

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM SERVICES LLC
Petitioner,

v.

SMART SPEAKER LLC
Patent Owner.

Patent No. 12,401,721

Inter Partes Review No. IPR2026-00276

**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 '721 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, 9-15, 17, 21-22, 24-31, 33.....	8
1. Summary of Lim	8
2. Claim 1	13
a. 1[PR]: “A first device for use with a Wireless Local Area Network (WLAN) in a building, the first device comprising”	13
b. 1[A]: “a WLAN transceiver for communicating over the WLAN;”	15
c. 1[B]: “a microphone for capturing human voice data;”	16
d. 1[C]: “and a sounder for converting an electrical energy to audible sound waves;”	18
e. 1[D]: “and one or more processors programmed with computer program instructions that, when executed, cause the first device to:”	19
f. 1[E]: “capture, using the microphone, first and second human voice data;”	20
g. 1[F]: “send, over the Internet via the WLAN, to a server device, the captured first and second human voice data;”	24

TABLE OF CONTENTS (continued)

	Page
h. 1[G]: “receive, from the server device via the WLAN over the Internet, first and second messages, in response to the sending of the respectively captured first and second human voice data;”	28
i. 1[H]: “sound, using the sounder, digital audio content in response to the received first message;”	35
j. 1[I][1]: “and send, to a second device over the WLAN, a control message that comprises a control data for controlling a first actuator in the second device”	38
k. 1[I][2]: “to directly or indirectly affect, change, or produce, a first physical phenomenon,”	43
l. 1[I][3]: “in response to the received second message,”	44
m. 1[J]: “wherein each of the first and second devices is addressable in the WLAN and in the Internet using a respective Internet Protocol (IP) address.”	46
3. Claims 5/6	46
4. Claim 9/29/30/33.....	47
5. Claims 10/11/12	47
6. Claims 13/14/15	48
7. Claims 17/21/22/24.....	49
8. Claim 25.....	49
9. Claims 26/27	50
10. Claim 28.....	51
11. Claim 31	51
B. Ground 2: Lim and Wittenberg Render Obvious Claim 2.....	52
1. Summary of Wittenberg.....	52

TABLE OF CONTENTS
(continued)

	Page
2. Claim 2.....	52
C. Ground 3: Lim and Gruber Render Obvious Claim 3-4, 11	53
1. Summary of Gruber	53
2. Claim 3.....	54
3. Claim 4.....	56
4. Claim 11	57
D. Ground 4: Lim and Benesty Render Obvious Claims 7-8	57
1. Summary of Benesty	57
2. Claims 7/8	58
E. Ground 5: Lim and Rossing Render Obvious Claims 9, 29, 33	59
1. Summary of Rossing.....	59
2. Claim 9.....	59
3. Claims 29/33	61
F. Ground 6: Lim and Goldsmith Render Obvious Claims 13-15	62
1. Summary of Goldsmith.....	62
2. Claim 13	62
3. Claim 14.....	63
4. Claim 15	63
G. Ground 7: Lim and Chatterjee Render Obvious Claim 18-20, 23	64
1. Summary of Chatterjee	64
2. Claims 18-20 and 23	64
H. Ground 8: Lim and Rosenberger Render Obvious Claims 26-28.....	67
1. Summary of Rosenberger	67
2. Claims 26/27	67
3. Claim 28	69

TABLE OF CONTENTS (continued)

	Page
I. Ground 9: Williams and Lim Render Obvious Claims 1, 5-6, 9-15, 17, 21-22, 24-29, 31, 33	69
1. Summary of Williams	69
2. A POSITA would have been motivated to combine Williams and Lim.....	71
3. Claim 1	72
a. 1[PR]: “A first device for use with a Wireless Local Area Network (WLAN) in a building, the first device comprising”	72
b. 1[A]: “a WLAN transceiver for communicating over the WLAN;”	74
c. 1[B]: “a microphone for capturing human voice data;”	75
d. 1[C]: “and a sounder for converting an electrical energy to audible sound waves;”	76
e. 1[D]: “and one or more processors programmed with computer program instructions that, when executed, cause the first device to:”	77
f. 1[E]: “capture, using the microphone, first and second human voice data;”	77
g. 1[F]: “send, over the Internet via the WLAN, to a server device, the captured first and second human voice data;”	80
h. 1[G]: “receive, from the server device via the WLAN over the Internet, first and second messages, in response to the sending of the respectively captured first and second human voice data;”	84
i. 1[H]: “sound, using the sounder, digital audio content in response to the received first message;”	88

TABLE OF CONTENTS (continued)

	Page
j. 1[I][1]: “and send, to a second device over the WLAN, a control message that comprises a control data for controlling a first actuator in the second device”	89
k. 1[I][2]: “to directly or indirectly affect, change, or produce, a first physical phenomenon,”	91
l. 1[I][3]: “in response to the received second message,”	91
m. 1[J]: “wherein each of the first and second devices is addressable in the WLAN and in the Internet using a respective Internet Protocol (IP) address.”	92
4. Claims 5/6/26/27	93
5. Claims 9/29/33	93
6. Claims 10/11/12	93
7. Claims 13/14/15	94
8. Claims 17/21/22/24	94
9. Claim 25	95
10. Claim 28	96
11. Claim 31	96
J. Ground 10: Williams, Lim, and Wittenberg Render Obvious Claim 2	98
K. Ground 11: Williams, Lim, and Gruber Render Obvious Claims 3-4, 11	98
1. Claims 3-4	98
2. Claim 11	99
L. Ground 12: Williams, Lim, and Rosenberger Render Obvious Claims 5-6, 26-27, 28	99
1. Claims 5/6/26/27	99

TABLE OF CONTENTS
(continued)

		Page
	2. Claim 28	100
M.	Ground 13: Williams, Lim, and Benesty Render Obvious Claims 7-8	101
N.	Ground 14: Williams, Lim, and Rossing Render Obvious Claims 9, 29, 33.....	101
O.	Ground 15: Williams, Lim, and Goldsmith Render Obvious Claims 13-15	102
P.	Ground 16: Williams, Lim, and Chatterjee Render Obvious Claim 18-20, 23	102
V.	NOTICES AND STATEMENTS.....	103

Exhibit List

Exhibit Description	Exhibit #
U.S. Patent No. 12,401,721 to Binder <i>et al.</i>	1001
Expert Declaration of Dr. Henry Houh	1002
File History of U.S. Patent No. 12,401,721	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

Exhibit Description	Exhibit #
Kenneth J. Christensen <i>et al.</i> , “The next frontier for communications networks: power management” (2004) (available at https://www.sciencedirect.com/science/article/abs/pii/S0140366404002427)	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
U.S. Patent No. 8,635,243 to Philips <i>et al.</i>	1023
U.S. Patent Pub. No. 2011/0313775 to Lalgand <i>et al.</i>	1024
U.S. Patent Pub. No. 2012/0253480 to Abe	1025
U.S. Patent No. 8,200,798 to Yadav <i>et al.</i>	1026
U.S. Patent No. 7,593,718 to Gorday <i>et al.</i>	1027
U.S. Patent Pub. No. 2006/0135067 to Dunko	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

Exhibit Description	Exhibit #
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
David Chen et al., “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)	1035
Yu-Shuo Chang et al., “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)	1036
Zhemin Tu et al., “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)	1037
Declaration of Tony Corsini on the scanned copies of pages of Jim Doherty <i>et al.</i> , “Wireless Home Networking Simplified” (Cisco Press, 2007).	1038
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).	1039

