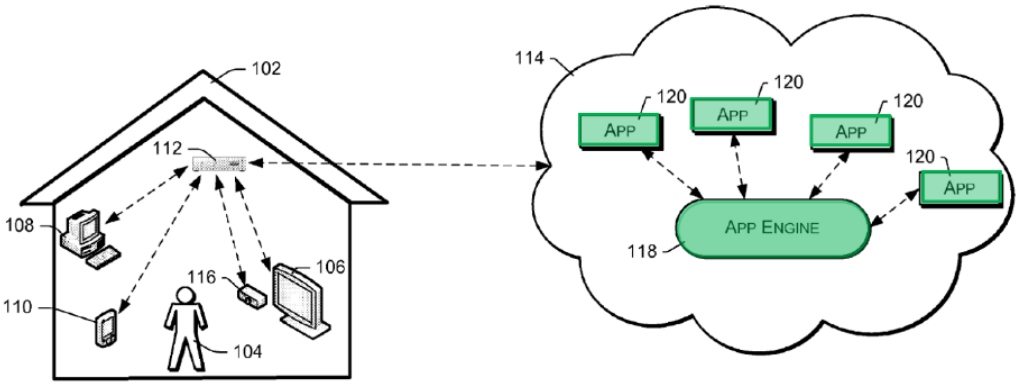


FIG. 1

86. As shown in FIG. 1, Williams discloses the in-home network in connected to cloud network 114. (EX1004, FIG. 1.) A POSITA would have understood that the server 400 is a cloud server, in other words, Internet-connected. Williams discloses that its ***server 400*** implementing the application engine 118 and applications 120 is ***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." (EX1004, [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." (EX1004, Abstract; *see also id*., [0007]-[0009].) Williams uses "cloud" and "Internet"

interchangeably. (*See, e.g.,* EX1004, [0001] ("Homes and other user premises are increasingly equipped with always-on ***Internet or 'cloud'*** connectivity.").)

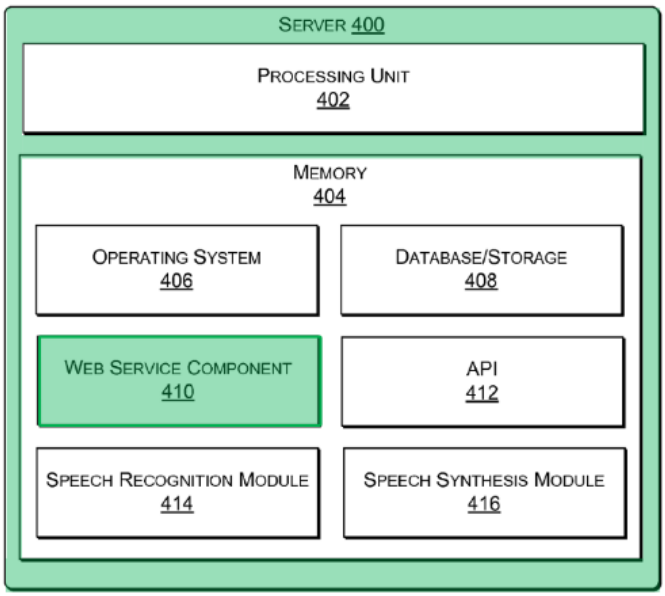


FIG. 4

87. As shown in FIG. 1, the application engine 118 and applications 120 are implemented in the wireless network 114. 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." (EX1004, [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." (EX1004, [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." (EX1004, [0021].)

88. 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 '710 patent. The '710 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.)

89. 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." (EX1004, [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 residential homes 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[.]" (EX1014, 277.) "[T]he 10th busiest web site in the world, Wikipedia, is hosted on 350 servers having 856 cores, spread across three ***data centers***." (*Id*., 282.) "Data centers and server clusters represent a highly visible 'hotspot' 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***." (EX1015, 1760.)

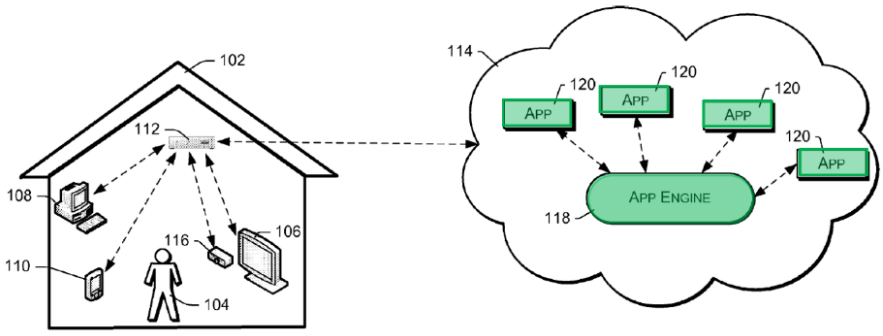


FIG. 1

90. 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." (EX1004, [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.

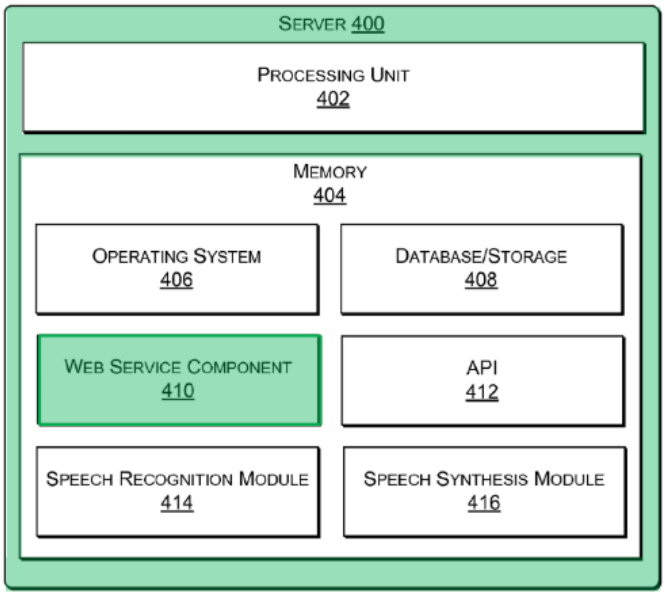


FIG. 4

91. 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." (EX1004, [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 may 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 the cloud server 400. 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." (EX1016, 664.)

e. 1[A]: "capturing, by a microphone in the client device, a first human voice data;"

92. Williams discloses capturing, by a microphone in the client device, a first human voice data. First, Williams discloses capturing of ***human voice data***. Williams discloses several exemplary user cases of providing services and functionalities in response to captured 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." (EX1004, [0048].) A POSITA would have understood that the human voice data is captured by the microphone of the client device.

93. 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." (EX1004, [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." (EX1004, [0050].) A POSITA would have understood that the human voice data is captured by the microphone of the client device.

94. Williams discloses a ***microphone*** in the ***client device***. With reference to FIG. 1, Williams discloses "[a]n ***audio interface device 116*** … having a ***microphone*** and a network interface." (EX1004, [0016]; *see also*, *id.*, [0007] ("network-enabled ***microphones*** can be placed at different activity centers within a home" to collect audio information and "microphones may be incorporated in small, self-contained units"); *id.*, [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).

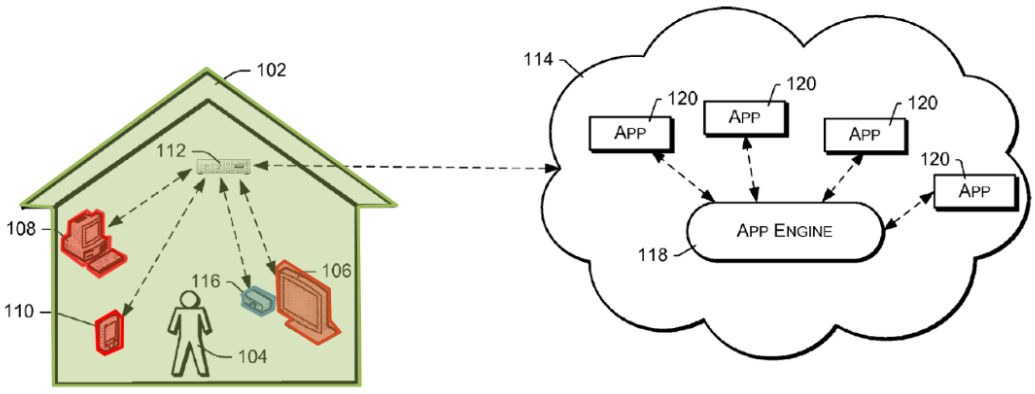


FIG. 1

95. Williams discloses capturing of a ***first human voice data*** by the microphone in the ***client device***. Williams discloses that the "***in-home interface device*** [] compris[es]: a ***microphone*** configured to ***receive*** audio from a nearby user[.]" (EX1004, 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.***" (EX1004, claims 1-3.)

96. Further, with reference to FIG. 2, 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***." (EX1004, [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." (EX1004, [0019].) That data (i.e., the audio-based information/ speech/ user commands) is captured and sent to the cloud. (*Id*., [0022], [0024]-[0025].)

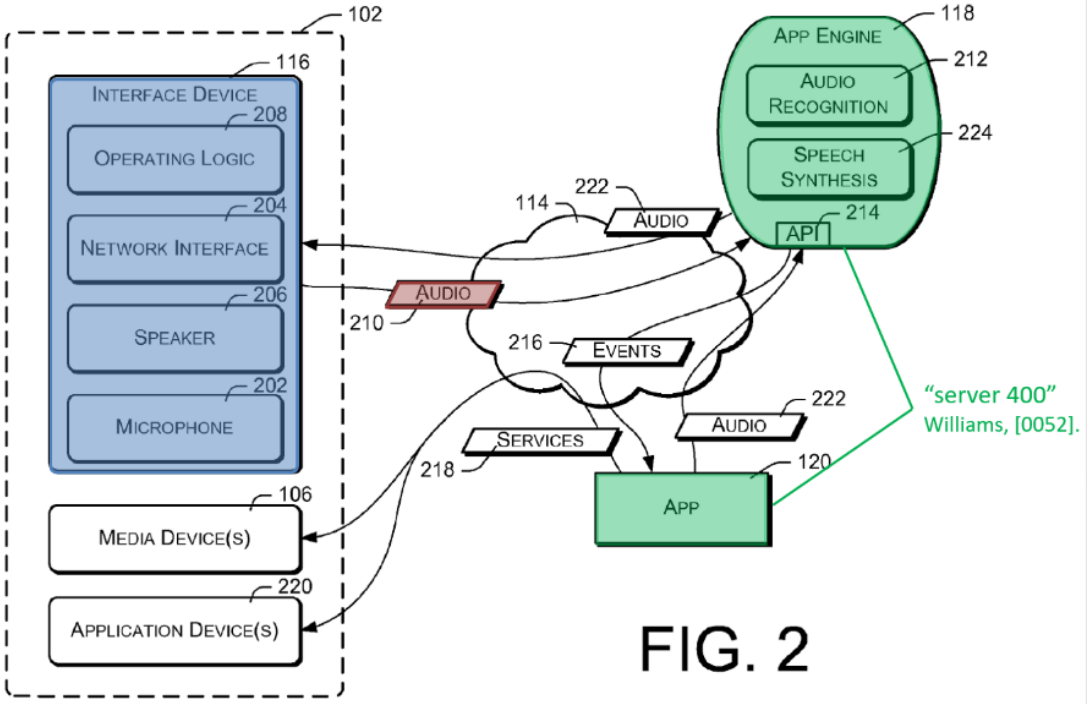


FIG. 2

97. 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.***" (EX1004, [0038] – [0039].)

f. 1[B]: "sending to the server, by the client device via the wireless network, the captured first human voice data;"

98. Williams discloses sending to the server, by the client device via the wireless network, the captured first human voice data. First, Williams discloses that the ***client device*** sends the captured ***first 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***. (*Supra* § IX.A.2.d.) 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." (EX1004, [0022], *see also id.*, [0020].)

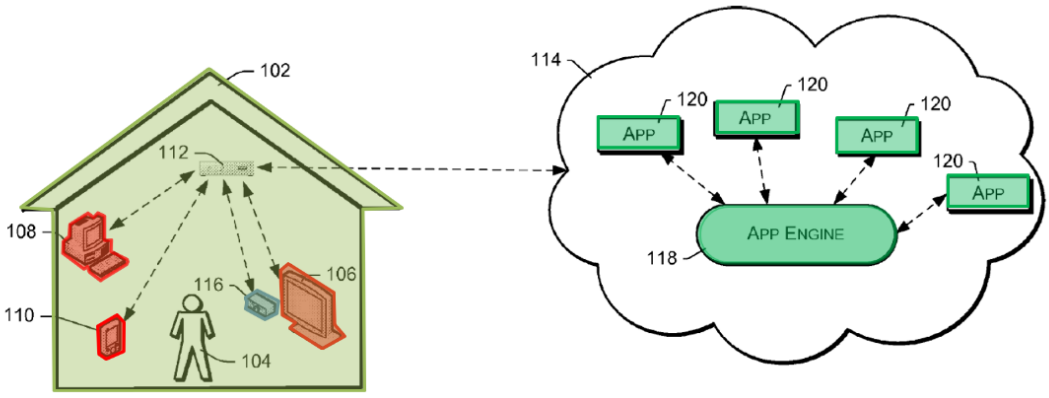


FIG. 1

99. 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." (EX1004, [0023].)

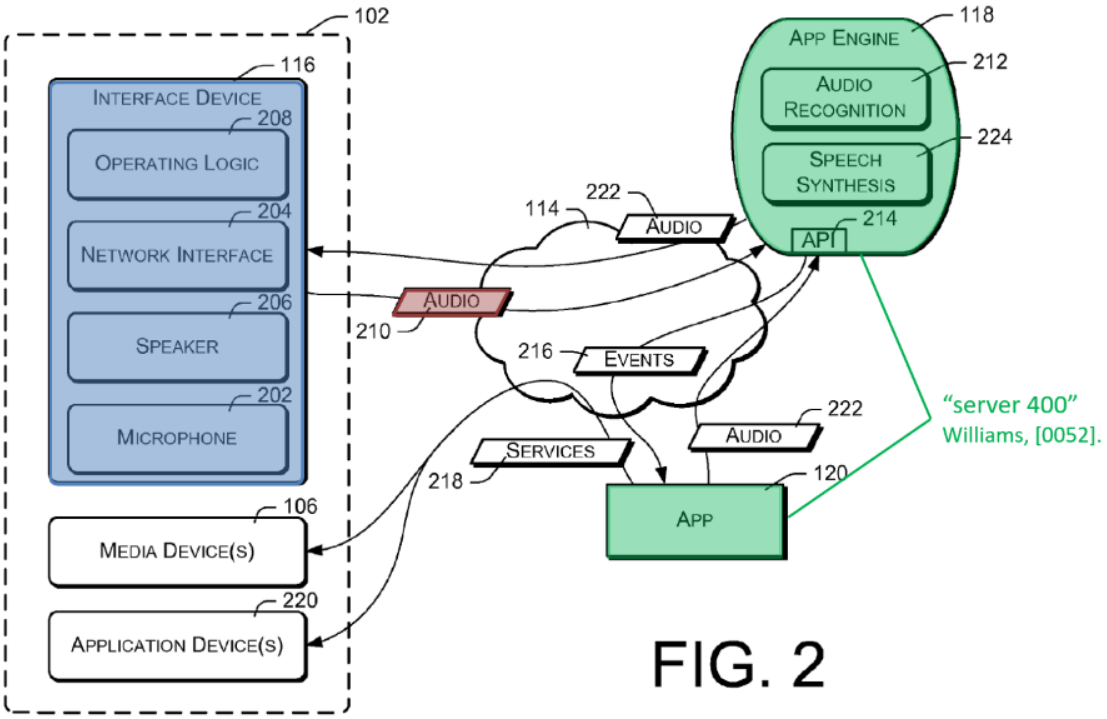


FIG. 2

100. 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." (EX1004, [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. (EX1004, [0038].) As shown in FIG. 2, ***audio 210*** is sent to the ***server*** by the ***interface device 116***.

101. Williams discloses that the ***client device*** sends the captured ***first human voice data*** to the ***server*** via the ***wireless network***. 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 and to the ***server*** via the Internet (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***. As mentioned above, such cloud-based communication might be implemented in many ways, and is not limited to the particular router-based configuration shown in FIG. 1. As an example, the interface devices might use ***cellular or other wireless data communications technologies*** to communicate with the application engine 118."

(EX1004, [0020]; *see also supra* § IX.A.2.b and IX.A.2.c.) As discussed above, the wireless network may be the in-home network (e.g., via router 112). Alternatively, the wireless network may be any in-home network, wireless portions of the network 114, or any combination thereof.

g. 1[C]: "receiving, by the server over the Internet, the captured first human voice data;"

102. Williams discloses receiving, by the server over the Internet, the captured first human voice data. First, Williams discloses that the ***server*** receives the captured ***first human voice data***. For example, "in 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." (EX1004, [0022], *see also id.*, [0020].) Thus, a POSITA would have understood that the cloud server 400 receives the captured human voice data over the Internet.

103. With reference to FIG. 2, Williams shows that "the application engine 118 [of ***server 400***] is configured to receive the ***audio-based information 210***." (EX1004, [0027].). Williams also discloses that the "***audio-based information 210*** received by the application engine 118 [of ***server 400***] may comprise a ***raw audio stream***," and "depending on the type or format of the received information, 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 … to identify or indicate the following types of information: [0028] spoken commands uttered by users." (EX1004, [0027]-[0028], [0034].)

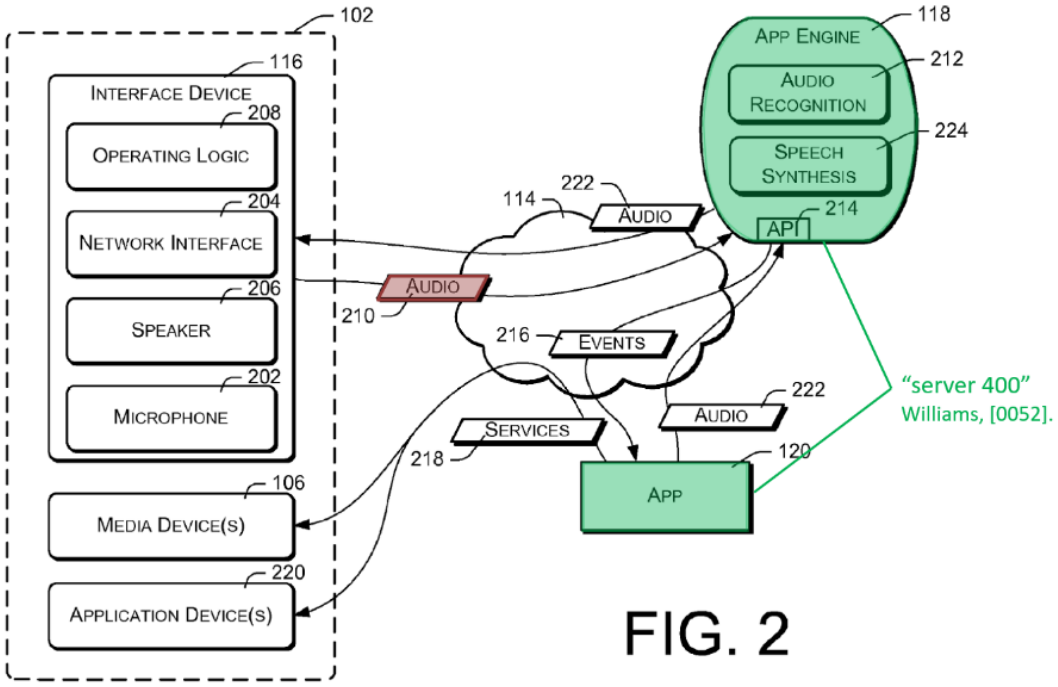


FIG. 2

104. Williams discloses that the "***[a]udio*** may be received by an in-home or on-premises interface, by a cloud-based service or application engine [running on a ***server***], by other cloud-based services and applications, or combinations of these entities." (EX1004, [0039]; s*ee also*, *id.*, [0024], [0038]).

105. Williams discloses that the ***server*** receives the captured ***first human voice data*** over the ***Internet***. As shown in FIG. 2, ***audio 210*** is received by the application engine 118 [running on ***server 400***] via the wireless network 114. As discussed, Williams discloses that "[t]he 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." (*Supra* § IX.A.2.d; EX1004, [0015].)

h. 1[D]: "processing, by the server, the captured first human voice data;"

106. Williams discloses that the ***server*** processes the captured ***first human voice data***. For example, "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 [run on ***server 400***], which can act on the information to provide or enhance services to the user 104." (EX1004, [0022].) As discussed above, a server device, ***server 400***, is used to implement application engine 118 and/or any of the applications 120. (*Supra* § IX.A.2.d; EX1004, [0052].)

107. Williams teaches that "the application engine 118 [of ***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". (EX1004, [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*** [.]" (EX1004, [0027] - [0028].)

108. Williams discloses that an exemplary processing is speech recognition. "[T]he 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." (EX1004, [0034].) "For example, ***voice recognition technologies*** can be used to convert user speech into text, which can in turn be parsed to identify particular commands or instructions. . . . Similar technologies can be used to compare audio against databases of other know media such as television shows, podcasts, movies, and so forth." (*Id.*, [0033].) Indeed, with reference to FIG. 4, Williams teaches that "an example ***server 400*** may comprise a processing unit 402 composed of one or more processors, and memory 404." (EX1004, [0053].) "The memory 404 may also include a ***speech recognition module*** 414 and a speech synthesis module 416 [.]" (EX1004, [0056]; *see also id.*, [0039] – [0040].)

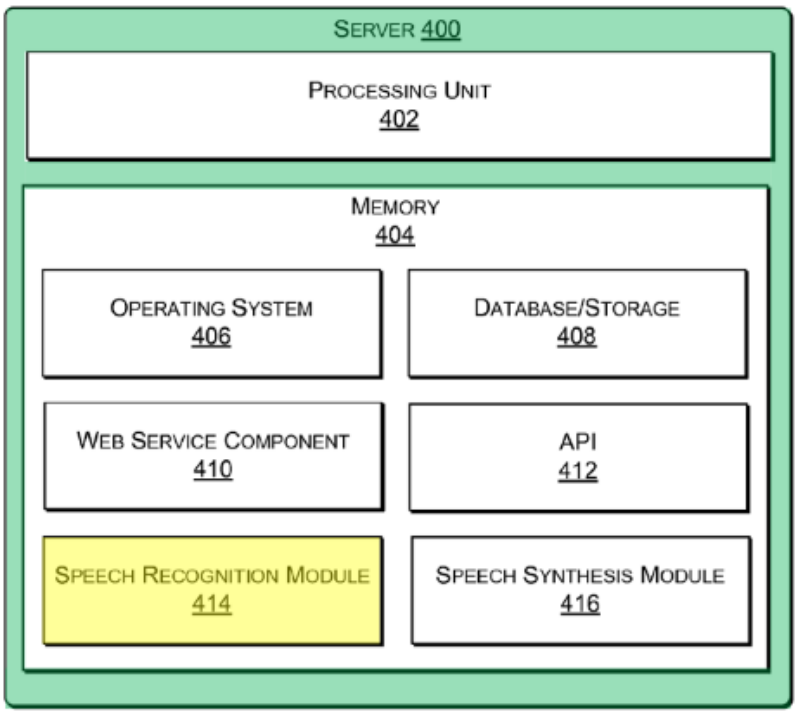


FIG. 4

109. Williams discloses that application engine 118 and/or any of the applications 120 of ***server 400*** can operate together to process the captured ***first human voice data*** via "an application interface 214 (also referred to as an application programming interface or API)" (EX1004, [0035].) "The API 214 allows an application 120 to obtain event descriptors that have been created by the application engine 118" and "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.*, [0035] - [0037].)

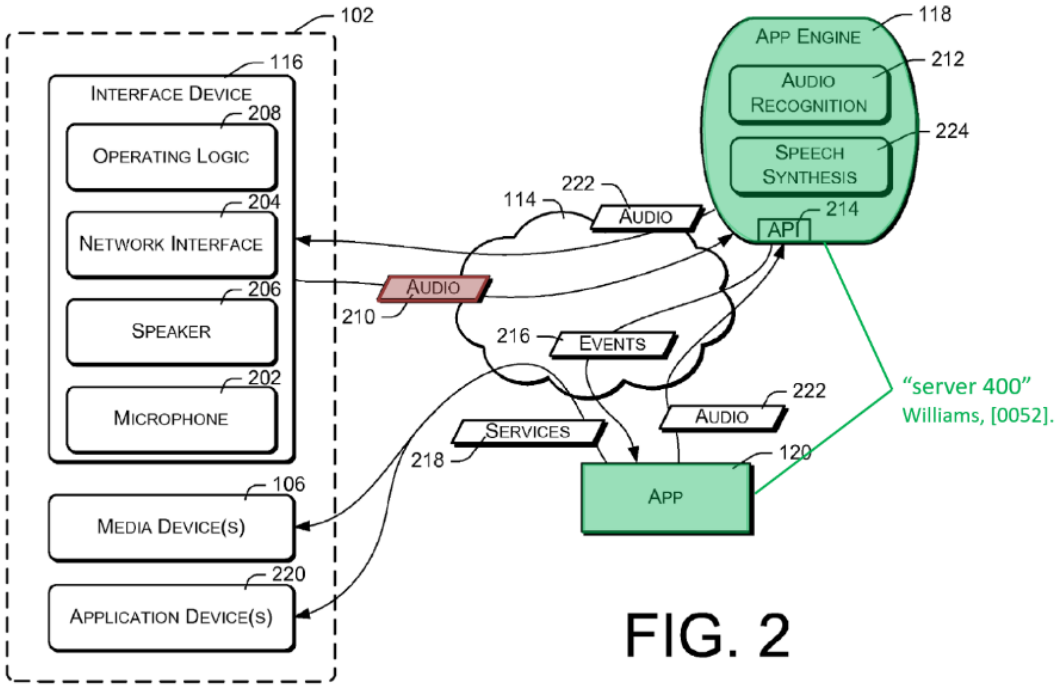


FIG. 2

i. 1[E]: "responsive to the processing, sending a first message, by the server to the client device over the Internet;"

110. Williams discloses that the ***server*** sends a ***first message*** to the ***client device*** over the Internet responsive to processing.

111. Williams discloses several examples of providing to the user an ***audio response*** generated by the ***cloud server*** in response to processing the received audio 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***." (EX1004, [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." (EX1004, [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***." (EX1004, [0045].) A POSITA would have understood that the audio response or information associated with the audio response sent to the client device disclose the "first message".

112. Williams discloses that the ***first message*** is sent from the ***server*** to the ***client device*** over the Internet. As discussed above, a server device, ***server 400***, implements the application engine 118 and/or any of the applications 120. (*Supra* § IX.A.2.d; EX1004, [0052].) William discloses, with reference to FIG. 2, that the "API 214 of the application engine 118 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." (EX1004, [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." (*Id.*)

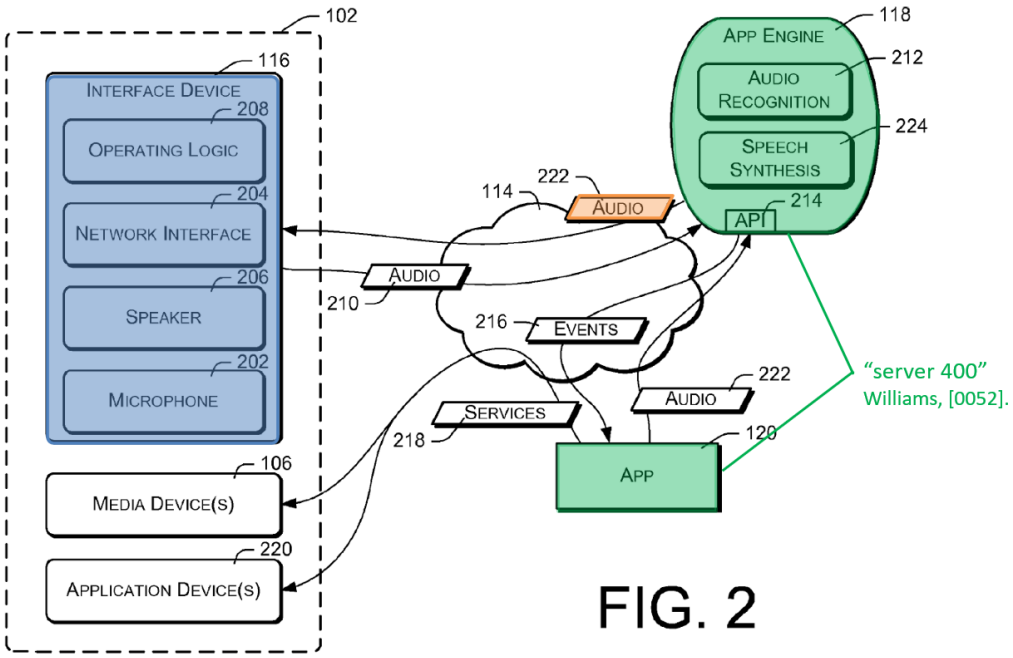


FIG. 2

113. Williams teaches that the first message is sent by the server "responsive to the processing" as recited. The processing is performed by the ***server device 400*** implementing the application engine 118 and/or any of the applications 120 and the first message is sent in response. "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 [run on ***server 400***], which can

act on the information to provide or enhance services to the user 104." (EX1004, [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." (*Id.*, [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 network-based applications." (EX1004, Claims 9 and 16.)

114. Williams discloses that the ***server*** sends the ***first message*** to the ***client device*** over the ***Internet***. As shown in FIG. 2, ***audio 222*** is sent by the ***cloud server*** via the wireless network 114 (i.e., "the Internet"). As discussed, Williams discloses that "[t]he 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." (*Supra* § IX.A.2.d; EX1004, [0015].)

j. 1[F]: "receiving, by the client device via the wireless network, the first message;"

115. Williams discloses receiving, by the client device via the wireless network, the first message. First, Williams discloses that the ***client device*** receives the ***first message***. With reference to FIG. 2, 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." (*Supra* § IX.A.2.i; EX1004, [0037].) As shown in FIG. 2, the ***interface device 116*** receives the ***audio 222*** for output by the speaker 206 of the ***interface device 116***. (*See also id.*, [0023] ("The interface device may also include a speaker 206, and operating logic 208.").

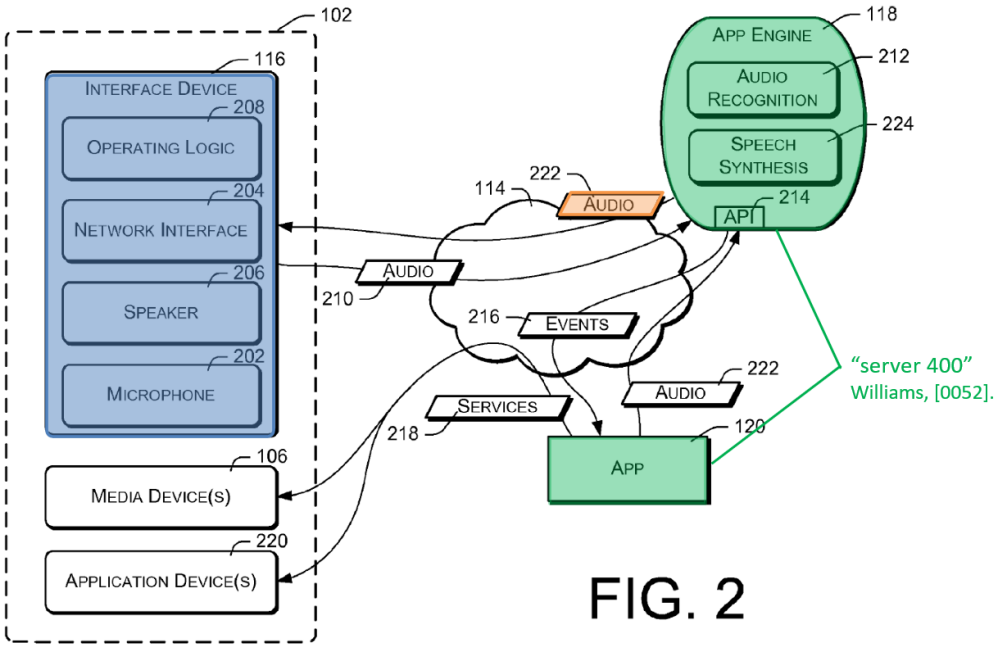


FIG. 2

116. Williams discloses that the ***client device*** receives the ***first message*** via the ***wireless network***. 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***." (EX1004, [0012].) Thus, it would have been obvious that the ***audio 222*** is received by the "***router or access point 112***" (which is connected to the ***cloud server*** via the Internet or "***public network 114***") and then wirelessly transmitted to the ***interface device*** via the home wireless network. (*Id.*) Williams discloses that other types of wireless networks can be used. "Although the devices within the home 102 are illustrated as connecting to the network 114 via the in-home router 112, network connectivity might be implemented in various other ways. For example, 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." (EX1004, [0013].)

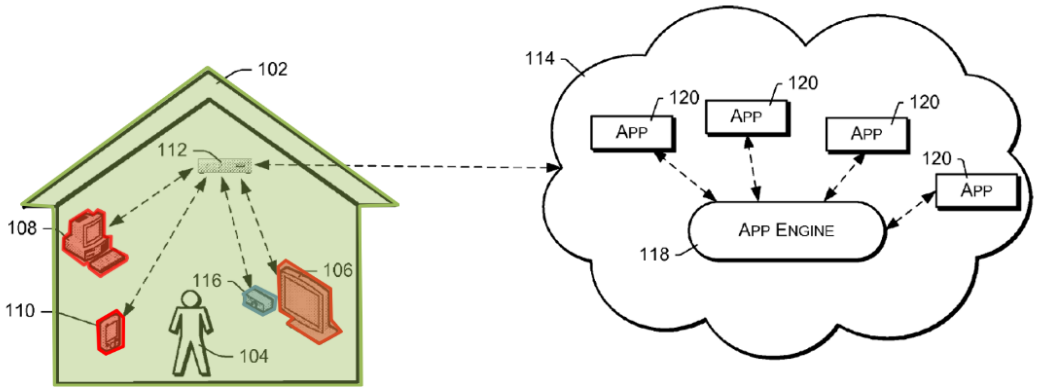


FIG. 1

117. Further, as shown in FIG. 2, the ***interface device 116*** receives the ***audio 222*** via the ***wireless network 114***. (*See also supra* § IX.A.2.b.)

k. 1[G]: "operating a first actuator in the client device in response to the received first message;"

118. Williams discloses operating a ***first actuator*** in the ***client device*** in response to the received ***first message***. First, Williams discloses a ***first actuator/speaker*** in the ***client device***. Williams discloses that "the ***interface device*** may also have a ***speaker*** or ***other form of audio output***." (EX1004, [0016].) With reference to FIG. 2, William discloses that "[t]he ***interface device*** may also include a ***speaker 206***, and operating logic 208." (EX1004, [0023].)

119. Williams discloses operating a ***first actuator/speaker*** in the ***client 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." (EX1004, [0037].)

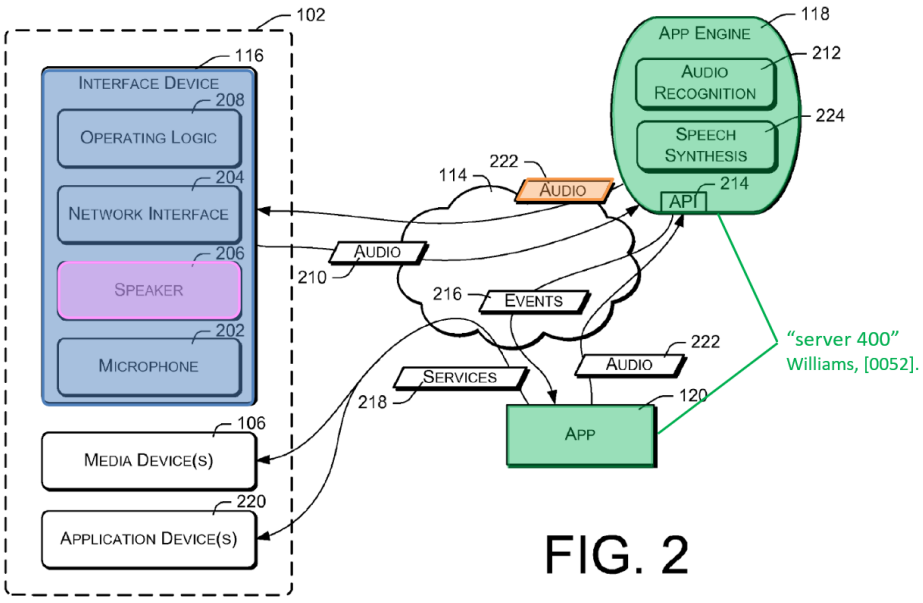


FIG. 2

120. The specification of the '710 patent itself states that a speaker has or can be an actuator. The '710 patent explains "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 ***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.) Considering that disclosure, a POSITA would have understood that a speaker is or has an actuator to produce sound.

121. At the time of the invention, it was well-known that a speaker is or has an actuator that 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." (EX1017, 122, 123.)

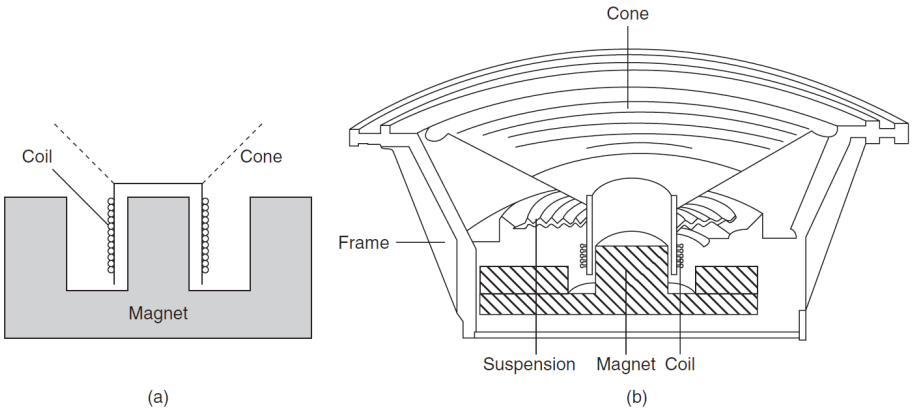


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

122. "[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." (EX1013, 763.) Thus, a POSITA would have understood that Williams discloses operating an actuator (e.g., ***speaker 206***) in the client device (e.g., ***interface device 116***) in response to the received first message (e.g., ***audio 222***).

123. Williams discloses several exemplary use cases of operating 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." (EX1004, [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***." (EX1004, [0048].)

124. 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." (EX1004, [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***." (EX1004, [0045].)

l. 1[H]: "capturing, by the microphone in the client device, a second human voice data;"

125. Williams discloses capturing, by the microphone in the client 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" (which is not conceded by Petitioner), Williams expressly discloses that its system can receive and process ***multiple*** human voice commands.

126. First, Williams discloses that each application can receive multiple human voice commands, for example, "[a] cloud-based application may be configured to received ***vocal commands*** from users and to perform operations based on those commands. For example, a user may speak a grocery order, and the spoken order may be interpreted using the voice recognition functionality described above." (EX1004, [0044].) "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. 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." (EX1004, [0036].)

127. Williams also discloses multiple applications, for example, "a ***plurality of cloud-based applications or services 120*** that interact with the application engine 118 and that base their services at least in part on information obtained or derived from the in-home audio interface devices 116 . . . 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." (EX1004, [0021].) This means

that Williams can capture multiple commands, which may be processed by different cloud-based applications or services.

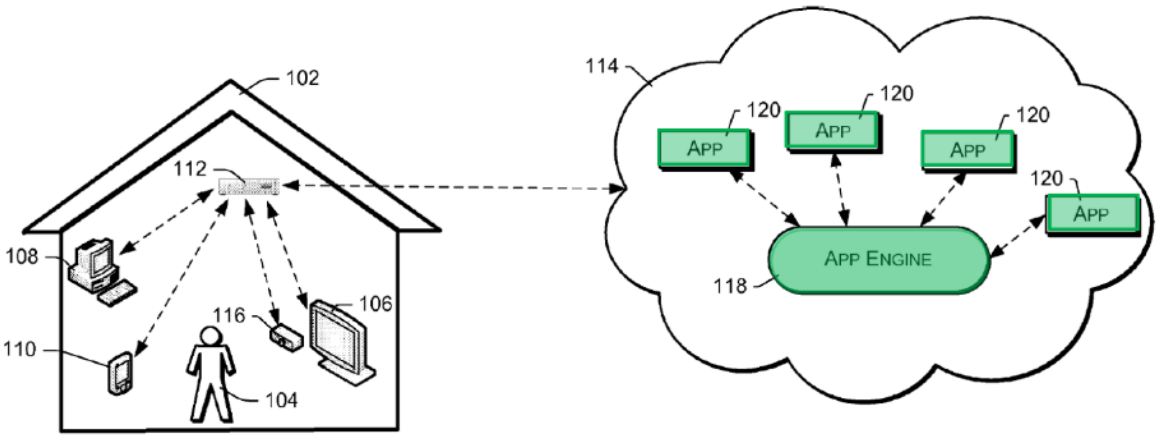


FIG. 1

128. Thus, a POSITA would understand that Williams' system can receive ***multiple*** human voice commands. Williams discloses various audio commands such as '***What's the recipe for deviled eggs***?' and inquiries "***about a flight or other event***." (EX1004, [0043], [0045], [0048], [0050].) Any of the human voice commands disclosed in Williams that cause an actuator (e.g., ***speaker 206***) in the client device (e.g., ***interface device 116***) to operate discloses the recited "first human voice data."

129. Further, Williams discloses human voice commands that can effectuate remote device control. 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." (EX1004, [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***." (EX1004, [0021]; *see also id.*, [0036].)

130. For example, in FIG. 1, "[w]ithin the home 102 is a user 104 and several devices associated with the user 104. The illustrated devices include a ***media device 106***, representing any one or more of various types of devices such as televisions, audio players, video players, and so forth. The illustrated devices also include a ***computer 108***, which represents one or more of various types of devices such as laptop computers, desktop computers, tablet computers, netbooks, other network-enabled devices, and so forth. A ***personal computing device 110*** may also be associated with the user, such as a smartphone, pager, PDA (personal digital assistant), book reader device, or other type of portable device, including various types of electronic communications devices." (EX1004, [0011].)

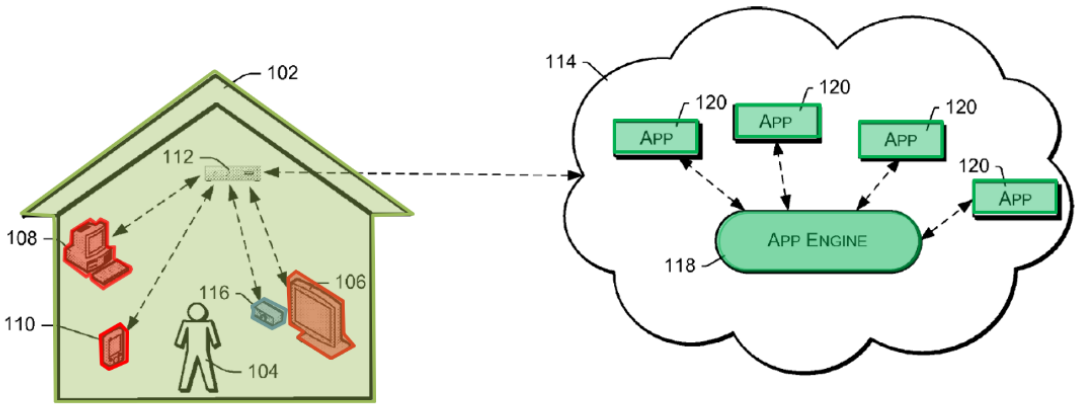


FIG. 1

131. Any of the ***human voice commands*** disclosed in Williams that cause an actuator in an in-home device (e.g., "***entertainment devices, game boxes, media devices, home automation and security devices, communications devices***") to operate discloses the recited "second human voice data." (*Infra* § IX.A.2.r.) A POSITA would have understood that the ***second human voice data*** (to the extent it is different from the first human voice data) is captured by the microphone in the client device and sent to the application engine 118 [run on ***server 400***] and application 120 [run on ***server 400***] in the same manner as shown in FIG. 2. Alternatively, the "audio 210" shown can include or be both the first and second human voice data because "audio 210" results in both first and second messages. 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"). (*E.g.*, EX1047, 2:12-16, 34:63-67.)

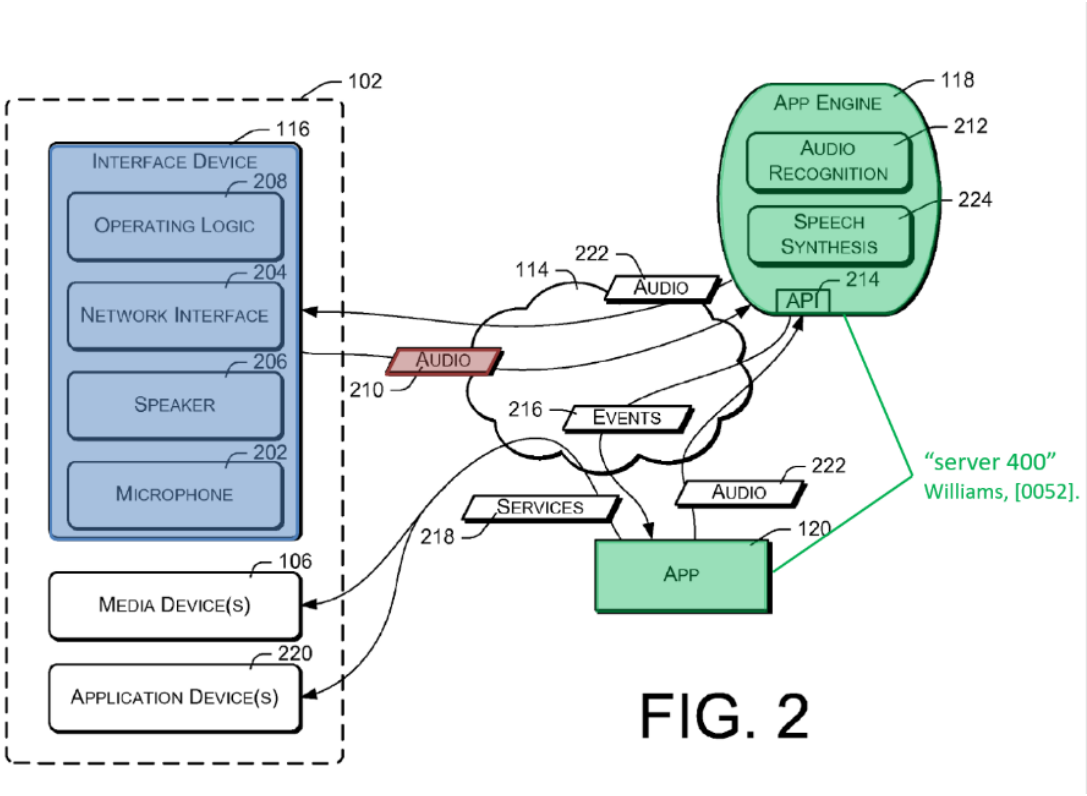


FIG. 2

m. 1[I]: "sending to the server, by the client device via the wireless network, the captured second human voice data;"

132. Williams discloses sending to the server, by the client device via the wireless network, the captured second human voice data. As discussed above, Williams discloses that the client device (e.g., ***interface device 116***) sends captured human voice data to the server (e.g., "***server 400*** that may be used to implement the functionality of the application engine 118 and/or any of the applications 120") via the wireless network. (EX1004, [0052]; *supra* § IX.A.2.f; *see also*, EX1004, [0020], [0022], [0024], [0038]-[0039].) In addition, Williams discloses sending,

receiving, and processing multiple human voice commands. (*Supra* § IX.A.2.l.) Thus, Williams discloses sending to the server, by the client device via the wireless network, the captured second human voice data.

n. 1[J]: "receiving, by the server over the Internet, the captured second human voice data;"

133. Williams discloses receiving, by the server over the Internet, the captured second human voice data. As discussed above, Williams discloses that the server (e.g., "***server 400*** that may be used to implement the functionality of the application engine 118 and/or any of the applications 120") receives the captured human voice data over Internet. (EX1004, [0052]; *supra* § IX.A.2.g; *see also*, EX1004, [0022], [0027], [0034], [0039].) In addition, as discussed above, Williams discloses sending, receiving, and processing multiple human voice commands. (*Supra* § IX.A.2.l). Thus, Williams discloses receiving, by the server over the Internet, the captured second human voice data.

o. 1[K]: "processing, by the server, the captured second human voice data;"

134. Williams discloses processing, by the server, the captured second human voice data. As discussed above, Williams discloses processing of the captured human voice data at by server (e.g., "***server 400*** that may be used to implement the functionality of the application engine 118 and/or any of the applications 120"). (EX1004, [0052]; *supra* § IX.A.2.h; *see also*, EX1004, [0022],

[0027], [0034], [0039].) The second human voice data can be one or more "spoken commands from a user regarding on-premises devices." (EX1004, [0051].) Williams discloses sending, receiving, and processing multiple human voice commands. (*Supra* § IX.A.2.l.) Thus, Williams discloses processing, by the server, the captured second human voice data.

