

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

**PETITION FOR *INTER PARTES* REVIEW
UNDER 35 U.S.C. §§ 311-319 AND 37 C.F.R. § 42.100 *et seq.***

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. THE '710 PATENT	2
A. Summary	2
B. Effective Filing Date	3
C. Prosecution History	3
D. Level of Ordinary Skill in the Art	4
III. CLAIM CONSTRUCTION	5
IV. PRECISE REASONS FOR RELIEF REQUESTED	5
A. Ground 1: Williams Renders Obvious Claims 1, 5-6, 16, 23, and 29	8
1. Summary of Williams	8
2. Claim 1	10
3. Claims 5 and 6	36
4. Claims 16 and 23	36
5. Claim 29	36
B. Ground 2: Williams and Wittenberg Render Obvious Claim 2	37
1. Summary of Wittenberg	37
2. Claim 2	37
C. Ground 3: Williams and Gruber Render Obvious Claims 3-4, 10-11	39
1. Summary of Gruber	39
2. Claim 3	39
3. Claim 4	42
4. Claims 10 and 11	42
D. Ground 4: Williams and Benesty Render Obvious Claims 7 and 8	44
1. Summary of Benesty	44
2. Claims 7 and 8	44

TABLE OF CONTENTS **(continued)**

	Page
E. Ground 5: Williams and Goldsmith Render Obvious Claims 12-15	46
1. Summary of Goldsmith.....	46
2. Claim 12.....	46
3. Claim 13.....	48
4. Claim 14.....	49
5. Claim 15.....	50
F. Ground 6: Williams and Chatterjee Render Obvious Claim 17-22	51
1. Summary of Chatterjee	51
2. Claims 17-22	52
G. Ground 7: Williams and Rosenberger Render Obvious Claims 24 and 25	54
1. Summary of Rosenberger	54
2. Claims 24 and 25.....	54
H. Ground 8: Williams and Rossing Render Obvious Claims 9 and 26-28.....	56
1. Summary of Rossing.....	56
2. Claim 9	57
3. Claims 26-28	59
I. Ground 9: Howard and Doherty Render Obvious Claims 1, 5, 6, 10-11, 13, 16-21, 23, and 29.	61
1. Summary of Howard.....	61
2. Summary of Doherty.....	63
3. Claim 1	66
4. Claims 5 and 6.....	98
5. Claims 10 and 11.....	99

TABLE OF CONTENTS **(continued)**

	Page
6. Claim 13	99
7. Claims 16-21 and 23	100
8. Claim 29	101
J. Ground 10: Howard, Doherty, and Wittenberg Render Obvious Claim 2	102
K. Ground 11: Howard, Doherty, and Gruber Render Obvious Claim 3-4	104
1. Claims 3-4	104
L. Ground 12: Howard, Doherty, and Benesty Render Obvious Claims 7-8	104
M. Ground 13: Howard, Doherty, and Goldsmith Render Obvious Claims 12, 14-15	105
N. Ground 14: Howard, Doherty and Chatterjee Render Obvious Claim 22	106
O. Ground 15: Howard, Doherty, and Rosenberger Render Obvious Claims 24-25	107
P. Ground 16: Howard, Doherty, and Rossing Render Obvious Claims 9, 26-28	107
V. NOTICES AND STATEMENTS	108

Exhibit List

Exhibit Description	Exhibit #
U.S. Patent No. 11,128,710 to Binder <i>et al.</i>	1001
Expert Declaration of Dr. Henry Houh	1002
File History of U.S. Patent No. 11,128,710	1003
U.S. Patent Pub. No. 2012/0198339 to Williams <i>et al.</i>	1004
U.S. Patent No. 7,062,339 to Howard <i>et al.</i>	1005
Jim Doherty <i>et al.</i> , “Wireless Home Networking Simplified” (Cisco Press, 2007).	1006
U.S. Patent No. 7,529,677 to Wittenberg	1007
U.S. Patent Pub. No. 2012/0265528 to Gruber <i>et al.</i>	1008
Jacob Benesty <i>et al.</i> , “Springer Handbook of Speech Processing” (Springer-Verlag Berlin Heidelberg, 2008).	1009
RESERVED	1010
U.S. Patent Pub. No. 2013/0052946 to Chatterjee <i>et al.</i>	1011
U.S. Patent No. 8,340,975 to Rosenberger	1012
Thomas D. Rossing <i>et al.</i> , “Springer Handbook of Acoustics” (Springer Science+Business Media, LLC, 2007).	1013
Michele Mazzucco <i>et al.</i> , “Profit-Aware Server Allocation for Green Internet Services” (2010 18 th Annual IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems) (available at https://ieeexplore.ieee.org/abstract/document/5581586)	1014

Exhibit Description	Exhibit #
Kenneth J. Christensen <i>et al.</i> , “The next frontier for communications networks: power management” (2004) (available at https://www.sciencedirect.com/science/article/abs/pii/S0140366404002427)	1015
Lidong Wang <i>et al.</i> , “MMSEmail: Delivering Emails to Mobile phone through MMS” (2009 IEEE International Conference on Multimedia Information Networking and Security) (available at https://ieeexplore.ieee.org/abstract/document/5370492)	1016
Ian Sinclair <i>et al.</i> , “Electronic and Electrical Servicing” (Elsevier Limited, 2 nd Edition, 2007)	1017
U.S. Patent No. 6,906,617 to Van der Meulen	1018
Ali Vedavarz <i>et al.</i> , “HVAC: Handbook of Heating, Ventilation and Air Conditioning for Design and Implementation” (Industrial Press Inc., 2007)	1019
U.S. Patent Pub. No. 2013/0132094 to Lim	1020
U.S. Patent Pub. No. 2011/0299412 to Diab <i>et al.</i>	1021
Robert J. Shimonski <i>et al.</i> , “Sniffer Pro Network Optimization and Troubleshooting Handbook” (Syngress Publishing, Inc., 2002).	1022
RESERVED	1023
U.S. Patent Pub. No. 2012/0253480 to Abe	1024
U.S. Patent No. 7,539,533 to Tran	1025
U.S. Patent No. 8,356,754 to Johnson <i>et al.</i>	1026
U.S. Patent No. 8,249,731 to Tran <i>et al.</i>	1027
P.A. Roberts <i>et al.</i> , “IP address management” (BT Technology Journal, Vol. 18, 2000) (available at https://link.springer.com/article/10.1023/A:1026749131441)	1028

Exhibit Description	Exhibit #
Jacob Fraden, “Handbook of Modern Sensors” (Springer Science+Business Media, LLC, 2010)	1029
U.S. Patent No. 8,183,995 to Wang <i>et al.</i>	1030
U.S. Patent No. 7,590,542 to Williams <i>et al.</i>	1031
U.S. Patent No. 7,492,787 to Ji <i>et al.</i>	1032
Declaration of Tony Corsini on the scanned copies of pages of Jim Doherty <i>et al.</i> , “Wireless Home Networking Simplified” (Cisco Press, 2007).	1033
Declaration of Catherine Vassikova on the scanned copies of pages of Jacob Benesty <i>et al.</i> , “Springer Handbook of Speech Processing” (Springer-Verlag Berlin Heidelberg, 2008).	1034
Declaration of Rachel J. Watters on the scanned copies of Chapter 1 and Appendix D of Andrea Goldsmith, “Wireless Communications” (Cambridge University Press, 2005).	1035
Declaration of Tony Corsini on the scanned copies of pages of Thomas D. Rossing <i>et al.</i> , “Springer Handbook of Acoustics” (Springer Science+Business Media, LLC, 2007).	1036
Declaration of Gordon MacPherson on the IEEE publications Michele Mazzucco <i>et al.</i> , “Profit-Aware Server Allocation for Green Internet Services” (2010 18 th Annual IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems) and Lidong Wang <i>et al.</i> , “MMSEmail: Delivering Emails to Mobile phone through MMS” (2009 IEEE International Conference on Multimedia Information Networking and Security).	1037
Declaration of Catherine Vassikova on the scanned copies of pages of Kenneth J. Christensen <i>et al.</i> , “The next frontier for communications networks: power management” (2004).	1038

Exhibit Description	Exhibit #
Declaration of Catherine Vassikova on the scanned copies of pages of Lidong Wang <i>et al.</i> , “MMSEmail: Delivering Emails to Mobile phone through MMS” (2009 IEEE International Conference on Multimedia Information Networking and Security).	1039
Declaration of Tony Corsini on the scanned copies of pages of Ian Sinclair <i>et al.</i> , “Electronic and Electrical Servicing” (Elsevier Limited, 2 nd Edition, 2007).	1040
Declaration of Catherine Vassikova on the scanned copies of pages of Ali Vedavarz <i>et al.</i> , “HVAC: Handbook of Heating, Ventilation and Air Conditioning for Design and Implementation” (Industrial Press Inc., 2007).	1041
RESERVED	1042
Declaration of Tony Corsini on the scanned copies of pages of Jacob Fraden, “Handbook of Modern Sensors” (Springer Science+Business Media, LLC, 2010)	1043
Declaration of Tony Corsini on the scanned copies of pages of P.A. Roberts <i>et al.</i> , “IP address management” (BT Technology Journal, Vol. 18, 2000)	1044
M. Needham, “Smart Home Market on Track for Rebound Thanks to Emerging Regions, According to IDC,” https://my.idc.com/getdoc.jsp?containerId=prUS52632124 , IDC Research (Oct. 1, 2024)	1045
U.S. Patent No. 8,200,798 to Yadav <i>et al.</i>	1046
U.S. Patent No. 8,635,243 to Philips <i>et al.</i>	1047

Claims Listing

Claim	Limitation #	Limitation
1	1[PR][1]	A method for operating multiple actuators in response to captured human voice data,
	1[PR][2]	for use with a client device and a controlled device in a building,
	1[PR][3]	each communicating over a wireless network, and
	1[PR][4]	for use with an Internet-connected server device external to the building, the method comprising:
	1[A]	capturing, by a microphone in the client device, a first human voice data;
	1[B]	sending to the server, by the client device via the wireless network, the captured first human voice data;
	1[C]	receiving, by the server over the Internet, the captured first human voice data;
	1[D]	processing, by the server, the captured first human voice data;
	1[E]	responsive to the processing, sending a first message, by the server to the client device over the Internet;
	1[F]	receiving, by the client device via the wireless network, the first message;
	1[G]	operating a first actuator in the client device in response to the received first message;
	1[H]	capturing, by the microphone in the client device, a second human voice data;
	1[I]	sending to the server, by the client device via the wireless network, the captured second human voice data;
	1[J]	receiving, by the server over the Internet, the captured second human voice data;
	1[K]	processing, by the server, the captured second human voice data;
	1[L]	responsive to the processing, sending a second message, by the server to the controlled device over the Internet;

Claim	Limitation #	Limitation
	1[M]	receiving, by the controlled device via the wireless network, the second message; and
	1[N]	operating a second actuator in the controlled device in response to the received second message.
2	2	The method according to claim 1, wherein the processing comprises performing a voice recognition algorithm for identifying the voice of a specific person.
3	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.
4	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.
5	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.
6	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.
7	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.
8	8	The method according to claim 7, wherein the

Claim	Limitation #	Limitation
		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.
9	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.
10	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.
11	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.
12	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.
13	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,

Claim	Limitation #	Limitation
		Institute of Electrical and Electronics Engineers (IEEE) IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac.
14	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.
15	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.
16	16	The method according to claim 1, wherein the controlled device is integrated in, is part of, or is entirely included in, an appliance.
17	17	The method according to claim 16, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation.
18	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.
19	19	The method according to claim 17, wherein the

Claim	Limitation #	Limitation
		appliance is a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker.
20	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.
21	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.
22	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.
23	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.
24	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.
25	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.
26	26	The method according to claim 1, wherein the first

Claim	Limitation #	Limitation
		actuator is a motion actuator that causes linear or rotary motion.
27	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.
28	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.
29	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.

Petitioner Amazon.com Services LLC respectfully petitions for *inter partes* review of claims 1-29 of U.S. Patent No. 11,128,710 (EX 1001) per 35 U.S.C. §§311-319 and 37 C.F.R. §42.100 *et seq.*

I. INTRODUCTION

One billion smart devices are projected to be shipped by 2028. (EX1045.) The industry that manufactures and sells these devices drives domestic manufacturing and supports many U.S. jobs. Ensuring that this industry remains vital and unencumbered by assertions of unpatentable claims is critical to the U.S. economy. Smart Speaker has asserted one such patent. The claimed invention of the '710 patent covers well-known and routine smart-home functionalities. Even though the PTAB had affirmed much of the independent claims as unpatentable during examination, the Examiner allowed the claims following minor amendments without explanation. The Director should institute review on the '710 patent to protect our robust patent system and a critical segment of our nation's economy.

The '710 patent relates to “system and method for server-based control.” (EX1001.) Claim 1 requires a client device that receives two voice inputs: (1) “first human voice data” that triggers operation of an actuator in the client device itself and (2) “second human voice data” that triggers operation of an actuator in a controlled device, such as a home appliance. *Id.*

The Board found the latter obvious during prosecution. (*See* EX1003, 389-421.) The former was added only after prolonged prosecution, resulting in an abrupt allowance without explanation. (EX1003, 9-14.) But devices capable of receiving and responding to audio commands were prevalent. (*E.g.*, EX1004, EX1005, EX1008.) So were systems that process two voice inputs in the exact recited manner. *Id.* And the recited network architecture simply reflected the conventional configuration of home wireless networks at the time. (*E.g.*, EX1003.)

The challenged claims are unpatentable for the reasons set forth herein, as supported by Dr. Houh. (EX1002.)

II. THE '710 PATENT

A. Summary

The patent's specification is a vague aggregation of off-the-shelf technologies at the time of the invention. It recites 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.) Various network technologies were described to be "known in the art." (*E.g.*, EX1001, 2:34-38; 6:42-43; 6:66-7:6; 9:35-38.) Each was well-known at the time, and the specification provides no technical details of how they are integrated.

Although the claims focus on “human voice data,” the specification provides sparse and scattered references to voice. (EX1002, ¶¶19-22.) Only one paragraph—out of 171 columns of written description—describes voice control over a network, and that paragraph’s description is at a high level of generality. It describes “a router,” “a first device in the building (or vehicle) . . . operative to transmit a digital data corresponding to the sensed voice to the router,” and “[a] second device may be operative to execute actuator commands received from the router.” (EX1001, 32:25-38.) Further, a “control server may be operative to produce actuator commands in response to [a] voice processor output” (*Id.*, 32:57-33:13.) FIG. 19 depicts a high-level flowchart of “controller involving voice processing.” (*Id.*, 55:64-65.)

B. Effective Filing Date

The ’710 patent was filed on July 23, 2017, and claims priority to a provisional filed on January 9, 2012, which Petitioner treats as the earliest effective filing date only for purposes of this proceeding. (EX1002, ¶18.)

C. Prosecution History

The application underwent lengthy examination, including 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 explanation. (EX1003, 9-14.)

Original claim 1 merely recited processing a single set of “human voice data.” (EX1003, 1304.) After four Office Actions, the Applicant amended the claims and appealed. The Board affirmed the Examiner’s rejections of all claims. (EX1003, 389-421.) Effectively, *the Board has found the second half of the issued claim 1 (1[H]-1[N]) to be obvious.* (*Id.*, EX1002, ¶¶36-38.)

After the PTAB decision, the Applicant further amended claim 1 to recite processing both “first voice data” and “second voice data,” which cause operation of first and second actuators within the controlled device. (EX1003, 344.) The Examiner issued two additional Office Actions. (EX1003, 229-248, 145-167.)

The Applicant then amended claim 1 again to merely recite that the first human voice data is processed to control a *client device* rather than the controlled device. (EX1003, 121-130.) Although minimal, the Examiner allowed the application following that distinction without explaining the abrupt shift. (EX1003, 9-14.)

D. Level of Ordinary Skill in the Art

A person of ordinary skill in the art (POSITA) in the field of the ’710 patent would have a bachelor’s degree in electrical engineering, computer engineering, or equivalent, and two years of experience working on remote device control and/or networked devices. An advanced graduate degree may make up for the work experience indicated above. (EX1002, ¶¶47-49.)

III. CLAIM CONSTRUCTION

Claim terms are presumed to carry their ordinary and customary meaning. The Board construes the terms only to the extent necessary to resolve the issue at hand: here, the invalidity of the claims under 35 U.S.C. §103. The Board need not construe any of the claims because, as set forth in the 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. (EX1002, ¶50.)

IV. PRECISE REASONS FOR RELIEF REQUESTED

Pursuant to 37 C.F.R. §42.104(b), Petitioner requests cancellation of the challenged claims based on the following references (EX1002, ¶¶51-61). The Examiner never found, and Applicant never cited, any of these references during prosecution. *Id.*

The following references are prior art *at least* under the following provisions.

Prior Art Reference	Abbreviation (Exhibit No.)	Prior Art Basis
U.S. Publication No. 2012/0198339	“Williams” (EX1004)	§102(e)
U.S. Patent No. 7,062,339	“Howard” (EX1005)	§102(a), (b), and (e)

Prior Art Reference	Abbreviation (Exhibit No.)	Prior Art Basis
Jim Doherty <i>et al.</i> , “Wireless Home Networking Simplified” (Cisco Press, 2007).	“Doherty” (EX1006)	§102(a) and (b)
U.S. Patent No. 7,529,677	“Wittenberg” (EX1007)	§102(a), (b), and (e)
U.S. Publication No. 2012/0265528	“Gruber” (EX1008)	§102(e)
Jacob Benesty <i>et al.</i> , “Springer Handbook of Speech Processing” (Springer-Verlag Berlin Heidelberg, 2008).	“Benesty” (EX1009)	§102(a) and (b)
Andrea Goldsmith, “Wireless Communications” (Cambridge University Press, 2005).	“Goldsmith” (EX1035)	§102(a) and (b)
U.S. Publication No. 2013/0052946	“Chatterjee” (EX1011)	§102(e)
U.S. Patent No. 8,340,975	“Rosenberger” (EX1012)	§102(e)
Thomas D. Rossing <i>et al.</i> , “Springer Handbook of Acoustics” (Springer Science+Business Media, LLC, 2007).	“Rossing” (EX1013)	§102(a) and (b)

The statutory grounds for the challenge of each claim are set forth below.

All of the statutory citations are pre-AIA.

Ground	35 U.S.C. §	Claims	Prior Art Reference(s)
1	103	1, 5-6, 16, 23, 29	Williams
2	103	2	Williams, Wittenberg
3	103	3-4, 10-11	Williams, Gruber
4	103	7-8	Williams, Benesty
5	103	12-15	Williams, Goldsmith
6	103	17-22	Williams, Chatterjee
7	103	24-25	Williams, Rosenberger
8	103	9, 26-28	Williams, Rossing
9	103	1, 5-6, 10-11, 13, 16-21, 23, 29	Howard, Doherty
10	103	2	Howard, Doherty, Wittenberg
11	103	3-4	Howard, Doherty, Gruber
12	103	7-8	Howard, Doherty, Benesty
13	103	12, 14-15	Howard, Doherty, Goldsmith
14	103	22	Howard, Doherty, Chatterjee
15	103	24-25	Howard, Doherty, Rosenberger
16	103	9, 26-28	Howard, Doherty, Rossing

Below, Petitioner discusses why the challenged claims are unpatentable under the statutory grounds raised, including how and where the prior art satisfies

each element of each challenged claim, as required by 37 C.F.R. §42.104(b)(4).¹

Petitioner's showing establishes that there is a reasonable likelihood that it will prevail on each ground of unpatentability to each challenged claim.

A. Ground 1: Williams Renders Obvious Claims 1, 5-6, 16, 23, and 29

1. Summary of Williams

Williams was filed on January 28, 2011, before the earliest provisional application filing date of the '710 patent, and thus is prior art under §102(e). (EX1004, 1.) The Examiner never found, and Applicant never cited, Williams during prosecution. In Williams, “[a] user may ask a question, such as ‘What’s the recipe for deviled eggs?’” and the system can “provide audio answers to spoken questions.” (EX1004, [0048].) Further, the system “may accept spoken commands from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections.” (*Id.*, [0051], [0001], [0008], [0009], [0024]; EX1002, ¶62.)

Referring to FIG. 1, “[a]n individual *interface device 116* may include a microphone 202 that is configured to detect ambient noise, sounds, and *speech*.”

¹ Petitioner also shows how each claim preamble is satisfied but does not concede any is limiting.

(EX1004, [0023] (emphasis added)².) 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]; EX1002, ¶¶62-63.) Device 112 is “a router or access point.” (*Id.*, [0011]-[0013].)

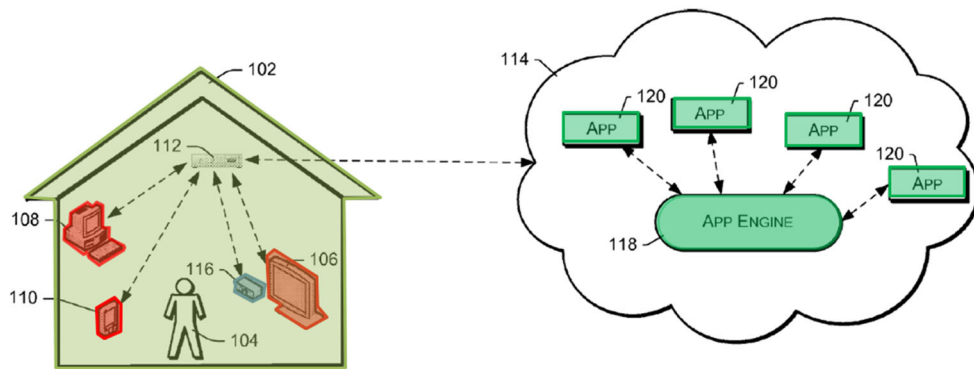


FIG. 1

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].) “In operation, the application engine 118 [of the *server 400*] receives audio-based information from the *interface devices 116.*” (*Id.*, [0022].) The server can “provide audio 222 for eventual transmission to the *interface device 116*, to be played on the speaker.” (*Id.*, [0037].) Further, Williams

² All emphasis and annotations are added unless otherwise stated.

discloses “allow[ing] a user to *control various on-premises devices and equipment, such as media equipment and home-automation devices.*” (EX1004, [0051]; EX1002, ¶¶64-66.)

2. Claim 1

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

Williams discloses a method for operating *multiple actuators* in response to captured *human voice data*. Williams “may accept *spoken commands* from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections.” (*Id.*, [0051]; EX1002, ¶¶67-68.) As explained herein, each of the “on-premises devices” contains actuators operable in response to captured human voice data. (*Infra* §IV.A.2.i, IV.A.2.n; EX1002, ¶68.)

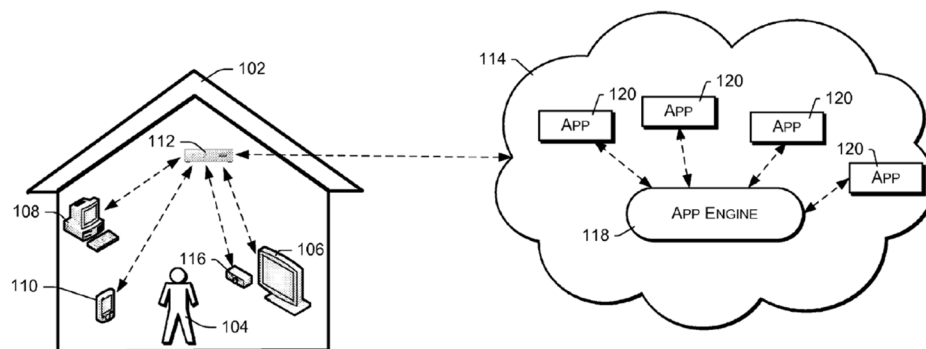


FIG. 1

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

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

(EX1002, ¶69.)

First, Williams discloses a *client device* in a *building*. For example, “[t]he *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].) The *device 116* communicates with and receives information from server 400 via network 114, and is thus a client device. (*Id.*, [0023], [0052]; EX1002, ¶70.)

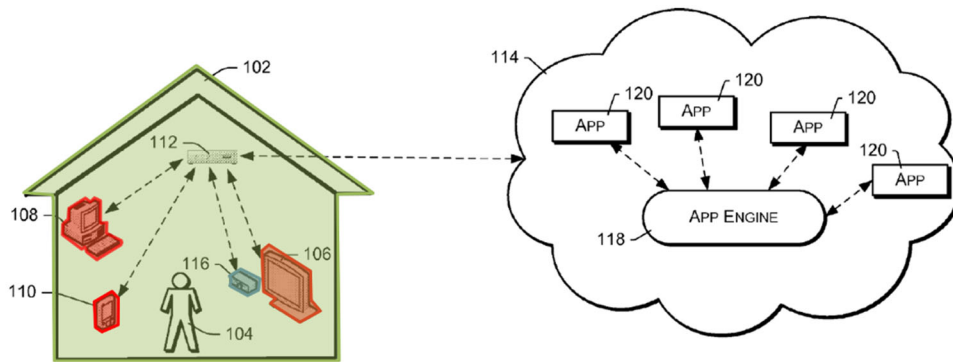


FIG. 1

FIG. 2 “is a block diagram illustrating additional details regarding the system of FIG. 1.” (EX1004, [0004].) Williams discloses that “[a]n individual *interface device 116* may include a microphone 202 that is configured to detect ambient noise, sounds, and speech.” (*Id.*, [0023]; EX1002, ¶71.)