Exhibit Description	Exhibit #
Declaration of Rachel J. Watters on the scanned copies of Chapter 1 and Appendix D of Andrea Goldsmith, “Wireless Communications” (Cambridge University Press, 2005).	1040
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).	1041
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).	1042
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).	1043
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).	1044
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).	1045
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).	1046
RESERVED	1047

Exhibit Description	Exhibit #
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)	1048
Declaration of Tony Corsini on the scanned copies of pages of Jacob Fraden, “Handbook of Modern Sensors” (Springer Science+Business Media, LLC, 2010)	1049
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)	1050
File History of U.S. Patent No. 12,316,706	1051
File History of U.S. Patent No. 12,010,174	1052
iPhone User Guide For iOS 5.0 Software, https://web.archive.org/web/20111107154723/http://support.apple.com/manuals/en/iPhone_User_Guide.pdf , Apple Support (November 7, 2011)	1053

Claims Listing

Claim	Limitation #	Limitation
1	1[PR]	A first device for use with a Wireless Local Area Network (WLAN) in a building, the first device comprising:
	1[A]	a WLAN transceiver for communicating over the WLAN;
	1[B]	a microphone for capturing human voice data;
	1[C]	and a sounder for converting an electrical energy to audible sound waves;
	1[D]	and one or more processors programmed with computer program instructions that, when executed, cause the first device to:
	1[E]	capture, using the microphone, first and second human voice data;
	1[F]	send, over the Internet via the WLAN, to a server device, the captured first and second human voice data;
	1[G]	receive, from the server device via the WLAN over the Internet, first and second messages, in response to the sending of the respectively captured first and second human voice data;
	1[H]	sound, using the sounder, digital audio content in response to the received first message;
	1[I][1]	and send, to a second device over the WLAN, a control message that comprises a control data for controlling a first actuator in the second device
	1[I][2]	to directly or indirectly affect, change, or produce, a first physical phenomenon,

Claim	Limitation #	Limitation
	1[I][3]	in response to the received second message,
	1[J]	wherein each of the first and second devices is addressable in the WLAN and in the Internet using a respective Internet Protocol (IP) address.
2	2	The first device according to claim 1, wherein the first device or the server device is configured for processing the first and second human voice data using a voice recognition algorithm for identifying a voice of a specific person.
3	3	The first device according to claim 1, further comprising a sensor that outputs sensor data that responds to a physical phenomenon, wherein the computer program instructions, when executed, further cause the first device to send to the server device, via the WLAN over the Internet, the sensor data, and wherein the first message is further responsive to the sent sensor data.
4	4	The first device according to claim 3, wherein the sensor is a thermoelectric sensor that responds to a temperature or to a temperature gradient of an object using conduction, convection, or radiation, or wherein the sensor is a photoelectric sensor that responds to a visible or an invisible light or gamma rays.
5	5	The first device according to claim 1, further comprising a second actuator that is configured for directly or indirectly affect, change, or produce, a second physical phenomenon.
6	6	The first device according to claim 5, wherein the second physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical

Claim	Limitation #	Limitation
		voltage, or electrical current.
7	7	The first device according to claim 1, further comprising an additional microphone for capturing human voice data, and wherein the capturing comprises capturing, by the additional microphone, the first and second human voice data.
8	8	The first device according to claim 7, wherein the microphone and the additional microphone are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array.
9	9	The first device according to claim 1, wherein the microphone is an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, or wherein the microphone comprises a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone.
10	10	The first device according to claim 1, wherein the IP address is stored in a volatile or non-volatile memory.
11	11	The first device according to claim 1, further addressed by a Media Access Control (MAC) layer address that is MAC-48, Extended Unique Identifier (EUI) EUI-48, or EUI-64 address type.
12	12	The first device according to claim 1, wherein the IP address is a static or dynamic Internet Protocol (IP) address that is IPv4 or IPv6 type address.
13	13	The first device according to claim 1, wherein the WLAN is according to, or is compatible with, an

Claim	Limitation #	Limitation
		Institute of Electrical and Electronics Engineers (IEEE) 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard.
14	14	The first device according to claim 1, wherein the WLAN uses an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.
15	15	The first device according to claim 1, further configured for communicating over a cellular telephone network that is a Third Generation (3G) network that uses, or is compatible with, Universal Mobile Telecommunications First device (UMTS), Wideband Code Division Multiple Access (W-CDMA) UMTS, High Speed Packet Access (HSPA), UMTS Time-Division Duplexing (TDD), CDMA2000 1×RTT, Evolution-Data Optimized (EV-DO), or Global First device for Mobile communications (GSM), Enhanced Data rates for GSM Evolution (EDGE) EDGE-Evolution.
17	17	The first device according to claim 1, wherein the second device is integrated in, is part of, or is entirely included in, a household appliance having a primary function.
18	18	The first device according to claim 17, wherein the primary functionality of the appliance is associated with food storage, handling, or preparation.
19	19	The first device according to claim 18, wherein the primary function of the appliance is heating food, and wherein the appliance is a microwave oven, an electric mixer, a stove, an oven, or an induction cooker.

Claim	Limitation #	Limitation
20	20	The first device according to claim 18, wherein the appliance is a refrigerator, a freezer, a food processor, a dishwasher, a food blender, a beverage maker, a coffeemaker, or an iced-tea maker.
21	21	The first device according to claim 17, wherein the primary function of the appliance is associated with environmental control, and the appliance is part of a Heating, Ventilation and Air Conditioning (HVAC) system.
22	22	The first device according to claim 21, wherein the primary function of the appliance is associated with temperature control, and wherein the appliance is an air conditioner or a heater.
23	23	The first device according to claim 17, wherein the primary function of the appliance is associated with cleaning, wherein the appliance primary function is associated with clothes cleaning and the appliance is a washing machine, or wherein the appliance is a vacuum cleaner.
24	24	The first device according to claim 17, wherein the appliance is an answering machine, a telephone set, a home cinema system, a High Fidelity (HiFi) system, a Compact Disc (CD) or Digital Video Disc (DVD) player, an electric furnace, a trash compactor, a smoke detector, a light fixture, or a dehumidifier.
25	25	The first device according to claim 1, further configured to operate within a building, and wherein the server device is external to the building.
26	26	The first device according to claim 1, further comprising an electric light source for converting electrical energy into light that emits visible or non-

Claim	Limitation #	Limitation
		visible light for illumination or indication, and the non-visible light is infrared, ultraviolet, X-rays, or gamma rays.
27	27	The first device according to claim 26, wherein the electric light source comprises a lamp, an incandescent lamp, a gas discharge lamp, a fluorescent lamp, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode.
28	28	The first device according to claim 27, wherein the sounding comprises playing digital audio content that is pre-recorded or synthesized, wherein the sounding comprises simulating the voice of a human being or generating music, or wherein the sounding comprises sounding a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.
29	29	The first device according to claim 1, wherein the sounder comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker.
30	30	The first device according to claim 1, wherein the sounder comprises an electric bell, a buzzer, a chime, a whistle, or a ringer.
31	31	The first device according to claim 1, further comprising an enclosure that houses the WLAN transceiver, the sounder, the microphone, and the one or more processors.
33	33	The first device according to claim 1, wherein the sounder is configured for converting an electrical energy to omnidirectional, unidirectional, or bidirectional pattern emitted, audible or inaudible,

Claim	Limitation #	Limitation
		sound waves.

