

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM SERVICES LLC
Petitioner,

v.

SMART SPEAKER LLC
Patent Owner.

Patent No. 12,010,174

Inter Partes Review No. IPR2026-00147

**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 '174 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	4
IV. PRECISE REASONS FOR RELIEF REQUESTED	5
A. Ground 1: Lim Renders Obvious Claims 1, 5-6, 10-12, 16, 20-21, 23, and 29	8
1. Summary of Lim	8
2. Claim 1	13
3. Claims 5/6	42
4. Claims 10/11/12	42
5. Claims 16/20/21/23	44
6. Claim 29	44
B. Ground 2: Lim and Wittenberg Render Obvious Claim 2	45
1. Summary of Wittenberg	45
2. Claim 2	45
C. Ground 3: Lim and Gruber Render Obvious Claim 3-4	46
1. Summary of Gruber	46
2. Claim 3	47
3. Claim 4	49
D. Ground 4: Lim and Benesty Render Obvious Claims 7-8	49
1. Summary of Benesty	49
2. Claims 7/8	50
E. Ground 5: Lim and Goldsmith Render Obvious Claims 13-15	52

TABLE OF CONTENTS
(continued)

	Page
1. Summary of Goldsmith.....	52
2. Claim 13.....	52
3. Claim 14.....	54
4. Claim 15.....	54
F. Ground 6: Lim and Chatterjee Render Obvious Claim 17-19, 22	56
1. Summary of Chatterjee	56
2. Claims 17-19, 22	56
G. Ground 7: Lim and Rosenberger Render Obvious Claims 24-25	58
1. Summary of Rosenberger	58
2. Claims 24/25	59
H. Ground 8: Lim and Rossing Render Obvious Claims 9, 26-28	61
1. Summary of Rossing.....	61
2. Claim 9	62
3. Claims 26-28	64
I. Ground 9: Williams and Lim Render Obvious Claims 1, 5-6, 10-12, 16, 20-21, 23, 29.	65
1. Summary of Williams	65
2. A POSITA would have been motivated to combine Williams and Lim.....	67
3. Claim 1	69
4. Claims 5/6	93
5. Claims 10/11/12	93
6. Claims 16/20/21/23	94
7. Claim 29	95
J. Ground 10: Williams, Lim, and Wittenberg Render Obvious Claim 2	95

TABLE OF CONTENTS
(continued)

	Page
1. Claim 2.....	95
K. Ground 11: Williams, Lim, and Gruber Render Obvious Claims 3-4.....	96
1. Claims 3-4.....	96
L. Ground 12: Williams, Lim, and Benesty Render Obvious Claims 7-8	97
1. Claims 7/8	97
M. Ground 13: Williams, Lim, and Goldsmith Render Obvious Claims 13-15	98
1. Claims 13-15	98
N. Ground 14: Williams, Lim, and Chatterjee Render Obvious Claim 17-19, 22.....	99
1. Claims 17-19, 22	99
O. Ground 15: Williams, Lim, and Rosenberger Render Obvious Claims 24-25	99
1. Claims 24/25	99
P. Ground 16: Williams, Lim, and Rossing Render Obvious Claims 9 and 26-28.....	100
1. Claims 9, 26-28	100
V. NOTICES AND STATEMENTS.....	100

Exhibit List

Exhibit Description	Exhibit #
U.S. Patent No. 12,010,174 to Binder	1001
Expert Declaration of Dr. Henry Houh	1002
File History of U.S. Patent No. 12,010,174	1003
File History of U.S. Patent No. 11,128,710	1004
U.S. Patent Pub. No. 2013/0132094 to Lim	1005
U.S. Patent Pub. No. 2012/0198339 to Williams <i>et al.</i>	1006
Jim Doherty <i>et al.</i> , “Wireless Home Networking Simplified” (Cisco Press, 2007).	1007
U.S. Patent No. 7,529,677 to Wittenberg	1008
U.S. Patent Pub. No. 2012/0265528 to Gruber <i>et al.</i>	1009
Jacob Benesty <i>et al.</i> , “Springer Handbook of Speech Processing” (Springer-Verlag Berlin Heidelberg, 2008).	1010
RESERVED	1011
U.S. Patent Pub. No. 2013/0052946 to Chatterjee <i>et al.</i>	1012
U.S. Patent No. 8,340,975 to Rosenberger	1013
Thomas D. Rossing <i>et al.</i> , “Springer Handbook of Acoustics” (Springer Science+Business Media, LLC, 2007).	1014
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)	1015

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)	1016
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)	1017
Ian Sinclair <i>et al.</i> , “Electronic and Electrical Servicing” (Elsevier Limited, 2 nd Edition, 2007)	1018
U.S. Patent No. 6,906,617 to Van der Meulen	1019
Ali Vedavarz <i>et al.</i> , “HVAC: Handbook of Heating, Ventilation and Air Conditioning for Design and Implementation” (Industrial Press Inc., 2007)	1020
U.S. Patent No. 7,062,339 to Howard <i>et al.</i>	1021
U.S. Patent Pub. No. 2011/0299412 to Diab <i>et al.</i>	1022
Robert J. Shimonski <i>et al.</i> , “Sniffer Pro Network Optimization and Troubleshooting Handbook” (Syngress Publishing, Inc., 2002).	1023
U.S. Patent No. 8,200,798 to Yadav <i>et al.</i>	1024
U.S. Patent Pub. No. 2012/0253480 to Abe	1025
U.S. Patent No. 7,539,533 to Tran	1026
U.S. Patent No. 8,356,754 to Johnson <i>et al.</i>	1027
U.S. Patent No. 8,249,731 to Tran <i>et al.</i>	1028
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)	1029
Jacob Fraden, “Handbook of Modern Sensors” (Springer Science+Business Media, LLC, 2010)	1030
U.S. Patent No. 8,183,995 to Wang <i>et al.</i>	1031

U.S. Patent No. 7,590,542 to Williams <i>et al.</i>	1032
U.S. Patent No. 7,492,787 to Ji <i>et al.</i>	1033
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)	1034
U.S. Patent No. 7,593,718 to Gorday <i>et al.</i>	1035
U.S. Patent Pub. No. 2006/0135067 to Dunko	1036
David Chen <i>et al.</i> , “Low-Cost Asset Tracking using Location-Aware Camera Phones” (Applications of Digital Image Processing XXXIII, Proceedings Volume 7798, 2010) (available at https://www.spiedigitallibrary.org/conference-proceedings-of-spie/7798/77980R/Low-cost-asset-tracking-using-location-aware-camera-phones/10.1117/12.862426.short)	1037
Yu-Shuo Chang <i>et al.</i> , “CSR: a Cloud-assisted Speech Recognition Service for Personal Mobile Device” (2011 International Conference on Parallel Processing) (available at https://ieeexplore.ieee.org/document/6047199)	1038
Zheming Tu <i>et al.</i> , “Speech Recognition Over the Internet Using JAVA” (1999 IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings, ICASSP99) (available at https://ieeexplore.ieee.org/document/758414)	1039
Declaration of Tony Corsini on the scanned copies of pages of Jim Doherty <i>et al.</i> , “Wireless Home Networking Simplified” (Cisco Press, 2007).	1040
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).	1041
Declaration of Rachel J. Watters on the scanned copies of Chapter 1 and Appendix D of Andrea Goldsmith, “Wireless Communications” (Cambridge University Press, 2005).	1042
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).	1043

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)	1044
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).	1045
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).	1046
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).	1047
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).	1048
RESERVED	1049
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)	1050
Declaration of Tony Corsini on the scanned copies of pages of Jacob Fraden, “Handbook of Modern Sensors” (Springer Science+Business Media, LLC, 2010)	1051
U.S. Patent No. 8,635,243 to Philips <i>et al.</i>	1052
Declaration of Tony Corsini on the scanned copies of pages of David Chen <i>et al.</i> , “Low-Cost Asset Tracking using Location-Aware Camera Phones” (Applications of Digital Image Processing XXXIII, Proceedings Volume 7798, 2010)	1053

Claims Listing

Claim	Limitation #	Limitation
1	1[PR][1]	A system for operating multiple actuators in response to captured human voice data,
	1[PR][2]	for use with a wireless network in a building and with a controlled device that comprises a second actuator in the building, the system comprising:
	1[A][1]	an Internet-connected server device external to the building configured for processing first and second human voice data,
	1[A][2]	and to produce first and second messages respectively in response to the processing;
	1[B]	a client device in the building configured to communicate with the controlled device and with the server device over the wireless network;
	1[C]	a microphone in the client device for capturing the first and second human voice data;
	1[D]	and a first actuator in the client device,
	1[E]	wherein the system is operative for sending to the server by the client device via the wireless network, the captured first and second human voice data,
	1[F][1]	and wherein the system is further operative for receiving from the server the first and second messages in response to the sending,
	1[F][2]	to operate the first actuator in the client device in response to the received first message,
	1[F][3]	and to operate the second actuator in the controlled device in response to the received second message.

Claim	Limitation #	Limitation
2	2	The system according to claim 1, wherein the processing comprises performing a voice recognition algorithm for identifying the voice of a specific person.
3	3	The system according to claim 1, wherein the client device further comprises a sensor that outputs sensor data that responds to a physical phenomenon, wherein the system is further configured for 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 system 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 system according to claim 1, wherein each of the first and second actuators is configured for directly or indirectly affecting, changing, producing, or creating a physical phenomenon.
6	6	The system 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 system 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.

Claim	Limitation #	Limitation
8	8	The system according to claim 7, wherein the multiple microphones are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array.
9	9	The system 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 system according to claim 1, wherein the client device is addressable in the wireless network or the Internet using an address stored in a volatile or non-volatile memory for uniquely identifying the respective device in the network.
11	11	The system 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 system 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.

Claim	Limitation #	Limitation
13	13	The system according to claim 1, wherein the wireless network is a Wireless Local Area Network (WLAN) that is according to, or base on, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, Institute of Electrical and Electronics Engineers (IEEE) IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac.
14	14	The system 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 system 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 system according to claim 1, wherein the controlled device is integrated in, is part of, or is entirely included in, a household appliance having a primary function.

Claim	Limitation #	Limitation
17	17	The system according to claim 16, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation.
18	18	The system 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 system according to claim 17, wherein the appliance is a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker.
20	20	The system according to claim 16, wherein the primary function of the appliance is associated with environmental control, and the appliance is part of a Heating, Ventilation and Air Conditioning (HVAC) system.
21	21	The system 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 system 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 system 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

Claim	Limitation #	Limitation
		compactor, a smoke detector, a light fixture, or a dehumidifier.
24	24	The system 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 system 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 system according to claim 1, wherein the first actuator is a motion actuator that causes linear or rotary motion.
27	27	The system 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 system 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 system 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

Claim	Limitation #	Limitation
		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. 12,010,174 (EX1001) 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. (EX1034.) 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 '174 patent covers well-known and routine smart-home functionalities. The Director should institute review on the '174 patent to protect our robust patent system and a critical segment of our nation's economy.

Review is particularly appropriate here, because the Examiner did not conduct a thorough examination. The Examiner allowed the claims without issuing any prior art-based rejections, relying on the allowance of a parent patent that the PTAB has already determined to recite mostly garden-variety limitations. (*Infra* §II.C.) As explained herein, devices capable of receiving and responding to audio commands (*e.g.*, for controlling home appliances) were prevalent. (*E.g.*, EX1005, EX1006, EX1009.) And the recited network architecture simply reflected the conventional configuration of home wireless networks at the time.

(*Id.*; EX1007.)

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

II. THE '174 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.) The specification provides no technical details of how these well-known parts 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; 32:63-64.) 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 ’174 patent purportedly claims priority to a provisional application 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 examination was not thorough. No prior art-based rejections were raised. (EX1003, 89-103.) The Examiner issued a single Office Action containing double-patenting rejections over parent patent U.S. Patent No. 11,128,710. (EX1003, 89-103.) After a terminal disclaimer was submitted, the Examiner allowed the claims citing the entirety of claim 1 as the reason for allowance without further explication. (*Id.*, 78-86, 38-40.)

If the Examiner was relying on the parent prosecution for a thorough prior art examination, he was mistaken. The examination of the parent ’710 patent was just as deficient. For example, the parent patent went through *ex parte* appeal, where the PTAB affirmed the rejections, finding that a client device that receives a voice command triggering operation of a *controlled device* to be obvious.

(EX1004, 389-421.) On remand, the Applicant then added a trivial requirement that a voice command trigger operation of the *client device* itself. (EX1004, 344-352, 121-130.) Despite the minimal and well-known addition, the Examiner abruptly allowed the application without offering any substantive explanation. (See EX1004, 8-14; EX1002, ¶¶35-39.)

Thus, beyond its own flawed examination, the examination of the '174 patent is flawed at least for the same reasons as the examination of its parent. (EX1002, ¶40.)

D. Level of Ordinary Skill in the Art

A person of ordinary skill in the art (POSITA) in the field of the '174 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, ¶¶46-48.)

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 should 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, ¶49.)

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, ¶¶50-60).

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. 2013/0132094	“Lim” (EX1005)	§102(e)
U.S. Publication No. 2012/0198339	“Williams” (EX1006)	§102(e)
Jim Doherty <i>et al.</i> , “Wireless Home Networking Simplified” (Cisco Press, 2007).	“Doherty” (EX1007)	§102(a) and (b)
U.S. Patent No. 7529677	“Wittenberg” (EX1008)	§102(a), (b), and (e)
U.S. Publication No. 2012/0265528	“Gruber” (EX1009)	§102(e)
Jacob Benesty <i>et al.</i> , “Springer Handbook of Speech Processing” (Springer-Verlag Berlin Heidelberg, 2008).	“Benesty” (EX1010)	§102(a) and (b)

Prior Art Reference	Abbreviation (Exhibit No.)	Prior Art Basis
Andrea Goldsmith, “Wireless Communications” (Cambridge University Press, 2005).	“Goldsmith” (EX1042)	§102(a) and (b)
U.S. Publication No. 2013/0052946	“Chatterjee” (EX1012)	§102(e)
U.S. Patent No. 8340975	“Rosenberger” (EX1013)	§102(e)
Thomas D. Rossing <i>et al.</i> , “Springer Handbook of Acoustics” (Springer Science+Business Media, LLC, 2007).	“Rossing” (EX1014)	§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, 10-12, 16, 20-21, 23, 29	Lim
2	103	2	Lim, Wittenberg
3	103	3-4	Lim, Gruber
4	103	7-8	Lim, Benesty
5	103	13-15	Lim, Goldsmith
6	103	17-19, 22	Lim, Chatterjee
7	103	24-25	Lim, Rosenberger
8	103	9, 26-28	Lim, Rossing
9	103	1, 5-6, 10-12, 16, 20-21, 23, 29	Williams, Lim
10	103	2	Williams, Lim, Wittenberg
11	103	3-4	Williams, Lim, Gruber
12	103	7-8	Williams, Lim, Benesty
13	103	13-15	Williams, Lim, Goldsmith
14	103	17-19, 22	Williams, Lim, Chatterjee
15	103	24-25	Williams, Lim, Rosenberger
16	103	9, 26-28	Williams, Lim, 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: Lim Renders Obvious Claims 1, 5-6, 10-12, 16, 20-21, 23, and 29

1. Summary of Lim

Lim was filed on November 17, 2011, and is prior art under §102(e). (EX1005, 1.) The Examiner never found, and Applicant never cited, Lim during prosecution. Lim relates to “controlling devices such as universal remote controls” and “appliances to be controlled.” (EX1005, [0007], [0001].) Lim discloses that “a *controlling device* may be provisioned with means to receive voice input from a user, *e.g., a microphone.*” (*Id.* (emphasis added)²) Lim’s *controlling device* can be device 100 (FIG. 1a) or smart device 110 in the form of a “smart phone” (FIG. 1b). (*Id.*, [0017]-[0018], [0023]; EX1002, ¶61.)

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

² All emphasis and annotations are added unless otherwise stated.

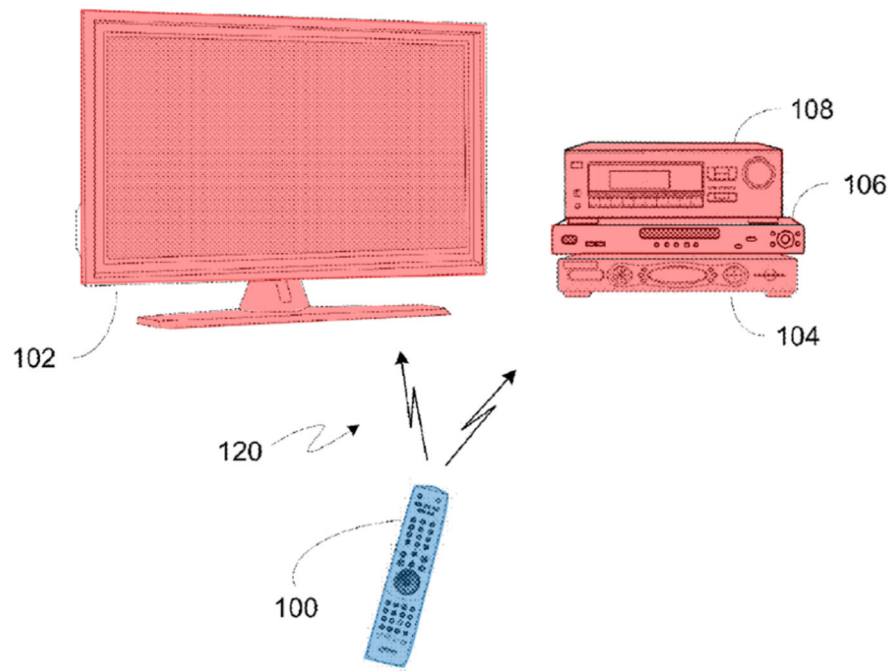


Figure 1a

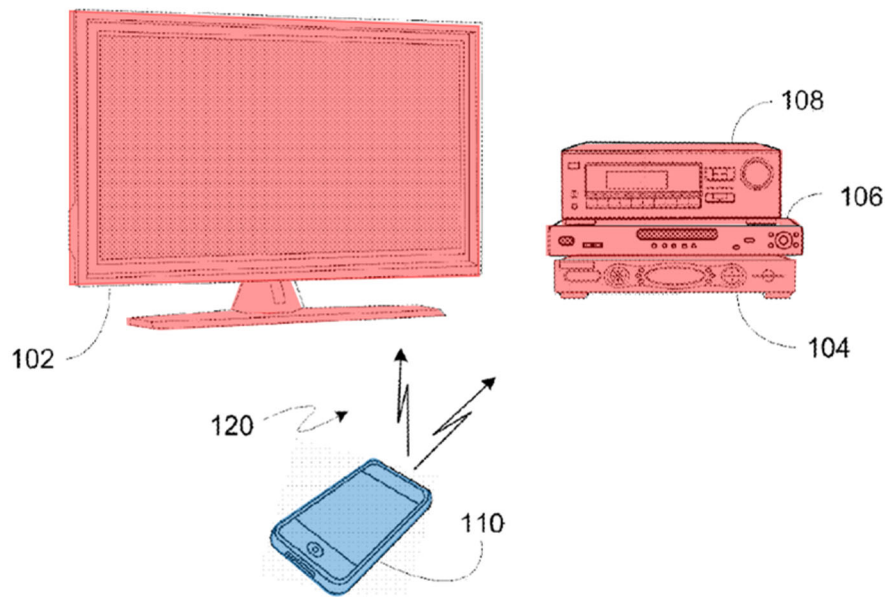


Figure 1b

Lim's *client device* receives speech commands and responds with a *visual or audio output*. For example, during a “setup and configuration” stage (e.g., FIG.

4), the *client device* can “prompt the user to speak the brand name of the appliance” and respond to user’s speech inputs with “subsequent prompts for voice input [that] may take the form of an illuminated LED and/or “speak” symbol, *a displayed text or image, a synthesized spoken message*” (*Id.*; EX1002, ¶62.)

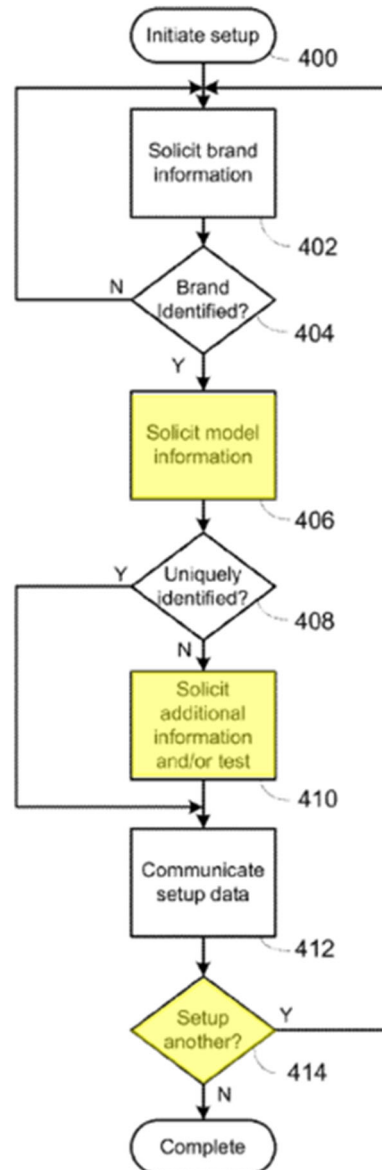


Figure 4 “Set up and configuration of controlling device”
Lim, [0025].

The *controlling device* can be used to control appliances. Specifically, the *controlling device* receives speech commands and transmits *control commands* to *appliances*, as shown in FIG. 6 below. (See EX1005, [0017], [0001]; EX1002, ¶¶63-64.)