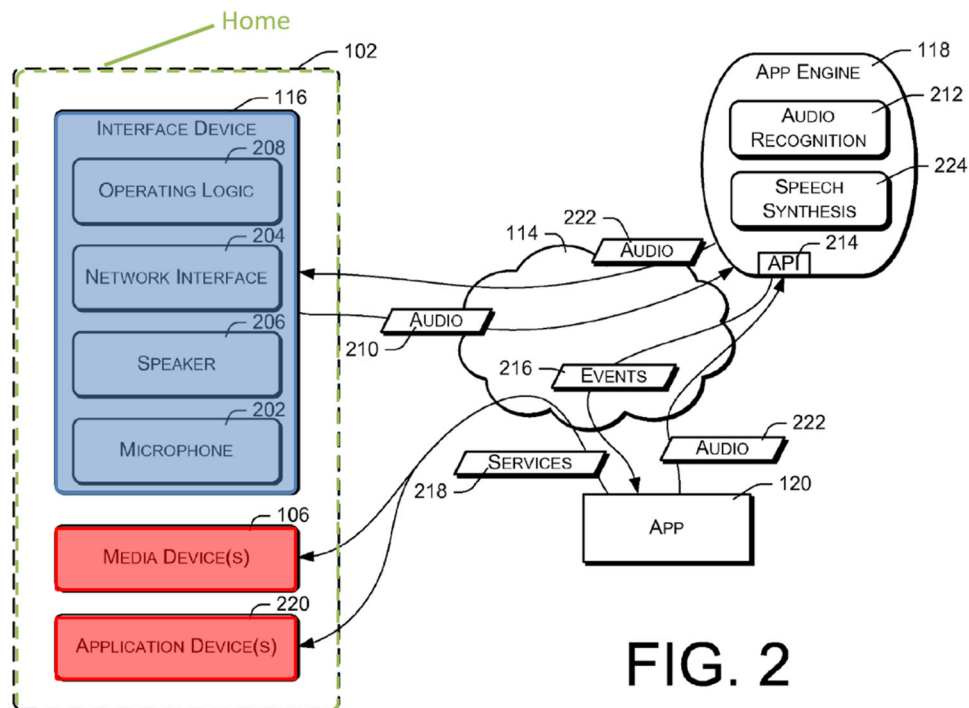


FIG. 2

Williams also discloses a *controlled device* in the *building*. With reference to FIG. 1, “[w]ithin the *home 102* is a user 104 and several devices . . . 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*” within the home. (EX1004, [0011], [0010]; EX1002, ¶72.)

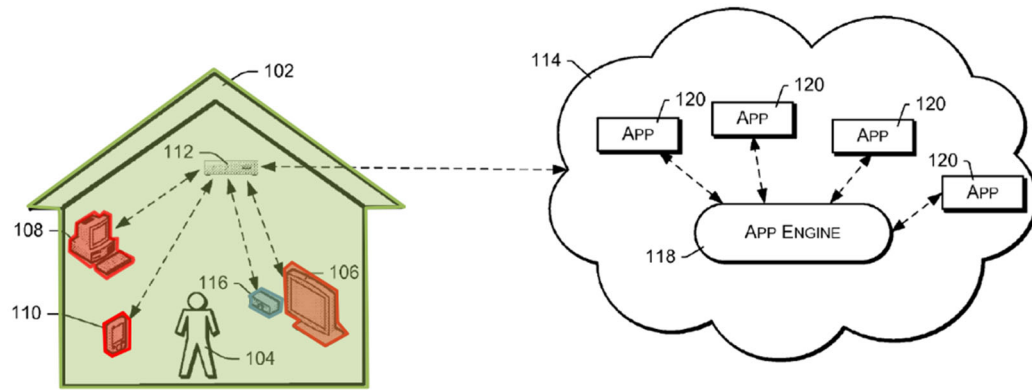


FIG. 1

The *devices 106, 108, and 110* are controlled devices. (EX1002, ¶73.)

Williams discloses “[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].) Referring to FIG. 2, “initiat[ing] or customiz[ing] 218 services 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]; EX1002, ¶74.)

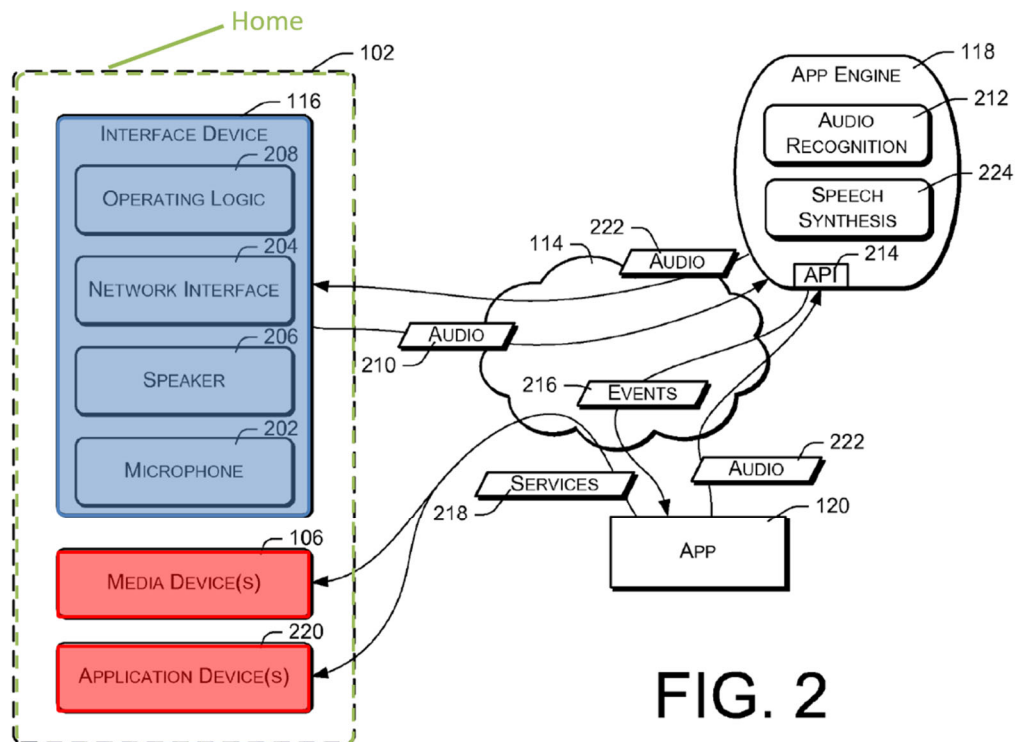


FIG. 2

As shown above, Williams discloses that the *client device* and the *controlled device* are in the same *building*. (EX1004, [0011], [0016], [0023], [0036]; EX1002, ¶¶75-78.)

c. 1[PR][3]: “each communicating over a wireless network,”

Williams discloses the *client device* and the *controlled device* communicate over a wireless network. Referring to FIG. 1, “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, [0011]-[0013].) Thus, *interface device 116* and *media devices 106/108/110* communicate over a wireless network through wireless

router/access point 112. 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, 28:7-10; EX1002, ¶¶79-80.)

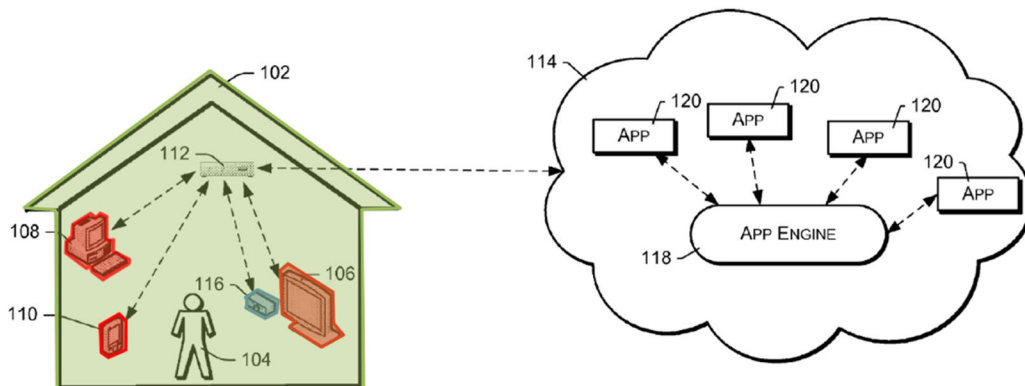


FIG. 1

Williams discloses other types of wireless networks. 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], [0020], [0023].) Further, network 114 “represents . . . the ‘cloud’” and can itself include “wireless networks”. (*Id.*, [0013]-[0015].) Any of Williams’ wireless networks, alone or in combination, discloses a “wireless network”. (EX1002, ¶¶80-83.)

Williams discloses *audio interface device 116* has “a microphone and a network interface” “with wireless networking capabilities”. (EX1004, [0007], [0016], [0020].) It would have been obvious that the *audio interface device 116* communicates over a wireless network. (EX1002, ¶84.)

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

Williams discloses the method is for use with an *Internet-connected server device* external to the *building*.

Williams discloses the in-home network is connected to cloud network 114. (EX1004, FIG. 1; EX1002, ¶¶85-86.) “The network 114 may thus include . . . *the Internet*, wireless networks, . . . cellular communications networks.” (EX1004, [0015], [0001].) The network 114 includes servers, such as “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], [0014], [0021].) “[T]he *server 400* may also comprise a *web service component 410*.” (*Id.*, [0055].) A POSITA would have understood that *server 400* is Internet-connected and the in-home wireless network (*e.g.*, via router 112) is connected to cloud *server 400* via the Internet. (EX1002, ¶¶85-88.)

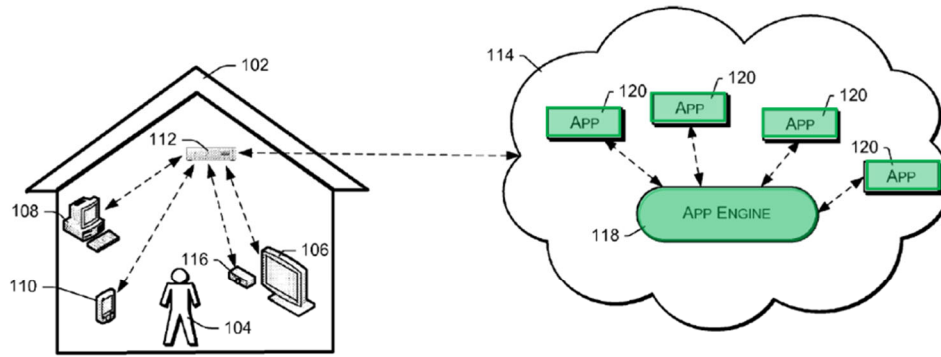


FIG. 1

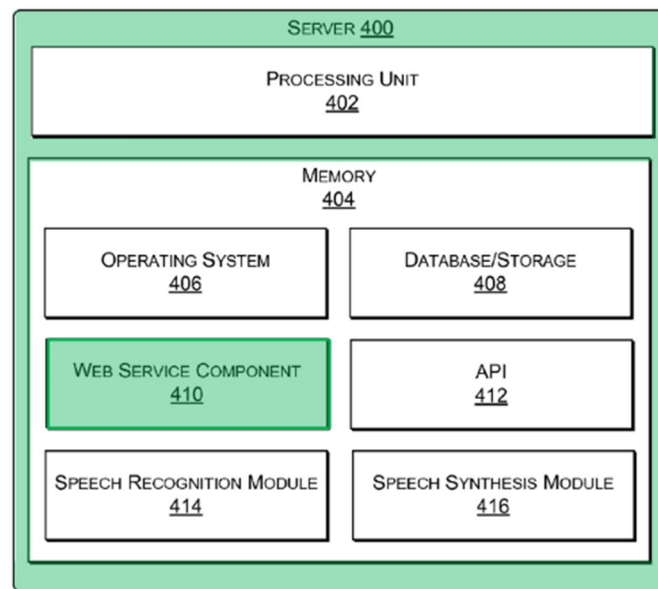


FIG. 4

This is consistent with the '710 patent, which explains that “[t]he system may comprise . . . an Internet-connected control server (referred herein also as ‘*cloud server*’ and ‘gateway server’) external to the building . . . and

communicatively coupled to the router over the Internet via the external network.”
(EX1001, 32:25-42.)

Williams’ *server 400* is external to the *building* (e.g., the user premise 102) as shown in FIG. 1. Further, “[s]ervers may be located together or separately, and organized as virtual servers, *server banks, and/or server farms.*” (EX1004, [0052].) A POSITA would have understood that “server banks” and the “server farms” are outside a residential home and that they are accessible via the Internet. (EX1002, ¶¶89-90; e.g., EX1014, 277; EX1015, 1760.)

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.” (EX1004, [0021].) A POSITA would have understood that this does not mean that the server itself is local. (EX1002, ¶91.) Rather, the “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]; see EX1016, 664.)

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

Referring to FIG. 2, “[a]n individual *interface device 116* may include a *microphone* 202 that is configured to detect ambient noise, sounds, and *speech.*” (EX1004, [0023], [0016], [0019] Claim 1-3 and 21.) The speech data is captured

by the microphone and sent to the cloud for processing. (*Id.*, [0022], [0024]-[0025]; EX1002, ¶¶92-97.)

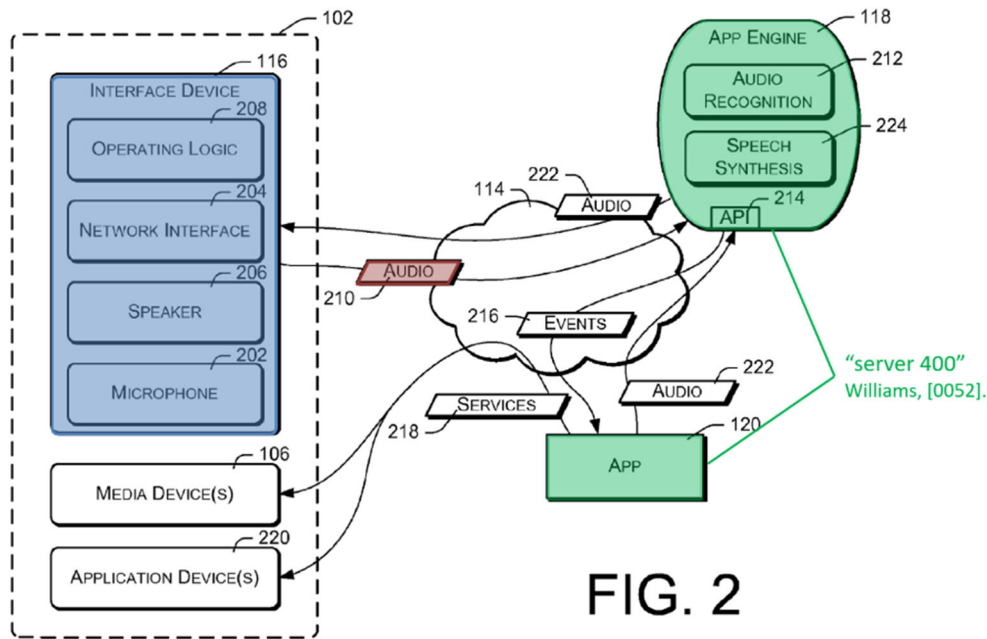
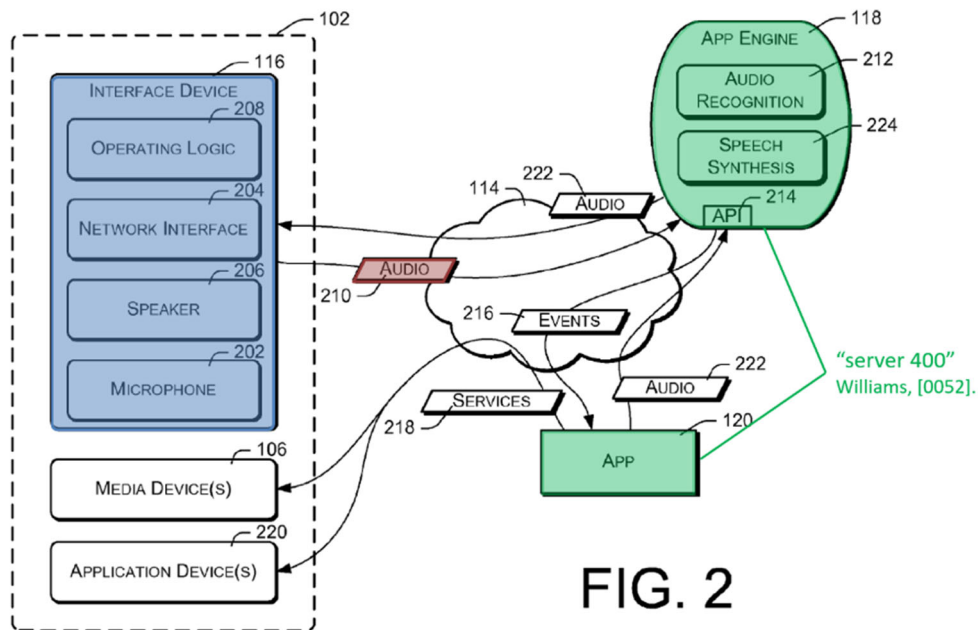


FIG. 2

f. 1[B] and 1[C]: “sending to the server, by the client device via the wireless network, the captured first human voice data; receiving, by the server over the Internet, the captured first human voice data;”

Williams discloses that the *client device* sends the captured *first human voice data* to the *server* and receiving it by the *server* over the Internet. As shown in FIG. 2, *interface device 116* sends *audio 210* (captured by microphone 202) to *server 400*. (EX1004, [0023]; *supra* §IV.A.2.d.) The *audio 210* includes human voice data such as “spoken commands uttered by users”. (EX1004, [0028], [0007], [0009].) “[T]he application engine 118 [of *server 400*] receives *audio-based*

information from the *interface devices 116*.” (EX1004, [0022], [0020], [0024]-[0028], [0033]-[0034], [0038]-[0039]; EX1002, ¶¶98-100, 102-104.)



Referring to FIG. 1, the *audio data* is sent by *interface device 116* to the router 112 via the in-home wireless network and then 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 the *server*], *via the router 112 and the network 114* [i.e., the Internet]” and may “use cellular or other wireless data communications technologies” to do so. (EX1004, [0001], [0007], [0011]-[0015], [0020], [0038], [0051]; *supra* §§IV.A.2.b, IV.A.2.c, IV.A.2.d; EX1002, ¶¶101, 105.)

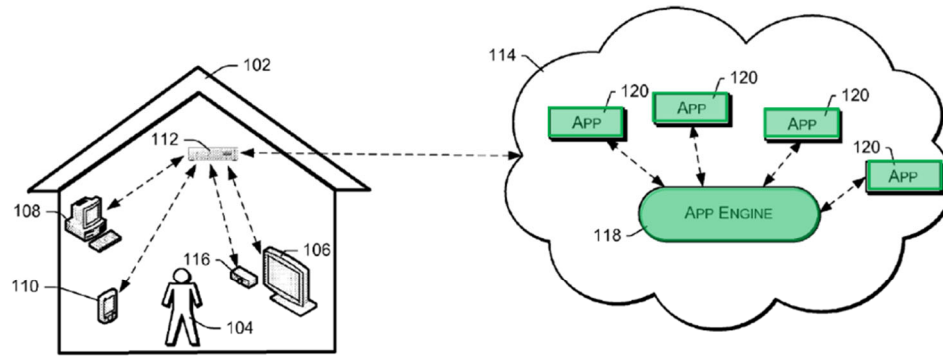


FIG. 1

g. 1[D]: “processing, by the server, the captured first human voice data;”

Williams discloses that the *server* processes the captured *first human voice data*. “[T]he application engine 118 [run on *server 400*] processes the *audio-based information* to determine various data about user activities, status, environment, commands, etc.” (EX1004, [0022], [0052]; *supra* §IV.A.2.d.) Williams discloses several examples of processing, such as “speech-to-text conversion”. (EX1004, [0034]-[0037], [0039]–[0040], [0053], [0056]; EX1002, ¶¶106-109.)

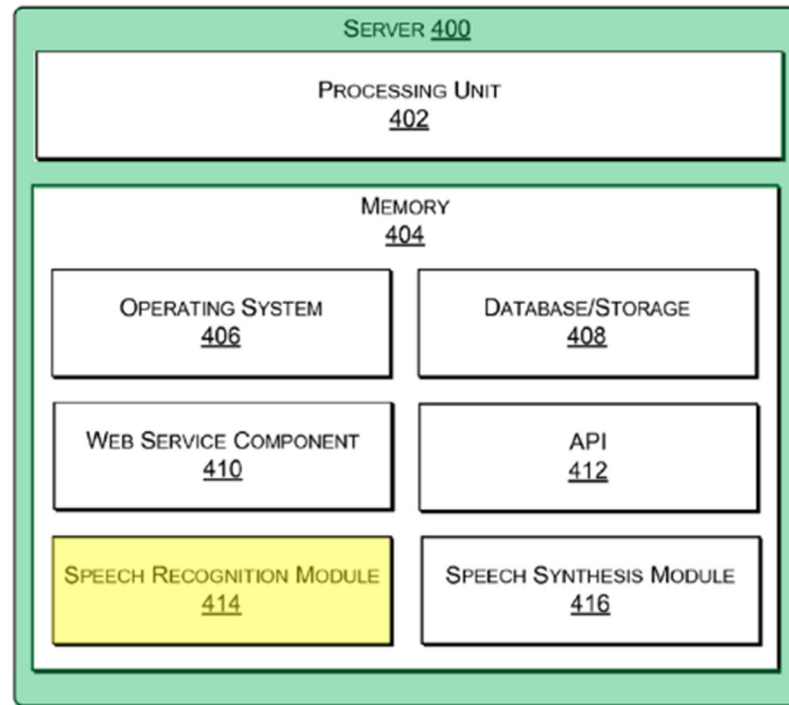


FIG. 4

h. 1[E] and 1[F]: “responsive to the processing, sending a first message, by the server to the client device over the Internet; receiving, by the client device via the wireless network, the first message;”

Williams discloses that the *server* sends the *first message* to the *client device* over the *Internet* and receiving the *first message* by the *client device* over the wireless network. (EX1002, ¶¶110, 115.)

Referring to FIG. 2, Williams’ *cloud server* “processes the *audio-based information* to determine various data about user activities, status, environment, commands, etc.” and “provide[s] *audio 222* for eventual transmission to the *interface device 116*, to be played on the speaker 206.” (EX1004, [0022], [0037].)

The system “may use the speech synthesis functionality of the system to provide *audio answers to spoken questions.*” (EX1004, [0048], [0045].) As shown in FIG. 2, *audio 222* is sent by the *cloud server* via the wireless network 114 (i.e., “the Internet”). (*Supra* §IV.A.2.d; EX1004, [0015], [0037], [0052].) Williams discloses several examples of providing to the user an *audio response* generated by the *cloud server*. (EX1004, [0048], [0050]; EX1002, ¶¶111-112, 115.)

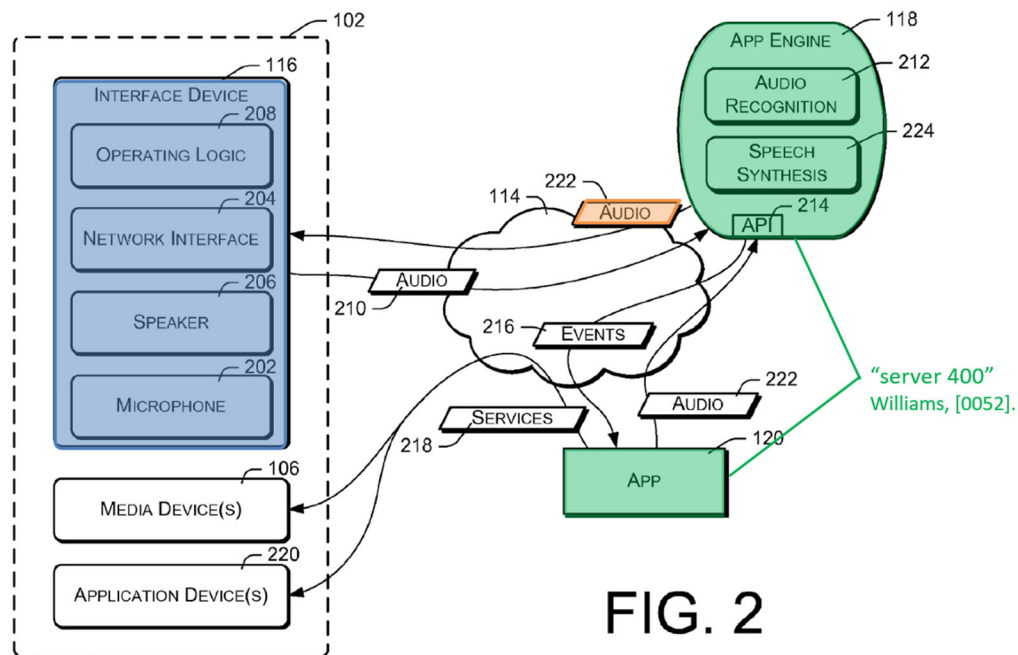


FIG. 2

Williams teaches that the first message is sent by the server “responsive to the processing” as recited. “The application engine 118 [run on *server 400*] *processes* the audio-based information to determine various data . . . 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]; EX1002, ¶113.)