135. For example, Williams discloses that the "***server 400*** that may be used to implement the functionality of the application engine 118 and/or any of the applications 120" processes captured human voice data. (EX1004, [0052].) "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 [run on ***server 400***], which can act on the information to provide or enhance services to the user 104." (EX1004, [0022].)

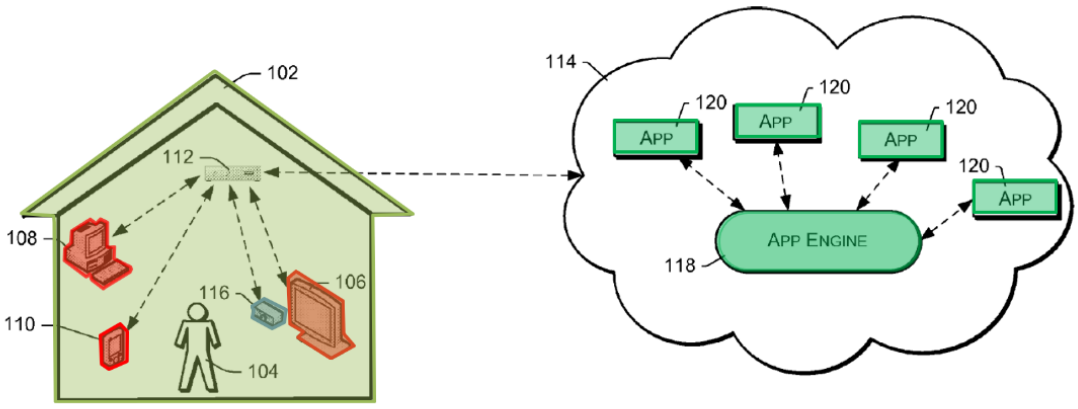


FIG. 1

136. With reference to FIG. 2, Williams discloses that "the application engine 118 [run on ***server 400***] is configured to receive the audio-based information 210. Depending on the type or format of the received information, 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." (EX1004, [0027].)

137. Williams discloses that the processing can be done by the combination of the application engine 118 and application(s) 120, which Williams discloses to be implemented by "the ***server 400***." (EX1004, [0052].) For example, "the 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. With appropriate authentication, the application 120 may query the application engine 118 through the API 214 to obtain and receive event descriptors or notifications 216 regarding events and conditions within the home 102. Authorization for an application 120 to receive information regarding a particular home or installation will typically be granted by a user associated with the home--only those applications authorized by the user are allowed to access the in-home information." (EX1004, [0035].)

138. FIG. 3 is "a flow diagram illustrating an example process that may be implemented within the environment of FIGS. 1 and 2." (EX1004, [0005].) With reference to FIG. 3, "[a]n action 302 comprises receiving audio-based information from premises associated with one or more users, such as a home, office, automobile, etc." and "[a]t 304, the audio-based information is ***processed*** to create data structures identifying current user-specific activities, events, status, environment, commands, and media within the user premises." (EX1004, [0039]-[0040].) "Processing 304 may include speech recognition, speech-to-text conversion, music recognition, and other types of audio recognition, and may result in text-based or code-based notifications or commands." (*Id.*) Additionally, "[a]t 306, the identified events, commands, and so forth are reported to applications that have been authorized by the user" and "[a]t 308, applications receive reported event and command notifications." (EX1004, [0041]-[0042].)

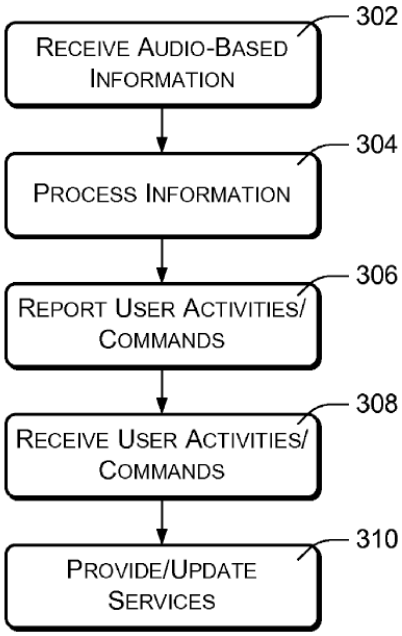


FIG. 3

p. 1[L]: "responsive to the processing, sending a second message, by the server to the controlled device over the Internet;"

139. Williams discloses responsive to the processing, sending a second message, by the server to the controlled device over the Internet. First, Williams discloses the ***server*** sends a ***second message*** to the ***controlled device***. 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***." (EX1004, [0036], *see also*, *id*., [0021], [0038], [0040].)

140. As shown in FIG. 2 below, William's "***server 400***. . . implement[ing] the functionality of the application engine 118 and/or any of the applications 120" sends *services 218* to ***media devices 106 and application devices 220***. (*See also* EX1004, [0052]-[0056].)

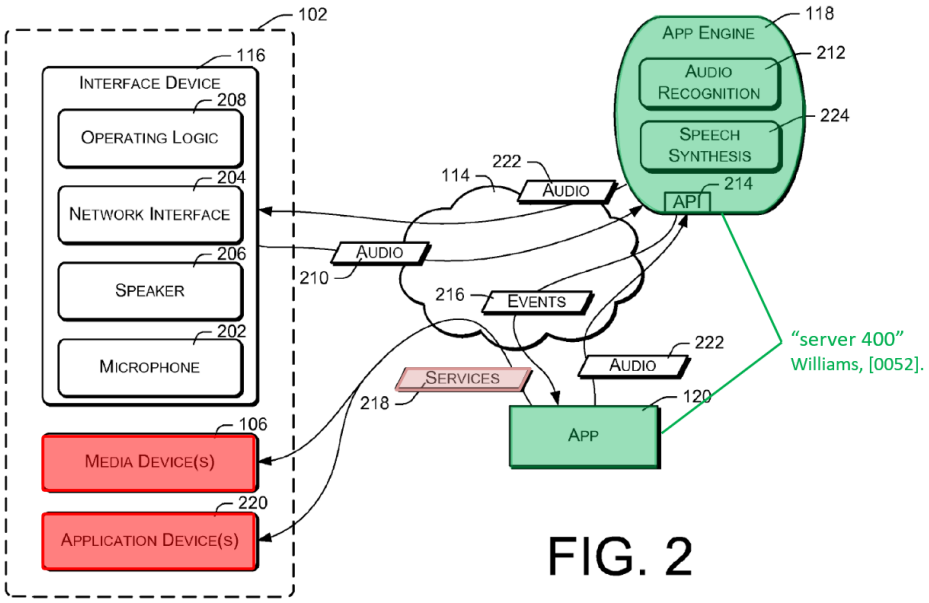


FIG. 2

141. 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." (EX1004, [0042].)

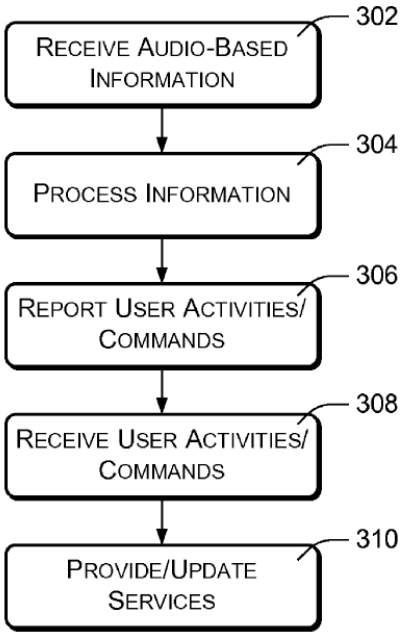


FIG. 3

142. A POSITA would have understood that providing *services 218* to a controlled device (e.g., ***media devices 106 and/or application devices 220***) involves sending information (second message) to the controlled device. Williams discloses that providing services includes "providing information." (EX1004, [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." (EX1004, [0036].) A POSITA would have understood that, in order to operate a controlled device, some information (e.g., a control message related to the service 218) is sent to the controlled 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." (EX1018, 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***." (*Id.*, 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***." (EX1007, claim 1; *see also*, EX1011, [0012].)

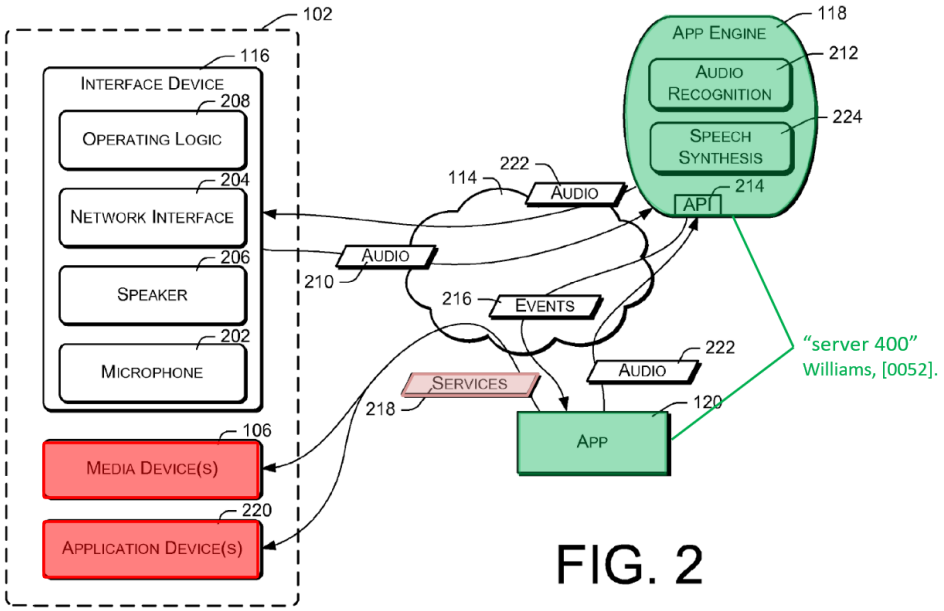


FIG. 2

143. Williams discloses the ***server*** sends a ***second message*** to the ***controlled device*** "responsive to the processing" as recited. With reference to FIG. 2, Williams discloses "the network-based services act ***in response to*** the spoken commands." (EX1004, claim 20.) Further, "[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." (EX1004, [0036].) Further, with reference to FIG. 3, Williams discloses that "[a]t 310, the applications 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." (EX1004, [0042]; *supra* IX.A.2.i.)

144. Williams discloses that the server sends a ***second message*** to the ***controlled device over the Internet***. As shown in FIG. 2, the ***services 218*** is sent via the wireless network 114 and received by ***media devices 106 and/or application devices 220***. As discussed, Williams discloses that "[t]he 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." (*Supra* § IX.A.2.d; EX1004, [0015].)

145. Referring to FIG. 1, it would have been obvious that such information is received by the router 112 (which is connected to the ***cloud server*** via the Internet or network 114) and then wirelessly transmitted to the ***controlled devices*** via the home wireless network.

q. 1[M]: "receiving, by the controlled device via the wireless network, the second message; and"

146. Williams discloses receiving, by the controlled device via the wireless network, the second message. As discussed above, Williams discloses that responsive to the processing, the server (e.g., "***server 400*** . . . implement[ing] the functionality of the application engine 118 and/or any of the applications 120") sends a second message (e.g., information related to ***services 218***) to the controlled

device (e.g., ***media devices 106 and application devices 220***) over the Internet. (*Supra* § IX.A.2.p.)

147. As shown in FIG. 2 below, William's ***media devices 106 and application devices 220*** receives *services 218* or information related to *services 218* from the "***server 400*** . . . implement[ing] the functionality of the application engine 118 and/or any of the applications 120". (*See also* EX1004, [0052]-[0056].) Williams discloses that, "the 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 ***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***". (EX1004, [0036]; *see also id.*, [0042].)

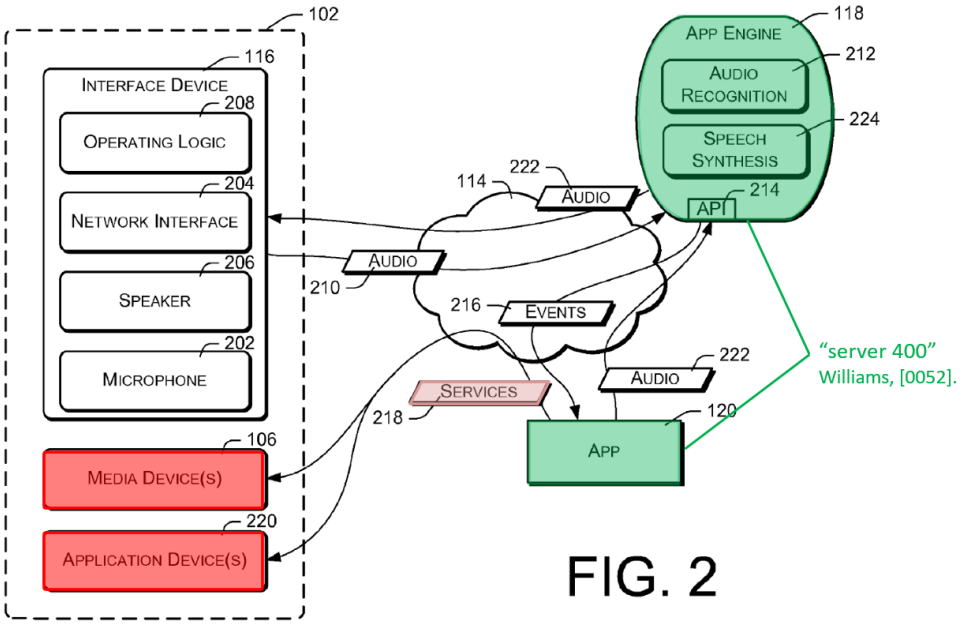


FIG. 2

148. A POSITA would have understood that providing ***services 218*** to a controlled device (e.g., ***media devices 106 and/or application devices 220***) involves having the controlled device receive information needed to provide the services. FIG. 2 expressly shows 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 "***services 218*** may be provided as web-based services that are accessible through web or Internet browsers, in conjunction with any suitable hardware. Services 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." (EX1004, [0036].) A POSITA would understand that in order to operate a controlled device, information related to the services 218 (e.g., a control message) is received by the controlled device. As discussed, for example, it was common knowledge that a control signal would be sent to an appliance to operate 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." (EX1018, 3:19-22.)

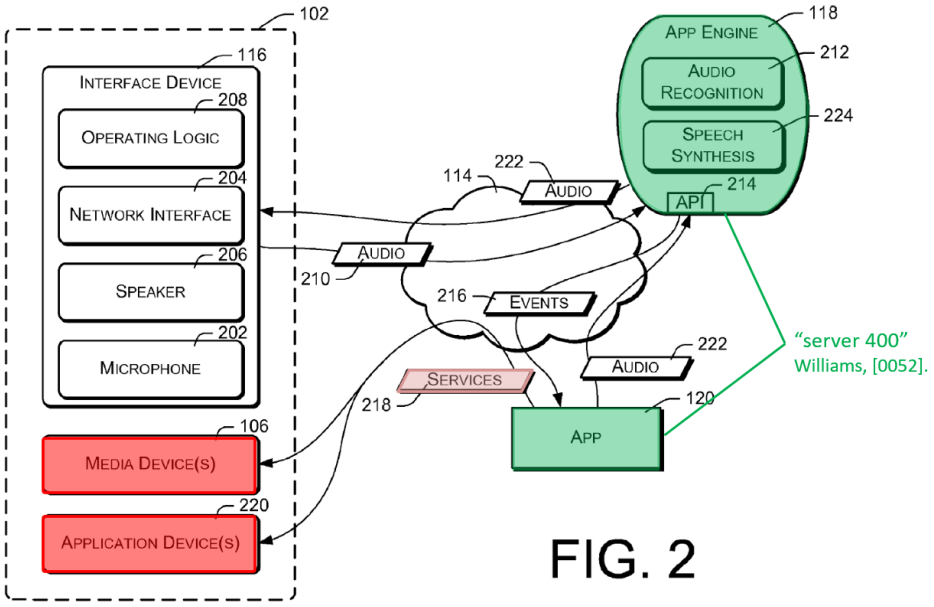


FIG. 2

149. Williams discloses that the ***controlled device*** receives the *second message* via the wireless network. As shown in FIG. 2, the ***media devices 106 and/or application devices 220*** receives the *services 218* or information related to *services 218* via the ***wireless network 114***. (*See also supra* § IX.A.2.b.)

150. Further, 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***." (EX1004, [0012].) Thus, the ***media devices 106 and/or application devices 220*** receives the *services 218* or information related to *services 218* via "***a router or access point 112***, which is in turn connected to ***a public network 114***". (*Id.*) Williams discloses that other types of wireless networks can be used. "Although the devices within the home

102 are illustrated as connecting to the network 114 via the in-home router 112, network connectivity might be implemented in various other ways. For example, 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." (EX1004, [0013].)

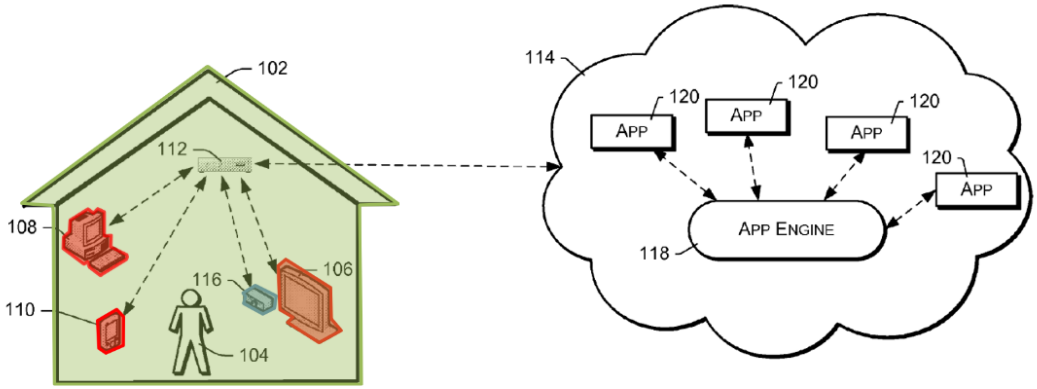


FIG. 1

r. 1[N]: "operating a second actuator in the controlled device in response to the received second message."

151. Williams discloses operating a second actuator in the controlled device in response to the received second message. As discussed above, Williams discloses the ***media devices 106 and/or application devices 220*** receives the *services 218* or information related to *services 218*. (*Supra* § IX.A.2.p - IX.A.2.q.)

152. Williams discloses that, in response to the received *services 218* or information related to *services 218*, actuators in the ***controlled devices*** are

operated. For example, "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. Such services may be provided through ***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***". (EX1004, [0036]; *see also id.*, [0042].) It would have been obvious that the second actuator (i.e., any of the actuators in the controlled devices) is operated in response to the information (second message) received from ***server 400***.

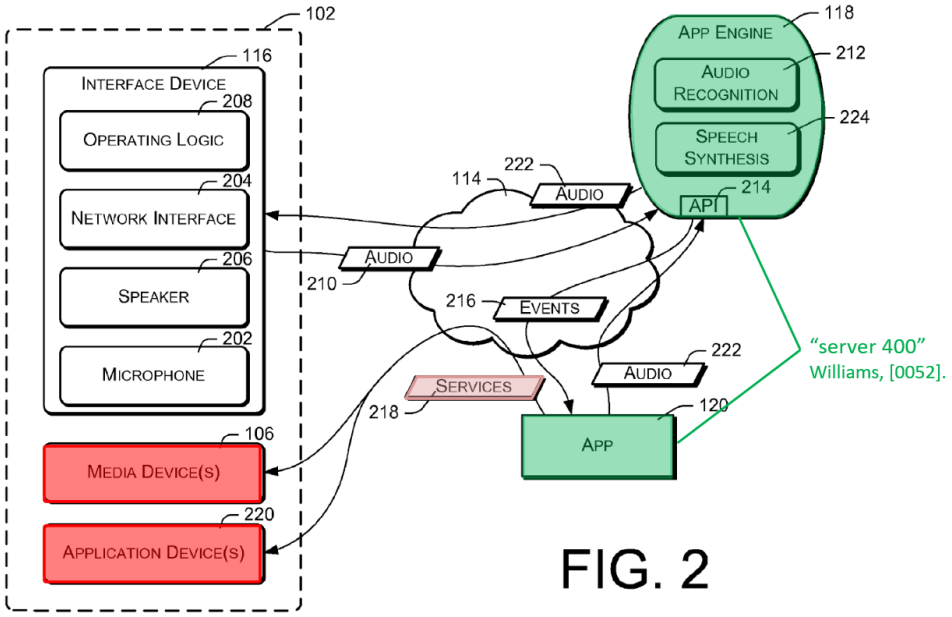


FIG. 2

153. Williams further 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***. 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." (EX1004, [0051]; *see also* claim 20.) "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." (EX1004, [0021].)

154. Williams discloses a variety of services that involve operation of actuators in ***controlled 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." (EX1004, [0036].)

155. A POSITA would have understood that "***control[ling]*** various ***on-premises devices and equipment***, such as ***media equipment*** and ***home-automation devices***" involves operating various "actuators" in media equipment and home-automation devices.

156. One example of the "second actuator" is an actuator providing visual information. The '710 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, 40:51-41-7, *see also, id.*, 109:41-46, 110:60-66; 144:64-145:1.) It would have been obvious that the second message would lead to operation of a display and/or a speaker of "***media device 106***, representing any one or more of various types of devices such as ***televisions, audio players, video players***" within the home. (EX1004, [0011], [0009]-[0010], [0021], [0036], [0054].)

157. 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. (EX1004, [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." (EX1004, [0021].) A POSITA would have understood that controlling media devices 106 involves operating display actuators in media devices to provide visual

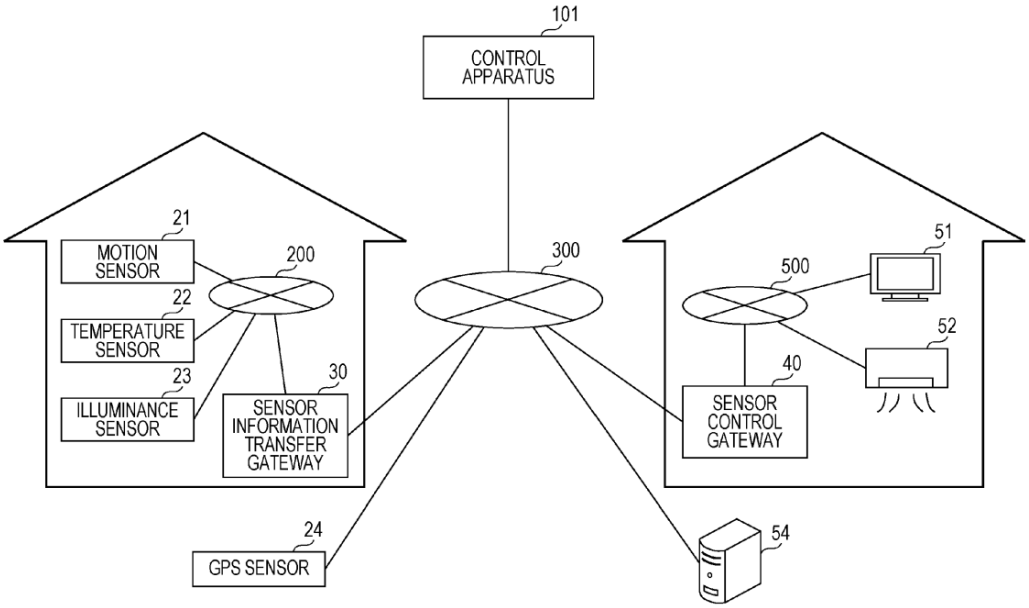
information. Further, Williams discloses that "the user may ask for directions to a certain location. The application may send a map, textual directions, and computerized routes to the user via the user's ***personal communications device***, or may send information to the user's ***navigation device*** regarding the desired location." (EX1004, [0049].) It is obvious to a POSITA that the "map, textual directions, and computerized routes" may be provided via the speaker and or displayed on a controlled device.

158. Another example of the "second actuator" is an actuator providing sound. 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. (EX1004, [0011]; *see also id.*, [0010].) Williams discloses "services may be provided through any of the media devices 106 located in the home." (EX1004, [0036].) For example, Williams discloses providing "content items that are supplied to consuming devices such as electronic reader 104." (EX1004, [0054].) A POSITA would have understood that controlling media devices 106 (e.g., to playback a media item) involves operating actuators in media devices to control sound. As discussed, it was common knowledge that a speaker has an actuator to produce sound by converting 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. (*Supra* IX.A.2.k.) The operation of speakers generally adopts the moving-coil principle, "[m]ost loudspeakers use the moving-coil principle … a signal current through the coil will then vibrate the diaphragm, producing sound waves." (EX1017, 122.)

159. This is consistent with prior findings of the Board in its Decision on Appeal affirming the Examiner's rejections of claims 1-30 of the '710 patent that an image display device such as a TV is a controlled 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." (EX1003, 396.)

FIG. 15



(EX1024, FIG. 15.)

160. Another example of the "second 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." (EX1004, [0021]; *see also id.*, [0036] and [0051].) A POSITA would have understood that controlling home automation and security devices involves operating motion actuators, such as automated doors or windows, automated lights, automated HVAC systems (e.g., fans), automated thermostats, automated cameras, automated locks, automated kitchen appliances. Indeed, at the

time of the invention, it was well-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 function." (EX1019, 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." (*Id.*)

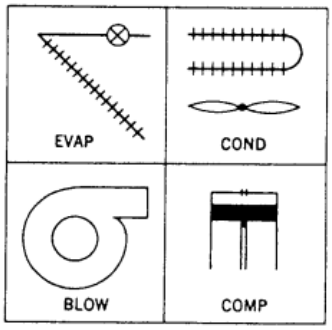


Fig 13-1. Self contained air conditioner

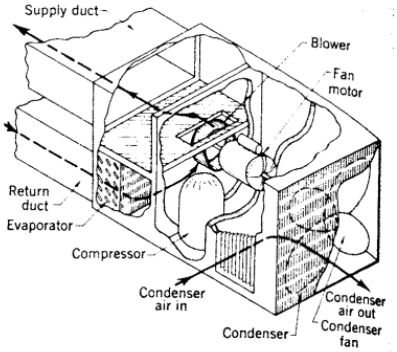


Fig 13-3. Horizontal self contained air conditioned unit

161. Further, as discussed, the Board agreed in its Decision on Appeal affirming the Examiner's rejections of claims 1-30 of the '710 patent that 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." (EX1003, 396.)

3. Claim 5: "The method according to claim 1, wherein each of the first and second actuators is directly or indirectly affecting, changing, producing, or creating a physical phenomenon."

162. Williams discloses this limitation. As discussed, Williams discloses operating a speaker (i.e., a first actuator) in the client device in response to the received first message. (*Supra* § IX.A.2.k.) Operating the speaker involves operating an actuator affecting, changing, producing, or creating sound or audio. (*Id.*)

163. Second, the controlled device disclosed in the Williams system includes "***on-premises devices and equipment***, such as ***media equipment*** and ***home-automation devices***." (*Supra* § IX.A.2.r.) The media devices disclosed includes "any one or more of various types of devices such as ***televisions, audio players, video players***." (EX1004, [0011]; *see also id.*, [0010].) As discussed, controlling media devices 106 (e.g., to playback a media item) involves operating sound actuators in media devices to affect, change, produce, or create sound. (*Supra* IX.A.2.k.) Further, controlling media devices 106 (e.g., to playback a media item) involves operating actuators in media devices (e.g., displays) to provide visual information, thus affecting, changing, producing, or creating light. (*Id.*) Further, controlling home automation and security devices involves operating motion actuators, thus affecting, changing, producing, or creating motion, temperature, humidity, etc. (*Id.*)

4. Claim 6: "The method 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. Williams discloses this limitation. As discussed above, Williams discloses a first actuator and a second actuator that directly or indirectly affect, change, produce, or create at least audio, light, motion, sound, temperature, humidity, etc. (*Supra* § IX.A.3.)

5. Claim 16: "The method according to claim 1, wherein the controlled device is integrated in, is part of, or is entirely included in, an appliance."

165. Williams discloses a controlled device that is integrated in, is part of, or is entirely included in, an appliance. As discussed, the controlled device 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.A.2.r; EX1004, [0051], [0011].) These devices are appliances, consistent with claim 23 of the '710 patent.

6. Claim 23: "The method according to claim 16, 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."

166. Williams discloses this limitation. As discussed, the controlled device in the Williams system are devices such as "***televisions, audio players, video***

players". (*Supra* § IX.A.2.r; EX1004, [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* EX1013, 746.)

7. Claim 29: "The method according to claim 27, wherein the operating of the first actuator comprises playing digital audio content that is pre-recorded or synthesized, or wherein the operating of the first actuator comprises simulating the voice of a human being or generating music, or wherein the operating of the first actuator comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

167. Williams discloses this limitation. 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.A.2.k; EX1004, [0037].) Further, "[t]he 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." (*Id*., *see also id*., [0045], [0048], and [0050].) As another 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***." (*Id*., [0046].) Thus, Williams discloses 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.

B. Ground 2: Williams and Wittenberg Renders Obvious Claim 2

1. Summary of Wittenberg

168. Wittenberg was filed on January 21, 2005, published on May 5, 2009, and is prior art under §102(a), (b), (e). (EX1007, 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." (EX1007, 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." (EX1007, Abstract, Claim 1.)

2. Claim 2: "The method according to claim 1, wherein the processing comprises performing a voice recognition algorithm for identifying the voice of a specific person."

169. Claim 1 is obvious in view of Williams. (*Supra IX.A.2*) Claim 2 is obvious in further view of Wittenberg. Wittenberg discloses, with reference to

FIG. 2, that "voice 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." (EX1007, 5:61-6:6.)

170. At the time of the invention, it would have been obvious to combine the teachings of Wittenberg with Williams. Both Wittenberg and Williams are analogous art to the '710 patent. Both 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." (EX1004, [0036].) 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* EX1007, Claim 1.) Thus, Williams and Wittenberg are from the same field of endeavor and reasonably pertinent to the same problem of controlling devices based on audio commands and ensuring security within the command/control system architecture.

171. A POSITA would have been motivated to combine Williams with 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." (EX1004, [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[.]" (EX1007, Claim 1.) The voice recognition

processing techniques can be combined with the teachings of Williams 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.

172. Wittenberg describes a security issue that unauthorized users may be able to control devices via audio commands. "For 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." (EX1007, 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." (EX1007, 5:61-6:6.)

173. 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." (EX1007, 6:7-13.)

174. The Williams-Wittenberg combination would, for example, result in a system in which Wittenberg's voice recognition capabilities are used to analyze human audio commands in Williams. For example, the Williams-Wittenberg combination can 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.

175. Because both Williams 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 in the same way. Further, it would have been obvious to try combining the teachings of Wittenberg and Williams. 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 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 . . . ." (EX1007, 6:1-3.)

C. Ground 3: Williams and Gruber Renders Obvious Claim 3-4, 10-11

1. Summary of Gruber

176. Gruber was filed on September 30, 2011, and is prior art under §102(e). (EX1008, 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." (EX1008, [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." (EX1008, [0013].)

177. 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." (EX1008, [0218] - [0225].)

2. Claim 3: "The method according to claim 1, wherein the client device further comprises a sensor that outputs sensor data that responds to a physical phenomenon, wherein the method further comprising sending to the server, by the client device via the wireless network, the sensor data, and wherein the first message is sent by the server in response to the sensor data."

178. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 3 is obvious in further view of Gruber. First, Gruber discloses that the client 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." (EX1008, [0218] - [0225].)

179. Gruber also discloses sending to the server, by the client 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." (EX1008, [0179].) Gruber discloses that the "virtual assistant 1002 also ***receives context information 1000*** . . ." (*Id*., [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)*** . . .)" (*Id*., [0080].)

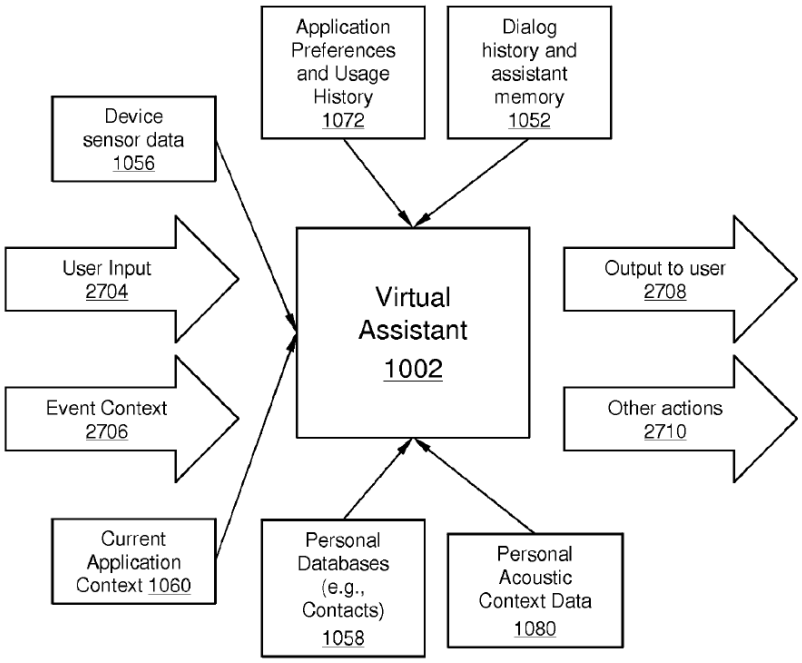


FIG. 1

180. "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***." (EX1008, [0366] – [0367].) Gruber may be implemented "using a web browser 1304A or other software application for enabling communication with servers 1340 via network 1361." (*Id*., [0090].) "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 . . ." (*Id*., [0100] – [0102], [0110] – [0117].)

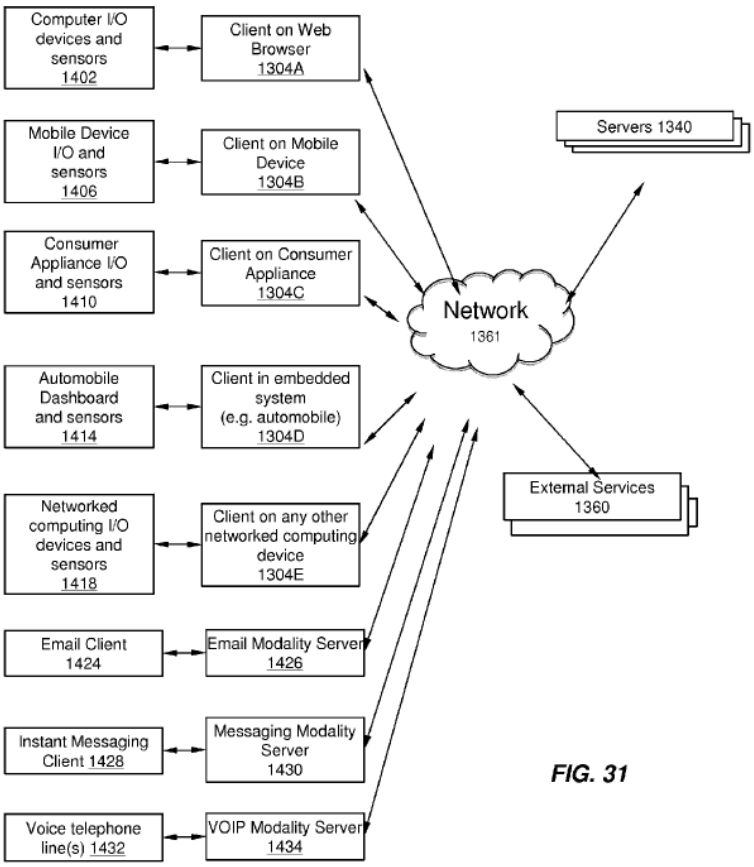


FIG. 31

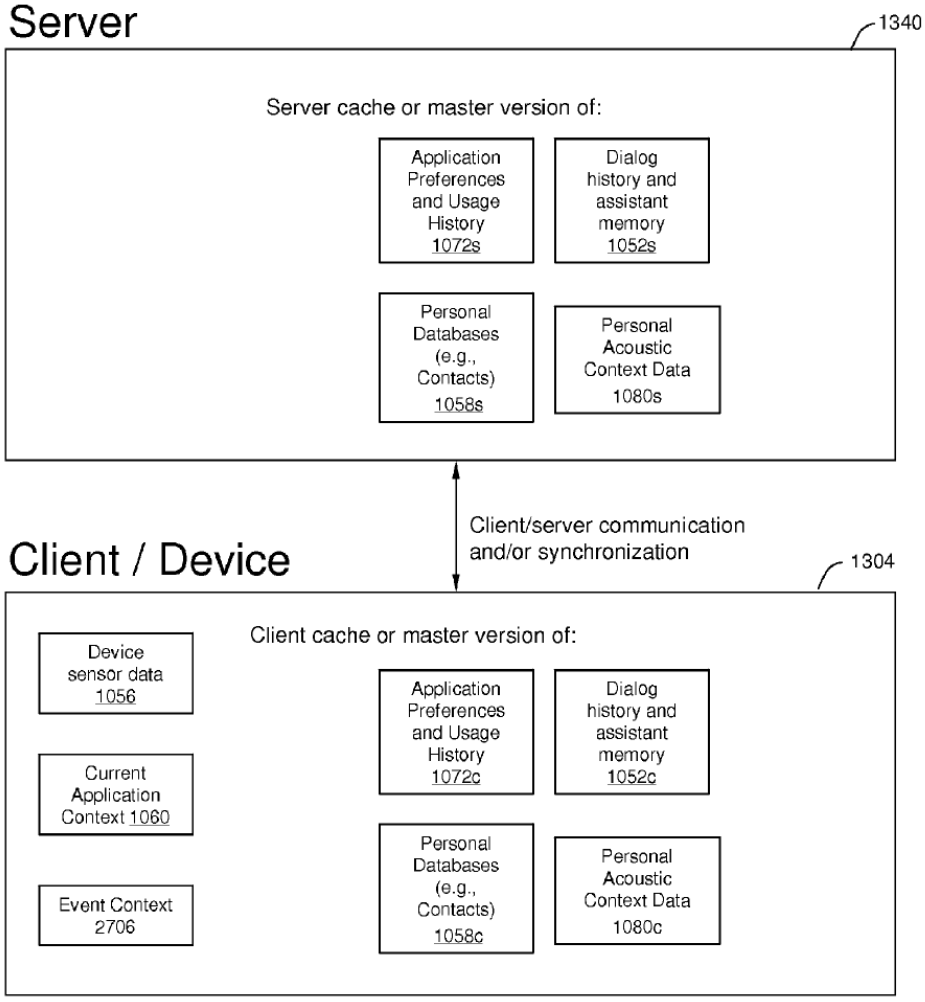


FIG. 6

181. It would have been obvious to a POSITA to combine the teachings of Williams and Gruber. Williams and Gruber are analogous to the '710 patent. All 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." (EX1004, [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." (EX1004, [0036].) 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." (EX1008, [0007].) Thus, Williams and Gruber are from the same field of endeavor and reasonably pertinent to the same technical problems of processing and responding to human speech commands (e.g., to provide services).

182. A POSITA would have been motivated to combine Williams with the speech processing techniques of Gruber to achieve improved speech recognition and with a reasonable expectation of success. 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. "A virtual assistant uses context information to supplement natural language or gestural input from a user. Context 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. Context can be used to constrain solutions during various phases of processing, including, for example, speech recognition, natural language processing, task flow processing, and dialog generation." (EX1008, Abstract.) The technique of using sensor data as contextual information can be combined with the teachings of William to improve the accuracy of speech processing, with a reasonable expectation of success.

183. Specifically, the Williams-Gruber combination would result in a system that can capture sensor data (e.g., temperature) using one or more sensors (e.g., as taught by Gruber) in the interface device and transmit such data and/or related information to Williams' server to supplement audio processing as taught by Williams and/or Gruber. Williams discloses a system that can capture sensor data (e.g., audio) using sensors (microphones) 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* EX1004, [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." (EX1008, [0013].) For example, the Williams-Gruber combination can capture sensor data in the audio interface device and transmit the captured sensor data to Williams' server 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). The sensor data can be particularly helpful for processing commands related to home appliances and other in-home devices. Combining the teachings of Williams and Gruber as such would yield predictable results of improving audio processing.

184. Incorporating Gruber's speech processing techniques into Williams 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 inputs. Because both Williams 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 Williams in the same way. Further, it would have been obvious to try combining the teachings of Williams 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 context data. 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 and Gruber.

185. 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 method 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."

186. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 4 is obvious in further view of Gruber. 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." (EX1008, [0225].) The contextual information from these sensors may include "the current light level, temperature, and other environmental measures."

(EX1008, [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: "the 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). (EX1029, 466.) "Thermometer[] . . . 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." (*Id*., 116.)

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

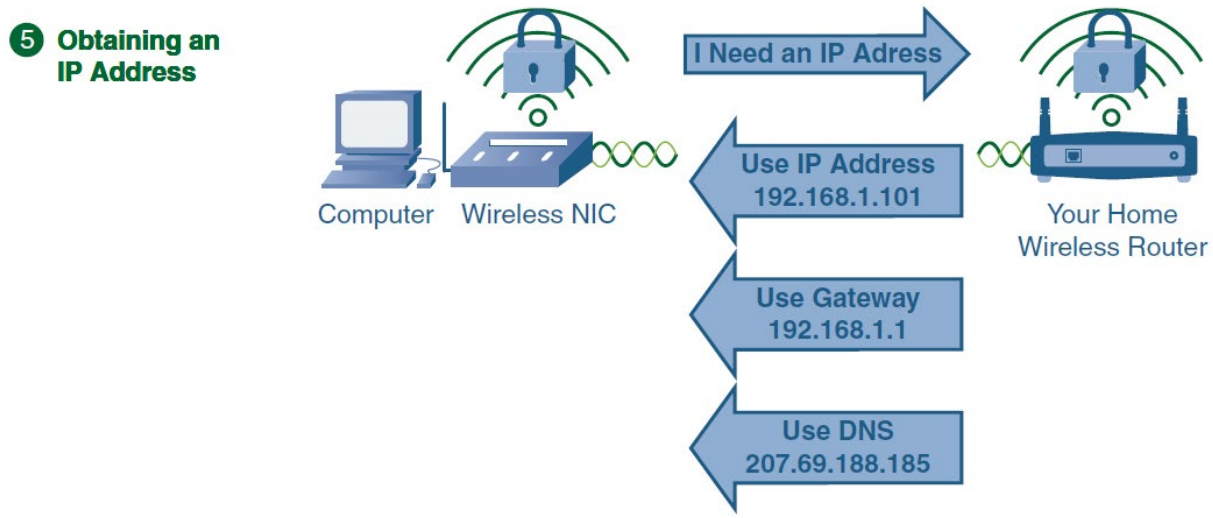
4. Claim 10: "The method according to claim 1, wherein the client device or the controlled device are addressable in the wireless network or the Internet using an address stored in a volatile or non-volatile memory of the respective device for uniquely identifying the respective device in the network."

188. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 10 is obvious in further view of Gruber. As discussed above, Williams discloses 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.C.2; EX1004, [0012] – [0013], [0020], Figure 1.) A POSITA would have

understood that this involves associating each device with a device identifier or address (which is locally stored) that uniquely identifies the device in the network. Indeed, at the time of the invention, it was well-known that networked devices are associated with IP addresses and/or MAC addresses that uniquely identifies the device in the network. "***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 your home network through a wireless signal; [s]ecuring the communications path between the wireless router and your computers and devices connected over wireless . . . " and "***each wireless card has a unique identifier called a MAC address***." (EX1006, 99, 166.) Each networked device is thus associated with a globally unique MAC address for identification of the device in the network.

189. A POSITA would have understood that the associated 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. "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." (*Id*., 108, 210, *see also*, *Id*., 175.) "Internet protocol (IP) addresses are the unique numbers used to identify individual connections to the Internet." (EX1028, 127.) Thus, the IP address assigned to each networked device is unique to the device in the network.



190. Further, as shown in the figure below, the networked devices can display the IP addresses, which means the IP addresses are locally stored. (*Id.*, 211, Figure 10-44, *see also, Id*., 265.)

Figure 10-44 Verifying IP Address Assignment with Windows XP



191. Gruber similarly 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.***" (EX1008, [0218] - [0225].)

192. It would have been obvious to a POSITA to combine the teachings of Williams and Gruber and use the device addresses (e.g., IP addresses, MAC addresses) for wireless communication and/or supplementing speech processing. (*Supra* § IX.C.2.)

5. Claim 11: "The method according to claim 10, wherein the address is a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type or wherein the address is a layer 3 address and is static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address."

193. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 11 is obvious in further view of Gruber. As discussed, Williams discloses 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.C.2, IX.C.4; EX1004, [0012] – [0013], [0020], Figure 1.) As discussed, Gruber discloses that a client device can have signatures such as "IP addresses" and "MAC addresses". (EX1008, [0218] - [0225].) Further, at the time of the invention, it was well-known that networked devices are associated with IP addresses and/or 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***." (EX1006, 99, 166.) "***[E]ach 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. This is called ***dynamic IP addressing***[.]" (*Id*., 210.)

194. In 2012, it was well-known that Internet Protocol (IP) addresses are layer 3 addresses, are either static or dynamic, and are IPv4 type addresses. "In the case of today's Internet, these addresses are Internet protocol version 4 (IPv4) addresses." (EX1028, 127.) "Internet Protocol (IP) is a Layer 3 protocol that

contains addressing and control information that allows packets to be routed." (EX1022, 11.) "The IP system in common use today is known as IPv4, for Internet Protocol version 4. A newer system, IPv6, or Internet Protocol version 6, has been developed and exists today in small deployments." (*Id*., 13.) Doherty discloses that "[t]he broadband networks typically assign a ***dynamic IP address*** to your broadband service. Dynamic means the IP address can change[.]" (EX1006, 277.) Doherty also discloses that the IP address may be a "***static IP address***, which means . . . your address never changes." (*Id*.) Further, at the time of the invention, MAC address based on MAC-48, EUI-48, and EUI-64 are well known. "Most MAC addresses use one of the IEEE standardized protocols; ***MAC-48, EUI-48®, and EUI-64®***." (EX1046, 2:50-59.)

195. It would have been obvious to a POSITA to combine the teachings of Williams and Gruber and use the device addresses for wireless communication and/or supplementing speech processing. (*Supra* § IX.C.2, IX.C.4.)

D. Ground 4: Williams and Benesty Renders Obvious Claims 7-8

1. Summary of Benesty

196. Benesty was published by at least December 11, 2007, and is prior art under §102(a), (b). (EX1034, 1.) The Examiner never found, and Applicant never cited, Benesty during prosecution. Benesty is a textbook titled "Springer Handbook of Speech Processing." (EX1009, Title.) Benesty provides an overview

of well-known speech processing techniques, including speech recognition, speaker recognition, speech enhancement via noise reduction, and microphone arrays.

197. Benesty discloses that "it 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." (EX1009, 1021.)

198. 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*.

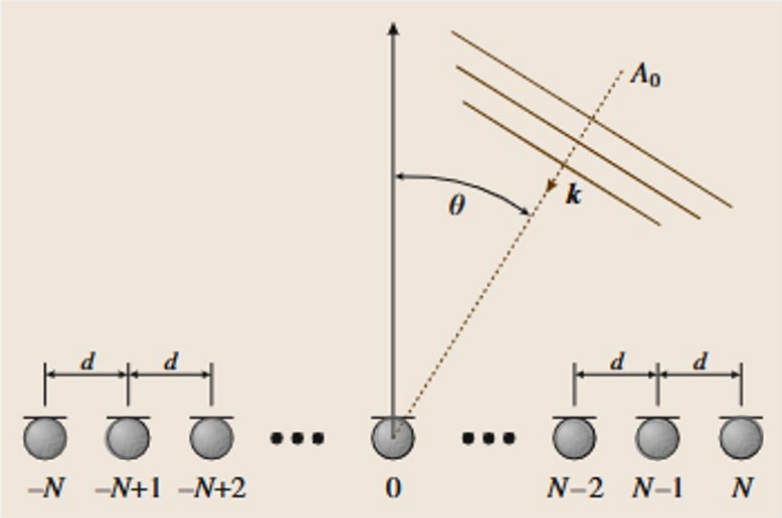


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N+1$ microphones spaced by d

199. 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 lead 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." (EX1009, 1044.)

200. 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 method according to claim 1, wherein the client device comprises multiple microphones, and wherein the capturing comprises capturing, by the multiple microphones in the client device, the human voice data."

201. Claim 1 is obvious in view of Williams. (*Supra IX.A.2*) Claims 7 and 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***." (EX1009, 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." (*Id*.)

202. 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*.