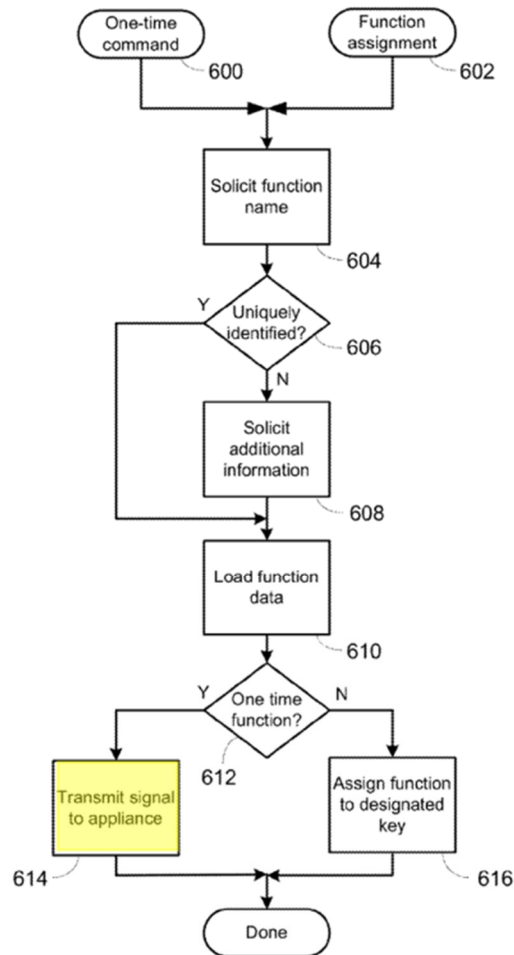


Figure 6 “Appliance command”
Lim, [0035].

In Lim, the speech commands are transmitted to a *remote server 500* for processing. (EX1005, [0027], [0034]-[0035].) Further, Lim discloses (1) the

server is accessible via the *Internet* and (2) the *controlling device* connects to the *Internet* via *WiFi* through a gateway device such as a router. (*Id.*; EX1002, ¶65.)

An exemplary network architecture is shown in FIG. 5b below. The “*remote server 500* is accessible via a network 502 such as the *Internet* or PSTN” and “access to such network [502] may be . . . via a *local network connection 508*, such as *WiFi or Bluetooth*, through a gateway device such as a router (not illustrated).” (EX1005, [0027]; EX1002, ¶¶65-66.)

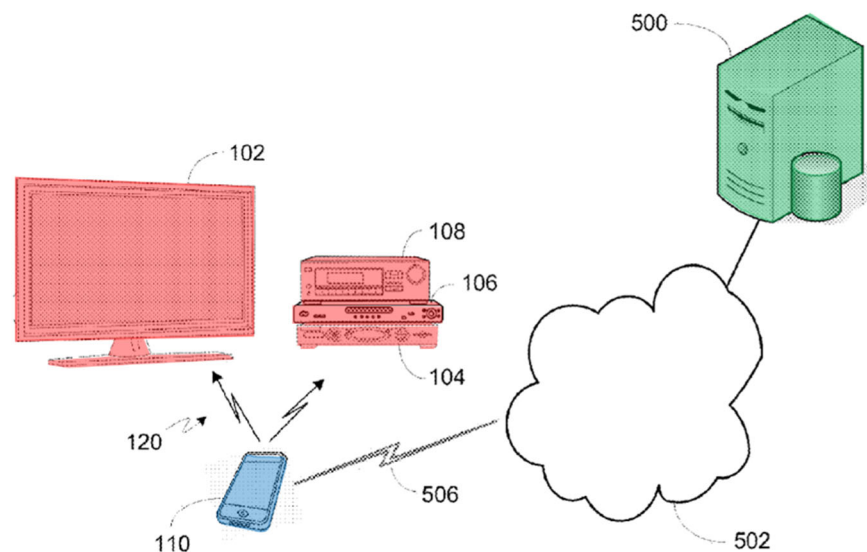


Figure 5b

Lim discloses that the *controlling device* has a microphone, a display, a speaker, various sensors, and WiFi/Bluetooth capabilities. (EX1005, [0023], FIG. 3; EX1002, ¶67.)

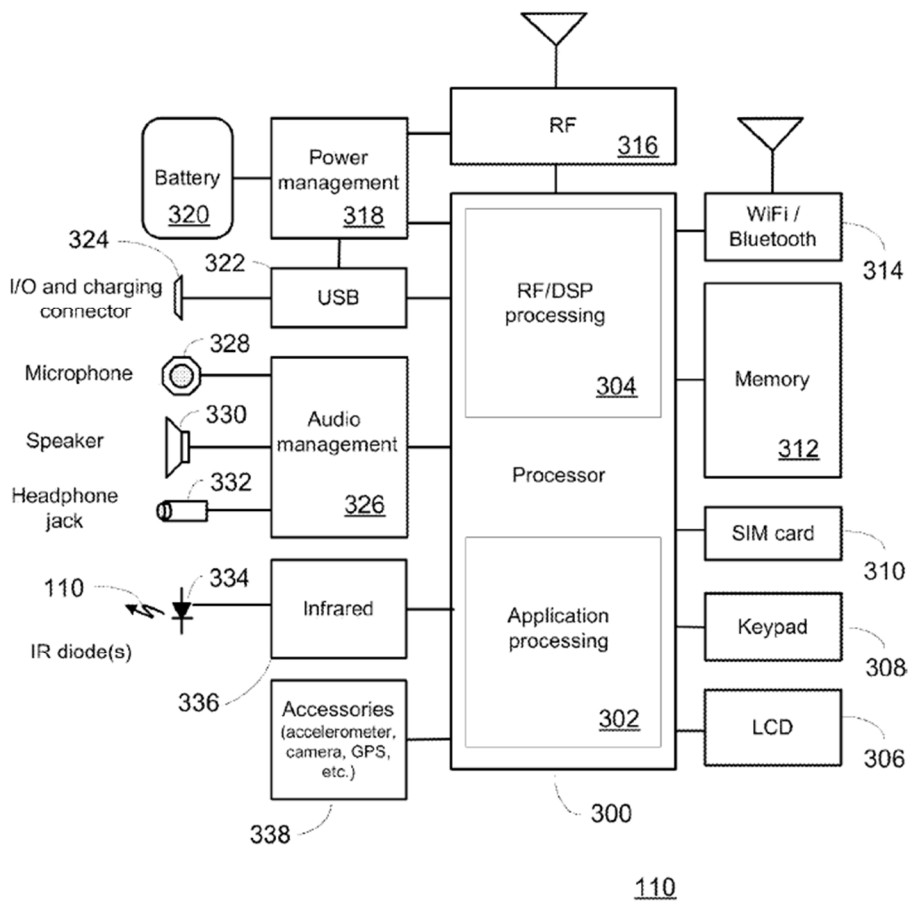


Figure 3

2. Claim 1

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

Lim’s system comprises a *client device*, a *server*, and a *controlled device*.

For example, Lim discloses “a *system* and method for setting up and configuring a *controlling device* . . . to command functions of one or more types of remotely *controllable appliances* of one or more manufacturers.” (EX1005, [0001]; EX1002, ¶¶68-69.)

Lim's system captures human voice data. "[A] *controlling device* may be provisioned with means to receive *voice input* from a user, [*e.g.*,] a *microphone*" (EX1005, [0007].) A POSITA would have understood that the system "captures" voice inputs via the microphone. (EX1002, ¶70.)

As explained below, each of Lim's "*controllable appliances*" and *client/controlling device* contains actuators, such as displays, speakers, motion actuators, lights, etc. (*Infra* §§IV.A.2.b, IV.A.2.g.) Lim's system operates these actuators in response to the captured human voice data. (*Infra* §IV.A.2.i, IV.A.2.j; EX1002, ¶71.)

b. 1[PR][2]: "for use with a wireless network in a building and with a controlled device that comprises a second actuator in the building, the system comprising:"

First, Lim's system is for use with a *wireless network* in a *building*. Referring to FIG. 5b below, Lim discloses that "all or part of said speech recognition processing [of human voice data captured by the *client/controlling device*] may be performed by programming resident on a *remote server 500* accessible via a network 502 such as the Internet" and "access to such network [502] may be . . . via a *local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router.*" (EX1005, [0027].) A POSITA would have understood what Lim describes is a home wireless network (e.g., home WiFi network), provided by a local router that local devices (e.g., *controllable*

appliances, *controlling device*) connect to for Internet access. WiFi and Bluetooth networks were known to be wireless, consistent with the description of the '174 patent. (EX1042, 600-601; EX1013, 7:48-51; EX1021, 5:22-26; EX1001, 152:18-32, 8:25-32; EX1002, ¶¶72-74.) Lim's wireless network reflects a common configuration for a home wireless network in 2012 and is also consistent with the '174 patent. (EX1005, [0027]; EX1007, 10-11, 100, 38-39, 228, 229; EX1001, 28:7-10; EX1002, ¶¶75-76.)

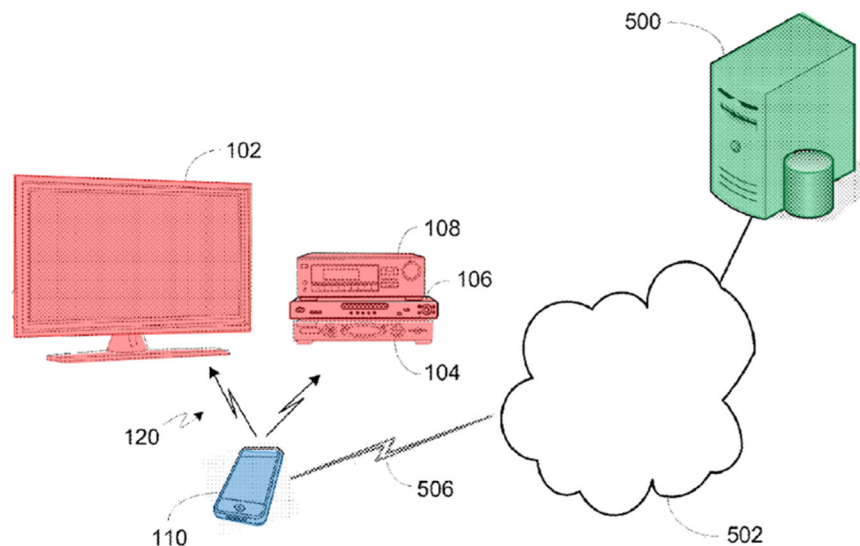


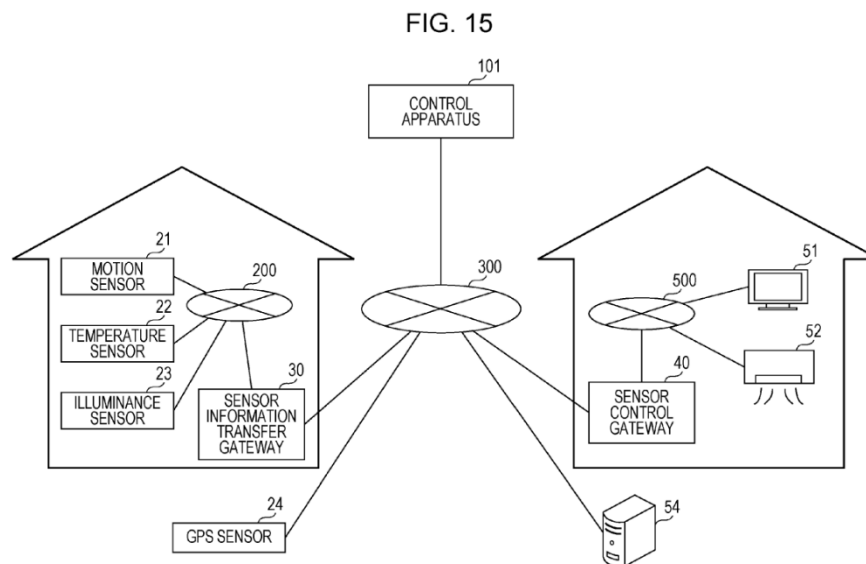
Figure 5b

Second, Lim discloses *controlled devices*. “[A] *controlling device 100* is configured to *control* various *controllable appliances*, which can include, *televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes (‘STBs’), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers.*” (EX1005, [0017].) These

“controllable appliances” are controlled devices. (EX1002, ¶77.)

Lim’s *controlled devices* have actuators. For example, the appliances (*e.g., television 102, DVD players*) have displays, which are actuators. (EX1014, 175; EX1004, 396; EX1025, FIG. 15, EX1002, ¶78.) The ’174 patent explains that “[a]n actuator may be a display for presentation of visual data or information, commonly on a screen.” (EX1001, 25:33-38, 40:51-41-7, 109:41-46, 110:60-66; 144:64-145:1.)

This is further consistent with prior findings of the PTAB in its Decision on Appeal in the parent ’710 patent, affirming the Examiner’s rejections. (EX1004, 396 (finding that “image display device 51 and air conditioner 52 [] are controlled devices with actuators”); EX1002, ¶79-80.)



(EX1025, FIG. 15.)

Lim's appliances (*e.g.*, **television 102, DVD players**) also have speakers, which are actuators (*e.g.*, operating using the moving coil principle shown below). (EX1002, ¶¶81-83; *e.g.*, EX1018, 122; EX1018, 123; EX1014, 763.) The '174 patent explains that "an actuator may be an *electroacoustic actuator, such as a sounder* . . . A loudspeaker (a.k.a. speaker) is a sounder[.]" (EX1001, 105:48-61.)

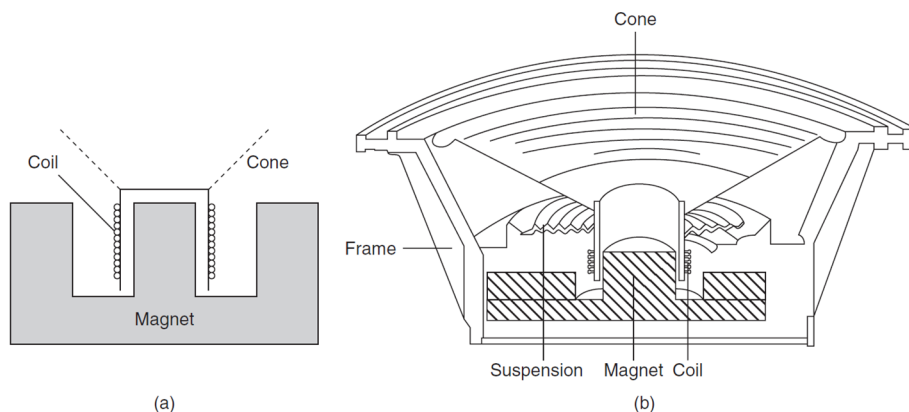


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

(EX1018, 123, Figure 10.8.)

Lim's appliances (*e.g.*, "**drapery**", "**fans**", "**HVAC systems**") have motion actuators. (EX1002, ¶¶84-87.) For example, it was well-known that HVAC has motion actuators, as shown below. (*Id.*; EX1020, 401.) Further, "**VCRs**", "**CD players**", and "**DVD players**" have motion actuators (*e.g.*, for spinning the disk). (EX1018, 226-227, 280, Figure 23.7.) The Board agreed that an air conditioner has actuators during prosecution. (EX1004, 396; EX1002, ¶85.)

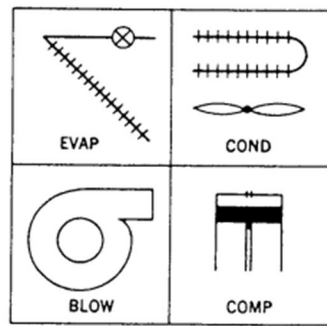


Fig 13-1. Self contained air conditioner

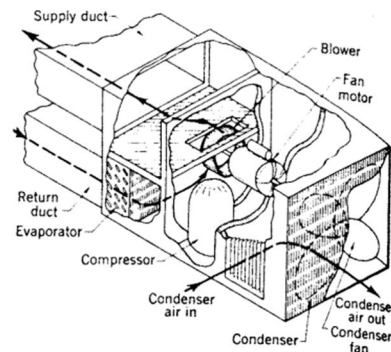


Fig 13-3. Horizontal self contained air conditioned unit

(EX1020, 401.)

Lim's *controlled devices* are (or certainly suggested to be) also in the user's *home*. It would have been obvious that "appliances" such as "*televisions*", "*home lighting*", "*drapery*", "*HVAC systems*", etc., would be placed in residential homes.

(EX1005, [0017]; EX1024, FIG. 1; EX1025, FIG. 15; EX1002, ¶88.)

c. 1[A][1]: "an Internet-connected server device external to the building configured for processing first and second human voice data,"

Lim discloses an *Internet-connected server device* external to the *building*. Lim discloses that "all or part of said speech recognition processing . . . may be performed by programming resident on a *remote server 500* accessible via a network 502 such as the *Internet* or PSTN." (EX1005, [0027].) Lim's remote server is consistent with the description of the cloud server in the '174 patent. (EX1001, 32:25-42.) A POSITA would understand and it would be obvious that the *remote server 500* is external to the user's *home* because it is "remote." (EX1002, ¶89-90; EX1005, [0007] ("remotely at a Web server"), [0024]

(differentiating between “remote server” and “local PC”).) In 2012, the benefits for having a cloud server (i.e., outside a user’s home) perform resource-intensive tasks such as speech processing were well-known and widely implemented. (EX1002, ¶¶91-92; EX1008, 8:26-32; EX1006, [0022], Figure 2; EX1009, [0080], [0100] – [0102], [0110] – [0117].)

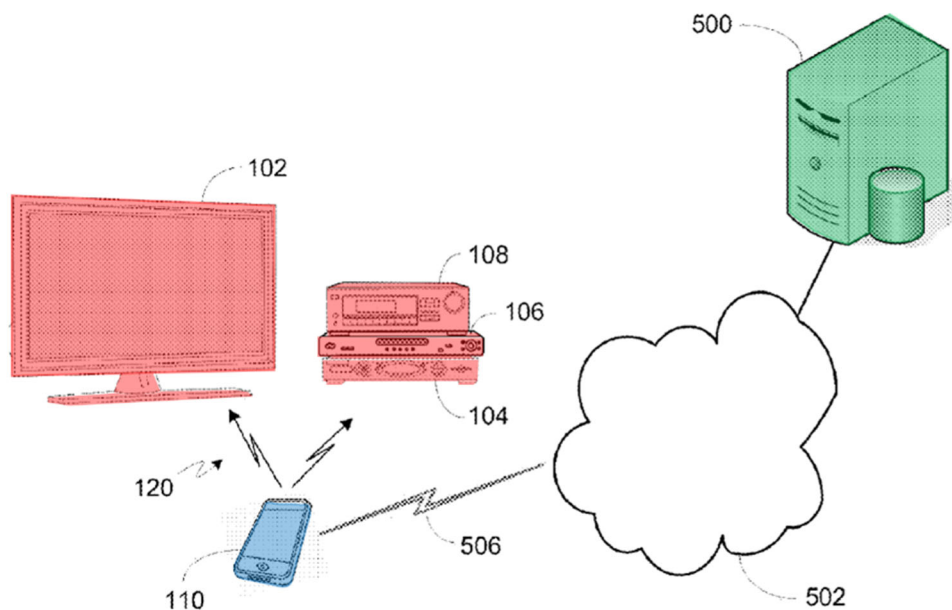


Figure 5b

Per the claim language of claim 1, “**first human voice data**” includes human voice data that causes an actuator in the **client device** to operate in the recited manner. (*Infra* §§IV.A.2.g, IV.A.2.i.) The human voice commands in Lim causing actuators (e.g., speaker, display) in its **controlling device** to operate in the recited manner thus disclose “**first human voice data.**” (EX1002, ¶93.)

Lim discloses many instances where the *client device* receives speech

commands (“*first human voice data*”) that causes its actuator (e.g., speaker, display) to operate. For example, Lim discloses a “setup and configuration” stage for the *client device* depicted in FIG. 4. (EX1005, [0025], [0013].) “[D]uring setup and configuration of the controlling device, a user may provide appliance identification via *spoken input*. . . .” (EX1005, [0025].) Lim discloses that these spoken inputs include “*speak[ing] the brand name of the appliance*,” “*spell[ing] the first few characters of the brand name*,” “*speak[ing] the name of a geographic region or a service provider*,” “*provid[ing] spoken appliance type and/or model number information*” and “*indicat[ing] if they wish to configure additional appliances*.” (*Id.*, [0027]-[0030], FIG. 4; EX1002, ¶94.) Lim discloses that each such voice data is sent to the *remote server 500* for “speech recognition processing”. (EX1005, [0027]; EX1002, ¶95.) Based on the processing, the *client device* provides a next prompt via “a displayed text or image”, “a synthesized spoken message”, etc. (*Id.*; EX1005, [0028]-[0030].)

Per the claim language of claim 1, “*second human voice data*” includes human voice data that causes an actuator (e.g., speaker, display, motion actuators) in the “*controlled appliances*” to operate in the recited manner. (*Infra* §IV.A.2.j; EX1002, ¶96-97.) In FIG. 6 below, “at step 604 the user may be prompted [by the *client device*] to *speak the desired function name*”, which “may be digitized and presented to a *speech recognition means for processing* . . . on a *remote server*

500. . . .” (EX1005, [0035], [0036]; EX1002, ¶98.) Based on the processing, the *client device* transmits commands to *appliances* in block 612. *Id.*

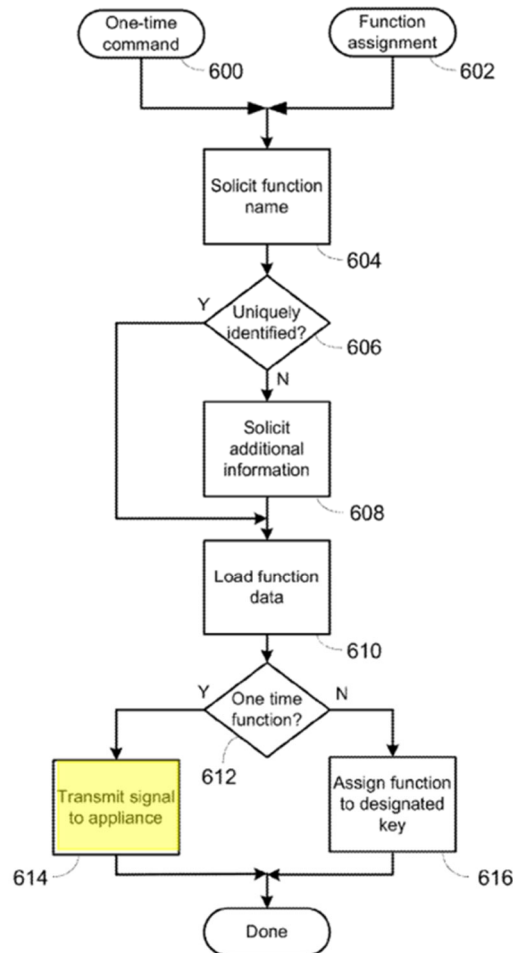


Figure 6 “Appliance command”
Lim, [0035].