FIG. 2 shows that the *interface device 116* receives the *audio 222*. It would have been obvious that the *audio 222* 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 *interface device* via the in-home wireless network. (EX1002, ¶¶114, 116-117.) Referring to FIG. 1, “the in-home devices connect to a router or access point 112, which is in turn connected to a public network 114. (EX1004, [0012], [0013].)

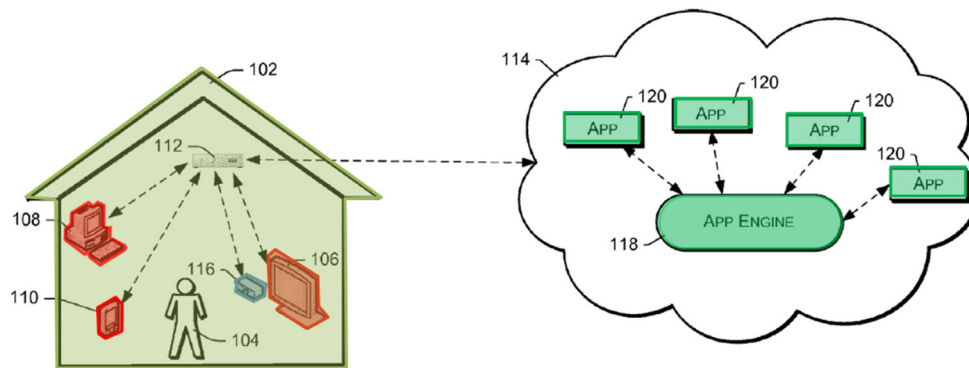


FIG. 1

i. 1[G]: “operating a first actuator in the client device in response to the received first message;”

The ‘710 patent explains “an actuator may be an *electroacoustic actuator*, *such as a sounder . . . A loudspeaker (a.k.a. speaker) is a sounder.*” (EX1001, 105:48-61.) Considering that disclosure, a POSITA would have understood that a

speaker is or has an actuator to produce sound. (EX1002, ¶¶120-122; e.g., EX1017, 122, EX1013, 763.).

Williams discloses operating a *first actuator/speaker* in the *client device* to produce sound in response to the received *first message*. For example, Williams discloses operating a *speaker* in the *client device* in response to the received *first message*. Specifically, the cloud server “provide[s] *audio 222* for eventual transmission to the *interface device 116*, to be played on the *speaker 206*.” (EX1004, [0037], [0045], [0048]; EX1002, ¶¶118-119, 123-124.)

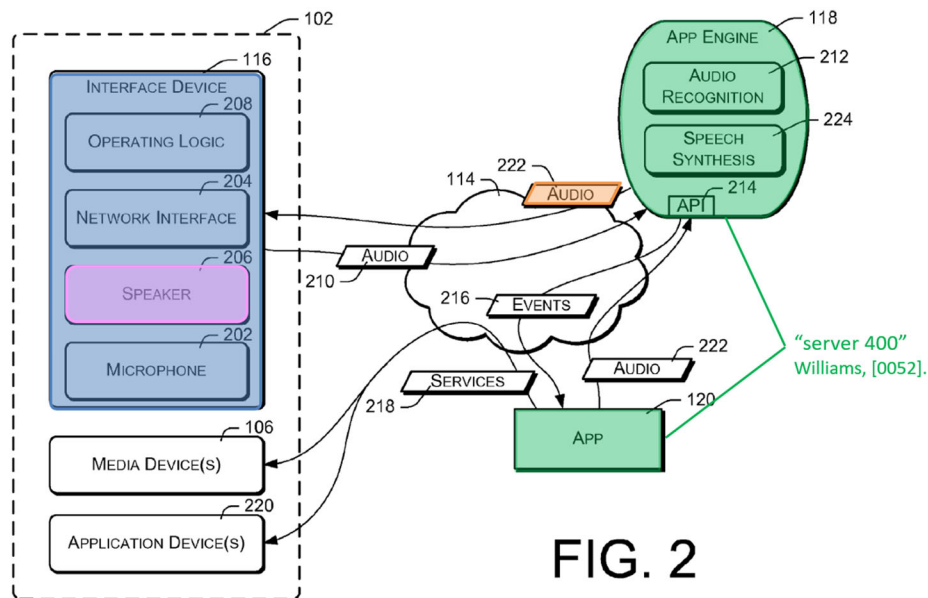


FIG. 2

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

Williams discloses capturing, by the microphone in the client device, a *second human voice data*. To the extent that the “*second human voice data*” is different from the “first human voice data”, Williams expressly discloses receiving

and processing *multiple* human voice commands. (EX1002, ¶¶125-128.) For example, “[a] cloud-based application may be configured to received *vocal commands* from users and to perform operations based on those commands” and “a *plurality of cloud-based applications or services 120 . . .* may communicate and operate in conjunction with a wide variety of network-connected devices.” (EX1004, [0044], [0021].)

Further, Williams discloses human voice commands that can effectuate remote device control. Williams “allow[s] a user to control various *on-premises devices and equipment, such as media equipment and home-automation devices.*” (EX1004, [0051].) 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].) Such spoken commands are “second human voice data.” (EX1002, ¶¶129-131.)

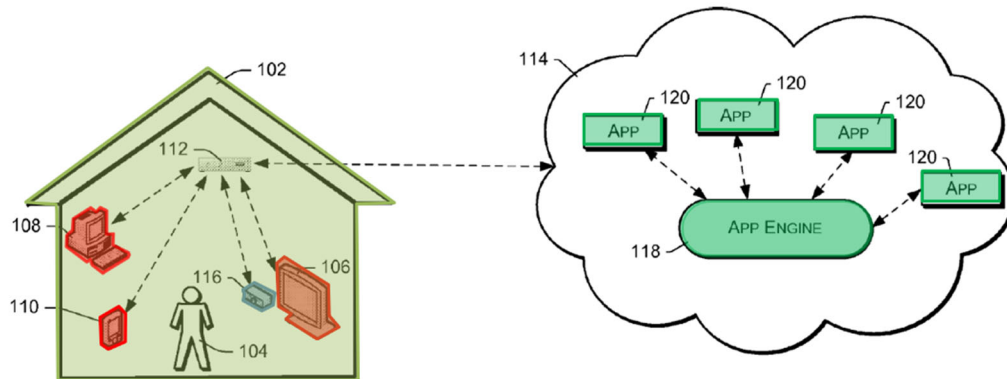
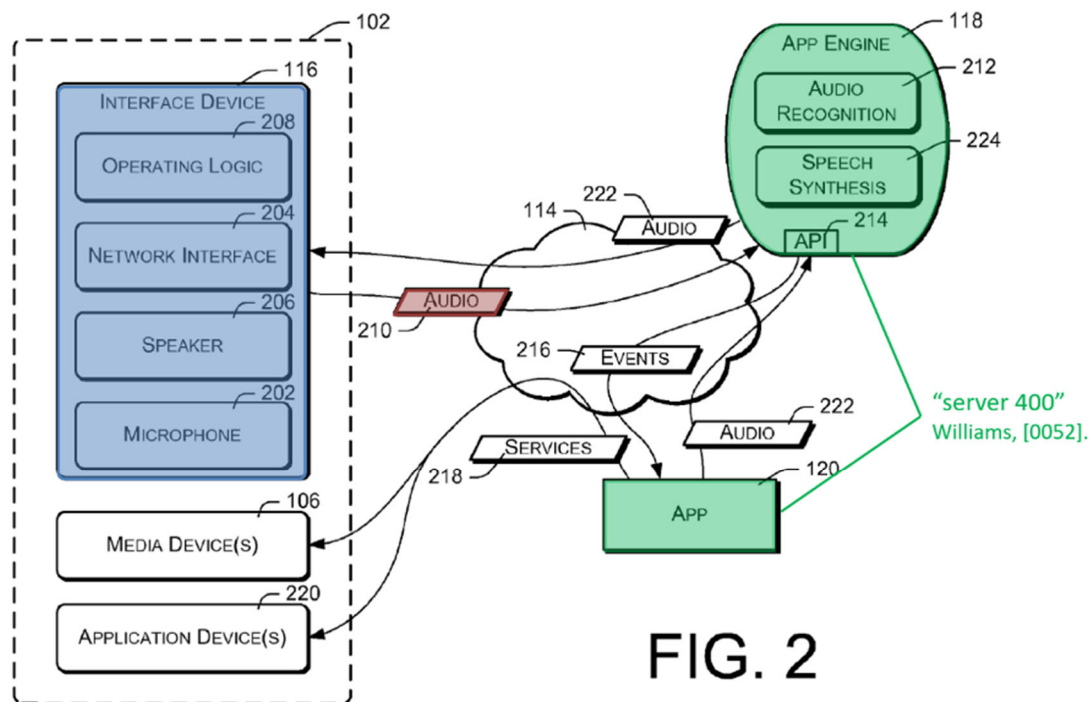


FIG. 1

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 and sent to the *server 400* for speech processing in the same manner as shown in FIG. 2.

(EX1002, ¶131.) Alternatively, the “audio 210” shown in FIG. 2 can include or be both the first and second human voice data because “audio 210” results in both first and second messages. (EX1002, ¶131.)



k. 1[I] and 1[J]: “sending to the server, by the client device via the wireless network, the captured second human voice data; receiving, by the server over the Internet, 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 via the wireless network

and the server receives the captured human voice data over Internet. (*Supra* §IV.A.2.f; EX1004, [0020], [0022], [0024], [0027], [0034], [0038]-[0039], [0052]; EX1002, ¶¶132-133.)

l. 1[K]: “processing, by the server, the captured second human voice data;”

Williams discloses processing, by the server, the captured second human voice data. The second human voice data can be one or more “*spoken commands* from a user regarding on-premises devices.” (EX1004, [0051].) Williams discloses processing of the captured human voice data at the server. (*Supra* §IV.A.2.g; EX1004, [0022], [0027], [0034], [0039], [0052].) Williams discloses sending, receiving, and processing multiple human voice commands. (*Supra* §IV.A.2.j; EX1002, ¶¶134-137.)

FIG. 3 is “a flow diagram illustrating an example process that may be implemented within the environment of FIGS. 1 and 2.” (EX1004, [0005].) “At 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, [0040], [0039], [0022], [0052]; EX1002, ¶138.)

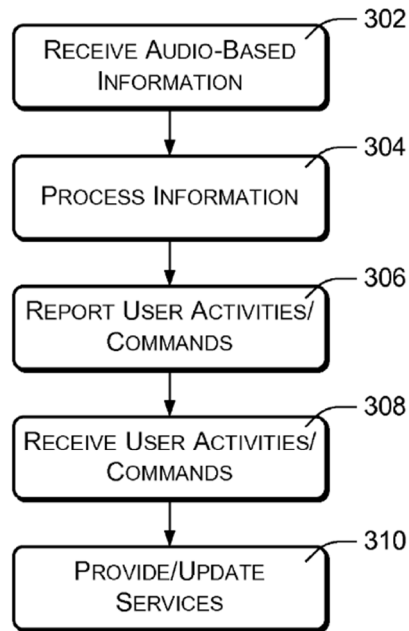


FIG. 3

m. 1[L] and 1[M]: “responsive to the processing, sending a second message, by the server to the controlled device over the Internet; receiving, by the controlled device via the wireless network, the second message;”

Williams discloses responsive to the processing, sending a *second message*, by the *server* to the *controlled device* over the Internet. Referring 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*” and that “[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], [0021], [0038], [0040], [0042], [0049], [0050]; EX1002, ¶¶139, 146-147.)

As shown in FIG. 2 below, William’s “server [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*. (EX1004, [0052]-[0056]; EX1002, ¶¶140, 147.)

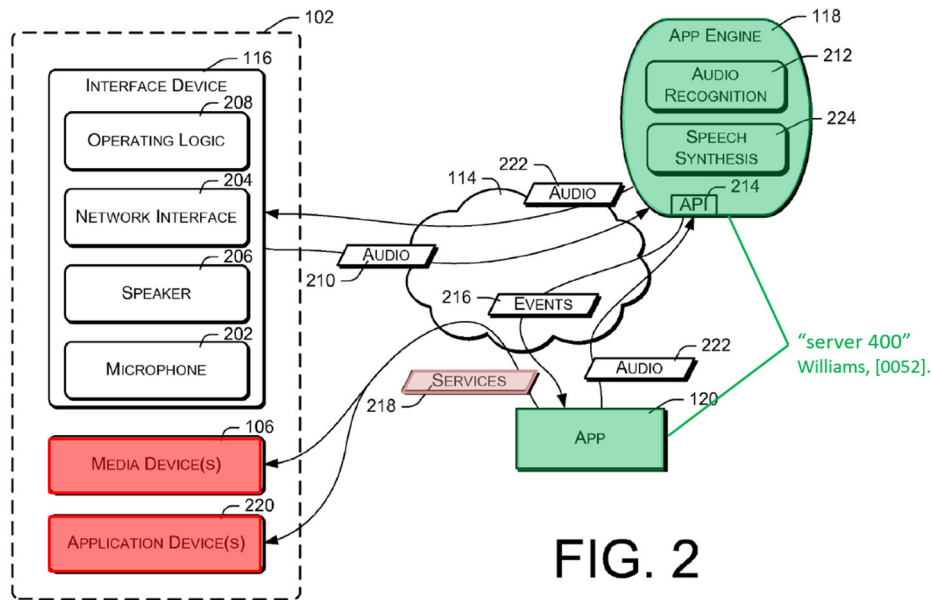


FIG. 2

Referring 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]; EX1002, ¶141.)

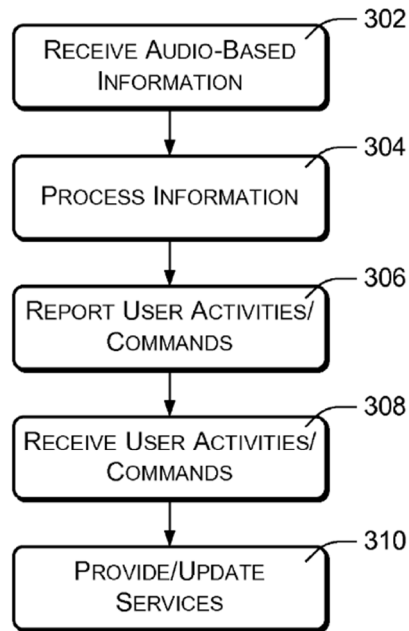


FIG. 3

A POSITA would have understood that providing *services 218* to a controlled device (e.g., *media devices 106, application devices 220*) involves sending information (second message) to the controlled device. (EX1002, ¶¶142, 148.) FIG. 2 expressly depicts that providing *services 218* involves transmitting information by the cloud-based application 120 [of the *server*] to the *media devices 106 and/or application devices 220*. A POSITA would have understood and found it obvious that, to operate a controlled device, information related to the services 218 (e.g., a control message) is sent to the controlled device, consistent with

common implementations. (E.g., EX1018, 3:19-22; EX1007, Claim 1; EX1011, [0012]; EX1002, ¶¶142, 148.)

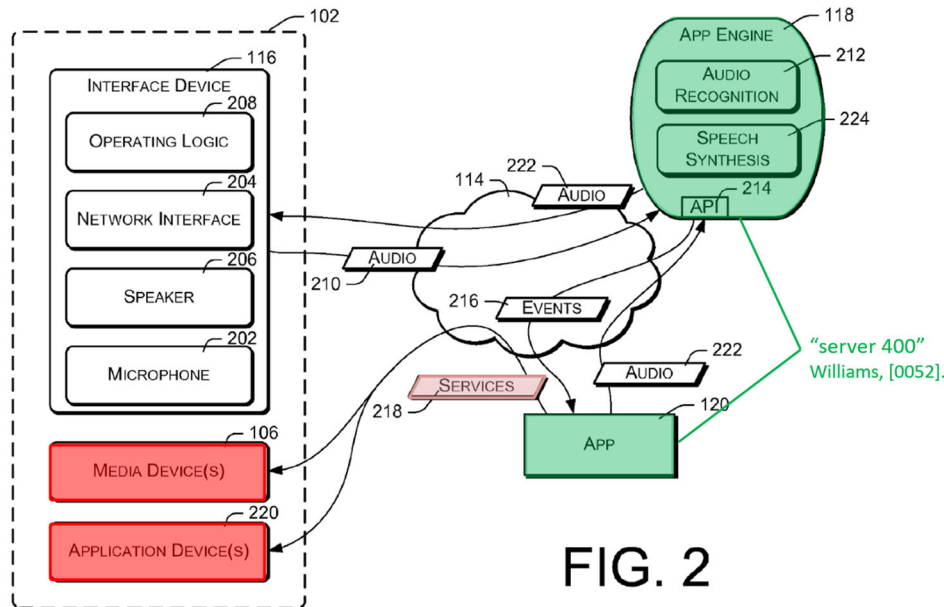


FIG. 2

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, [0042]; *supra* §IV.A.2.h; EX1002, ¶143.)

Williams discloses that the server sends a **second message** to the **controlled device over the Internet** and the **controlled device** receives the **second message** via the **wireless network**. As shown in FIG. 2, the **services 218** is sent via the wireless network 114 (e.g., the Internet) and received by **media devices 106 and/or application devices 220**. (*Supra* §IV.A.2.d; EX1004, [0015]; EX1002, ¶¶144, 149-150.)

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 in-home wireless network (e.g., through the home router). (EX1002, ¶¶145, 150.)

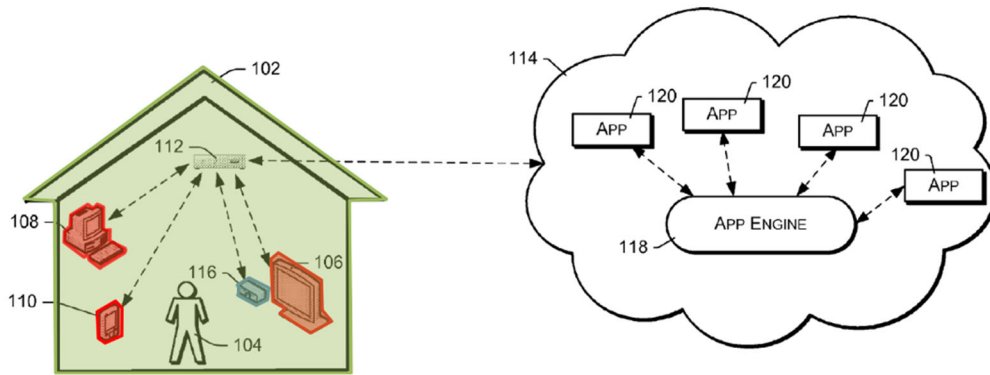


FIG. 1

n. 1[N]: “and operating a second actuator in the controlled device in response to the received second message.”

Williams discloses “allow[ing] a user to *control* various *on-premises devices and equipment*, such as *media equipment* and *home-automation devices*.”

(EX1004, [0051], claim 20; EX1002, ¶¶151-155.)

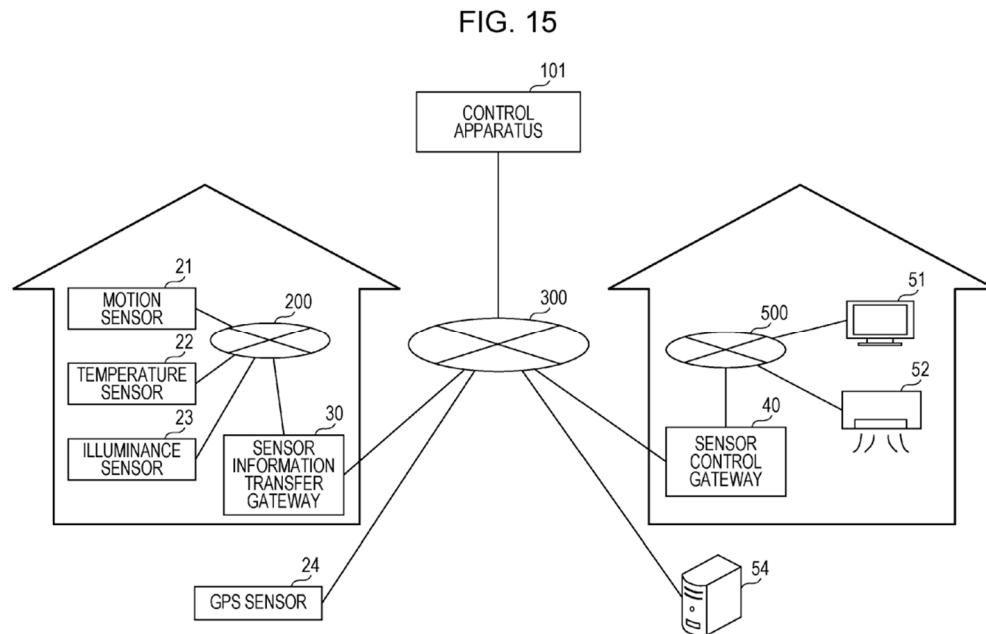
The ‘710 patent explains that a display is an actuator. “An actuator may be a display for presentation of visual data or information, commonly on a screen.” (EX1001, 25:33-38.) So is a speaker. (*Supra* §IV.A.2.i.) It would have been obvious that the second message would lead to operation of a display 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.

(EX1002, ¶¶156-158; EX1004, [0011], [0009]-[0010], [0021], [0036], [0054].)

Williams also discloses other use cases, such as “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].)

This is consistent with prior findings of the Board in its Decision on Appeal affirming the Examiner’s rejections 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; EX1002, ¶159.)



(EX1024, FIG. 15.)

Further, Williams discloses that “the applications 120 may communicate and *operate* in conjunction with . . . *home automation and security devices*.”

(EX1004, [0021], [0036], [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. (EX1002, ¶160; *e.g.*, EX1019, 401.) Indeed, the Board agreed in its Decision on Appeal 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; EX1002, ¶161.) It would have been obvious that the second actuator in the *controlled device* is operated in response to the information (second message) received from the *server 400*. (EX1002, ¶¶152-155.)

3. Claims 5 and 6

Operating speakers/home automation devices/media devices, as disclosed by Williams, involves operating an actuator affecting, changing, producing, or creating sound or audio, light, motion, temperature, humidity, etc. (*Supra* §§IV.A.2.i, IV.A.2.n; EX1002, ¶¶162-166.)

4. Claims 16 and 23

The controlled device in Williams includes devices such as “*televisions, audio players, video players*” and “*home-automation devices*”. (*Supra* §IV.A.2.n; EX1004, [0051], [0011]; EX1002, ¶¶165-166.) These are “appliances” consistent with claim 23. *Id.*

These also disclose a “home cinema system” recited in claim 23. (EX1002, ¶166.) Further, it would have been obvious to a POSITA that Williams’ “video players” and “audio players” include DVD players and CD players, as CDs had been released commercially as early as 1982. (EX1013, 746; EX1002, ¶166.)

5. Claim 29

Williams teaches “playing audio 222 on the *speaker 206*.” (*Supra* §§IV.A.2.i and IV.A.3; EX1004, [0037]; EX1002, ¶167.) 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,” and “an application may respond ... by using the *speech synthesis functionality* of the system to *query the user regarding movie selections and time*” (*Id.*, EX1004, [0037], [0045]-[0046] and [0050].)

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

1. Summary of Wittenberg

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.) Wittenberg discloses “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, Abstract, Claim 1; EX1002, ¶168.)

2. Claim 2

Claim 1 is obvious in view of Williams. (*Supra* §IV.A.2.) Claim 2 is obvious in further view of Wittenberg. (EX1002, ¶169.) Wittenberg discloses that “[v]oice recognition processor analyzes audio signals received at the remote processor to determine whether the person issuing the voice command is

authorized to issue the command (operation 250 in FIG. 2). The processing involves comparing the characteristics (*e.g.*, frequency spectrum, speech patterns, etc.) of the audio signal to stored voice information to ***determine whether the audio signal voice matches that of an authorized user.***” (EX1007, 5:61-6:6; EX1002, ¶169.)

Williams and Wittenberg 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. (EX1004, [0036]; EX1007, Claim 1; EX1002, ¶170.)

The Williams-Wittenberg combination, for example, would determine if the person issuing the voice data is authorized to make the command as a part of processing at the server as taught by Wittenberg. (EX1002, ¶174.) A POSITA would have been motivated to combine Williams with the voice recognition of Wittenberg, for example, to achieve greater security and with a reasonable expectation of success. (EX1002, ¶¶170-175.) Williams discloses that “[a]uthorization and authentication safeguards are employed to ensure that only entities authorized by users are able to access this information.” (EX1004, [0008].) Wittenberg teaches the benefits for including voice recognition during processing as “[v]oice 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.)

Wittenberg recognizes the benefits of applying its techniques in home appliances.

(EX1007, 6:7-13.) Wittenberg explains that “[a]lgorithms and processing techniques for performing voice recognition are well known” (EX1007, 6:1-3; EX1002, ¶175.)

C. Ground 3: Williams and Gruber Render Obvious Claims 3-4, 10-11

1. Summary of Gruber

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]; EX1002, ¶¶176-177.)

2. Claim 3

Claim 1 is obvious in view of Williams. (*Supra* §IV.A.2) Claims 3/4, 10/11 are obvious in further view of Gruber. (EX1002, ¶178, 186, 188, 193; *see supra* §IV.A.2.) 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 . . . 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.” (*Id.*, EX1008, [0218]-[0225].)

Gruber further discloses that the contextual information provided by sensors can be transmitted to a speech processing server to process human voice data. “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], [0168], [0080]; EX1002, ¶¶178-180.)