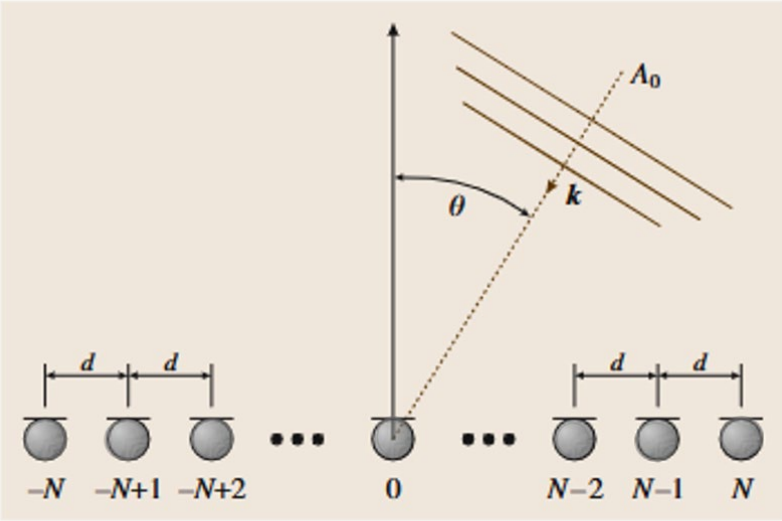


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

203. It would have been obvious to a POSITA to apply Benesty's techniques of including multiple microphones to Williams' client device . Williams and Benesty are analogous to the '710 patent. Both Williams and Benesty are related to capturing and processing human audio data. Williams is directed to hands-free audio communication of a user with a client 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." (EX1004, [0003].) 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." (EX1009, 1021.) Thus, Williams and Benesty are from the same field of endeavor and reasonably pertinent to the same technical problem of providing hands-free audio communication between users and devices.

204. The Williams-Benesty combination would result in a system where multiple microphones as taught by Benesty are incorporated in Williams's client device to capture human voice data to improve the quality of captured audio. For example, Benesty's microphone array(s) may be included in Williams' audio interface device to capture the first and second human voice data. Because Williams and Benesty relate to audio capturing and processing, a POSITA would be motivated to use the known techniques of Benesty to improve Williams' 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.

205. Combining the teachings of Williams 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, as there is a finite number of identified, predictable resolutions, without requiring much 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 method according to claim 7, wherein the multiple microphones 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."

206. 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 lead 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." (EX1009, 1044.)

207. Benesty also discloses that "alternatively, 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." (EX1009, 1044.)

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

E. Ground 5: Williams and Goldsmith Renders Obvious Claims 12-15

1. Summary of Goldsmith

209. Goldsmith was published by at least January 20, 2006, and is prior art under §102(a), (b). (EX1035, 2.) The Examiner never found, and Applicant never cited, Goldsmith during prosecution. Goldsmith is a textbook titled "Wireless Communications." (EX1035, Title.) Goldsmith provides an overview of wireless communication techniques, including the wireless systems available at the time of publication in 2005.

210. 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. (EX1035, 8-21.)

211. 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." (EX1035, 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 were developed to provide higher data rates than 802.11b. The ***IEEE 802.11a wireless LAN*** standard . . . the other standard, ***802.11g***." (*Id*., 15-16.)

212. Goldsmith discloses that both Bluetooth, and 802.11b WLAN are ISM Band II frequencies. (EX1035, 22-23.)

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

213. 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." (EX1035, 12-13.) "The ***third generation (3G) cellular systems are based on a wideband CDMA standard***." *Id.* "***Enhanced Data Services for GSM Evolution (EDGE)***, further increases data rates using a high-level modulation format combined with FEC coding." *Id.* "***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)***." (EX1035, 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.

214. 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 12: "The method according to claim 1, wherein the wireless network is a Wireless Personal Area Network (WPAN), that is according to, or based on, Bluetooth™ or Institute of Electrical and Electronics Engineers (IEEE) 802.15.1-2005 standards, or wherein the WPAN is a wireless control network that is according to, or based on, Zigbee™, IEEE 802.15.4-2003, or Z-Wave™ standards."

215. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 12 is obvious in further view of Goldsmith. Williams discloses devices communicating over a wireless network. For example, 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***." (EX1004, [0012].)

216. Williams also discloses that "interface devices might use ***cellular or other wireless data communications technologies*** to communicate with the application engine 118," 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." (EX1004, [0020], [0013].) Specifically, Williams points out that all devices in the home, the cloud-based application engine, and cloud-based applications may communicate using ***existing networking protocols and techniques***. "[V]arious network communications between the components of FIG. 2 may be performed using existing networking protocols and techniques." (EX1004, [0038].)

217. Goldsmith discloses Bluetooth and Zigbee as two wireless communication systems designed to be compliant with the IEEE 802.15 standards, a family of Wireless Personal Area Networks (WPANs).

218. First, Goldsmith discloses both Bluetooth and Zigbee as wireless communication systems that can be embedded in electronic devices for smart homes. "As radios decrease their cost and power consumption, it becomes ***feasible***

to embed them in more types of electronic devices, which can be used to create smart homes, sensor networks, and other compelling applications. Two radios have emerged to support this trend: ***Bluetooth and Zigbee***." (EX1035, 19.) "The Bluetooth standard was developed jointly by 3 Com, Ericsson, Intel, IBM, Lucent, Microsoft, Motorola, Nokia, and Toshiba. The standard has ***now been adopted by over 1300 manufacturers, and many consumer electronic products incorporate Bluetooth***, including wireless headsets for cell phones, wireless USB or RS232 connectors, wireless PCMCIA cards, and wireless settop boxes." (EX1035, 20.)

219. Then, Goldsmith provides details of Bluetooth and Zigbee and discloses that they are designed to be compliant with the IEEE 802.15 standards (WPANs). Goldsmith "summarizes the main characteristics of Zigbee, Bluetooth, and UWB, which have emerged to support a wide range of short distance wireless network applications. These specifications are designed to be ***compliant with the IEEE 802.15 standards***, a family of IEEE standards for short distance wireless networking called ***Wireless Personal Area Networks (WPANs)***." (EX1035, 600-601.)

220. Goldsmith provides details of Bluetooth, "Bluetooth radios provide ***short range connections between wireless devices*** along with rudimentary networking capabilities. The Bluetooth standard is based on a tiny microchip incorporating a radio transceiver that is built into digital devices. The transceiver

takes the place of a connecting cable for devices such as cell phones, laptop and palmtop computers, portable printers and projectors, and network access points. Bluetooth is mainly for short range communications, e.g. from a laptop to a nearby printer or from a cell phone to a wireless headset. Its normal range of operation is 10 m (at 1 mW transmit power), and this range can be increased to 100 m by increasing the transmit power to 100 mW. The system operates in the unlicensed 2.4 GHz frequency band, hence it can be used. worldwide without any licensing issues. The Bluetooth standard provides 1 asynchronous data channel at 723.2 Kbps. In this mode, also known as Asynchronous Connection-Less, or ACL, there is a reverse channel with a data rate of 57.6 Kbps. The specification also allows up to three synchronous channels each at a rate of 64 Kbps. This mode, also known as Synchronous Connection Oriented or SCO, is mainly used for voice applications such as headsets, but can also be used for data. These different modes result in an aggregate bit rate of approximately 1 Mbps. Routing of the asynchronous data is done via a packet switching protocol based on frequency hopping at 1600 hops per second. There is also a circuit switching protocol for the synchronous data" (EX1035, 19-20.)

221. Goldsmith also discloses the specifics of Zigbee as a wireless communication system. "The ZigBee radio specification is designed for lower cost and power consumption than Bluetooth. ***The specification is based on the IEEE***

802.15.4 standard. ***The radio operates in the same ISM band as Bluetooth***, and is capable of connecting 255 devices per network. The specification supports data rates of up to 250 Kbps at a range of up to 30 m. These data rates are slower than Bluetooth, but in exchange the radio consumes significantly less power with a larger transmission range. The goal of ZigBee is to provide radio operation for months or years without recharging, thereby targeting applications such as sensor networks and inventory tags." (EX1035, 20.)

222. At the time of the invention, routers having Bluetooth capabilities were well-known. "A wireless router such as 802.11 router, 802.16 router, WiFi router, WiMAX router, ***Bluetooth router***, X10 router can be connected to the mesh network." (EX1025, 59:21-24.) "[T]he wireless network router 31 may be a wireless LAN router, such as an IEEE 802.11x or ***Bluetooth router***." (EX1026, 4:25-27.) "[T]he wireless router comprises one of: 802.11 router, 802.16 router, WiFi router, WiMAX router, ***Bluetooth router***, ×10 router." (EX1027, 4:20-22.) Further, Bluetooth access points/gateways with connection to the Internet were well-known. For example, a patent application filed on March 29, 2002, discloses a mobile computing device "coupled, wirelessly through a Bluetooth network interface card (NIC) to ***Bluetooth access point*** 130, which, in turn, is coupled to the wired ***Ethernet*** backbone 120." (EX1032, 1:67-2:3.) A POSITA would have understood that, by coupling to the Ethernet backbone, the Bluetooth access point

is connected to the Internet. For example, access to "a remote server 500 accessible via a network 502 such as ***the Internet***," "may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router***." (EX1020, [0027].)

223. It would have been obvious for a POSITA to apply Goldsmith's disclosure of various types of wireless networking techniques and protocols to Williams. Both Williams and Goldsmith are related to wireless communication technologies. Williams expressly discloses usage of "cellular or other wireless data communications technologies" between interface device and application engine and "existing networking protocols and techniques" for various network communications between the components of FIG. 2 (which includes the client device, the controlled device, and the server device). (EX1004, [0020], [0038], [0012]-[0013].) Goldsmith discloses various existing wireless connection systems in a chapter titled "1.4 Current Wireless Systems" (EX1035, 8-21.) Thus, Williams and Goldsmith are from the same field of endeavor and reasonably pertinent to the same technical problem of providing wireless communication between devices.

224. It would have been obvious for a POSITA to look to Goldsmith for conventional wireless networking techniques and protocols that can be applied to Williams. A POSITA would have looked to Goldsmith's disclosure of Bluetooth

and Zigbee as two conventional wireless communication protocols and understood that the wireless data communications between devices as disclosed in Williams can be implemented by Bluetooth and Zigbee in the same way (i.e., the interface device and controlled device may communicate wirelessly via Bluetooth or Zigbee). For example, the interface device and controlled device may communicate via Bluetooth with a Bluetooth router or another device that can manage and relay Bluetooth connections, and the router/device can communicate with the cloud server over the Internet. Goldsmith itself discloses that it is feasible to embed Bluetooth and Zigbee in electronic devices to create "smart homes." (EX1035, 19.) A POSITA would be motivated to use the known wireless networking techniques of Bluetooth or Zigbee to improve the usability of Williams and expand its applicability and range in the same way. Combining the teachings of Williams and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

225. Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (e.g., Bluetooth and Zigbee) with the teachings of Williams, without requiring much 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 such as explained by Goldsmith. There was a finite

number of potential wireless network communications protocols used for a variety of different coverage areas, such as infrared, Wi-Fi, 3G, 4G, UWB, and GSM, that were commonly available and can be easily applied in Williams at the time of the invention. The known benefits and feasibility of including a known wireless networking technique such as Bluetooth and Zigbee as taught by Goldsmith to connect in-home devices for creating smart home would prompt the application of Goldsmith in the same field based on design incentives and other market forces.

3. Claim 13: "The method according to claim 1, wherein the wireless network is a Wireless Local Area Network (WLAN) that is according to, or base [sic] 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."

226. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 13 is obvious in further view of Goldsmith. 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.

227. Goldsmith discloses that wireless LAN was well-known and widely adopted. "Wireless LANs provide high-speed data 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 U.S. operate with 80 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 150m. 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.)

228. 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 20-70 Mbps data rates. Since 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, "the other standard, ***802.11g***, also uses multicarrier modulation and can be used in

either the 2.4 GHz and 5 GHz bands with speeds of up to 54 Mbps. ***Many wireless LAN cards and access points support all three standards to avoid incompatibilities***." (EX1040, 16.)

229. As discussed above, since Williams discloses that the devices in the home (e.g., interface devices and in-home devices) may communicate with the network 114 through cellular data technologies, or through other types of wired or wireless techniques, and further discloses that the various network communications between components in FIG. 2 may be performed using existing networking protocols and techniques, a POSITA would have looked to Goldsmith which discloses current wireless network communication systems for wireless network communications systems to apply to Williams. (*Supra* § IX.E.2.) As discussed, Williams and Goldsmith are from the same field of endeavor and reasonably pertinent to the same technical problem of providing wireless communications between devices. 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 in the same way. Combining the teachings of Williams and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

230. A POSITA would have looked to Goldsmith's disclosure of WLAN based on IEEE 802.11b, IEEE 802.11a, and IEEE 802.11g, and understood that the

wireless data communications between devices as disclosed in Williams can be WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g, especially since 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 "many wireless LAN cards and access points support all three standards [802.11a, 802.11b, 802.11g] to avoid incompatibilities." *Id*.

231. 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, without requiring much effort and with a reasonable expectation of success. 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. 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.

4. Claim 14: "The method according to claim 1, wherein the wireless network uses a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

232. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 14 is obvious in further view of Goldsmith. Goldsmith states, "[a]ll wireless LAN standards in the U.S. ***operate in unlicensed frequency bands***. ***The primary unlicensed bands are the ISM bands*** at 900 MHz, 2.4 GHz, and 5.8 GHz, and the Unlicensed National Information Infrastructure (U-NII) band at 5 GHz." (EX1035, 15.) "The ZigBee radio specification is designed for lower cost and power consumption than Bluetooth. The specification is based on the IEEE 802.15.4 standard. The ***radio operates in the same ISM band as Bluetooth***, and is capable of connecting 255 devices per network." (EX1035, 20.)

233. Indeed, Goldsmith further summarizes in the table below the unlicensed radio frequency spectrum allocations in the U.S., which expressly shows that both Bluetooth and 802.11b WLANs are ISM Band II frequencies. (EX1035, 22-23.)

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

234. As discussed above, a POSITA would be motivated to combine the teachings of Goldsmith and Williams and apply the well-known wireless communication techniques such as Bluetooth, Zigbee, and WLAN over 802.11b as taught in Goldsmith to Williams. (*Supra* §§ IX.E.2 - IX.E.3.)

5. Claim 15: "The method according to claim 1, wherein the wireless network is 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, or wherein the cellular telephone network is a Fourth Generation (4G) network that uses 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."

235. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 15 is obvious in further view of Goldsmith. For example, 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.

236. Goldsmith discloses that "the ***third generation (3G) cellular systems are based on a wideband CDMA standard*** developed within 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." (EX1035, 13.)

237. 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." (EX1035, 12.)

238. Goldsmith discloses an enhancement to GSM called EDGE. "A more fundamental enhancement, ***Enhanced Data Services for GSM Evolution (EDGE)***, further increases data rates using a high-level modulation format combined with FEC coding. This modulation is more sensitive to fading effects, and EDGE uses adaptive techniques to mitigate this problem. Specifically, EDGE defines six different modulation and coding combinations, each optimized to a different value of received SNR. The received SNR is measured at the receiver and fed back to the transmitter, and the best modulation and coding combination for this SNR value is used. The IS-54 and IS-136 systems currently provide data rates of 40-60 Kbps by aggregating time slots and using high-level modulation. This evolution of the IS-

136 standard is called IS-136HS (high-speed). The IS-95 systems support higher data using a time-division technique called high data rate (HDR)." (EX1035, 13.)

239. 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." (EX1035, 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.

240. 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***." (EX1030, claims 14-16, 18.)

241. 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. As discussed above, Williams and Goldsmith are from the same field of endeavor and reasonably pertinent to the same technical problem of providing wireless communications between devices. Williams discloses that "interface devices might use ***cellular*** or other wireless data ***communications technologies*** to communicate with the application engine 118," 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***." (EX1004, [0020], [0013].) Goldsmith discloses conventional cellular data technologies. (EX1035, 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 in the same way. (*Supra* § IX.E.2.) Combining the teachings of Williams and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

242. 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, without requiring much 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, as explained by Goldsmith. 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. 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.

F. Ground 6: Williams and Chatterjee Renders Obvious Claim 17-22

1. Summary of Chatterjee

243. Chatterjee was filed on August 23, 2011, and is prior art under §102(e). (EX1011, 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." (EX1011, Abstract.)

244. 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." (EX1011, [0001].)

245. 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.)." (EX1011, [0012].)

246. 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***." (*Id*.)

2. Claim 17: "The method according to claim 16, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation."

247. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claims 17-22 are obvious in further view of Chatterjee. 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***." (EX1004, [0051].) Williams notes that "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." (EX1004, [0012].) For example, 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." (*Id*.,[0021]; *see also id.*, [0007]).

248. 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." (EX1011, [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.)." (EX1011, [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*.)

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

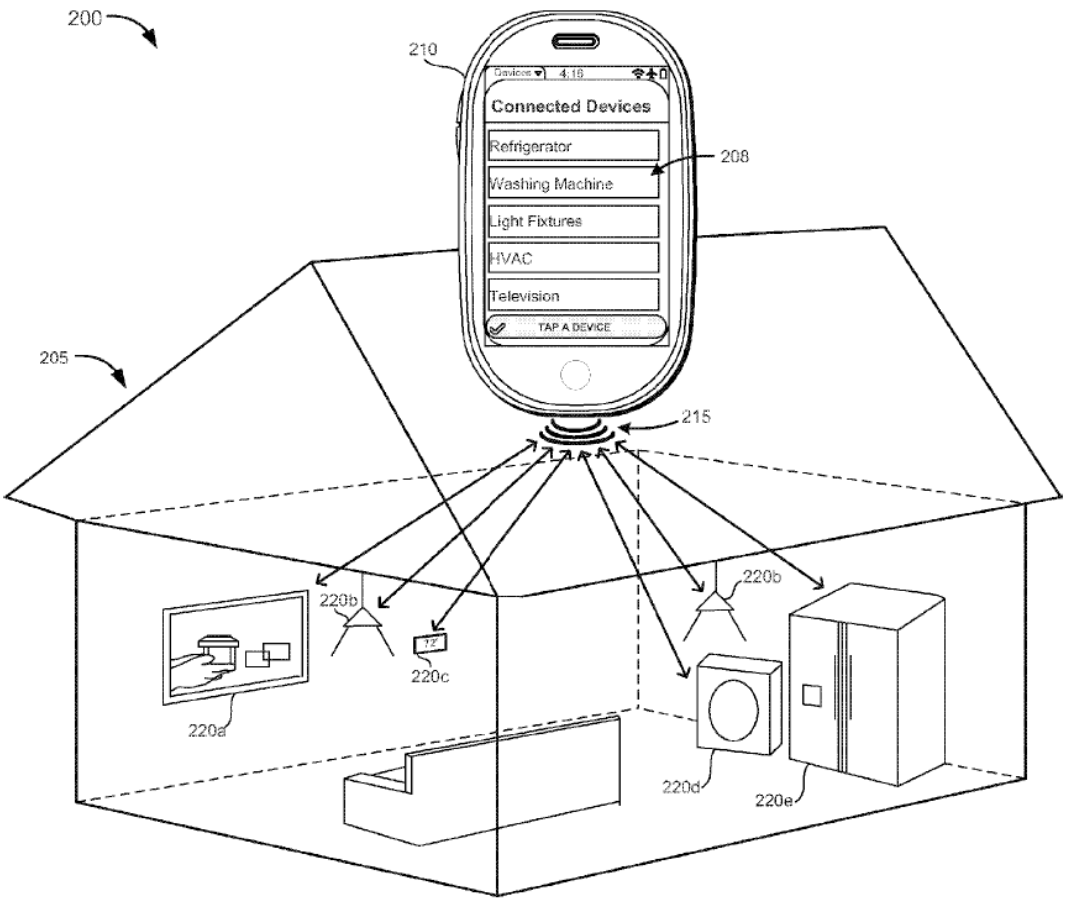


FIG. 2

250. It would have been obvious to a POSITA to combine the teachings of Williams and Chatterjee. Both Williams and Chatterjee are analogous art to the '710 patent. (EX1001, Claim 1, 135:54-60; EX1004, [0007], [0051].) All relate to a home automation system for remotely controlling home appliances. Chatterjee discloses "a system and method for providing home automation using a mobile device." (EX1011, Abstract.) Thus, Williams 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 would improve and expand the capabilities of Williams in the same way. Combining the teachings of Williams and Chatterjee as such would yield predictable results of enabling home automation.

251. Because Williams explicitly provides that its system can be used to control ***home automation devices*** and Chatterjee enumerates exemplary home automation devices, it would have been obvious to try combining the teachings of Williams and Chatterjee such that Williams can be used to control the home automation devices of Chatterjee.

252. Prior art such as Chatterjee has successfully demonstrated that the controlled smart appliances can be an oven, a dishwasher, light fixtures, a refrigerator, and HVAC, and thus, there would have been a reasonable expectation of success for combining Williams 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 such as Chatterjee. 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 to use Williams' techniques to control home appliances such as an oven, a dishwasher, light fixtures, a refrigerator, and HVAC as disclosed by Chatterjee. The known benefits and

feasibility of remote controlling home appliances such as oven, a dishwasher, light fixtures, a refrigerator, and HVAC as disclosed by Chatterjee would prompt the application of Chatterjee in the same field based on design incentives and other market forces.

3. Claim 18: "The method according to claim 17, 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."

253. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 18 is obvious in further view of Chatterjee. 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." (EX1011, [0001].)

254. 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.)." (EX1011, [0012].)

255. It would have been obvious to a POSITA to combine the teachings of Williams and Chatterjee. (*Supra* § IX.F.2.)

4. Claim 19: "The method according to claim 17, 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."

256. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 19 is obvious in further view of Chatterjee. 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." (EX1011, [0001].)

257. 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.)." (EX1011, [0012].)

258. 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***." (EX1011, [0014].)

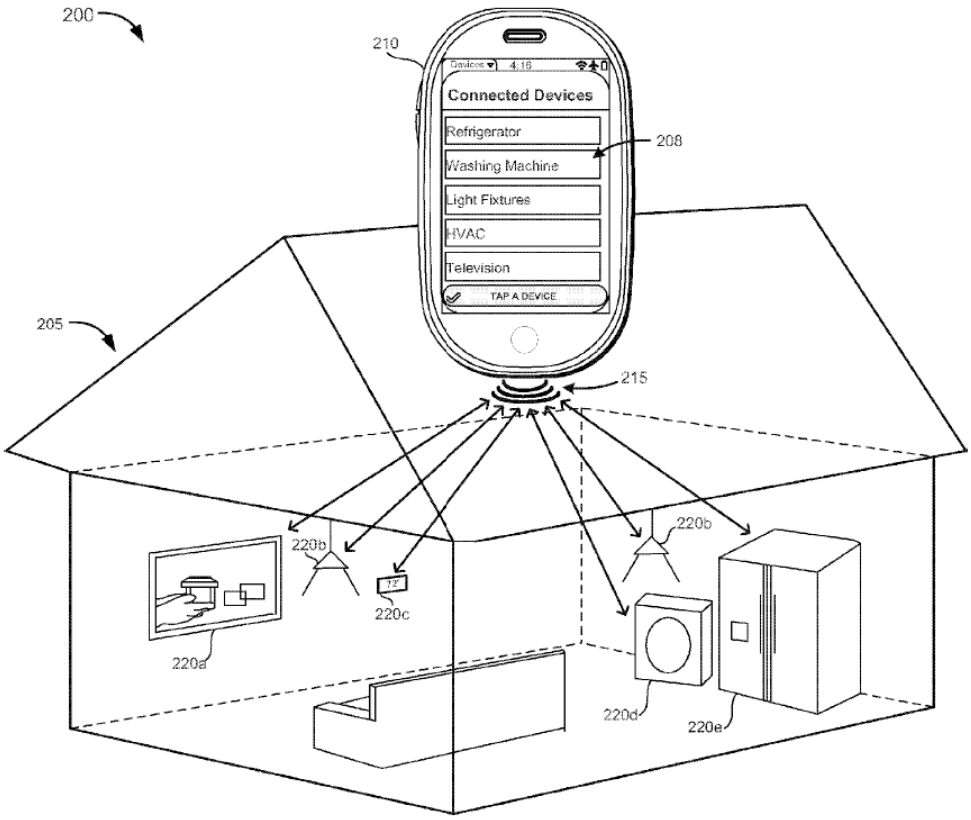


FIG. 2

259. It would have been obvious to a POSITA to combine the teachings of Williams and Chatterjee. (*Supra* § IX.F.2.)

5. Claim 20: "The method according to claim 16, wherein the primary function of the appliance is associated with environmental control, and the appliance is part of an Heating, Ventilation and Air Conditioning (HVAC) system."

260. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 20 is obvious in further view of Chatterjee. 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." (EX1011, [0001].)

261. 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.)." (EX1011, [0012].)

262. Further, Chatterjee discloses, with reference to FIG. 2, that the controlled home appliance includes a HVAC appliance. "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." (EX1011, [0014].)

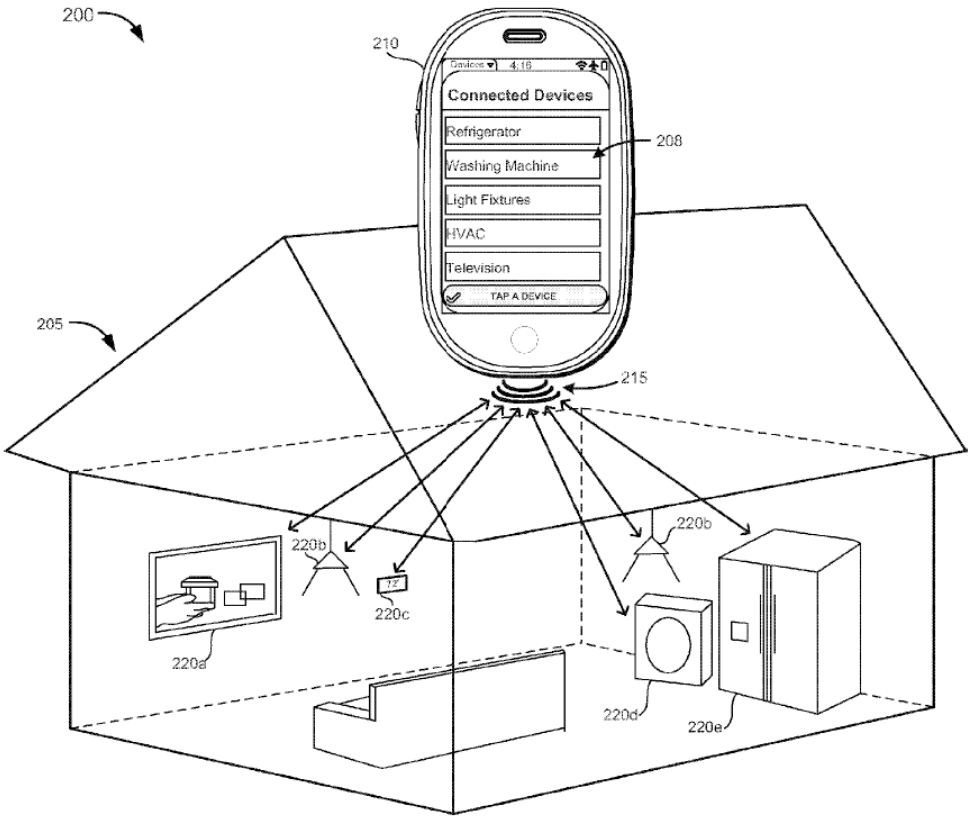


FIG. 2

263. It would have been obvious to a POSITA to combine the teachings of Williams and Chatterjee. (*Supra* § IX.F.2.)

6. Claim 21: "The method according to claim 20, wherein the primary function of the appliance is associated with temperature control, and wherein the appliance is an air conditioner or a heater."

264. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 21 is obvious in further view of Chatterjee. 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." (EX1011, [0001].)

265. Chatterjee specifically discloses that the controlled home appliance includes an HVAC appliance for cooling/heating the home environment like air conditioning. 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.)." (EX1011, [0012].)

266. Further, Chatterjee discloses, with reference to FIG. 2, that the controlled home appliance includes a HVAC appliance with temperature displayed such as an air conditioner. "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." (EX1011, [0014].)

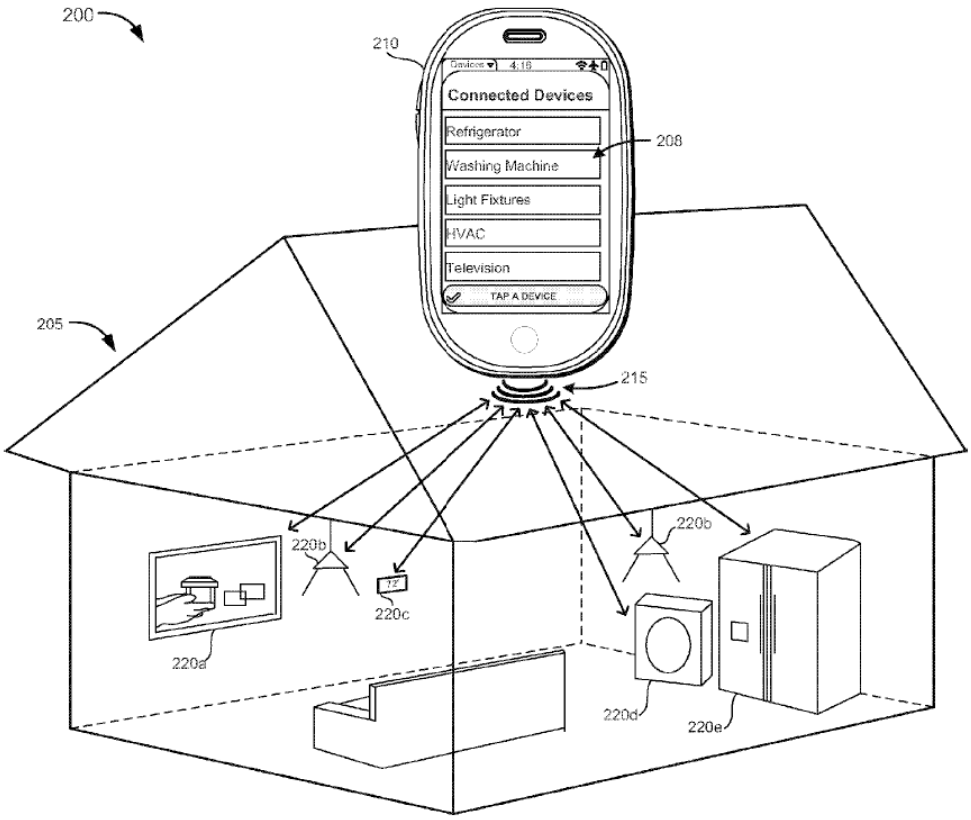


FIG. 2

267. It would have been obvious to a POSITA to combine the teachings of Williams and Chatterjee. (*Supra* § IX.F.2.)

7. Claim 22: "The method according to claim 16, 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."

268. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 22 is obvious in further view of Chatterjee. 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." (EX1011, [0001].)

269. 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." (EX1011, [0014].)

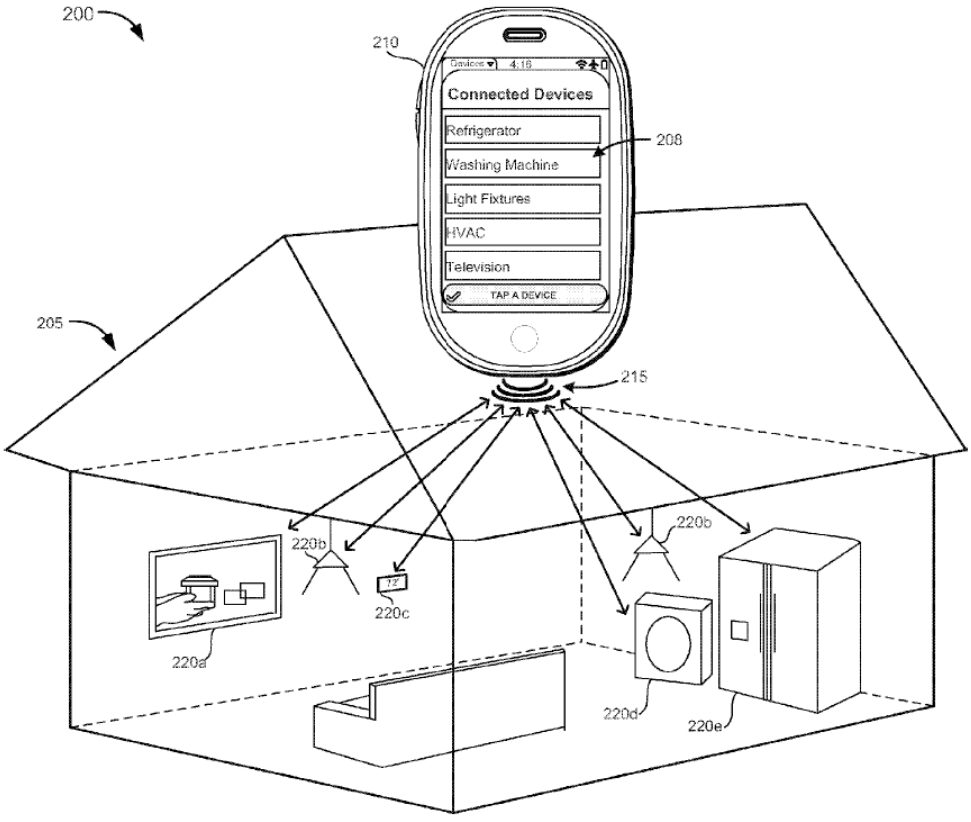


FIG. 2

270. It would have been obvious to a POSITA to combine the teachings of Williams and Chatterjee. (*Supra* § IX.F.2.)

G. Ground 7: Williams and Rosenberger Renders Obvious Claims 24-25

1. Summary of Rosenberger

271. Rosenberger was filed on October 4, 2011, and is prior art under §102(e). (EX1012, 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." (EX1012, 1:7-13.)

272. 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***." (EX1012, 8:8-16.)

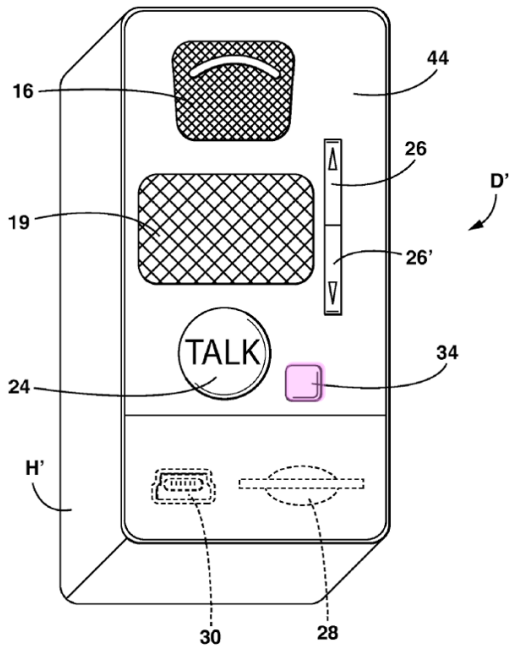


FIG. 2B

273. Rosenberger discloses that the multi-colored LED status light may 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." (EX1012, 8:20-32.)

2. Claim 24: "The method according to claim 1, wherein the first actuator is 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."

274. Claim 1 is obvious in view of Williams. (*Supra IX.A.2*) Claims 24-25 are obvious in further view of Rosenberger. 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***." (EX1012, 8:8-16.)

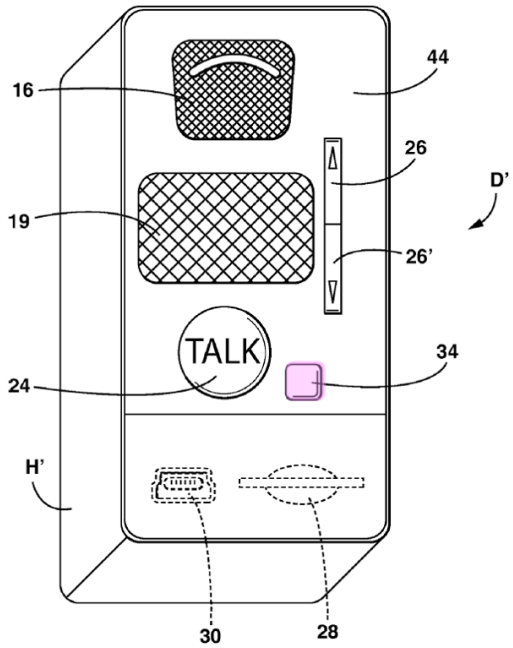


FIG. 2B

275. 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***." (EX1012, 8:20-32.)

276. At the time of the invention, it would have been obvious to combine the teachings of Rosenberger with Williams. Both Rosenberger and Williams are analogous art to the '710 patent. All relate to processing audio information such as human speech for controlling devices or providing services through devices. (*See* EX1004, [0027] and [0036].) 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." (EX1012, 1:7-13.) Thus, Rosenberger and Williams are from the same field of endeavor and reasonably pertinent to the same goal of processing and responding to audio commands.

277. 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. 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 Williams to improve the usability of Williams' interface device, with a reasonable expectation of success.

278. Specifically, the Williams-Rosenberger combination would result 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-Rosenberger combination can 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 Rosenberger.

279. The technique of incorporating an indicator light in the audio interface device disclosed in the Williams-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 cue to prompt users to provide the "first human voice data" and "second human voice data" to control the device. Because both Williams and

Rosenberger relate 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' 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.

280. Further, it would have been obvious to try combining the teachings of Williams 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/modes and prompting users to provide speech input, and thus, there would have been a reasonable expectation of success for combining Williams 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 25: "The method according to claim 24, 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."

281. As discussed, Rosenberger discloses a multi-colored LED status light in the device that visually indicate 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***." (EX1012, 8:8-16.)

282. It would have been obvious to a POSITA to combine the teachings of Williams and Rosenberger. (*Supra* § IX.G.2.)

H. Ground 8: Williams and Rossing Renders Obvious Claims 9, 26-28

1. Summary of Rossing

283. Rossing was published by at least July 23, 2007, and is prior art under §102(a), (b). (EX1036, 1.) The Examiner never found, and Applicant never cited, Rossing during prosecution. Rossing is a textbook titled "Springer Handbook of Acoustics." (EX1013, Title.) Rossing discloses how conventional microphones and speakers operate.

284. 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***." (EX1013, 757.)

285. 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." (EX1013, 758.) "A popular modern variant is the electret condenser microphone, perfected in 1962 by Sessler and West of Bell Laboratories." (EX1013, 759.)

286. Rossing discloses additionally, the piezoelectric microphones and carbon microphones. (EX1013, 758.)

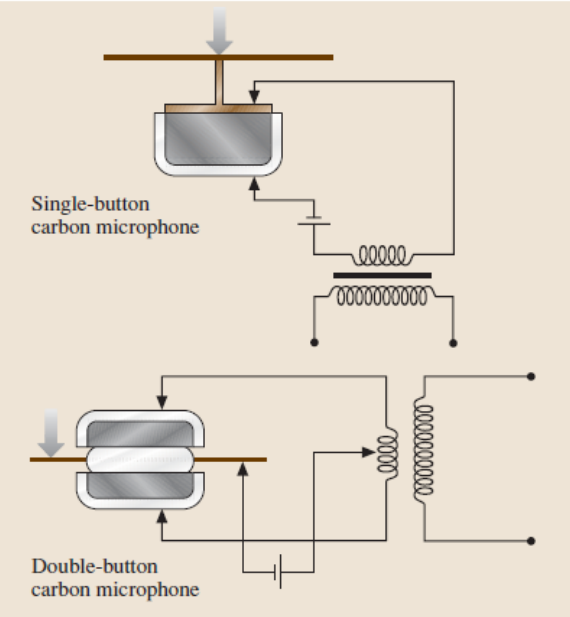


Fig. 18.11 A carbon microphone (After [18.25])

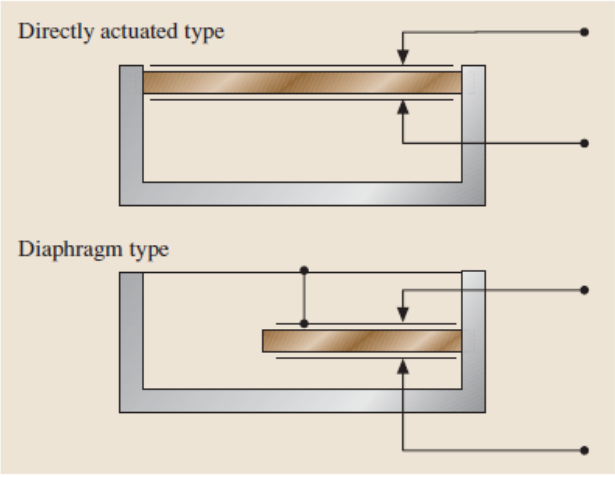


Fig. 18.12 Piezoelectric microphones (After [18.29])

287. Rossing also discloses how a conventional speaker operates. Rossing discloses that "in 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." (EX1013, 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." (EX1013, 903.)

288. 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 "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***." (EX1013, 765.)

289. Further, Rossing discloses that there can be different types of speakers including electromagnetic, piezoelectric, and electrostatic. "Many 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." (EX1013, 763.)

290. 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 Williams.

2. Claim 9: "The method 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."

291. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 9 is obvious in further view of Rossing. A POSITA would understand that the claim refers to a standard microphone. Williams discloses use of such a ***microphone*** in the ***client device***, consistent with the description of the '710 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." (EX1004, [0016]; *see also*, *id.*, [0007] ("network-enabled ***microphones*** can be placed at different activity centers within a home" "to collect audio information" and "microphones may be incorporated in small, self-contained units"); *id.*, [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).

292. 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." (EX1013, 757.)

293. Rossing also discloses other types 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." (EX1013, 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." (*Id.*)

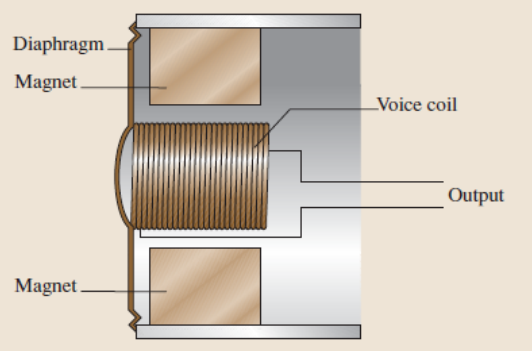


Fig. 18.10 A dynamic microphone uses a coil of wire attached to a diaphragm, suspended in a magnetic field (After [18.24])

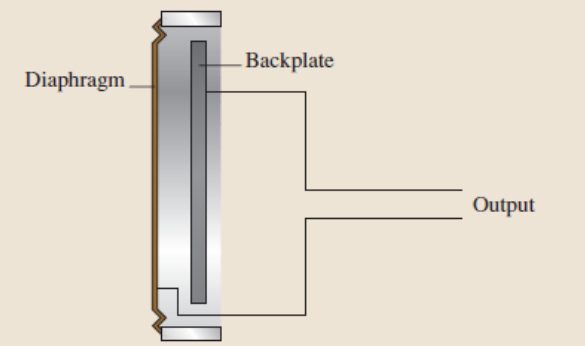


Fig. 18.9 A condenser microphone consists of a very thin diaphragm suspended parallel to a backplate (After [18.24])

294. 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." (EX1013, 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." (*Id.*)

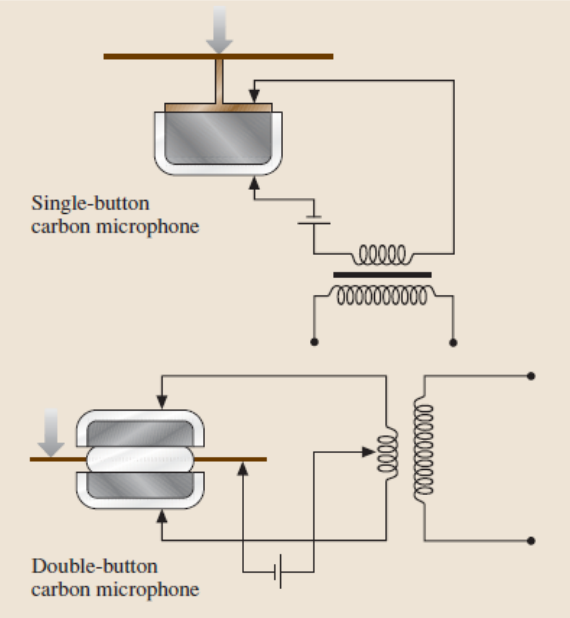


Fig. 18.11 A carbon microphone (After [18.25])

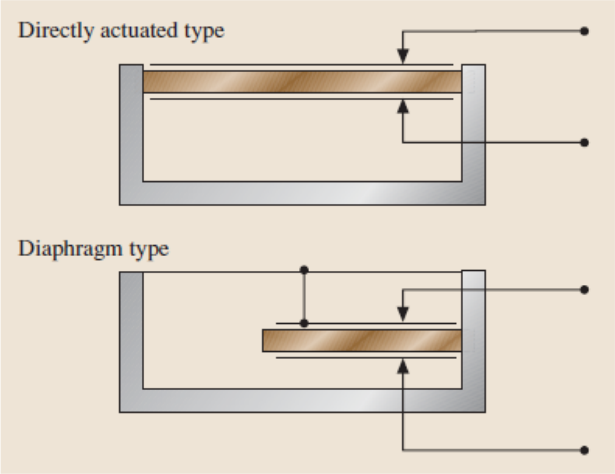


Fig. 18.12 Piezoelectric microphones (After [18.29])

295. 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." (EX1013, 759.)

296. Rossing also discloses that omnidirectional microphones are conventional. Rossing discloses that "for measurement of the lateral energy fraction (LEF), the early part (up to 80ms) 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_l is the impulse response pressure recorded with a figure-of-eight microphone, whereas h is captured through the (usual) omnidirectional microphone." (EX1013, 309.)

297. Both Williams and Rossing are related to operating microphone(s) and speaker(s). (EX1004, [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. (EX1013, 757-758, 763, 765, 903.) Thus, Williams and Rossing are from the same field of endeavor and reasonably pertinent 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.

298. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Williams, 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, 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 requiring much 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 such as Rossing. 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.

3. Claim 26: "The method according to claim 1, wherein the first actuator is a motion actuator that causes linear or rotary motion."

299. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 26 is obvious in further view of Rossing. As discussed, Williams teaches operating an

actuator (e.g., ***speaker 206***) in the client device (e.g., ***interface device 116***) in response to the received first message (e.g., ***audio 222***). (*Supra* § IX.A.2.k.) For example, Williams discloses several exemplary use cases of operating the speaker "[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***." (EX1004, [0048].)

300. Rossing discloses that a speaker is or comprises a motion actuator that causes linear motion. First, Rossing explains that a speaker is a motion actuator. Rossing explains that a speaker converts electrical energy to mechanical energy. "In 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." (EX1013, 763.) As loudspeakers convert electrical energy to mechanical energy resulting in movement of a diaphragm, it is clear that loudspeakers are motion actuators.

301. Second, Rossing discloses that a speaker is or contains a motion actuator that causes linear motion. Rossing discloses that "in the low-frequency range, the ***linear motion of an electrodynamical loudspeaker***, for example, can be described by a single-degree-of-freedom oscillator." (EX1013, 903.) Further, Rossing describes that the dynamic or moving coil loudspeaker move axially. "Ernst W. Siemens was the first to describe the ***dynamic or moving-coil loudspeake*r**, with a circular coil of wire in a magnetic field and supported so that ***it could move axially***." (EX1013, 18.) Thus, Rossing discloses that an electrodynamical loudspeaker itself is a motion actuator that causes linear motion.

302. As discussed, Williams and Rossing are from the same field of endeavor and reasonably pertinent 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 related to speakers for implementing Williams. Further, it would have been obvious to try combining Rossing with the teachings of Williams. (*Supra* § IX.H.2.)

4. Claim 27: "The method according to claim 1, wherein the first actuator is a sounder for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves."

303. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 27 is obvious in further view of Rossing. A POSITA would have understood that this claim refers to a speaker. As discussed above, William teaches operating a

speaker 206 consistent with descriptions of the '710 patent. (*Supra* § IX.A.2.k; EX1001, 105:48-52, 105:61-106:4; EX1004, [0037].) As discussed above, at the time of the invention, it was common knowledge that a speaker is a sounder for converting an electrical energy to sound waves. For example, the actuator 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." (EX1017, 122.)

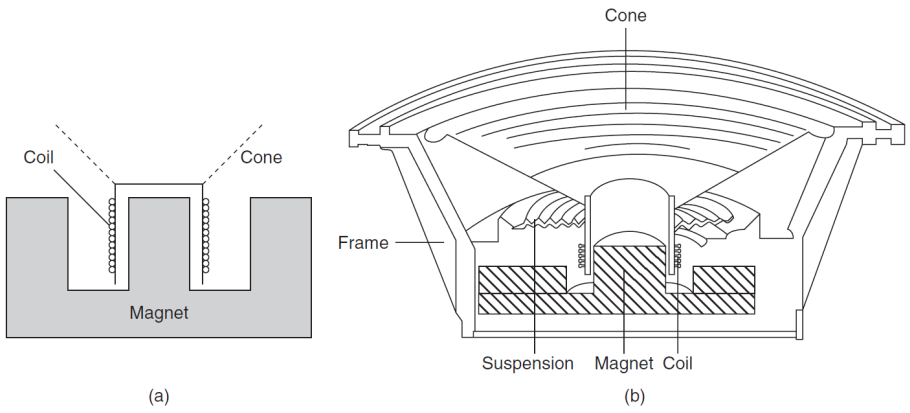


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

304. "[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." (EX1013, 763.)