Petitioner Amazon.com Services LLC respectfully petitions for *inter partes* review of claims 1-15, 17-31, 33 of U.S. Patent No. 12,401,721 (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, 1.) 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 '721 patent covers well-known and routine smart-home functionalities. The Director should institute review on the '721 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 never issued any prior art-based rejections, relying on the allowance of a parent patent that the PTAB has already determined to mostly recite garden-variety limitations. (EX1003, 203-228; EX1004, 389-421.) As explained herein, devices capable of receiving/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 herein, as supported by Dr. Houh. (EX1002.)

II. THE '721 PATENT

A. Summary

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

Like the claims, 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, 94:3-21; 96:10-97:29; 99:5-21; 100:7-101:5; 104:8-29; 105:48-106:53; 109:42-110:25.) 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.) A paragraph in the

Summary Section describes voice control over a network 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.) 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 '721 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 Examiner did not issue any prior art-based rejections and the application issued on August 26, 2025, less than 6 months from filing. (EX1003, 160-162, 366-369.) The Examiner issued a single Office Action containing double-patenting rejections with respect to parent patents U.S. Patent Nos. 11,128,710; 12,010,174; and 12,316,706. (EX1003, 203-228.) After a terminal disclaimer was submitted, the Examiner allowed the claims, citing entire claim 1 as the reason for allowance. (*Id.*, 170-176.)

But the examination of the parent patents was just as deficient. The earliest patent (the '710 patent) went through *ex parte* appeal, where the PTAB affirmed the rejections, finding that a client device that receives a voice command, which triggers operation of a ***controlled device***, to be obvious. (EX1004, 389-421.) On remand, the Applicant then added a trivial requirement of a voice command that triggers operation of the ***client device*** itself. (EX1004, 344-350, 121-130.) Despite the minimal and well-known addition, the Examiner abruptly allowed the application without tailoring the search strategy. (See EX1004, 8-14; EX1002, ¶¶42-50.). Following the grant of the '710 patent, the '174 and '706 patents sailed through examination without any prior art-based Office Action. *Id.*

D. Level of Ordinary Skill in the Art

A person of ordinary skill in the art (POSITA) in the field of the '721 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, ¶¶56-58.)

III. CLAIM CONSTRUCTION

Claim terms are presumed to carry their ordinary and customary meaning. The Board construes the terms only to the extent necessary to resolve the issue at hand: here, the invalidity of the claims under 35 U.S.C. §103. The Board need not

construe any of the claims because, as set forth in the Petition, the claims are invalid under 35 U.S.C. §103 under any reasonable interpretation, and thus, no claim construction is necessary to resolve any issue before the Board. (EX1002, ¶¶59.)

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, ¶¶60-70).

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)

Prior Art Reference	Abbreviation (Exhibit No.)	Prior Art Basis
Jacob Benesty <i>et al.</i> , “Springer Handbook of Speech Processing” (Springer-Verlag Berlin Heidelberg, 2008).	“Benesty” (EX1010)	§102(a) and (b)
Andrea Goldsmith, “Wireless Communications” (Cambridge University Press, 2005).	“Goldsmith” (EX1040)	§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, 9-15, 17, 21-22, 24-31, 33	Lim
2	103	2	Lim, Wittenberg
3	103	3-4, 11	Lim, Gruber
4	103	7-8	Lim, Benesty
5	103	9, 29, 33	Lim, Rossing
6	103	13-15	Lim, Goldsmith
7	103	18-20, 23	Lim, Chatterjee
8	103	26-27, 28	Lim, Rosenberger
9	103	1, 5-6, 9-15, 17, 21-22, 24-29, 31, 33	Williams, Lim
10	103	2	Williams, Lim, Wittenberg
11	103	3-4, 11	Williams, Lim, Gruber
12	103	5-6, 26-28	Williams, Lim, Rosenberger
13	103	7-8	Williams, Lim, Benesty
14	103	9, 29, 33	Williams, Lim, Rossing
15	103	13-15	Williams, Lim, Goldsmith
16	103	18-20, 23	Williams, Lim, Chatterjee

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, 9-15, 17, 21-22, 24-31, 33

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 discloses that “a *controlling device* may be provisioned with means to receive voice input from a user, e.g., **a microphone**.” (EX1005, [0007] (emphasis added).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, ¶71.)

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.

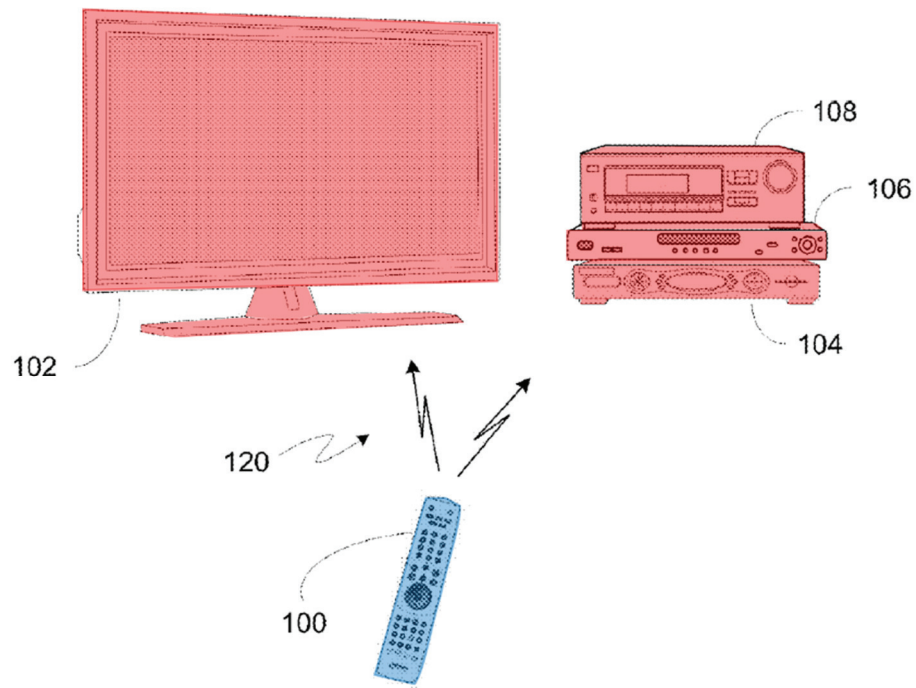


Figure 1a

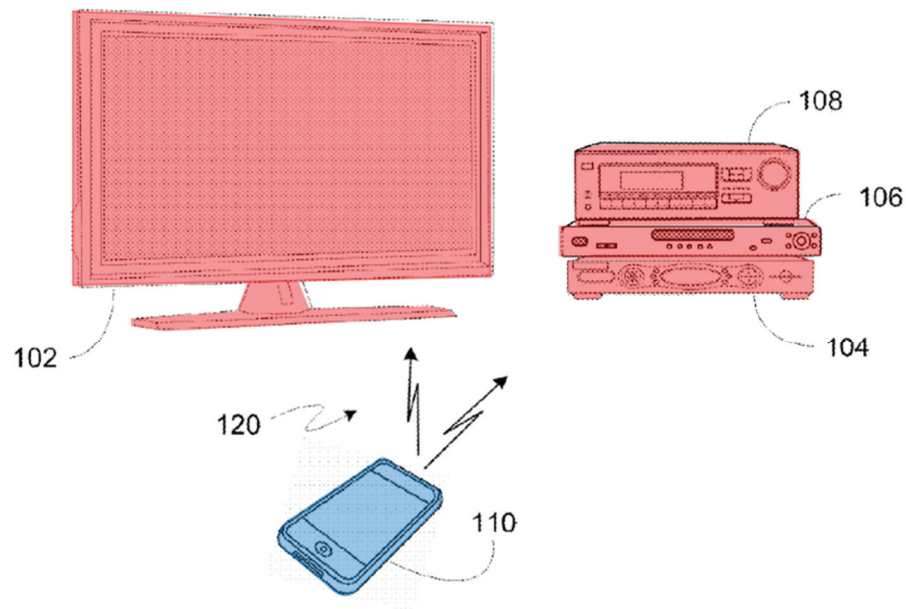


Figure 1b

Lim's *controlling device* receives speech commands and responds with an *audio output*. For example, during a “setup and configuration” stage (e.g. FIG. 4),