To the extent that the first and second human voice data are captured/processed together (*e.g.*, as parts of one query) (which is not conceded), Lim certainly suggests and a POSITA would find it obvious that a user command related to operation of an appliance (*e.g.*, FIG. 6) would also trigger a visual/audio

feedback on the controlling device like it does in FIG. 4 to provide confirmation.
(EX1002, ¶144.)

d. 1[A][2]: “and to produce first and second messages respectively in response to the processing;”

As discussed, Lim discloses that the *Internet-connected server device* processes the *first* and *second human voice data*. (*Supra* §IV.A.2.c; EX1005, [0027], [0035]; EX1002, ¶¶99-100.) It would have been obvious that *speech processing results* (i.e., *first message* and *second message*) are produced by the server. Indeed, at the time of the invention, remote speech processing servers routinely transmit results related to speech recognition back to the requesting device. (*E.g.*, EX1009; EX1038, 307; EX1039, 2367; EX1002, ¶101.)

Indeed, in the “setup and configuration” stage in FIG. 4, after its *remote server* processes a spoken input, a next prompt is provided by the *client device* accordingly. For example, at step 402, a user’s voice input specifying “*brand name of the appliance*” may be presented to *remote server 500* for speech recognition processing. (EX1005, [0027].) “At step 404, *if the speech recognition means is unable to uniquely identify the desired brand*, step 402 may be repeated in order to solicit additional identifying information from the user.” (EX1005, [0028]; EX1002, ¶102.)

Thus, Lim’s next prompt is based on *speech processing result(s)* (i.e., first

message) produced by the *server*. (EX1002, ¶101-102.) The *speech processing result(s)* may be the recognized speech (*e.g.*, text strings), a result of analyzing the recognized speech (*e.g.*, semantic processing to determine meaning and adequacy of the speech), and/or information associated with the next prompt (*e.g.*, text, audio, information to display, information triggering a visual or auditory output such as flashing a light indicator or playing a sound). (*Id.*; *see, e.g.*, EX1009; EX1038, 307; EX1039, 2367.)

Lim discloses other instances of *first messages* produced by the *remote server 500* for the *client device* to determine the next prompt. With reference to FIG. 4 below, the *remote server 500* process spoken inputs and produces *speech processing result(s)*, for example, for determining “[i]f the initially provided model information is insufficient to uniquely identify the appliance” in block 406 or “if [the user] wish to configure additional appliances” in block 414. (EX1005, [0029]-[0030].) It would have been obvious that such *speech processing result(s)* are produced by the *server* “in response to the processing” because the *first message* is generated based on the *server*’s processing of first human voice data. (EX1002, ¶103-105.)

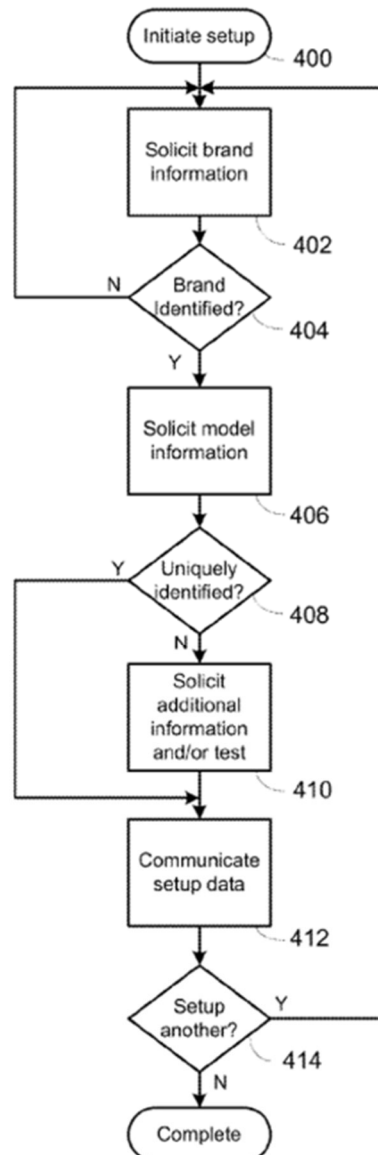


Figure 4 “Set up and configuration of controlling device”
Lim, [0025].

Similarly, Lim’s *remote server* produces *speech processing result(s)* after processing *second human voice data*. For example, referring to FIG. 6 below, after receiving voice input on “*function name*” at step 604, “*[o]nce desired function has been uniquely identified*”, at step 610 the indicated command code

data may be provisioned to the controlling device.” (EX1005, [0036].) Thus, the second message would be the *speech processing result(s)* from the speech recognition server to determine whether to perform step 610. (EX1002, ¶¶106-107.) The second message can be the recognized speech (*e.g.*, text strings), a result of further analyzing the recognized speech, etc. (EX1002, ¶107.) It would have been obvious to a POSITA that the *second message* is produced by the *server* “in response to the processing” because the *speech processing result(s)* is generated based on the server’s processing of the second human voice data. (*Id.*; EX1002, ¶110.)

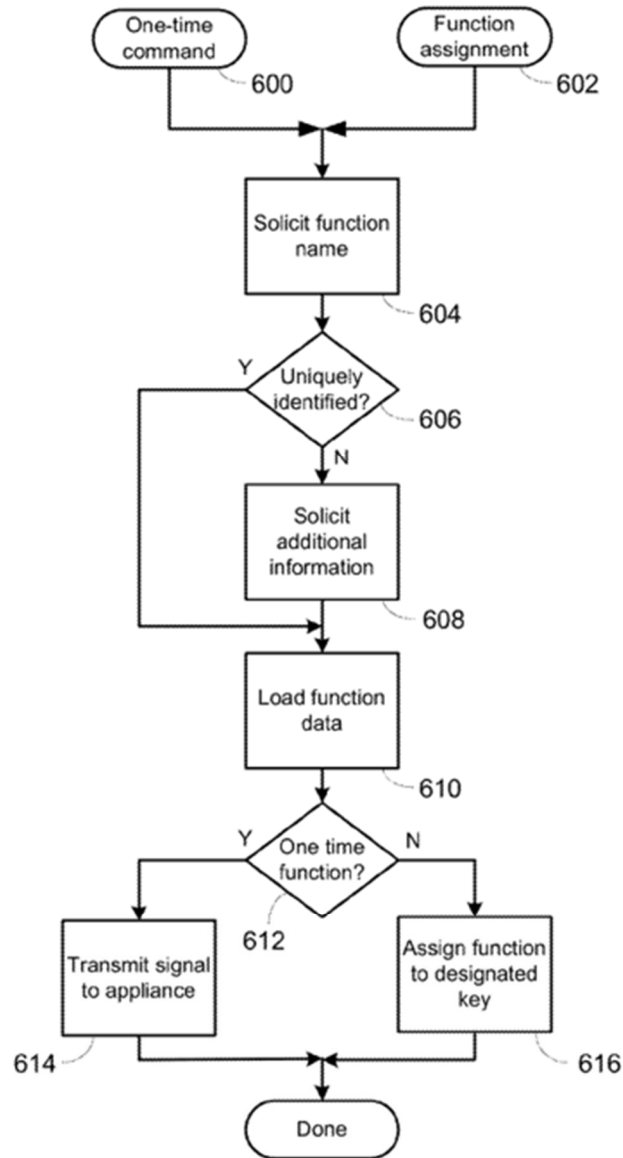


Figure 6 “Appliance command”
Lim, [0035].

Another example of the “*second message*” is “*command code data*” produced by the *server* in block 610 after processing *second human voice data*. “Once the desired function has been uniquely identified, at step 610 the *indicated*

command code data may be provisioned to the *controlling device*. . . such provisioning may comprise downloading the *data* itself, supply of a *pointer* to an entry in a local database, etc.” (EX1005, [0036].) The command code data defines “command values for use in transmitting command signals to controllable appliances” and is downloaded “from an *off-platform database of codesets* located . . . on a *remote server 500*[.]” (EX1005, [0020], [0030].) The library of codesets “may be located externally at a local PC or *remote server device* accessible to *smart device 110* for the download of *only those code sets necessary for operation* of the configured appliances[.]” (EX1005, [0022], [0024]; EX1002, ¶¶108-109.)

It would have been obvious that such “*command code data*” is produced by the *server* “in response to the processing” because it is obtained and transmitted based on the processing of the spoken input. (EX1002, ¶110; EX1005, [0020].)

e. 1[B]: “a client device in the building configured to communicate with the controlled device and with the server device over the wireless network;”

First, Lim discloses a *controlling device (e.g., smartphone)* that is a *client device*. (EX1002, ¶111-112.) “[A] *controlling device* may be provisioned with means to receive voice input from a user, [e.g.,] a microphone”. (EX1005, [0007], [0017].) Lim discloses that the *controlling device* may be “a *smart device 110* as shown in FIGs. 1b and 5b. (*Id.*, [0018], [0023]; EX1002, ¶112.)

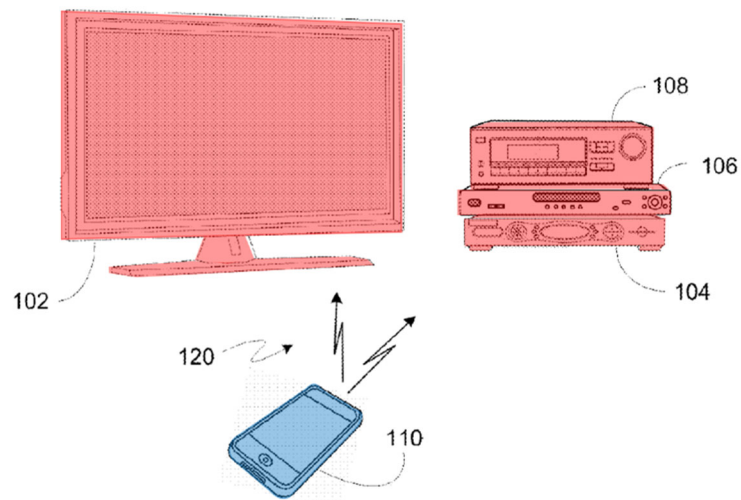


Figure 1b

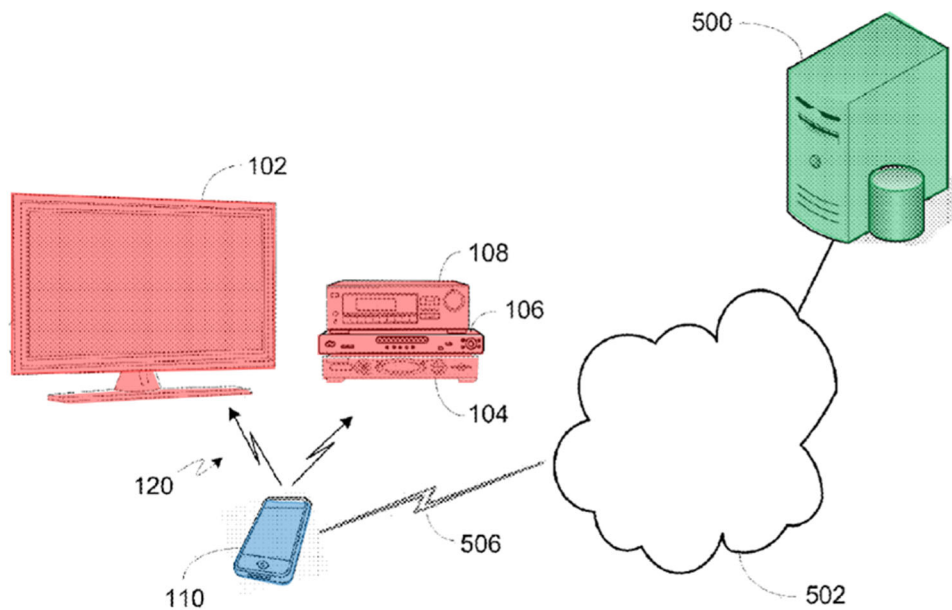


Figure 5b

Lim's *controlling device* is a client device because it communicates with a server (*remote server 500*) to, for example, provide the captured spoken inputs for processing. (EX1005, [0027]; EX1002, ¶¶113-114.) Further, Lim's *client device* is connected to a *wireless network* (e.g., home WiFi) via a router, which in turn

accesses the *remote server 500* over the Internet. (*Supra* §IV.A.2.b; EX1042, 600-601; EX1013, 7:48-51; EX1021, 5:22-26; EX1001, 152:18-32; 8:25-32.)

Lim’s *client device* also communicates with the *controlling device* over the *wireless network*. Lim discloses that a “*controlling device 100* is configured to control various *controllable appliances, such as a television 102* . . . the *controlling device 100* may be capable of transmitting commands to the *appliances*, using any convenient IR, RF, Point-to-Point, or *networked protocol 120*.” (EX1005, [0017].) As shown in FIG. 1b below, *client/controlling device 110* transmits commands to the *appliances* via *networked protocol 120*. (EX1002, ¶115.) Lim discloses that the *smart device 110* “may include . . . *WiFi and/or Bluetooth wireless interface(s) 314*[.]” (EX1005, [0023].) Further, in 2012, home appliances that are WiFi-enabled were well-known. (EX1012, [0018], EX1007, 321.) Thus, a POSITA would have understood that Lim’s *client device* would communicate with *appliances* over a WiFi wireless network, that is, the same *wireless network* that the *client device* uses to communicate with the *server*. (EX1002, ¶116.)

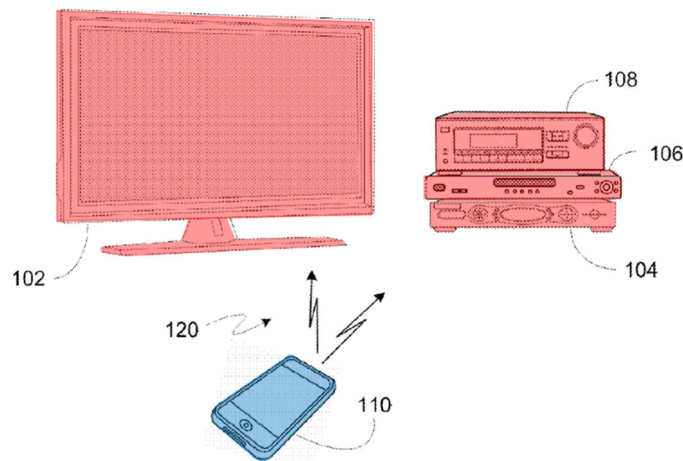
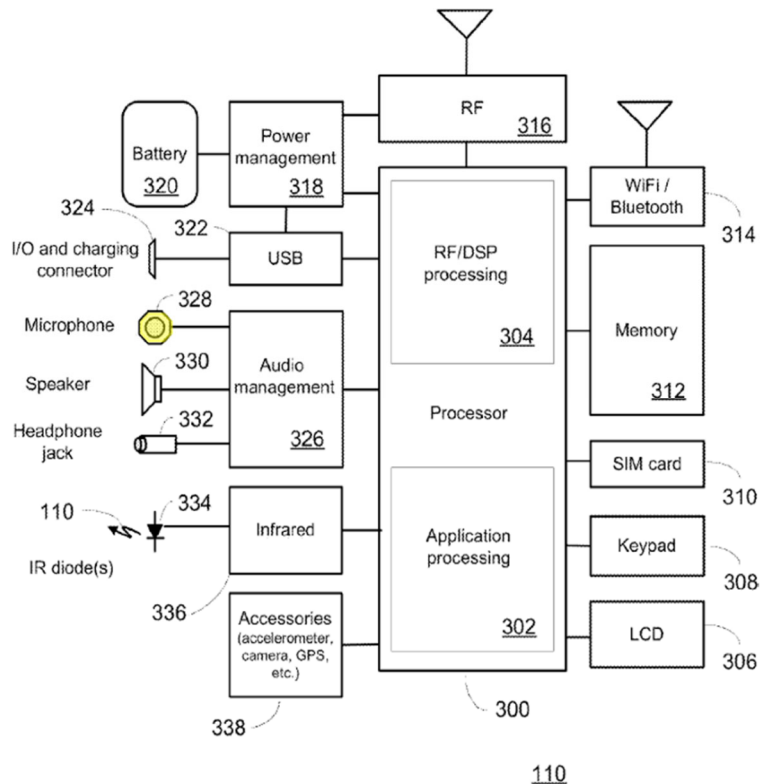


Figure 1b

Lim discloses or certainly suggests that the *client device* is in the same *building*. Lim discloses that the *smart device 110* can be “a smart phone” and other portable devices. (EX1005, [0018].) It would have been obvious that these are devices that users can carry with them, including when the user is at *home* to conveniently control home appliances. (EX1002, ¶117; EX1021, 1:53-58, 2:19-26.)

f. 1[C]: “a microphone in the client device for capturing the first and second human voice data;”

Lim’s *client/controlling device*, “illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated *microphone 328*, speaker 330, and headphone jack 332 [.]” (EX1005, [0023]; [0019], EX1002, ¶120.)



It would have been obvious that the *first human voice data* and *second human voice data* are captured by the microphone. (EX1002, ¶120.) When “the *controlling device* [] prompt[s] the user to speak the *brand name of the appliance* to be added to the configuration of the controlling device,” a POSITA would have understood that the *client device* uses its *microphone 224/338* to capture that voice input. (*Id.*; *Supra* §IV.A.2.c; EX1005, [0027], FIG. 4.) So are subsequent voice inputs such as the “*desired function name*” (i.e., “*second human voice data*”). (EX1005, [0035].)

g. 1[D]: “and a first actuator in the client device,”

Lim’s *client device* has a *speaker* and a *display*. Referring to FIG. 3 below,

the *smart device* may include “an *LCD display 306*; . . . an audio management system 326 with associated microphone 328, *speaker 330*[.]” (EX1005, [0023], [0019].) As discussed above, both speakers and displays are actuators. (*Supra* §IV.A.2.b; EX1002, ¶¶119-120.)

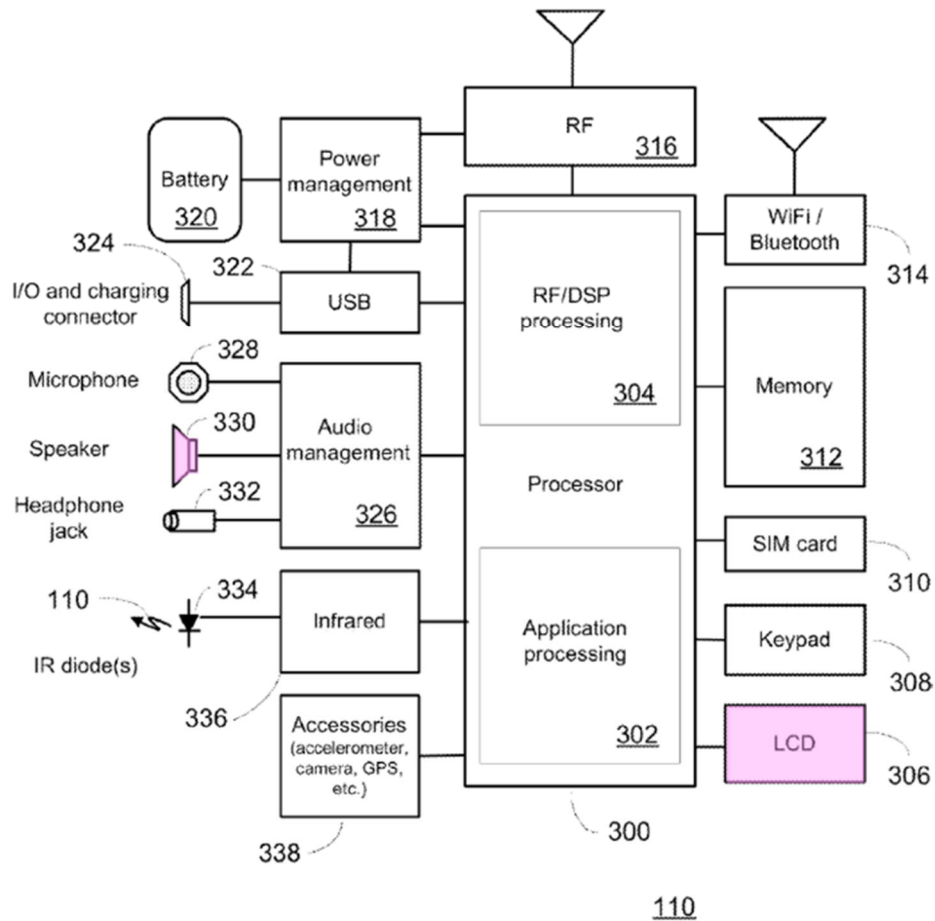


Figure 3

Moreover, Lim’s *controlling device 100* “may include, . . . , a means 202 to provide feedback to the user (*e.g.*, *one or more LEDs*[.]” (EX1005, [0019].) LEDs are actuators. The ’174 patent explains an “actuator may be a light source

that emits visible or non-visible light . . . The light source may be an electric light source . . . based on Solid-State Lighting (SSL) such as a *Light Emitting Diode (LED)*[.]” (EX1001, 39:10-21.) It would have been obvious that *smart device 110* is similarly equipped with *LEDs*. (EX1002, ¶121.)

h. 1[E]-1[F][1]: “wherein the system is operative for sending to the server by the client device via the wireless network, the captured first and second human voice data,” “and wherein the system is further operative for receiving from the server the first and second messages in response to the sending,”

Lim’s *client device* sends the captured human voice data to the *server* via the *wireless network*. (EX1002, ¶¶122-123.)

Regarding the *first human voice data*, Lim discloses that “at step 402 the *controlling device* may prompt the user to speak the *brand name of the appliance* . . . The resulting *spoken syllables* may be digitized and *presented* to a *speech recognition means* . . . resident on a *remote server 500* accessible via a network 502 such as the Internet.” Further, “access to such network . . . may be via a *local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router.*” (EX1005, [0027]; EX1002, ¶124.)

It would have been obvious to a POSITA that other instances of “*first human voice data*” are similarly sent to the server for processing, including all of the voice inputs described with respect to FIG. 4 below: “*speak[ing] the brand name of the appliance*”, “*spell[ing] the first few characters of the brand name*”,

“*speaking the name of a geographic region or a service provider*”, etc.