305. 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***." (EX1013, 765.) It would have been obvious to combine Williams and Rossing. (*Supra* § IX.H.2.)

5. Claim 28: "The method according to claim 27, 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, or wherein the sounder comprises an electric bell, a buzzer, a chime, a whistle, or a ringer."

306. Claim 1 is obvious in view of Williams. (*Supra* § IX.A.2) Claim 28 is obvious in further view of Rossing. A POSITA would have understood that this claim refers to a speaker. As discussed, Williams teaches operating an actuator (e.g., ***speaker 206***) consistent with descriptions of the '710 patent in the client device (e.g., ***interface device 116***) in response to the received first message (e.g.,

audio 222). (*Supra* § IX.A.2.k; EX1001, 105:48-52, 105:61-106:4; EX1004, [0037].)

307. 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***." (EX1013, 763.) "In 1929, E. W. Kellogg filed a patent for an ***electrostatic loudspeaker***, consisting of panels not unlike large condenser microphones." (*Id*.) Rossing also discloses using "***small loudspeaker units*** (***piezoelectric or electro magnetic***) and the microphones 1/4" or even 1/8" condenser types." (EX1013, 321-322.)

308. As discussed, Williams and Rossing are from the same field of endeavor and reasonably pertinent 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 related to speakers for implementing Williams. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional speakers to Williams' speaker, as various speaker features can improve William'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.

309. Further, it would have been obvious to try combining the technical features of Rossing with the teachings of Williams, 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 requiring much effort and with a reasonable expectation of success. (*Supra* § IX.H.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 Williams system. (EX1013, 321-322, 763.)

I. Ground 9: Howard and Doherty Renders Obvious Claims 1, 5-6, 10-11, 13, 16-21, 23, 29.

1. Summary of Howard

310. Howard was filed on May 9, 2002, and published on June 13, 2006, before the earliest provisional application filing date of the ’710 patent, and thus qualifies as prior art under §102(a), (b), and (e). (EX1005, 1.) The Examiner never found, and Applicant never cited, Howard during prosecution. Howard relates generally to “a method and system that allows an operator to easily and readily control home activities without having to directly interface with or use a

PC." (EX1005, 1:53-56.) Howard discloses "a portable device readily carried by the operator that can be utilized to communicate with a server located at a fixed location. Utilizing the portable device, the operator can program and control various components and activities." (EX1005, 1:58-64.) FIG. 1(a) and 1(b) depict exemplary portable devices. (EX1005, 2:38-43.) Each portable device has a microphone 3 and a speaker 4. (EX1005, 3:24-27, 3:54-62.)

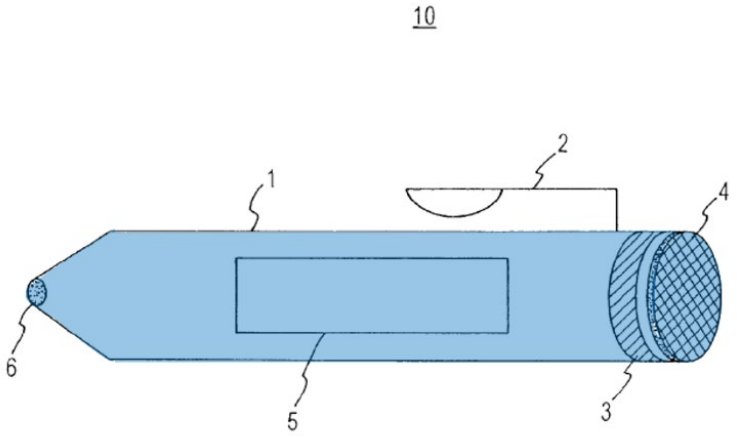


FIG. 1(a)

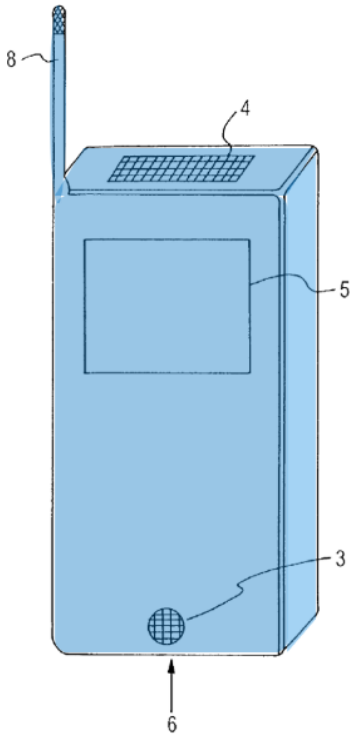


FIG. 1(b)

311. FIG. 9 illustrates an exemplary workflow. As shown, the user can speak audio commands such as "TV", "[c]hannel", "[v]olume", "[a] little small" to the ***portable device***. These audio commands are transmitted to and processed by a ***server***. (EX1005, 11:20-38.) As shown below, when the user utters "TV", the ***server*** can process the audio command and generate a message (***"What Can I do for you?"***), which is transmitted to and played back (using the speaker 4) at the ***portable device***. (*Id.*, 11:27-47.) If the user utters "[a] little small", the ***server*** can also process and audio command and generate ***control signals***, which are transmitted to a ***VTR device*** to change the volume. (EX1005, 11:20-12:14.)

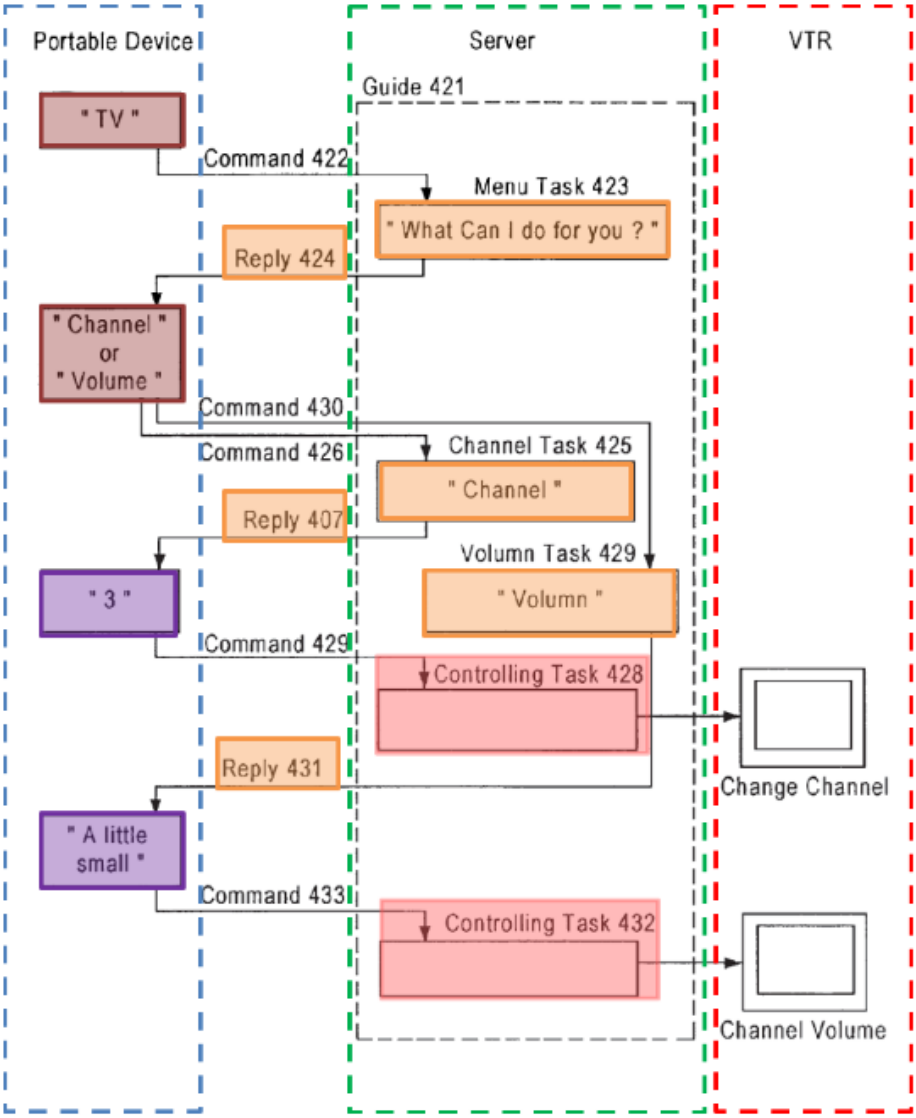


FIG. 9

312. Howard discloses a variety of electrical appliances that can be remotely controlled using the portable device. For example, Howard discloses "controlling apparatus connected on the home network 304. Such apparatus include, but are not limited to, (***a) audio/video equipment, for instance TVs, VCRs, DVDs, CD players, (b) electrical appliances, such as light systems, refrigerators, microwave ovens, heating and cooling systems, (c) computers, (d)***

home automation systems, (e) fire alarm systems, (f) burglar alarm systems, (g) telephone systems, and (h) intercom systems."

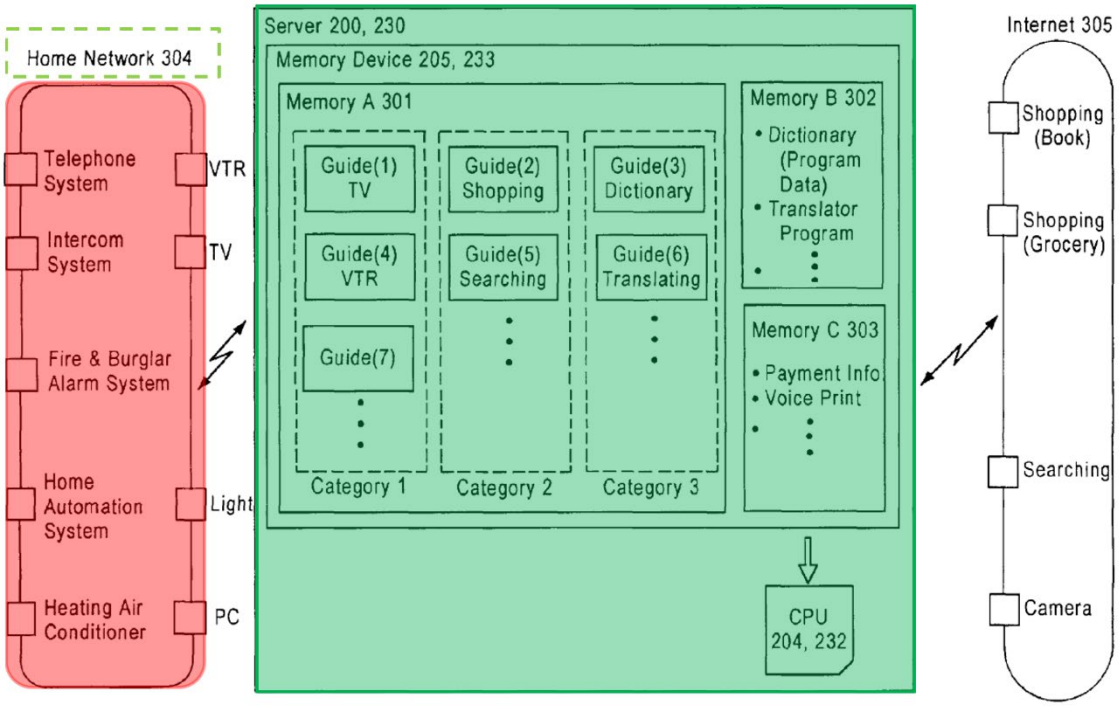


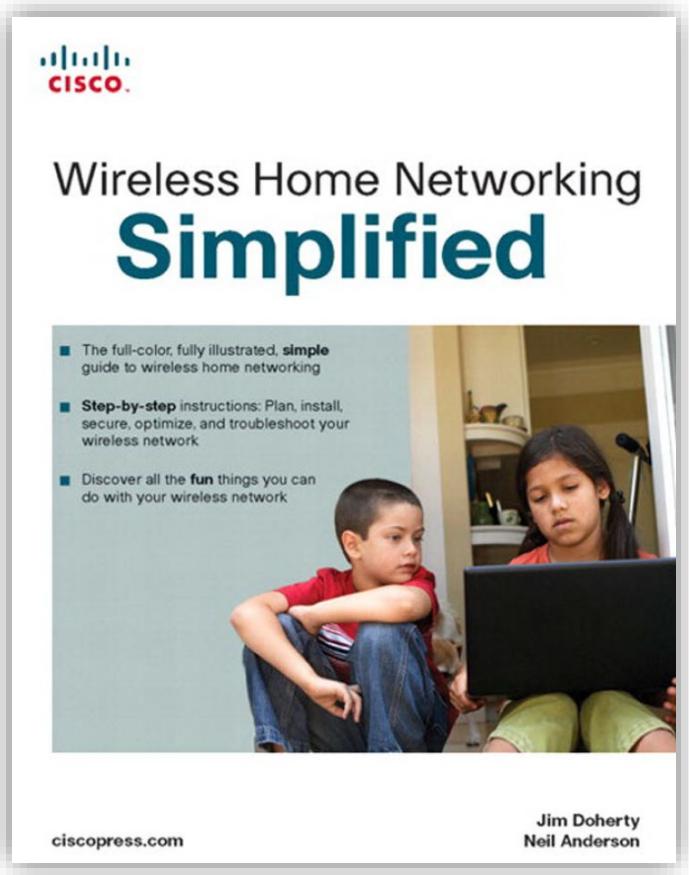
FIG. 6

313. Howard discloses that the portable device and the electrical appliances each communicates over a wireless network, and that the server is connected to the Internet. For example, Howard discloses that "the data and the control signals utilized to control the components coupled to the ***home network*** can be packetized and transmitted utilizing known schemes, such as ***BlueTooth***, which is the IEEE 802.15.1 standard for ***wireless personal area networks***." (EX1005, 5:22-26.) In addition, Howard discloses that "home appliances are being connected together and controlled via a ***home network***, which typically take different forms ranging

from a simple ***network hub*** device to a more complex hub and switching system with ***broadband connection to the Internet***." (EX1005, 1:18-22.) Further, "[b]oth the home network 304 and the ***Internet*** websites 305 are accessible by the server 200, 230" and "the server 200 is also coupled to the ***Internet*** 305 via any suitable connection." (EX1005, 8:40-42, 9:50-52.)

2. Summary of Doherty

314. Doherty was published by at least January 16, 2007, and is prior art under §102(a) and (b). (EX1033, 1.) The Examiner never found, and Applicant never cited, Doherty during prosecution. Doherty is a book titled "Wireless Home Networking Simplified". (EX1006, Title.) Doherty teaches readers where to buy gear and how to set up wireless networks at home using conventional products. "***With the explosion of popularity in wireless networking***, there is a proportional number of people who do not understand the technology, and a similar number of people who think they understand it (but don't) and are all too willing to dispense advice . . . This book is written for all of you out there who want to get the benefits of wireless networking but don't feel like you have the technical background to set it up yourselves. We are confident that if you read this book and follow the steps we lay out, you will have a much better understanding of the technology and, most importantly, a working wireless network." (EX1006, Introduction, Chapter 6.)



315. Doherty is a textbook published by Cisco teaching readers how to set up wireless networks. A POSITA would have looked to Doherty for implementing wireless networking at home.

316. Doherty discloses that wireless networking has been around for well before 2007. "Now, some folks may be a little dubious about using a 'new' technology, but the ability to provide wireless networking actually has been around for a long time." (EX1006, 5.)

317. Doherty discloses that "in a wired network, data is sent and received between computers and routers using pieces of cable called Ethernet. In a wireless

network, air replaces cables, and the same computer-to-computer communications are transmitted and received over the air using radio waves," and "[i]f it gets too technical, just remember one thing: wireless networking is just a way to send stuff like web pages to and from your computer without a cable." (EX1006, 3.)

318. Doherty discloses the use of a ***router*** or a switch. "Eventually, the topology that was (primarily) settled on was the bus topology. A bus topologoy has all the computers connected to a common cable, which is referred to as the bus. ***The common 'cable' connecting all the computers today is located inside the router or switch***, but it works the same way. Other topologies are still in use but ***most of the time LANs are connected using a bus topology***." (EX1006, 6; *see also id*., 62.)

Figure 6-4 Linksys Wireless-G Broadband Router (WRT54G)



These materials have been reproduced by Cisco Press with the permission of Cisco Systems, Inc.
COPYRIGHT ©2007 CISCO SYSTEMS, INC. ALL RIGHTS RESERVED

Figure 6-5 Linksys Compact Wireless-G Broadband Router (WRT54GC)



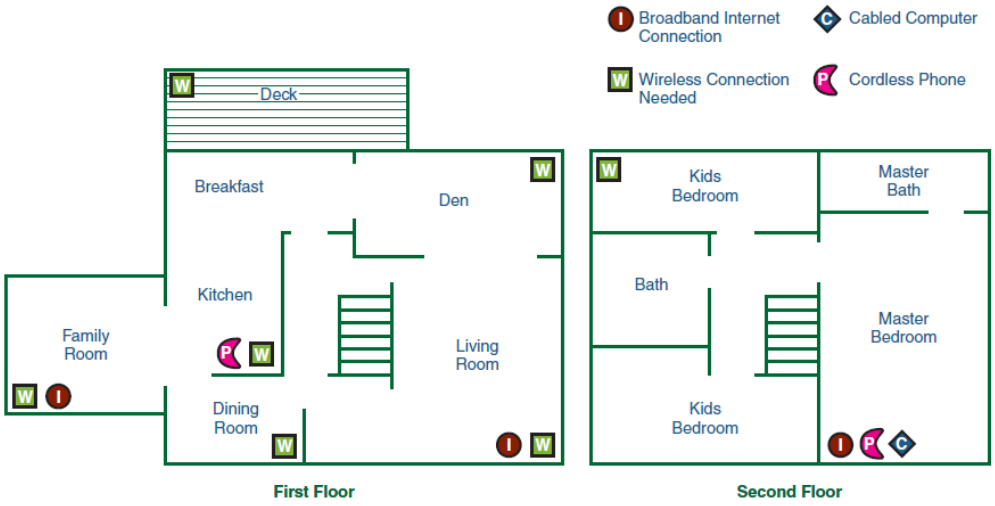
These materials have been reproduced by Cisco Press with the permission of Cisco Systems, Inc.
COPYRIGHT ©2007 CISCO SYSTEMS, INC. ALL RIGHTS RESERVED

319. "For all its magic, the technology that drives wireless communication is no more than a fancy walkie-talkie in a really small package. The wireless device—whether it is a card you plug in or ***electronics that are built right into the computer—is a transmitter and receiver, sending and receiving radio signals*** . . . The really good news is that the range of these wireless devices is good enough that you can put the ***wireless router in your home*** and provide access throughout most of your home and probably into areas of your yard. This is all done with minimal power consumption so that the battery on your wireless laptop does not drain too fast." (EX1006, 10.)

320. Doherty explains the concept of wireless LAN (802.11) protocol, also known as Wi-Fi, and discloses 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." (EX1006, 11.)

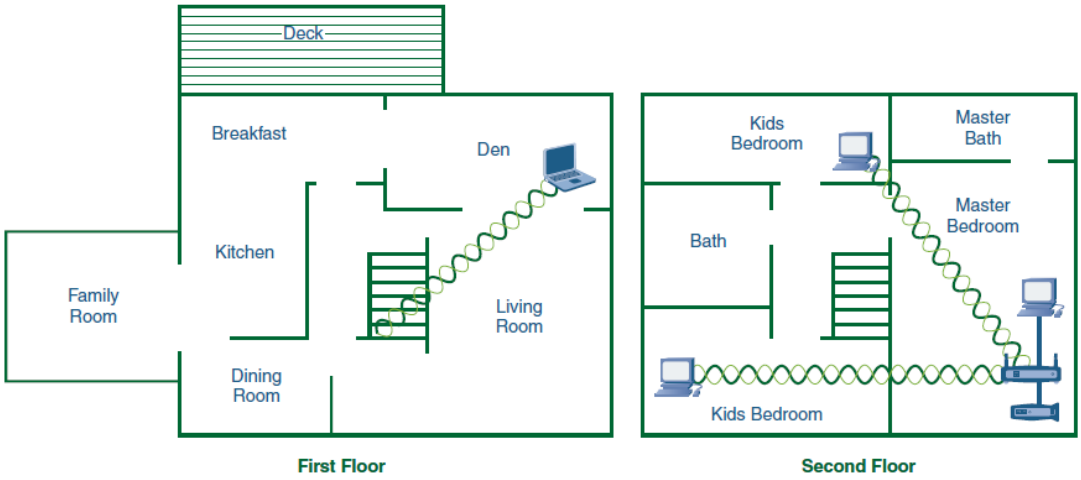
321. Doherty discloses, with reference to Figures 4-3 and 4-4, that the wireless router allows connection from multiple devices. "Finally, put a mark in the rooms or areas where you can conceive of using a ***laptop computer and other wireless devices, including video cameras, printers, video game consoles***, and so on. (This would include bathrooms, which we admit to using our computers in.) In Figure 4-3, we used a green square with a W in it to denote where we want to have a wireless network connection, like a laptop." (EX1006, 38.)

Figure 4-3 Sample Sketch of a House



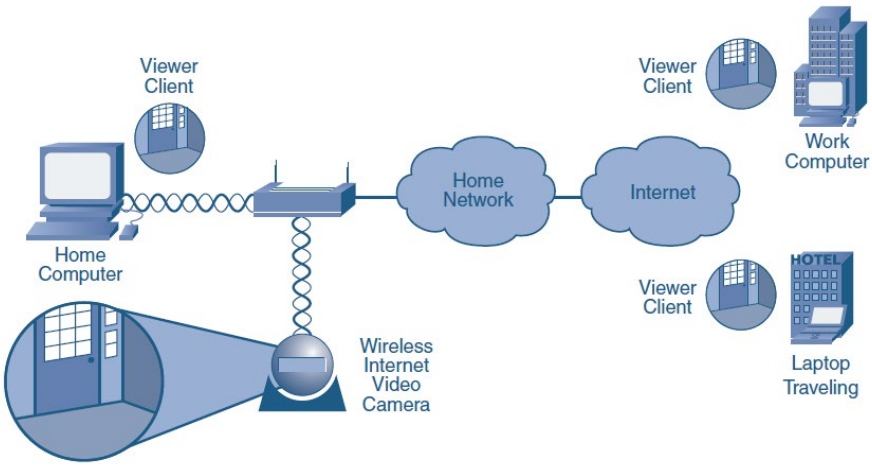
322. "Because the daughter's laptop and the son's desktop computer are in other bedrooms of the house, and not easy to run Ethernet cables to, ***these computers will have wireless connections***. The last computer is the work laptop, which is used primarily in the den, but could also be used in the living room, breakfast nook, and on the back deck. ***Wireless is definitely the way to go for this laptop as well***. By jotting these requirements into the sketch, we have what's shown in Figure 4-4." (EX1006, 39.)

Figure 4-4 Adding Wired and Wireless Computers to the Sketch



323. In addition to laptops and desktop computers, Doherty discloses that wireless video and entertainment devices such as wireless video cameras, and digital video recorder may be connected to the wireless network through the wireless router. "Also popular recently have been very inexpensive wireless cameras that can be placed anywhere in your house . . . No computer is required for them to operate, so they can be located anywhere in your house (or outside, for that matter) where there is an electrical outlet. Images from the camera are viewable with any computer from inside your house, or from anywhere outside your house where there is Internet service. In Figure 13-1, we placed a wireless Internet video camera in a house with a view of the front door." (EX1006, 261.)

Figure 13-1 Home Surveillance Using a Wireless Internet Video Camera



324. Doherty, published in 2007, also discloses more and more devices are connecting to the wireless network. "In 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." (EX1006, 321.)

3. Claim 1

a. 1[PR][1]: "A method for operating multiple actuators in response to captured human voice data,"

325. Howard discloses a method for operating multiple actuators in response to captured ***human voice data***. Howard discloses a "portable device includes: . . . a microphone for receiving a first ***audio data***[.]" (EX1005, Abstract, Claim 1.) Howard states that "audio data . . . is usually ***human voice or speech***." (EX1005, 4:10-13.)

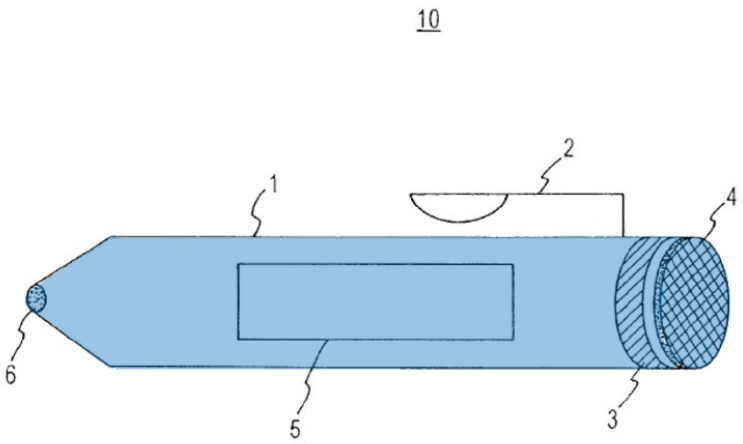


FIG. 1(a)

326. For example, with reference to FIG. 9, Howard discloses multiple human voice data. "In operation, when the ***word 'TV'*** is input into microphone 3 in the portable device 10, in the same manner as set forth above, the data is encoded by audio coder 111 and transmitted to the server 200 via the transmitter 113." (EX1005, 11:20-23.) "Upon being queried as to the desired channel, assuming the ***operator states the desired channel***, this audio data is encoded and sent to the server, where it is determined to correspond to command 429." (EX1005, 11:57-61; *see also* FIGs. 8 and 10-13(d).)

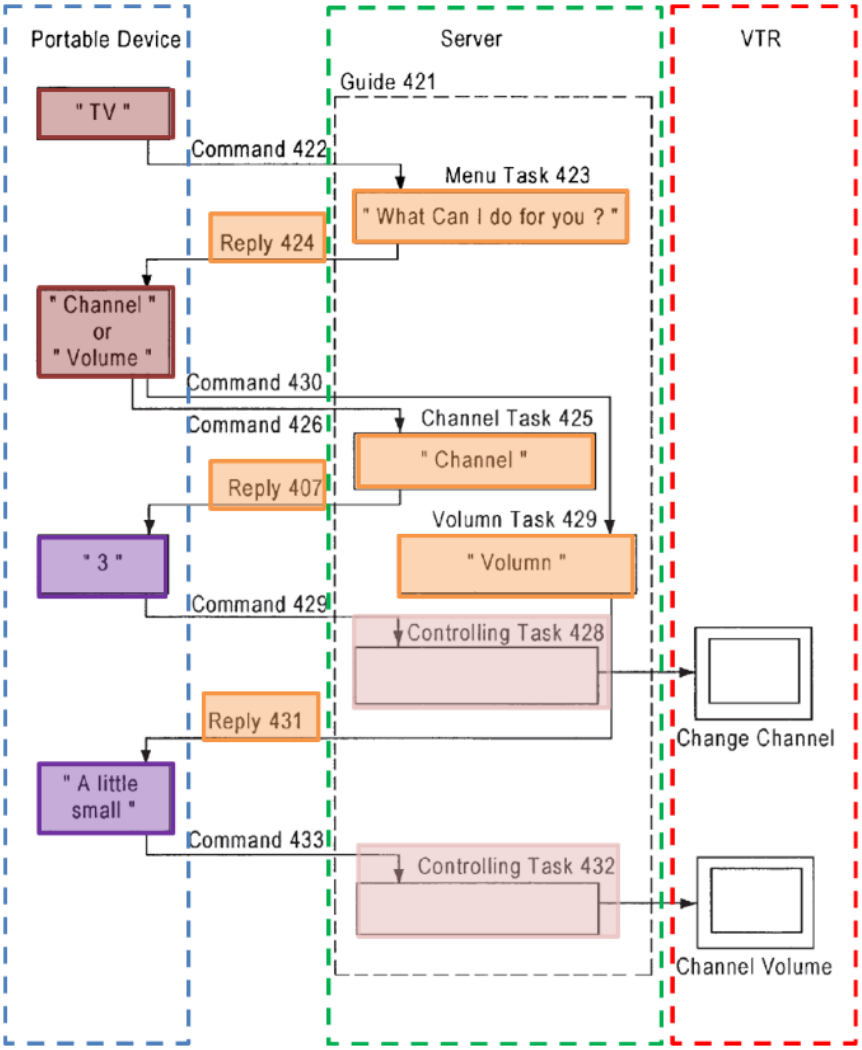


FIG. 9

327. As another example, Howard discloses that "the portable device and control system allows for many such types of operations by simply placing one command. … Similarly, by simply ***stating a command*** into the microphone, the operator can control any component coupled to the home network (e.g., tune a TV to channel 8), or access any website via the Internet connection." (EX1005, 16:13-25.)

328. As described herein, Howard operates multiple actuators, including actuators in its portable device, "audio/video equipment", "electrical appliances", "home automation systems", etc. (EX1005, 9:14-23, FIGS. 8-9; *infra* IX.A.2.k and IX.I.3.r.)

b. 1[PR][2]: "for use with a client device and a controlled device in a building,"

329. Howard discloses a ***client device*** and a ***controlled device*** in a ***building***.

330. First, Howard discloses a ***client device***. For example, Howard discloses an "control system includes a ***portable device*** and a server." (EX1005, 1:65-66.) Howard discloses that "the control system of the present invention includes a ***portable device*** readily carried by the operator that can be utilized to communicate with a server at a fixed location. Utilizing the ***portable device***, the operator can program and control various components and activities." (EX1005, 1:59-64.) FIGs. 1(a) and 1(b) illustrates "exemplary configuration[s] of a portable device." (EX1005, 2:38-43.)

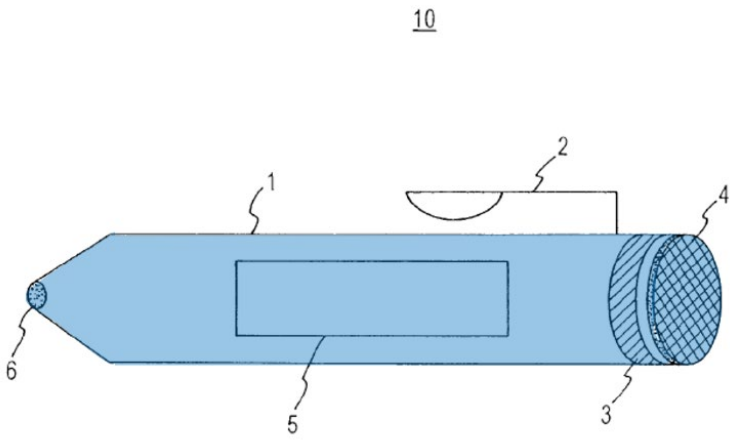


FIG. 1(a)

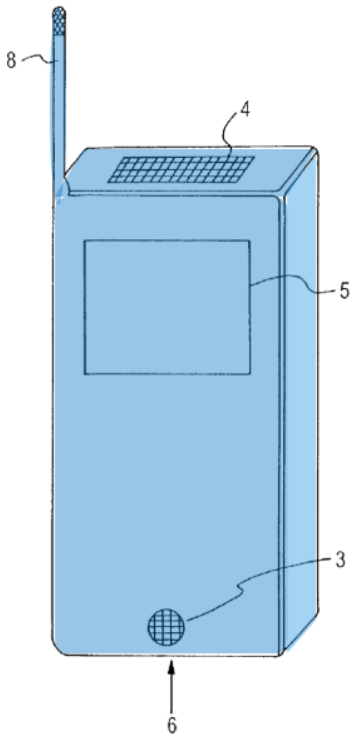


FIG. 1(b)

331. With reference to FIG. 2, Howard discloses that "the ***portable device 10*** comprises both a transmitter system 101 and a receiver system 102." (EX1005, 4:2-4.) Howard discloses that the ***portable device 10*** communicates wirelessly with ***server*** 200. (*Id.*, 4:2-24.) Thus, Howard's portable device is a client device.

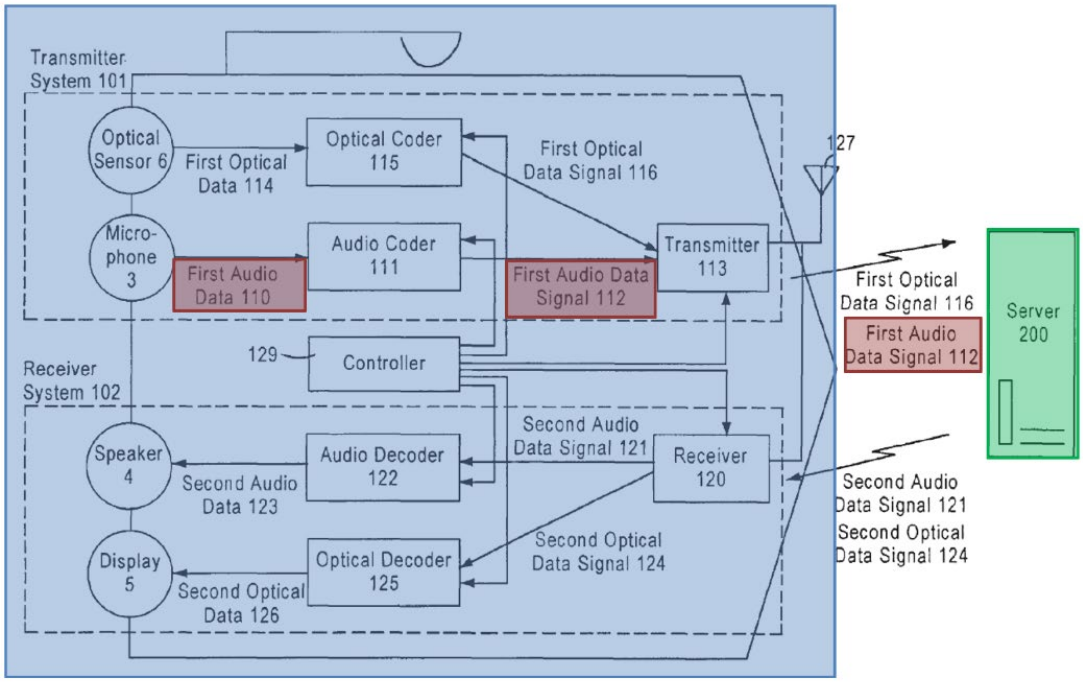


FIG. 2

332. Second, Howard discloses a ***controlled device***. Howard discloses that "by simply stating a command into the microphone, the operator can ***control*** any ***component coupled to the home network*** (e.g., tune a TV to channel 8), or access any website via the Internet connection." (EX1005, 16:21-25.) Howard further explained that by "utilizing the home gateway, it is possible to control a wide variety of activities/components within a home. For example, these activities/components include various ***entertainment network devices, such as TV,***

stereo, and video equipment. In addition, homes usually have various ***'life support systems', such as, heating systems, air conditioning systems and lighting and security systems***. These life support systems can also be connected to the home gateway and controlled via the home gateway." (EX1005, 1:18-30; *see also*, *id.*, 9:46-50 ("allow for control of the various ***household components and appliances***.").) Thus, these devices can be controlled and are thus controlled devices.

333. With reference to FIG. 6, Howard discloses "controlling apparatus connected on the home network 304. Such apparatus include, but are not limited to, ***(a) audio/video equipment, for instance TVs, VCRs, DVDs, CD players, (b) electrical appliances, such as light systems, refrigerators, microwave ovens, heating and cooling systems, (c) computers, (d) home automation systems, (e) fire alarm systems, (f) burglar alarm systems, (g) telephone systems, and (h) intercom systems***." (EX1005, 9:13-23.) Howard discloses that "utilizing the home gateway, it is possible to control a wide variety of activities/components within a home." (EX1005, 1:22-24.)

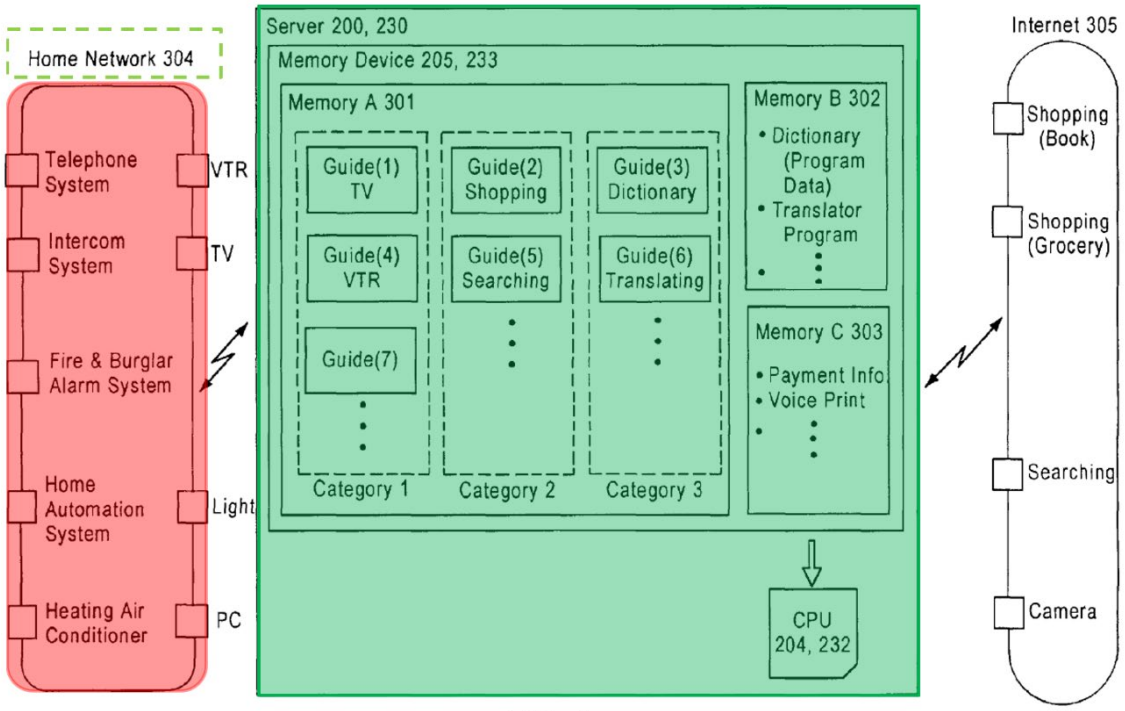


FIG. 6

334. Howard discloses that the ***client device*** and the ***controlled device*** are in a ***building***. Howard is clear that the ***controlled devices*** are within a ***home***. Howard "relates generally to a ***portable device*** for communicating with a computer server, especially a home sever [sic], which provides for and/or controls various activities ***in a home***, including for example, ***home appliances***, purchasing of goods, etc. The present invention also relates to a method and system, which includes the ***portable device*** and server, for controlling such ***home activities***." (EX1005, 1:7-14.) Further exemplary disclosures of Howard include: "***home*** appliances," "control of the various ***household*** components and appliances," "control various activities within a ***home***," "controlling various ***home*** related

activities." (EX1005, 1:7-11, 1:22-24, 1:33-36, 9:49-50.) A POSITA would have understood that a home is or is part of a building.

335. Howard further discloses that the ***client device*** (i.e., Howard's "portable device") is in the ***building***. Howard states that the "***portable device*** [can be] readily carried by the operator" and that the system "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." (EX1005, 1:53-58; *see also*, *id.*, 2:19-26.)

c. 1[PR][3]: "each communicating over a wireless network, and"

336. Howard discloses that each of the ***client device*** and the ***controlled device*** communicates over a wireless network.

337. First, Howard discloses that the ***client device*** communicates over a wireless network. Howard discloses "a ***portable device*** comprising . . . a ***transmitter*** for transmitting at least the audio command or the optical command to the ***server***." (EX1005, claims 1 and 9; *see also*, *id.*, 1:65-2:7 and 4:57-63). With reference to FIG. 1, Howard discloses that "the portable device 10 also comprises a controller (not shown in FIG. 1(a)) which functions to control the operation of the foregoing components, and an ***antenna*** (also not shown in FIG. 1(a)) which functions to allow for ***wireless communication*** between the ***portable device 10*** and

a ***server 200*** forming part of the home gateway. In one embodiment, the ***antenna*** is incorporated into the design of the clip 2." (EX1005, 3:41-48.)

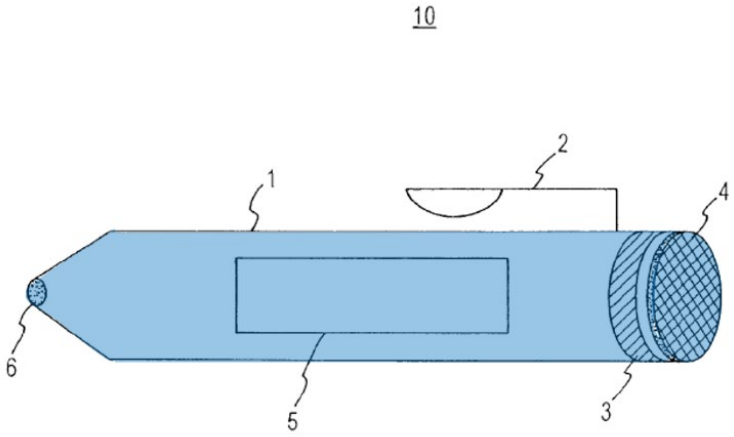


FIG. 1(a)

338. With reference to FIG. 2, Howard discloses that "the ***portable device 10*** comprises both a transmitter system 101 and a receiver system 102" and that "the transmitter system 101 includes the microphone 3, which functions to receive audio data 110, which is usually human voice or speech. The microphone 3 is coupled to an audio coder 111 (e.g., CODEC) and outputs the audio data 110 to the audio coder 111. The audio coder 111 functions to packet/code or convert the audio data 110 to modulated audio data signals 112, which are suitable for ***wireless transmission*** to the ***server 200***." (EX1005, 4:2-4, 4:10-18.) Further, "[t]he audio coder 111 is coupled to a ***transmitter 113***, which functions to receive modulated audio data signals 112 from the audio coder 111 and to transmit the modulated

audio data signals to the ***server 200*** via an ***antenna 127***." (EX1005, 4:20-24; *see also*, *id.*, 6:25-30). Thus, Howard shows that the ***portable device*** communicates over wireless network, for example, via a transmitter and an antenna.

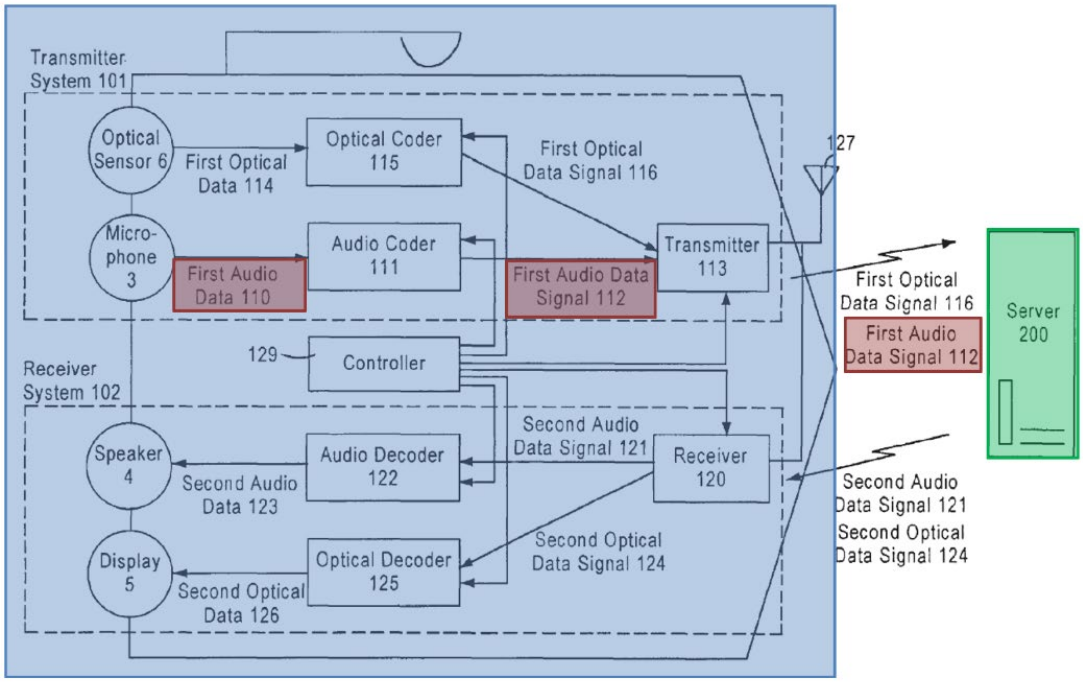


FIG. 2

339. Further, with reference to FIG. 4, "the communication device 201 (e.g. data transceiver) is capable of performing ***wireless and/or optical transmission of data/commands*** with the ***portable device 10***." (EX1005, 6:32-36.)

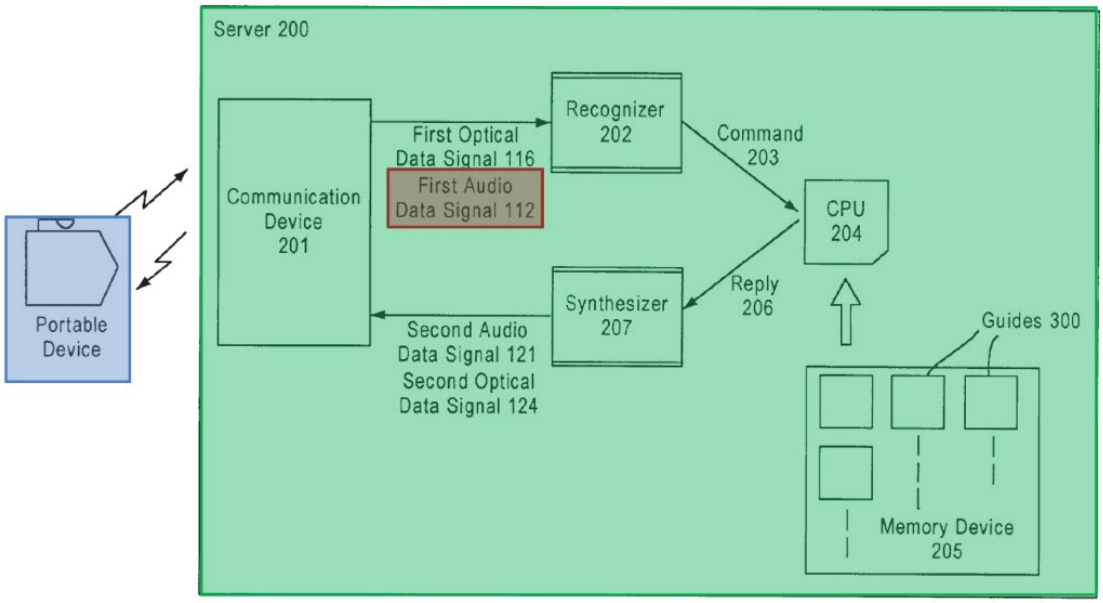


FIG. 4

340. Second, Howard discloses that the ***controlled device*** communicates over a wireless network. Howard discloses that "the data and the control signals utilized to control the ***components*** coupled to the ***home network*** can be packetized and transmitted utilizing known schemes, such as ***BlueTooth***, which is the IEEE 802.15.1 standard for ***wireless personal area networks***." (EX1005, 5:22-26.) In addition, Howard discloses that "***home appliances*** are being connected together and controlled via a ***home network***, which typically take different forms ranging from a ***simple network hub device*** to a more complex hub and switching system with ***broadband connection to the Internet***." (EX1005, 1:18-22; *see also id.*, FIGS. 8-13(d); 10:64-11:5, 11:48-65; 11:66-12:14.)