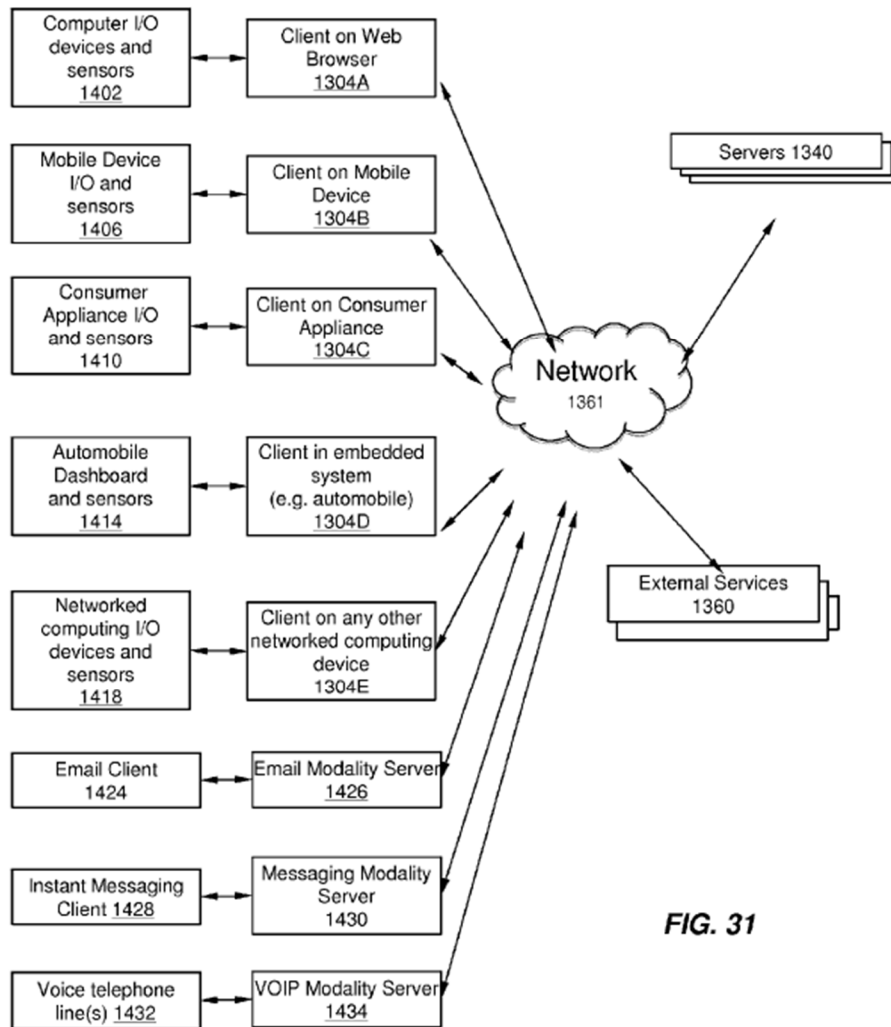


FIG. 31

Williams and Gruber are analogous art to the '710 patent. Both relate to processing and responding to human speech commands. (EX1004, [0027], [0036]; EX1008, [0007]; EX1002, ¶181.)

The Williams-Gruber combination, for example, would capture sensor data (*e.g.*, temperature) using one or more sensors (*e.g.*, as taught by Gruber) in the audio interface device and transmit such data to Williams' server to supplement audio processing as taught by Williams and/or Gruber. (EX1004, [0013]; EX1008,

[0013]; EX1002, ¶¶182-183.) A POSITA would have been motivated to combine Williams and Gruber to include the speech processing techniques of Gruber, for example, to achieve improved speech recognition and with a reasonable expectation of success. (EX1002, ¶¶181-185.) Gruber teaches additional speech processing techniques of using contextual information obtained from device sensors to “supplement natural language or gestural input from a user,” “clarify the user’s intent”, and “reduce[] the need for the user to provide excessive clarification input.” (EX1008, Abstract.)

3. Claim 4

Gruber discloses various sensors including “*light sensor, thermometer, . . . and the like.*” (EX1008, [0225].) The contextual information from these sensors may include “the *current light level, temperature, and other environmental measures.*” (EX1008, [0222].) A POSITA would have understood that photoelectric sensors are common light sensors and thermoelectric sensors are common temperature sensors. (EX1029, 116, 466; EX1002, ¶186.) It would have been obvious to a POSITA to combine the teachings of Williams and Gruber with a reasonable expectation of success. (*Supra* §IV.C.2; EX1002, ¶187.)

4. Claims 10 and 11

Williams discloses that the interface device 116 and home devices communicate with a wireless network (*e.g.*, via a router). (*Supra* §IV.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. (EX1002, ¶¶188-189, 193.) The address may be an assigned IP address to the device (*e.g.*, assigned by the router), the router's IP address, a combination of the two, or a globally unique IP address. (EX1002, ¶189.) In 2012, it was well-known that networked devices are associated with IP addresses that are layer 3, either static or dynamic, of IPv4 type, and/or MAC addresses based on MAC-48 or EUI-48, both of which uniquely identify devices in the network. (EX1002, ¶¶189-190, 194; *e.g.*, EX1006, 99, 108, 166, 210-211, 175, 265, 277; EX1028, 127; EX1046, 2:50-59; EX1022, 11, 13.)

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]; EX1002, ¶191.)

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 with a reasonable expectation of success. (*Supra* §IV.C.2; EX1002, ¶¶192, 195.)

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

1. Summary of Benesty

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, a textbook titled “Springer Handbook of Speech Processing”, provides an overview of well-known speech processing techniques. (EX1009, 1021; EX1002, ¶¶196-199.)

Benesty was published by Springer and was well-known among POSITAs. (EX1002, ¶200.) A POSITA would have looked to Benesty for microphone-related techniques to, for example, improve the audio signal capturing and processing in Williams with a reasonable expectation of success. (*Id.*)

2. Claims 7 and 8

Claim 1 is obvious in view of Williams. (*Supra* §IV.A.2) Claims 7/8 are obvious in further view of Benesty. (EX1002, ¶¶201-202; *see supra* §IV.A.2.) Benesty discloses that “[i]t is shown that *microphone arrays* have the ability to offer directional gains that can *significantly improve the quality of signal pickup in reverberant and noisy environments.*” (EX1009, 1021.)

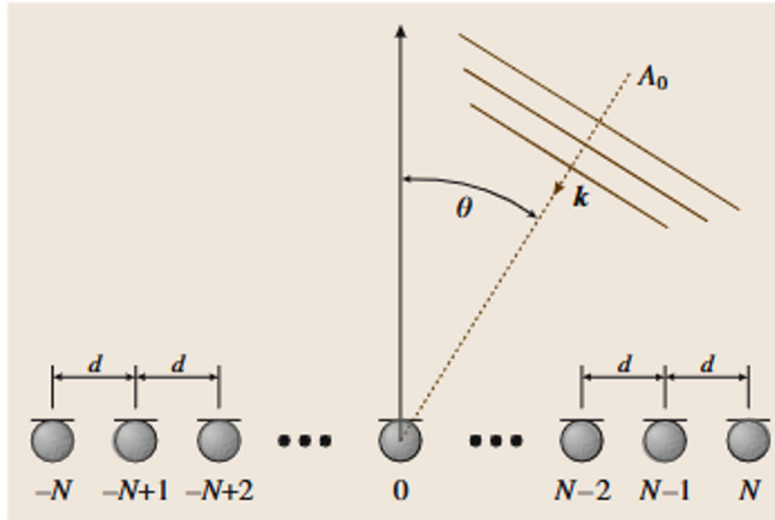


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

Benesty discloses that “[f]or speech source localization, microphone array beam scanning and *time delay estimation* (TDE)-based localization methods are the two most widely used approaches . . . steering a microphone array beamformer and scanning across a room for the highest-energy output leads to an *estimate of the direction* of an active speech source.” (EX1009, 1044.) Benesty also discloses that “a set of relative time differences of arrival (TDOAs) for different microphone pairs are calculated.” (EX1009, 1044; EX1002, ¶¶206-207.)

Williams and Benesty are analogous art to the '710 patent. Both relate to capturing and processing audio. (EX1004, [0003], Figure 1; EX1009, 1021.) A POSITA would have been motivated to combine Williams and Benesty to include the microphone array techniques of Benesty, for example, to improve audio capturing and processing and with a reasonable expectation of success.

(EX1002, ¶¶203-205, 208.) 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 and process human voice data, yielding the predictable results of effectively improving audio capturing, especially in a room where the microphone is not placed near the mouth of the user. (EX1002, ¶¶204-205.)

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

1. Summary of Goldsmith

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, a textbook titled “Wireless Communications,” was published by Cambridge University Press and was well-known among POSITAs. (EX1035, Title; EX1002, ¶¶209-214.) A POSITA would have looked to Goldsmith for wireless communication systems, protocols, and techniques of applying them. *Id.*

2. Claim 12

Claim 1 is obvious in view of Williams. (*Supra* §IV.A.2) Claims 12-15 are obvious in further view of Goldsmith. (EX1002, ¶¶215, 226, 232, 235; *see supra* §IV.A.2.) Williams discloses devices (*e.g.*, interface devices and in-home devices) communicating over a wireless network, using “*cellular data technologies or*

other wireless data communications technologies” and “existing networking protocols and techniques.” (EX1004, Figure 1, [0012], [0013], [0020], [0038]; EX1002, ¶¶215-216.)

Goldsmith “summarizes the main characteristics of Zigbee, Bluetooth, and UWB...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, 601-602.) “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; EX1002, ¶¶217-219.)

“Bluetooth radios provide *short range connections between wireless devices* along with rudimentary networking capabilities.” (EX1035, 19.) “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*.” (EX1035, 20; EX1002, ¶¶220-221.)

At the time of the invention, routers having Bluetooth capabilities and Bluetooth access points/gateways with connection to the Internet were well-

known. (EX1025, 59:21-24; EX1026, 4:25-27; EX1027, 4:20-22; EX1032, 1:67-2:3.) For example, access to a remote Internet-connected server may be “via a local network connection 508, such as WiFi or *Bluetooth*, through a *gateway device such as a router*.” (EX1020, [0027]; EX1002, ¶222.)

It would have been obvious for a POSITA to combine Goldsmith and Williams and look to Goldsmith for conventional wireless networking techniques and protocols that can be applied to Williams, which expressly discloses usage of “cellular or other wireless data communications technologies”. (EX1002, ¶¶223-225; EX1004, Figure 1, Figure 2, [0012], [0013], [0020], [0038].) For example, Goldsmith discloses that it is feasible to embed Bluetooth and Zigbee in electronic devices to create “smart homes.” 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. (EX1002, ¶225.)

3. Claim 13

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. (EX1035, 15; EX1002, ¶¶226-227.)

Further, “[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

operates with 300 MHz of spectrum in the 5 GHz U-NII band . . . 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.***” (EX1035, 16; EX1002, ¶228.)

It would have been obvious to combine Williams and Goldsmith with reasonable expectation of success. (*Supra* §IV.E.2; EX1002, ¶¶229-231.)

4. Claim 14

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.” (EX1035, 15.) “The ZigBee radio specification . . . is based on the IEEE 802.15.4 standard. The ***radio operates in the same ISM band as Bluetooth.***” (EX1035, 20; EX1002, ¶232.)

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; EX1002, ¶233.)

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, 802.11a WLANs)	5.25–5.35 GHz
U-NII band III (long-range outdoor systems, 802.11a WLANs)	5.725–5.825 GHz

As discussed above, a POSITA would be motivated to combine Williams and Goldsmith. (*Supra* §§IV.E.2 - IV.E.3; EX1002, ¶234.)

5. Claim 15

Goldsmith discloses the recited conventional wireless network communication protocols, for example, “[t]he *third generation (3G) cellular systems are based on a wideband CDMA standard.*” (EX1035, 12-13.) “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.” *Id.* “*Enhanced Data Services for GSM Evolution (EDGE)*, further increases data rates using a high-level modulation format combined with FEC coding.” (*Id.*; EX1002, ¶¶235-238.)

In addition, Goldsmith also discloses usage of W-CDMA as a 3G standard. “*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; EX1002, ¶229.)

Table D.4: 3G digital cellular phone standards							
Parameter	cdma2000				W-CDMA		
	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.

At the time of the invention, home appliances with cellular connectivity were well-known. (EX1030, claims 14-16, 18; EX1002, ¶240.)

It would have been obvious to combine Williams and Goldsmith. (*See supra* §IV.E.2; EX1002, ¶¶241-242.)

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

1. Summary of Chatterjee

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; EX1002, ¶¶243-246.)

2. Claims 17-22

Claim 1 is obvious in view of Williams. (*Supra* §IV.A.2) Claims 17-22 are obvious in further view of Chatterjee. (EX1002, ¶247.) Williams discloses controlling various on-premises devices and equipment such as media equipment and home-automation devices. (EX1004, [0051], [0012], [0007].) Similarly, Chatterjee describes that a variety of 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 120 α -120c or an integrated unit of the home appliance device 120 α -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.*” (EX1011, [0012]; EX1002, ¶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].) “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]; EX1002, ¶¶248, 253-254, 256-257, 260-261, 264-265.)

Further, FIG. 2 depicts a “smart refrigerator 220e”, “HVAC 120c”, and “a smart dish washer or washing machine 220d.” (EX1011, [0014]; EX1002, ¶¶249, 258, 262, 266.)

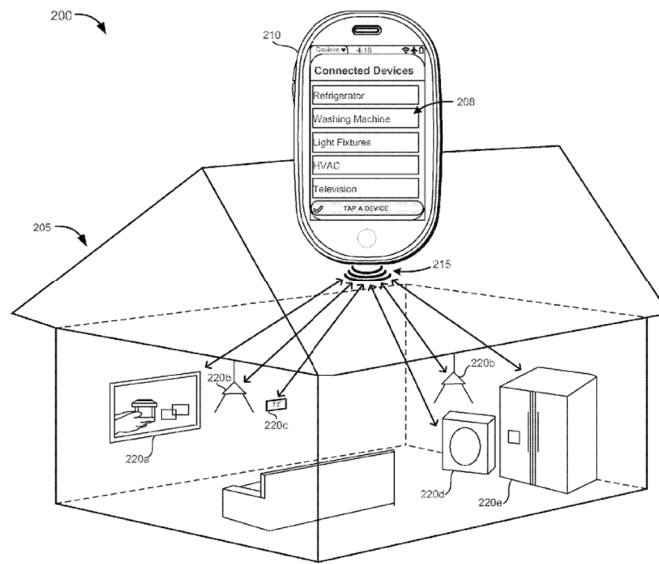


FIG. 2

Williams and Chatterjee analogous art to the '710 patent. (EX1011, Abstract; EX1004, [0007], [0051]; EX1001, Claim 1, 135:54-60.) A POSITA would have been motivated to combine Williams with Chatterjee's teachings regarding control of automated devices, for example, to expand the capabilities of Williams with a reasonable expectation of success. (EX1002, ¶¶250-252, 255, 259, 263, 267, 270.)

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

1. Summary of Rosenberger

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.) “Preferably, device D also includes ... a *multi-colored LED status light 34 for visually indicating various device states and modes.*” (EX1012, 8:8-16; EX1002, ¶¶271-273.)

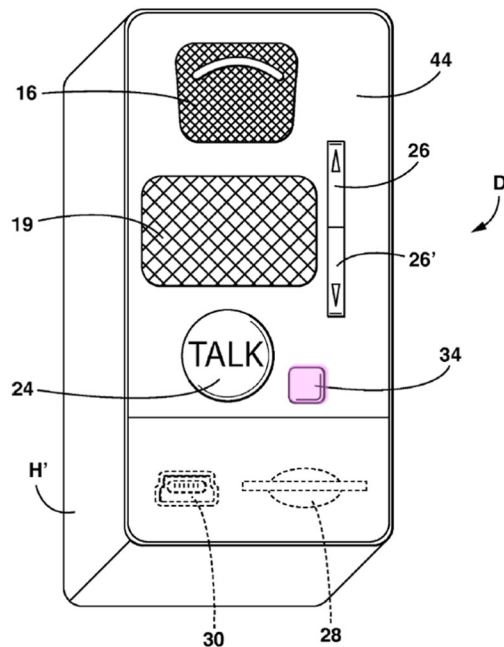


FIG. 2B

2. Claims 24 and 25

Claims 24-25 are obvious in further view of Rosenberger. (EX1002, ¶¶274, 281; *see supra* §IV.A.2.) Rosenberger discloses 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.*” (*Id.*, EX1012, 8:8-16.)

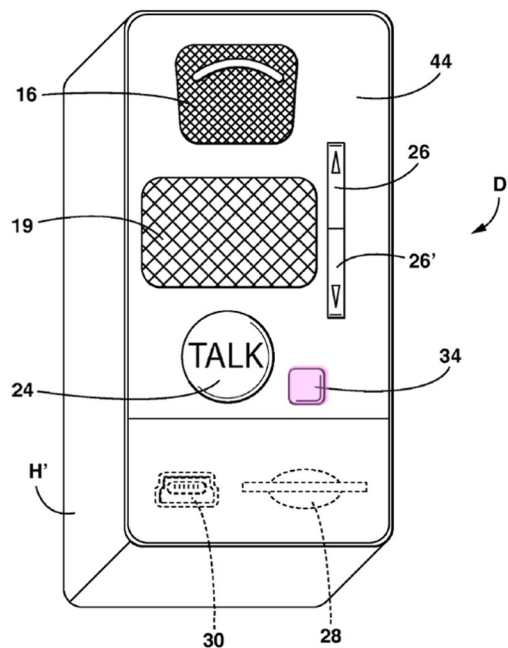


FIG. 2B

Rosenberger discloses that “[u]pon receiving and recognizing a speech trigger phrase . . . 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; EX1002, ¶275.)

Rosenberger and Williams are analogous art to the ’710 patent. (EX1002, ¶276.) Both relate to processing and responding to audio commands. (EX1004, [0027], [0036]; EX1012, 1:7-13.) A POSITA would have been motivated to combine Williams with Rosenberg to include Rosenberg’s 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, for example, to improve the usability of the device, and with a reasonable expectation of success. (EX1002, ¶¶276-280, 282.)

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

1. Summary of Rossing

Rossing was published by at least July 23, 2007, and is prior art under §102(a) and (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. (EX1002, ¶283-289.) Rossing was published by Springer and was well-known among POSITAs. (EX1002, ¶290.) A POSITA would have looked to Rossing for detailed specifics and techniques of using microphones and speakers in the system disclosed in Williams. *Id.*

2. Claim 9

Claim 1 is obvious in view of Williams. (*Supra* §IV.A.2.) Claims 9, 26-28 are obvious in further view of Rossing. (EX1002, ¶¶291, 299, 303, 306; *see supra* §IV.A.2.)

A POSITA would have understood that the claim refers to a standard microphone. (EX1002, ¶291.) Williams discloses use of such a **microphone** in the *client device*, consistent with the description of the '710 patent. (*Id.*; EX1004, [0016], [0007], [0019], claims 1-3 and 21; EX1001, claim 9, 17:47-54.)

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” and that “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; EX1002, ¶292.)

Rossing also discloses other variety of microphones such as the “ribbon microphone and the condenser microphone.” (EX1013, 758.) Rossing discloses piezoelectric microphones and carbon microphones. (EX1013, 758; EX1002, ¶293-294.)

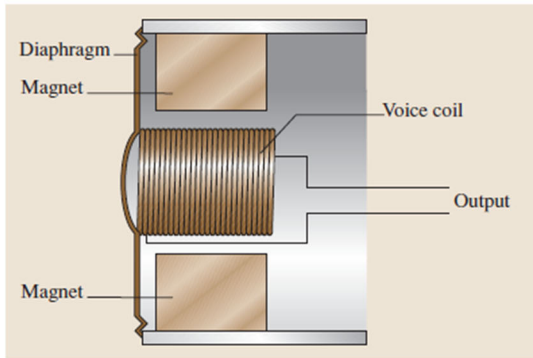


Fig. 18.10 A dynamic microphone uses a coil of wire attached to a diaphragm, suspended in a magnetic field (After [18.24])

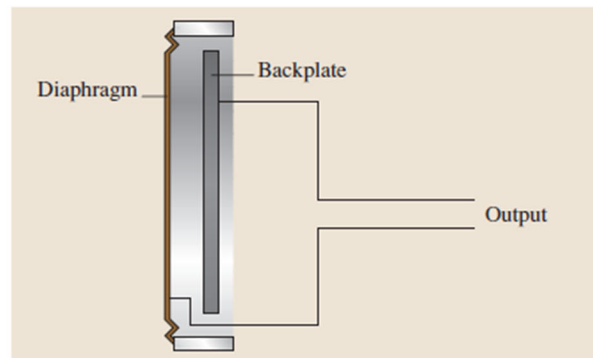


Fig. 18.9 A condenser microphone consists of a very thin diaphragm suspended parallel to a backplate (After [18.24])

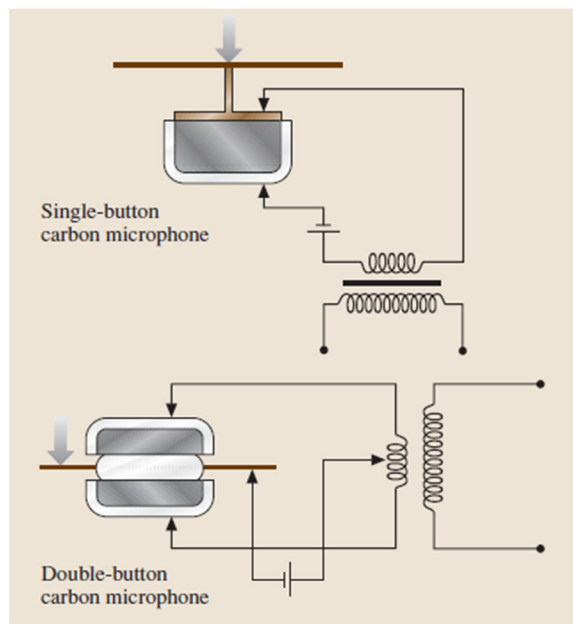


Fig. 18.11 A carbon microphone (After [18.25])

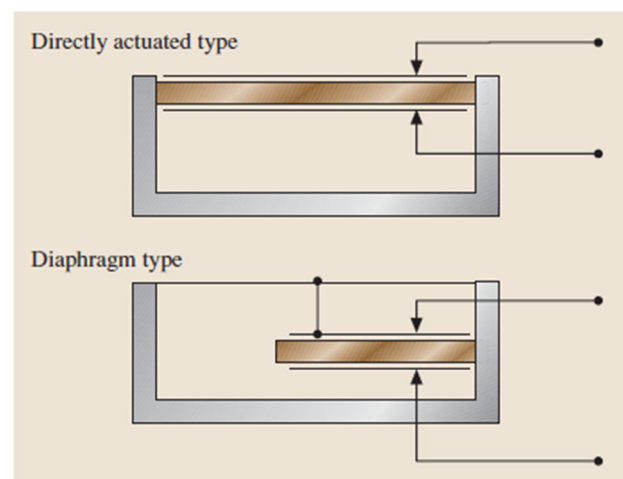


Fig. 18.12 Piezoelectric microphones (After [18.29])

Further, Rossing discloses that “[a] popular modern variant is the *electret condenser microphone*, perfected in 1962 by Sessler and West of Bell Laboratories . . . Such microphones also usually exhibit good sample-to-sample consistency, making them *practical for multi-microphone arrays*.” (EX1013, 759.) Rossing also discloses “an *ordinary omnidirectional microphone*”. (EX1013, 309; EX1002, ¶¶295-296.)

Rossing is a textbook providing basic information how microphones and speakers operate. (EX1002, ¶297.) Thus, a POSITA would have been motivated to combine Rossing and Williams and look to Rossing for techniques and details regarding these components for implementing Williams. (EX1002, ¶¶297-298.) For example, it would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Williams and with a reasonable expectation of success. (EX1002, ¶298.)

3. Claims 26-28

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* §IV.A.2.i; EX1004, [0048]; EX1002, ¶299.)

Regarding claim 26, Rossing discloses that a speaker is a motion actuator where “the electrical energy is converted to mechanical energy, specifically the movement of a diaphragm.” (EX1013, 763.) Further, Rossing discloses that loudspeakers have “linear motion” (EX1013, 903) and the “dynamic or moving-coil loudspeaker” has a coil that can “move axially.” (EX1013, 18; EX1002, ¶¶300-301.)

Regarding claims 27 and 28, a POSITA would have understood that the claim refers to a speaker. (EX1002, ¶¶303, 306.) William teaches operating a *speaker 206* consistent with descriptions of the '710 patent. (*Supra* §IV.A.2.i;

EX1001, 105:48-52, 105:61-106:4; EX1004, [0037]; EX1002, ¶¶303, 306.) 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. (EX1017, 122; EX1013, 763; EX1002, ¶¶303-304.)