(EX1005, [0028]-[0030], FIG. 4; EX1002, ¶125.)

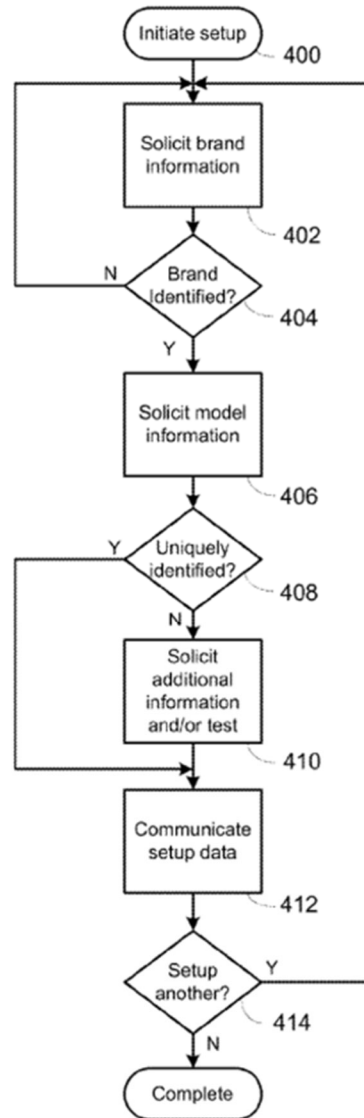


Figure 4 “Set up and configuration of controlling device”
Lim, [0025].

Similarly, for the *second human voice data*, Lim discloses that, “at step 604 the user may be prompted to speak the *desired function name*. . . . The *resulting spoken syllables* may be digitized and *presented* to a speech recognition means for

processing”, which “may be performed by programming resident on a *remote server* accessible via a network such as the Internet or PSTN.” (EX1005, [0035], [0034].) It would have been obvious that the “*second human voice data*” is sent over the same *wireless network* (e.g., home WiFi). (EX1002, ¶126.)

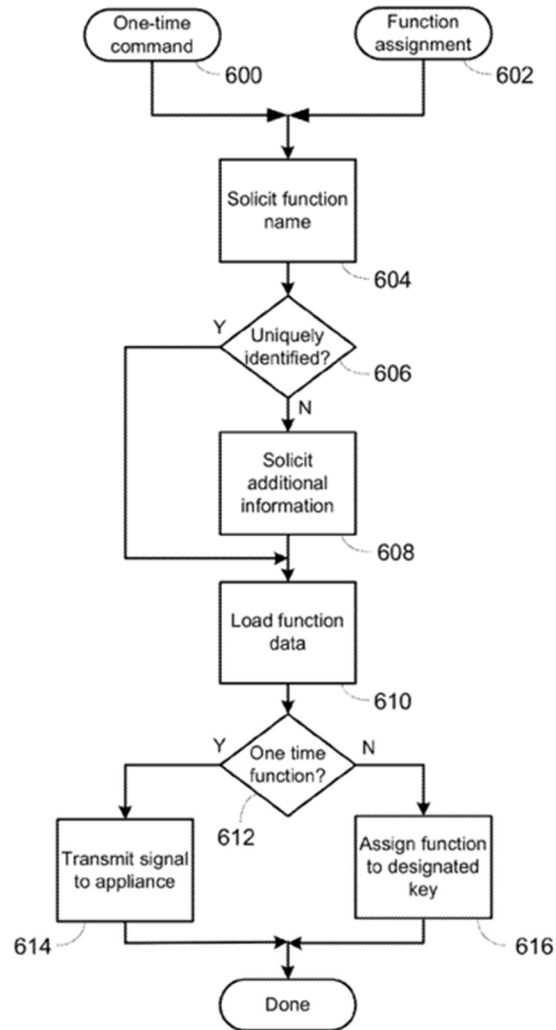


Figure 6 “Appliance command”
Lim, [0035].

As discussed above, in response to the processing, the *remote server 500*

produces speech processing results (the *first message* and *second message*).

(*Supra* §IV.A.2.d; EX1002, ¶¶127-130.) Further, it would have been obvious to a POSITA that the system (*e.g.*, the *client device*) receives these messages to output the next prompt. (EX1005, [0028], [0036]; EX1002, ¶¶129-130.)

It would have been obvious that such *first message* and *second message* are received by the *client device* from the *server* “in response to the sending of the captured first and second human voice data” because the messages are generated based on the *server*’s processing of human voice data. (EX1002, ¶131.)

i. 1[F][2]: “to operate the first actuator in the client device in response to the received first message,”

As discussed, Lim’s *client device* has actuators including a *speaker*, a *display*, and *LEDs*. (*Supra* §§IV.A.2.g, IV.A.2.b; EX1005, [0019], EX1005, [0023], FIGs. 2 and 3; EX1002, ¶¶132-133.)

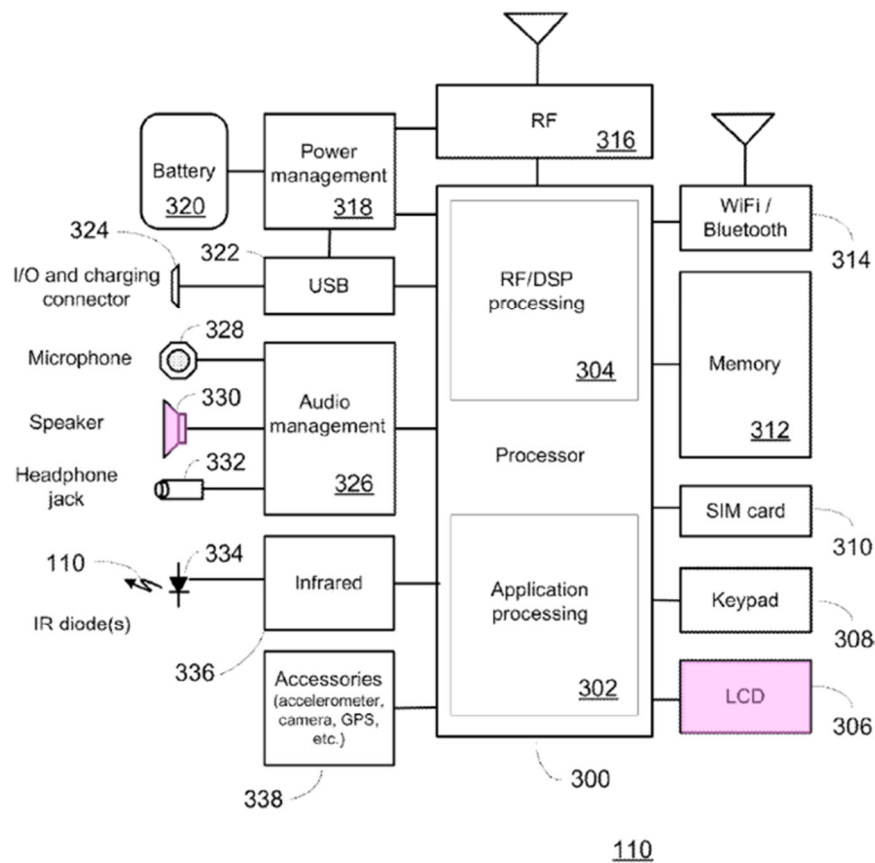
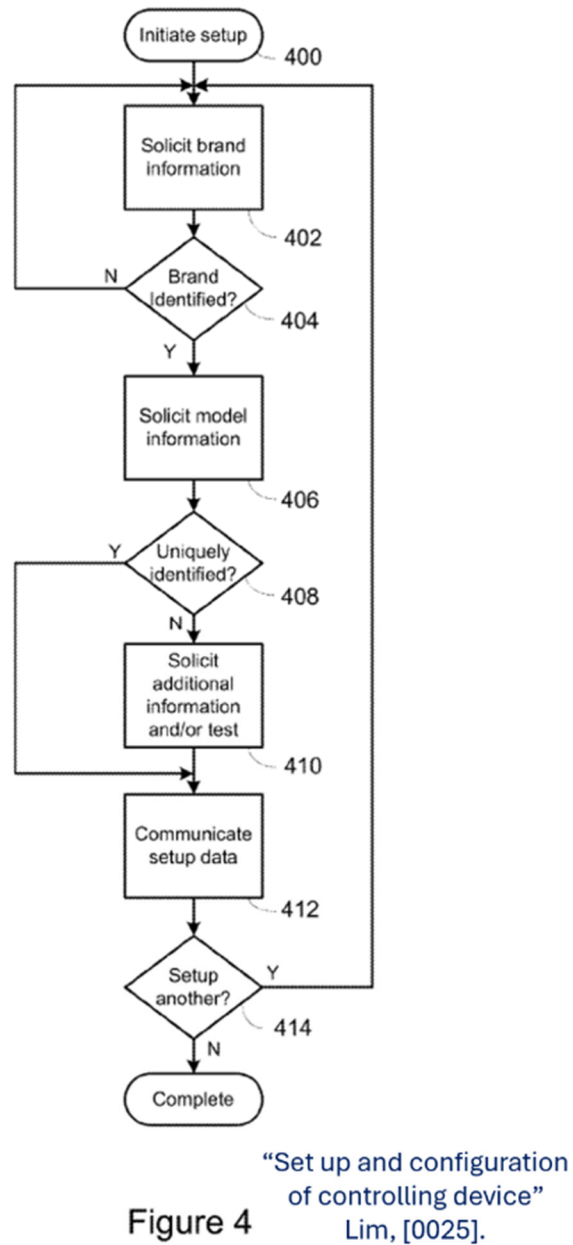


Figure 3

Lim discloses operating a *first actuator* in the *client device* in response to the received *first message*. In response to *speech processing result(s)* from the *server 500*, the *client device* provides a next prompt in “the form of an *illuminated LED* and/or ‘*speak*’ *symbol*, a *displayed text or image*, a *synthesized spoken message* (e.g., where controlling device *user feedback means 202* comprises speech synthesis means and a *speaker*, or where a smart device app has access to *speaker 330*), etc.” (EX1005, [0027].) Lim discloses that the operation of actuators is based on *speech processing result(s)*. “[I]f the speech recognition

means is unable to uniquely identify the desired brand” in block 404, the controlling device may “*solicit* additional identifying information from the user,” “*present* [] a short list of possible brand identities,” “*prompt* [user] to spell the first few characters of the brand name,” and “*request* [user] to speak the name of a geographic region or a service provider.” (*Id.*, [0028], [0036].) “*Once a brand has been successfully identified*, at step 406 the user may be *prompted* to provide spoken appliance type and/or model number information.” (*Id.*, [0029]; EX1002, ¶¶134-135.) Additional prompts are provided in blocks 410 and 414. (*Id.*; EX1005, [0029]-[0030].)



It would have been obvious that Lim operates its *display* and *speaker* in the *client device* in response to *first message* because the *speech processing result(s)* is needed to determine what next prompt to display or otherwise output. (EX1002, ¶136.)

j. 1[F][3]: “and to operate the second actuator in the controlled device in response to the received second message.”

As discussed, the *second message* can be speech processing result(s) that allows the *client/controlling device* to determine whether to send a command, thus operating the controlled device. (*Supra* §IV.A.2.d.) In FIG. 6, responsive to the *second message*, the *client device* obtains function data locally or remotely (block 610) and sends (block 614) the *command* to the *appliance*. (EX1002, ¶¶137-139.) “[A]n app may serve to allow use of the *smart phone* to *transmit commands* suitable for operation of *entertainment and other appliances*.” (EX1005, [0024], [0047].)

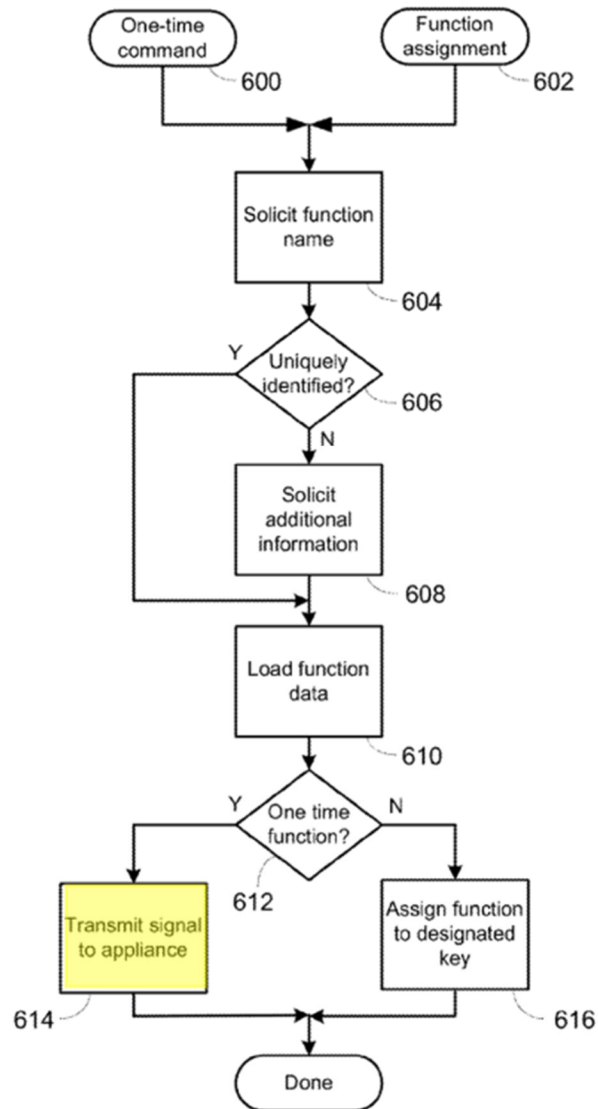


Figure 6 "Appliance command"
Lim, [0035].

Additionally or alternatively, the *second message* may be the function data or "codesets" produced by the server and downloaded by the client device in block 610. (*Supra* §IV.A.2.d; EX1005, [0024], [0030], [0036], EX1002, ¶140.)

Responsive to the received *function data*, the *client device* sends (block 614) the

command to operate the *appliance*. *Id.*

Lim discloses that the commands can be for controlling a *DVD player* to “[p]ower on,” “[f]ast forward,” “[p]ause,” “[o]pen/close.” (EX1005, [0031].) The command can “cause a device such as a *TV set or STB* to tune from a currently selected channel to the immediately previous channel[.]” (EX1005, [0033].) These operations involve operating actuators such as displays, speakers, and motion actuators. (*Supra* §IV.A.2.b; EX1001, 105:48-52, 105:61-106:4; EX1014, 763; EX1018, 280, Figure 23.7.) A POSITA would have understood that the actuators in Lim’s “*fans, HVAC systems, thermostats*” are operated in response to the *second message* in the same manner. (EX1005, [0017]; EX1002, ¶¶141-144.)

3. Claims 5/6

As discussed, the “first actuator” and the “second actuator” disclosed in Lim include: speakers, displays, LED lights, HVAC-related motion actuators, VCR-related motors etc. (*Supra* §§IV.A.2.b, IV.A.2.g; EX1002, ¶¶145-148, 151.) They are configured to directly or indirectly affect, change, produce, or create audio and/or light. *Id.* HVAC-related motors are configured to directly or indirectly affect, change, produce, or create motion, temperature, and/or humidity. (*Supra* §IV.A.2.b.) VCR-related motors cause motion of the drum head in the VCR. (*Id.*; EX1002, ¶¶149-150.)

4. Claims 10/11/12

As discussed, Lim discloses its client device is WiFi-enabled and connects to the Internet via a wireless network. (*Supra* §§IV.A.2.b, IV.A.2.c) A POSITA would have understood that this wireless network involves associating each device with a device identifier or address that uniquely identifies the device in the network. (EX1002, ¶¶152, 156-157; EX1007, 99, 108, 210, 175.) 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, ¶153.) 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 that are of the MAC-48, EUI-48, or EUI-64 type. (EX1029, 127, EX1007, 277, EX1027, 2:50-59; EX1023, 11, 13; EX1002, ¶¶158-159.)

It was common knowledge that a networked device locally stores the associated IP addresses. (EX1007, 211, Figure 10-44, 265.) Lim's *client/controlling device* includes "ROM", "RAM", may include, "a non-volatile read/write memory", etc. (EX1005, [0021], [0023].) Thus, a POSITA would have understood Lim's client device (*e.g.*, *smart device 110*) locally stores the IP addresses. (EX1002, ¶¶154-155.)

Further, Lim discloses its client device is Bluetooth-enabled. (EX1005, [0027].) It was common knowledge that Bluetooth is a wireless personal area network (WPAN) and based on the IEEE 802.15 standard. (EX1042, 601-602.) In

2012, routers having Bluetooth capabilities and Bluetooth access points/gateways with connection to the Internet were well-known. (EX1026, 59:21-24; EX1027, 4:25-27; EX1028, 4:20-22; EX1033, 1:67-2:3; EX1002, ¶¶160-162.)

5. Claims 16/20/21/23

In Lim, the “*controllable appliances* may include . . . *televisions, VCRs, DVRs, DVD players, cable or satellite converter set-top boxes (‘STBs’), amplifiers, CD players, game consoles, home lighting, drapery, fans, HVAC systems, thermostats, personal computers*, etc.” (*Supra* §IV.A.2.b; EX1005, [0017].) These disclose the recited “household appliance” in claim 16, as well as the recited “HVAC system,” “air conditioner or a heater,” “home cinema system,” “CD or DVD player,” and “light fixture” recited in claims 20-21 and 23. (EX1002, ¶¶163-166.)

6. Claim 29

Lim’s *client device* plays a digital audio content that is synthesized human speech, which comprises words, phrases, and sentences. (*Supra* §§IV.A.2.i, IV.A.2.d; EX1002, ¶¶167-169.) “This and subsequent prompts [by the *client device*] for voice input may take the form of . . . a *synthesized spoken message* (e.g., where controlling device user feedback means 202 comprises *speech synthesis means* and a *speaker*, or where a smart device app has access to *speaker 330*), etc.” (*Id.*; EX1005, [0027].)

B. Ground 2: Lim 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). (EX1008, 1.) The Examiner never found, and Applicant never cited, Wittenberg during prosecution. Wittenberg is related to “a technique for remotely processing a local audio command to control a local device.” (EX1008, Abstract, Claim 1; EX1002, ¶170.)

2. Claim 2

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2) Claim 2 is obvious in further view of Wittenberg. (EX1002, ¶171.) 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.*” (EX1008, 5:61-6:6; EX1002, ¶171.)

At the time of the invention, it would have been obvious to combine the teachings of Wittenberg with Lim. Lim and Wittenberg are analogous art to the ’174 patent. Both relate to remotely processing audio information such as speech for controlling devices or providing services through devices. (EX1005, [0007],

[0027]; EX1008, Claim 1; EX1002, ¶172.)

The Lim-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, ¶176.) A POSITA would have been motivated to combine Lim with the voice recognition of Wittenberg to achieve greater security and with a reasonable expectation of success. (EX1002, ¶¶173-177.) For example, Wittenberg expressly 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.” (EX1008, 5:61-6:6.) Wittenberg recognizes the benefits of applying its techniques in home appliances. (EX1008, 6:7-13; EX1002, ¶177.)

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

1. Summary of Gruber

Gruber was filed on September 30, 2011, and is prior art under §102(e). (EX1009, 1.) The Examiner never found, and Applicant never cited, Gruber during prosecution. Gruber relates to “virtual assistants, and more specifically to mechanisms for improving interpretation and processing of commands provided to such an assistant.” (EX1009, [0007]; EX1002, ¶¶178-179.)

2. Claim 3

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2) Claims 3-4 are obvious in further view of Gruber. (EX1002, ¶¶180, 189.) 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.*; EX1009, [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.*” (EX1009, [0366]–[0367], [0179], [0168], [0080]; EX1002, ¶¶181-182.)

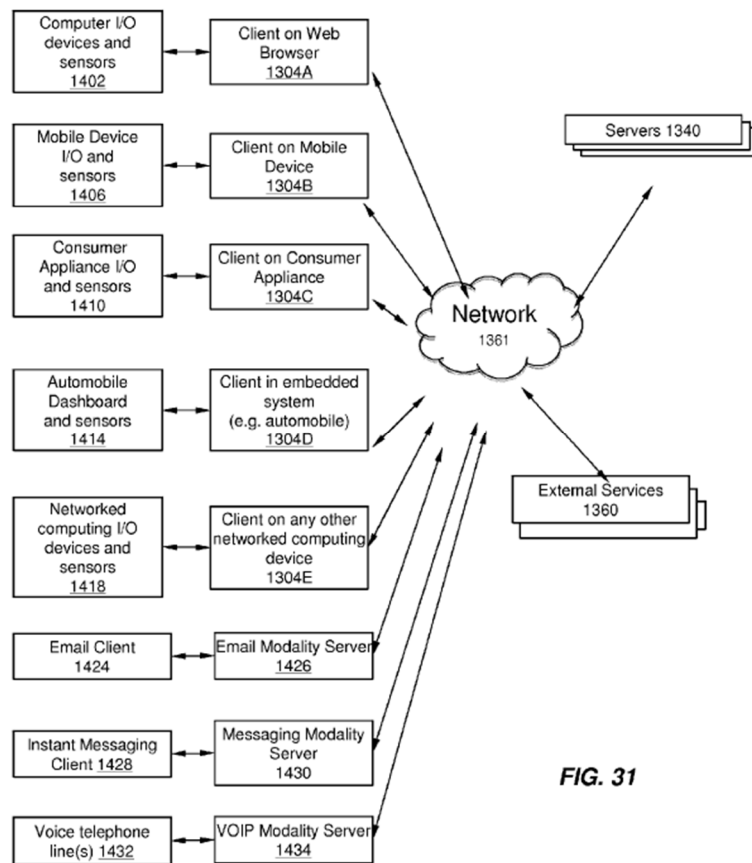


FIG. 31

Lim and Gruber are analogous art to the '174 patent. Both relate to processing and responding to human speech commands. (EX1005, [0007], [0027]; EX1009, [0007]; EX1002, ¶183.)

For example, the Lim-Gruber combination would capture sensor data (*e.g.*, temperature, location) using one or more sensors (*e.g.*, as taught by Gruber/Lim) in the controlling device and transmit such data and/or related information to Lim's server (*e.g.*, remote server 500), which can use the captured sensor data as contextual information as taught by Gruber to supplement speech processing. (EX1002, ¶186.)