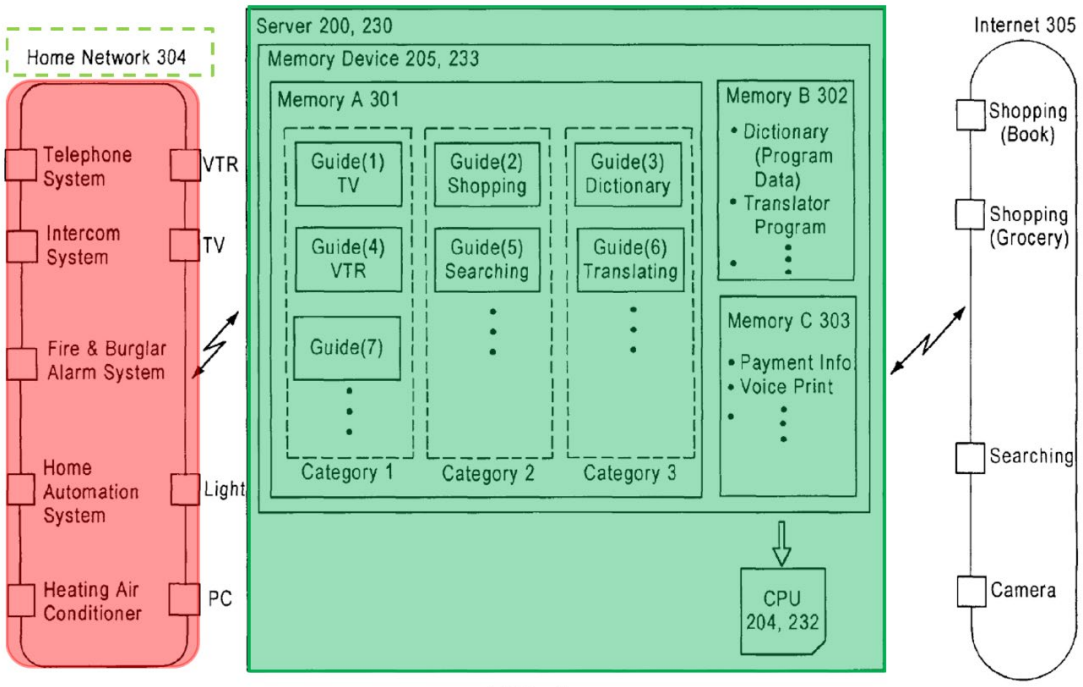
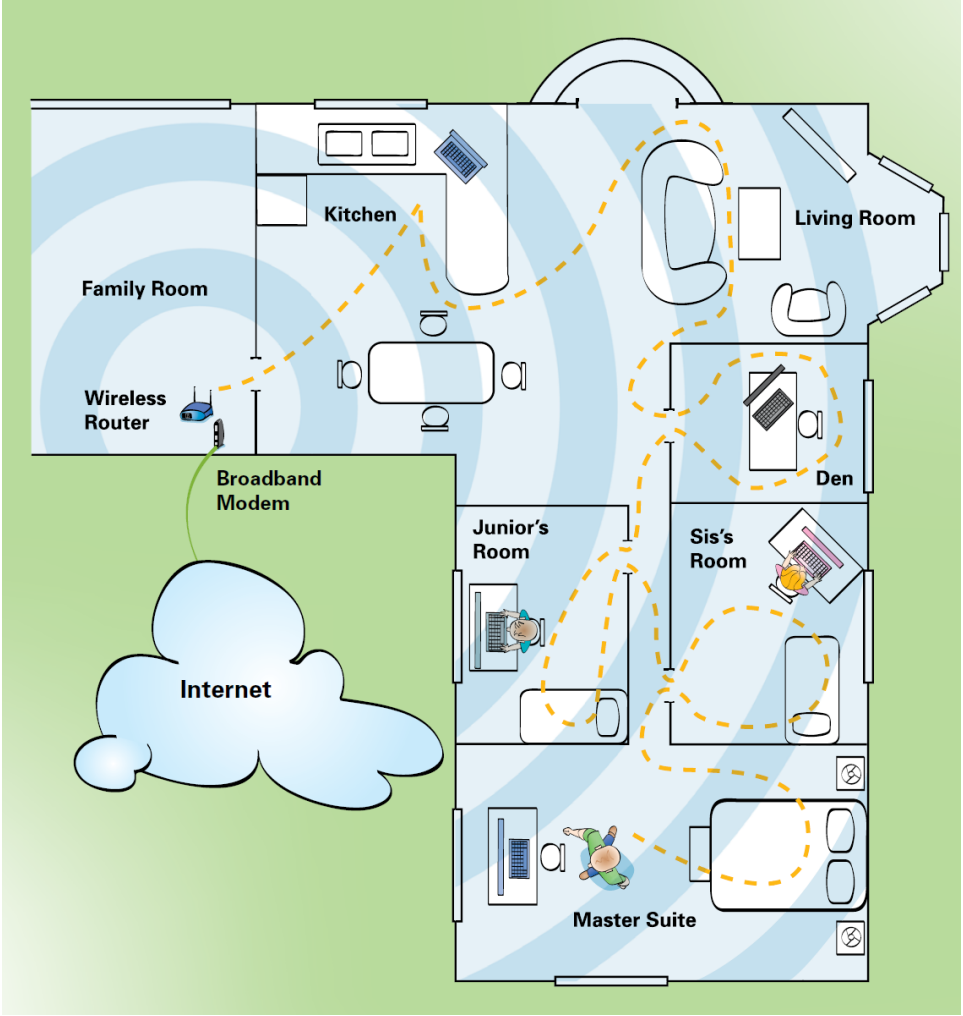


FIG. 6

341. Thus, Howard discloses that each of the ***portable device*** and the ***home appliances and other devices*** communicates wirelessly with the ***server***. To the extent that the claim requires the same wireless network (which is not conceded), a POSITA would have understood that the same wireless network (e.g., a local wireless network) can be used to enable wireless communication in the data path between the ***portable device*** and the ***server*** and in the data path between the ***home appliances and other devices*** the ***server***. It would have been obvious to a POSITA, in light of the teaching of Howard, that the ***portable device*** and the ***home appliances and other devices*** communicates with the same wireless network in the home, for example, via a wireless router.

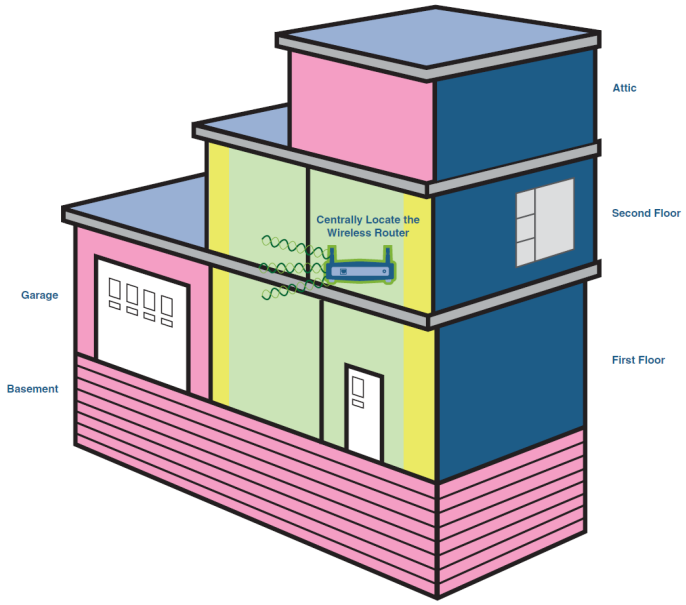
342. Further, Doherty discloses that devices at a home can communicate over a wireless network via a wireless router, which is connected to the Internet. Doherty discloses that a "wireless router acts as a central communication routing point" between multiple devices and the Internet. (EX1006, 11.) Devices located anywhere in a home can send and receive information to and from the Internet by their wireless connection to the router. Doherty explains the concept of wireless LAN (802.11) protocol, also known as Wi-Fi, and discloses 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." (EX1006, 11, see also, *Id*., 100.)

343. Doherty's exemplary figures show that home devices are connected wirelessly to a wireless router, which is connected to the Internet or the cloud. (EX1006, 100, 96, 38-39, 261, FIGS. 7-28.)



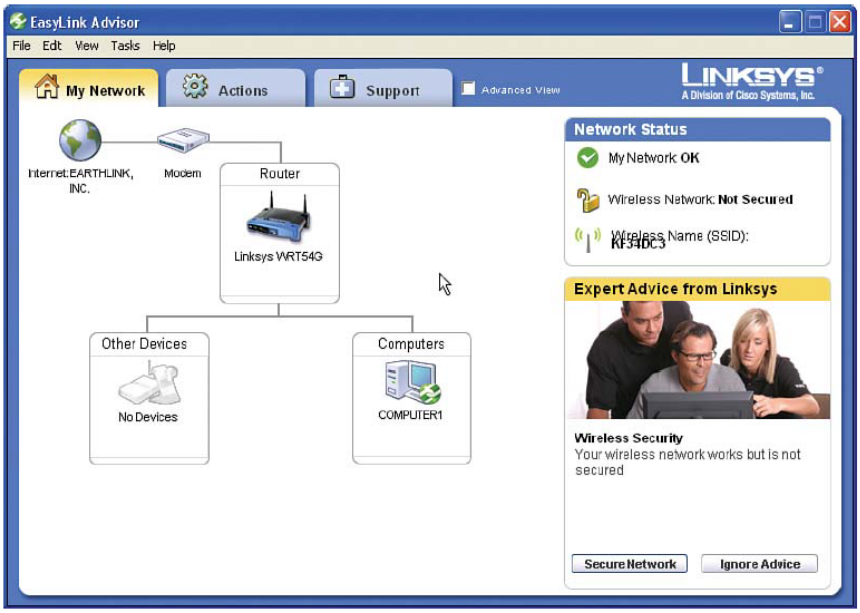
344. Doherty discloses placing the wireless router at home. "Choose a site for the router in a central location in your house, considering all three dimensions (see Figure 11-9)." (*Id*., 228, 229.)

Figure 11-9 Choose a Central Location for the Wireless Router



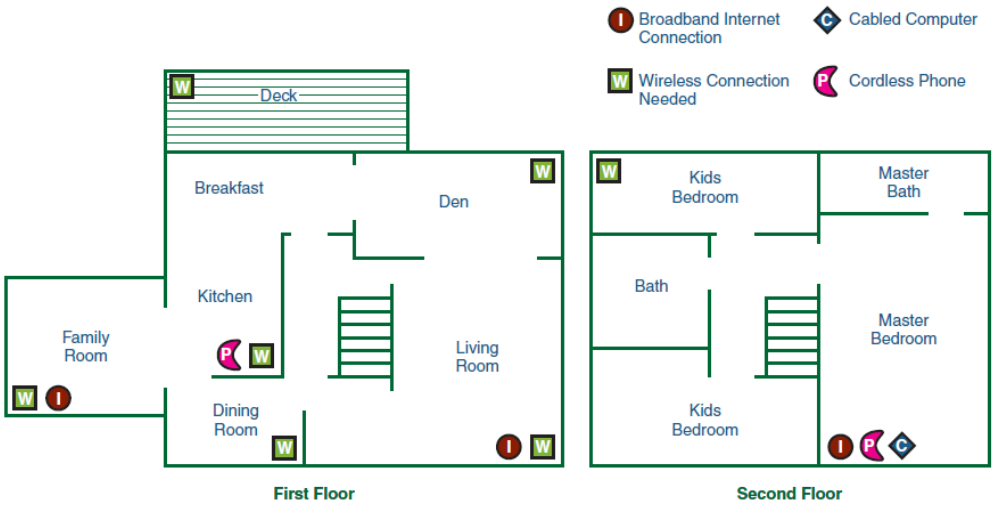
345. "Figure 7-28 shows what the initial ***network*** should look like if things are working properly, including ***connectivity to the Internet***." (*Id*., 96; *see also id*., Figure 7-28.)

Figure 7-28 EasyLink Graphical View of Your Network



346. Doherty discloses, with reference to Figures 4-3 and 4-4, that the wireless router allows connection from multiple devices. "In Figure 4-3, we used a blue diamond with a C in it to denote where we want to have a wired network connection, like a desktop computer. Finally, put a mark in the rooms or areas where you can conceive of using a ***laptop computer and other wireless devices, including video cameras, printers, video game consoles***, and so on. (This would include bathrooms, which we admit to using our computers in.) In Figure 4-3, we used a green square with a W in it to denote where we want to have a wireless network connection, like a laptop." (EX1006, 38.)

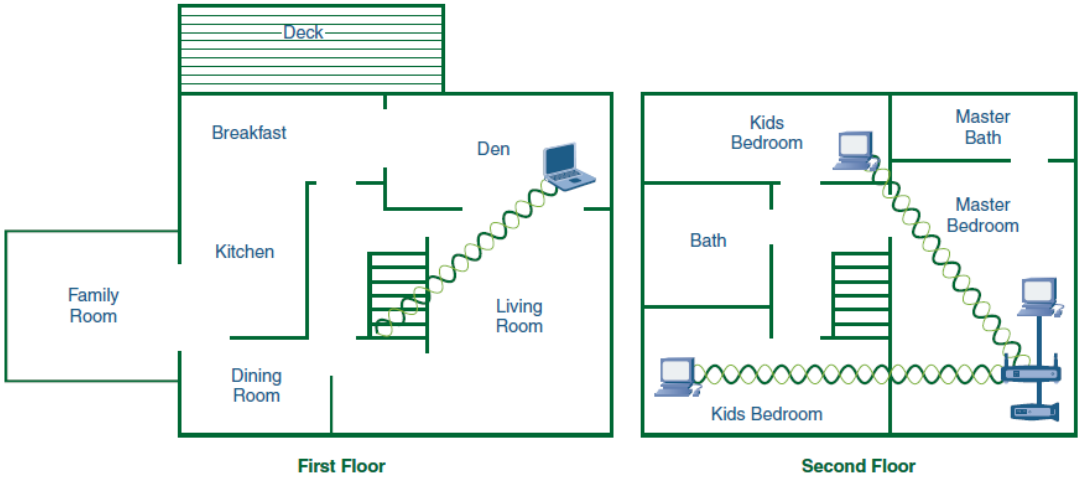
Figure 4-3 Sample Sketch of a House



347. "Because the daughter's laptop and the son's desktop computer are in the other bedrooms of the house, and not easy to run Ethernet cables to, ***these computers will have wireless connections***. The last computer is the work laptop, which is used primarily in the den, but could also be used in the living room,

breakfast nook, and on the back deck. ***Wireless is definitely the way to go for this laptop as well***. By jotting these requirements into the sketch, we have what's shown in Figure 4-4." (EX1006, 39.)

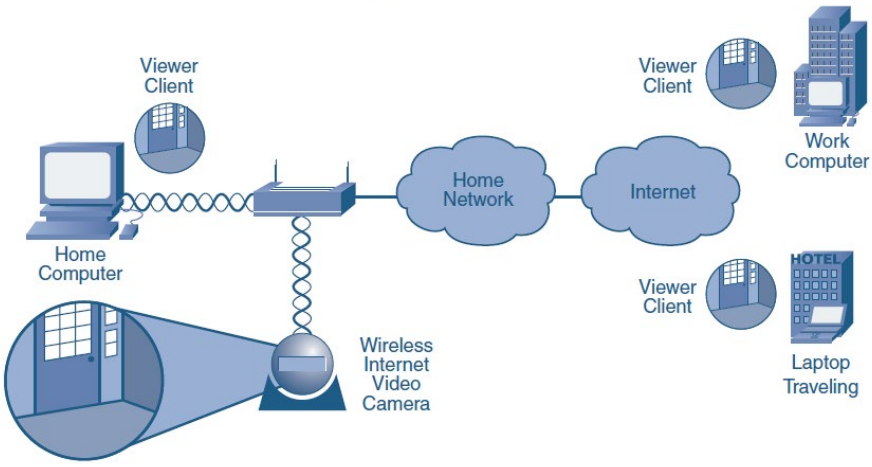
Figure 4-4 Adding Wired and Wireless Computers to the Sketch



348. In addition to laptops and desktop computers, Doherty discloses that wireless video and entertainment devices such as wireless video cameras, and digital video recorder may be connected to the wireless network through the wireless router as disclosed above. "Also popular recently have been very inexpensive wireless cameras that can be placed anywhere in your house. The ability to use either of these video cameras for live surveillance while you are not at home is limited. However, newer products like the Linksys wireless Internet video cameras have several distinct advantages: No computer is required for them to operate, so they can be located anywhere in your house (or outside, for that matter) where there is an electrical outlet. Images from the camera are viewable

with any computer from inside your house, or from anywhere outside your house where there is Internet service. In Figure 13-1, we placed a wireless Internet video camera in a house with a view of the front door." (EX1006, 261.)

Figure 13-1 Home Surveillance Using a Wireless Internet Video Camera



349. Doherty, published in 2007, also states, "[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." (EX1006, 321.) Indeed, at the time of the invention, it was prevalent to control home appliances wirelessly. (*See e.g.*, EX1007, 2:58-64, 3:15-29; EX1020, [0027]; EX1021, [0071], EX1011, [0012]-[0013], EX1012, 7:32-60, 9:37-56.)

350. It would have been obvious to a POSITA to combine the teachings of Doherty and Howard. Howard and Doherty are analogous art to the '710 patent. All are from the same field of endeavor and reasonably pertinent to the same

technical problems of providing wireless device connectivity and/or control at home. Howard describes a method and system for remotely controlling home appliances through audio commands. As discussed, Howard expressly discloses that the portable device wirelessly transmits the captured human voice data to a server (EX1005, Claim 1, Claim 9, 1:65-2:7, 3:41-48, 4:20-24, 4:57-63, *see also, id.*, 6:25-30), the server processes the data and wirelessly transmits messages back to the portable device (*id.*, 6:32-36.), and to controlled home appliances wirelessly (*id.*, 1:18-22; 5:22-26; *see also id.*, FIGS. 8-13(d); 10:64-11:5, 11:48-65; 11:66-12:14.). Further, Howard discloses that the server is Internet-connected. (*Infra* § IX.A.2.d) Doherty discloses widely adopted wireless connection techniques used in residential homes that allow devices at home to wirelessly connect to a router, which connects to the Internet. (EX1006, Title.)

351. A POSITA would have been motivated to combine Howard with Doherty such that (1) its ***portable device*** and ***home appliances and other devices*** can communicate with a wireless network via a wireless router, and (2) its ***server 200*** is a cloud server accessible via the Internet outside the user's home. For example, Howard's ***portable device*** and any of the ***home devices*** can send and receive information (i.e., captured human voice data, messages) to/from the ***server 200*** in the cloud (i.e., the Internet) via the wireless router as taught by Doherty. The speech processing server is in the cloud and communicates with the router via

the Internet. Thus, in the Howard-Doherty combination, Doherty's conventional techniques of connecting home devices to a wireless router for Internet access is applied to Howard.

352. Regarding (1), Doherty teaches that home devices can be connected via the wireless router, stating that "[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." (EX1006, 321.) In fact, at the time of the invention, it was prevalent to control home appliances wirelessly. (*E.g.*, EX1007, 2:58-64, 3:15-29; EX1020, [0027]; EX1021, [0071], EX1011, [0012]-[0013], EX1012, 7:32-60, 9:37-56.

353. Further, Howard itself recognized the same benefit of reducing resource requirement of the local device by placing the audio processing task on the server. Howard discloses that "the analysis of the optical data signals 116 and/or the audio data signals 112, which is required to identify the guide and command corresponding to the particular data signal, is ***assigned to the server 200***. Accordingly, in this embodiment, ***the circuit components and the power requirements of the portable device 10 are minimized***." (EX1005, 7:59-65.)

354. Indeed, the benefits for having a server in the cloud 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." (EX1007, 8:26-32; *see also*, EX1004, [0022], Figure 2; EX1008, [0080], [0100] – [0102], [0110] – [0117].)

355. Moreover, the Howard-Doherty system is compatible with the workflow of Howard, where the portable device wirelessly transmits the captured human voice data to an Internet-connected server (EX1005, Claim 1, Claim 9, 1:65-2:7, 3:41-48, 4:20-24, 4:57-63, 6:25-30), which processes the data and wirelessly transmit messages back to the portable device (*id*., 6:32-36.) and/or control home appliances (*id*., 1:18-22, 5:22-26, FIGS. 8-13(d), 10:64-11:5, 11:48-65, 11:66-12:14).

356. Combining Doherty with Howard would improve the wireless connectivity of Howard in the same way. Combining them would yield predictable results of enhancing Howard's capability for speech processing and device control while achieving efficient resource utilization on local devices and reliable network connectivity. The combination would also yield predictable results of conveniently

connecting devices wirelessly to the Internet-connected server, especially as Doherty discloses techniques that are designed to be used at home by laymen. "With the explosion of popularity in wireless networking, there is a proportional number of people who do not understand the technology . . . This book is written for all of you out there who want to get the benefits of wireless networking but don't feel like you have the technical background to set it up yourselves. We are confident that if you read this book and follow the steps we lay out, you will have a much better understanding of the technology and, most importantly, a working wireless network." (EX1006, Introduction.)

357. Further, it would have been obvious to try combining the teachings of Howard and Doherty, as there was a finite number of identified and predictable solutions for wirelessly connecting home devices and the Internet, such as through direct cellular data connection between the devices with the remote server. The known benefits of Doherty would prompt its application in the same field based on design incentives and other market forces.

d. 1[PR][4]: "for use with an Internet-connected server device external to the building, the method comprising:"

358. Howard discloses its method is for use with an Internet-connected server device external to the building. First, Howard discloses an ***Internet-connected server device***. With reference to FIG. 6, Howard discloses that, "the ***server 200*** is also coupled to the ***Internet 305*** via any suitable connection."

(EX1005, 9:50-52.) In addition, Howard discloses that "both the home network 304 and the ***Internet*** websites 305 are accessible by ***the server 200, 230***," and listed "examples of apparatus or websites accessible by the ***server 200*** via the ***Internet*** connection 304 [.]" (EX1005, 8:40-42, 9:52-56; *see also id.*, 15:61-16:25.) Further, "the server 200 . . . [can] retrieve information from any of the devices connected to the server 200 ***over the network or the Internet***." (EX1005, 6:52-58.)

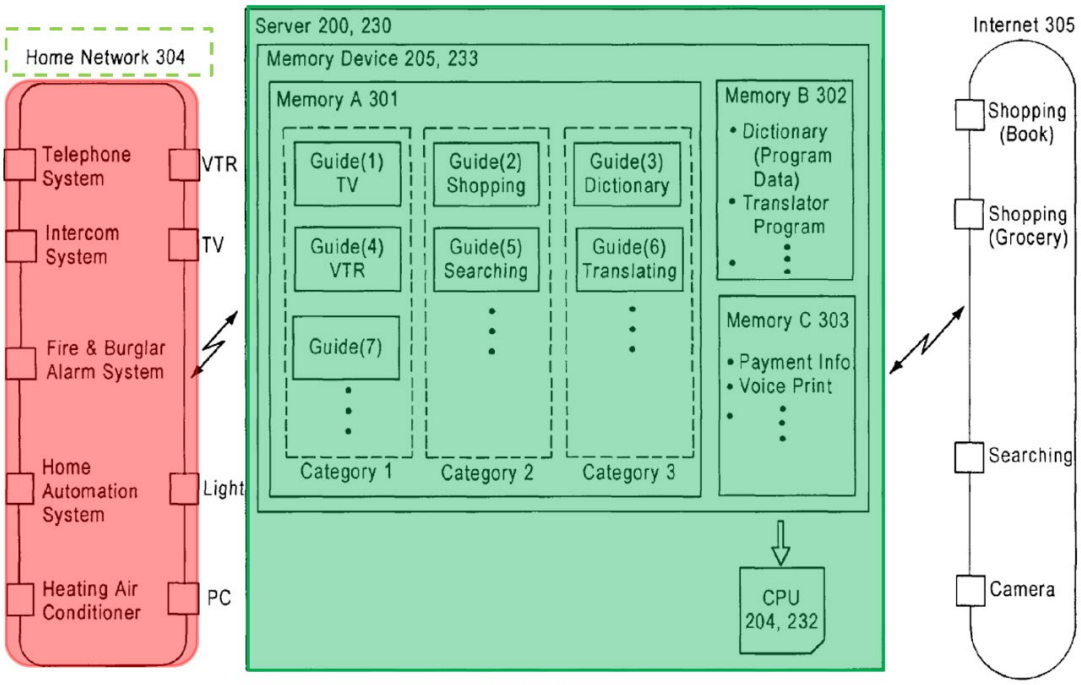


FIG. 6

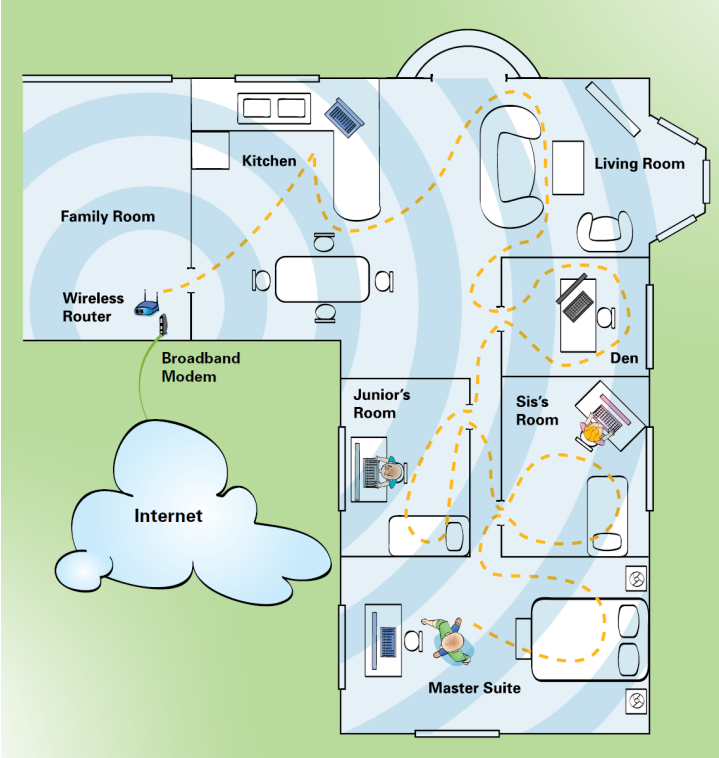
359. Further, a POSITA would have understood that Howard's server can be outside the home. While Howard mentions "home server" in some instances, Howard does not limit its system to a server at the user's home. To the contrary, Howard states that "[t]he following detailed description relates to a novel portable

device for use with a ***computer/server***, which allows an operator to control various activities within a home." (EX1005, 3:8-10.) Indeed, none of the claims in Howard limits the location of the server to home. (*See id.*, claims 1-21.)

360. At the time of the invention, remote servers (e.g., servers in the cloud) were well-known and readily available, particularly for providing computational resources for processing audio. For example, Wittenberg discloses "transmitting the audio signal to a ***remote site***; performing speech recognition processing on the audio signal at the ***remote site***," and "***the remote processor, such as a server***, performs speech recognition processing on the audio signal to determine the command to be carried out." (EX1007, Abstract, 2:50-53.) At the time of the invention, the benefits for having a server in the cloud 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." (EX1007, 8:26-32; *see also*, EX1004, [0022], Figure 2; EX1008, [0080], [0100] – [0102], [0110] – [0117].) Howard itself recognized the same benefit of reducing resource requirement of the local device by placing the audio processing task on the server. Howard discloses that "the analysis of the optical data signals 116

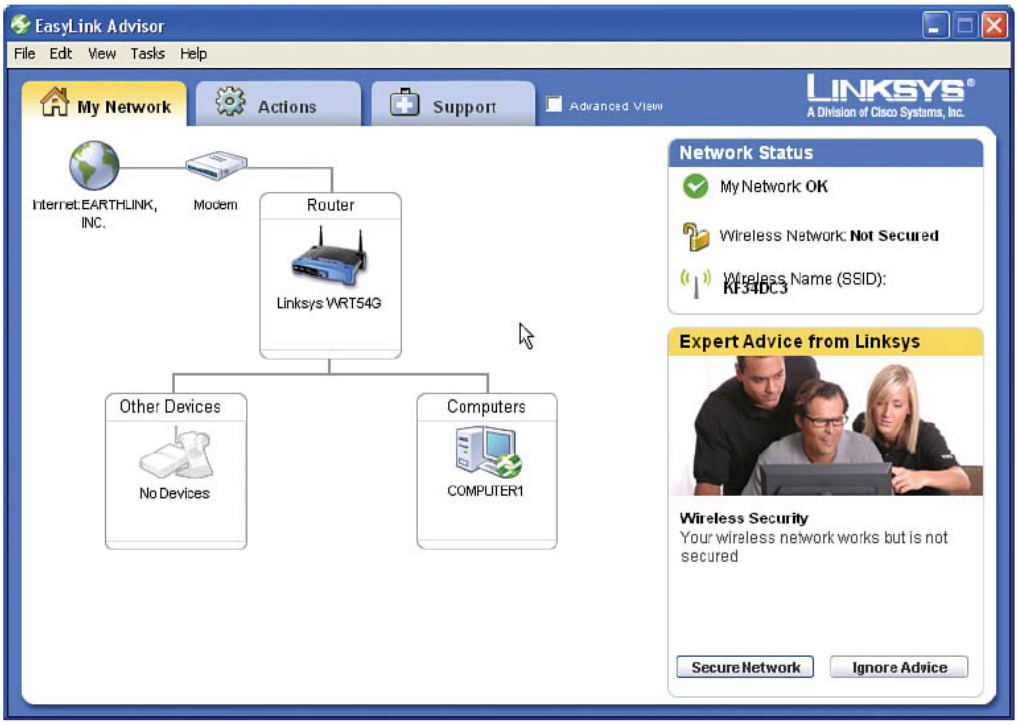
and/or the audio data signals 112, which is required to identify the guide and command corresponding to the particular data signal, is ***assigned to the server 200***. Accordingly, in this embodiment, ***the circuit components and the power requirements of the portable device 10 are minimized***." (EX1005, 7:59-65.) Howard further discloses that a home network may take the form of "a more complex hub and switching system with broadband connection to the ***Internet***." (EX1005, 1:18-22.) Thus, it would have been obvious that Howard's server can be a remote server outside the user's residential building accessible via the Internet and communicates with portable devices and home appliances via a wireless network provided by a wireless router.

361. For the same reasons, it would have been obvious to combine Howard and Doherty such that the speech processing server is remote and accessible via the Internet (i.e., in the cloud). Doherty discloses a wireless router at home acting as a "central communication routing point" between multiple devices and the Internet. (EX1006, 11, 100.) It would have been obvious to a POSITA to combine the teachings of Howard and Doherty, including implementing the speech processing server in the cloud rather than at home.



362. "Figure 7-28 shows what the initial ***network*** should look like if things are working properly, including ***connectivity to the Internet***." (*Id*., 96.) Figure 7-28 also shows that other devices and computers may connect to the router which connects to the Internet. (*Id*., Figure 7-28.)

Figure 7-28 EasyLink Graphical View of Your Network



363. The resulting system consists of the portable device sending data (e.g., voice data) wirelessly to the router, which routes it to the server in the cloud via the Internet. Same with the home devices. The server in the cloud sends data (e.g., messages) to the router via the Internet, which routes the data wirelessly to the portable device or the home devices.

e. 1[A]: "capturing, by a microphone in the client device, a first human voice data;"

364. Howard discloses capturing by a microphone in the client device, a first human voice data. First, Howard discloses a microphone in the ***client device***. Howard discloses a "***portable device*** comprising: a body; a ***microphone*** for receiving first audio data[.]" (EX1005, Claims 1 and 9.) Further, Howard discloses

that "the ***portable device*** includes: (1) a body; (2) a ***microphone*** for receiving a first audio data." (EX1005, Abstract, 1:65-2:6.) FIG. 1(a) depicts, "the ***portable device 10*** comprises a clip 2, a ***microphone*** 3, a speaker 4, a display device 5, and an optical sensor 6." (EX1005, 3:25-27*; see also id.*, FIG. 1(b).)

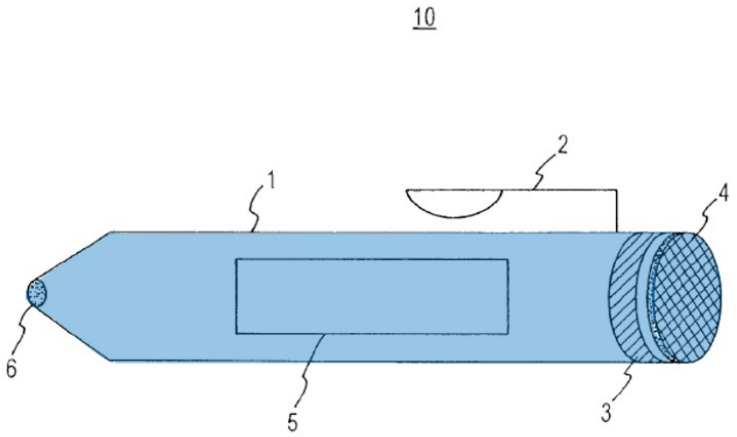


FIG. 1(a)

365. Howard discloses capturing of a ***first human voice data*** by the microphone in the ***client device***. Howard discloses a "***portable device*** comprising: a body; a ***microphone*** for receiving first audio data," "wherein the ***first audio data*** corresponds to a ***human voice***." (EX1005, Claims 1, 4, and 9.) Thus, Howard's ***portable device*** captures human voice data by its microphone 3. (EX1005, 3:24-27, 3:54-62

366. Further, with reference to in FIG. 2, Howard discloses that "the ***portable device 10*** comprises both a transmitter system 101 and a receiver system

102. … the transmitter system 101 includes the ***microphone*** 3, which functions to receive ***audio data 110***, which is usually ***human voice or speech***." (EX1005, 4:2-4, 4:10:12.) The human voice data is captured by the microphone and sent to the server. (*Id.*, 4:2-4, 4:10-24; *see also*, *id.*, 3:41-48; 6:25-30; 10:4-9, 10:35-41; 11:20-23, claims 1 and 9.)

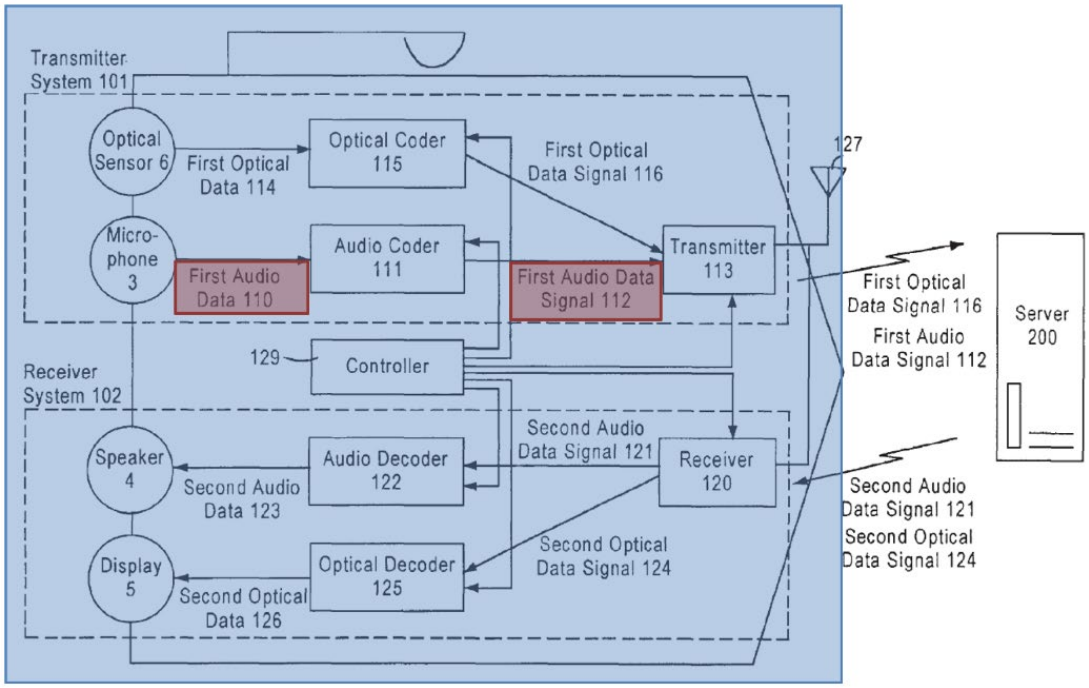


FIG. 2

367. As another example, with reference to FIG. 8, "[u]pon input of a ***human voice stating 'VTR'*** to the ***microphone*** 3 contained in the ***portable device 10***, the audio coder 115 contained in the portable device 10 functions to encode this audio data, which is then transmitted to the server 200 by the transmitter 113 of the portable device 10 in the manner detailed above;" "In response to the foregoing inquiry, it is assumed that the ***operator states 'Record'***. Thus, ***audio data***

is received by the microphone 3 and coupled to the audio coder 115, and is then processed in the same manner as set forth above. Specifically, the data is encoded by the audio coder 115 and then transmitted to the server 200 by the transmitter 113 of the portable device 10." (EX1005, 10:4-9, 10:35-38.)

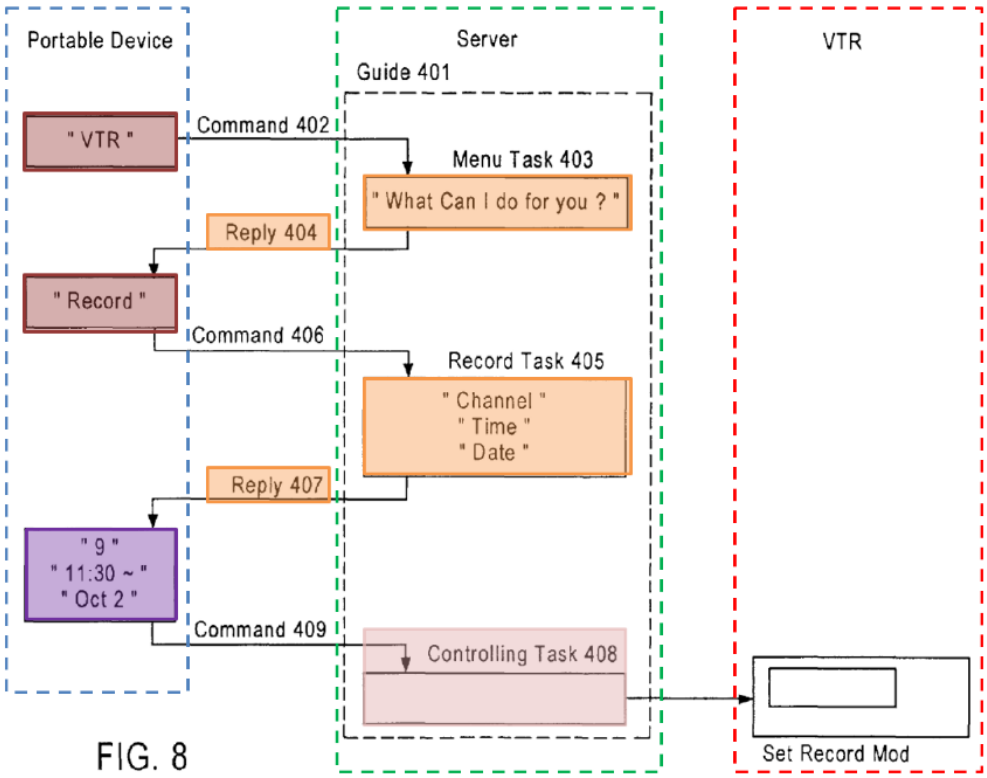


FIG. 8

368. As yet another example, with reference to FIG. 9, Howard discloses that "in operation, when the ***word 'TV'*** is input into ***microphone*** 3 in the ***portable device 10***". (EX1005, 11:20-23.) Further, "assuming the operator selects 'channel', by ***speaking channel*** into the ***microphone*** 3, this data is encoded and sent to the server 200, 230, where it is determined to correspond to command 426." (EX1005, 11:50-53.)

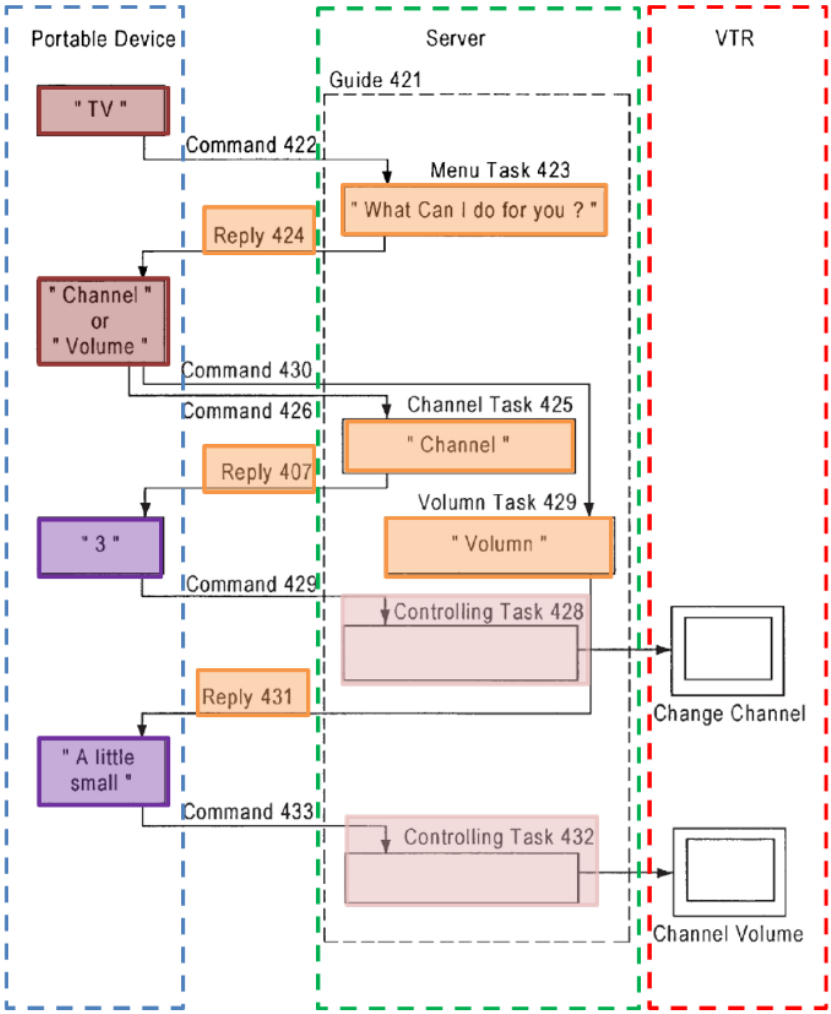


FIG. 9

f. 1[B]: "sending to the server, by the client device via the wireless network, the captured first human voice data;"

369. Howard discloses sending to the ***server***, by the ***client device*** via the wireless network, the captured ***first human voice data***.

370. First, Howard discloses that the ***client device*** sends the captured ***first human voice data*** to the ***server***. With reference to FIG. 9, Howard discloses that "in operation, when the ***word 'TV'*** is input into microphone 3 in the ***portable device 10***, in the same manner as set forth above, the data is encoded by audio

coder 111 and ***transmitted to the server 200 via the transmitter 113***." (EX1005, 11:20-23.) Further, "assuming the operator selects 'channel', by ***speaking channel*** into the microphone 3, this data is encoded and ***sent to the server 200, 230***, where it is determined to correspond to command 426." (EX1005, 11:50-53.) "Continuing this example, assuming the operator ***selected 'volume'***, this data would be encoded and ***transmitted to the server 200, 230***, where it is determined to correspond to command 430." (EX1005, 11:66-12:2.)

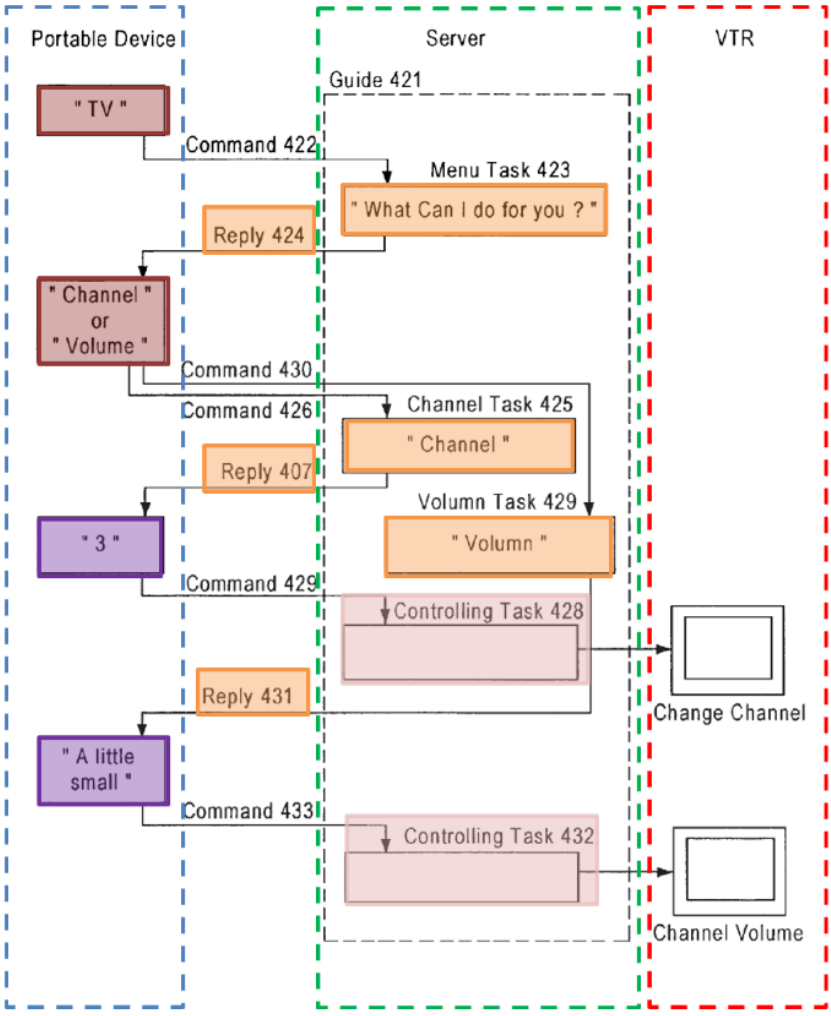


FIG. 9

371. Further, Howard discloses that the ***client device*** sends the captured ***first human voice data*** to the ***server*** "via the wireless network." With reference to FIG. 2, Howard discloses that "the ***portable device 10*** comprises both a transmitter system 101 and a receiver system 102 … the transmitter system 101 includes the microphone 3, which functions to receive ***audio data 110***, which is usually ***human voice or speech***. The microphone 3 is coupled to an audio coder 111 (e.g., CODEC) and outputs the ***audio data 110*** to the audio coder 111. The audio coder 111 functions to packet/code or convert the ***audio data 110*** to modulated ***audio data signals 112***, which are suitable for ***wireless transmission*** to the ***server 200***. Examples of acceptable modulation schemes to be implemented by the audio coder 111 include, but are not limited to sigma delta code, PCM, etc. The audio coder 111 is coupled to a ***transmitter 113, which functions to receive modulated audio data signals 112 from the audio coder 111 and to transmit the modulated audio data signals to the server 200 via an antenna 127***." (EX1005, 4:2-4, 4:10-24; *see also*, *id.*, 3:41-48; 6:25-30; 10:4-9, 10:35-41; claims 1 and 9.) In the Howard-Doherty system, Howard's ***portable device 10*** sends voice data via a wireless network provided by a wireless router to Howard's ***server***. (*Supra* §IX.I.3.c; EX1006, 11, 38-39, 96, 100, 228-229, 261.) Further, it would have been obvious that Howard itself discloses this architecture. (*Supra* §§ IX.I.3.c- IX.I.3.d.)

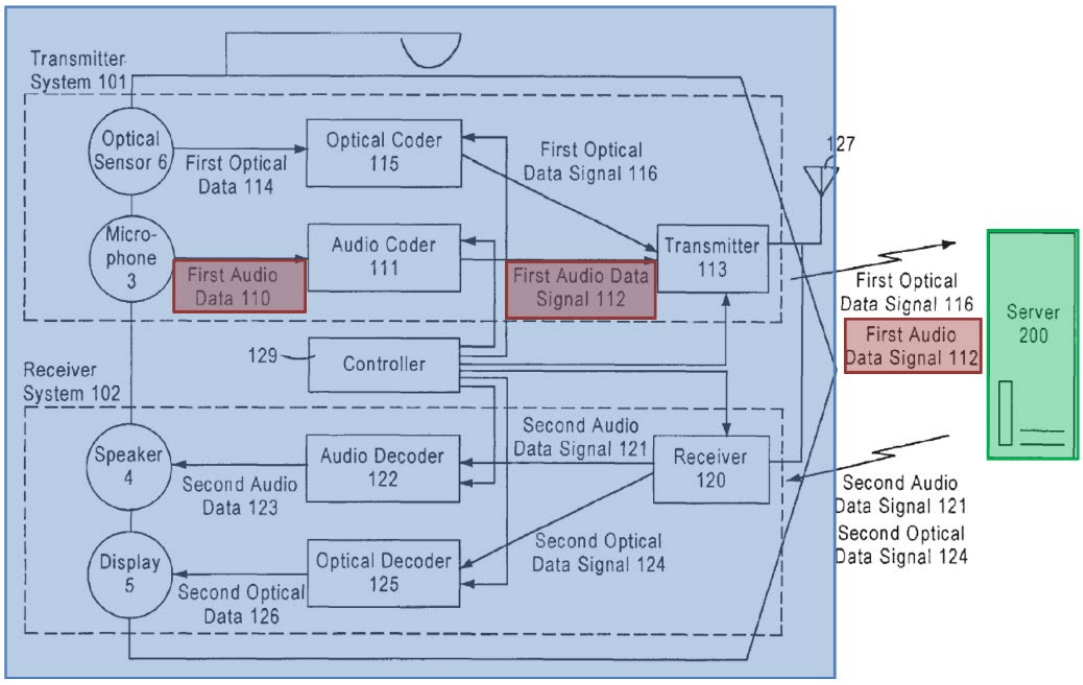


FIG. 2

g. 1[C]: "receiving, by the server over the Internet, the captured first human voice data;"

372. Howard discloses that the ***server*** receives the captured ***first human voice data*** over the Internet. Howard discloses the "portable device compris[es] . . . a transmitter for transmitting at least the ***audio command*** or the optical command to the ***server***" and "the ***server*** compris[es]: a communication device for ***receiving*** at least ***first audio data signals*** or the first optical data signals transmitted by said transmitter[.]" (EX1005, Claim 9.)