the *controlling device* can “prompt the user to speak the brand name of the appliance” in block 402.” (EX1005, [0027].) As shown in the yellow blocks, the *controlling device* responds 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, ¶72.)

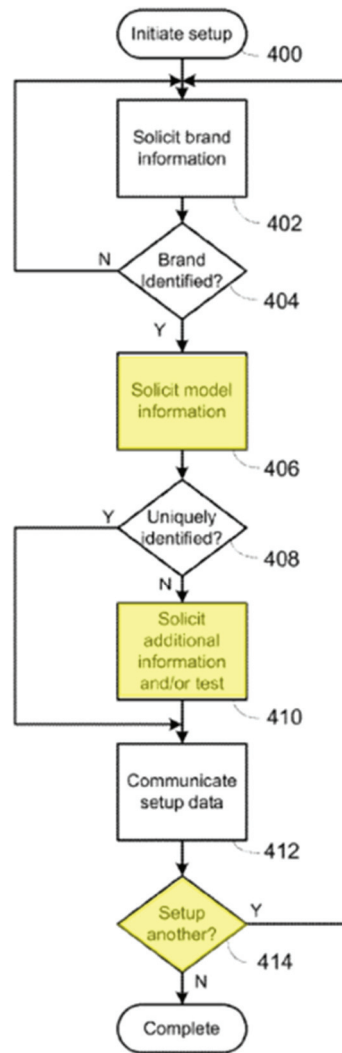


Figure 4 “Set up and configuration of controlling device”
Lim, [0025].

The *controlling device* can control appliances. Specifically, the *controlling device* receives speech commands and transmits *control commands* to *appliances* (e.g., *second devices*), as shown in FIG. 6 below. (EX1005, [0017], [0001]; EX1002, ¶¶73-74.)

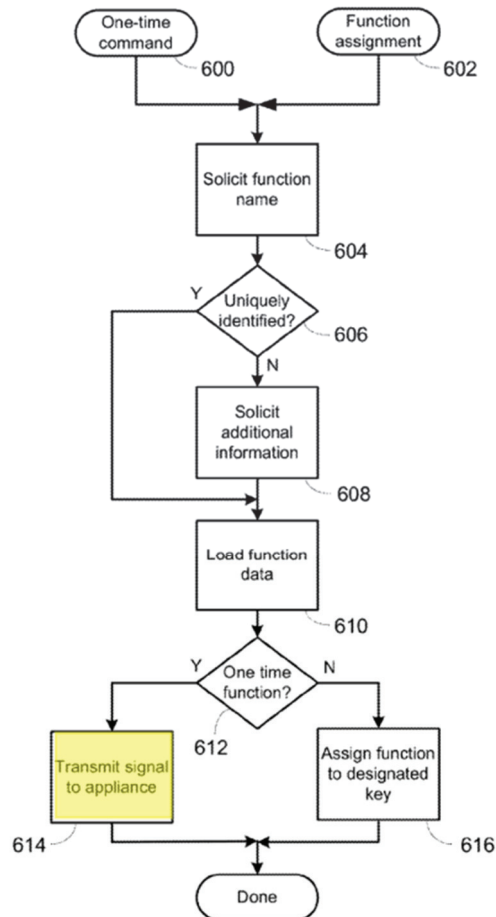


Figure 6 "Appliance command"
Lim, [0035].

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, ¶¶75-76.) As shown in FIG. 5b below, the “*remote server 500* [is] accessible via a network 502 such as the *Internet* or PSTN.” Lim discloses that “access to such network [502] may be . . . via a *local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router* (not illustrated).” (*Id.*; EX1005, [0027].)

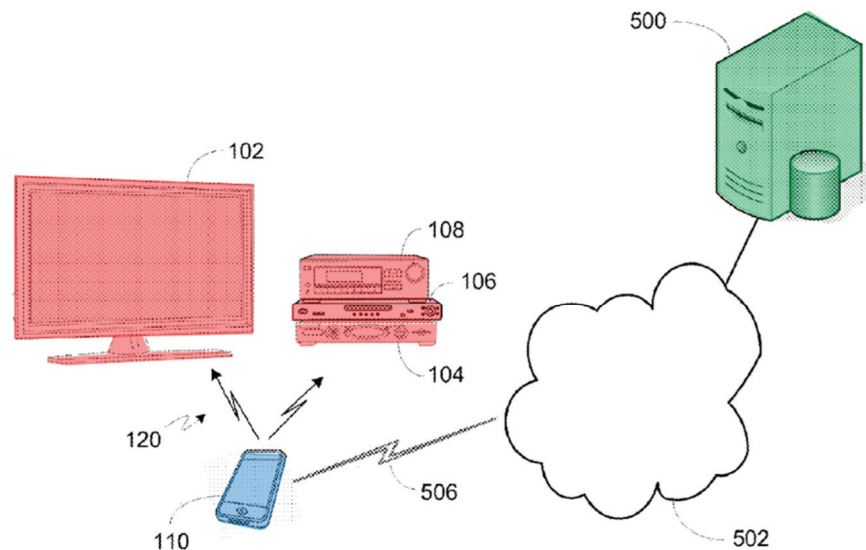


Figure 5b

Lim discloses that the *controlling device* has a microphone, a display, a speaker, various sensors, one or more processors, and WiFi/Bluetooth capabilities. (EX1005, [0023], FIG. 3; EX1002, ¶77.)