A POSITA would have been motivated to combine Lim with the speech processing techniques of Gruber to achieve improved speech recognition and with a reasonable expectation of success. (EX1002, ¶¶183-188.) For example, 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.” (EX1009, Abstract.) Lim itself discloses that the client device may be a smart phone that includes sensors such as GPS and accelerometer. (EX1005, [0023].)

3. Claim 4

Gruber discloses various sensors including “an accelerometer, compass, GPS unit, . . . *light sensor, thermometer*, . . . and the like.” (EX1009, [0225].) The contextual information from these sensors may include “the *current light level, temperature, and other environmental measures*.” (EX1009, [0222].) A POSITA would have understood that photoelectric sensors are common light sensors and thermoelectric sensors are common temperature sensors. (EX1030, 116, 466; EX1002, ¶189.) It would have been obvious to a POSITA to combine the teachings of Lim and Gruber. (*Supra* §IV.C.2; EX1002, ¶190.)

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

1. Summary of Benesty

Benesty was published by at least December 11, 2007, and is prior art under §102(a), (b). (EX1041, 1.) The Examiner never found, and Applicant never cited, Benesty during prosecution. Benesty, a textbook titled “Springer Handbook of Speech Processing”. (EX1010, Title.) Benesty provides an overview of well-known speech processing techniques. (EX1010, 1021; EX1002, ¶¶191-194.)

Benesty was published by Springer and was well-known among POSITAs. (EX1002, ¶195.) A POSITA would have looked to Benesty for microphone-related techniques to, for example, implement/improve the audio signal capturing and processing in Lim. *Id.*

2. Claims 7/8

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2) Claims 7/8 are obvious in further view of Benesty. (EX1002, ¶196, 201.) Benesty discloses that, “[i]t is shown that *microphone arrays* have the ability to offer directional gains that can *significantly improve the quality of signal pickup in reverberant and noisy environments.*” (EX1010, 1021.)

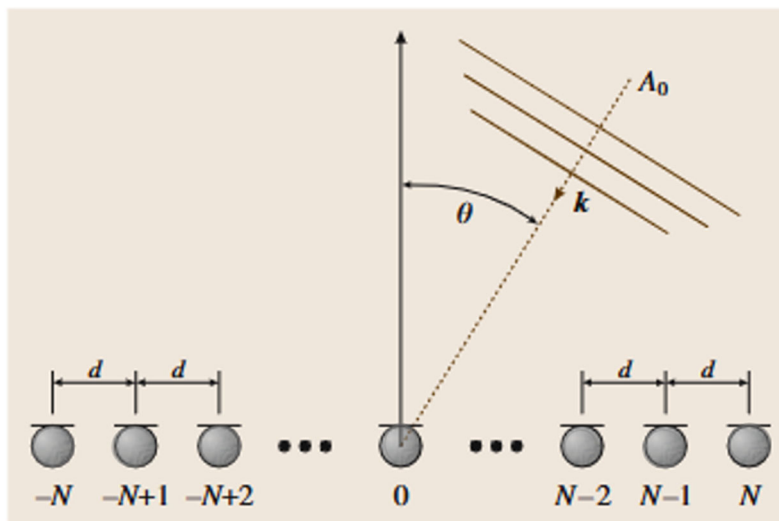


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

Benesty discloses that “for speech source localization, microphone array beam scanning and *time delay estimation* (TDE)-based localization methods are the two most widely used approaches . . . 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.” (EX1010, 1044.) Benesty also discloses that “a set of relative time differences of arrival (TDOAs) for different microphone pairs are calculated.” (*Id.*; EX1002, ¶¶201-202.)

Lim and Benesty are analogous art to the '174 patent. Both relate to capturing and processing human audio data. (EX1005, [0007]; EX1010, 1021.) A POSITA would have been motivated to combine Lim and Benesty to include the microphone array techniques of Benesty to improve audio capturing and processing and with a reasonable expectation of success. (EX1002, ¶¶198-200,

203.) For example, the Lim-Benesty combination would result in a system where multiple microphones as taught by Benesty are incorporated in Lim's client device to capture and process human voice data, yielding the predictable results of effectively improve audio capturing, especially in a room where the microphone is not placed near the mouth of the user. (EX1002, ¶199.)

E. Ground 5: Lim and Goldsmith Render Obvious Claims 13-15

1. Summary of Goldsmith

Goldsmith was published by at least January 20, 2006, and is prior art under §102(a), (b). (EX1042, 2.) The Examiner never found, and Applicant never cited, Goldsmith during prosecution. Goldsmith, a textbook titled "Wireless Communications," and published by Cambridge University Press, was well-known among POSITAs. (EX1042, Title; EX1002, ¶¶204-209.) A POSITA would have looked to Goldsmith for wireless communication systems, protocols, and techniques of applying them. *Id.*

2. Claim 13

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2) Claims 13-15 are obvious in further view of Goldsmith. (EX1002, ¶210.)

Lim discloses that the client device communicates with the *remote server 500* over the *Internet* (i.e., network 502), and *via WiFi*, a form of WLAN, through a gateway device such as a *router*. (*Supra* §§IV.A.2.b, IV.A.2.c; EX1005, [0027],

Figure 5b.; EX1002, ¶211.) Consistent with the description of the '174 patent, a POSITA would have understood that WiFi is a form of WLAN based on the IEEE 802.11 standard group comprising 802.11a/b/g standard. (EX1042, 600-601; EX1013, 7:48-51; EX1001, 8:25-32; EX1002, ¶212.)

Further, Goldsmith discloses a WLAN based on IEEE 802.11b as one of the common wireless network communication protocols well before the priority date. (EX1042, 15; EX1002, ¶¶213-214.)

“Two additional standards in the 802.11 family were developed to provide higher data rates than 802.11b. The *IEEE 802.11a wireless LAN* standard operates with 300 MHz of spectrum in the 5 GHz U-NII band . . . the 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.*” (EX1042, 16; EX1002, ¶215.)

Lim and Goldsmith are analogous art to the '174 patent. Both relate to wireless communications between devices. Goldsmith discloses various existing wireless connection systems in a chapter titled “1.4 Current Wireless Systems” (EX1042, 8-21.) It would have been obvious for a POSITA to look to Goldsmith for conventional wireless networking techniques and protocols that can be applied to Lim, which expressly discloses usage of Bluetooth and WiFi. (EX1005, [0027];

EX1002, ¶¶216-219.) Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (*e.g.*, WLAN over IEEE 802.11b, 802.11a, and/or 802.11g, 3G) with the teachings of Lim without requiring much effort and with a reasonable expectation of success. (EX1002, ¶219.)

3. Claim 14

Goldsmith summarizes in the table below the unlicensed radio frequency spectrum allocations in the U.S., which expressly shows that 802.11b WLANs are ISM Band II frequencies. (EX1042, 22-23; EX1002, ¶220.)

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 the teachings of Goldsmith and Lim. (*Supra* §IV.E.2; EX1002, ¶221.)

4. 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.*” (EX1042, 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, ¶¶222-225.)

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)*.” (EX1042, 599-600; EX1002, ¶226.)

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.

In 2012, home appliances with cellular connectivity were well-known. (EX1031, claims 14-16, 18; EX1002, ¶227.)

It would have been obvious for a POSITA to combine Lim and Goldsmith.

(*Supra* §IV.E.2; EX1002, ¶¶228-229.)

F. Ground 6: Lim and Chatterjee Render Obvious Claim 17-19, 22

1. Summary of Chatterjee

Chatterjee was filed on August 23, 2011, and is prior art under §102(e). (EX1012, 1.) The Examiner never found, and Applicant never cited, Chatterjee during prosecution. Chatterjee generally relates to “a system and method for providing home automation using a mobile device.” (EX1012, Abstract; EX1002, ¶¶230-233.)

2. Claims 17-19, 22

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2) Claims 17-19, 22 are obvious in further view of Chatterjee. (EX1002, ¶234.)

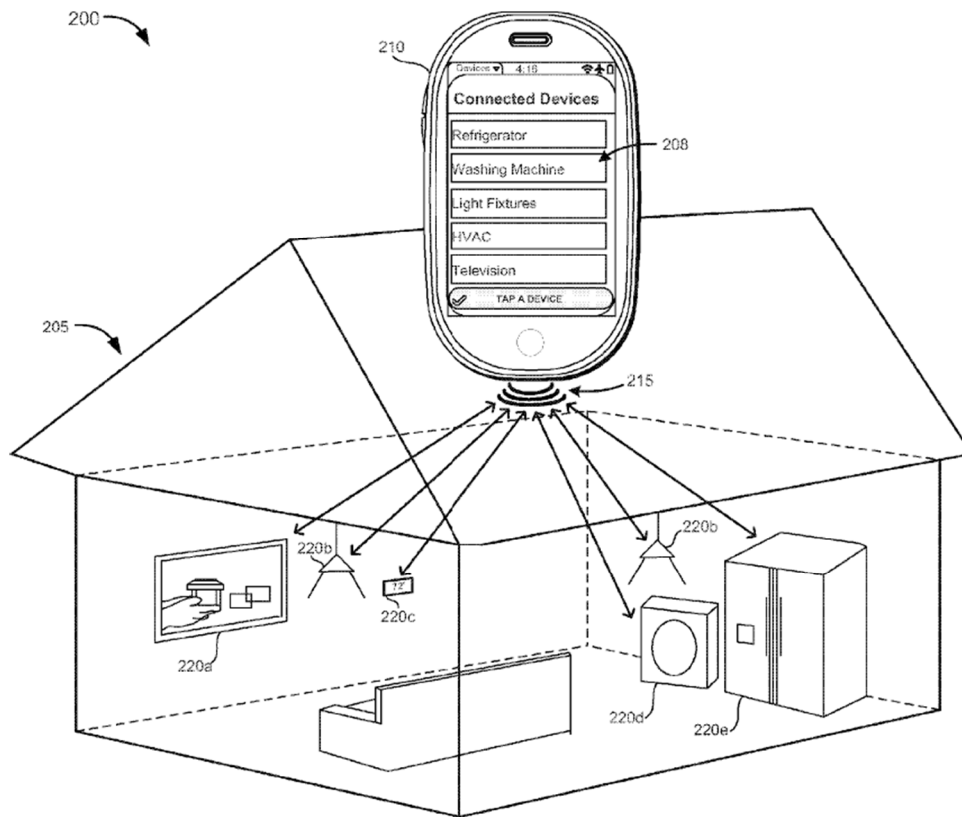
Lim discloses controlling various appliances. (EX1005, [0017].) Similarly, Chatterjee describes that a variety of appliances can be controlled via a device controller integrated into the appliance devices. “Device controller 125a-125c may be a device or module electronically coupled to the home appliance device 120α-120c or an integrated unit of the home appliance device 120α-120, and is configured to execute programming instructions *and control operations on the home appliance device such as controlling the temperature of a HVAC or toggling the light system on and off.*” (EX1012, [0012]; EX1002, ¶¶235-236.)

Chatterjee discloses that “*home automation systems* provide centralized

control for *lighting, HVAC (heating, ventilation and air conditioning), appliances (such as refrigerators and dishwashers)*, and similar systems.”

(EX1012, [0001].) “As will be used herein, home appliance devices represent any electrical or mechanical appliance, system, or machine used to perform a household function (*e.g., oven for cooking, dishwasher for cleaning, light fixtures for illumination, refrigerator for storage, and HVAC for cooling/heating the home environment*, etc.).” (EX1012, [0012]; EX1002, ¶¶236, 241-242, 244-245, 248-249.)

Further, FIG. 2 below depicts a “smart refrigerator 220*e*”, “HVAC 120*c*”, and “a smart dish washer or washing machine 220*d*.” (EX1012, [0014]; EX1002, ¶¶237, 246.)

**FIG. 2**

Lim and Chatterjee are analogous art to the '174 patent. (EX1001, Claim 1, 135:54-60; EX1005, [0017], [0007]; EX1012, Abstract.) A POSITA would have been motivated to combine Lim with Chatterjee's teachings regarding control of automated devices, therefore expanding the capabilities of Lim with a reasonable expectation of success. (EX1002, ¶238-240, 243, 247, 250.)

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

1. Summary of Rosenberger

Rosenberger was filed on October 4, 2011, and is prior art under §102(e). (EX1013, 1.) The Examiner never found, and Applicant never cited, Rosenberger during prosecution. Rosenberger relates to “improvements in speech-recognizing control devices that respond to oral commands from a system user to produce a desired effect or result.” (EX1013, 1:7-13, 8:8-16; EX1002; ¶¶251-253.)

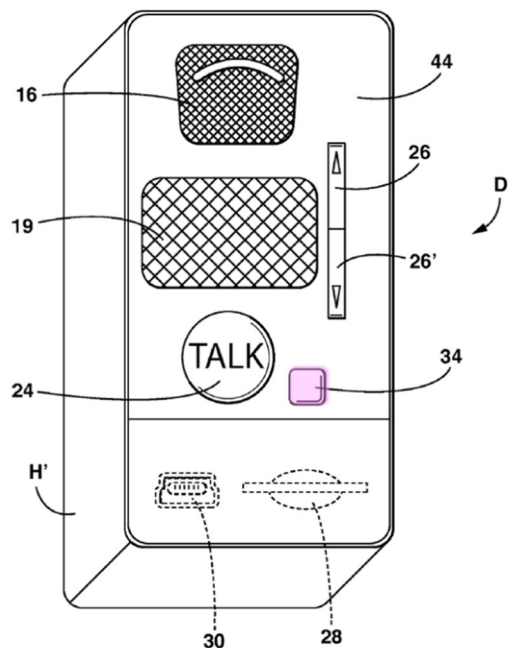


FIG. 2B

2. Claims 24/25

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2.) Claims 24-25 are obvious in further view of Rosenberger. (EX1002, ¶254.) 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., EX1013, 8:8-16.)

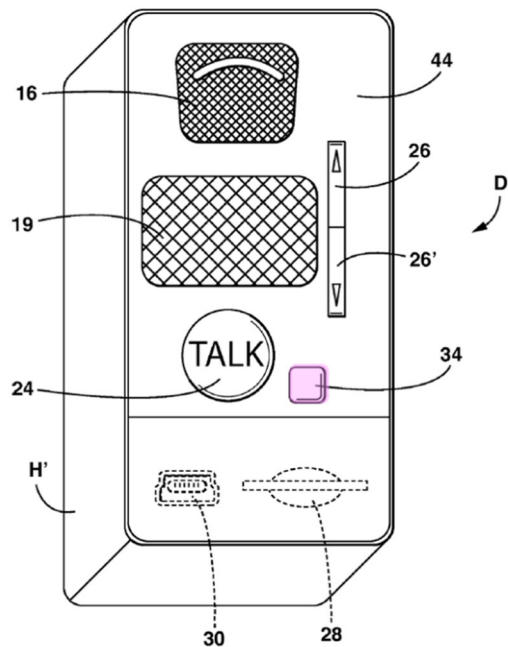


FIG. 2B

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.*” (EX1013, 8:20-32; EX1002, ¶255.)

As discussed, Lim’s *client device* includes *LEDs*. (EX1005, [0019]; *supra* §IV.A.2.g.) A POSITA would have understood that these are electric light sources

for converting an electrical energy into light that emits a visible light, consistent with the description of the '174 patent. (EX1001, 39:10-21; EX1018, 174; EX1002, ¶256.)

A POSITA would have been motivated to combine Lim with Rosenberger to include Rosenberger's status light in the audio interface device, which turns on/off and/or changes color to indicate different device states and modes to improve the usability of the device, as taught by Lim/Rosenberger, and with a reasonable expectation of success. (EX1002, ¶¶259-260.)

It would have been obvious to combine the teachings of Rosenberger with Lim. (EX1002, ¶¶257-261.) Lim and Rosenberger are analogous art to the '174 patent and both relate to processing and responding to audio commands. (EX1005, [0007]; EX1013, 1:7-13.) Lim explicitly discloses that "prompts for voice input may take the form of an *illuminated LED*". (EX1005, [0027].)

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

1. Summary of Rossing

Rossing was published by at least July 23, 2007, and is prior art under §102(a), (b). (EX1043, 1.) The Examiner never found, and Applicant never cited, Rossing during prosecution. Rossing is a textbook discussing "Audio and Electroacoustics." (EX1014, Title.) Rossing discloses how conventional microphones and speakers operate. (EX1002, ¶¶265-271.) Rossing was published

by Springer and was well-known among POSITAs. (EX1002, ¶272.) A POSITA would have looked to Rossing for detailed specifics and techniques of using microphones and speakers in the system disclosed in Lim. *Id.*

2. Claim 9

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2) Claims 9, 26-28 are obvious in further view of Rossing. (EX1002, ¶273.)

A POSITA would understand that the claim refers to a standard microphone. (EX1002, ¶274.) Lim discloses use of such a *microphone* in the *controlling device*. (*Id.*; EX1005, [0019], [0023]; EX1001, claim 9, 17:47-54.EX1002)

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.” (EX1014, 757; EX1002, ¶275.)

Rossing also discloses other variety of microphones such as “the ribbon microphone and the condenser microphone.” (EX1014, 758.) Rossing discloses piezoelectric microphones and carbon microphones. (*Id.*; EX1002, ¶¶276-277.)

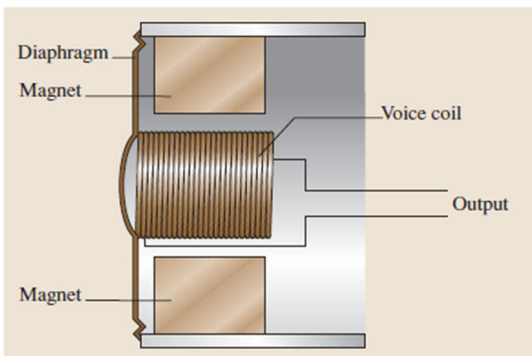


Fig. 18.10 A dynamic microphone uses a coil of wire attached to a diaphragm, suspended in a magnetic field (After [18.24])

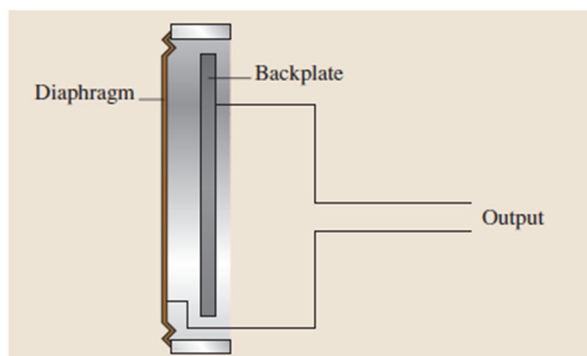


Fig. 18.9 A condenser microphone consists of a very thin diaphragm suspended parallel to a backplate (After [18.24])

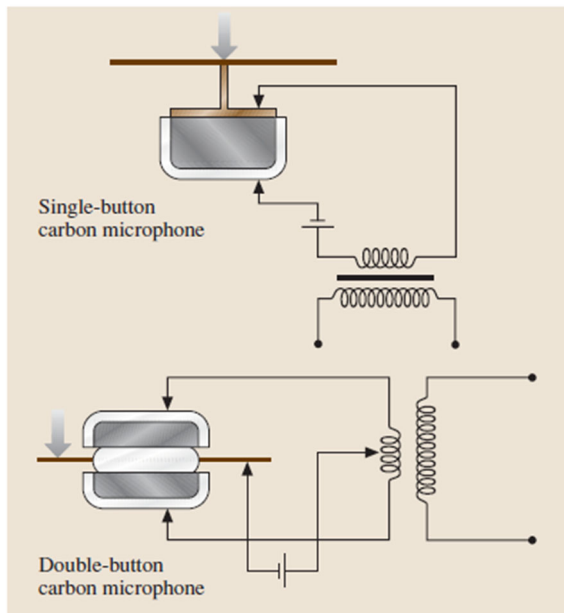


Fig. 18.11 A carbon microphone (After [18.25])

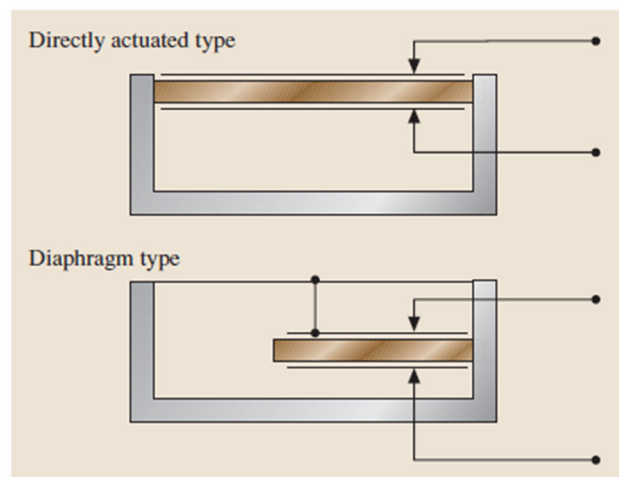


Fig. 18.12 Piezoelectric microphones (After [18.29])

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*.” (EX1014, 759.) Rossing also discloses “an ordinary omnidirectional microphone.” (EX1014, 309; EX1002, ¶¶278-279.)

Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1014, 757-758, 763, 765, 903.) It would have been obvious to a POSITA to apply Rossing’s teachings related to conventional components to Lim and with a reasonable expectation of success. (EX1002, ¶¶280-281, 285, 287, 291-292.)

3. Claims 26-28

As discussed, Lim teaches operating an actuator (*e.g.*, *speaker or speaker 303*). (*Supra* §§IV.A.2.i, IV.A.2.g; EX1001, 105:48-52, 105:61-106:4; EX1005, [0027]-[0029]; EX1002, ¶¶282, 286, 288-289.)

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.” (EX1014, 763.) Further, Rossing discloses that loudspeakers have “linear motion” (EX1014, 903) and the “dynamic or moving-coil loudspeaker” has a coil that can “move axially.” (EX1014, 18; EX1002, ¶¶283-284.)

Regarding claims 27 and 28, a POSITA would have understood that the claim refers to a speaker. (EX1002, ¶¶286, 288.) Lim teaches operating a *speaker or speaker 303* consistent with the disclosure of the ’174 patent. (*Supra* §§IV.A.2.i, IV.A.2.g; EX1005, [0037]; EX1001, 105:48-52, 105:61-106:4.) In 2012, a POSITA would have understood that a speaker is a sounder for converting

an electrical energy to sound waves. (EX1018, 122; EX1014, 763; EX1002, ¶286.)