373. For example, with reference to FIG. 9, Howard discloses that "in operation, when the ***word 'TV'*** is input into microphone 3 in the portable device 10, in the same manner as set forth above, the data is encoded by audio coder 111 and transmitted to the ***server 200*** via the transmitter 113. ***Upon receipt of the data,***

the data analyzer 202 [of server 200], utilizing the speech recognition software, functions to interpret the language fragment and determine if the received audio data corresponds to any preprogrammed language fragments." (EX1005, 11:20-23.) Further, "assuming the operator selects 'channel', by ***speaking channel*** into the microphone 3, this data is encoded and sent to the ***server 200, 230***, where it is determined to correspond to command 426. ***Upon receipt of the command 426, the program associated with guide 421 functions to generate reply 427 (task 425)***, which is an audio question requesting the operator to indicate what channel should be selected." (EX1005, 11:50-53; *see also id.*, 11:66-12:2.)

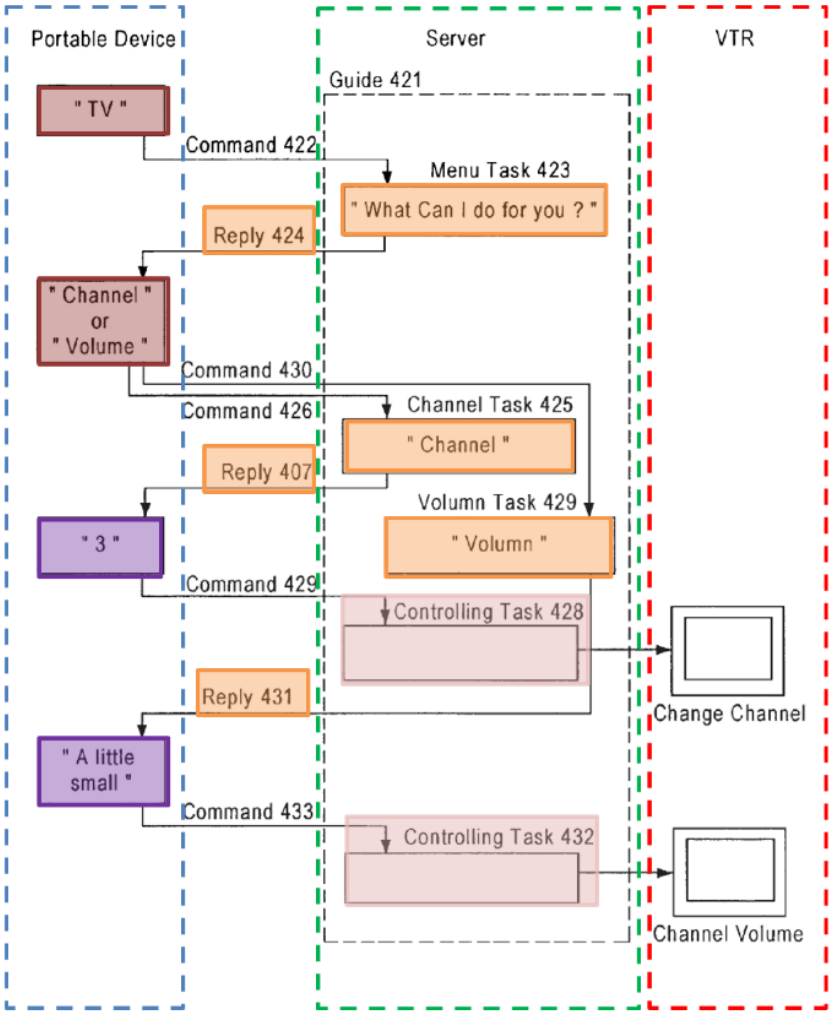


FIG. 9

374. With reference to FIG. 4, Howard discloses that "the ***server 200*** includes a communication device 201, a data analyzer 202, a CPU 204, a data synthesizer 207 and an activity guide 300. The communication device 201 functions to receive ***audio data signals 112*** and/or optical data signals 116 sent by the transmitter 113 in the portable device 10[.]" (EX1005, 6:25-30.)

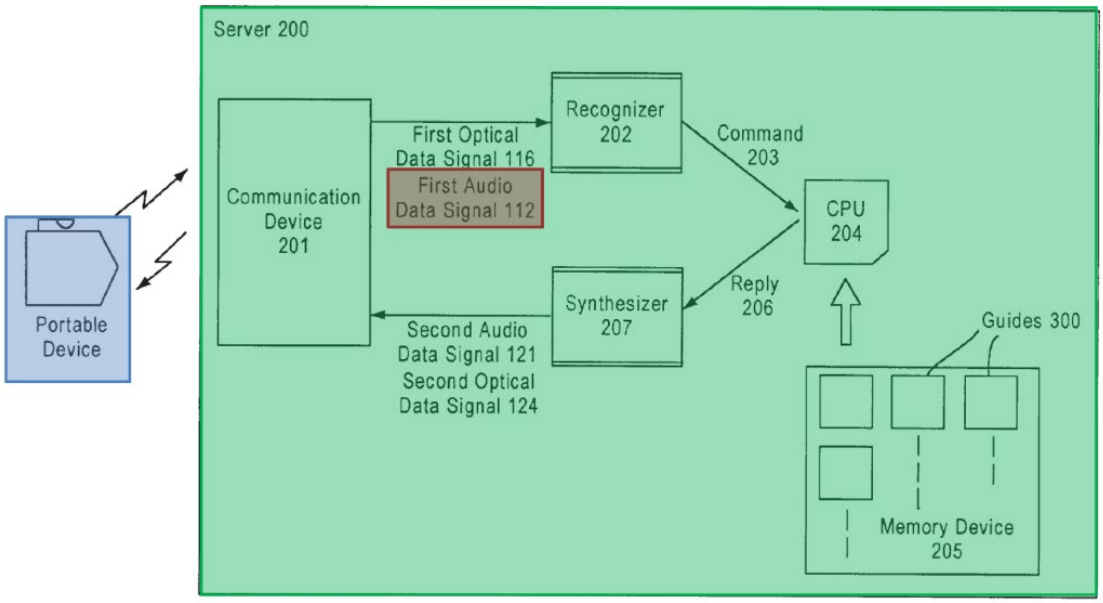


FIG. 4

375. Howard discloses in a few example applications of the invention that the audio data transmitted by the client device to the server is received by the server: "Upon input of a ***human voice stating 'VTR'*** to the microphone 3 contained in the portable device 10, the audio coder 115 contained in the portable device 10 functions to encode this ***audio data***, which is then transmitted to the ***server 200*** by the transmitter 113 of the portable device 10 in the manner detailed above. Upon receipt of the ***transmitted data*** by the ***server 200***, the ***audio data*** is coupled to the data analyzer 202 via the communication device 201." (EX1005, 10:4-12.) *See also*, EX1005, 11:48-53, 11:66-12:2 (data transmitted to the server "where it is determined to correspond to command").

376. Howard discloses that the ***server*** receives the captured ***first human voice data*** over the Internet. As discussed above, Howard's ***server 200*** is

connected to the Internet. (*Supra* § IX.I.3.c - IX.I.3.d). Thus, a POSITA would have understood that Howard discloses or at least suggests an implementation in which the ***first human voice data*** is transmitted to the ***server 200*** via the Internet.

377. As discussed above, in the Howard-Doherty system, Howard's ***portable device*** sends data (e.g., voice data) via a wireless network provided by a wireless router and Howard's ***server***, which is a remote server in the cloud, receives the data over the Internet. (*Supra* § IX.I.3.c - IX.I.3.d; EX1006, 11, 38-39, 96, 100, 228-229, 261.) Further, it would have been obvious that Howard itself discloses this architecture. (*Supra* § IX.I.3.c - IX.I.3.d.)

h. 1[D]: "processing, by the server, the captured first human voice data;"

378. Howard discloses processing, by the ***server***, the captured ***first human voice data***. With reference to FIG. 9, Howard discloses that the server processes the user command captured, "in operation, when the ***word 'TV'*** is input into microphone 3 in the portable device 10, in the same manner as set forth above, the data is encoded by audio coder 111 and transmitted to the ***server 200*** via the transmitter 113. ***Upon receipt of the data, the data analyzer 202 [of the server 200], utilizing the speech recognition software, functions to interpret the language fragment and determine if the received audio data corresponds to any preprogrammed language fragments.*** In the current example, the received ***audio data 'TV'*** matches the preprogrammed 'TV' language fragment, which

corresponds to guide 421 and command 422. Accordingly, the CPU 204 retrieves guide 421 and begins execution of the program associated therewith, the first step of which requires inputting the phrase 'TV' (command 422) into the guide 421." (EX1005, 11:20-23; *see also id.*, FIGs. 8 and 10-13(d); 11:50-53; 10:10-22; 10:42-51, 10:64-11:1, 11:50-53, 11:58-61, 11:66-12:2, 12:7-10.)

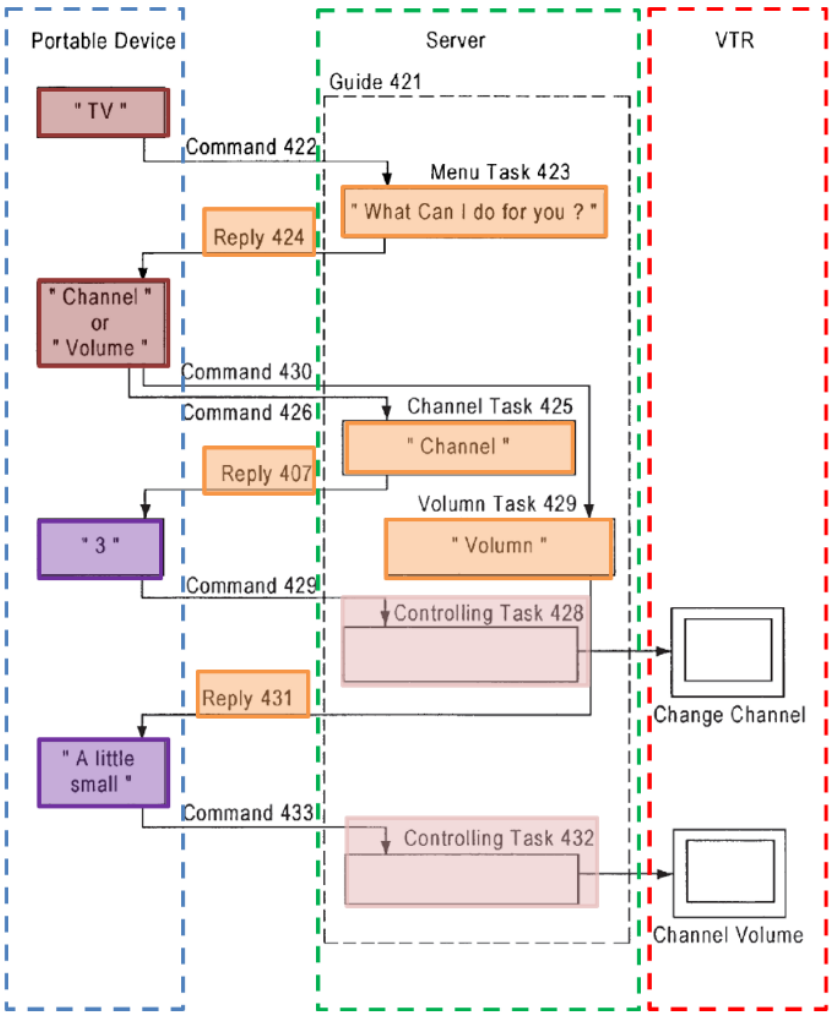


FIG. 9

379. Further, with reference to FIG. 4, Howard discloses that "the ***analysis*** of the optical data signals 116 and/or the ***audio data signals 112***, which is required

to identify the guide and command corresponding to the particular data signal, is assigned to the ***server 200***. Accordingly, in this embodiment, the circuit components and the power requirements of the portable device 10 are minimized." (EX1005, 7:59-65.)

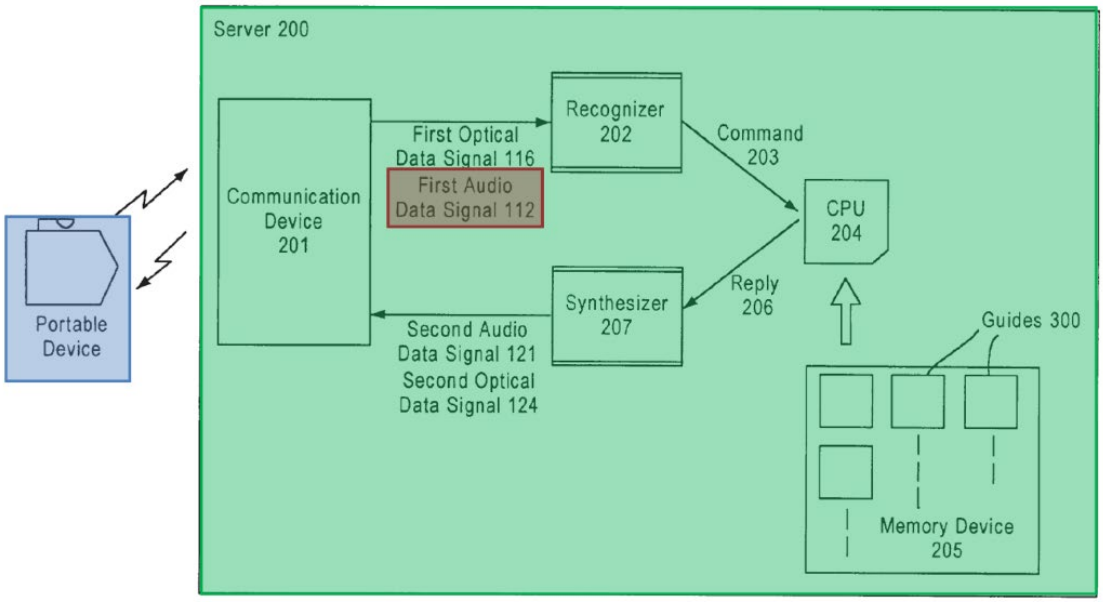


FIG. 4

380. In addition, Howard discloses that "the data analyzer 202 contained in the ***server 200*** functions to ***convert*** the incoming ***audio*** and optical ***data*** into a corresponding command. More specifically, in the given embodiment, the data analyzer 202 incorporates the ***speech recognition system*** noted above to convert the ***incoming audio data*** into a corresponding command. This speech recognition system has the ability of limiting the natural language dialog domain to a particular device, system or discourse, allowing for a more accurate interpretation of the meaning of the ***operator's input speech***. The data analyzer 202 interprets the

meaning of the ***operator's input speech*** and provides appropriate information (e.g., commands) to the ***server 200*** [.]" (EX1005, 6:43-58; *see also id.*, 6:59-67.)

i. 1[E]: "responsive to the processing, sending a first message, by the server to the client device over the Internet;"

381. Howard discloses responsive to the processing, sending a ***first message***, by the ***server*** to the ***client device*** over the Internet.

382. First, with reference to FIG. 9, Howard discloses that the ***server*** processes the captured voice command and, in response to the processing, sends a ***first message*** to the ***client device***. "Upon receipt of the data, the data analyzer 202, utilizing the speech recognition software, functions to interpret the language fragment and determine if the received audio data corresponds to any preprogrammed language fragments. In the current example, the received ***audio data 'TV'*** matches the preprogrammed 'TV' language fragment, which corresponds to guide 421 and command 422. Accordingly, the CPU 204 retrieves guide 421 and begins execution of the program associated therewith, the first step of which requires inputting the phrase 'TV' (command 422) into the guide 421. Upon receipt of command 422, the program associated with guide 421 functions to generate ***reply 424*** (task 423), which is an ***audio question requesting the operator to indicate what her/she would like done***. In the given example, the ***reply*** sent back to the operator states '***What can I do for you***'. The data representing reply 424 is generated by the CPU 204 and coupled to the data synthesizer 207 such that

it can be encoded in a manner suitable for transmission to the ***portable device 10*** via the communication device 201. Upon receipt of the encoded reply data by the receiver 120 of the ***portable device 10***, the data is coupled to the audio decoder 122, and then coupled to the speaker 4 so as to be reproduced for the user." (EX1005, 11:20-23; *see also id.*, FIGs. 8 and 10-13(d); 10:15-29; 10:47-60, 11:27-43, 11:50-58, 11:66-12:6.)

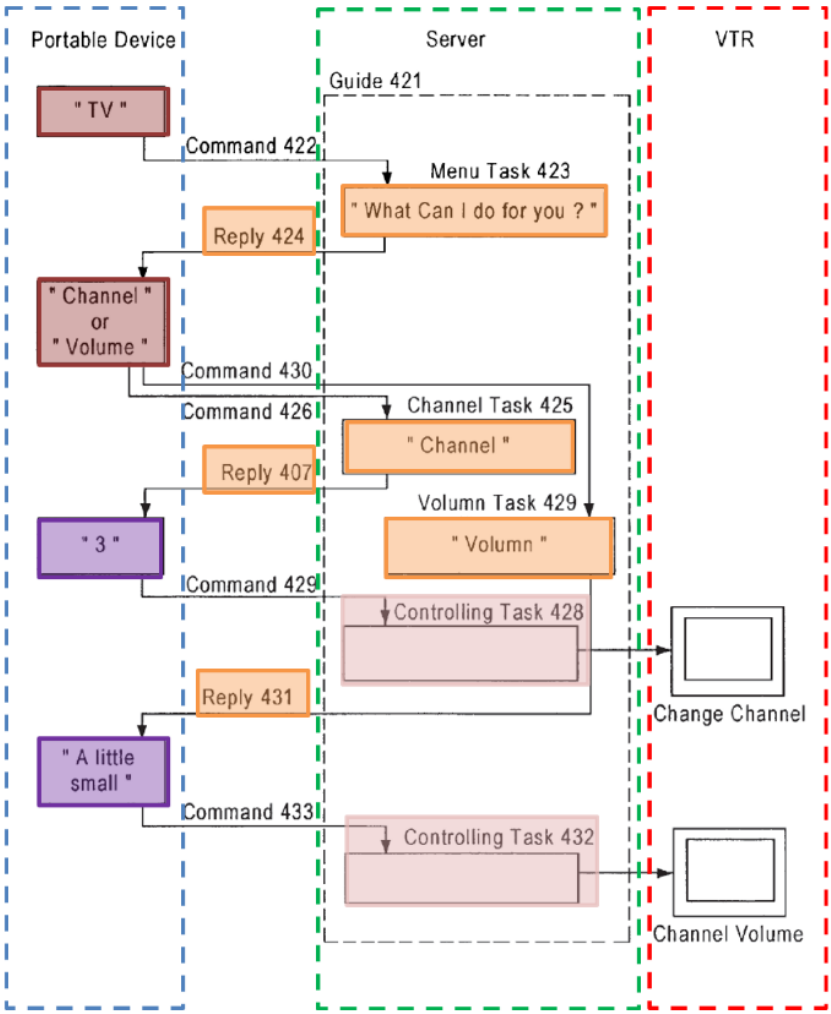


FIG. 9

383. Further, "assuming the operator selects 'channel', by ***speaking channel*** into the microphone 3, this data is encoded and sent to the ***server 200, 230***, where it is determined to correspond to command 426. Upon receipt of the command 426, the program associated with guide 421 functions to generate ***reply 427*** (task 425), which is an audio question requesting the operator to indicate what channel should be selected. The ***reply 427*** is then forwarded to the operator." (EX1005, 11:50-53.)

384. Howard further discloses that "in the event there is no matching 'language fragment', various options are available. For example, the ***server 200*** can generate an ***audio*** and/or optical ***signal*** to inform the user that the last instruction was not defined or not understood, or that the user should try again." (EX1005, 7:5-9.) The sending of the first message (i.e., ***audio*** and/or optical ***signal***) is in response to the processing of the command (which determines which command to correspond to and whether there is a matching 'language fragment').

385. FIG. 4 also discloses sending a ***first message***, by the ***server*** to the ***client device***, as shown below. Howard discloses that "[i]f the application corresponding to the activity guide 300 selected by the CPU 204 generates a ***response or data*** that needs to be relayed to the user, CPU 204 generates the necessary ***reply 206*** and sends the ***reply 206*** to the data synthesizer 207. The data synthesizer 207 functions to convert the ***reply 206*** to ***audio data signals 121*** and/or

optical data signals 124. The output of the data synthesizer 207 is coupled to the communication device 201, which functions to modulate the ***audio signals 121*** and the optical data signals 124 and then transmit the signals to the ***portable device 10***." (EX1005, 7:43-53; *see also id.*, 8:37-54; 8:59-61, 8:62-9:1.)

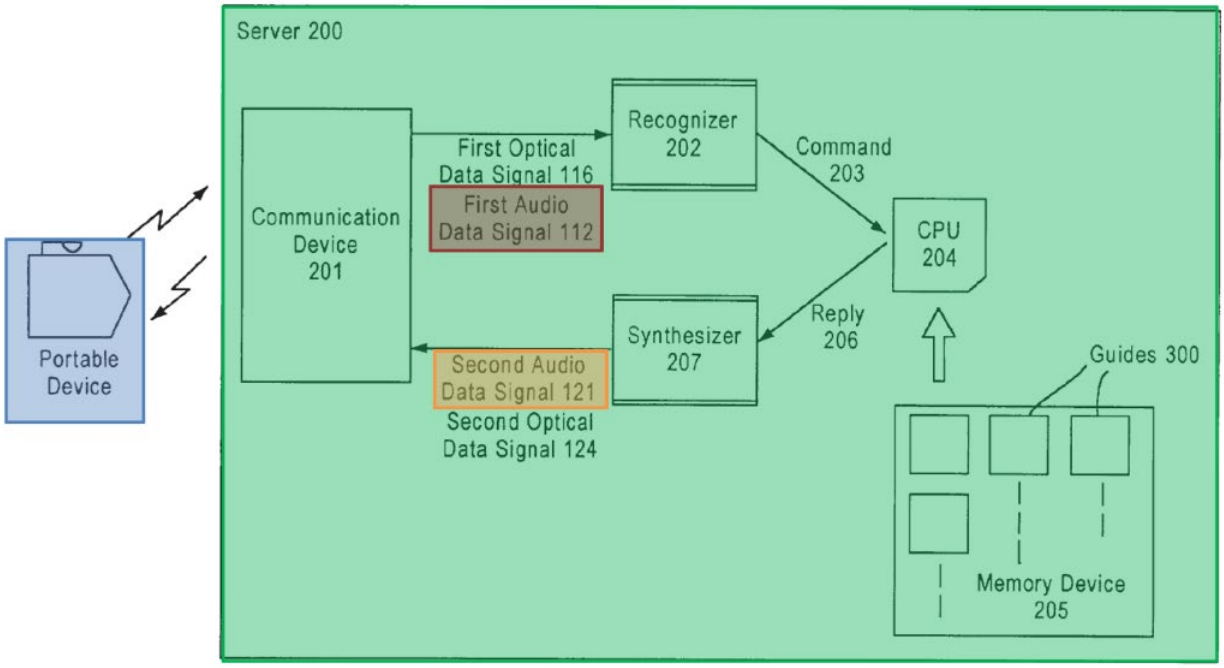


FIG. 4

386. Further, Howard discloses that the ***server*** sends a ***first message*** to the ***client device*** over the Internet responsive to processing. As discussed above, Howard's ***server 200*** is connected to the ***Internet***. (*Supra* § IX.I.3.c - IX.I.3.d). Thus, a POSITA would have understood that Howard discloses or at least suggests an implementation in which the ***first message*** is transmitted to the ***server 200*** via the Internet.

387. As discussed above, in the Howard-Doherty system, Howard's ***server***, which is a remote server in the cloud, sends data over the Internet to the wireless router, which routes the data wirelessly to the portable device or the home devices. (*Supra* §§ IX.I.3.c - IX.I.3.d; EX1006, 11, 38-39, 96, 100, 228-229, 261.) Further, it would have been obvious that Howard itself discloses this architecture. (*Supra* §§ IX.I.3.c - IX.I.3.d.)

j. 1[F]: "receiving, by the client device via the wireless network, the first message;"

388. Howard discloses receiving, by the ***client device*** via the wireless network, the ***first message***.

389. Howard explicitly discloses that "[t]he ***portable device*** . . . compris[es]: a receiver for receiving an ***audio reply*** from the server; an audio synthesizer for converting the ***audio reply*** to second audio data; and a speaker for outputting the second audio data." (EX1005, Claim 2.)

390. With reference to FIG. 9, Howard discloses that ***client device*** receives the ***first message***. "Upon receipt of command 422, the program associated with guide 421 functions to generate ***reply 424*** (task 423), which is an ***audio question requesting the operator to indicate what her/she would like done***. In the given example, the ***reply*** sent back to the operator states '***What can I do for you***'. The data representing reply 424 is generated by the CPU 204 and coupled to the data synthesizer 207 such that it can be encoded in a manner suitable for transmission to

the ***portable device 10*** via the communication device 201." (EX1005, 11:35-43; *see also id.*, FIGs. 8 and 10-13(d); 11:50-53; 10:23-33; 10:52-63, 11:35-47, 11:53-58, 12:2-6 .)

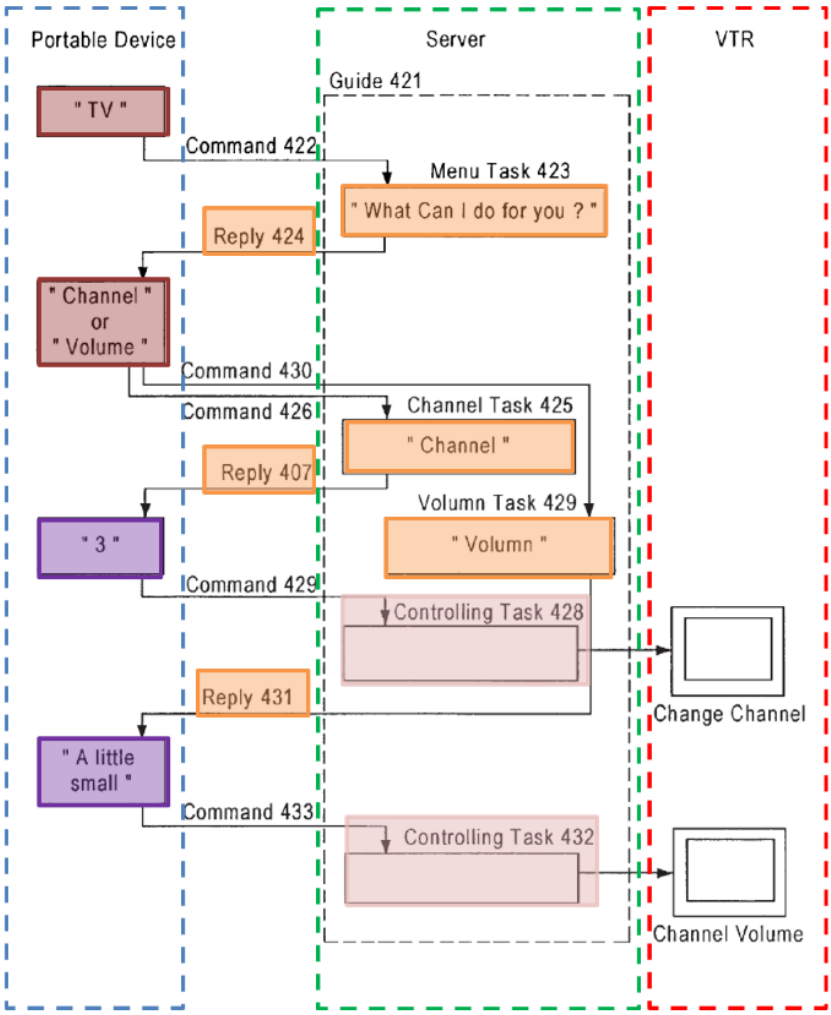


FIG. 9

391. In addition, as discussed previously, Howard teaches that "the ***portable device 10*** comprises both a transmitter system 101 and a receiver system 102." (EX1005, 4:2-4.) Howard further teaches that "[t]he receiver system 102 comprises a receiver 120 for receiving incoming ***audio data signals 121*** and/or

incoming optical data signals 124, which are transmitted by the server 200 and coupled to the receiver 120 via antenna 127." (EX1005, 4:39-43.)

392. As discussed above, in the Howard-Doherty system, Howard's ***server***, which is a remote server in the cloud, sends data over the Internet to the wireless router and Howard's ***portable device*** receives the data via the wireless network provided by the wireless router. (*Supra* § IX.I.3.c; EX1006, 11, 38-39, 96, 100, 228-229, 261.)

393. Further, a POSITA would have understood that Howard itself discloses or at least suggests an implementation in which the ***first message*** is transmitted by the ***server 200*** over the Internet and received by the ***client device*** via "the wireless network." As shown in FIG. 2, the ***portable device 10*** comprises a transmitter 113 and an antenna 127 for receiving audio data signals wirelessly from the ***server 200***. (*Supra* § IX.I.3.c) Howard discloses its portable device has an "antenna . . . which functions to allow for ***wireless communication*** between the ***portable device 10*** and a ***server 200*** forming part of the home gateway." (*Supra* § IX.I.3.c; EX1005, 3:41-48.) Further, Howard's ***server 200*** is connected to the ***Internet***. (*Supra* § IX.I.3.d.) Thus, Howard teaches the receipt of such reply (***first message***) by the portable device 10 from the server 200 or 230 through the wireless network.

k. 1[G]: "operating a first actuator in the client device in response to the received first message;"

394. Howard discloses operating a ***first actuator*** in the ***client device*** in response to the received ***first message***. First, Howard discloses a ***first actuator*** in the ***client device***. For example, Howard discloses that "[t]he ***portable device*** of claim 1 further comprising: a receiver for receiving an ***audio reply*** from the server; an audio synthesizer for converting the ***audio reply*** to second audio data; and a ***speaker for outputting the second audio data***." (EX1005, Claims 2 and 9.) As another example, Howard teaches that "[t]he ***portable device 10*** comprises a clip 2, a microphone 3, a ***speaker 4***, a display device 5, and an optical sensor 6." (EX1005, 3:25-27.)

395. With reference to FIG. 9, Howard discloses operating a ***speaker*** in response to the received ***first message*** sent to the ***client device***. Howard discloses that "[u]pon receipt of the ***encoded reply data*** by the receiver 120 of the ***portable device 10***, the data is coupled to the audio decoder 122, and then coupled to the ***speaker 4*** so as to be reproduced for the user." (EX1005, 11:27-47; *see also id.*, FIGs. 8 and 10-13(d); 10:52-63; 11:53-58; 11:38-47.)

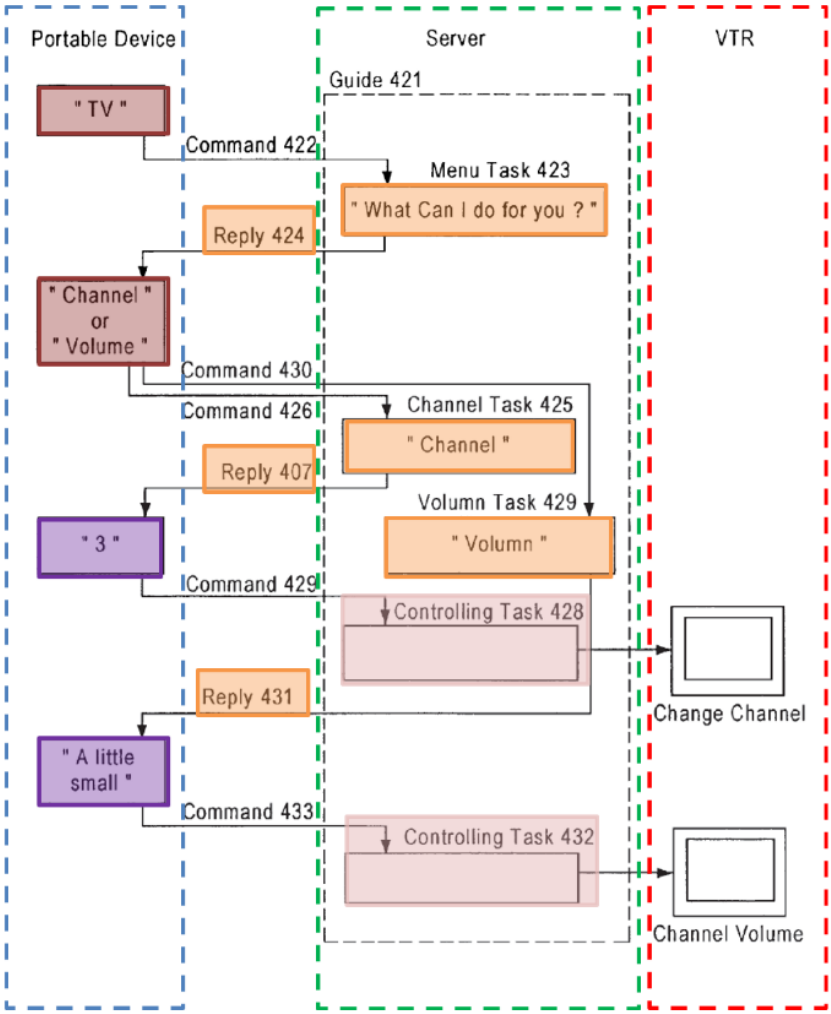


FIG. 9

396. Howard further discloses, with reference to FIG. 2, that "the receiver system 102 comprises a receiver 120 for receiving ***incoming audio data signals 121*** and/or incoming optical data signals 124, which are transmitted by the server 200 and coupled to the receiver 120 via antenna 127. In the event an audio signal is transmitted by the server 200, the receiver 120 couples the ***audio data signal 121*** to an audio decoder 122, which functions to unpacket/decode or convert the ***audio***

signals 121 to audio data 123. The output of the audio decoder 122 is coupled to the *speaker 4*, which functions to output the audio data 123." (EX1005, 4:39-49.)

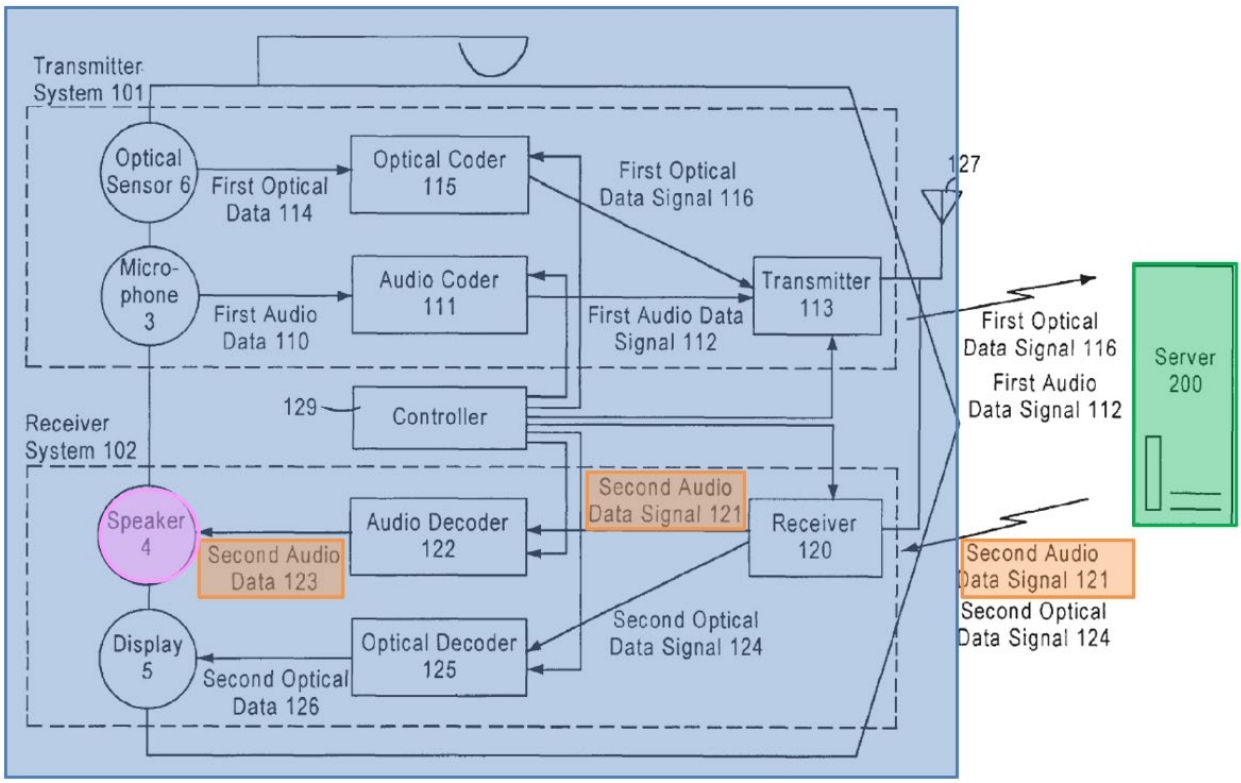


FIG. 2

397. As discussed, operating a speaker involves operating an actuator. (*Supra* § IX.A.2.k.) At the time of the invention, a POSITA would have understood that a speaker has an actuator to produce sound. (*Id.*) For example, the actuator 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. The operation of speakers can be generally summarized as below: "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." (EX1017, 122.)

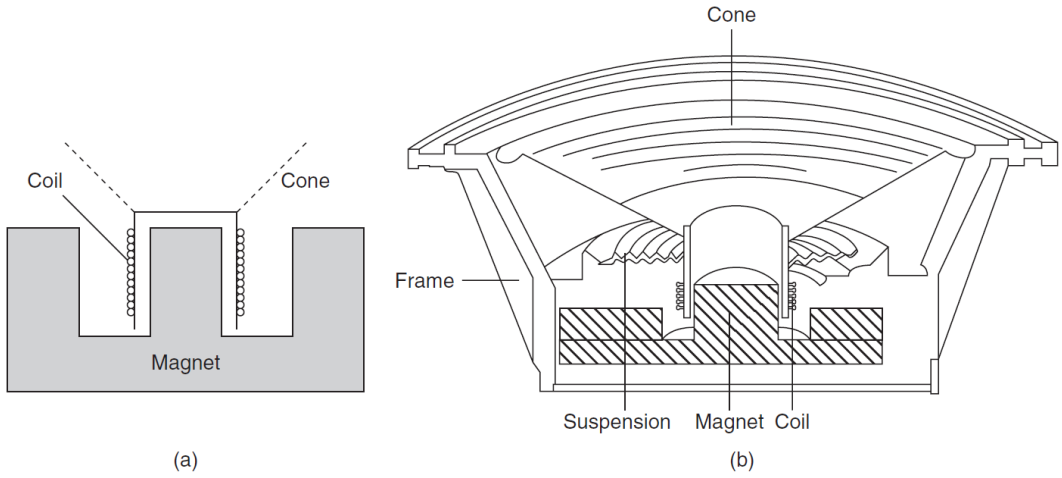


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

398. "[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." (EX1013, 763.) Thus, a POSITA would have understood that Howard discloses operating an actuator (e.g., ***speaker 4***) in the client device (e.g., ***portable device 10***) in response to the received first message (e.g., ***reply 424, audio signals 121***). Indeed, the specification of the '710 patent itself, states that a speaker as an actuator. (*See* EX1001, 105:48-52; 105:61-106:4.)

399. Another example of the recited "first actuator" is a display for providing visual information. Howard discloses a ***display screen*** in the ***client device***. For example, with reference to FIG. Howard discloses that "the ***portable device 10*** comprises a clip 2, a microphone 3, a speaker 4, a ***display device 5***, and an optical sensor 6. The clip 2 is disposed on the side of the body 1 and functions to clip the portable device 10 to, for example, the shirt pocket of the user." (EX1005, 3:25-27.) "The ***display device 5***, for example an ***LCD device***, is placed in the center portion of the ***portable device 10***." (EX1005, 3:34-36.)

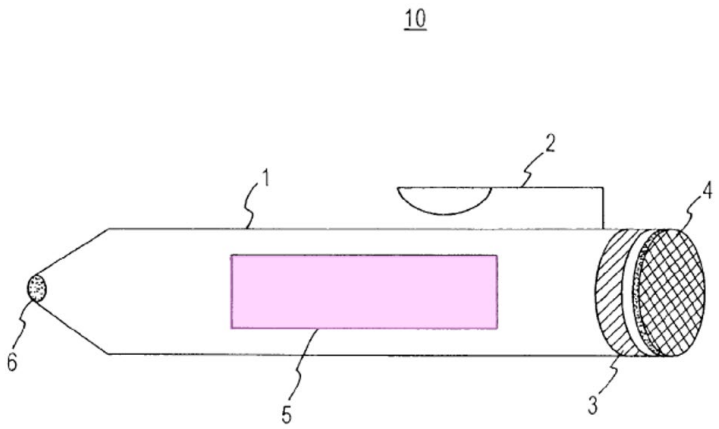


FIG. 1(a)

400. With reference to FIG. 8, Howard discloses operating a ***display*** in response to the received ***first message*** sent to the ***client device***, "upon receipt of command 402, the program associated with guide 401 functions to generate ***reply 404*** (task 403), which is an ***audio question 'What can I do for you?'***. The data

representing reply 404 is generated by the CPU 204, 232 and coupled to the data synthesizer 207 such that it can be encoded in a manner suitable for transmission to the portable device 10 via the communication device 201. Upon receipt of the encoded reply data by the receiver 120 of the portable device 10, the data is coupled to the audio decoder 122, and then coupled to the speaker 4 so as to be reproduced for the operator. Alternatively, the ***reply 404*** can be displayed on the ***display 5*** as a visual message." (EX1005, 10:23-34.)

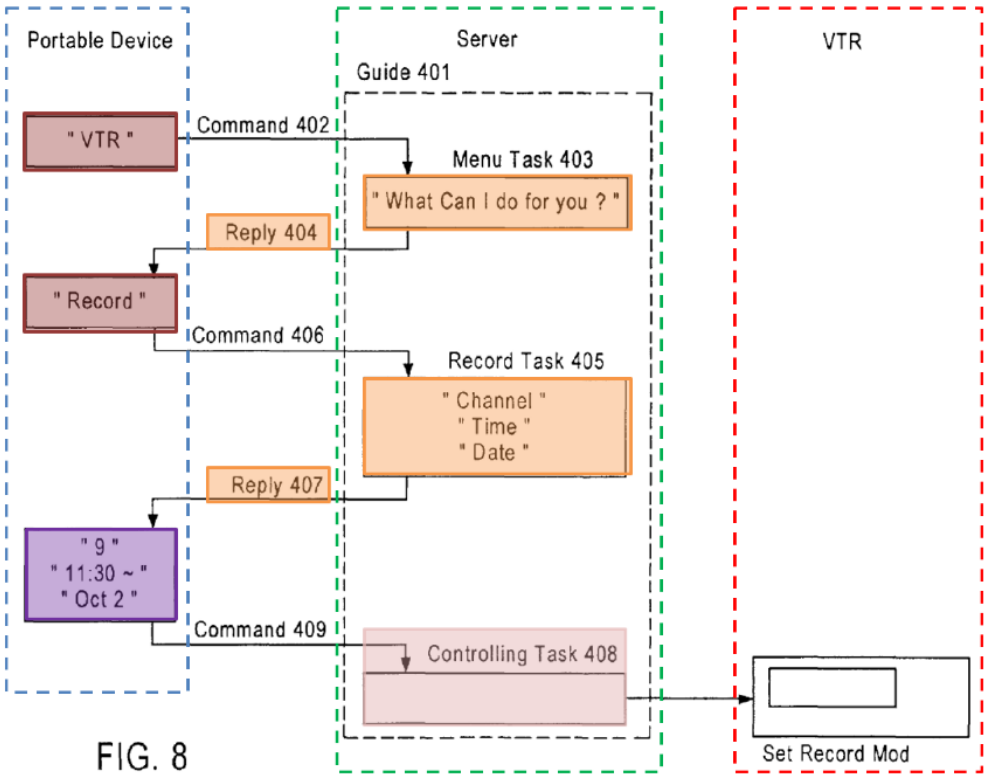


FIG. 8

401. Further, Howard discloses, with reference to FIG. 2, that "in the event an ***optical signal*** is transmitted by the server 200, the receiver 120 couples the optical data signal 124 to an optical decoder 125, which functions to

unpacket/decode or convert the optical data signals 124 to optical data 126. The output of the optical decoder 125 is coupled to the ***display device 5***, which functions to display the optical data 126 as visible text or graphics." (EX1005, 4:50-56.)

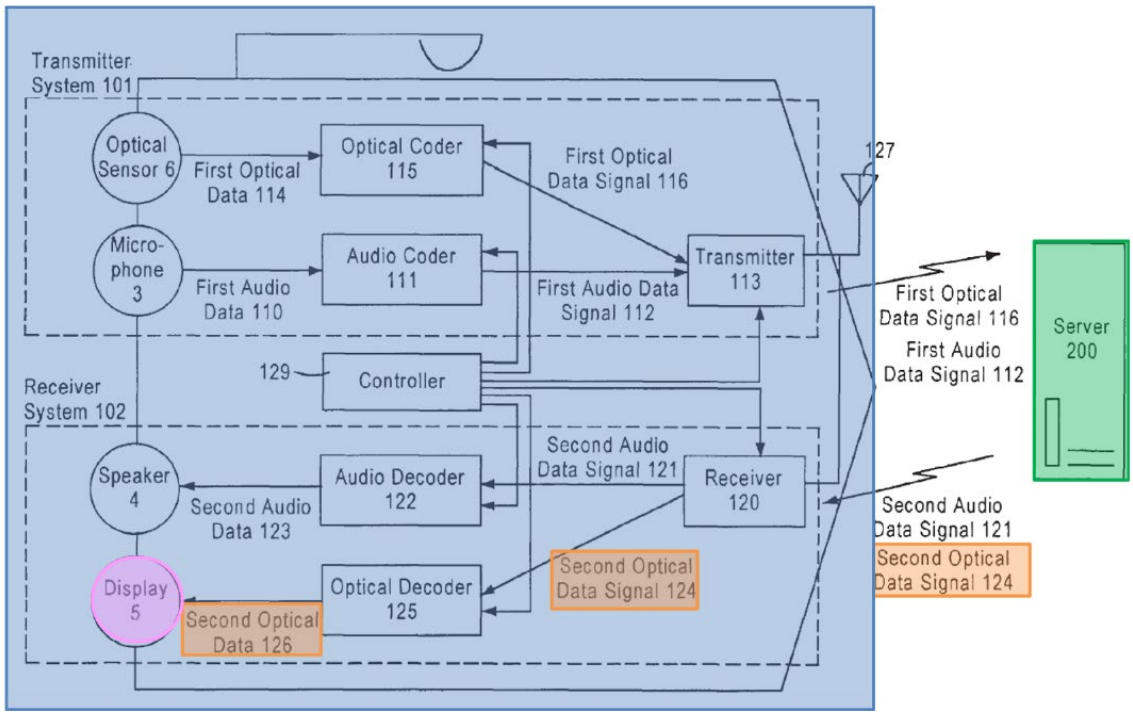


FIG. 2

402. Howard discloses several exemplary use cases of operating the display in response to the received ***first message***. For example, "upon receiving reply 484, the operator speaks the word "baseball" into the microphone 3. This data is encoded and sent to the server 200, 230, where it is determined to correspond to command 486. Upon receipt of the command 486, the program associated with guide 481 functions to generate ***reply 487*** (task 485), which is an ***audio answer stating the definition of baseball***. It is again noted that the answer

can also be provided in a visual format utilizing the ***display device*** contained in the ***portable device 10***." (EX1005, 14:40-49.)

403. As discussed, operating a display includes operating an actuator. (*Supra* §IX.A.2.k.) At the time of the invention, a POSITA would have understood that a display for presentation of visual data or information can be an actuator. For example, a display may be a liquid crystal display (LCD), when electric current passes through the LCD, the liquid crystals align and operate like a shutter to block out light, thereby producing images on the screen. LCDs make use of the following principles: "[t]hey depend on a liquid crystal material that polarizes light when an electrical potential is applied. The display is made by sandwiching this liquid crystal material between transparent metal electrodes deposited on glass substrates (Figure 15.6a), one or both of which are transparent and polarizing. Light passing through it will be polarized by the glass polarizing filter and so the light can be switched on or off by altering the voltage across the liquid crystal, which changes its polarization. The amount of power that is needed is very small, and the brightness of the display is just the brightness of the light that is used." (EX1017, 175.) Thus, a POSITA would have understood that Howard discloses operating an additional actuator (e.g., ***display device 5***) in the client device (e.g., ***portable device 10***) in response to the received first message (e.g., ***reply 404, optical data signals 124***).