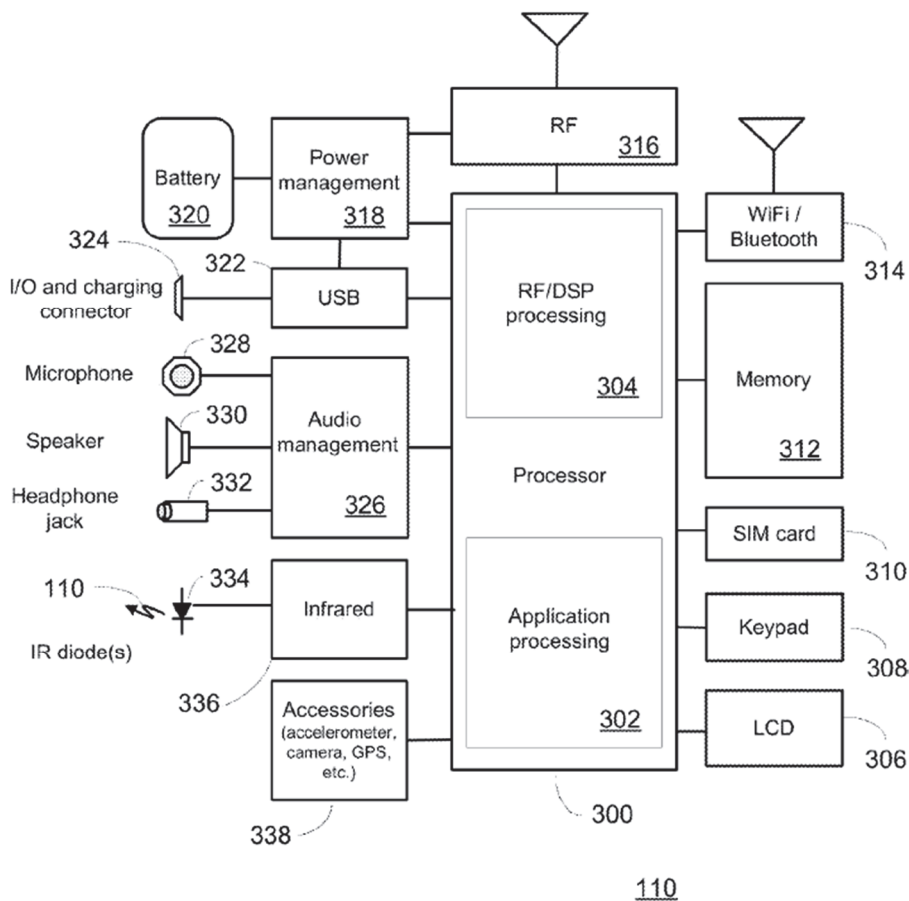


Figure 3

2. Claim 1

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

Lim’s *controlling device* is a *first device*. (EX1002, ¶¶78-79.) “[A] *controlling device* may be provisioned with means to receive voice input from a user, e.g., a microphone”. (EX1005, [0007], [0017].) Lim’s *controlling device* may be “a *smart device 110*” as shown in FIGs. 1b and 5b. (*Id.*, [0018], [0023].)

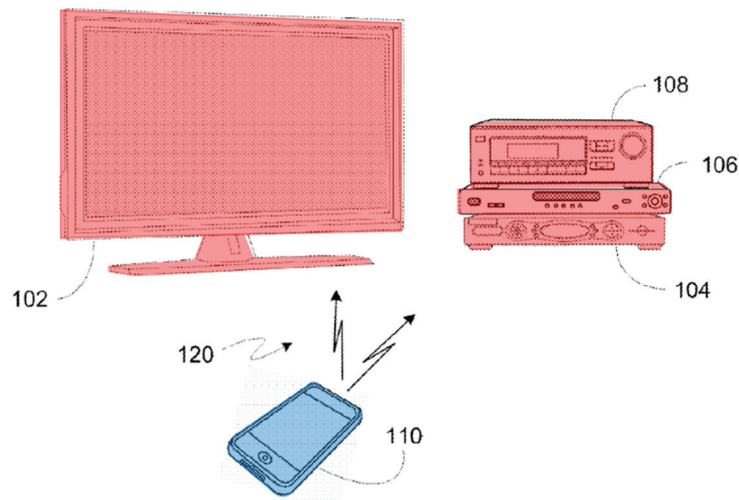


Figure 1b

Lim's *first/controlling device* is for use with a **WLAN**. Referring to FIG. 5b, Lim's *first/controlling device* uses “speech recognition processing [that] may be performed by programming resident on a **remote server 500** accessible via a network 502 such as the Internet” and “access to such network may be . . . via a local network connection 508, such as **WiFi** or Bluetooth, *through a gateway device such as a router* (not illustrated).” (EX1005, [0027]; EX1002, ¶80.) It was common knowledge that WiFi is a type of WLAN. (EX1040, 600-601; EX1013, 7:48-51; EX1002, ¶81.) The '721 patent confirms the same. (*Id.*; EX1001, 8:25-32.)

Lim discloses or certainly suggests that the *first device* is in a **building**. Lim discloses that the *smart device 110* can be “a smart phone” and other portable devices. (EX1005, [0018].) Lim certainly suggests, and a POSITA would find it obvious, that these are devices that users can carry with them, including when the

user is at *home*, to conveniently control home appliances from home. (EX1002, ¶82; EX1021, 1:53-61, 2:19-26.)

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

Lim’s *first/controlling device* has a *WLAN transceiver* for communicating over the WLAN. (EX1002, ¶83.)

Lim’s *controlling device 100* “may include . . . Zigbee, *WiFi*, or Bluetooth *transceiver*[.]” (EX1005, [0019], [0027].) It was common knowledge that WiFi is a type of WLAN. (*Supra* §IV.A.2.a; EX1040, 600-601; EX1013, 7:48-51; EX1001, 8:25-32; EX1002, ¶84.) Thus, Lim’s “WiFi [] transceiver” is a WLAN transceiver. *Id.*

A POSITA would have found it obvious that the *smart device 110* similarly has a WLAN transceiver. Indeed, referring to FIG. 3, Lim discloses that the *smart device 110* “may include . . . *WiFi and/or Bluetooth wireless interface(s) 314*[.]” (EX1005, [0023].) A POSITA would have understood that a “WiFi wireless interface” refers to the hardware and software that enable WiFi communication, including a WiFi transceiver. (EX1022, [0067], EX1002, ¶85.) Indeed, in 2012, mobile phones commonly included WiFi transceivers. (*Id.*; EX1035, 3; EX1027, 1:20-23.)