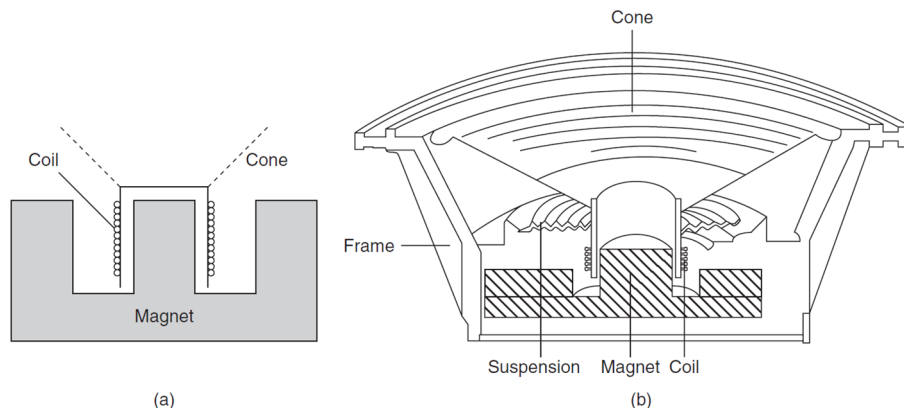


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

(EX1018, 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. (EX1014, 765; EX1002, ¶287.) 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.*” (EX1014, 763; EX1002, ¶290.) “In 1929, E. W. Kellogg filed a patent for an *electrostatic loudspeaker* . . .” *Id.* It would have been obvious to combine Lim and Rossing. (*Supra* §IV.H.2; EX1002, ¶¶285, 287, 291-292.)

I. Ground 9: Williams and Lim Render Obvious Claims 1, 5-6, 10-12, 16, 20-21, 23, 29.

1. Summary of Williams

Williams was filed on January 28, 2011, and is prior art under §102(e). (EX1006, 1.) The Examiner never found, and Applicant never cited, Williams during prosecution. 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.” (EX1006, [0046].) 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, ¶293.)

Referring to FIG. 1 below, “[a]n individual *interface device 116* may include a microphone 202 that is configured to detect ambient noise, sounds, and *speech*.” (EX1006, [0023].) Further, “within 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*.” (EX1006, [0011]; EX1002, ¶294.)

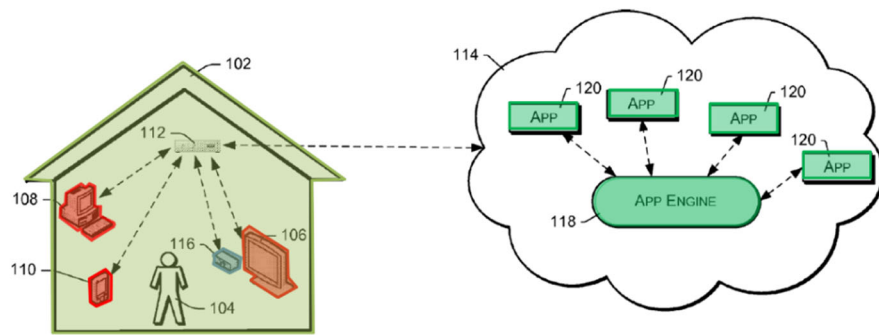


FIG. 1

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.” (EX1006, [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 discloses “allow[ing] a user to *control various on-premises devices and equipment, such as media equipment and home-automation devices*.” (EX1006, [0051]; EX1002, ¶¶295-296.)

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

At the time of the invention, it would have been obvious for a POSITA to combine Williams and Lim. (EX1002, ¶297.) Williams and Lim are analogous art to the '174 patent and all relate to processing and responding to human speech

commands. (*Id.*; EX1006, [0027], [0036]; EX1005, [0007].)

A POSITA would have been motivated to combine Williams with Lim. (EX1002, ¶¶297-301.) Williams itself discloses claim 1. To the extent that limitation 1[B] requires actual data transmission—beyond mere hardware configuration—between the client and controlled devices (which is not conceded), Williams-Lim discloses this. For example, the Williams-Lim combination may perform the recited steps as disclosed in Williams, but in 1[B] and 1[F][3], Williams’ *server 400* sends information (*e.g.*, related to services 218) for controlling controlled devices to the *client device* (rather than to the *controlled devices*), which in turn transmits commands to the *controlled devices* as taught by Lim. (EX1002, ¶¶298-299.) By adopting Lim’s teaching, the server only needs to establish and maintain connection with the *client device*, rather than maintaining numerous individual connections with the controlled devices respectively. This combination would predictably simplify network management and improve reliability and security, as well as making the system easier to scale, troubleshoot, and manage. The combination would also result in easier and more secure firewall management. *Id.*

The combination would yield predictable results of enhancing the reliability and scalability of the system by having a centralized, local hub (*i.e.*, the client device) that receives control information from the remote server and transmits

commands to the controlled device. (EX1002, ¶¶299-301.) It would have been obvious to try combining the communication method described in Lim with Williams, as there is a finite number of identified, predictable solutions for transmitting messages and control signals from a remote server to home appliances, without requiring much effort and with a reasonable expectation of success. *Id.* The known benefits of Lim would prompt its application in the same field based on design incentives and other market forces. *Id.*

3. Claim 1

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

Both Williams and Lim disclose this. For example, Williams’ system comprises a *client device* (e.g., an interface device), a *server*, and a *controlled device* (e.g., *media devices, home automation and security devices*). (EX1006, [0016], [0023], [0011], [0010], [0027], [0007].) Williams’ system “accept[s] *spoken commands* from a user regarding *on-premises devices*, and may respond by interacting with such devices through available network connections.” (*Id.*, [0051]; EX1002, ¶¶302-303.)

Each of Williams’ *controlled devices* and *client device* contains actuators, such as displays, speakers, motion actuators, etc. (*Infra* §§IV.I.3.b, IV.I.3.g.) And Williams’ system operates these actuators in response to the captured human voice data. (*Infra* §IV.I.3.i, IV.I.3.j; EX1002, ¶304.)

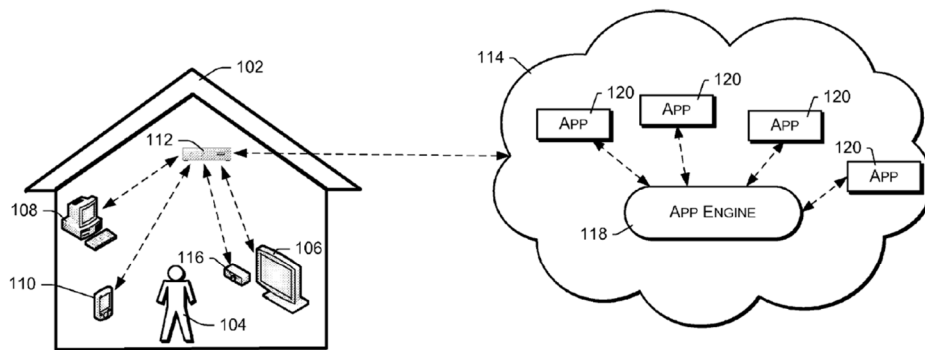


FIG. 1

b. 1[PR][2]: “for use with a wireless network in a building and with a controlled device that comprises a second actuator in the building, the system comprising:”

Both Williams and Lim disclose this. For example, Williams’ system is for use with a *wireless network* in a *building*. Referring to FIG. 1 below, “the *in-home* devices connect to *a router or access point 112*, which is in turn connected to a public network 114. Connection *between in-home devices and the router 112 may be through wired or wireless technologies.*” (EX1006, [0011]-[0013].) As shown, *interface device 116* and *media devices 106/108/110* communicate over a wireless network through wireless router/access point 112, thus forming a wireless network within *home 102*. (EX1002, ¶¶305-306; EX1006, [0020], [0023].) Williams’ wireless network reflects a common configuration for a wireless network in a *home* in 2012 and is also consistent with the ’174 patent. (EX1007, 11, 100; EX1001, 28:7-10; EX1002, ¶¶307-308.)

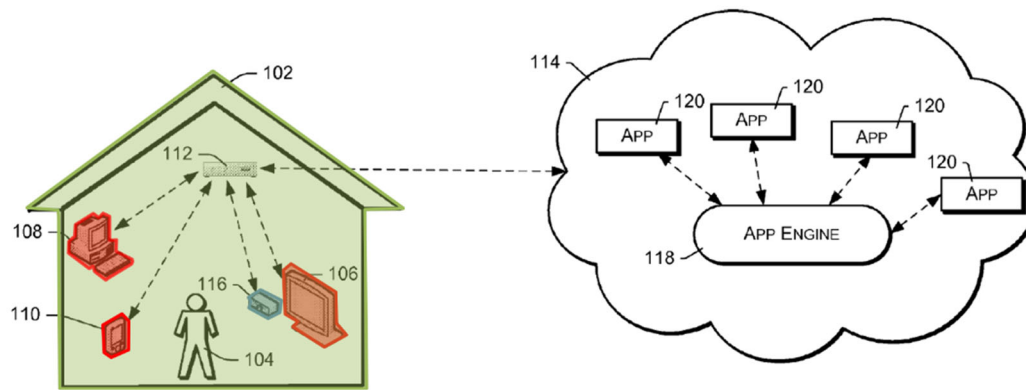


FIG. 1

Second, Williams' system has a *controlled device* that comprises a second actuator in the same *building*. With reference to FIG. 1 below, "[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. (EX1006, [0011], [0010].) Williams discloses that the *controlled device* is also in the *building* (e.g., *home 102*). (EX1006, [0011], [0016], [0023], [0036], FIGs. 1 and 2; EX1002, ¶¶309-311, 317-318.)

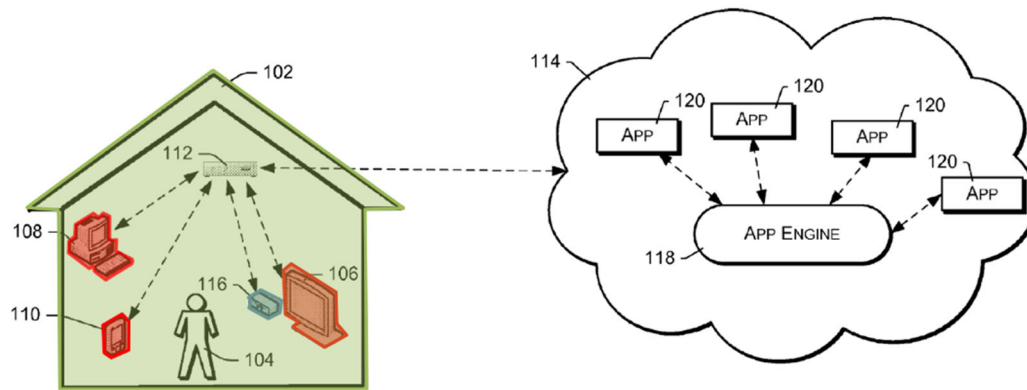


FIG. 1

The *devices 106, 108, and 110* are controlled devices because they can be controlled. (EX1002, ¶312.) Williams discloses “an application might be designed to allow a user to **control** various *on-premises devices and equipment*, such as *media equipment* and *home-automation devices*.” (EX1006, [0051], [0036], [0049]; EX1002, ¶312.)

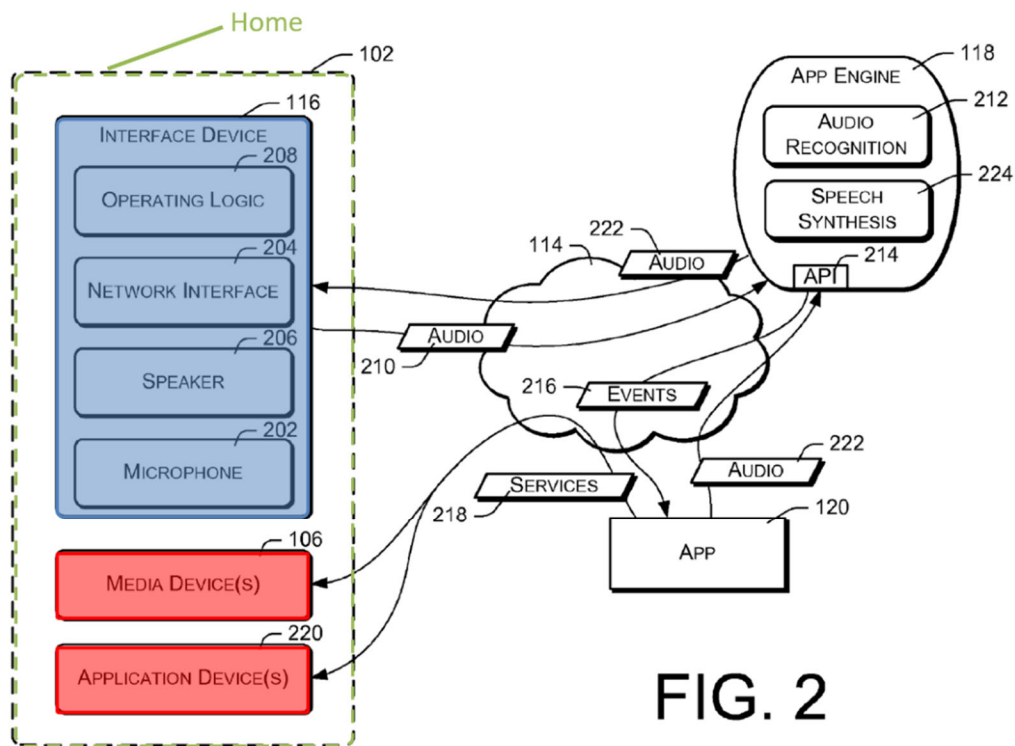


FIG. 2

It would have been obvious that Williams' controlled devices (*e.g.*, “**televisions, audio players, video players**”) contain displays and/or speakers. (EX1006, [0011], [0009]-[0010], [0021], [0036], [0054]; EX1004, 396; EX1025, FIG. 15; EX1002, ¶¶313-314.) Displays and speakers are actuators. (*Id.*; *Supra* §IV.I.3.b; EX1001, 40:51-41-7, 25:33-38, 109:41-46, 110:60-66; 144:64-145:1.)

Further, Williams discloses that “the applications 120 may communicate and *operate* in conjunction with . . . **home automation and security devices.**” (EX1006, [0021], [0036], [0051].) A POSITA would have understood that home automation and security devices (*e.g.*, automated doors or windows, automated

lights, automated locks, automated cameras, automated HVAC systems) have motion actuators. (E.g., EX1020, 401; EX1004, 396; EX1002, ¶¶315-316.)

c. 1[A][1]: “an Internet-connected server device external to the building configured for processing first and second human voice data,”

Both Williams and Lim disclose this limitation. First, Williams discloses an *Internet-connected server device* external to the *building*. Williams’ “network 114 may thus include . . . *the Internet*. . . .” (EX1006, [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.” (EX1006, [0052], [0014], [0021].) “[T]he *server 400* may also comprise a *web service component 410*.” (*Id.*, [0055]; EX1002, ¶¶319-320.)

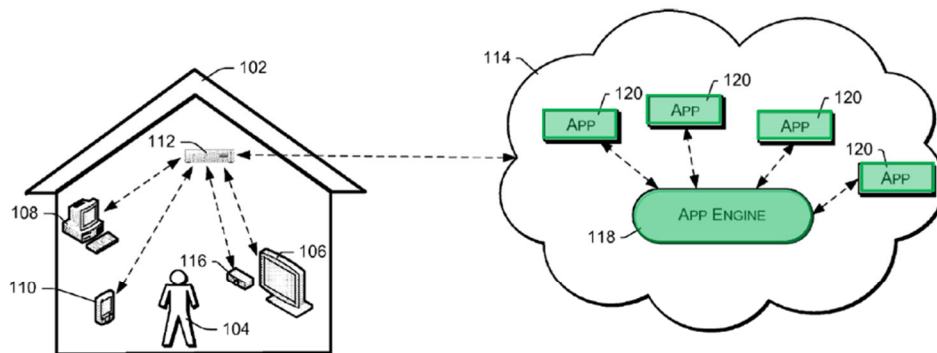


FIG. 1

As shown in FIG. 1 above, the *server 400* is in the cloud and is connected to the in-home WLAN via router 112. (FIG. 1, EX1002, ¶322.) This is consistent

with the '174 patent. (EX1001, 32:25-42.)

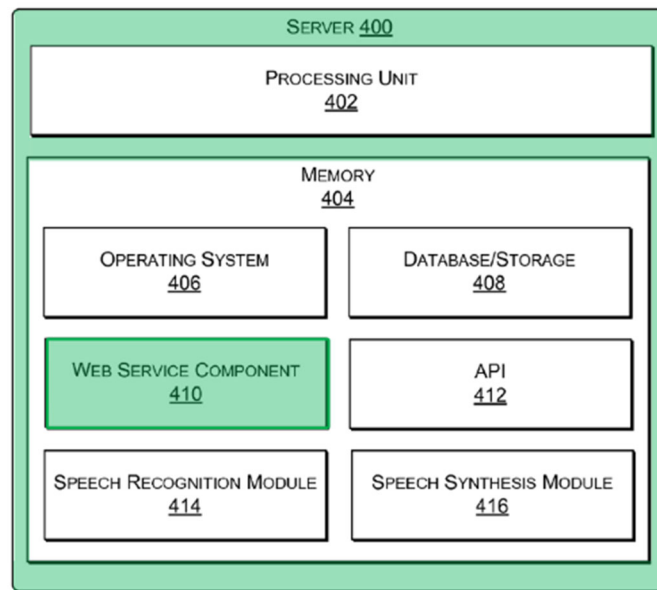


FIG. 4

Williams' *server 400* is external to the *building* (e.g., the user premise 102) as shown in FIG. 1 above. Indeed, Williams states that “[s]ervers may be located together or separately, and organized as virtual servers, *server banks*, and/or *server farms*”, which a POSITA would have understood to be outside residential homes. (EX1006, [0052]; EX1015, 277; EX1016, 1760; EX1002, ¶¶323-325.)

Second, Williams discloses that the *Internet-connected server device* is configured for *processing first human voice data* and *second human voice data*. Williams' system “provide[s] audio answers to *spoken questions*”, such as “*What's the recipe for deviled eggs?*”, an *inquiry “about a flight or other event”*,

etc. (EX1006, [0048], [0050].) Referring to FIG. 2 below, Williams depicts exemplary human voice data such as “*audio-based information 210*” captured by microphone 202. (*Id.*, [0022], [0024]-[0025]; *infra* §IV.I.3.f; EX1002, ¶¶326-329.)

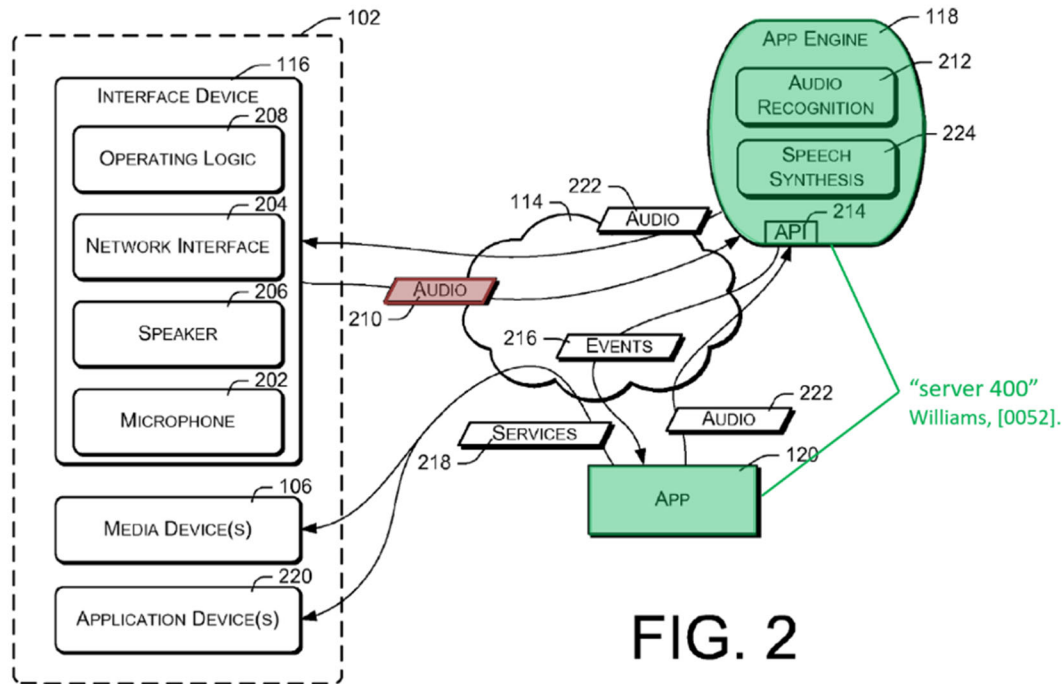


FIG. 2

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.” (EX1006, [0022], [0052].) Williams discloses several examples of processing, such as “speech-to-text conversion”. (EX1006, [0034]-[0037], [0039]-[0040], [0053], [0056]; EX1002, ¶¶330-333.)

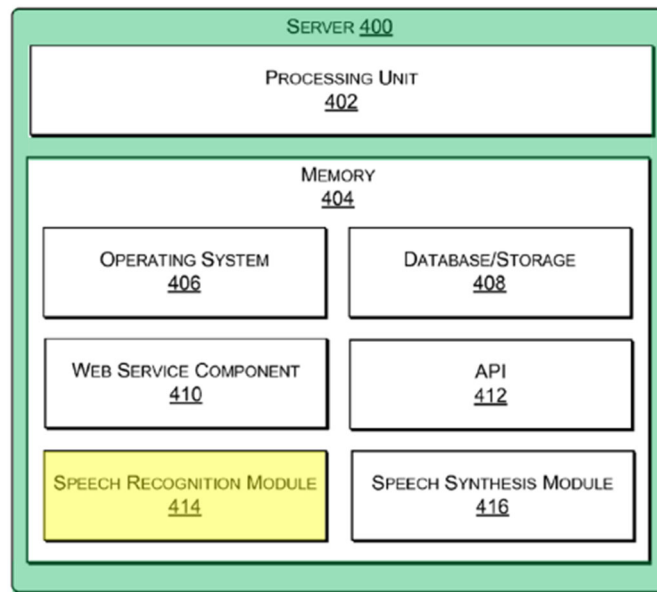


FIG. 4

To the extent that the recited “*second human voice data*” is different from the recited “*first human voice data*”, Williams is clear that its system can receive and process *multiple* human voice commands. (EX1006, [0044], [0036], [0021]; EX1002, ¶¶334, 338.)