404. Indeed, the specification of the '710 patent itself, states multiple times that a display has 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, 40:51-41-7, *see also, id*., 109:41-46, 110:60-66.)

l. 1[H]: "capturing, by the microphone in the client device, a second human voice data;"

405. Howard discloses capturing of a ***second human voice data*** by the microphone in the client device. To the extent that the recited "second human voice data" is different from the recited "first human voice data", Howard is clear that it can receive and process multiple human voice commands. As shown in FIG. 9, Howard discloses receiving different user utterances. (EX1005, 11:20-23; *see also id.*, 11:50-53.) The first two voice commands ***("TV", "Channel", "Volume"***) disclose the "first human voice data." The last two voice commands

(***"3", "A little small"***) disclose the "second human voice data." (*See also*

(EX1005, 10:64-11:5; 11:58-65; 12:7-14.)

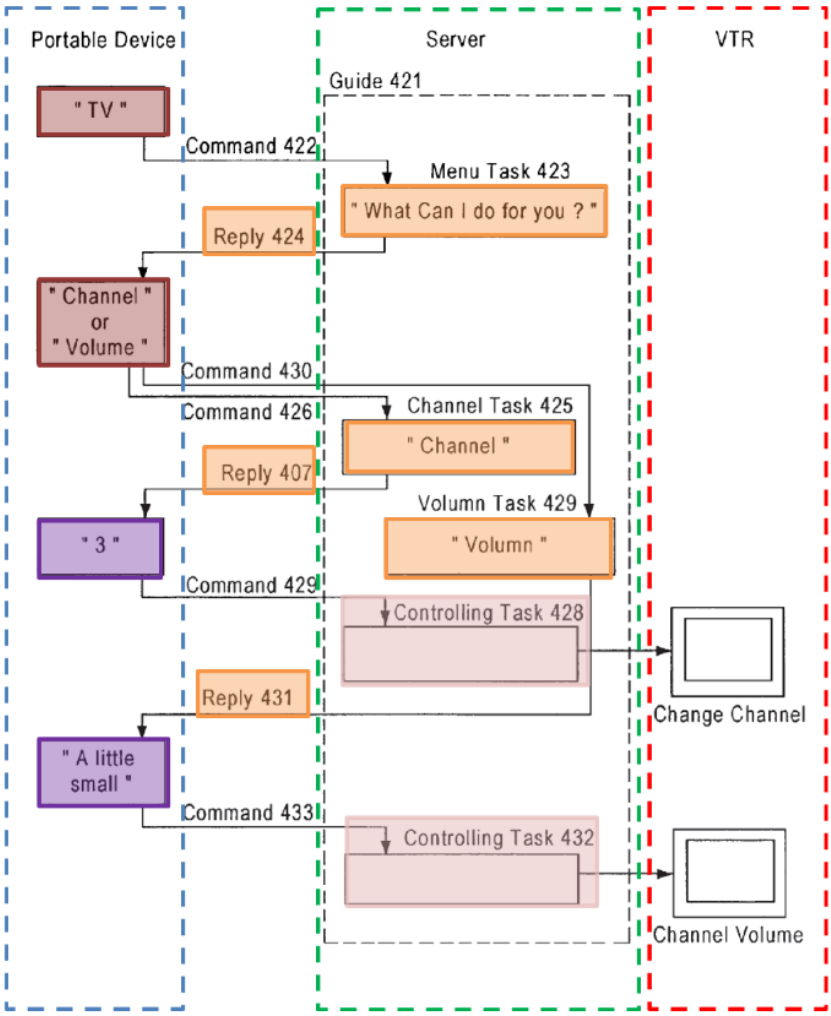


FIG. 9

406. Similarly, as shown in FIG. 8, Howard discloses receiving different user utterances. The first two voice commands (***"VTR", "Record"***) disclose the "first human voice data." The last voice command (***"9" "11:30~" "Oct 2"***) disclose the "second human voice data."

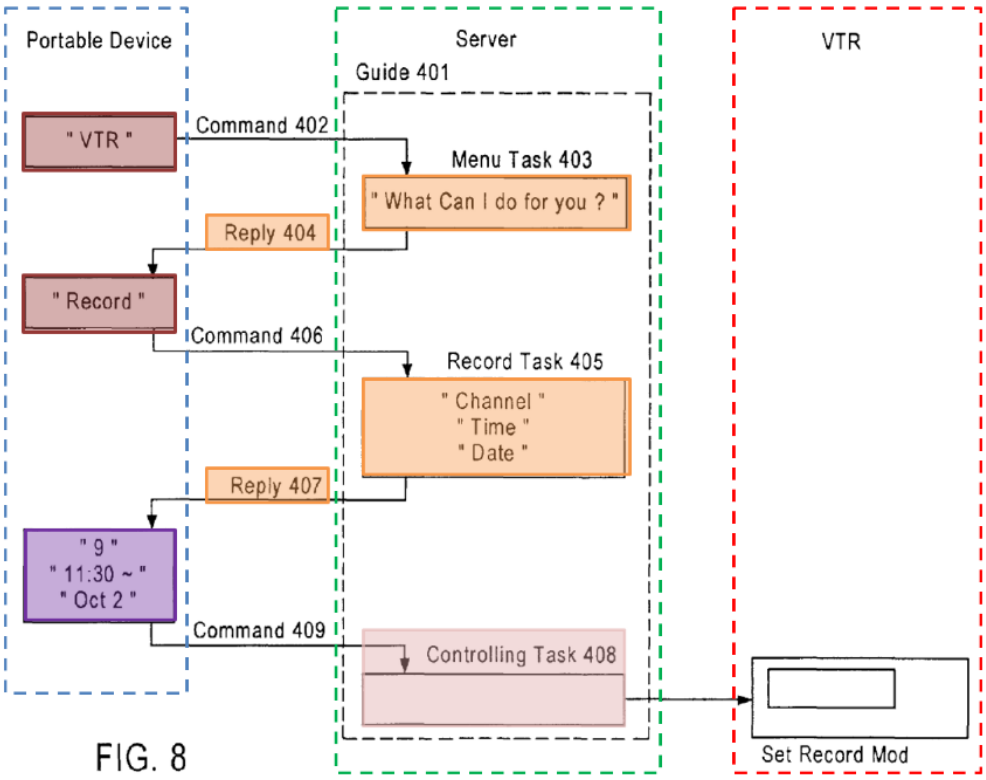


FIG. 8

407. Any of the human voice commands disclosed in Howard that cause an actuator in an in-home device in the home network to operate discloses the recited "second human voice data." (*Infra* IX.I.3.r.)

m. 1[I]: "sending to the server, by the client device via the wireless network, the captured second human voice data;"

408. As discussed, the Howard-Doherty system discloses this limitation. The Howard-Doherty system discloses that the client device sends the captured human voice data to the server via the wireless network. (*Supra* § IX.I.3.f). In addition, as discussed above, Howards discloses sending, receiving, and processing different utterances and human voice commands. (*Supra* § IX.I.3.l). Thus, Howard-Doherty discloses sending to the server, by the client device wirelessly

through a router, the captured second human voice data. It would have been obvious that Howard itself discloses this limitation.

n. 1[J]: "receiving, by the server over the Internet, the captured second human voice data;"

409. As discussed, the Howard-Doherty system discloses this limitation. As discussed above, the Howard-Doherty system discloses that the server receives the captured human voice data over the Internet via the router. (*Supra* § IX.I.3.g). In addition, as discussed above, Howards discloses sending, receiving, and processing different utterances and human voice commands. (*Supra* § IX.I.3.l). Thus, Howard-Doherty discloses receiving by the server over the Internet, the captured second human voice data. It would have been obvious that Howard itself discloses or at least suggests this limitation.

o. 1[K]: "processing, by the server, the captured second human voice data;"

410. As discussed, the Howard-Doherty system discloses this limitation. As discussed above, Howard discloses that the server processes the captured human voice data. (*Supra* § IX.I.3.f). In addition, as discussed above, Howards discloses sending, receiving, and processing different utterances and human voice commands. (*Supra* § IX.I.3.l). Thus, it would have been obvious that Howard itself discloses this limitation of processing, by the server, the captured second human voice data.

p. 1[L]: "responsive to the processing, sending a second message, by the server to the controlled device over the Internet;"

411. Howard discloses responsive to the processing, sending a ***second message***, by the ***server*** to the ***controlled device*** over the Internet.

412. FIG. 9 depicts that the server comprises "activity guide stored in memory A 301 associated with controlling/programming a TV." (EX1005, 11:14-16.) "Upon receipt of command 429, the program associated with guide 421 [of the ***server***] functions to generate ***control signals (task 428)***, which are transmitted to the ***TV*** via the home network 304 and utilized to program the TV to the desired channel." (EX1005, 11:58-65.) Further, "the program associated with guide 421 functions to generate ***control signals (task 432***), which are transmitted to the ***TV*** via the home network 304 and utilized to adjust the volume of the ***TV***." (EX1005, 11:61-65, 12:11-14; *see also* FIG. 8.) Howard discloses that the ***server*** functions to generate ***control signals (task 428)*** responsive to processing of the command 429 specifying the TV volume. Thus, the sending of the control signals is responsive to the processing of the commands by the server.

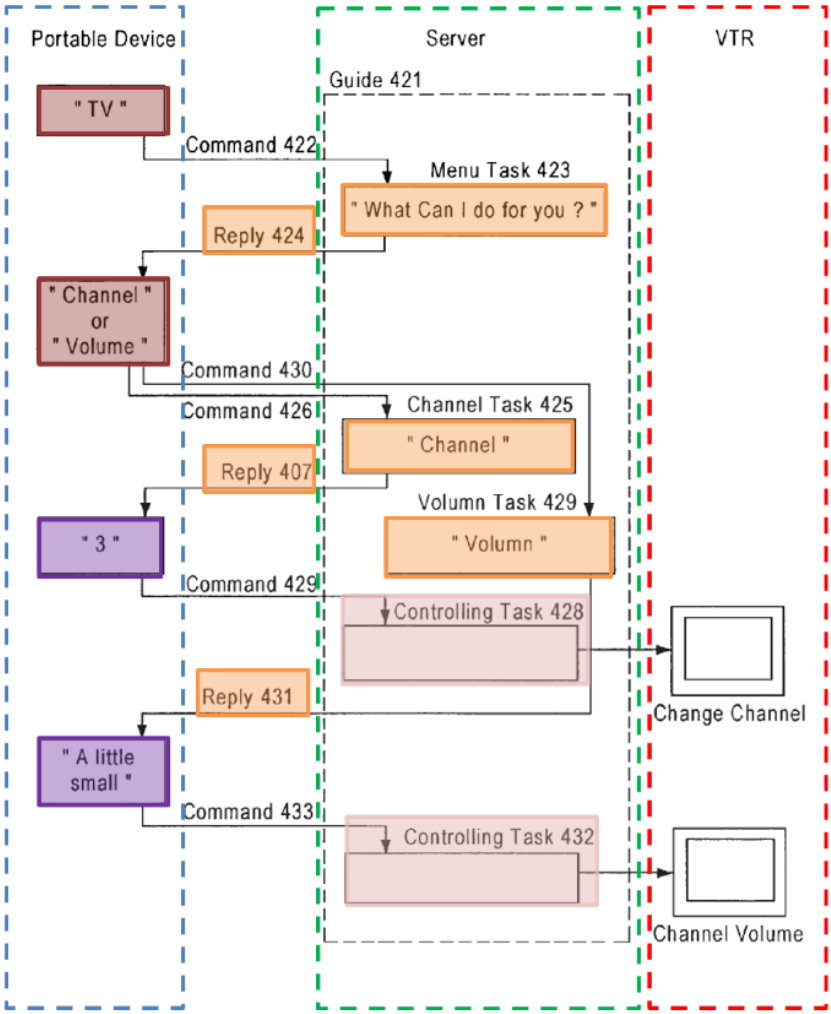


FIG. 9

413. With reference to FIG. 6, "the memory device 205 of the ***server 200*** … is divided into at least 3 areas, namely, memory A 301, memory B 302, and memory C 303. In the given embodiment, memory A 301 functions to store the activity guides related to ***controlling electrical components*** coupled to the home network 304." (EX1005, 8:42-47.)

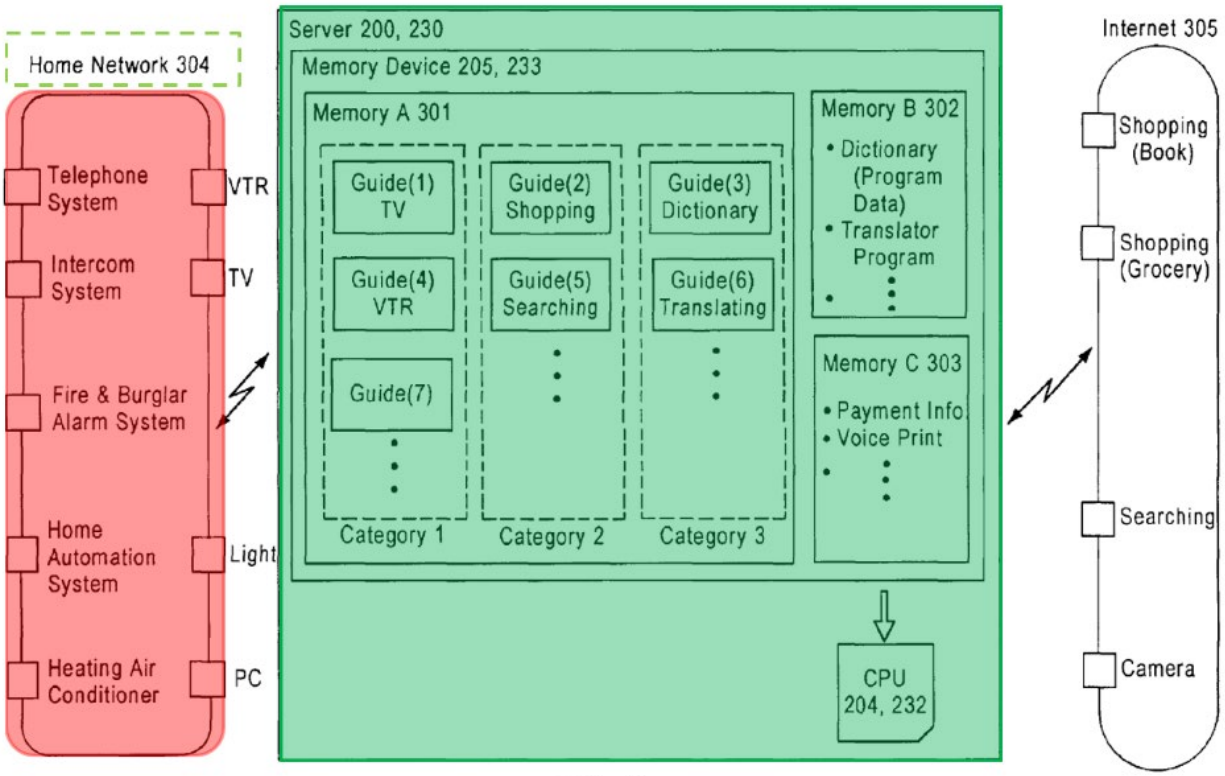


FIG. 6

414. Thus, Howard discloses an activity guide stored in memory A 301 located in the memory device 205 of the ***server 200*** generating ***control signals*** ("second message") for the server to send to the ***TV*** ("controlled device") in response to processing.

415. Second, Howard discloses that the ***server*** sends a ***second message*** to the ***controlled device*** over the Internet. As shown in FIG. 6, and disclosed by Howard with reference to FIG. 8 and FIG. 9, ***control signals*** are sent by the server to the controlled device via the home network 304. As discussed above, Howard discloses that a home network may take the form of "a more complex hub and

switching system with broadband connection to the ***Internet***." (*Supra* § IX.I.3.c; EX1005, 1:18-22.) And as discussed above, server 200 is connected to the Internet and home network 304. (*Supra* § IX.I.3.c - IX.I.3.d.) Thus, Howard discloses or at least suggests an implementation in which the ***server*** sends a ***second message*** to the ***controlled device*** over the Internet.

416. The Howard-Doherty system discloses that the server in the cloud sends data (e.g., messages) to the wireless router via the Internet, which routes the data wirelessly to the portable device or the home devices. (*Supra* § IX.I.3.c-IX.I.3.d; EX1006, 11, 38-39, 96, 100, 228-229, 261.)

417. Further, it would have been obvious to a POSITA that Howard itself discloses an implementation in which the ***server*** sends a ***second message*** to the ***controlled device*** over the Internet and the ***controlled device*** receives it via a wireless network. (*Supra* § IX.I.3.d.)

q. 1[M]: "receiving, by the controlled device via the wireless network, the second message; and"

418. Howard discloses receiving, by the ***controlled device*** via the wireless network, the ***second message***.

419. First, Howard discloses that the ***controlled device*** receives the ***second message***. The ***control signals*** depicted in FIG. 8 and FIG. 9 are transmitted to and received by the ***VCR*** and ***TV*** to be "utilized to program the VCR," "utilized to program the TV to the desired channel," or "utilized to adjust the volume of the

TV." (EX1005, 11:1-5, 11:61-65, 12:11-14.) Howard discloses that, "[u]pon receipt of command 429, the program associated with guide 421 functions to generate ***control signals*** (task 428), which are ***transmitted to the TV*** via the home network 304 and utilized to program the TV to the desired channel." (EX1005, 11:58-65.) As another example, "the program associated with guide 421 functions to generate ***control signals (task 432***), which are ***transmitted to the TV*** via the home network 304 and utilized to adjust the volume of the ***TV***." (EX1005, 11:61-65, 12:11-14.)

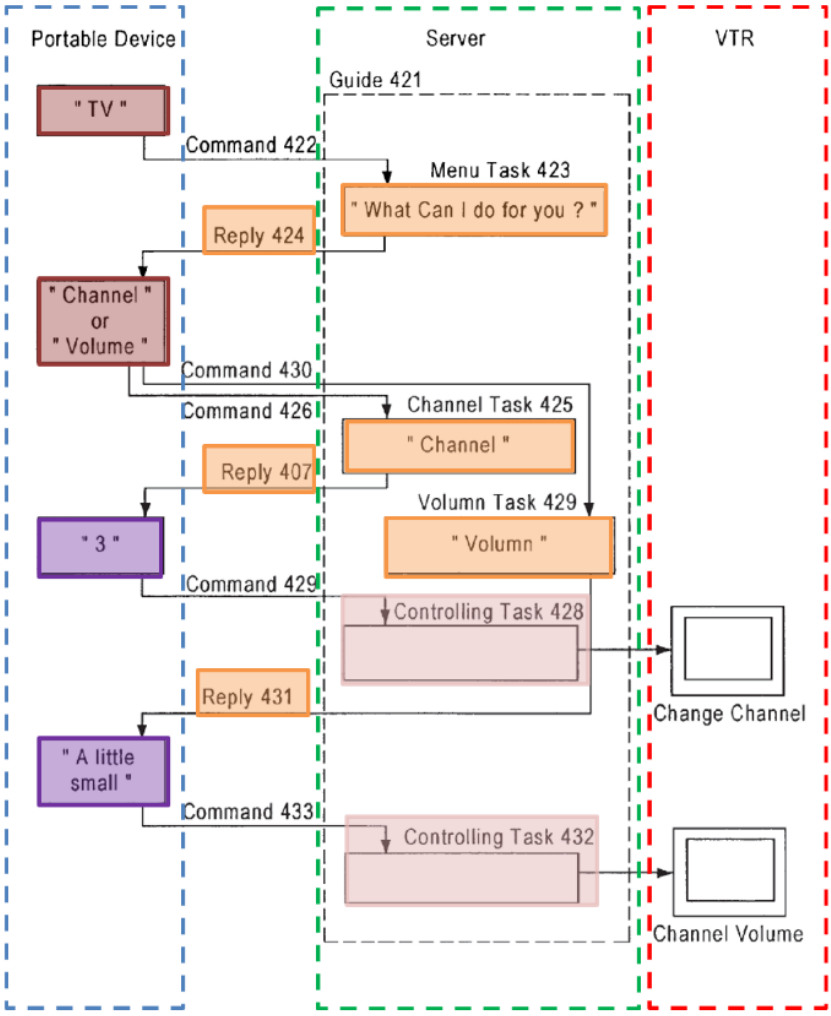


FIG. 9

420. Further, "the program associated with guide 401 functions to execute task 408, which generates ***control signals***, which are ***transmitted to the VCR*** via the home network 304 and are utilized to program the ***VCR***." (EX1005, 11:1-5)

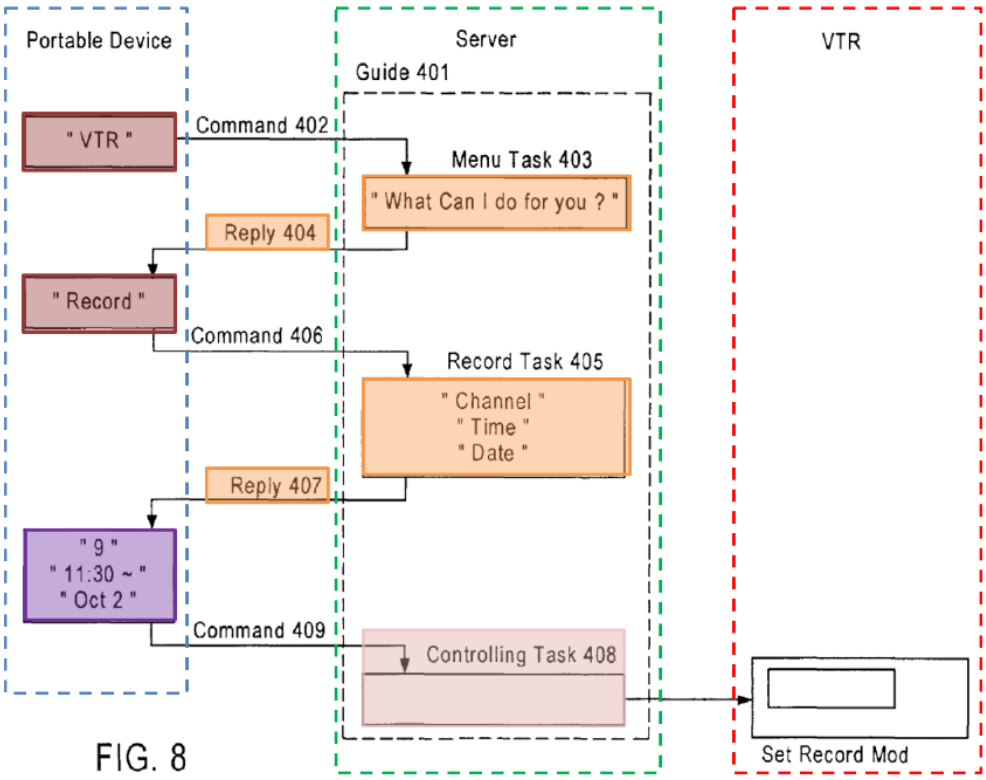


FIG. 8

421. Thus, the ***control signals*** so depicted in FIG. 8 and FIG. 9 are transmitted to and received by the ***VCR*** and ***TV*** to be "utilized to program the VCR," "utilized to program the TV to the desired channel," or "utilized to adjust the volume of the TV." (EX1005, 11:1-5, 11:61-65, 12:11-14.)

422. Howard discloses that the ***controlled device*** receives the ***second message*** via wireless network. As shown in FIG. 6, and disclosed by Howard with reference to FIG. 8 and FIG. 9, ***control signals*** are sent by the server to the controlled device via the home network 304. Further, as discussed above, Howard teaches that the control signals utilized to control the components coupled to the home network (*the controlled device*) can be packetized and transmitted using

wireless personal area networks. (*Supra* § IX.I.3.c). In the Howard-Doherty system, the server in the cloud sends data via the Internet to the wireless router, which routes the data wirelessly to the portable device or the home devices, the home devices thus receive the data wirelessly through the wireless router. (*Supra* §IX.I.3.c- IX.I.3.d; EX1006, 11, 38-39, 96, 100, 228-229, 261.)

423. Further, it would have been obvious to a POSITA that Howard itself discloses or at least suggests an implementation in which the ***server*** sends a ***second message*** to the ***controlled device*** over the Internet and the ***controlled device*** receives it via a wireless network. (*Supra* § IX.I.3.d.)

r. 1[N]: "operating a second actuator in the controlled device in response to the received second message."

424. Howard discloses operating a second actuator in the ***controlled device*** in response to the received ***second message***. The ***control signals*** depicted in FIG. 8 and FIG. 9 are transmitted to and received by the ***VCR*** and ***TV*** to be "utilized to program the VCR," "utilized to program the TV to the desired channel," or "utilized to adjust the volume of the TV." (EX1005, 11:1-5, 11:61-65, 12:11-14.)

425. First, Howard discloses, with reference to FIG. 9, that an actuator in the controlled device is operated in response to the received second message: "Upon being queried as to the desired volume, assuming the operator states the desired volume change (e.g., higher or lower), this audio data is encoded and sent to the server 200, 230, where it is determined to correspond to command 433.

Upon receipt of command 433, the program associated with guide 421 functions to generate ***control signals (task 432***), which are transmitted to the ***TV*** via the home network 304 and utilized to ***adjust the volume of the TV***." (EX1005, 11:61-65, 12:11-14.)

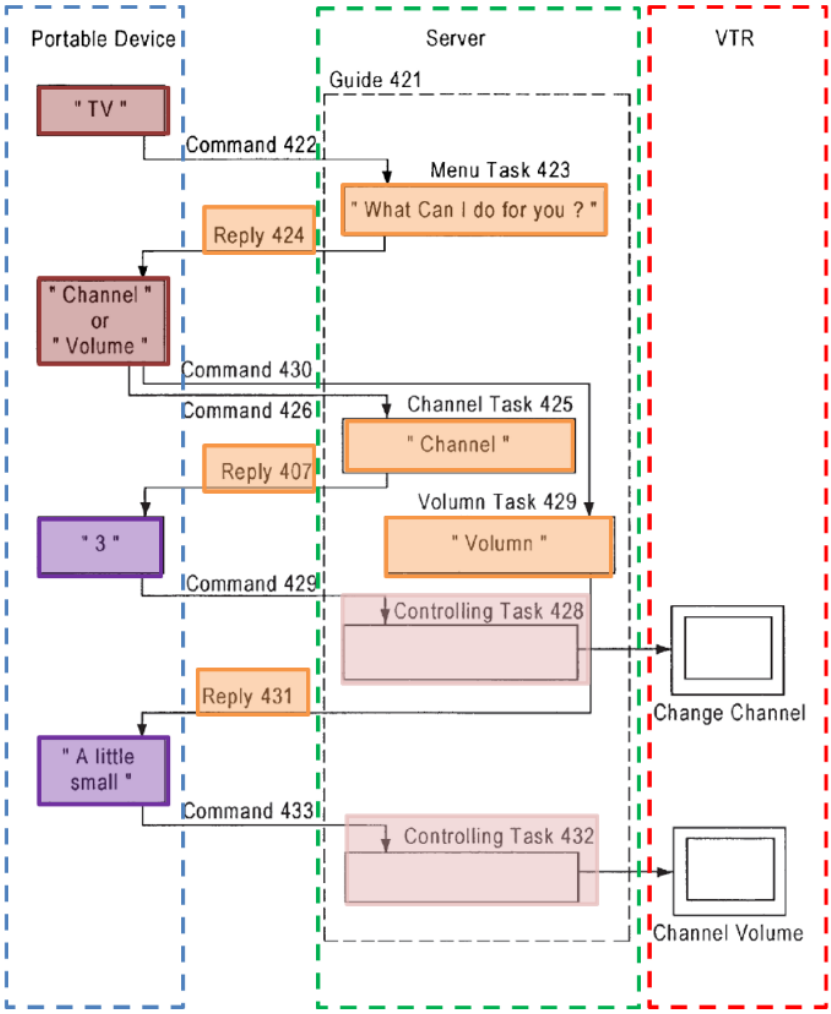


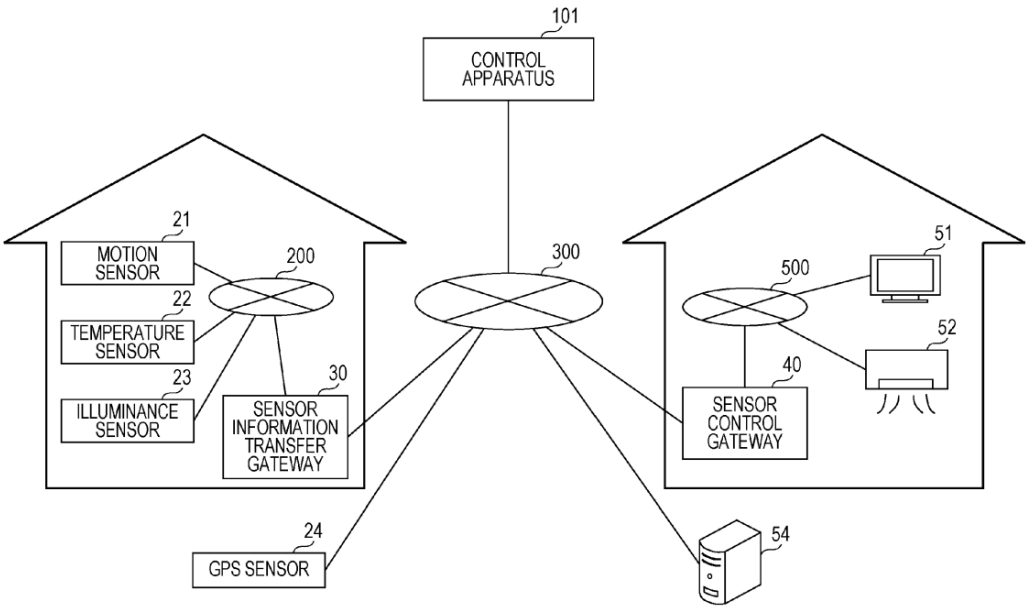
FIG. 9

426. A POSITA would have understood that adjusting the volume of the TV and changing channels involve operating one or more ***displays*** and ***speakers*** (i.e., actuators) in the TV. Thus, by programming or controlling the TV to adjust

the volume, the actuators in the form of speakers in the TV are operated to output sound at different levels of intensity. (*Supra* § IX.A.2.k.) It would have been obvious that the display and speaker of the ***TV*** is operated in response to the ***control signals*** (second message) received from the ***server***. (EX1005, 11:61-65, 12:11-14, FIG. 9.)

427. Indeed, as discussed, the Board agreed in its Decision on Appeal affirming the Examiner's rejection of claims 1-30 of the '710 patent that an image display device such as a TV 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." (EX1003, 396.)

FIG. 15



428. Howard further discloses with reference to FIG. 8, an example application of the system in controlling/programming a VCR. "Upon receipt of command 409, the program associated with guide 401 functions to execute task 408, which generates ***control signals***, which are transmitted to the ***VCR*** via the home network 304 and are utilized to ***program the VCR***." (EX1005, 10:64-11:5.) FIG. 8 labels as an action "Set Record Mod."

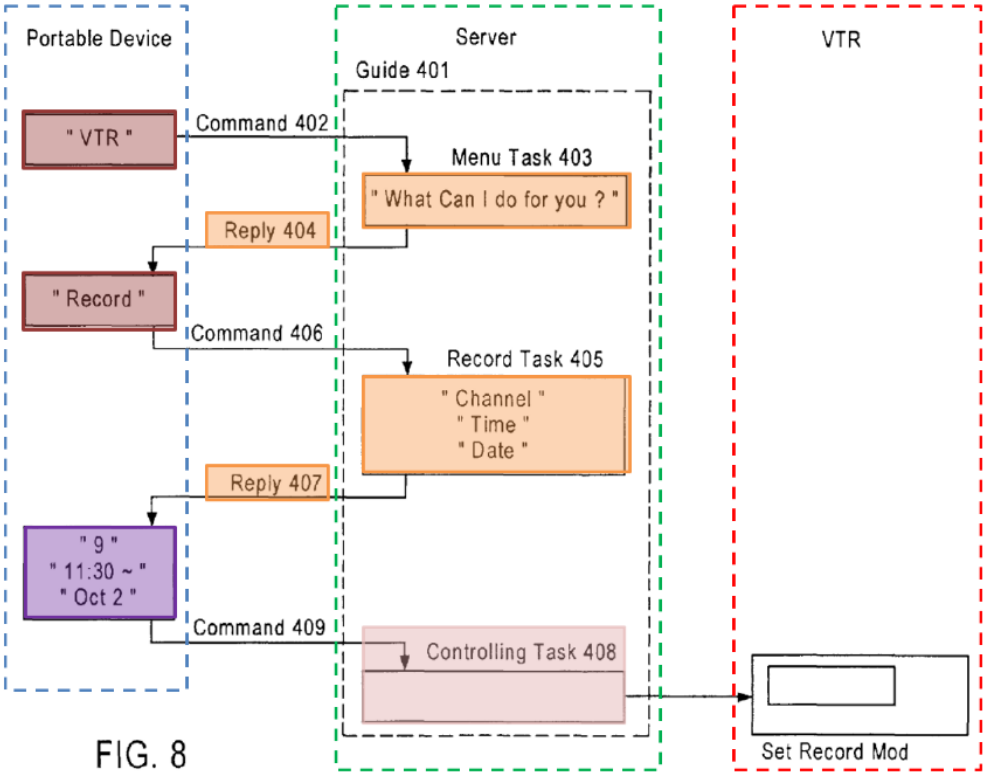


FIG. 8

429. Further, recording media content using an VCR involves operating actuators. For example, VCRs contain actuators that spins the recording/reading head to precisely read the tracks on the tape for 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." (EX1017, 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." (EX1017, 226-227.) Thus, operating the VCR to record in response to control signals discloses this limitation.

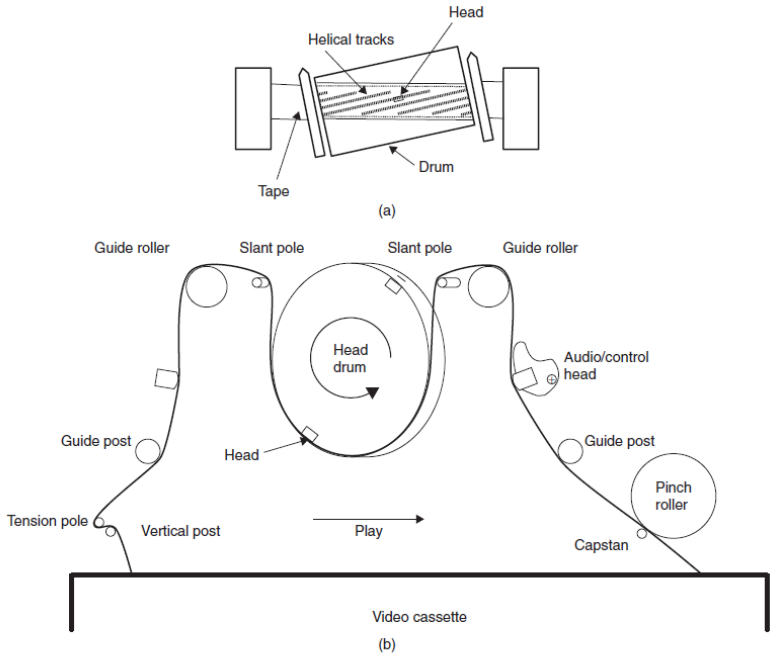


Figure 19.6 Tape transport of typical helical scan video recorder: (a) tracks on tape, and (b) tape path

430. Howard discloses a variety to actuators in controlled devices that are operated in response to messages from the server. For example, "guides containing tasks related to controlling apparatus connected on the home network 304. Such apparatus include, but are not limited to, (a) ***audio/video equipment, for instance TVs, VCRs, DVDs, CD players, (b) electrical appliances, such as light systems, refrigerators, microwave ovens, heating and cooling systems, (c) computers, (d) home automation systems, (e) fire alarm systems, (f) burglar alarm systems, (g) telephone systems, and (h) intercom systems***." (EX1005, 9:14-23.) A POSITA would have understood that controlling these apparatuses and systems involves operating actuators. For example, CD and DVD drives use motors and 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." (EX1017, 280, Figure 23.7.)

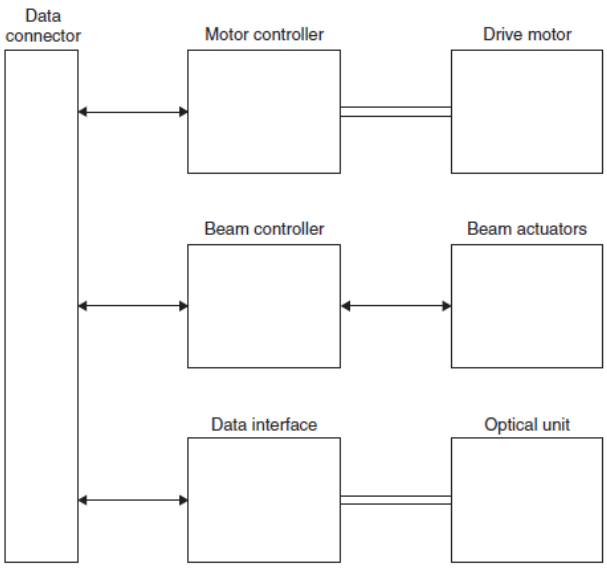


Figure 23.7 Block diagram, simplified, for a CD-ROM drive

431. For example, HVAC systems operate actuators. Heating and cooling systems include actuators for regulating airflow and controlling temperature. 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 function." (EX1019, 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." (*Id.*)

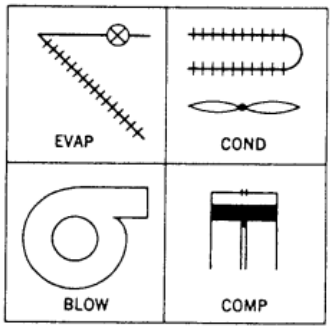


Fig 13-1. Self contained air conditioner

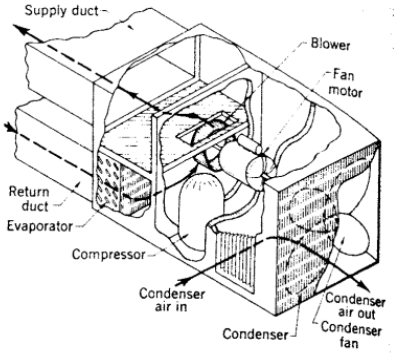


Fig 13-3. Horizontal self contained air conditioned unit

432. Further, as discussed above, the Board agreed in its Decision on Appeal affirming the Examiner's rejection of Claims 1-30 that an air conditioner is a controlled device with actuators. (EX1003, 396.)

4. Claim 5: The method according to claim 1, wherein each of the first and second actuators is directly or indirectly affecting, changing, producing, or creating a physical phenomenon.

433. The Howard-Doherty system discloses this limitation. As discussed, Howard discloses operating a speaker (i.e., a first actuator) in the client device in response to the received first message. (*Supra* § IX.I.3.k; EX1005, 11:27-47, Claims 2 and 9.) As discussed, operating the speaker involves operating an actuator affecting, changing, producing, or creating sound or audio. (*Supra* § IX.I.3.k.)

434. As discussed, Howard discloses operating a ***display screen*** (i.e., a first actuator) in the ***client device***, in response to the received first message. (*Supra* § IX.I.3.k; EX1005, 10:23-34.) A display screen is another example of the recited "first actuator", an actuator providing visual information. (*Supra* § IX.I.3.k.)

Operating the display screen involves operating an actuator affecting, changing, or producing light. (*Id*.)

435. Second, the controlled device disclosed in the Howard-Doherty system includes audio/video equipment such as DVDs, CD players, VCRs, electrical appliances such as light systems, heating and cooling systems, and home automation systems (*Supra* § IX.I.3.r; EX1005, 9:14-23.) As discussed, controlling audio/video equipment like TVs, DVDs, CD players (e.g., to play a media item) involves operating sound actuators in media devices to affect, change, produce, or create sound. (*Supra* IX.A.2.k, IX.I.3.r.) Further, controlling media devices 106 (e.g., to playback a media item) involves operating actuators in media devices (e.g., displays) to provide visual information, thus affecting, changing, producing, or creating light. (*Id.*) Further, controlling home automation and security devices involves operating motion actuators, thus affecting, changing, producing, or creating motion, temperature, humidity, etc. (*Id.*)

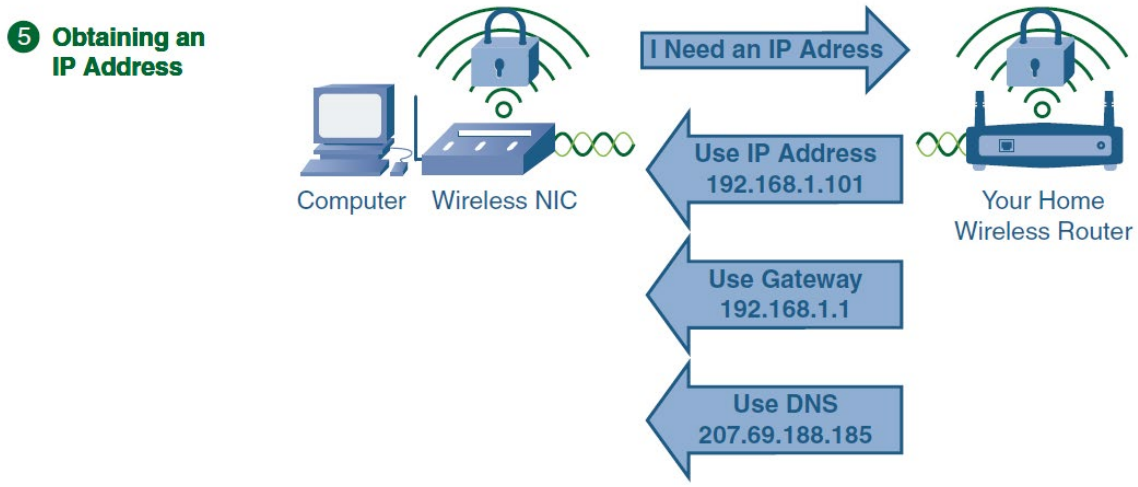
5. Claim 6: The method according to claim 5, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current.

436. Howard-Doherty discloses this limitation. As discussed above, Howard discloses a first actuator and a second actuator that directly or indirectly affect, change, product, or create at least audio, light, motion, sound, temperature, humidity, etc. (*Supra* § IX.I.4.)

6. Claim 10: "The method according to claim 1, wherein the client device or the controlled device are addressable in the wireless network or the Internet using an address stored in a volatile or non-volatile memory of the respective device for uniquely identifying the respective device in the network."

437. As discussed above, the Howard-Doherty system discloses that the portable device and in home devices each communicate wirelessly with a router that routes the data (i.e., audio) to the remote server over the Internet. (*Supra* § IX.I.3.c, IX.I.3.d; EX1006, 11, 38-39, 96, 100, 228-229, 261.) As discussed, a POSITA would have understood that this involves associating each device with a device identifier or address (which is locally stored) that uniquely identifies the device in the network. Indeed, at the time of the invention, it was well-known that networked devices are associated with IP addresses and/or MAC addresses. "***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 your home network through a wireless signal; [s]ecuring the communications path between the wireless router and your computers and devices connected over wireless . . . " and "***each wireless card has a unique identifier called a MAC address***." (EX1006, 99, 166.) Each networked device is thus assigned with a globally unique MAC address.

438. "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." (*Id*., 108, 210, *see also*, *Id*., 175.) Thus, the networked devices are assigned with IP addresses that are unique to the device in the wireless network for identification. A POSITA would have understood that the associated 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.



439. Further, as shown in the figure below, Doherty discloses that the networked devices can display the IP addresses, which means the IP addresses are locally stored. (*Id.*, 211, Figure 10-44, *see also, Id*., 265.)

Figure 10-44 Verifying IP Address Assignment with Windows XP



7. Claim 11: "The method according to claim 10, wherein the address is a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type or wherein the address is a layer 3 address and is static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address."

440. As discussed above, the Howard-Doherty system discloses that the portable device and in home devices each communicate wirelessly with a router that routes the data (i.e., audio) to the remote server over the Internet. (*Supra* § IX.I.3.c, IX.I.3.d; EX1006, 11, 38-39, 96, 100, 228-229, 261.) As discussed above, Doherty discloses and it was well-known that each of the networked devices are associated with a globally unique MAC address and an IP address that uniquely identify the devices in the wireless network. (*Supra* § IX.I.6.) "***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***." (EX1006, 99, 166.) "***[E]ach 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." (*Id*., 210)

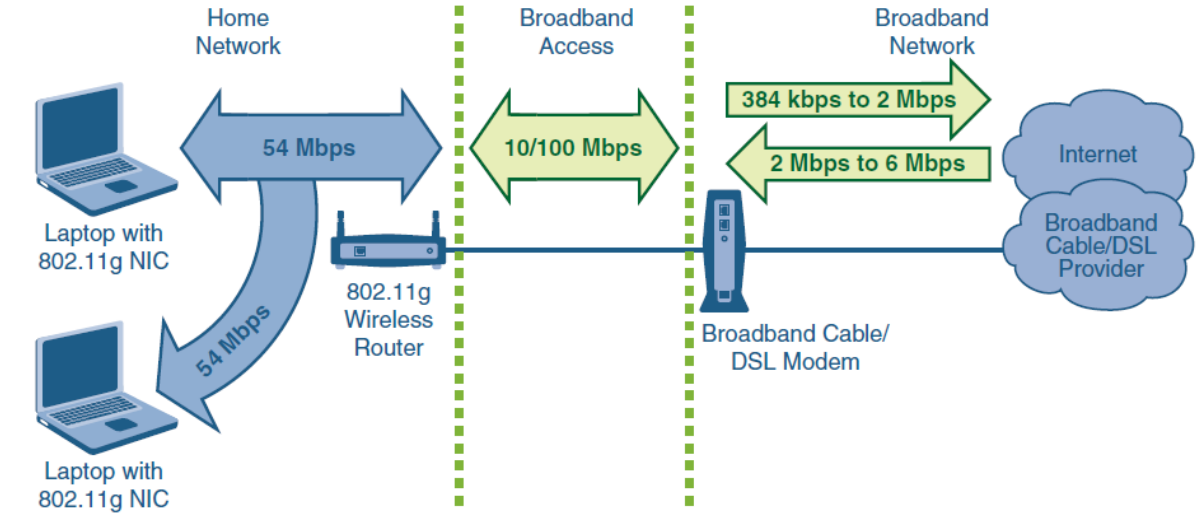
441. In 2012, it was well-known that Internet Protocol (IP) addresses are layer 3 addresses, static or dynamic, and are IPv4 addresses. "In the case of today's Internet, these addresses are Internet protocol version 4 (IPv4) addresses." (EX1028, 127.) "Internet protocol (IP) addresses are the unique numbers used to identify individual connections to the Internet." (*Id*.) "Internet Protocol (IP) is a Layer 3 protocol that contains addressing and control information that allows packets to be routed." (EX1022, 11.) "The IP system in common use today is known as IPv4, for Internet Protocol version 4. A newer system, IPv6, or Internet Protocol version 6, has been developed and exists today in small deployments." (*Id*., 13.) Doherty discloses that "[t]he broadband networks typically assign a *dynamic IP address* to your broadband service. Dynamic means the IP address can change[.]" (EX1006, 277.) Doherty also discloses that the IP address may be a "*static IP address*, which means . . . your address never changes." (*Id*.) Further, at the time of the invention, MAC address based on MAC-48, EUI-48, and EUI-64 are well known. "Most MAC addresses use one of the IEEE standardized protocols; ***MAC-48, EUI-48®, and EUI-64®***." (EX1046, 2:50-59.)

8. Claim 13: "The method according to claim 1, wherein the wireless network is a Wireless Local Area Network (WLAN) that is according to, or base 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."

442. Howard discloses that both the portable device and the controlled device communicates over wireless network. "The portable device 10 also comprises a controller (not shown in FIG. 1(a)) which functions to control the operation of the foregoing components, and an ***antenna*** (also not shown in FIG. 1(a)) which functions to allow for ***wireless communication*** between the ***portable device 10*** and a ***server 200*** forming part of the home gateway. In one embodiment, the ***antenna*** is incorporated into the design of the clip 2." (EX1005, 3:41-48.) Further, Doherty discloses that devices communicate with the wireless router over a wireless LAN (802.11) protocol, also known as Wi-Fi. "802.11b was developed at the same time as the A standard. Because it was less expensive to develop, ***802.11b is the standard that most vendors first developed for home networking***. As a result, 802.11b is also the first standard that enjoyed widespread adoption and is ***still widely used in home networking*** and Internet hotspots (Internet cafés and public wireless locations) today." (EX1006, 18.) Doherty also discloses the "802.11g standard". "The 802.11g standard was developed in 2003 and borrowed from the best of the two previous standards (A and B) . . . This standard is also backward compatible with the B standard." (*Id.*, 19.) Doherty discloses that "a

NIC for the 802.11b standard will certainly function with a wireless router for the 802.11g standard . . . Similarly, a NIC for the 802.11g standard will work just fine with a wireless router for the 802.11b standard." (*Id*., 27.) For example, Doherty discloses with reference to Figure 12-1 "an 802.11g home wireless network" that connects to the Internet comprising a 802.11g wireless router and laptops with 802.11g NIC. (*Id*., 239, 241)

Figure 12-1 Networking Speeds Inside and Outside the Home Network



443. As discussed, it would have been obvious to combine the teachings of Howard-Doherty. (*Supra* § IX.I.3.c - IX.I.3.d.)