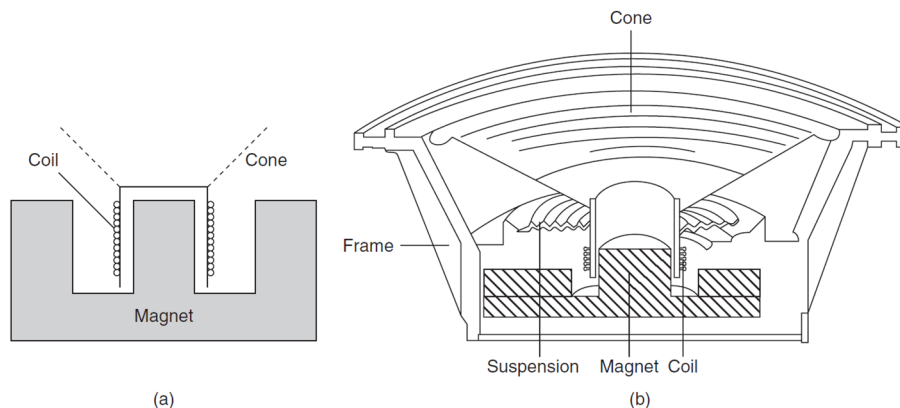


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

(EX1017, 123, Figure 10.8.)

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. (EX1002, ¶305; EX1013, 765.) Rossing states that “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; EX1002, ¶307.) “In 1929, E. W. Kellogg filed a patent for an *electrostatic loudspeaker . . .*” (*Id.*) It would have been obvious to combine Williams and Rossing. (EX1002, ¶¶302, 305, 308-309.)

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

1. Summary of Howard

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 is prior art under §102(a), (b), and (e). (EX1005, 1.) The Examiner never found, and Applicant never cited, Howard during prosecution. Howard discloses “a portable device readily carried by the operator. . .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; EX1002, ¶310.)

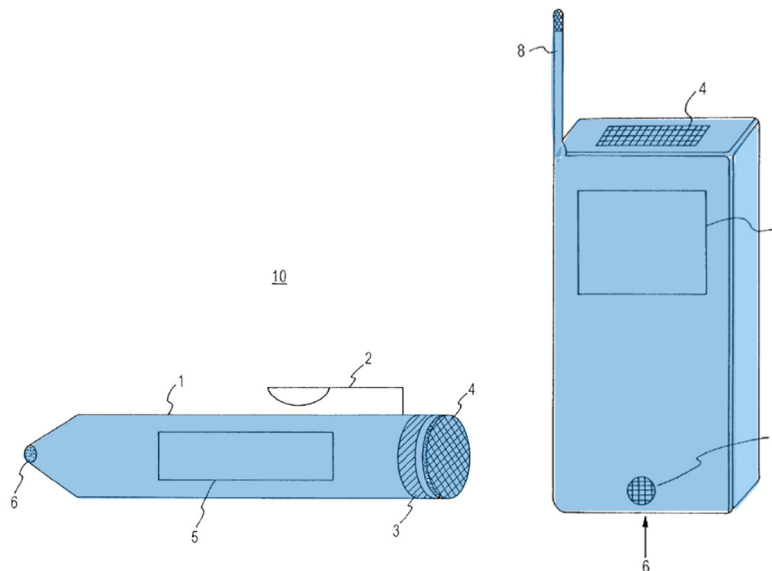


FIG. 1(a)

FIG. 1(b)

FIG. 9 illustrates an exemplary workflow. As shown, a user can utter commands to the *portable device*. These commands are transmitted to and processed by a *server*. (EX1005, 11:20-38.) When the user utters “TV”, the *server* can process the command and generate a message (“*What Can I do for you?*”), which is transmitted to and played back at the *portable device*. (*Id.*, 11:27-47.) If the user utters “[a] little small”, the *server* can process the command and generate *control signals*, which are transmitted to a *VTR device* to change the volume. (EX1005, 11:20-12:14; EX1002, ¶¶311-313.)

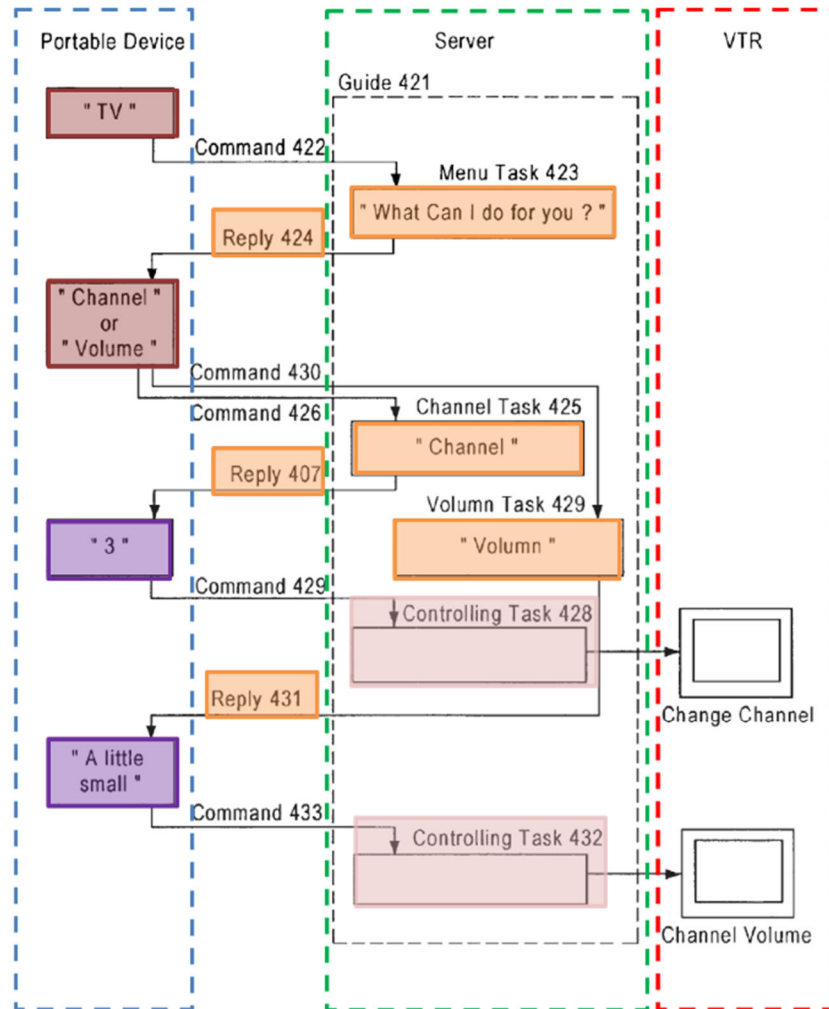


FIG. 9

2. Summary of Doherty

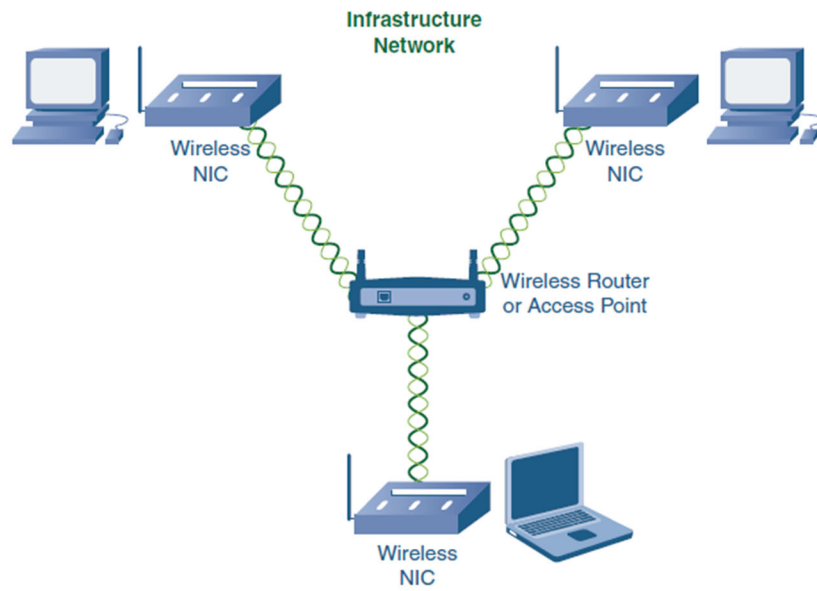
Doherty was published by at least January 16, 2007, and is prior art under §102(a) and (b). (EX1033, 1.) The Examiner never found, and Applicant never cited, Doherty during prosecution. Doherty is a textbook titled “Wireless Home Networking Simplified” published by Cisco. (EX1006, Title.) Doherty teaches readers to set up wireless networks at home using conventional products.

(EX1006, Introduction, Chapter 6; EX1002, ¶314.) A POSITA would have looked to Doherty for wirelessly connecting in-home devices. (EX1002, ¶315.)



Doherty discloses that a wireless router provides a wireless network for in-home devices, such as “daughter’s laptop” and “son’s desktop computer” in the same house. (EX1006, 39.) Further, the router connects the wireless network with the Internet: “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; EX1002, ¶¶316-324.)

Figure 4-2 Peer-to-Access Point Wireless Network



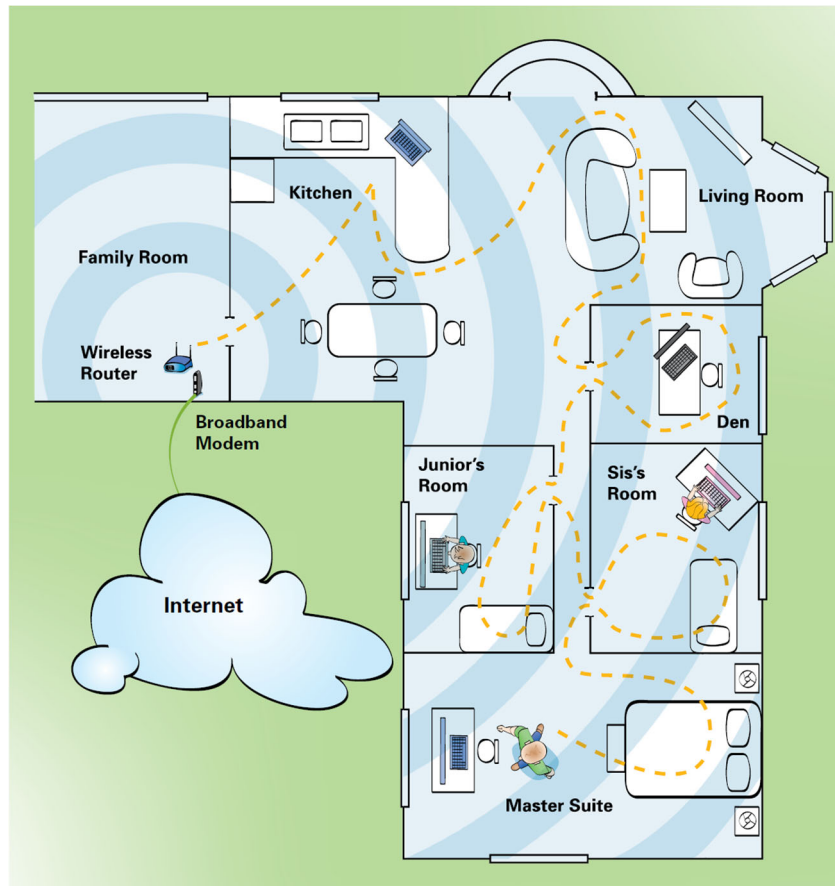
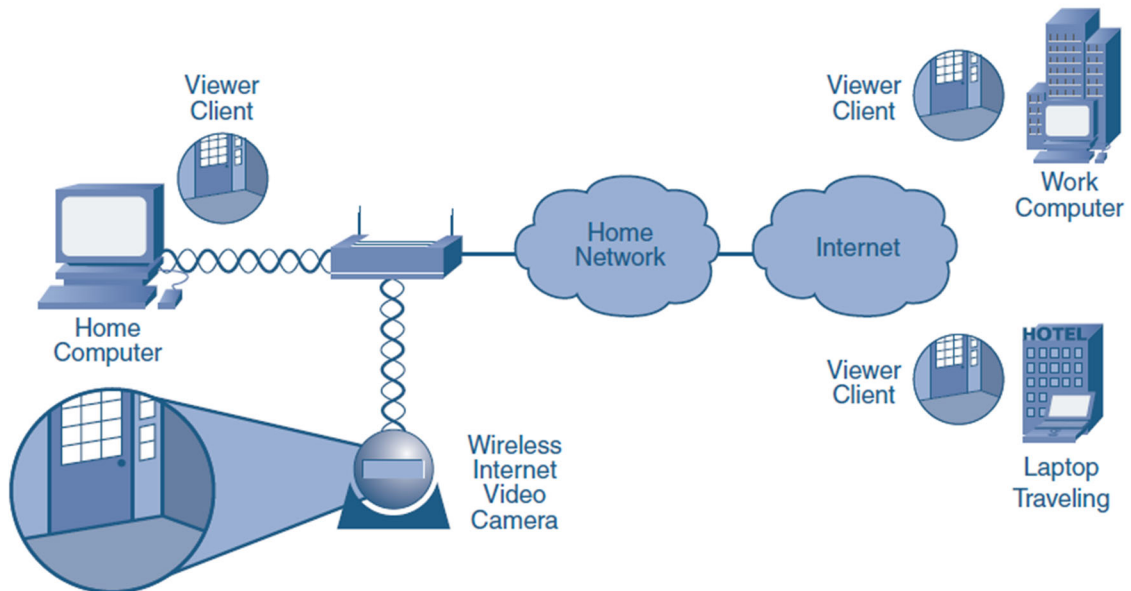


Figure 13-1 Home Surveillance Using a Wireless Internet Video Camera



3. Claim 1

a. 1[PR][1]

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.) 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* IV.I.3.i and IV.I.3.m; EX1002, ¶¶325-328.)

b. 1[PR][2]

Howard discloses a *client device* and a *controlled device* in a *building*. (EX1002, ¶329.)

For example, FIGs. 1(a) and 1(b) illustrates “exemplary configuration[s] of *a portable device*.” (EX1005, 2:38-43.) Referring to FIG. 2, “the *portable device 10* comprises both a transmitter system 101 and a receiver system 102” and communicates wirelessly with *server* 200. (EX1005, 4:2-24.) Thus, Howard’s portable device is a client device. (EX1002, ¶¶330-331.)

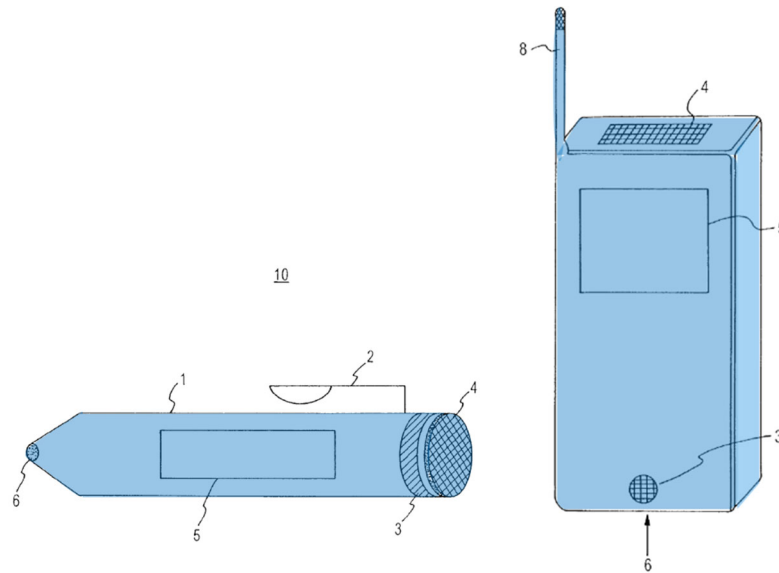


FIG. 1(a)

FIG. 1(b)

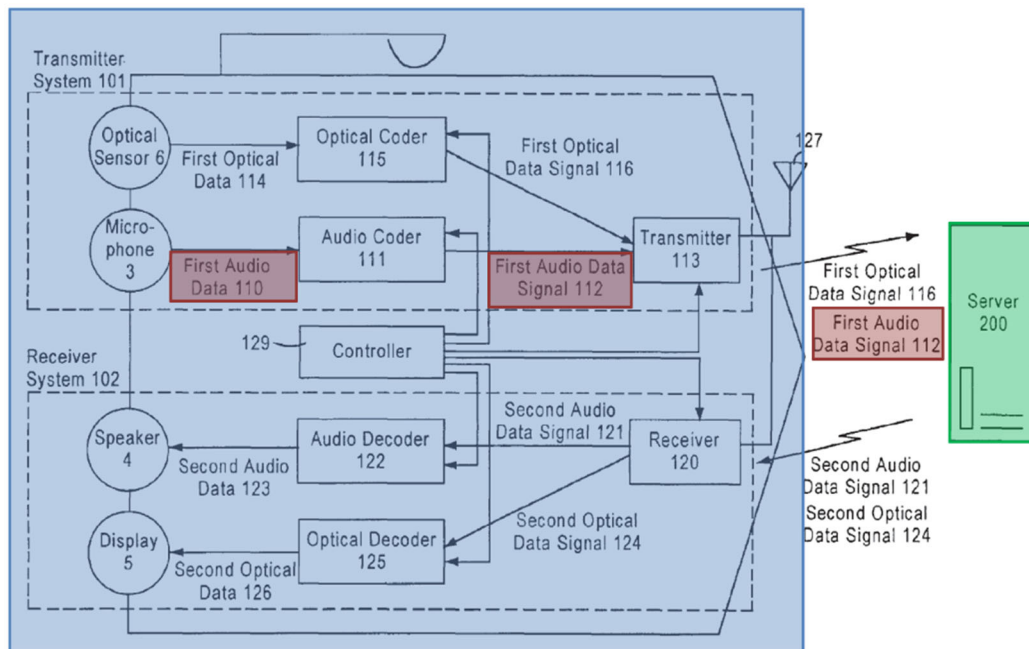


FIG. 2

Howard discloses a **controlled device**. “[B]y 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.) “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.*” (Id., 1:18-30.) Additional examples are shown in FIG. 6. (EX1005, 9:13-23; EX1002, ¶¶332-333.)

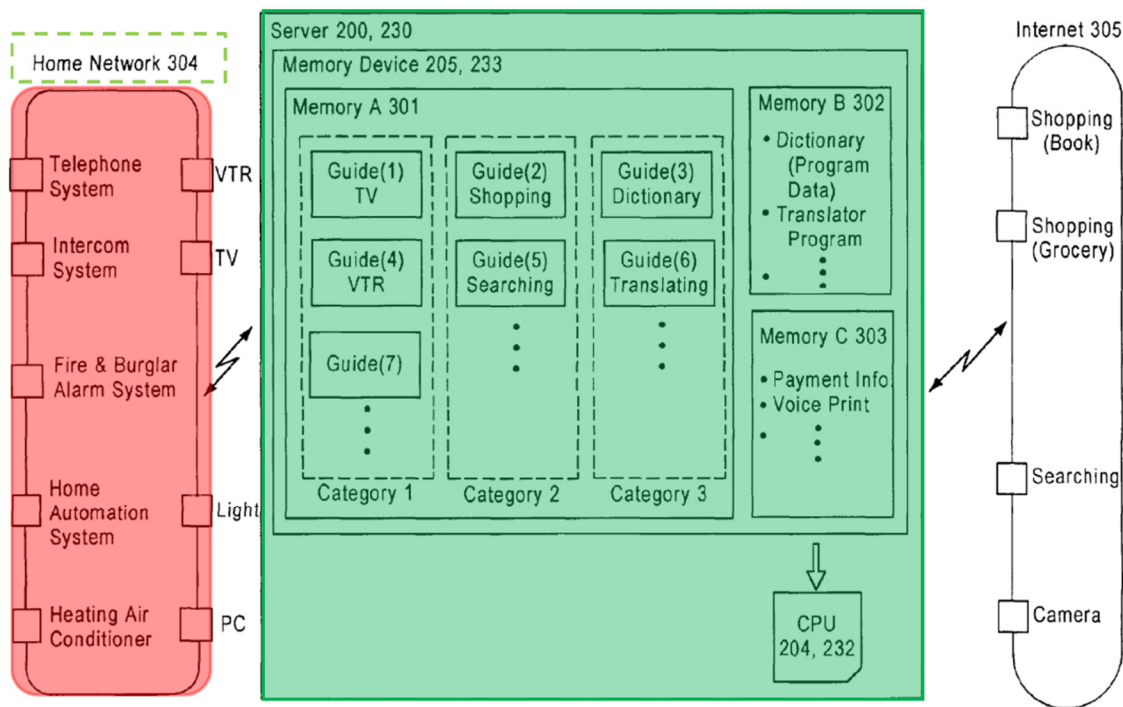


FIG. 6

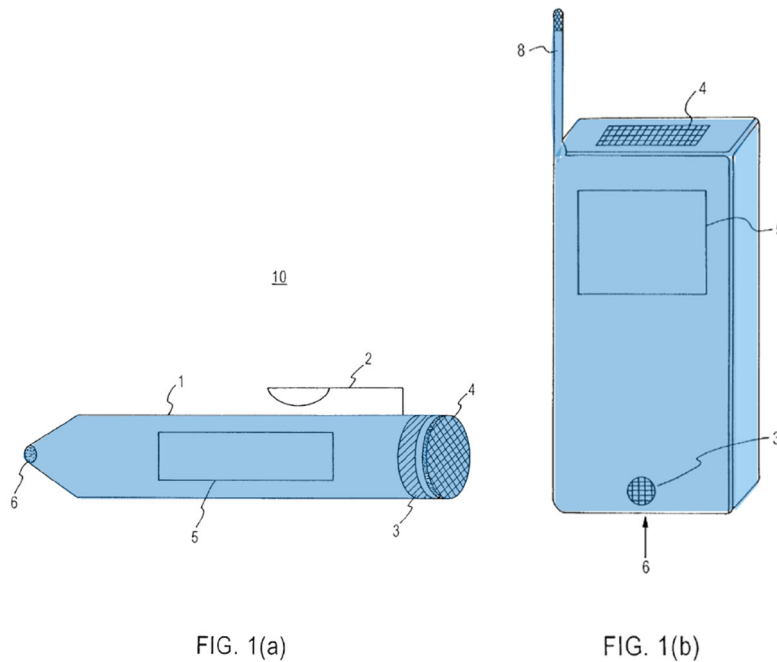
Howard discloses that the *client device* and the *controlled device* are in a *building*. Howard “relates generally to . . . control[] various activities *in a home*, including for example, *home appliances*.” (EX1005, 1:7-14.) Howard states that the “*portable device* [can be] readily carried by the operator” and that the system

“allows the operator to program and control the home gateway *from essentially any area of the home.*” (EX1005, 1:53-58, 2:19-26; EX1002, ¶¶334-335.)

c. 1[PR][3]

Howard discloses that each of the *client device* and the *controlled device* communicates over a wireless network. (EX1002, ¶336.)

“[T]he *portable device 10* also comprises . . . 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.” (EX1005, 3:41-48, claims 1 and 9, 1:65-2:7 and 4:57-63; EX1002, ¶337.)



Referring FIG. 2, “the *portable device 10* comprises both a transmitter system 101 and a receiver system 102.” (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, 6:25-30; EX1002, ¶338.) The wireless communication is also shown in FIG. 4. (EX1005, 6:32-36; EX1002, ¶¶338-339.)

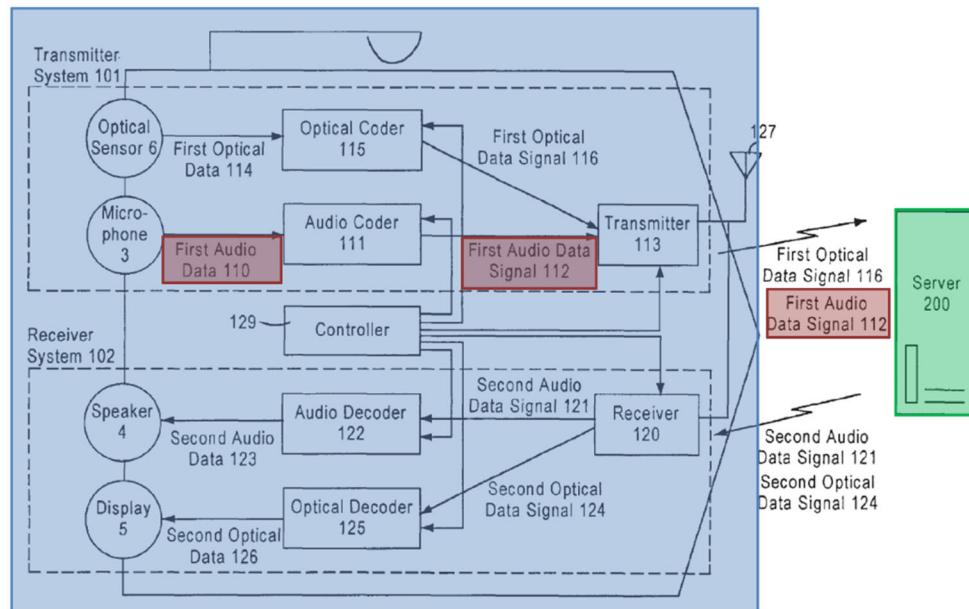


FIG. 2

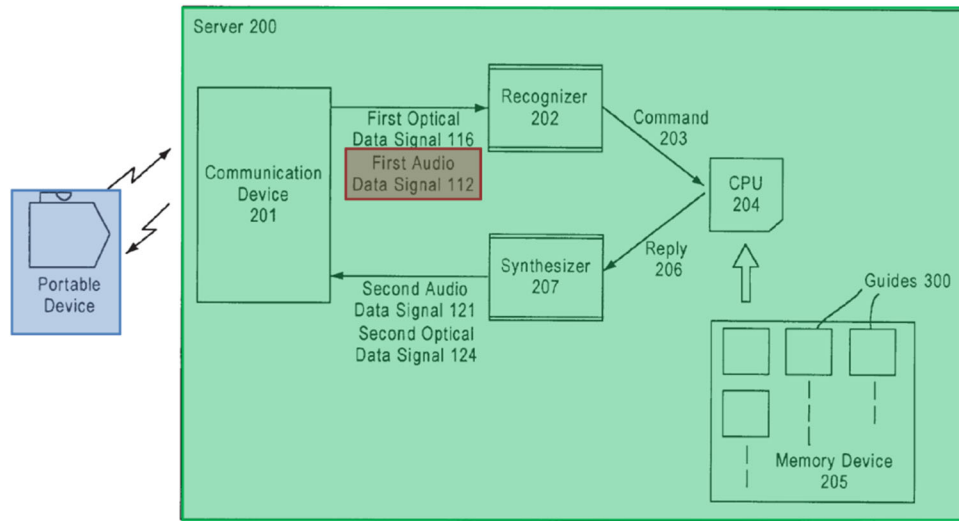


FIG. 4

Howard discloses that the *controlled device* communicates over a wireless network. “[T]he 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; EX1002, ¶340.)