Further, per the claim language, any voice command captured by the microphone that causes an actuator in the *controlled devices* to operate in the recited manner would disclose the “*second human voice data*.” (*Infra* §IV.I.3.j; IV.I.3.h; EX1002, ¶335.) Williams “allow[s] a user to control various *on-premises devices and equipment, such as media equipment and home-automation devices*.” (EX1006, [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.” (EX1006, [0051].) Such spoken commands are “*second human voice data*.” (EX1002, ¶336.)

A POSITA would have understood that the “*second human voice data*” (to the extent it is different from the first human voice data) is processed in the same manner as shown in FIG. 2 (e.g., sent to the *server 400* for speech processing). (EX1002, ¶337.) Alternatively, the “audio 210” shown can include or be both the first and second human voice data because “audio 210” results in both first and second messages. *Id.*

d. 1[A][2]: “and to produce first and second messages respectively in response to the processing;”

Both Williams and Lim disclose this limitation. As discussed, Williams’ server processes the first and second human voice data. (*Supra* §IV.I.3.c.) Williams discloses that the *first message* and *second message* are produced by the *server* in response to the processing. 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.” (EX1006, [0022], [0037], [0048], [0045].) Williams discloses several examples of “*audio answers* to *spoken questions*” produced by the *cloud server* in response to processing *first human voice data*. (EX1006, [0045], [0048], [0050];

EX1002, ¶¶339-342.) For example, “[t]he application may respond to a *spoken command to order movie tickets* by using the speech synthesis functionality of the system to *query the user regarding movie selections and times*.” (*Id.*; EX1006, [0045].)

The *server* also produces a *second message* in response to the processing. The “*second human voice data*” is “processed by the application engine 118 [of the *server*] to derive event descriptors” and “[t]he application 120 [of the *server*] may use the information from the event notifications 216 to initiate or customize *services 218*.” (EX1006, [0034], [0035], [0036].) “Such *services* may be provided through any of the *media devices 106* located in the home, as well as through other *application devices 220, including computers 108, personal devices 110, and other devices available to users*.” (EX1006, [0036], [0021], [0038], [0040], [0042], [0049], [0050], FIG. 3.) A POSITA would have understood, and it was common knowledge, that *information related to the service 218* (e.g., a control message) is sent to a controlled device to operate it wirelessly. (EX1002, ¶¶343-347; EX1019, 3:19-22, 6:23-27; EX1008, claim 1; EX1012, [0012].)

Because the messages are produced as a result of speech processing of human voice data, it would have been obvious to a POSITA that Williams’ *Internet-connected server* (e.g., server 400) produces the *first* and *second message* in response to the *processing*. ((EX1002, ¶¶342, 348.)

e. 1[B]: “a client device in the building configured to communicate with the controlled device and with the server device over the wireless network;”

Both Williams and Lim disclose this limitation. For example, Williams discloses a *client device* in the *building*. For example, “*the user’s home 102* may be equipped with one or more on-premises *audio monitoring devices 116* . . . having a microphone and a network interface”. (EX1006, [0016].) The *interface devices 116* communicates with cloud-based application service or engine 118 [of *server 400*] and thus is a client device. (*Supra* §IV.I.3.c; EX1006, [0020], [0023]; EX1002, ¶349-351.)

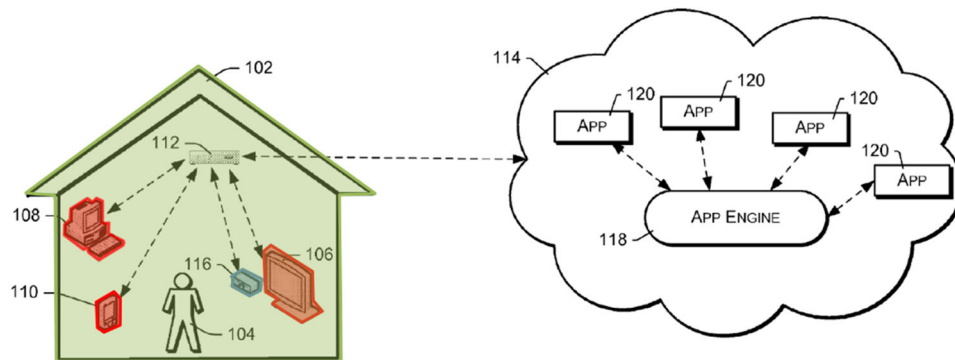


FIG. 1

FIG. 2 below “is a block diagram illustrating additional details regarding the system of FIG. 1.” (EX1006, [0004].) As shown, the *interface device 116* communicates with the *server*. (EX1002, ¶¶351-352.)

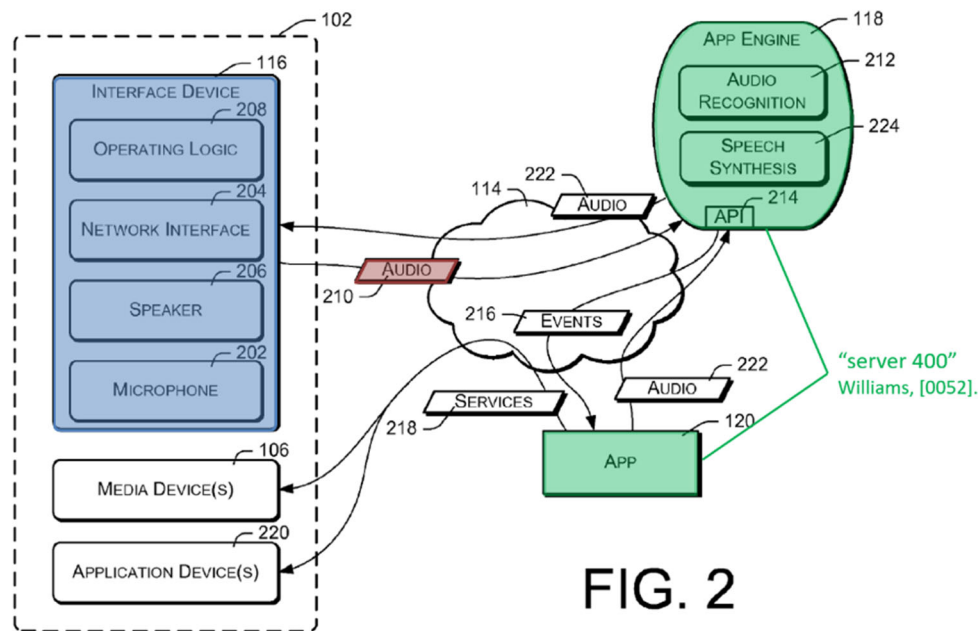


FIG. 2

Williams' *client device* is configured to communicate with the *Internet-connected server device* over the *wireless network*, which is formed by the wireless router/access point 112 within *home 102* as discussed above. (EX1002, ¶352; EX1006, [0020], [0023].) As shown in FIG. 1, Williams' *interface device 116* "communicate[s] with and transmit environmental information to a cloud-based application service or engine 118 [run on *server 400*], *via the router 112* and the network 114." (EX1006, [0020], [0018].)

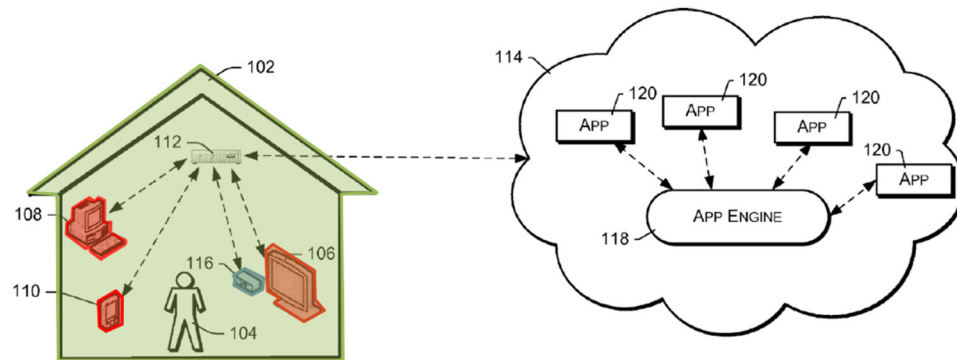


FIG. 1

Further, as discussed and shown in FIG. 1 above, Williams’ *client device* and *controlled device* (i.e., in-home devices) are connected to the same *wireless network* through the same *wireless router/access point 112*. (EX1006, [0012], FIG. 1.) By connecting to the same wireless network, Williams’ *client device* and *controlled device* are configured to communicate with each other over it. (EX1007, 11, 100; EX1002, ¶353.)

To the extent that the claim requires actual data transmission between *client device* and the *controlled device*, Williams-Lim discloses this. The Williams-Lim system can transmit control commands from the *client device* to the *controlled device* over the *wireless network*. (*Supra* §IV.I.2; EX1005, [0027]; EX1002, ¶354.)

f. 1[C]: “a microphone in the client device for capturing the first and second human voice data;”

Both Williams and Lim disclose this limitation. For example, referring to FIG. 2, Williams disclose that “[a]n individual *interface device 116* may include a *microphone 202* that is configured to detect ambient noise, sounds, and *speech*.” (EX1006, [0023], [0016], [0019] Claim 1-3 and 21; EX1002, ¶¶355-357.)

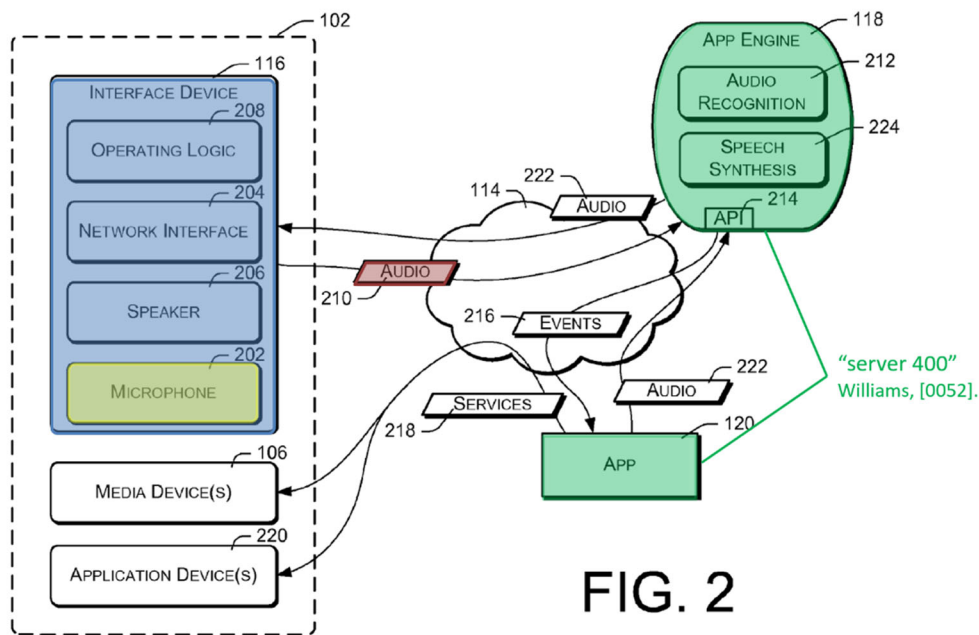


FIG. 2

It would have been obvious that the *first human voice data* and *second human voice data* are captured by the microphone, for example, because the “*interface device 116* is configured to send *audio-based information 210* to the application engine 118 [run on server 400], based on *ambient audio* detected by *microphone 202* of the *interface device 116*”. (EX1006, [0024]; EX1002, ¶358.) So are subsequent voice inputs by user such as “*spoken commands* from a user regarding on-premises devices.” (EX1006, [0051].)

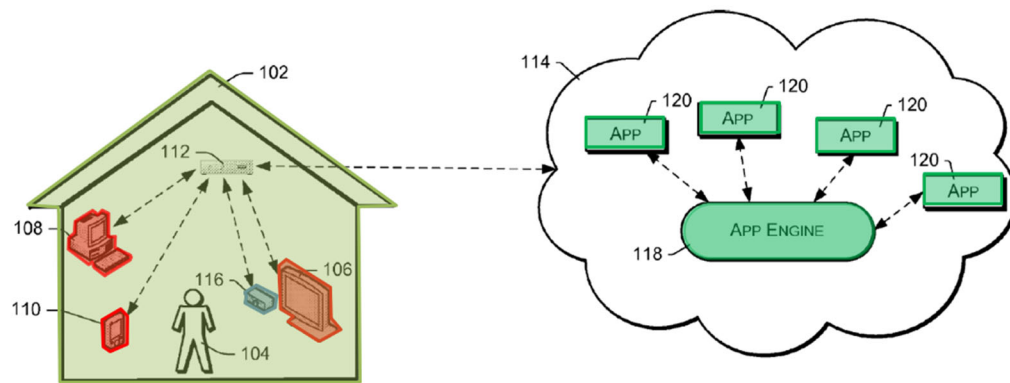


FIG. 1

g. 1[D]: “and a first actuator in the client device,”

Both Williams and Lim disclose this limitation. For example, Williams discloses that “the *interface device* may also have a *speaker* or *other form of audio output*”. (EX1006, [0016], [0023], FIG. 2; EX1002, ¶¶359-360.) The speaker is an actuator. (EX1002, ¶361; e.g., EX1018, 122, EX1014, 763.)

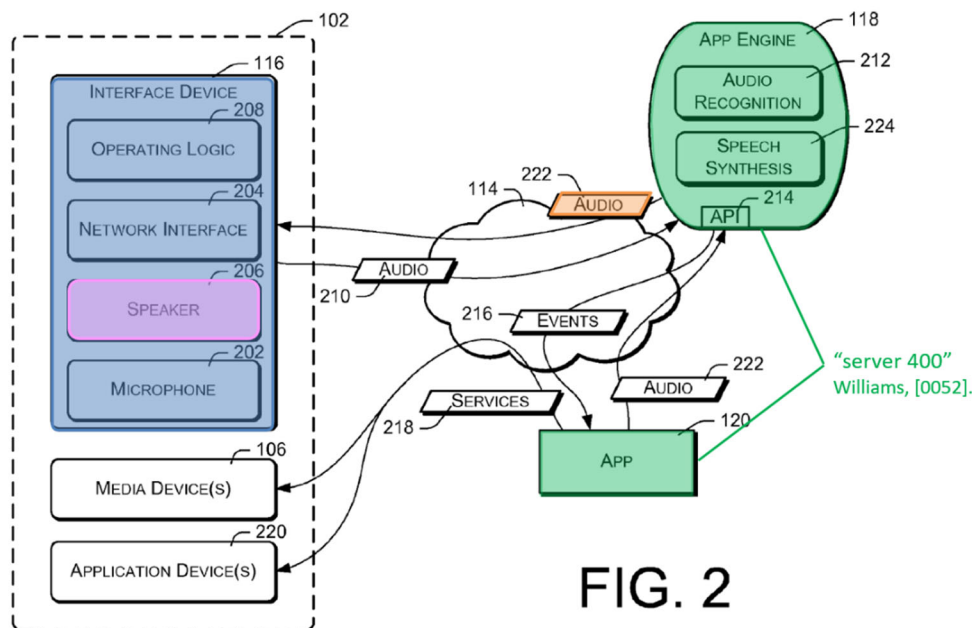
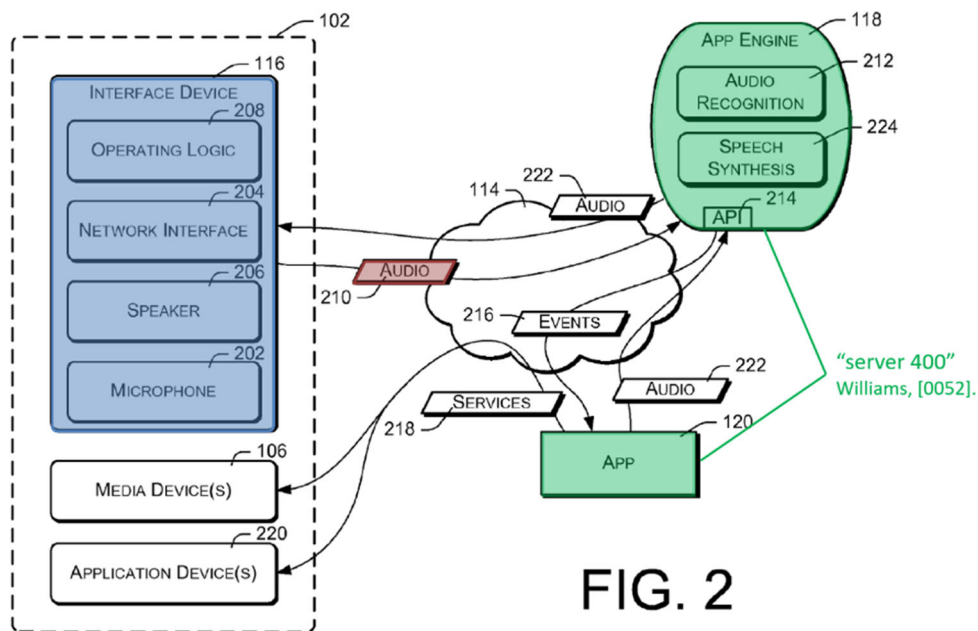


FIG. 2

h. 1[E]–1[F][1]: “wherein the system is operative for sending to the server by the client device via the wireless network, the captured first and second human voice data,” “and wherein the system is further operative for receiving from the server the first and second messages in response to the sending,”

Both Williams and Lim disclose this limitation. For example, Williams’ system discloses that the *client device* sends the captured human voice data to the *server* over the *wireless network*. FIG. 2 below shows how an exemplary voice command (either first or second human voice data) is processed. As shown, *interface device 116* sends *audio 210* (captured by microphone 202) to *server 400*. (EX1006, [0023].) “[T]he application engine 118 [of *server 400*] receives *audio-based information* from the *interface devices 116*.” (EX1006, [0022], [0020], [0024]–[0028], [0033]–[0034], [0038]–[0039]; EX1002, ¶¶362–366.) It would have been obvious to a POSITA that the *second human voice data* (e.g., “spoken

commands from a user regarding on-premises devices,” an “audio-based information”) is similarly sent to the *server* for processing. (EX1006, [0028], [0007], [0009], [0051]; EX1002, ¶¶367, 370.) Alternatively, the “audio 210” sent to and processed by *server 400* can include or be both the first and second human voice data, which result in first and second messages. *Id.*



The data is sent via the wireless network. Referring to FIG. 1 below, it would have been obvious that the human voice data is sent by *interface device 116* to the router 112 over the *in-home wireless network* and then to the *server* over 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]”. (EX1006, [0001], [0007], [0011]-[0015], [0020], [0038], [0051]; *supra* §IV.I.3.b; EX1002, ¶¶368-369.)

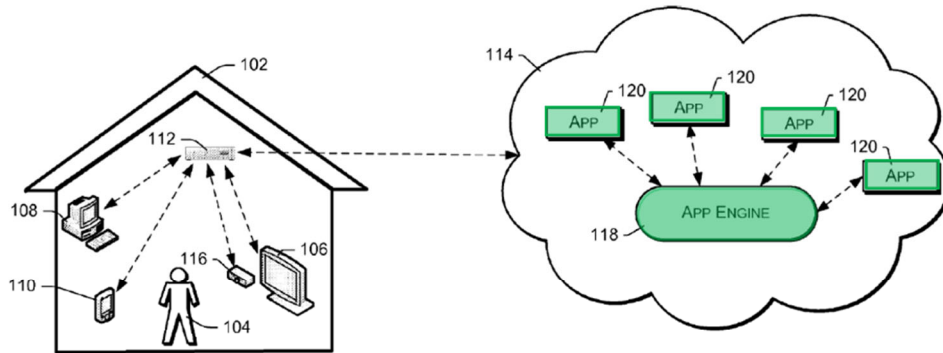


FIG. 1

William discloses, with reference to FIG. 2 below, that the *first* and *second message* generated by the *server* are sent to and received by the system (*client device* and *controlled device* respectively). (EX1002, ¶¶371-372.)

Referring to FIG. 2 below, Williams’ *interface device 116* sends “*audio-based information 210* to the application engine 118 [of the *cloud server*]” and 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.” (EX1006, [0022], [0024], [0037], [0048], [0045].) Williams discloses several examples of *interface device 116* providing an *audio*

response generated by the *cloud server*. (EX1006, [0045], [0048], [0050]; EX1002, ¶¶373, 375.)

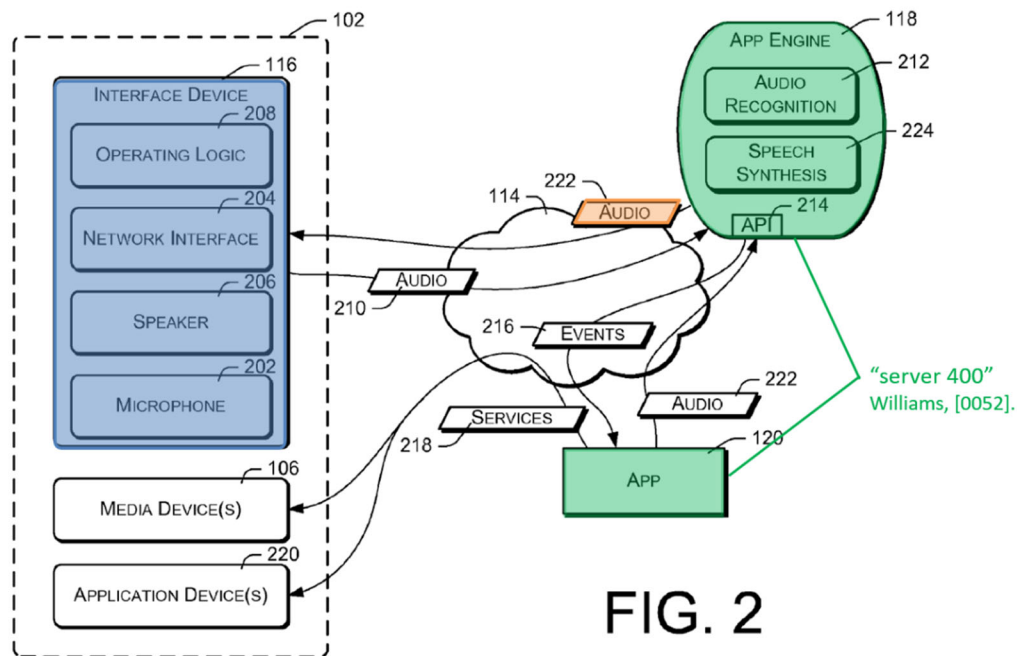


FIG. 2

It would have been obvious that the *first message* is received “in response to the sending of the captured *first . . . human voice data*” because the *first message* is generated as the result of speech processing of the *first human voice data*. (EX1002, ¶374.)