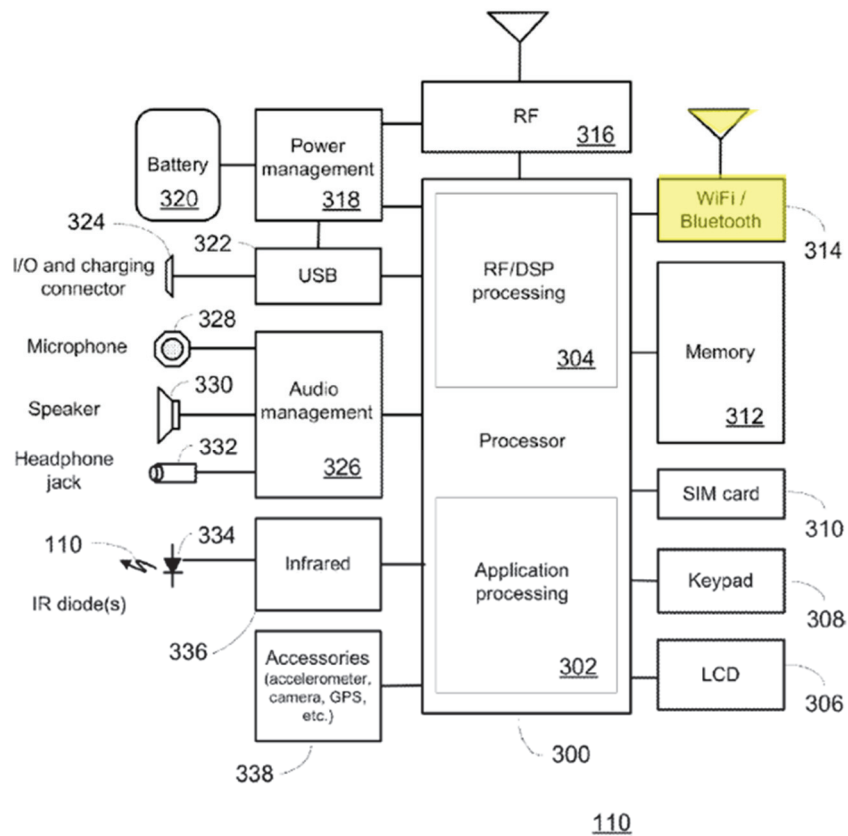
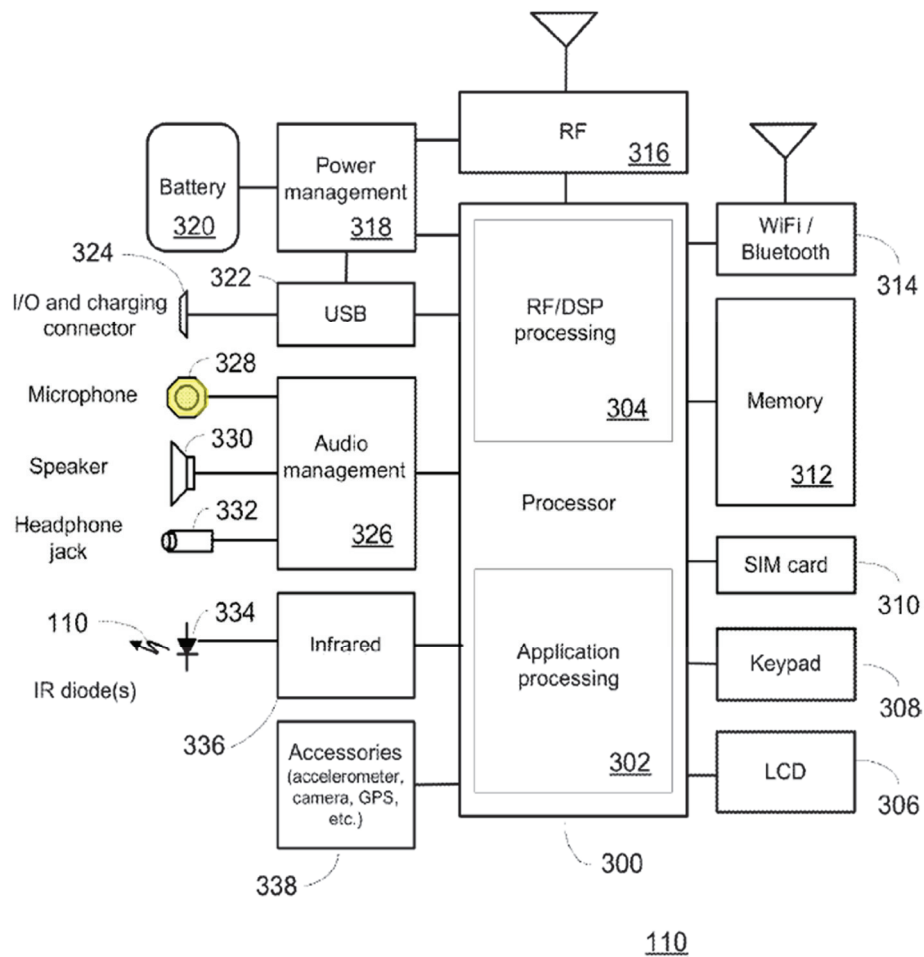


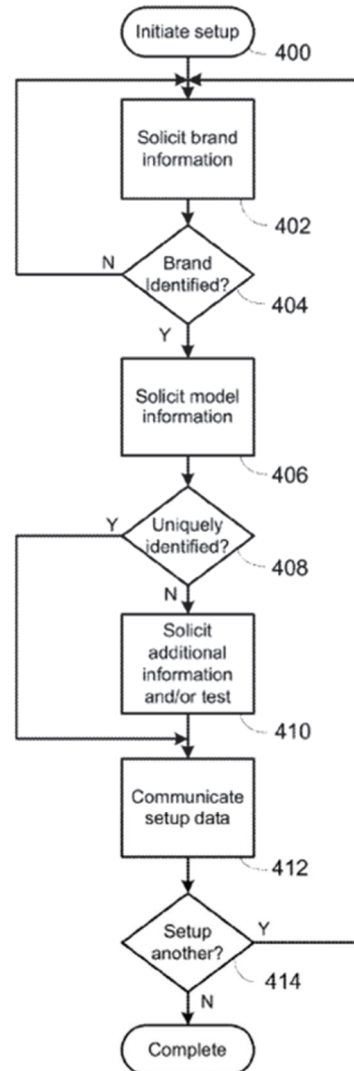
Figure 3

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

Lim’s *first/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, ¶86.)



Lim's microphone in the *first/controlling device* captures human voice data. For example, referring to FIG. 4, "a user may provide appliance identification via *spoken input*. . . ." (EX1005, [0025], [0013], [0027]-[0030], FIG. 4.) "The resulting spoken syllables may be digitized and presented [by the *first/controlling device*] to a speech recognition means" (EX1005, [0027].) It would have been obvious to a POSITA that the *controlling device*'s microphone captures these voice inputs. (EX1002, ¶¶87-88.)



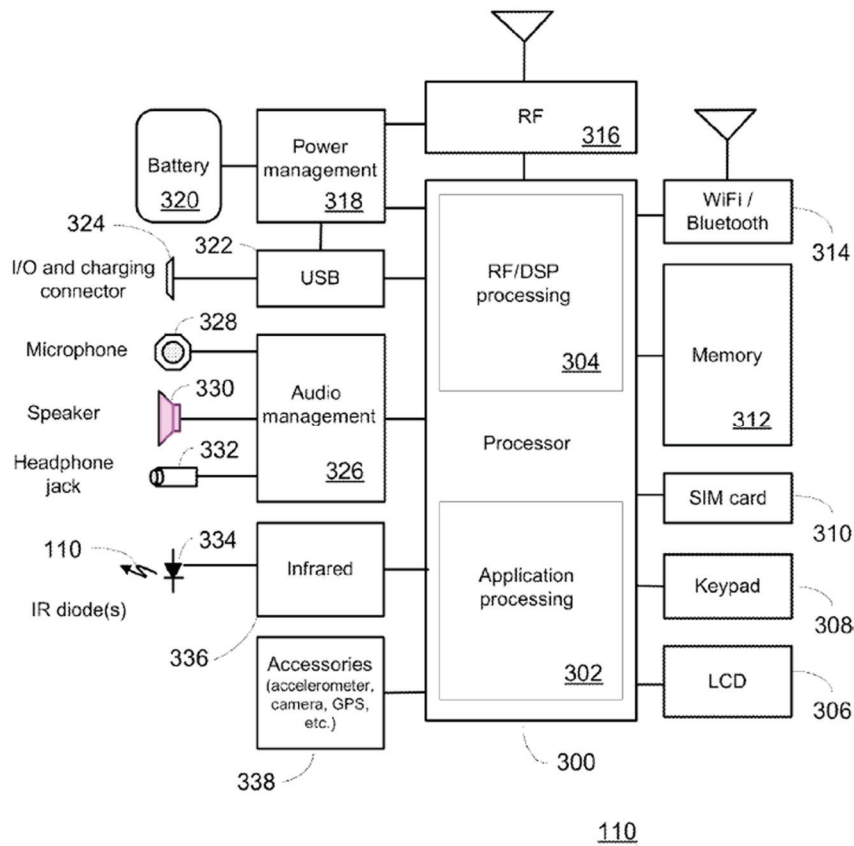
“Set up and configuration
of controlling device”
Lim, [0025].