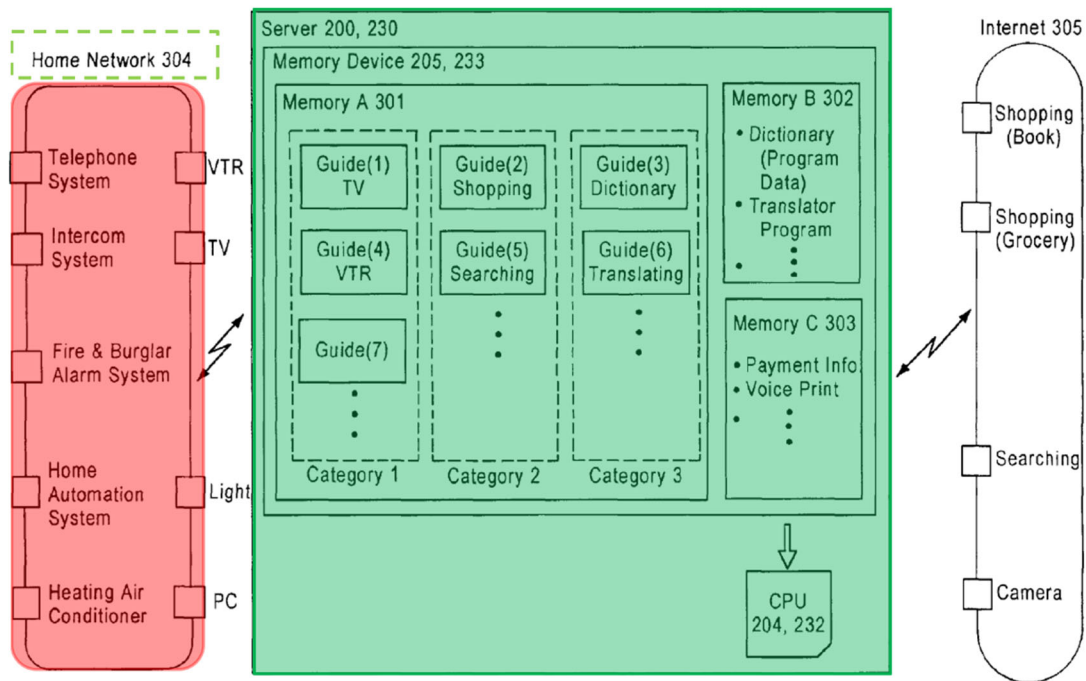


FIG. 6

Thus, Howard discloses that each of the *portable device* and the *home appliances and other devices* communicates wirelessly with the *server*. It would have been obvious to a POSITA 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. (EX1002, ¶341.)

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 the central communication routing point” and home devices “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).” (EX1006, 11; EX1002, ¶342.) Doherty’s

exemplary figures show that in-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; EX1002, ¶¶343-349.)

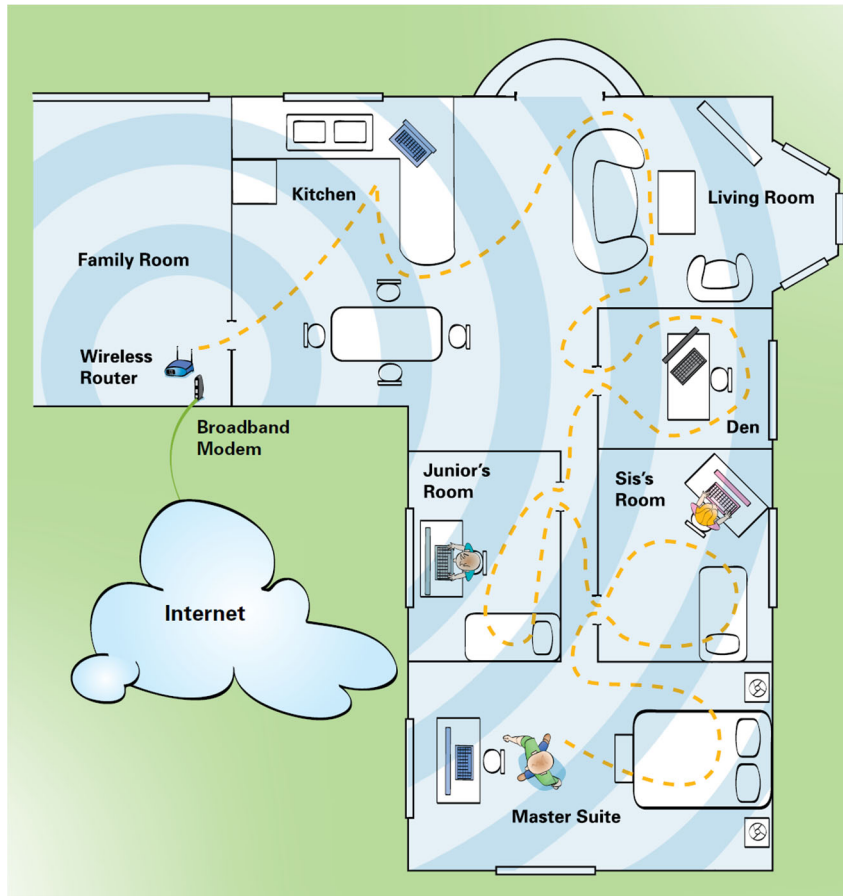


Figure 7-28 EasyLink Graphical View of Your Network

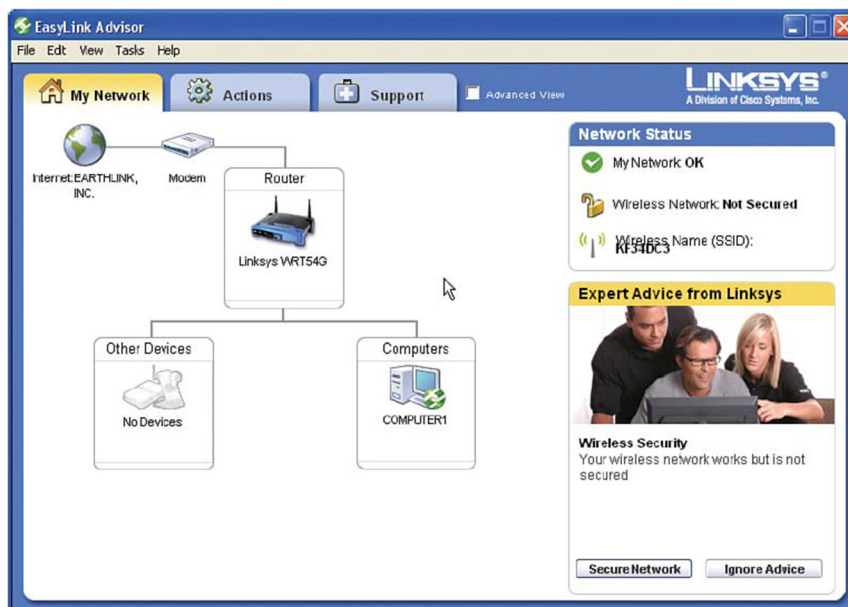
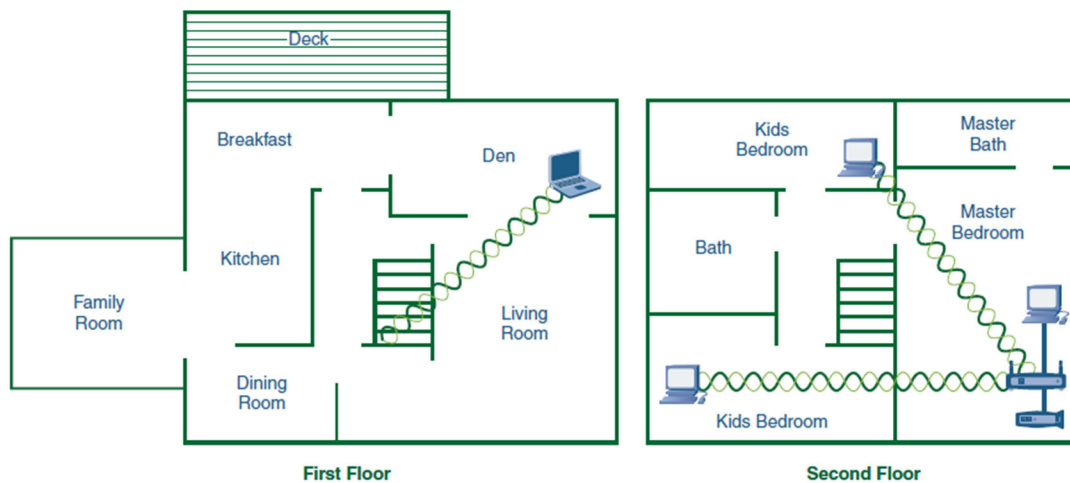


Figure 4-4 Adding Wired and Wireless Computers to the Sketch



Howard and Doherty are analogous art to the '710 patent. All relate to providing wireless device connectivity and/or control at home. (EX1005, Claim 1, Claim 9; EX1006, Title; EX1002, ¶350.)

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. (EX1002, ¶351.) For example, Howard's *portable device* and *home devices* can send/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. *Id.* 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. *Id.*

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; EX1002, ¶352.)

Regarding (2), Howard recognized the benefit of reducing resource requirement of local devices by placing the audio processing task on the server. “The analysis of the optical data signals 116 and/or the audio data signals 112 . . .

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; EX1002, ¶353.)

Indeed, the benefits for having a server in the cloud perform resource-intensive tasks such as speech processing were well-known. (EX1002, ¶354.) “By performing complex speech and voice recognition 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].)

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; EX1002, ¶355).

Thus, the combination would yield predictable results of enhancing Howard’s capabilities for speech processing and device control while achieving

efficient resource utilization on local devices and reliable network connectivity. (EX1002, ¶¶356-357). It would have been obvious to try combining the conventional techniques described in Doherty with Howard, as there is a finite number of identified, predictable solutions to achieve wireless connectivity and improve speech processing, without requiring much effort and with a reasonable expectation of success. *Id.* The known benefits of Doherty would prompt its application in the same field based on design incentives and other market forces. (*Id.*)

d. 1[PR][4]

Howard discloses an *Internet-connected server device*. Referring to FIG. 6, Howard discloses that, “the *server 200* is also coupled to the *Internet 305* via any suitable connection.” (EX1005, 9:50-52; *see also id.*, 6:52-58; EX1002, ¶358.)

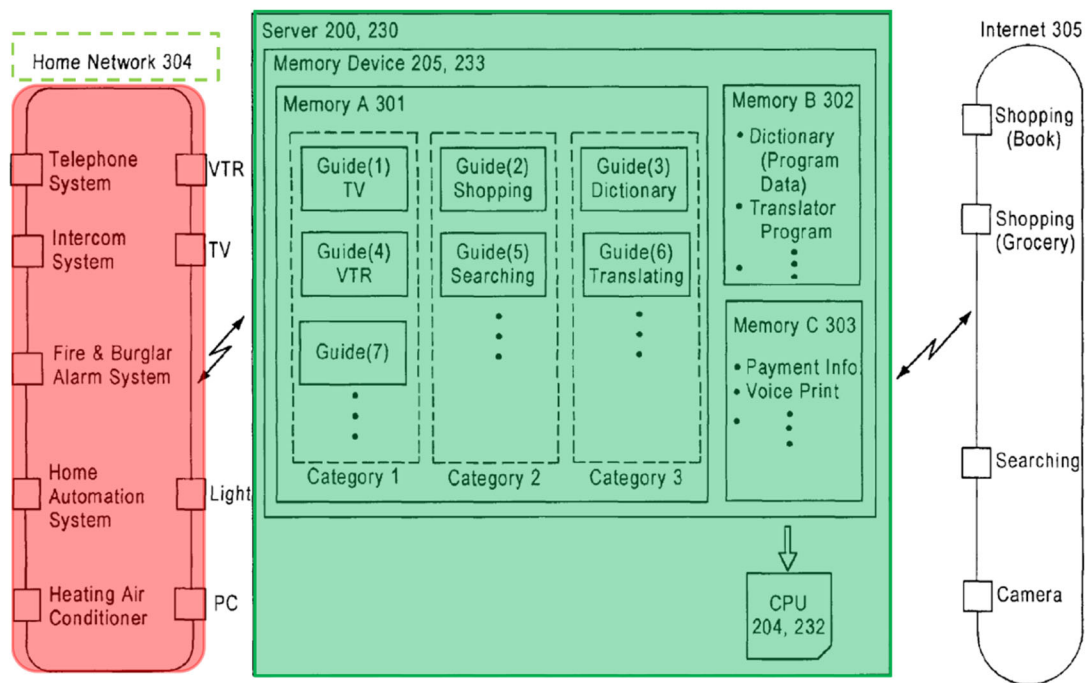


FIG. 6

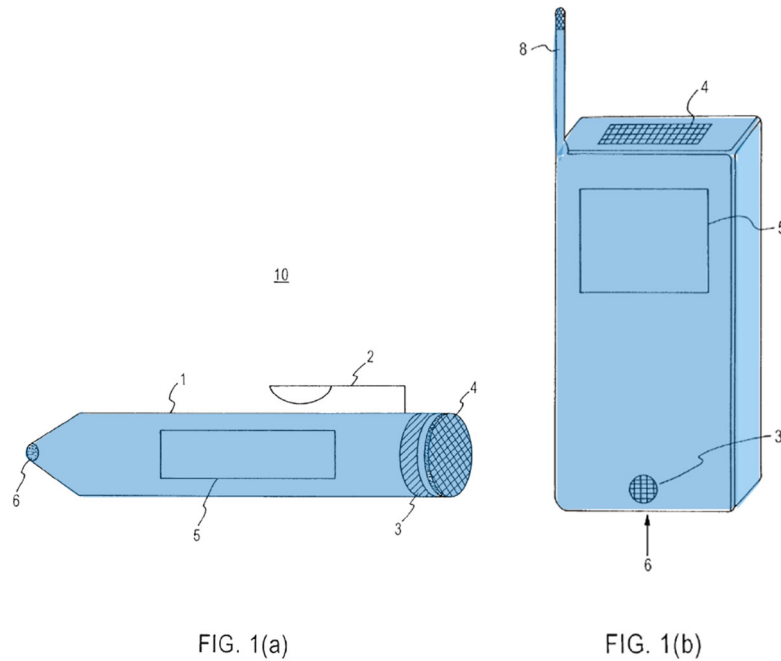
A POSITA would have understood that Howard's server can be outside the home. (EX1002, ¶359.) While Howard mentions "home server" in some instances, Howard does not limit its system to a server at the user's home. (*Id.*; EX1005, 3:8-10, claims 1-21.) At the time of the invention, remote servers (*e.g.*, cloud servers) were well-known and readily available, particularly for providing computational resources for processing audio. (EX1002, ¶360; *see* EX1007, Abstract, 2:50-53.) 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 would 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. (EX1002, ¶360.)

Further, the Howard-Doherty system discloses an Internet-connected server device external to the user's home. (*Supra* §IV.I.3.c.) It would have been obvious to implement the speech processing server in the cloud rather than at home. (*Id.*; EX1006, 11, 100; EX1002, ¶¶361-363.)

e. 1[A]

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, 3:25-27.) Thus, Howard's *portable device* captures human voice data by its microphone 3. (EX1005, 3:24-27, 3:54-62; EX1002, ¶¶364-365.)



Referring to FIG. 2, “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; EX1002, ¶¶366-368.) The human voice data is captured by the microphone and sent to the server. (*Id.*, 4:2-4, 4:10-24; EX1002, ¶¶366-368.)

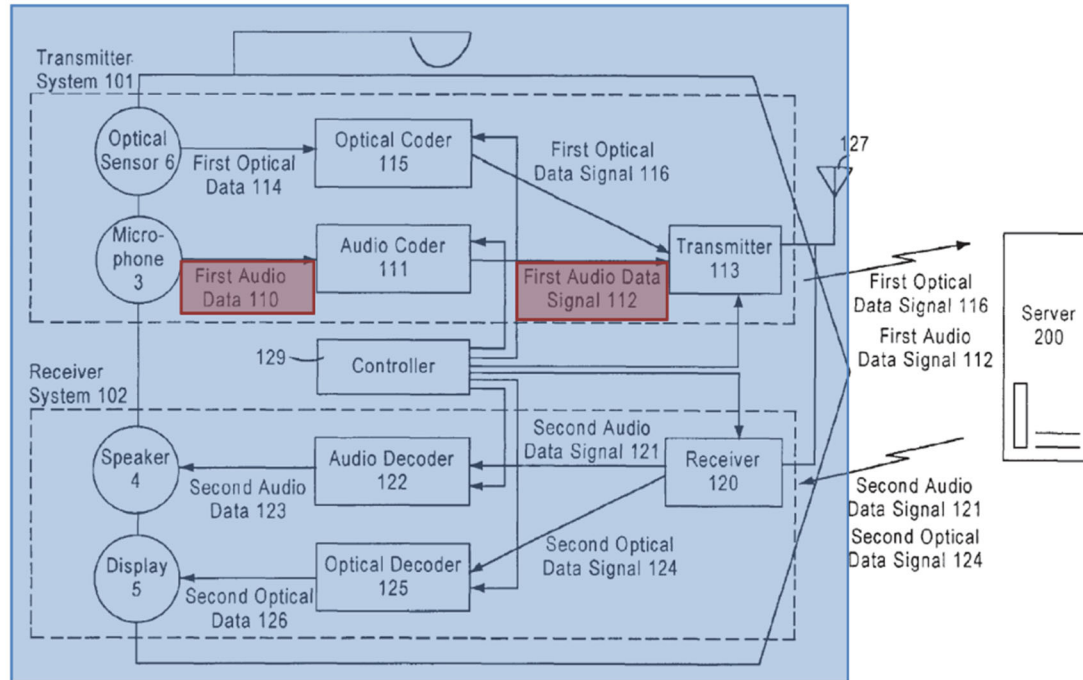


FIG. 2

f. 1[B] and 1[C]

Howard discloses sending to the *server*, by the *client device* via the wireless network, the captured *first human voice data* and receiving it by the *server* over the Internet. (EX1002, ¶¶369, 372.)

Referring to FIG. 9, “[i]n 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; EX1002, ¶¶370, 373, 375.)

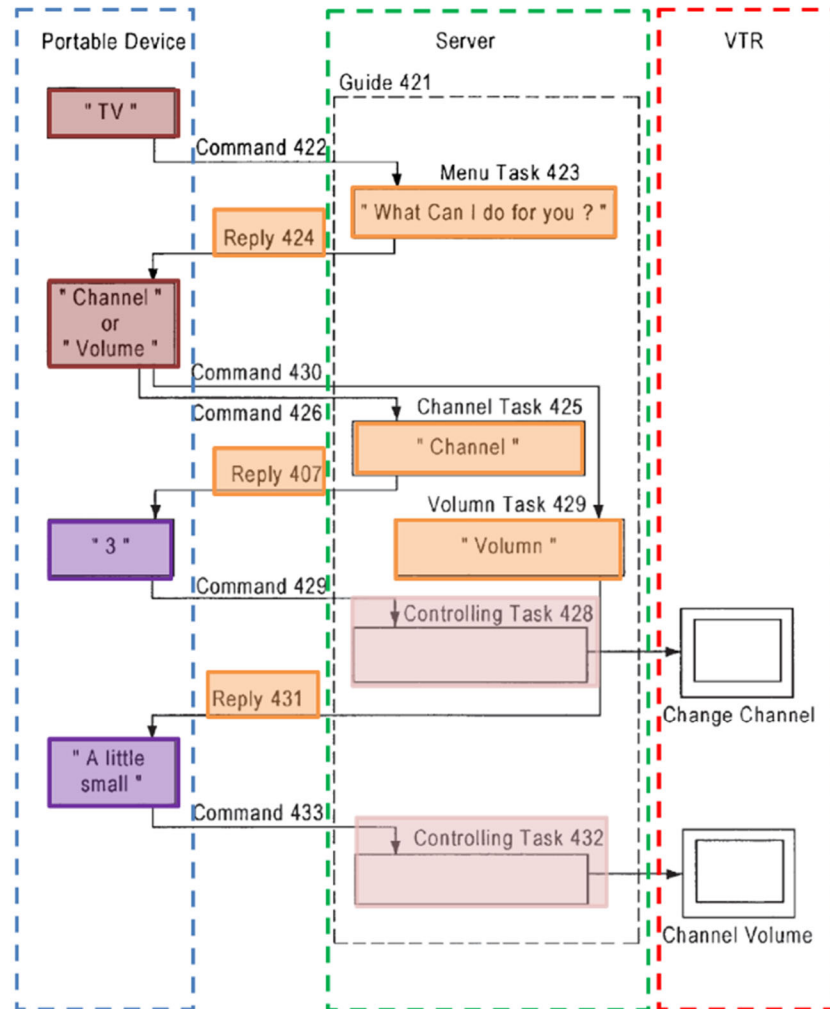


FIG. 9

Referring to FIG. 2, “[t]he 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, 4:10-24; *see also, id.*, 3:41-48; 6:25-30; 10:4-9, 10:35-41; claims 1 and 9; EX1002, ¶371.)

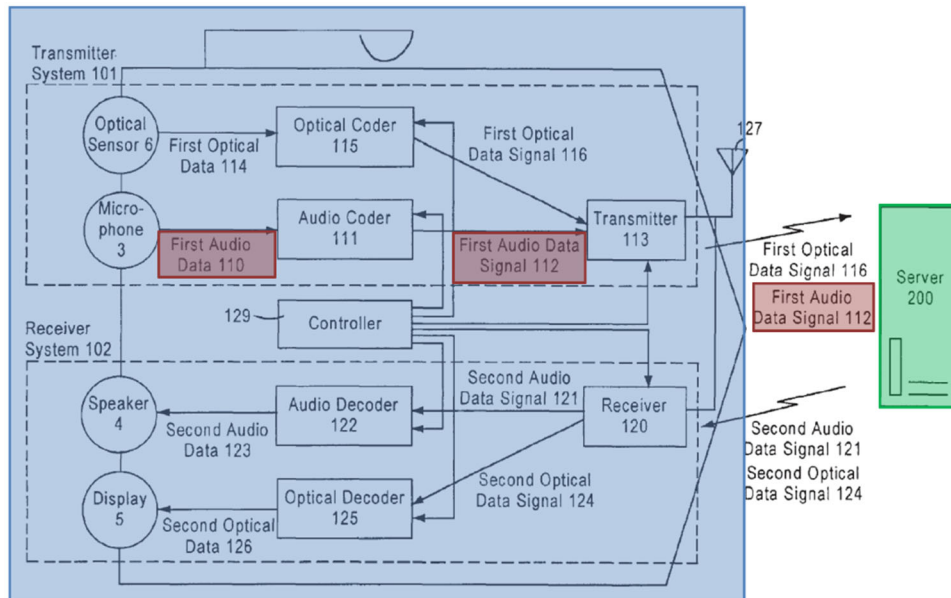


FIG. 2

Referring to FIG. 4, “the *server 200* includes a communication device 201 . . . 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, Claim 9; EX1002, ¶374.)

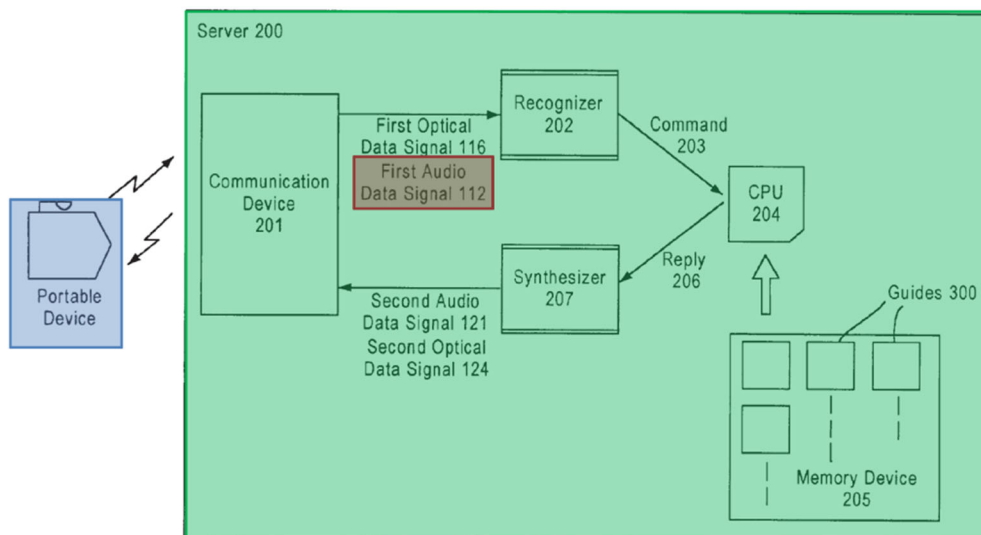


FIG. 4

In the Howard-Doherty system, Howard's *portable device* sends 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* §IV.I.3.c; EX1006, 11, 38-39, 96, 100, 228-229, 261; EX1002, ¶¶371, 376-377.) Further, it would have been obvious that Howard itself discloses or at least suggests this architecture. (*Id.*; *Supra* §§IV.I.3.c-IV.I.3.d.)

g. 1[D]

Howard discloses processing, by the *server*, the captured *first human voice data*. Referring to FIG. 9, “[u]pon 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.” (EX1005, 11:20-27, FIGs. 4, 8 and 10-13(d), 7:59-65, 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; EX1002, ¶¶378-380.)