As shown in FIGs. 1 and 2, the *audio 222* is sent by the *cloud server* over the network 114 (i.e., “the Internet”) and received by the *interface device 116* over the in-home wireless network (via router 112). (*Supra* §§IV.I.3.b, IV.I.3.c; EX1006, FIG. 1, [0015], [0037], [0052]; EX1002, ¶376.)

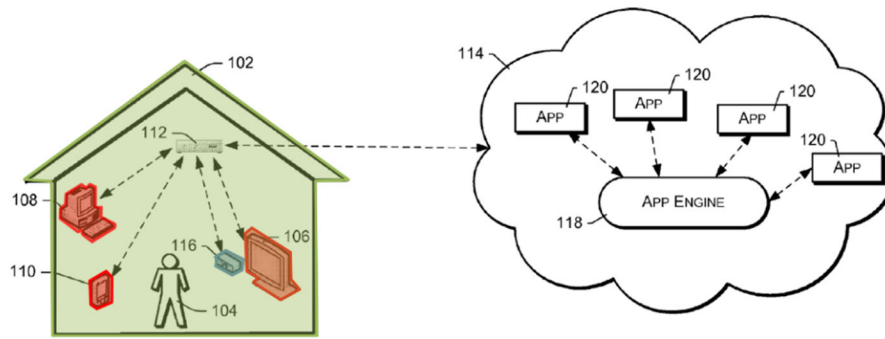


FIG. 1

Williams’ *controlled device* receives the *second message* from the *server* in response to the *interface device*’s sending of the captured *second human voice data*. As discussed, the processing of audio 210 can be equally mapped to the processing of the *second human voice data*. (EX1002, ¶377.) Williams discloses that the *server* “processes” it and as a result “may . . . initiate or customize *services 218*” which “may be provided through any of the *media devices 106* located in the home, as well as through other *application devices 220*.” (EX1006, [0027], [0036]; [0021], [0038], [0040], [0042], [0049], [0050], FIG. 3; EX1002, ¶377.)

As shown in FIG. 2 below, William’s *media devices 106 and application devices 220* receives *services 218* from *server 400* over the wireless network 114 (e.g., the Internet) and over the in-home wireless network (via router 112 in FIG. 1). (EX1006, [0012], [0013], [0052]-[0056]; EX1002, ¶¶378-379.) 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, ¶380.) A POSITA would have understood that, to operate a controlled device, *information related to the services 218* (e.g., a control message) is sent to and received by the *controlled device*, consistent with common implementations. (E.g., EX1019, 3:19-22; EX1008, Claim 1; EX1012, [0012]; EX1002, ¶380.)

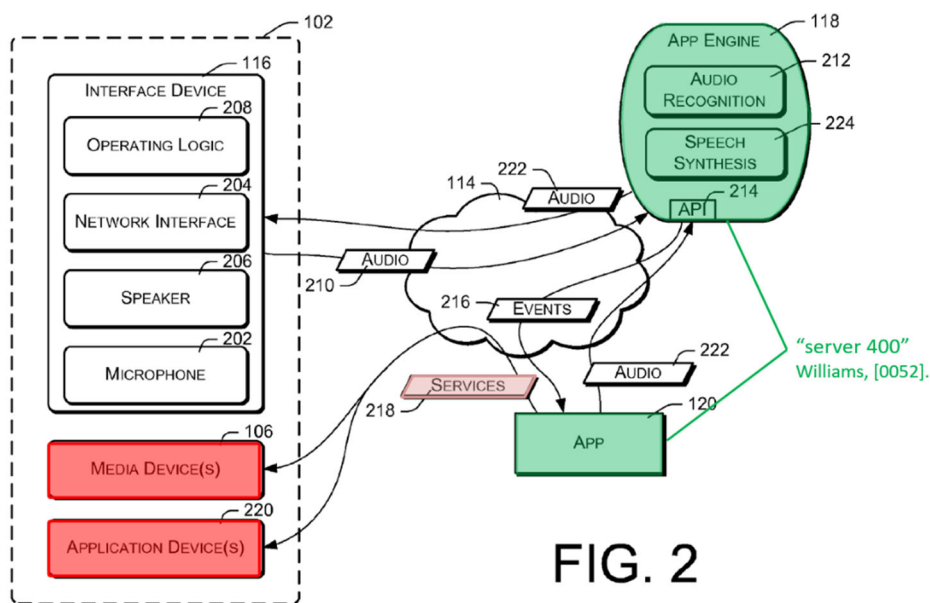


FIG. 2

Further, the Williams-Lim combination discloses this limitation. Though, in the Williams-Lim system, the *client device*, rather than *controlled device*, receives the *second message* from the *server* and uses it to control the *controlled device*. (*Supra* §IV.I.2; EX1002, ¶381.)

i. 1[F][2]: “to operate the first actuator in the client device in response to the received first message,”

Both Williams and Lim disclose this limitation. For example, Williams

discloses that the cloud server “provide[s] **audio 222** for eventual transmission to the **interface device 116**, to be played on the **speaker 206**.” (EX1006, [0037], [0045], [0048]; EX1002, ¶¶382-387.)

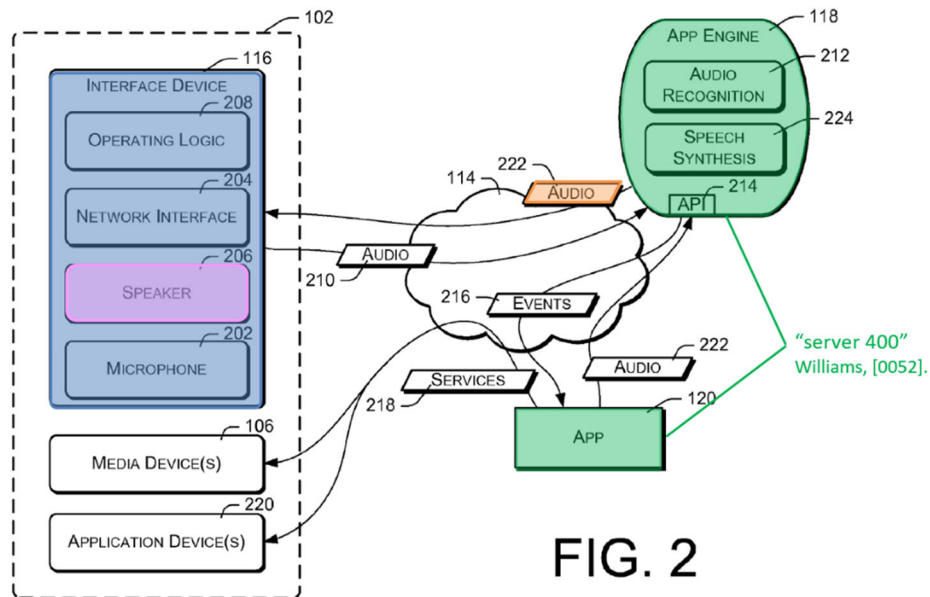
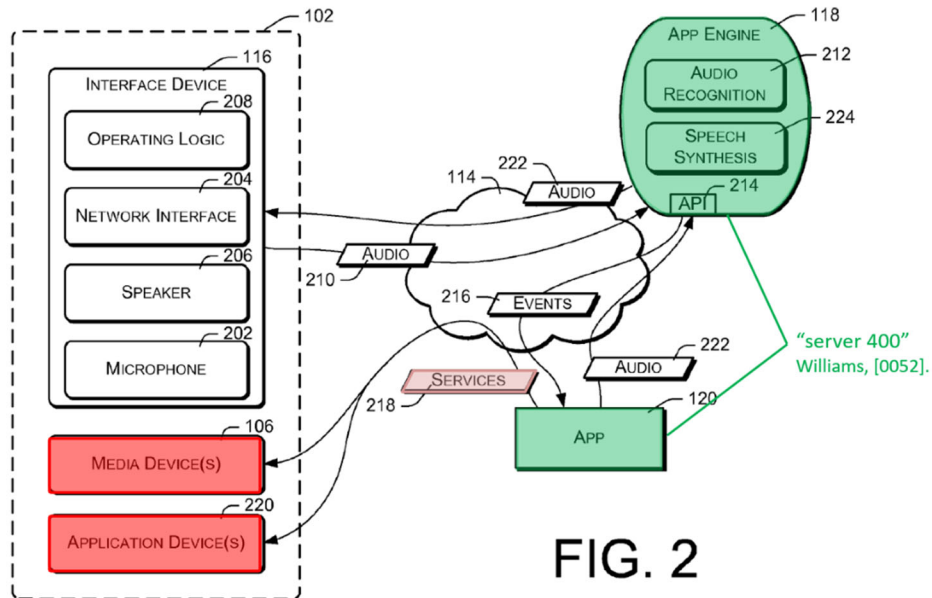


FIG. 2

j. 1[F][3]: “and to operate the second actuator in the controlled device in response to the received second message.”

Both Williams and Lim disclose this limitation. For example, Williams discloses the **media devices 106 and/or application devices 220** receive the **services 218** or information related to **services 218**. (*Supra* §IV.I.3.h; EX1006, [0042], FIG. 2.) For example, “the application 120 may use the information from the event notifications 216 to initiate or customize **services 218** on behalf of users associated with the home 102. Such services may be provided through **media devices 106** . . . as well as through other **application devices 220**.” (EX1006, [0036], [0042].) It would have been obvious that the second actuator (i.e., any of

the actuators in the controlled devices) is operated in response to the *information* (second message) received from *server 400*. (EX1002, ¶¶388-392.)



As discussed, Williams’s controlled devices have actuators such as displays, speakers, and motion actuators. (*Supra* §IV.I.3.b.) A POSITA would have understood that “*control[ling]* various *on-premises devices and equipment*, such as *media equipment* and *home-automation devices*” involves operating various “actuators” in media equipment and home-automation devices, for example, to playback a media item, or change the settings or turn on/off the home automation devices. (EX1002, ¶393.)

Further, the Williams-Lim combination discloses this limitation. (EX1002, ¶394.) In the Williams-Lim system, the *client device* receives the *second message* from the *server* and transmits commands to the *controlled device* to operate the

controlled device's actuators ("second actuators"). (*Supra* §IV.I.2.)

4. Claims 5/6

Both Williams and Lim disclose this. (*Supra* §IV.A.3.) For example, 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.I.3.b, IV.I.3.g, IV.I.3.i, IV.I.3.j; EX1002, ¶¶395-400.)

5. Claims 10/11/12

As discussed, Williams discloses its interface device 116 and home devices communicates with a wireless network (*e.g.*, via a wireless router which connects to the Internet). (*Supra* §§IV.I.3.b, IV.I.3.e; EX1006, [0012]–[0013], [0020], Figure 1; EX1002, ¶¶401, 406.) 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. (EX1007, 99, 166, 108, 210, 175; EX1002, ¶¶402-403, 407.) 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. *Id.* 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 globally unique MAC addresses that are of MAC-48, EUI-48 or EUI-64 type. (EX1029, 127; EX1007, 277, EX1027, 2:50-59;

EX1023, 11, 13; EX1002, ¶¶408-409.)

Williams discloses that the *interface device* may include “one or more *programmable processing units, associated memory*[.]” (EX1006, [0023].) As discussed, it was common knowledge that a networked device locally stores the associated IP addresses. (*Supra* §IV.A.4; EX1007, 211, Figure 10-44, 265.) A POSITA would have understood that the “associated memory” in Williams’ client device would locally store the IP addresses. (EX1002, ¶¶404-405; *e.g.*, EX1007, 211, Figure 10-44, 265.)

Lim discloses claims 10-11. (*Supra* §IV.A.4.) Further, Lim discloses its client device is Bluetooth enabled and connects to the Internet via a wireless network. (EX1005, [0027].) It was common knowledge that Bluetooth is a wireless personal area network (WPAN) and based on the IEEE 802.15 standard. (EX1042, 601-602.) It would have been obvious to a POSITA that the Williams-Lim combination discloses as least claim 12. (*Supra* §IV.I.2; EX1002, ¶410.)

6. Claims 16/20/21/23

The controlled device in Williams includes devices such as “*televisions, audio players, video players*” and “*home-automation devices*”. (*Supra* §IV.I.3.b; EX1006, [0051], [0011]; EX1002, ¶411.) These are “appliances”. *Id.* These also disclose a “home cinema system” recited in claim 23. (EX1002, ¶414.) 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. (EX1014, 746; EX1002, ¶414.)

Lim discloses controlling appliances such as “*fans, HVAC systems, thermostats*”. (*Supra* §IV.A.5; EX1005, [0017].) It would have been obvious to a POSITA that the Williams-Lim combination discloses at least claims 20/21. (*Supra* §IV.I.2; EX1002, ¶¶412-413.)

7. Claim 29

Both Williams and Lim disclose this limitation. (*Supra* §IV.A.6.) For example, Williams teaches “playing audio 222 on the *speaker 206*.” (*Supra* §§IV.I.3.i, IV.I.3.g; EX1006, [0037]; EX1002, ¶415.) Further, “the audio may be in the form of text, in which *speech synthesis or text-to-speech conversion* may be performed by the application engine 118,” and “an application may respond ... by using the *speech synthesis functionality* of the system to *query the user regarding movie selections and time*” (*Id.*, EX1006, [0037], [0045]-[0046] and [0050].)

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

1. Claim 2

Claim 1 is obvious in view of Williams-Lim. (*Supra* §IV.I.3.) Claim 2 is obvious in further view of Wittenberg. (EX1002, ¶416.)

Wittenberg discloses claim 2. (*Supra* §IV.B.2; EX1002, ¶417.)

It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Wittenberg. (EX1002, ¶418-423.) For example, Williams-Lim and Wittenberg are analogous art to the '174 patent and relate to remotely processing audio information such as speech for controlling devices or providing services through devices. (EX1006, [0036]; EX1005, [0007]; EX1008, Claim 1; EX1002, ¶418.)

The Williams-Lim-Wittenberg combination would result in a system in which Wittenberg's voice recognition capabilities are used to analyze human audio commands in Williams-Lim, 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, ¶422.)

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

1. Claims 3-4

Claim 1 is obvious in view of Williams-Lim. (*Supra* §IV.I.3.) Claims 3/4 are obvious in further view of Gruber. (EX1002, ¶424.)

Gruber discloses claims 3-4. (*Supra* §§IV.C.2, IV.C.3; EX1002, ¶¶425, 431.)

It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Gruber. (EX1002, ¶426-431, 433.) For example, Williams, Lim, and Gruber are analogous art to the '174 patent and all relate to processing

and responding to human speech commands. (EX1006, [0027], [0036]; EX1005, [0007]; EX1009, [0007]; EX1002, ¶426.)

The Williams-Lim-Gruber combination would capture sensor data (*e.g.*, temperature) using one or more sensors (*e.g.*, as taught by Lim/Gruber) in the Williams-Lim's client device and transmit such data to Williams-Lim's server to supplement audio processing as taught by Williams-Lim and/or Gruber. (EX1006, [0013]; EX1009, [0013]; EX1002, ¶429.) Combining the teachings of Williams-Lim and Gruber as such would, for example, yield predictable results of improving audio processing. (EX1002, ¶430.)

L. Ground 12: Williams, Lim, and Benesty Render Obvious Claims 7-8

1. Claims 7/8

Claim 1 is obvious in view of Williams-Lim. (*Supra* §IV.I.3.) Claims 7/8 are obvious in further view of Benesty. (EX1002, ¶434.)

Benesty discloses claims 7-8. (*Supra* §IV.D.2; EX1002, ¶¶435, 439.)

It would have been obvious to a POSITA to combine Benesty with the Williams-Lim system. (EX1002, ¶436-438, 440.) For example, Williams, Lim, and Benesty are analogous art and all relate to providing hands-free audio communication between users and devices. (EX1006, [0003], Figure 1; EX1005, [0007]; EX1010, 1021.) The Williams-Lim-Benesty combination would result in a system where multiple microphones as taught by Benesty are incorporated in

Williams-Lim's client device to capture human voice data, yielding the predictable results of effectively improving audio capturing. (EX1002, ¶437.)

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

1. Claims 13-15

Claim 1 is obvious in view of Williams-Lim. (*Supra* §IV.I.3.) Claims 13-15 are obvious in further view of Goldsmith. (EX1002, ¶441.)

Goldsmith discloses claim 13-15. (*Supra* §§IV.E.2, IV.E.3, IV.E.4; EX1002, ¶¶441, 446, 448.)

It would have been obvious for a POSITA to look to Goldsmith for conventional wireless networking techniques and protocols that can be applied to Williams-Lim. (EX1002, ¶¶442-445, 447, 450-451.) For example, Williams expressly discloses usage of “cellular or other wireless data communications technologies”, and Lim expressly discloses usage of “local network connection” and “networked protocol 120” such as WiFi which implements 802.11 WLAN. (EX1002, ¶442; EX1006, Figure 1, Figure 2, [0012], [0013], [0020], [0038]; EX1005, [0027], Figure 5b; EX1042, 599-600; EX1013, 7:48-51.) Further, it would have been obvious to try combining the wireless communication techniques of Goldsmith (*e.g.*, 802.11b WLAN and cellular 3G network) with the teachings of Williams-Lim without requiring much effort and with a reasonable expectation of success. (EX1002, ¶¶445, 451.)

N. Ground 14: Williams, Lim, and Chatterjee Render Obvious Claim 17-19, 22

1. Claims 17-19, 22

Chatterjee discloses claims 17-19, 22. (*Supra* §IV.F.2; EX1002, ¶¶453, 457, 459, 461.) It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Chatterjee. (EX1002, ¶¶454-456, 458, 460, 462.) For example, Williams-Lim and Chatterjee are analogous art to the '174 patent and all relate to controlling home appliances such as home-automation devices remotely. (*Supra* §IV.F.2; EX1006, [0051], [0012], [0007]; EX1005, [0007], [0017]; EX1012, Abstract; EX1001, Claim 1, 135:54-60.) Combining Chatterjee with Williams-Lim would improve and expand the capabilities of Williams-Lim in the same way. (EX1002; ¶456.)

O. Ground 15: Williams, Lim, and Rosenberger Render Obvious Claims 24-25

1. Claims 24/25

Rosenberger discloses claims 24-25. (*Supra* §G.2.)

It would have been obvious to combine the teachings of Rosenberger with Williams-Lim. (EX1002, ¶¶465-469, 472.) For example, Rosenberger, Williams, and Lim are analogous art to the '174 patent and all relate to processing and responding to audio commands. (EX1006, [0027], [0036]; EX1005, [0007], [0027]; EX1013, 1:7-13.) A POSITA would have been motivated to combine

Williams-Lim with Rosenberger to include Rosenberger's status light in the audio interface device, which turns on/off and/or changes color to indicate different device states and modes to improve the usability of the device, as taught by Lim/Rosenberger, and with a reasonable expectation of success. (EX1002, ¶¶467.)

P. Ground 16: Williams, Lim, and Rossing Render Obvious Claims 9 and 26-28

1. Claims 9, 26-28

Rossing discloses claim 9 and 26-28. (*Supra* §§IV.H.2, IV.H.3; EX1002, ¶¶474, 478, 480, 483.)

A POSITA would have understood that the claim 9 refers to a standard microphone and claims 27-28 refers to a standard speaker. (EX1002, ¶¶475, 480-481, 483.) Williams and Lim both disclose use of such a *microphone* and *speaker* in the *client device*, consistent with the description of the '174 patent. (*Id.*; EX1006, [0016], [0007], [0019], [0037], [0048], claims 1-3 and 21; EX1005, [0007], [0019]; EX1001, claim 9, 17:47-54, 105:48-52, 105:61-106:4.) Rossing is a textbook on how microphones and speakers operate. (EX1002, ¶¶476, 482, 484-485.) Thus, a POSITA would have been motivated to look to Rossing for techniques and details regarding these components for implementing Williams-Lim. (EX1002, ¶¶476-477, 479, 486.)

V. NOTICES AND STATEMENTS

Pursuant to 37 C.F.R. § 42.8(b)(1), Amazon.com Services LLC ("Amazon")

is the real party-in-interest for Petitioner.

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

- *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	Backup Counsel
Alex S. Yap Email: AYap@mofo.com Registration No. 60,609 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Tel: 213-892-5200 Fax: 213-892-5454	Mehran Arjomand Email: MArjomand@mofo.com Registration No.: 48,231 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Tel: (213) 892-5200 Fax: (213) 892-5454 Yunan Yuan Email: AYuan@mofo.com Registration No.: 73,637 MORRISON & FOERSTER LLP 755 Page Mill Road Palo Alto, CA 94304-1018 Tel: (650) 813-5600 Fax: (650) 494-0792

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), Petitioner certifies that the '174 patent is available for *inter partes* review and that Petitioner is 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**.

Dated: November 21, 2025

Respectfully submitted,

By /s/ Alex S. Yap

Alex S. Yap, Reg. No. 60,609
MORRISON & FOERSTER LLP
707 Wilshire Blvd., Suite 6000
Los Angeles, CA 90017
Email: AYap@mofo.com
Tel: 213-892-5200

Counsel for Petitioner

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

I hereby certify that this Petition for *Inter Partes* Review has 13,986 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, 2025

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. 12,010,174.

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

Dated: November 21, 2025

/s/ Alex S. Yap
Alex S. Yap