Figure 4

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

Lim’s *first/controlling device* has a *sounder*. Referring to FIG. 3, the *controlling device* includes “an audio management system 326 with associated microphone 328, *speaker 330*[.]” (EX1005, [0023], [0019].) A POSITA would have understood that Lim’s *speaker 330* is a sounder for converting an electrical

energy to audible sound waves. (EX1002, ¶¶89-93; EX1018, 122-123, Figure 10.8; EX1014, 763.) The '721 patent confirms the same. (EX1001, 105:61-106:4.)



Additionally, Lim’s *controlling device 100* “may include . . . [a] *buzzer*[.]” (EX1005, [0019].) It would have been obvious that *smart device 110* is similarly equipped with a buzzer. (EX1002, ¶94.) A buzzer is another example of a sounder, consistent with the '721 patent. (*Id.*; EX1001, 106:54-62.)

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

Lim discloses this element. For example, *controlling device 100* includes “a

processor 200” coupled to memories, which “may include *executable instructions* . . . that are intended to be *executed by the processor 200 to control the operation of the controlling device 100*.” (EX1005, [0019]-[0020].) Similarly, as shown in in FIG. 3 below, *controlling device 110* includes “processing means 300.” (EX1005, [0023].) A POSITA would have understood that the “processing means 300” would also be programmed with computer program instructions that, when executed, cause the *controlling device 110* to perform the various operations disclosed in Lim. (*Infra* §§ IV.A.2.f-IV.A.2.i; EX1002, ¶¶95-97.)

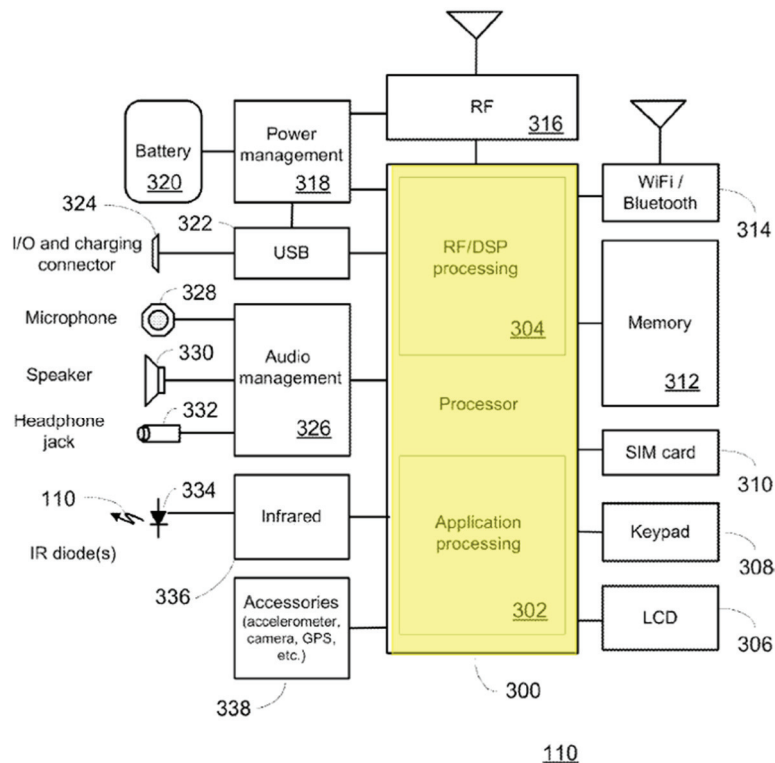


Figure 3

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

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

In Lim, the voice commands that trigger auditory responses by the first/controlling device are the recited “*first human voice data*”, and the voice commands that cause sending of a control message by the first/controlling device are the recited “*second human voice data*”, as discussed below. (EX1002, ¶¶98-99.)

Per the claim language of claim 1, “*first human voice data*” includes human voice data that causes the *sounder* to sound digital audio content in the recited manner. (*Infra* §IV.A.2.i; EX1002, ¶100.) Lim discloses many such voice commands (“*first human voice data*”). For example, referring to FIG. 4, “a user may provide appliance identification via *spoken input*. . . .” (EX1005, [0025].) These spoken inputs include “*speak[ing] the brand name of the appliance*,” “*spell[ing] the first few characters of the brand name*,” “*indicat[ing] if they wish to configure additional appliances*”, etc. (*Id.*, [0027]-[0030], FIG. 4.) It would have been obvious to a POSITA that the *first/controlling device*’s microphone captures these voice inputs. (EX1002, ¶¶101-102.) Further, these voice commands cause the speaker of the *first/controlling device* to provide an audio response, as discussed below. (*Id.*; *supra* §IV.A.2.i; EX1005, [0027]-[0030].)

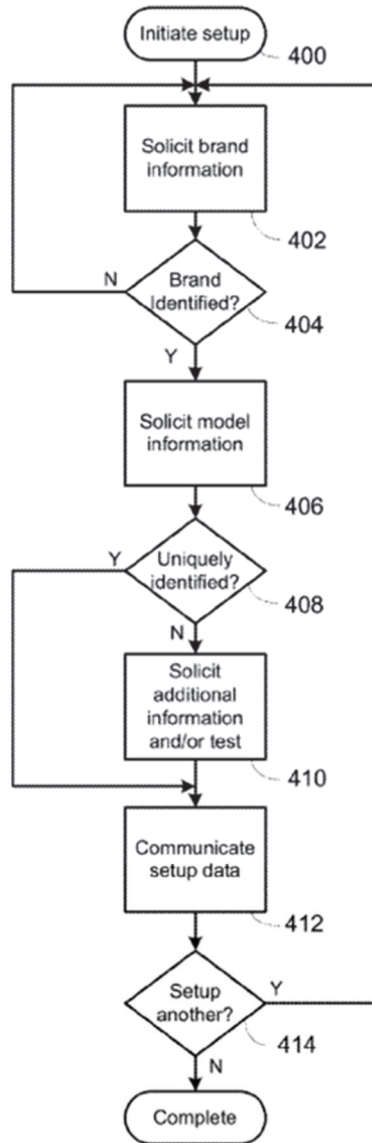


Figure 4 “Set up and configuration of controlling device”
Lim, [0025].

Per the claim language of claim 1, “*second human voice data*” includes human voice data that causes sending of a control message in the recited manner. (*Infra* §§IV.A.2.j-IV.A.2.l; EX1002, ¶¶103-105.) In FIG. 6 below, “[a]t step 604 the user may be prompted to *speak the desired function name*,” which “may be

digitized and presented to a *speech recognition means for processing* . . . on a *remote server 500*. . . .” (EX1005, [0035]-[0036].) Based on the processing, the *first/controlling device* “transmit[s] the desired command to the target appliance” at block 614. *Id.*