444. As discussed, the Howard-Doherty combination would result in a system in which the portable device sends data (e.g., voice data) wirelessly to the router which routes it to the server in the cloud via the Internet, thus, the server receives the voice data over the Internet via the router, and the server in the cloud

sends data (e.g., messages) to the router via the Internet, which routes the data wirelessly to the portable device or the home devices. (*Supra* § IX.I.3.c - IX.I.3.d.)

9. Claim 16: The method according to claim 1, wherein the controlled device is integrated in, is part of, or is entirely included in, an appliance.

445. Howard-Doherty discloses a controlled device that is integrated in, is part of, or is entirely included in, an appliance. As discussed, the controlled device in the Howard-Doherty system are various household components and appliances including "***entertainment network devices, such as TV, stereo, and video equipment*** like a ***VCR***, and various ***'life support systems', such as, heating systems, air conditioning systems and lighting and security systems***." (*Supra* § IX.I.3.p - IX.I.3.r; EX1005, 1:18-30, *see also, id*., 9:13-23.) These are all appliances.

10. Claim 17: The method according to claim 16, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation.

446. Howard-Doherty discloses this limitation. As discussed, the controlled device in the Howard-Doherty system are electrical appliance such as "**refrigerators**, **microwave ovens**". (*Supra* § IX.I.3.r; EX1005, 9:13-23, Figure 6.) These disclose appliances whose primary functionality is associated with food storage, handling, or preparation.

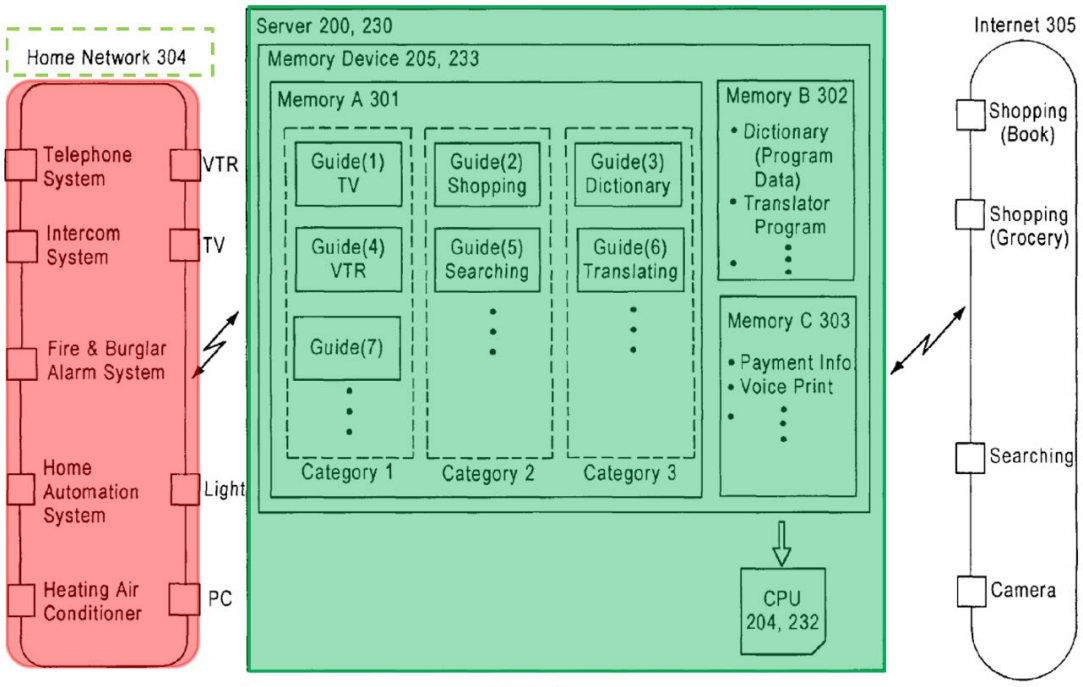


FIG. 6

11. Claim 18: The method according to claim 17, 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.

447. Howard-Doherty explicitly discloses this limitation. As discussed, the controlled device in the Howard-Doherty system are electrical appliances such as "***microwave ovens***". (*Supra* § IX.I.3.r; EX1005, 9:13-23, Figure 6.) This disclose appliances whose primary function is heating food.

12. Claim 19: The method according to claim 17, 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.

448. Howard-Doherty explicitly discloses this limitation. As discussed, the controlled device in the Howard-Doherty system are electrical appliances such as "***refrigerators***". (*Supra* § IX.I.3.r; EX1005, 9:13-23, Figure 6.)

13. Claim 20: The method according to claim 16, wherein the primary function of the appliance is associated with environmental control, and the appliance is part of an Heating, Ventilation and Air Conditioning (HVAC) system.

449. Howard-Doherty discloses this limitation. As discussed, the controlled device in the Howard-Doherty system are electrical appliance such as "***heating and cooling systems***". (*Supra* § IX.I.3.r; EX1005, 9:13-23, Figure 6, *see also, id.*, EX1005, 1:26-30.) These disclose appliances whose primary functionality is associated with environmental control.

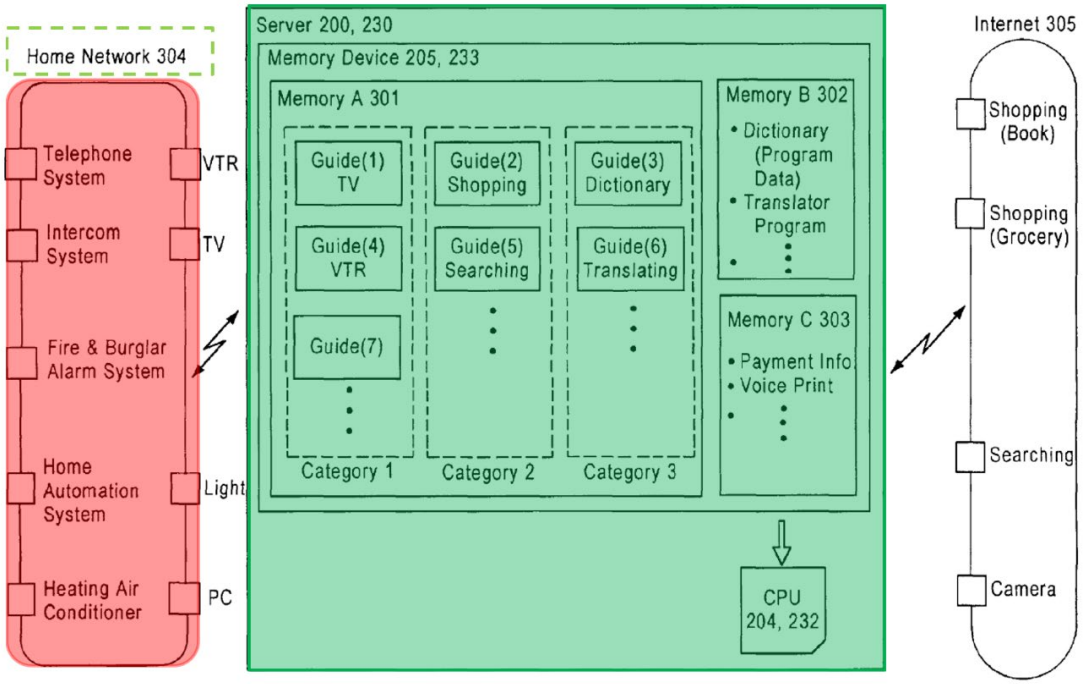


FIG. 6

14. Claim 21: The method according to claim 20, wherein the primary function of the appliance is associated with temperature control, and wherein the appliance is an air conditioner or a heater.

450. Howard-Doherty discloses this limitation. As discussed, the controlled device in the Howard-Doherty system are electrical appliance such as "***heating and cooling systems***". (*Supra* § IX.I.3.r; EX1005, 9:13-23, Figure 6, *see also, id.*, EX1005, 1:26-30.) These disclose appliances whose primary functionality is associated with temperature control.

451. Specifically, as shown in FIG. 6, "Heating Air Conditioner" can be connected on the Home Network 304 and be controlled through the Howard-Doherty system.

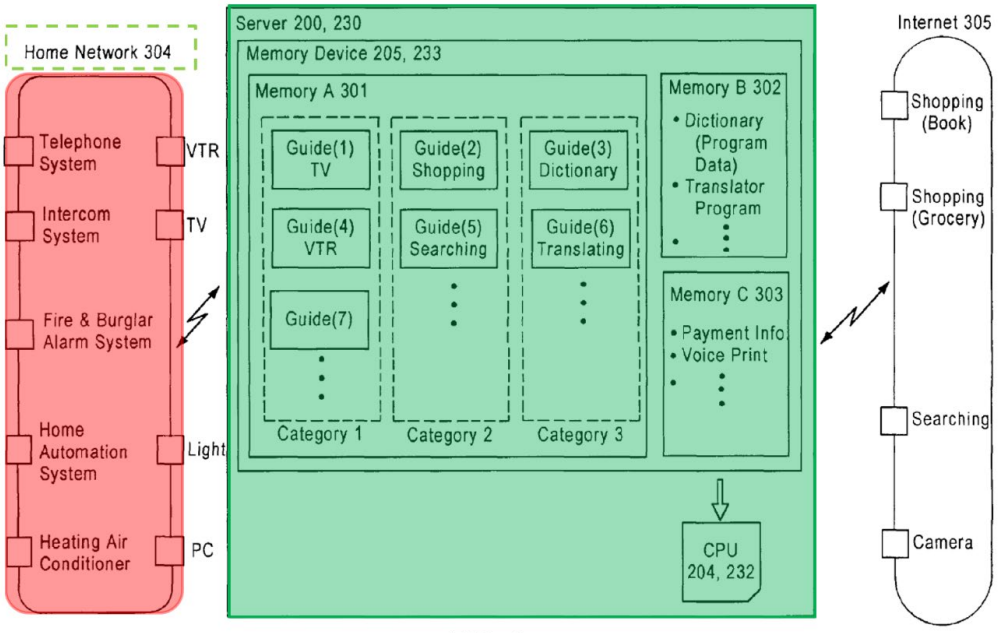


FIG. 6

15. Claim 23: The method according to claim 16, 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.

452. Howard-Doherty explicitly discloses this limitation. As discussed, Howard discloses remotely controlling "(a) audio/video equipment, for instance TVs, VCRs, DVDs, CD players, (b) electrical appliances, such as light systems, telephone systems." (*Supra* § IX.I.3.r; EX1005, 9:13-23, Figure 6, *see also, id.*, EX1005, 1:26-30.)

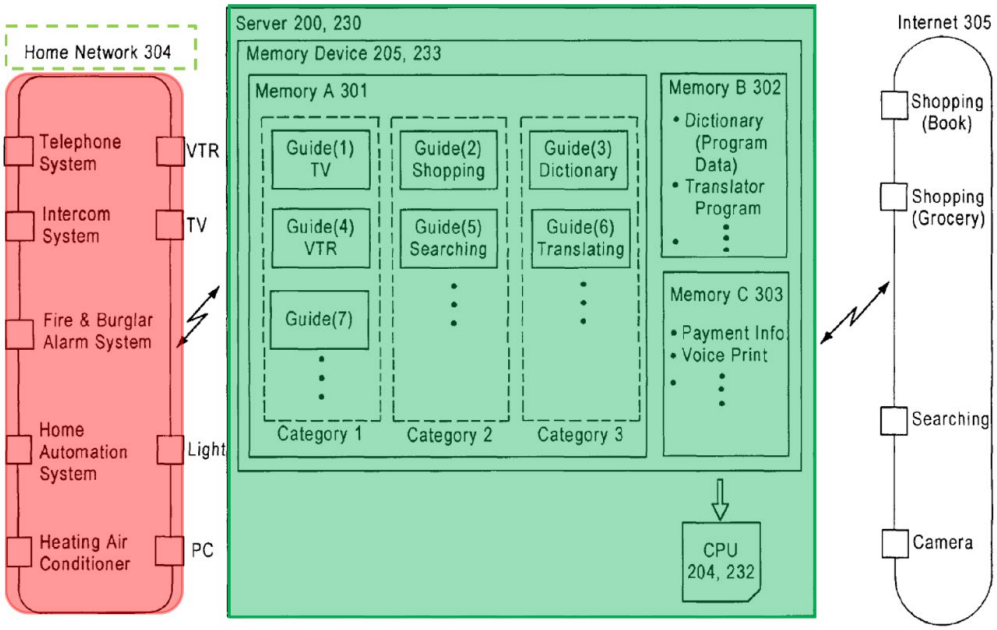


FIG. 6

16. Claim 29: The method according to claim 27, wherein the operating of the first actuator comprises playing digital audio content that is pre-recorded or synthesized, or wherein the operating of the first actuator comprises simulating the voice of a human being or generating music, or wherein the operating of the first actuator comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.

453. Howard-Doherty discloses this limitation. As discussed above, Howard discloses operating a ***speaker 4*** in the client device in response to the first message received from the server. (*Supra* § IX.I.3.k; EX1005, 11:20-23.) For example, Howard discloses, with reference to FIG. 9, in an exemplary application of controlling/programming a TV that "Upon receipt of command 422, the program associated with guide 421 functions to generate ***reply 424*** (task 423), which is an ***audio question requesting the operator to indicate what her/she would like done***. In the given example, the ***reply*** sent back to the operator states '***What can I do for you***'. The data representing reply 424 is generated by the CPU 204 and coupled to the data ***synthesizer*** 207 such that it can be encoded in a manner suitable for transmission to the ***portable device 10*** via the communication device 201. Upon receipt of the encoded reply data by the receiver 120 of the ***portable device 10***, the data is coupled to the audio decoder 122, and then coupled to the ***speaker 4*** so as to be reproduced for the user." (*Id.*)

454. The audio reproduced for the user at the ***speaker 4*** in the client device also involves "***reply 407***, which is an ***audio question requesting the operator to input the 'Channel, Time and Date' of the program to be recorded***," or "***reply 427*** (task 425), which is an ***audio question requesting the operator to indicate what channel should be selected***," or "***reply 431*** (task 429), which is an ***audio***

question requesting the operator to indicate if the volume should be lower or higher." (EX1005, 10:52-60; 11:50-58; 12:2-6.)

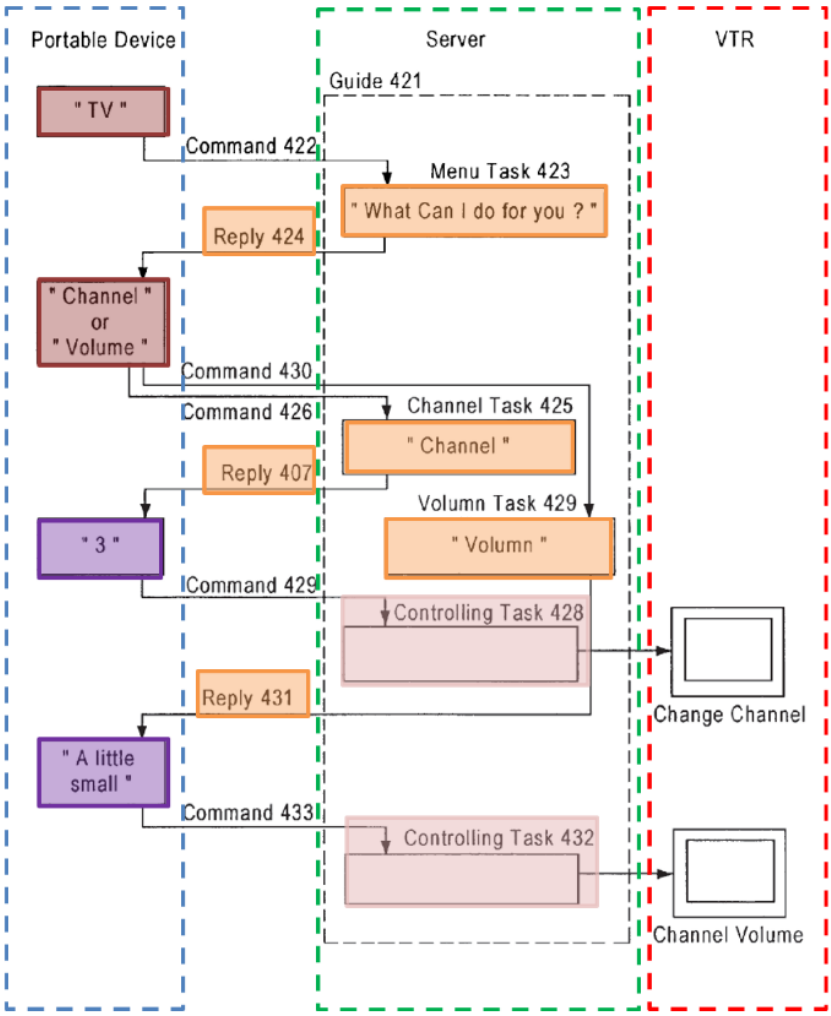


FIG. 9

455. Thus, Howard-Doherty discloses 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.

J. Ground 10: Howard, Doherty, and Wittenberg Render Obvious Claim 2

1. Claim 2: "The method according to claim 1, wherein the processing comprises performing a voice recognition algorithm for identifying the voice of a specific person."

456. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 2 is obvious in further view of Wittenberg. Howard also discloses performing a voice recognition algorithm for identifying the voice of a specific person. Howard discloses performing voice recognition during processing to identify an operator's voice print. "For example, the program executes a task 502 that generates a reply 503 as the message 'Please tell me your name'. Assuming the operator states his name into the microphone 3, ***voice print matching task 504*** is then executed so as to ***compare the voice print of the operator***, which is forwarded by the portable device 10, ***with the operator's voice print that was prestored in memory C 303*** in advance during an initialization process. If the voice prints match, the voice print matching task 504 places the order and sends payment information to the shopping website through the Internet connection and completes the purchase." (EX1005, 14:59-15:3; Figure 13(a).) Thus, Howard discloses performing voice recognition during processing to identify whether the operator's voice print matches prestored memory when payment information is involved.

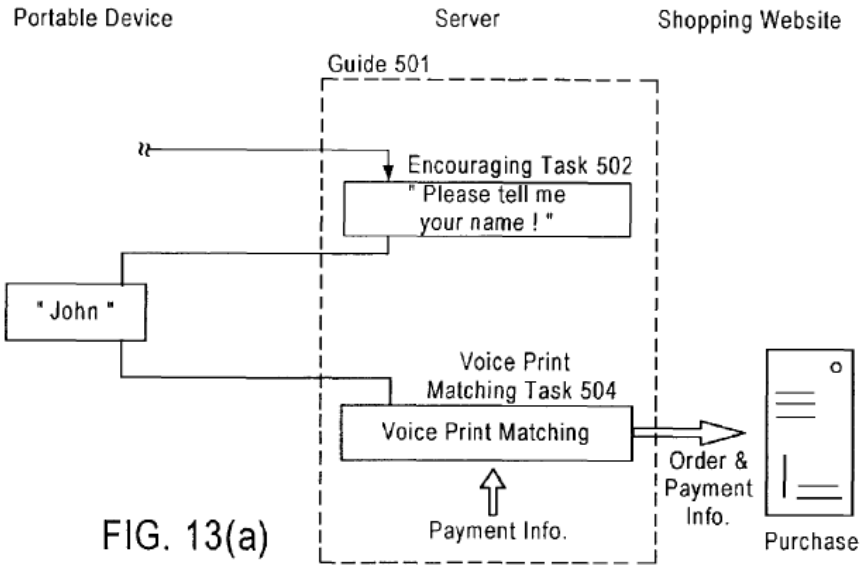


FIG. 13(a)

457. Further, Wittenberg discloses claim 2. (*Supra* § IX.B.2.)

458. It would also have been obvious to a POSITA to combine the teachings of Howard-Doherty and Wittenberg. Both the Howard-Doherty system and Wittenberg relate to processing audio information such as speech received from users for controlling devices. Further, Howard discloses a voice print matching task to identify the identity of the operator that gave the voice command. Thus, Howard-Doherty and Wittenberg are analogous art, from the same field of endeavor and relate to the same goal of controlling devices based on audio commands and providing security to the command/control system architecture.

459. As discussed, Wittenberg teaches additional processing techniques, first performing speech recognition processing and then performing voice recognition processing. (EX1007, Claim 1.) Wittenberg describes the security

issue that unauthorized users may be able to control devices via audio commands. (*Id.*, 2:21-30.) Wittenberg then expressly teaches the benefits for including voice recognition during processing to address this 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." (*Supra* § IX.B.2; EX1007, 5:61-6:6.)

460. Wittenberg recognizes the benefits of applying its techniques in numerous circumstances, from government/military, commercial, and private sectors. (*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." (EX1007, 6:7-13.)

461. The Howard-Doherty-Wittenberg combination results in a system in which Wittenberg's voice recognition capabilities are used to analyze human audio commands in the Howard-Doherty system, for example, to determine if the person issuing the voice data is authorized to make the command. Combining them as such would yield predictable results of enhancing 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 yield the predictable result of improving the security of the control and command systems and prevent control by unauthorized users.

462. Because both the Howard-Doherty system 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 to improve the security of the Howard-Doherty system in the same way.

463. Further, it would have been obvious to try combining the teachings of Howard-Doherty and Wittenberg. There are only a finite number of identified and predictable solutions for user identification, 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 processing techniques and algorithm to the processing of human voice data 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

Howard-Doherty and Wittenberg. Moreover, at the time of the invention, algorithms and processing techniques for performing voice recognition are known. "Algorithms and processing techniques for performing voice recognition are well known . . . ." (EX1007, 6:1-3.)

K. Ground 11: Howard, Doherty, and Gruber Render Obvious Claim 3-4

1. Claim 3: "The method according to claim 1, wherein the client device further comprises a sensor that outputs sensor data that responds to a physical phenomenon, wherein the method further comprising sending to the server, by the client device via the wireless network, the sensor data, and wherein the first message is sent by the server in response to the sensor data."

464. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 3 is obvious in further view of Gruber.

465. Gruber discloses claim 3. (*Supra* § IX.C.2.)

466. It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Gruber. The Howard-Doherty system and Gruber both relate to processing and responding to human speech commands. Howard discloses "a method and system that allows an operator to easily and readily control home activities without having to directly interface with or use a PC." (EX1005, 1:53-56.) Howard discloses that such control takes the form of audio commands. Howard explains an example application of the system to control/program a TV in FIG. 9. FIG. 9 illustrates that, first, the user ***speaks*** to a microphone in a ***portable***

device, then such audio is encoded and transmitted to the server for processing: "in operation, when the ***word 'TV'*** is input into microphone 3 in the ***portable device 10***, in the same manner as set forth above, the data is encoded by audio coder 111 and transmitted to the ***server 200*** via the transmitter 113." (EX1005, 11:20-23.) 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." (EX1008, [0007].) Thus, the Howard-Doherty system and Gruber 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).

467. Gruber teaches additional speech processing techniques of using contextual information obtain from device sensors to clarify the user's intent, reduce numbers of candidate interpretations and constrain solutions during speech recognition and natural language processing. "A virtual assistant uses context information to supplement natural language or gestural input from a user. Context 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. Context can be used

to constrain solutions during various phases of processing, including, for example, speech recognition, natural language processing, task flow processing, and dialog generation." (EX1008, Abstract.) The technique of using sensor data as contextual information can be combined with the teachings of the Howard-Doherty system to improve the accuracy of the speech processing, with a reasonable expectation of success.

468. Specifically, the Howard-Doherty-Gruber combination would result in a system that can capture sensor data (e.g., temperature) using sensors (as taught by Howard and/or Gruber) in the portable device and transmit the captured sensor data and/or related information to Howard-Doherty's server (i.e., server 200 in the cloud) to supplement speech processing. As discussed, Howard-Doherty discloses a system that can capture sensor data (e.g., audio or light) using sensors (microphones or light sensors) and transmit the captured sensor data to the server 200. For example, Howard discloses that "the transmitter system 101 also includes the ***optical sensor 6 for detecting optical data 114*** . . . The output of the optical sensor 6 is coupled to the input of an optical coder 115. Similar to the audio coder 111, the optical coder 115 packets/codes or converts the optical data 114 to modulated optical data signals 116, which are suitable for ***wireless transmission to the server 200*** . . . the transmitter 112 also functions to transmit the modulated optical data signals 116 to the server 200 via the antenna 127." (EX1005, 4:26-

38.) 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." (EX1008, [0013].) For example, the Howard-Doherty-Gruber combination can capture sensor data in the portable device and transmit the captured sensor data wirelessly to a router that routes the data to Howard-Doherty's remote server in the cloud via the Internet to supplement processing, Howard-Doherty's remote server (e.g., server 200) 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). Combining the teachings of Howard-Doherty and Gruber as such would yield predictable results of improving audio processing, particularly for commands for controlling home appliances and other in-home devices.

469. Incorporating Gruber's speech processing techniques into the Howard-Doherty system also 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 inputs. Because both the Howard-Doherty system 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 the Howard-Doherty system in the same way. Further, it would have been obvious to try combining the teachings of Howard-Doherty and Gruber. There are only a finite number of identified and predictable solutions for processing voice data and improving accuracy of the processing without requiring for excessive clarifications from users, 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 such as Gruber. Prior art such as Gruber has successfully demonstrated that device sensors can gather contextual information that would be useful in supplementing natural language input by users, and thus, there would have been a reasonable expectation of success for combining Howard-Doherty and Gruber.

470. 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 method 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."

471. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 4 is obvious in further view of Gruber.

472. Gruber discloses claim 4. (*Supra* § IX.C.3.)

473. It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Gruber. (*Supra* § IX.K.1.)

L. Ground 12: Howard, Doherty, and Benesty Render Obvious Claims 7-8

1. Claim 7: " The method according to claim 1, wherein the client device comprises multiple microphones, and wherein the capturing comprises capturing, by the multiple microphones in the client device, the human voice data."

474. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 7 is obvious in further view of Benesty.

475. Benesty discloses claim 7. (*Supra* § IX.D.2.)

476. It would have been obvious to a POSITA to apply Benesty's techniques of including multiple microphones to Howard-Doherty's client device. Both the Howard-Doherty system and Benesty are related to capturing and processing human audio data. Howard is directed to hands-free audio communication of a user with a client device, for the audio to be sent to a remote server for processing and to receive messages (audio reply or control signals) at the client device or the controlled device in response to the processing. Howard relates generally to "a method and system that allows an operator to easily and

readily control home activities without having to directly interface with or use a PC." (EX1005, 1:53-56.) Howard discloses "a portable device readily carried by the operator that can be utilized to communicate with a server located at a fixed location. Utilizing the portable device, the operator can program and control various components and activities." (EX1005, 1:58-64.) Each portable device has a microphone 3 and a speaker 4. (EX1005, 3:24-27, 3:54-62.)

477. 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." (EX1009, 1021.) Thus, Howard, Doherty, and Benesty are analogous art, from the same field of endeavor and all relate to the same technical problem of providing hands-free audio communication between users and devices.

478. The Howard-Doherty-Benesty combination would result in a system where multiple microphones as taught by Benesty are incorporated in Howard-Doherty's client device to capture human voice data and improve the quality of captured audio in spite of room reverberance and noise in the environment. For example, Benesty's microphone array(s) may be included in the Howard-Doherty's

portable device to capture the first and second human voice data. Because the Howard-Doherty system and Benesty relates audio capturing and processing, a POSITA would be motivated to use the known techniques of Benesty to improve the Howard-Doherty system's ability to capture and process audio, especially in a room where the microphone is not placed near the mouth of the user.

479. Combining the teachings of Howard-Doherty and Benesty as such would yield predictable results of effectively improving audio capturing. Further, it would have been obvious to try combining the technical features of Benesty with the teachings of Howard, as there is a finite number of identified, predictable resolutions, without requiring much 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 such as Benesty. 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 method according to claim 7, wherein the multiple microphones 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."

480. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 8 is obvious in further view of Benesty.

481. Benesty discloses claim 8. (*Supra* §§ IX.D.2 - IX.D.3.)

482. It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Benesty. (*Supra* § IX.L.1.)

M. Ground 13: Howard, Doherty, and Goldsmith Render Obvious Claims 12, 14-15

1. Claim 12: The method according to claim 1, wherein the wireless network is a Wireless Personal Area Network (WPAN), that is according to, or based on, Bluetooth™ or Institute of Electrical and Electronics Engineers (IEEE) 802.15.1-2005 standards, or wherein the WPAN is a wireless control network that is according to, or based on, Zigbee™, IEEE 802.15.4-2003, or Z-Wave™ standards.

483. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 12 is obvious in further view of Goldsmith.

484. Goldsmith discloses claim 12. (*Supra* § IX.E.2.)

485. Howard-Doherty discloses "the portable device 10 also comprises a controller (not shown in FIG. 1(a)) which functions to control the operation of the foregoing components, and an ***antenna*** (also not shown in FIG. 1(a)) which functions to allow for ***wireless communication*** between the ***portable device 10*** and a ***server 200*** . . ." (EX1005, 3:41-48.) Further, with reference to FIG. 4, "in the preferred embodiment, the communication device 201 (e.g. data transceiver) is

capable of performing ***wireless and/or optical transmission of data/commands*** with the ***portable device 10***." (EX1005, 6:32-36.)

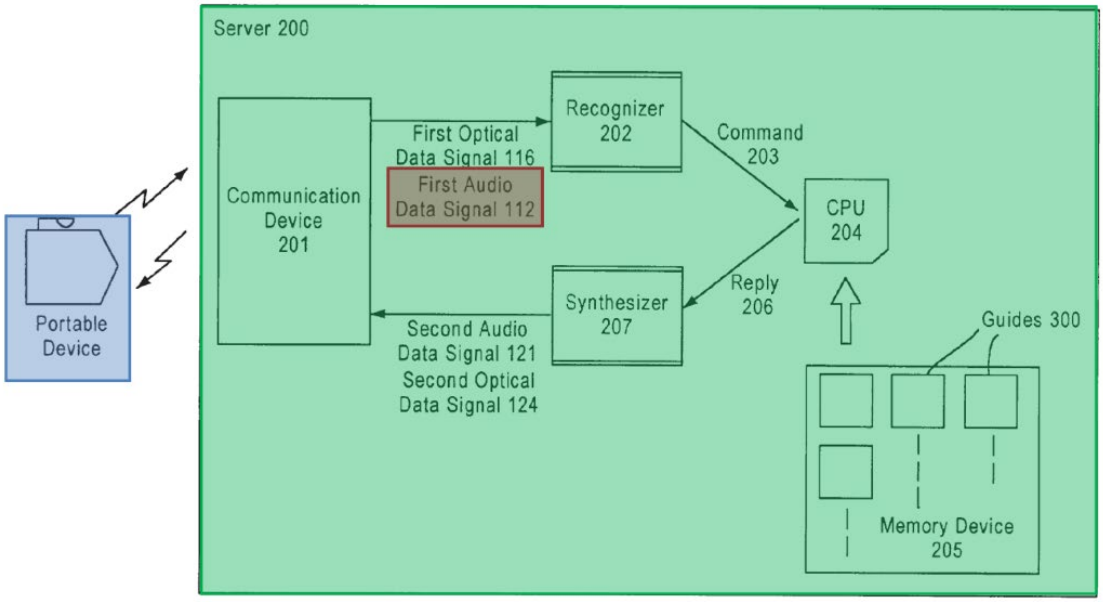


FIG. 4

486. Further, Howard-Doherty discloses that "the data and the control signals utilized to control the components coupled to the home network can be packetized and transmitted utilizing known schemes, such as BlueTooth, which is the IEEE 802.15.1 standard for wireless personal area networks." (*Supra* § IX.I.3.c; EX1005, 5:22-25.)

487. At the time of the invention, Bluetooth routers are well-known. "A wireless router such as 802.11 router, 802.16 router, WiFi router, WiMAX router, ***Bluetooth router***, X10 router can be connected to the mesh network." (EX1025, 59:21-24.) "[T]he wireless network router 31 may be a wireless LAN router, such as an IEEE 802.11x or ***Bluetooth router***." (EX1026,

4:25-27.) "[T]he wireless router comprises one of: 802.11 router, 802.16 router, WiFi router, WiMAX router, ***Bluetooth router***, ×10 router." (EX1027, 4:20-22.) Further, Bluetooth access points/gateways with connection to the Internet were well-known. For example, a patent application filed on March 29, 2002, discloses a mobile computing device "coupled, wirelessly through a Bluetooth network interface card (NIC) to ***Bluetooth access point*** 130, which, in turn, is coupled to the wired ***Ethernet*** backbone 120." (EX1032, 1:67-2:3.) A POSITA would have understood that, by coupling to the Ethernet backbone, the Bluetooth access point is connected to the Internet.

488. It would have been obvious for a POSITA to look to Goldsmith for conventional wireless networking techniques and protocols that can be applied to Howard-Doherty. Howard-Doherty and Goldsmith are analogous art, from the same field of endeavor and all relate to the same technical problems of wireless communications. (EX1005, 3:41-48, 5:22-25; EX1035; 8-21.) Goldsmith itself discloses that it is feasible to embed Bluetooth and Zigbee in electronic devices to create "smart homes." A POSITA would have looked to Goldsmith's disclosure of Bluetooth and Zigbee as two conventional wireless communication protocols and understood that the wireless data communications between devices as disclosed in Howard can be implemented by Bluetooth and Zigbee in the same way (e.g., the interface device and controlled device may communicate with the server 400 (i.e.,

application engine 118 and applications 120) wirelessly via Bluetooth or Zigbee). For example, the interface device and controlled device may communicate via Bluetooth with a Bluetooth router or another device that can manage and relay Bluetooth connections, and the router/device can communicate with the cloud server over the Internet. Combining Goldsmith with Howard-Doherty would improve and expand the capabilities/range of Howard-Doherty in the same way. Combining the teachings of Howard-Doherty and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system. Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (e.g., Bluetooth and Zigbee) with the teachings of Howard-Doherty without requiring much 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 as explained by Goldsmith. There was a finite number of potential wireless network communications, such as infrared, Wi-Fi, 3G, 4G, UWB, and GSM, that were commonly available and can be easily applied in Howard-Doherty at the time of the invention. The known benefits and feasibility of including a known wireless networking technique such as Bluetooth and Zigbee as taught by Goldsmith to connect in-home devices for creating smart

home would prompt the application of Goldsmith in the same field based on design incentives and market forces.

2. Claim 14: "The method according to claim 1, wherein the wireless network uses a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

489. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 14 is obvious in further view of Goldsmith.

490. As discussed, Howard-Doherty discloses wireless communications (i.e., Bluetooth and 802.11b WLAN) between devices and the server. (*Supra* § IX.I.3.c; EX1005, 5:22-25.)

491. Further, as discussed, the Howard-Doherty combination would result in a system in which the portable device sends data (e.g., voice data) wirelessly over 802.11b/g to the router which routes it to the server in the cloud via the Internet, thus, the server receives the voice data over the Internet via the router, and the server in the cloud sends data (e.g., messages) via the Internet to the router, which routes the data wirelessly over 802.11b/g to the portable device or the home devices. (*Supra* §§ IX.I.3.c - IX.I.3.d.)

492. Goldsmith and Doherty disclose this limitation. Both Bluetooth and WLAN over 802.11b and 802.11g are unlicensed radio frequency bands that are ISM bands. "Wireless networks operate today in what are called unlicensed frequencies . . . The common public unlicensed frequency bands (in the United

States) are 900 MHz, 2.4 GHz, and 5.1 GHz . . . 802.11b, 802.11g, and 802.11n networks are operate in the 2.4-GHz band." (EX1006, 229-230.) Further, Goldsmith expressly shows that both Bluetooth and 802.11b WLANs are ISM Band II frequencies. (EX1035, 22-23.)

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

493. As discussed above, a POSITA would be motivated to combine Howard-Doherty and Goldsmith and apply the well-known wireless communication techniques such as Bluetooth, Zigbee, and 802.11b WLAN as taught in Goldsmith to Howard-Doherty. (*Supra* § IX.M.1.)

3. Claim 15: "The method according to claim 1, wherein the wireless network is 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, or wherein the cellular telephone network is a Fourth Generation (4G) network that uses 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."

494. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 15 is obvious in further view of Goldsmith.

495. Goldsmith discloses claim 15. (*Supra* § IX.E.5.)

496. Howard specifically discloses that its portable device's "transmitter 113, receiver 120 . . . utilize the same data communication/transfer techniques as utilized for ***data transmission in cellular phones***." (EX1005, 5:16-21.)

497. As discussed, 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," "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***." (EX1030, claims 14-16, 18.)

498. 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 Howard-Doherty. As

discussed above, Howard-Doherty and Goldsmith are analogous art, from the same field of endeavor and all relate to the same technical problem of providing wireless communications between devices. Howard-Doherty discloses wireless communication between devices and server. (*Supra* §§ IX.I.3.c, IX.M.2; EX1005, 3:41-48, 5:22-25.) Goldsmith discloses conventional cellular data technologies. (EX1035, 12-13, 8-21.) 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 Howard-Doherty in the same way. (*Supra* § IX.E.2.) Combining the teachings of Howard-Doherty and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

499. Further, it would have been obvious to try combining the technical features of wireless communication techniques of Goldsmith via 3G over W-CDMA (UMTS), GSM, and EDGE with the teachings of Howard-Doherty as there is a finite number of identified and predictable solutions for network communications, such as infrared, Bluetooth, and Wi-Fi, that were commonly seen in the industry and were easy to apply to Howard-Doherty. The known benefits and feasibility of including a known wireless networking technique such as 3G using W-CDMA (UMTS), GSM, and EDGE as taught by Goldsmith to connect in-

home devices with remote Internet-connected servers would prompt the application of Goldsmith in the same field based on design incentives and market forces.

N. Ground 14: Howard, Doherty and Chatterjee Render Obvious Claim 22

1. Claim 22: "The method according to claim 16, 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."

500. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 22 is obvious in further view of Chatterjee.

501. Chatterjee discloses claim 22. (*Supra* § IX.F.7.)

502. It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Chatterjee. Both Howard-Doherty and Chatterjee relate to a home automation system for remotely controlling home appliances. As discussed, the controlled device in the Howard-Doherty system are devices such as "***TV, stereo, and video equipment*** or ***'life support systems', such as, heating systems, air conditioning systems and lighting and security systems***." (*Supra* § IX.I.3.b; EX1005, 1:18-30; *see also*, *id.*, 9:46-50 ("allow for control of the various ***household components and appliances***.").)

503. Further, with reference to FIG. 6, Howard discloses "controlling apparatus connected on the home network 304. Such apparatus include, but are not limited to, ***(a) audio/video equipment, for instance TVs, VCRs, DVDs, CD***

players, (b) electrical appliances, such as light systems, refrigerators, microwave ovens, heating and cooling systems, (c) computers, (d) home automation systems, (e) fire alarm systems, (f) burglar alarm systems, (g) telephone systems, and (h) intercom systems." (EX1005, 9:13-23.) Chatterjee discloses "a system and method for providing home automation using a mobile device." (EX1011, Abstract.) Thus, both Howard-Doherty and Chatterjee are analogous art, from the same field of endeavor, and all relate to the same technical problems of satisfying users' needs to control home appliances remotely. Combining Chatterjee with Howard-Doherty would improve and expand the capabilities of Howard-Doherty in the same way. Combining the teachings of Howard-Doherty and Chatterjee as such would yield predictable results of enabling home automation.

504. Because Howard explicitly provides that the Howard-Doherty system can be used to control any component coupled to the home network such as ***home automation systems*** and Chatterjee enumerates exemplary home automation devices, it would have been obvious to try combining the teachings of Howard-Doherty and Chatterjee.

505. 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 Chatterjee has successfully demonstrated that the controlled device can be a washing machine or a dishwasher, and thus, there would have been

a reasonable expectation of success for combining Howard-Doherty and Chatterjee. 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 Howard-Doherty to use Howard-Doherty's techniques to control home appliances such as a washing machine or a dishwasher as disclosed by Chatterjee. The known benefits and feasibility of remote controlling home appliances such as a washing machine or a dishwasher as taught by Chatterjee would prompt the application of Chatterjee in the same field based on design incentives and market forces.

O. Ground 15: Howard, Doherty, and Rosenberger Render Obvious Claims 24-25

1. Claim 24: "The method according to claim 1, wherein the first actuator is 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."

506. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 24 is obvious in further view of Rosenberger.

507. Rosenberger discloses claim 24. (*Supra* § IX.G.2)

508. At the time of the invention, it would have been obvious to combine the teachings of Rosenberger with Howard-Doherty. Both Rosenberger and Howard-Doherty relate to processing audio information such as human speech for controlling devices. (*See* EX1005, 1:53-56 and 11:20-23.) 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." (EX1012, 1:7-13.) Thus, Rosenberger and Howard-Doherty are analogous art, from the same field of endeavor, and all relate to the same goal of processing and responding to audio commands. A POSITA would be motivated to use the known techniques of Rosenberger to improve the usability of the device in the same way.

509. 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. 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 Howard-Doherty to improve the usability of the Howard-Doherty's portable device, with a reasonable expectation of success.

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

511. Incorporating an indicator light in the portable device disclosed in the Howard-Doherty-Rosenberger combination yields 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 Howard-Doherty and Rosenberger relate to capturing human voice data for controlling devices, a POSITA would be motivated to use the known techniques of Rosenberger to improve the usability of the device 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.

512. Further, it would have been obvious to try combining the teachings of Howard-Doherty 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 Howard-Doherty and Rosenberger. The known benefits of Rosenberger would prompt the application of Rosenberger in the same field based on design incentives and other market forces.

2. Claim 25: "The method according to claim 24, 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."

513. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 25 is obvious in further view of Rosenberger.

514. Rosenberger discloses claim 25. (*Supra* § IX.G.2)

515. It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Rosenberger. (*Supra* § IX.O.1.)

P. Ground 16: Howard, Doherty, and Rossing Render Obvious Claims 9, 26-28

1. Claim 9: "The method 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."

516. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 9 is obvious in further view of Rossing.

517. Rossing discloses claim 9. (*Supra* § IX.H.2)

518. As discussed, Howard discloses a microphone in the ***client device***. Howard discloses a "***portable device*** comprising: a body; a ***microphone*** for

receiving first audio data[.]" (EX1005, Claims 1 and 9.) Further, Howard discloses that "the ***portable device*** includes: (1) a body; (2) a ***microphone*** for receiving a first audio data." (EX1005, Abstract, 1:65-2:6.) With reference to FIG. 1(a), "the ***portable device 10*** comprises a clip 2, a ***microphone*** 3, a speaker 4, a display device 5, and an optical sensor 6." (EX1005, 3:25-27*; see also id.*, FIG. 1(b).) (*Supra* § IX.I.3.e.)

519. Both Howard-Doherty and Rossing relate to operating microphone(s) and speaker(s). Howard discloses a microphone and a speaker 4 in the client device. (EX1005, 3:25-27; claims 2 and 9, 3:25-27.) Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1013, 757-758, 763, 765, 903.) Thus, Howard-Doherty and Rossing are analogous art, 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 conventional components for implementing the Howard-Doherty system.

520. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional microphones to Howard-Doherty's microphone, as various microphone features can improve Howard-Doherty'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

other market forces. Further, it would have been obvious to try combining Rossing with the teachings of Howard-Doherty, 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 requiring much 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 seen in the industry and easy to manufacture and integrate into the Howard-Doherty system.

2. Claim 26: "The method according to claim 1, wherein the first actuator is a motion actuator that causes linear or rotary motion."

521. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 26 is obvious in further view of Rossing.

522. Rossing discloses claim 26. (*Supra* § IX.H.3)

523. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Howard-Doherty's speaker. As discussed, Howard-Doherty and Rossing are analogous art, from the same field of endeavor, and all relate to the same technical problems of operating microphone(s) and

speaker(s). As discussed, Howard discloses operating a ***speaker*** in response to the received ***first message*** sent to the ***client device***. (*Supra* § IX.I.3.k; EX1005, claims 2 and 9, 3:25-27.) Rossing is a textbook providing basic information on microphones and speakers. (EX1013, 757-758, 763, 765, 903.) Thus, a POSITA would have been motivated to look to Rossing for techniques and details related to conventional components for implementing the Howard-Doherty system. Further, it would have been obvious to try combining the technical features of Rossing with the teachings of Howard-Doherty. (*Supra* § IX.P.1.)

3. Claim 27: "The method according to claim 1, wherein the first actuator is a sounder for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible, sound waves."

524. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 27 is obvious in further view of Rossing.

525. As discussed, a POSITA would have understood that the claim refers to a speaker. (*Supra* § IX.H.4.) As discussed above, Howard discloses operating a speaker in the client device in response to the received first message. For example, FIG. 2 depicts a "***speaker 4***, which functions to output the audio data 123." (*Supra* § IX.I.3.k.)

526. Further, Rossing discloses claim 27. (*Supra* § IX.H.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. (*Supra* §

IX.H.4.) For example, the actuator 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." (EX1017, 122.)

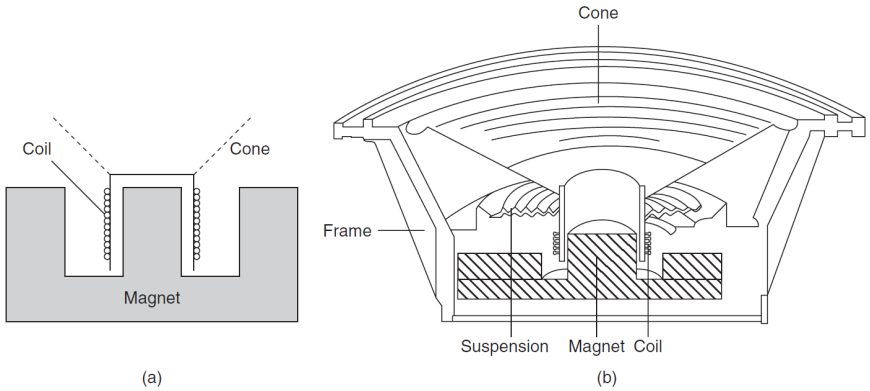


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

527. Further, in light of Rossing, a POSITA would understand that a speaker emits sound waves that are of at least omnidirectional pattern. (*Supra* § IX.H.4.) "***Sound 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***." (*Id*., EX1013, 765.)

4. Claim 28: "The method according to claim 27, 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, or wherein the sounder comprises an electric bell, a buzzer, a chime, a whistle, or a ringer."

528. Claim 1 is obvious in view of Howard and Doherty. (*Supra* § IX.I.3) Claim 28 is obvious in further view of Rossing.

529. As discussed, a POSITA would have understood that the claim refers to a speaker. (*Supra* § IX.H.4.)

530. Further, Rossing discloses claim 28. (*Supra* § IX.H.5)

531. 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. "[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***." (EX1013, 763.) "In 1929, E. W. Kellogg filed a patent for an ***electrostatic loudspeaker***, consisting of panels not unlike large condenser microphones." (EX1013, 321-322.) Rossing also discloses using "***small loudspeaker units*** (***piezoelectric or electro magnetic***) and the microphones 1/4" or even 1/8" condenser types." (*Id*.)

532. As discussed, Howard-Doherty and Rossing are analogous art, from the same field of endeavor, and all relate to the same technical problems of operating microphone(s) and speaker(s). Howard discloses operating a ***speaker*** in response to the received ***first message*** sent to the ***client device***. (*Supra* § IX.I.3.k; EX1005, claims 2 and 9, 3:25-27.) Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1013, 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 Howard-Doherty.

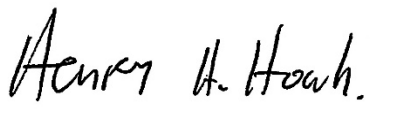
533. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Howard-Doherty's speaker, as various speaker features can improve Howard-Doherty's speaker in the same way and yield predictable results. The known benefits of various features would prompt the application of Rossing in the same field based on design incentives and other market forces. Further, it would have been obvious to try combining the technical features of Rossing with the teachings of Howard-Doherty, 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 requiring much effort and with a reasonable expectation of success. (*Supra* § IX.H.5.) 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 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 Howard-Doherty system. (EX1013, 321-322, 763.)

* * *

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: November 21, 2025



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

Amazon
EX1002 Appendix A Page 3

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

Amazon
EX1002 Appendix A Page 4

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"

Amazon
EX1002 Appendix A Page 5

- 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.

Amazon
EX1002 Appendix A Page 6

- "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 – 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

Amazon
EX1002 Appendix A Page 7

        - 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
- 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, 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

Amazon
EX1002 Appendix A Page 9