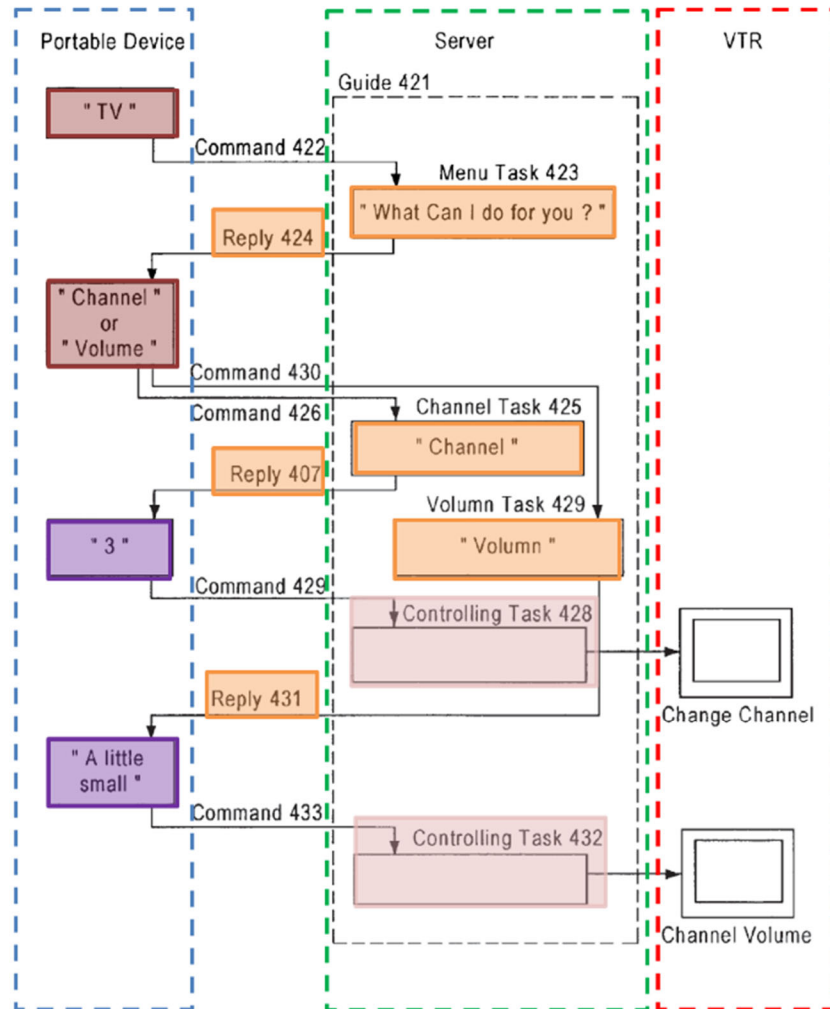


FIG. 9

h. 1[E] and 1[F]

Howard discloses responsive to the processing, sending a *first message*, by the *server* to the *client device* over the Internet and receiving the *first message* by the *client device* over the wireless network. (EX1002, ¶¶381, 388.)

Referring 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 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*’.’” (EX1005, 11:35-40, FIGs. 8 and 10-13(d), 10:15-29, 10:47-60, 11:27-43, 11:50-58, 11:66-12:6; EX1002, ¶¶382, 389-391.)

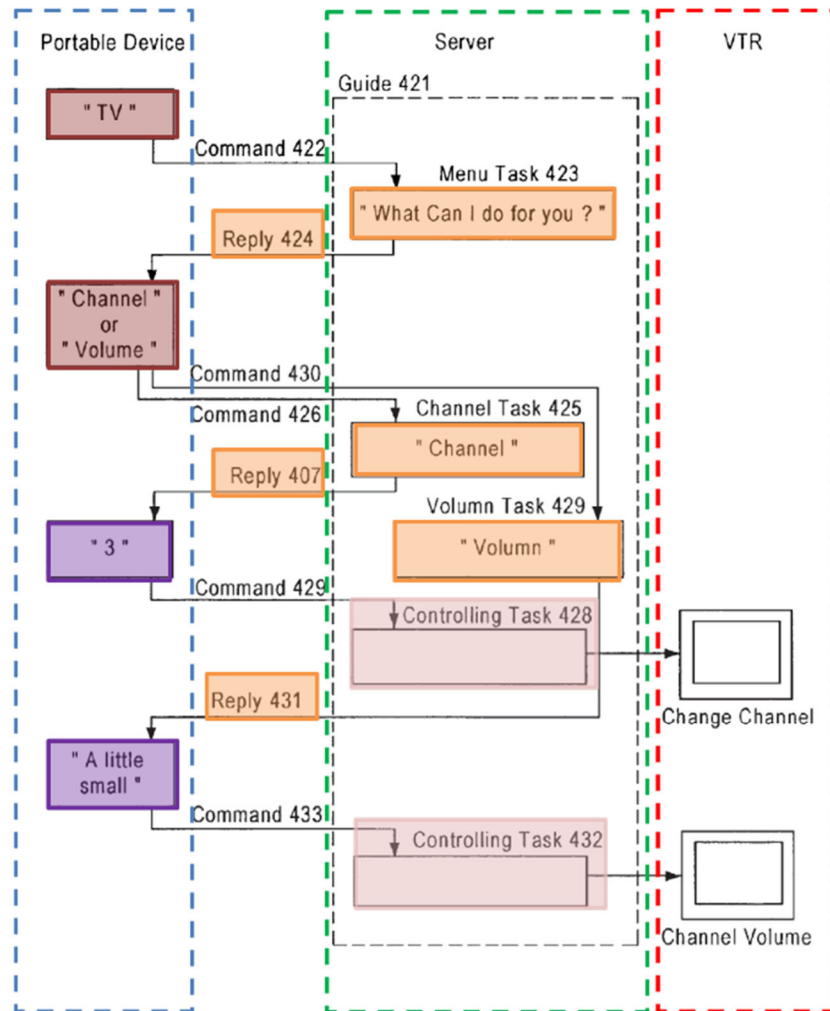


FIG. 9

“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, FIGs. 8 and 10-13(d), 4:39-43, 11:50-53, 10:23-33, 10:52-63, 11:35-47, 11:53-58, 12:2-6, Claim 2; EX1002, ¶¶382-383.) Further, “[i]n 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 is in response to the processing of the command. (EX1002, ¶384.)

FIG. 4 also discloses sending a *first message*, by the *server* to the *client device*, as shown below. (EX1005, 7:43-53, 8:37-54, 8:59-61, 8:62-9:1; EX1002, ¶385.)

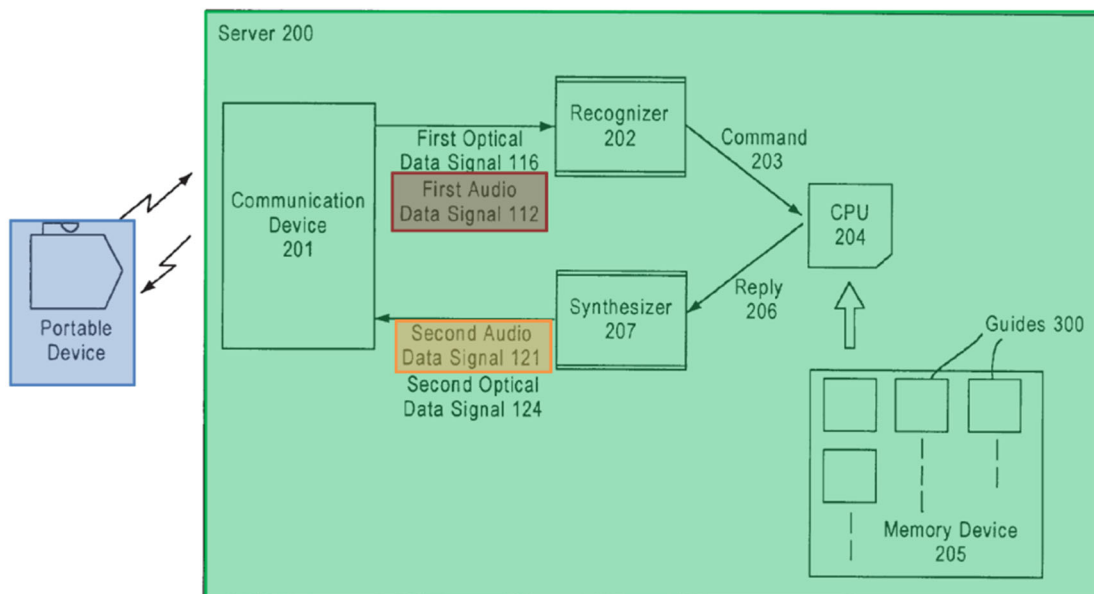


FIG. 4

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* §IV.I.3.c; EX1006, 11, 38-39, 96, 100, 228-229, 261; EX1002, ¶¶386-387, 392.)

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. (EX1002, ¶¶386-387, 393.) Howard's 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." (EX1005, 3:41-48, FIG. 2.) Further, Howard's *server 200* is connected to the *Internet*. (*Supra* §IV.I.3.d.)

i. 1[G]

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, FIGs. 2, 8 and 10-13(d), 3:25-27, 4:39-49, 10:52-63, 11:53-58, 11:38-47; EX1002, ¶¶394-396.) As discussed, operating a speaker involves operating an actuator. (*Supra* §IV.A.2.i; EX1002, ¶¶397-398.)

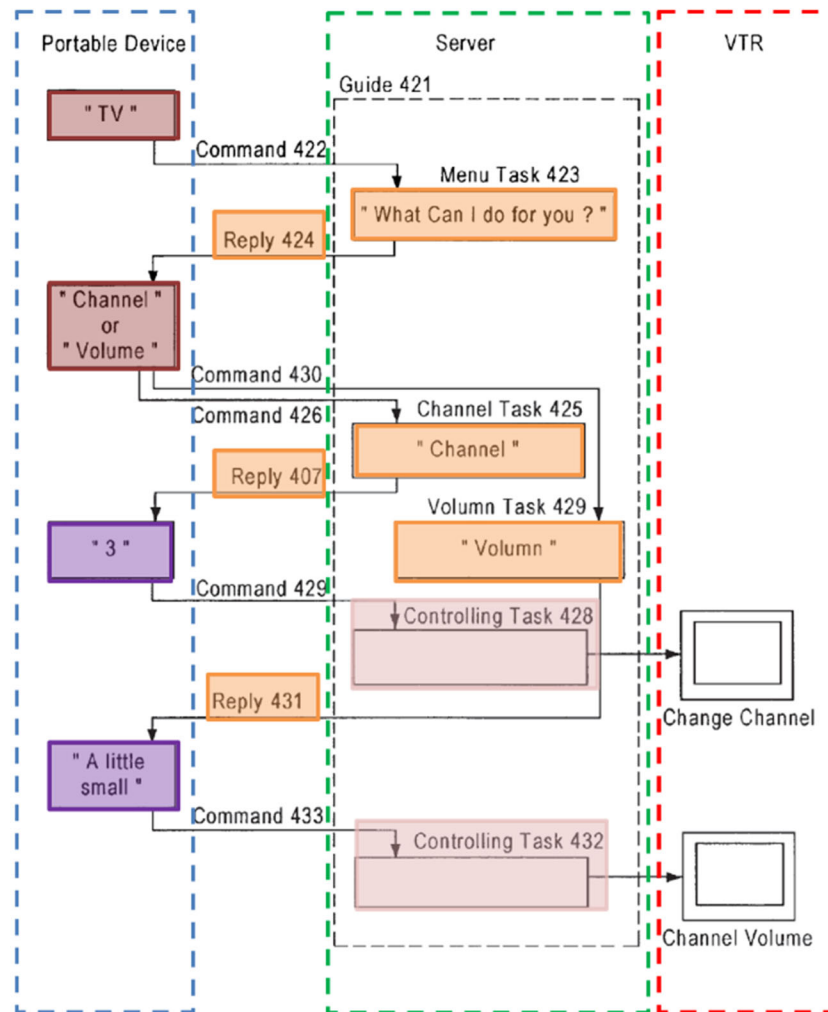


FIG. 9

Another example of “first actuator” is a display. (EX1002, ¶399; *e.g.*, EX1001, 25:33-38, EX1017, 175.) Howard discloses that “the *portable device 10* comprises a clip 2, a microphone 3, a speaker 4, a *display device 5*.” (EX1005, 3:25-27.) Howard discloses “the answer can also be provided in a visual format utilizing the *display device* contained in the *portable device 10*.” (EX1005, 14:40-49.) Referring to FIG. 8, “[u]pon 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?”* . . . Alternatively, the *reply 404* can be displayed on the *display 5* as a visual message.” (EX1005, 10:23-34; EX1002, ¶¶399-404.)

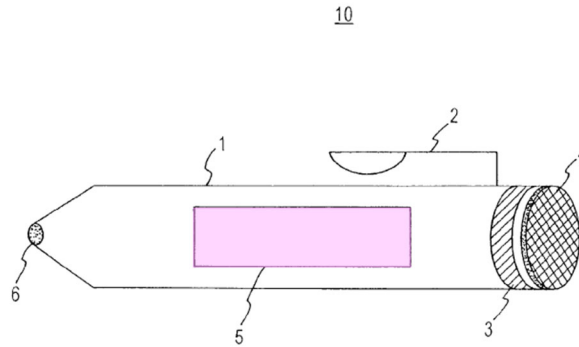


FIG. 1(a)

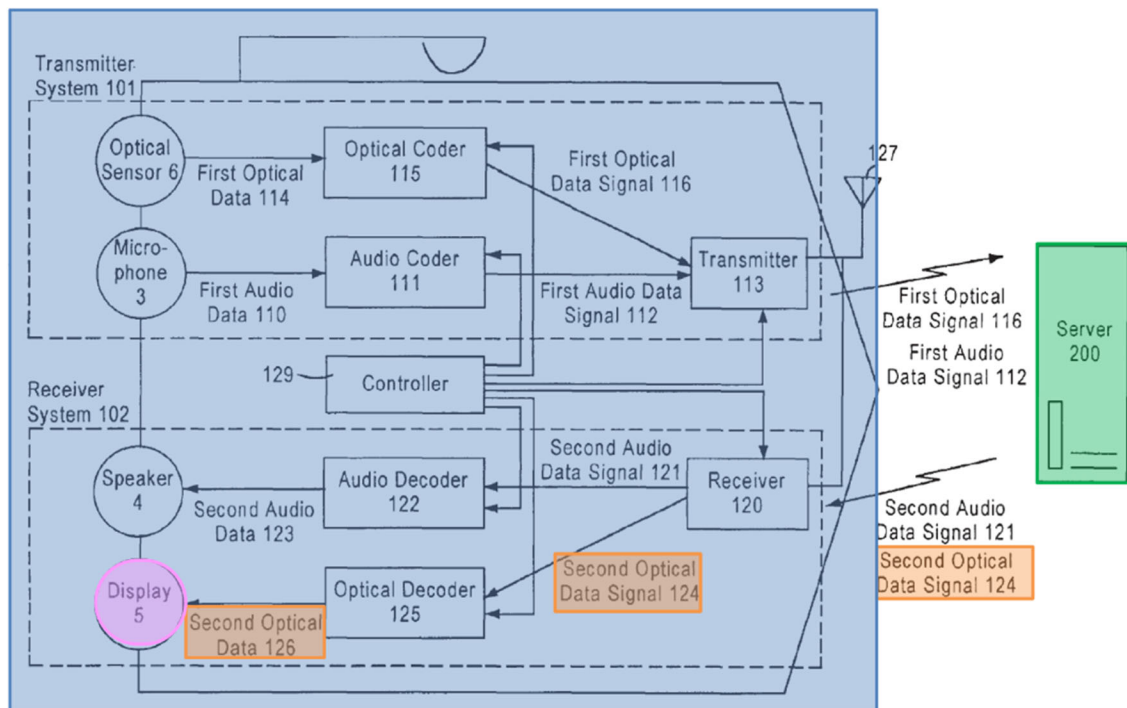
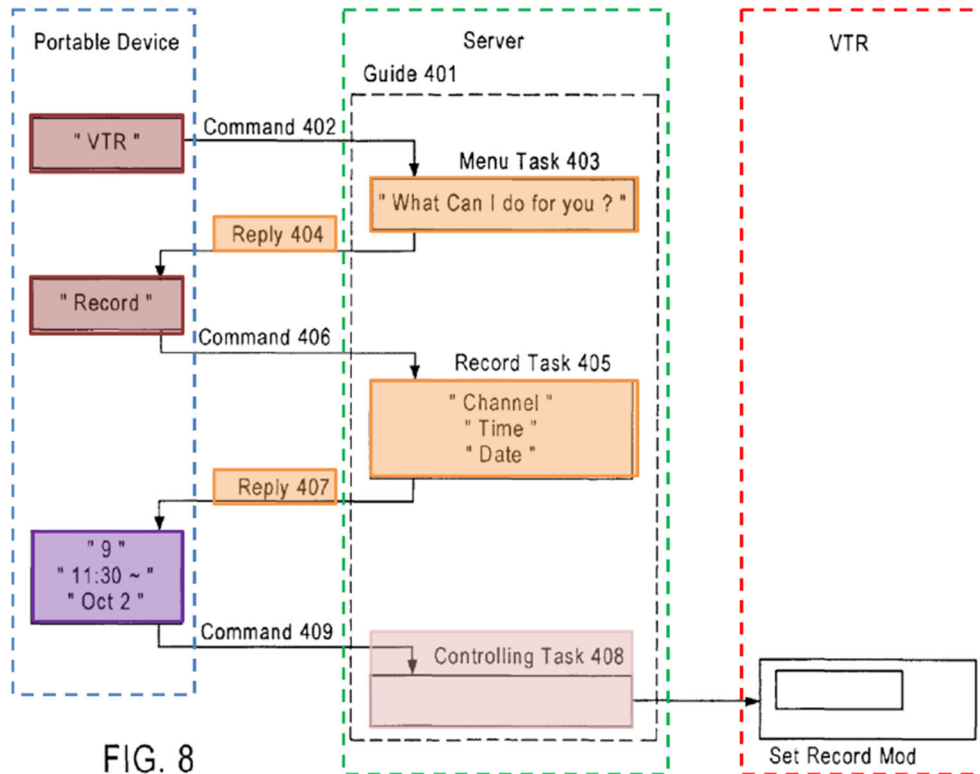


FIG. 2



j. 1[H]

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; EX1002, ¶¶405-407.)

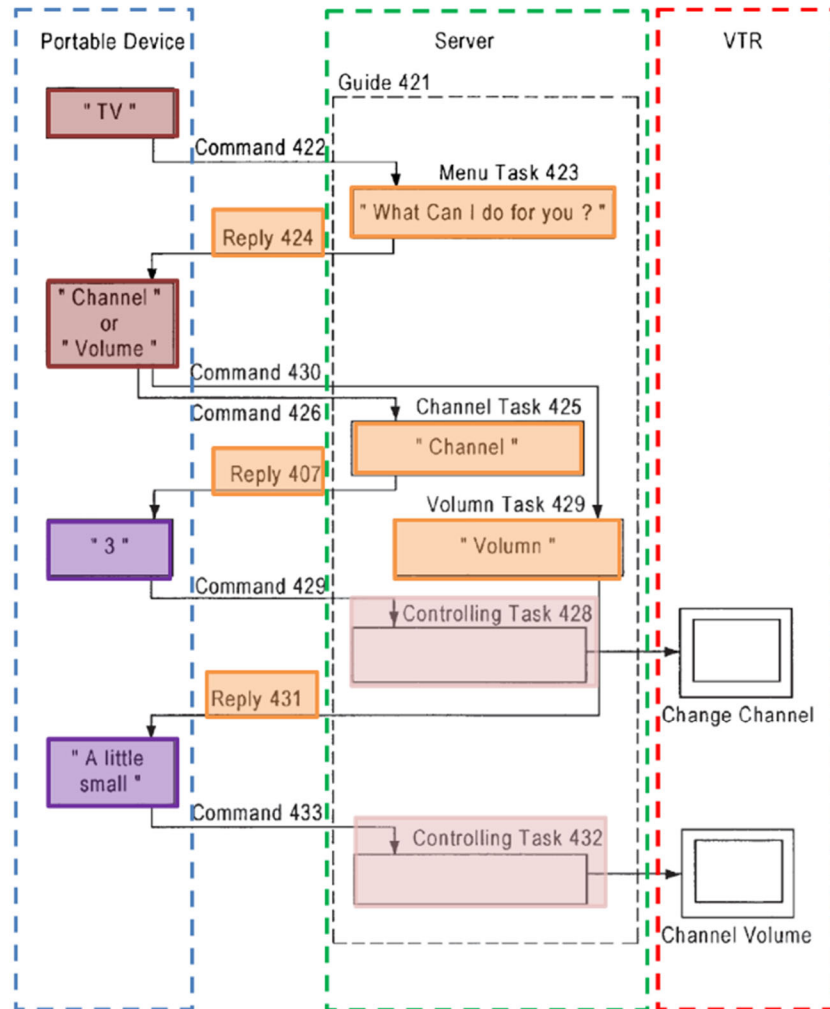


FIG. 9

k. 1[I], 1[J], and 1[K]

Like the first human voice data, the Howard-Doherty system discloses sending to the server the captured second human voice data by the client device via the wireless network, receiving it by the server over the Internet, and processing it by the server. (*Supra* §§IV.I.3.f-IV.I.3.g; EX1002, ¶¶408-410.) It would have been obvious to a POSITA that Howard itself discloses or at least suggests these limitations. *Id.*

I. 1[L] and 1[M]

Howard discloses responsive to the processing, sending a *second message*, by the *server* to the *controlled device* over the Internet and receiving it at the *controlled device* via the wireless network. (EX1002, ¶¶411, 418.)

Referring to FIG. 9, “[u]pon receipt of command 429 [specifying TV channel], 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.) Also, responsive to the command specifying TV volume, the *server* “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. 8.) The sending of the control signals is responsive to the processing of the commands by the server. (EX1002, ¶¶412, 419-421.)

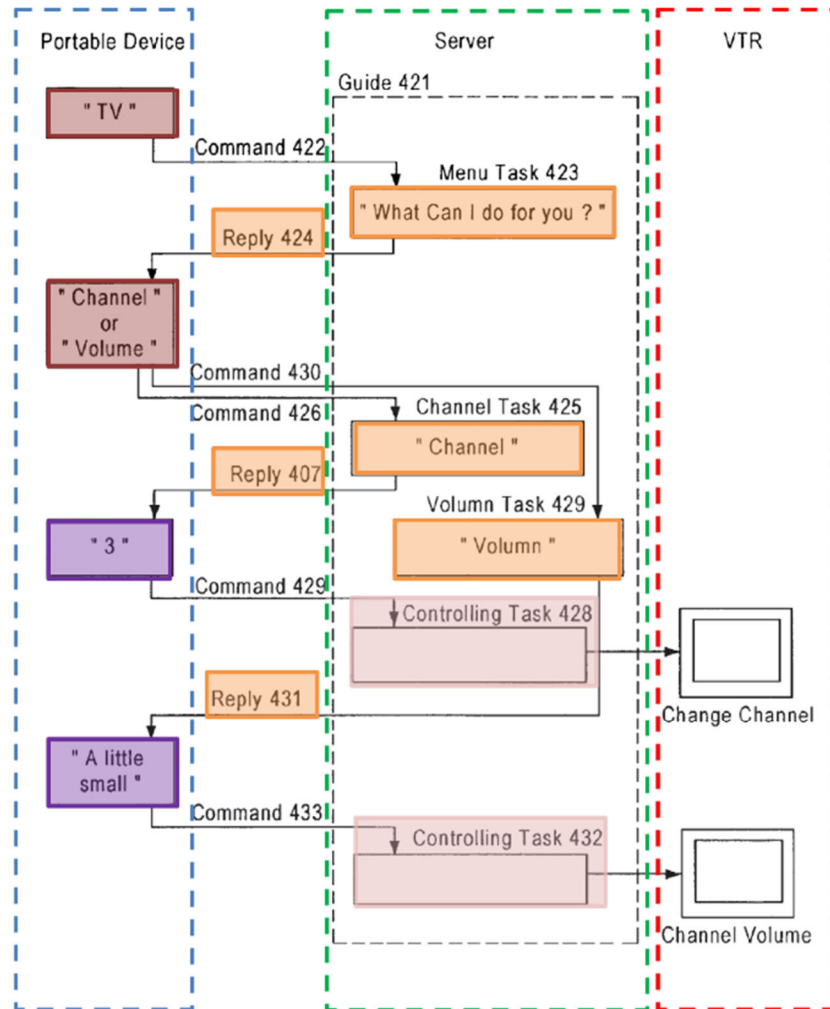


FIG. 9

Referring to FIG. 6, “memory A 301 [of *server* 200] functions to store the activity guides related to *controlling electrical components* coupled to the home network 304.” (EX1005, 8:42-47; EX1002, ¶¶413-414.)

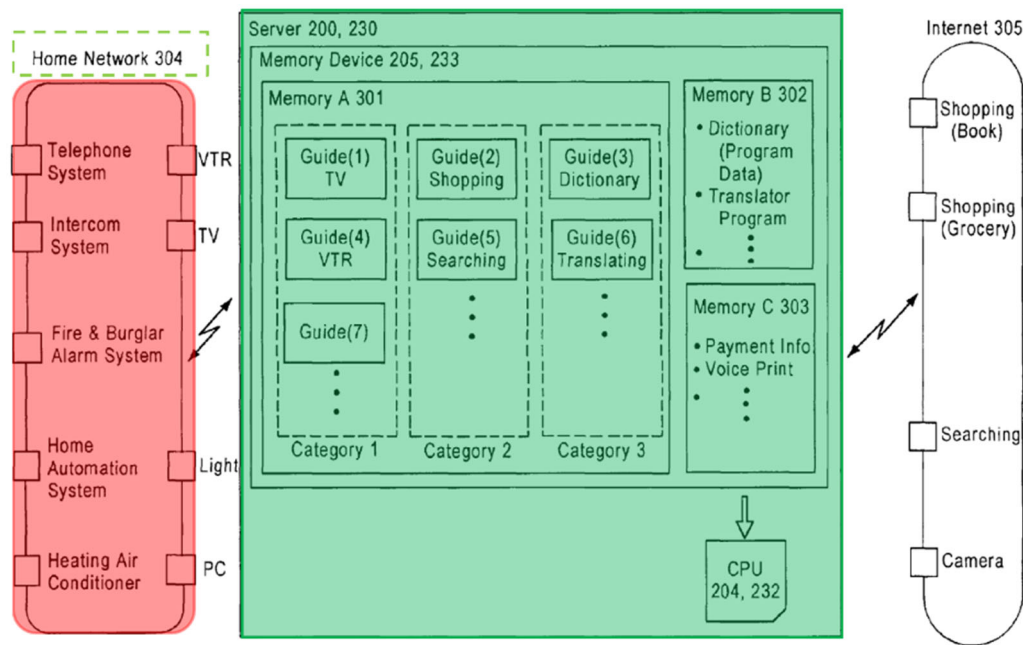


FIG. 6

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* §IV.I.3.c; EX1006, 11, 38-39, 96, 100, 228-229, 261; EX1002, ¶¶415-416, 422.)

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* §§IV.I.3.c, IV.I.3.d; EX1002, ¶¶417, 423.)

m. 1[N]

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; EX1002, ¶¶424-425, 428.)

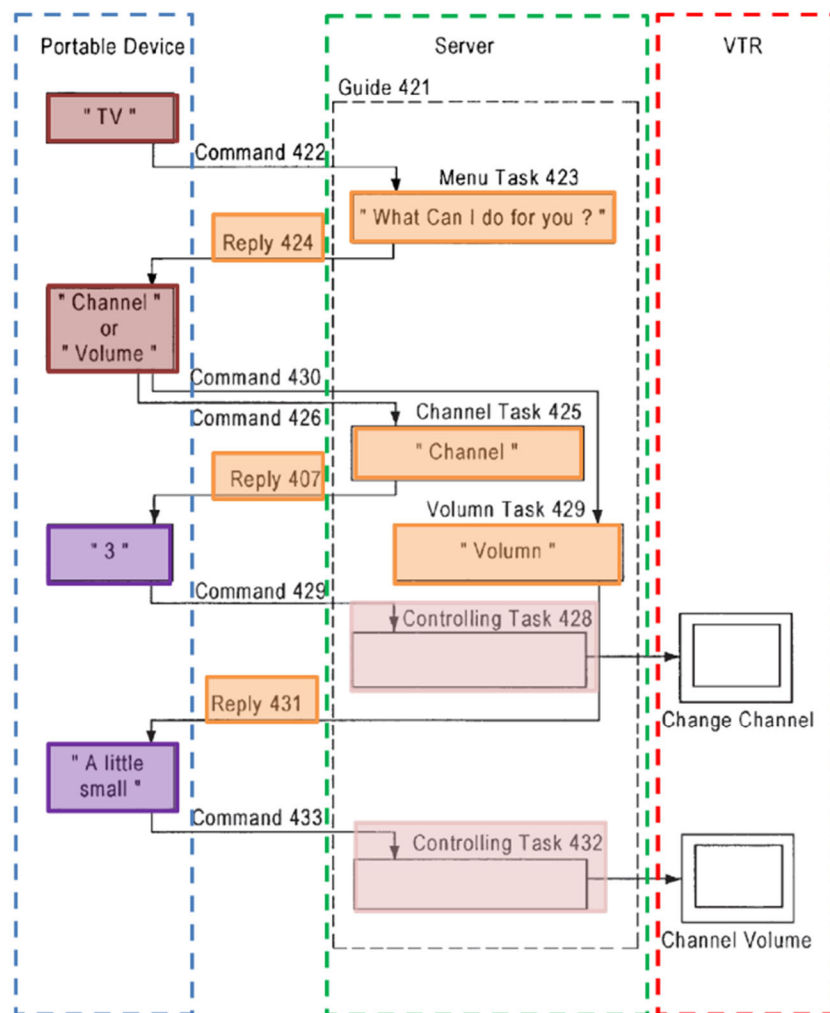


FIG. 9

Adjusting the volume of the TV and changing channels involve operating one or more *displays* and *speakers* (i.e., actuators) in the TV. (*Supra* §IV.A.2.i.) Further, recording media content using an VCR involves operating actuators (e.g., motors) that spins the recording/reading head. (EX1017, 226-227; EX1002, ¶429). It would have been obvious that the display and speaker of the *TV* are operated in response to the *control signals* (second message) received from the *server*. (EX1002, ¶¶426-427.)

Howard discloses a variety to actuators in controlled devices that “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 systems involves operating actuators. (E.g., EX1019, 401; EX1003, 396; EX1017, 280; EX1002, ¶¶430-432.)

4. Claims 5 and 6

Howard discloses operating speakers, displays, video equipment, electrical appliances. (*Supra* §§IV.I.3.i, IV.I.3.m; EX1005, 9:14-23, 11:27-47, 10:23-34, Claims 2 and 9.) These actuators can affect, change, produce, or create sound or audio, light, motion, temperature, humidity, etc. (*Id.*; EX1002, ¶¶433-436.)

5. Claims 10 and 11

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* §§IV.I.3.c - IV.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 that uniquely identifies the device in the network. (*Supra* §IV.C.4; EX1002, ¶¶437-438, 440-441.) The 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. (EX1002, ¶438.) Doherty discloses, and it was well-known, that networked devices are associated with IP addresses that are layer 3, either static or dynamic, and of IPv4 type, and/or globally unique MAC addresses of MAC-48, EUI-48, or EUI-64 type. (*Id.*; EX1006, 99, 166, 108, 175, 265, 210, 211, 277, Figure 10-44; EX1028, 127; EX1046, 2:50-59; EX1022, 11, 13.) Doherty discloses that the networked device can display the IP addresses, which means they are locally stored. (EX1006, 211, Figure 10-44; EX1002, ¶439.)

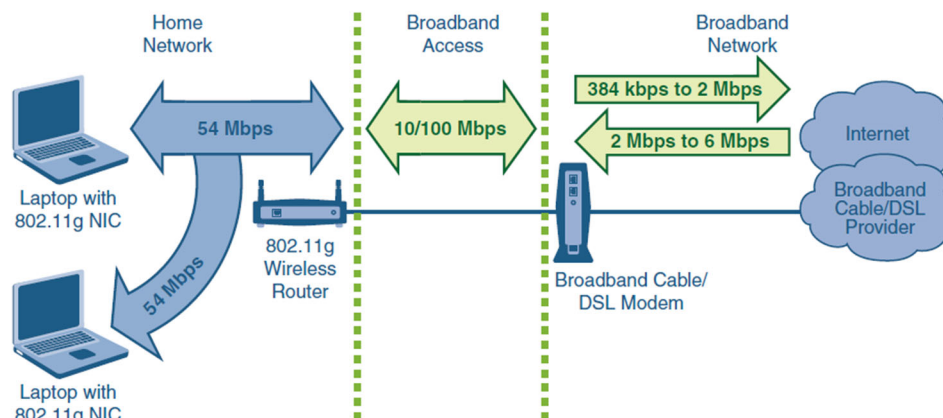
6. Claim 13

Howard discloses that the portable device and the controlled device communicates over a wireless network. (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. Howard states that “**802.11b is the standard that most vendors first developed for home networking.**” (EX1006, 18.)

Doherty also discloses the “802.11g standard”. (*Id.*, 19; EX1002, ¶¶442-444.)

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



7. Claims 16-21 and 23

The controlled device in the Howard-Doherty system are various household 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* §§IV.I.3.1-IV.I.3.m; EX1005, 1:18-30, 9:13-23; EX1002, ¶¶445-452.)

Howard discloses remotely controlling “(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**”. (*Supra* §IV.I.3.m; EX1005, 9:13-23, EX1005, 1:26-30.)

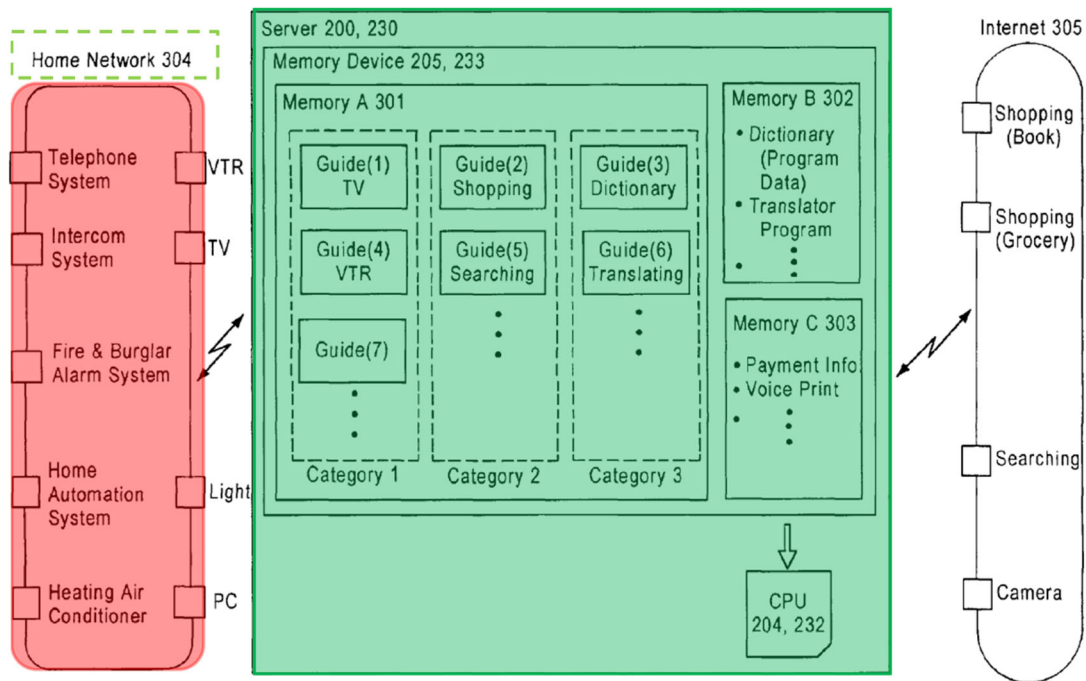


FIG. 6

8. Claim 29

Referring to FIG. 9, “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, 10:52-60, 11:40-46, 12:2-6; EX1002, ¶¶453-455.)

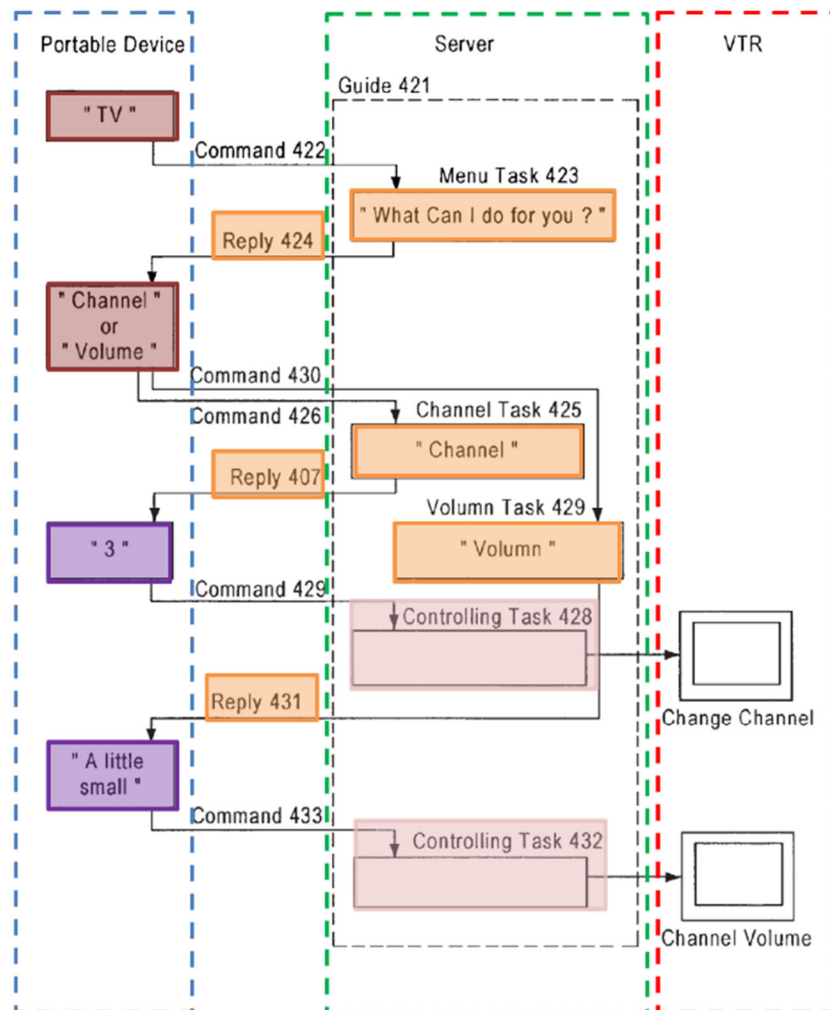
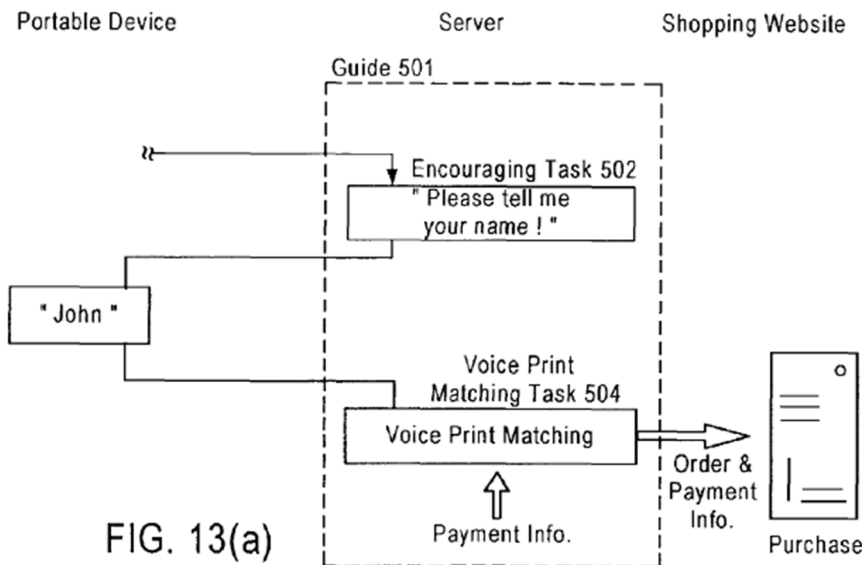


FIG. 9

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

Howard discloses performing voice recognition during processing to identify an operator's voice print "*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.*" (EX1005, 14:59-15:3; Figure 13(a); EX1002, ¶456.)



Further, Wittenberg discloses claim 2. (*Supra* §IV.B.2; EX1002, ¶457.)

It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Wittenberg. (EX1002, ¶458-463.) For example, all are analogous art and relate to controlling devices based on audio commands and providing security to the command/control system. (EX1005, 11:20-23; 14:59-15:3, 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; EX1007, Claim 1, 2:21-30, 2:25-31; 5:61-6:6, 6:7-13.)

The Howard-Doherty-Wittenberg combination would result in a system in which Wittenberg's voice recognition capabilities are used to analyze human audio commands in Howard-Doherty, 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. (EX1002, ¶¶461-463.)

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

1. Claims 3-4

Gruber discloses claims 3-4. (*Supra* §IV.C.2; EX1002, ¶¶464-465, 471-472.)

It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Gruber. (EX1002, ¶¶466-470.) For example, all are analogous art and relate to processing and responding to human speech commands (*e.g.*, to provide services). (*See* EX1005, 1:53-56, 11:20-23; EX1008, [0007].)

The Howard-Doherty-Gruber combination would capture sensor data (*e.g.*, light, temperature) using sensors (as taught by Howard and/or Gruber) in the portable device and transmit such data wirelessly to a router that routes the data to Howard-Doherty's server (*i.e.*, server 200 in the cloud) via the Internet to supplement processing as taught by Howard-Doherty and/or Gruber. (*See* EX1005, 4:26-38; EX1008, Abstract, [0013]; EX1002, ¶468.) Combining the teachings of Howard-Doherty and Gruber as such would yield predictable results of improving audio processing. (*Supra* §IV.C.2; EX1002, ¶¶467-468, 473.)

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

Benesty discloses claims 7-8. (*Supra* §IV.D.2; EX1002, ¶¶474-475, 480-481.)

It would have been obvious to a POSITA to combine Benesty with the Howard-Doherty system. (EX1002, ¶¶476-479.) For example, Howard, Doherty, and Benesty are analogous art and all relate to providing hands-free audio communication between users and devices. (EX1005, 1:53-56, 1:58-64, 3:24-27, 3:54-62; EX1009, 1021.)

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, yielding the predictable results of effectively improving audio capturing. (*Supra* §IV.D.2; EX1002, ¶¶478-479, 482.)

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

Howard discloses “wireless communication between the portable device 10 and a server 200” that its portable device’s “transmitter 113, receiver 120 . . . utilize the same data communication/transfer techniques as utilized for *data transmission in cellular phones*” and 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* §IV.I.3.c; EX1005, 3:41-48, 6:32-36; 5:16-21; 5:22-25; EX1002, ¶¶485-486, 496.)

Goldsmith discloses claims 12 and 14-15. (EX1006, 229-230; *supra* §IV.E.2; EX1002, ¶¶483-484, 489-492, 494-495.) At the time of the invention, routers having Bluetooth capabilities and Bluetooth devices (*e.g.*, Bluetooth access points/gateways) with connection to the Internet were well-known. (EX1025, 59:21-24; EX1026, 4:25-27; EX1027, 4:20-22; EX1032, 1:67-2:3; EX1002, ¶487.) Further, home appliances with cellular connectivity were known. (EX1030, claims 14-16, 18; EX1002, ¶497.) It would have been obvious to try combining Goldsmith with Howard-Doherty. (EX1002, ¶¶488, 493, 498-499.) For example, it would have been obvious for a POSITA to look to Goldsmith for conventional wireless networking techniques and protocols (*e.g.*, Bluetooth, Zigbee, WLAN based on IEEE 802.11a,b,g, cellular telephone 3G network) that can be applied to Howard-Doherty, especially as Goldsmith discloses that it is feasible to embed Bluetooth and Zigbee in electronic devices to create “smart homes.” (EX1035, 19; EX1005, 3:41-48, 5:22-25; EX1002, ¶488.)

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

Chatterjee discloses claim 22. (*Supra* §IV.F.2; EX1002, ¶¶500-501.)

It would have been obvious to a POSITA to combine the teachings of Howard-Doherty and Chatterjee. (EX1002, ¶502-505.) For example, Howard-Doherty and Chatterjee are analogous art and all relate to controlling home appliances remotely (*Supra* §IV.I.3.b; EX1005, 9:13-23, 1:18-30, *see also, id.*,

9:46-50; EX1011, Abstract.) Combining Chatterjee with Howard-Doherty would improve and expand the capabilities of Howard-Doherty in the same way.

(EX1002, ¶503.)

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

Rosenberger discloses claims 24-25. (*Supra* §IV.G.2; EX1002, ¶¶506-507, 513-514.) It would have been obvious to combine the teachings of Rosenberger with Howard-Doherty. (EX1002, ¶¶508-512, 515.) For example, Rosenberger and Howard-Doherty are analogous art and all relate to processing and responding to audio commands. (*See* EX1005, 1:53-56 and 11:20-23; EX1012, 1:7-13.) 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. A POSITA would be motivated to use the known techniques of Rosenberger to improve the usability of the device in the same way. (EX1002, ¶¶508-510.)

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

Rossing discloses claim 9 and 26-28. (*Supra* §§IV.H.2, IV.H.3; EX1002, ¶¶516-517, 521-522, 524, 526-530.)

Howard discloses a microphone and a speaker in the *client device*. (*See Supra* §§IV.I.3.e, IV.A.2.i; EX1005, Claims 1 and 9, Abstract, 1:65-2:6, 3:25-27;

see also id., FIG. 1(b).) Rossing is a textbook on how microphones and speakers operate. Thus, a POSITA would have been motivated to look to Rossing for techniques and details regarding microphones and speakers for implementing the Howard-Doherty system. (EX1002, ¶¶518-520, 523, 525-527, 531-533.)

V. NOTICES AND STATEMENTS

Pursuant to 37 C.F.R. §42.8(b)(1), Amazon.com Services LLC is the real party-in-interest for Petitioner.

Pursuant to 37 C.F.R. §42.8(b)(2), Petitioner identifies the following related matter.

- *Smart Speaker LLC v. Amazon.com Services LLC*, No. 2-25-cv-00707 (E.D. Tex.)

Pursuant to 37 C.F.R. §42.8(b)(3), Petitioner identifies the following counsel (and a power of attorney accompanies this Petition).

Lead Counsel	First Backup Counsel
Alex S. Yap Registration No.: 60609 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Tel: (213) 892-5200 Fax: (213) 892-5454 E-Mail: AYap@mofo.com	Mehran Arjomand Registration No.: 48231 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Tel: (213) 892-5200 Fax: (213) 892-5454 E-Mail: MArjomand@mofo.com

	Second Backup Counsel
	Yunan Yuan Registration No.: 73637 MORRISON & FOERSTER LLP 755 Page Mill Road Palo Alto, CA 94304-1018 Tel: (650) 813-5600 Fax: (650) 494-0792 E-Mail: AYuan@mofo.com

Pursuant to 37 C.F.R. § 42.8(b)(4), service information for lead and back-up counsel is provided above. Petitioner consents to electronic service by email to Amazon_SmartSpeaker_IPRs@mofo.com.

Pursuant to 37 C.F.R. § 42.104(a), Petitioners certify that the '710 patent is available for *inter partes* review and that Petitioners are not barred or estopped from requesting an *inter partes* review challenging the patent claims on the grounds identified in this Petition.

The USPTO is authorized to charge any required fees, including the fee as set forth in 37 C.F.R. § 42.15(a) and any excess claim fees, to Deposit Account No. **03-1952** referencing Docket No. **41951-00008.26**.

IPR2026-00145

Dated: November 21, 2025

Respectfully submitted,

By /s/ Alex S. Yap

Alex S. Yap

Registration No.: 60,609

MORRISON & FOERSTER LLP

707 Wilshire Blvd. Suite 6000

Los Angeles, CA 90017

Tel: (213) 892-5200

Fax: (213) 892-5454

E-Mail: AYap@mofo.com

Counsel for Petitioner

Certification of Word Count (37 C.F.R. § 42.24)

I hereby certify that this Petition for *Inter Partes* Review has 13,673 words (as counted by the “Word Count” feature of the Microsoft Word™ word-processing system), exclusive of “a table of contents, a table of authorities, mandatory notices under § 42.8, a certificate of service or word count, or appendix of exhibits or claim listing.”

Dated: November 21, 2015

By /s/ Alex S. Yap

Alex S. Yap

Certificate of Service (37 C.F.R. § 42.6(e)(4) and 42.205(a))

I hereby certify that the attached Petition for *Inter Partes* Review and supporting materials was filed via the P-TACTS system on November 21, 2025, and served on November 22, 2025, by Priority Mail Express®, on the Patent Owner at the correspondence address indicated for U.S. Patent No. 11,128,710.

May Patents Ltd.
c/o Dorit Shem-Tov
P.O.B 7230
Ramat-Gan 5217102
Israel

Dated: November 21, 2025

/s/ Alex S. Yap
Alex S. Yap