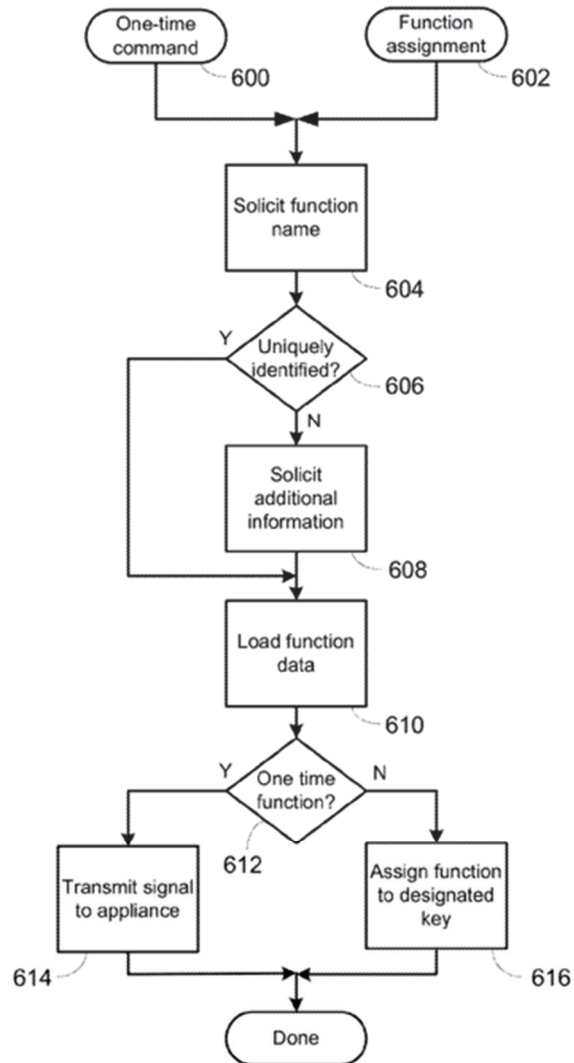


Figure 6 “Appliance command”
Lim, [0035].

Alternatively, to the extent that the first and second human voice data are

captured/sent together (*e.g.*, as parts of one query) (which is not conceded), a POSITA would find it obvious that a user command to operate an appliance (*e.g.*, speaking the desired function name in FIG. 6) in Lim can include multiple parts or phrases, such as a trigger phrase (“*first human voice data*”) and the instruction (“*second human voice data*”), consistent with common implementations. (EX1023, 2:12-16, 34:63-67; EX1013, 8:20-32; EX1024, [0045], [0047]; EX1002, ¶¶106-107.)

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

Lim discloses a *server device*. “All or part of said speech recognition processing . . . may be performed by programming resident on a *remote server 500*[.]” (EX1005, [0027]; EX1002, ¶108.)

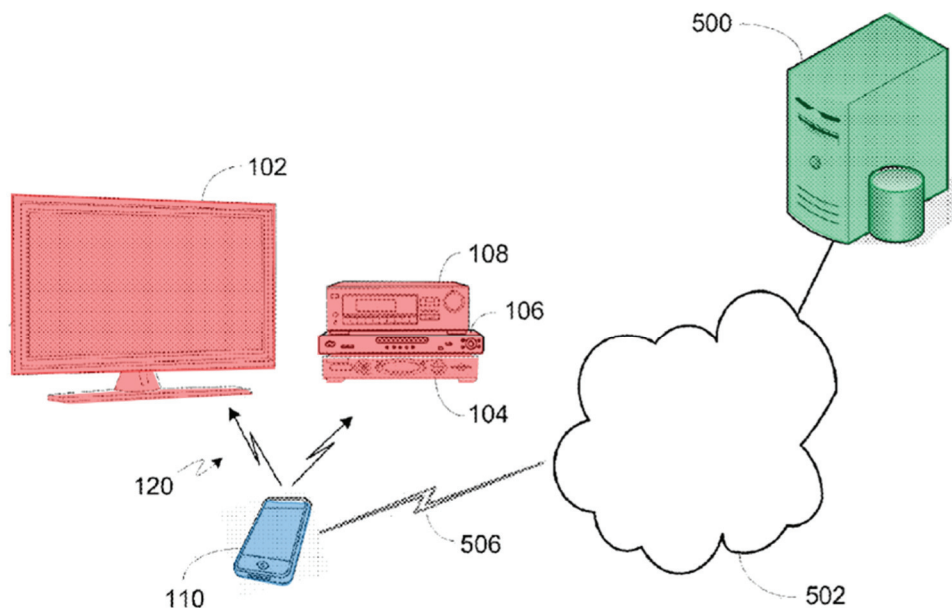
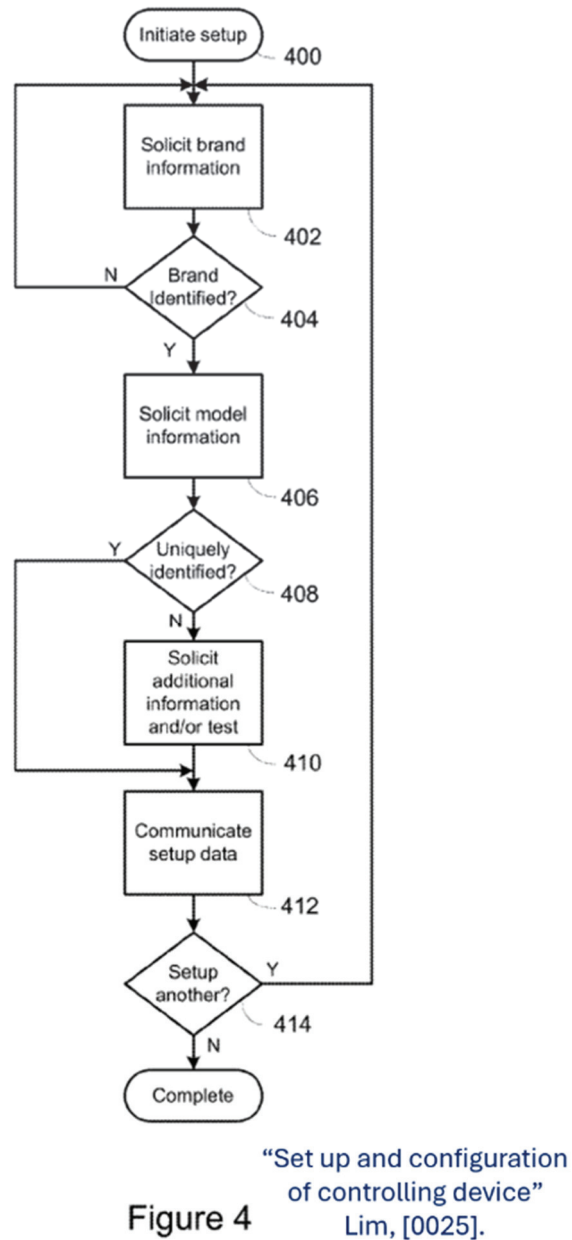


Figure 5b

Lim's *first/controlling device* communicates to the *remote server 500* over the *Internet via the WLAN*. “[A]ccess to such network [i.e., ‘a network 502 such as the Internet or PSTN’] may be . . . via a *local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router* (not illustrated).” (EX1005, [0027].) In other words, Lim's *first/controlling device* connects to a local WiFi network (i.e., WLAN) via a router, which in turn accesses the *remote server 500* over the Internet. This was a common configuration for a home network in 2012. (EX1002, ¶¶109-111; EX1007, 11, 100.)

Lim's *first/controlling device* sends the captured human voice data to the *server* over the Internet via the WLAN. (EX1002, ¶¶112-114.) Regarding the *first human voice data*, for example, “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.” (EX1005, [0027].) Other instances of “*first human voice data*” are similarly sent and processed. (EX1005, [0027]-[0030], FIG. 4; EX1002, ¶114.)



The *second human voice data* is similarly sent. Referring to FIG. 6 below, “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, [0034]-[0035]; EX1002, ¶115.)

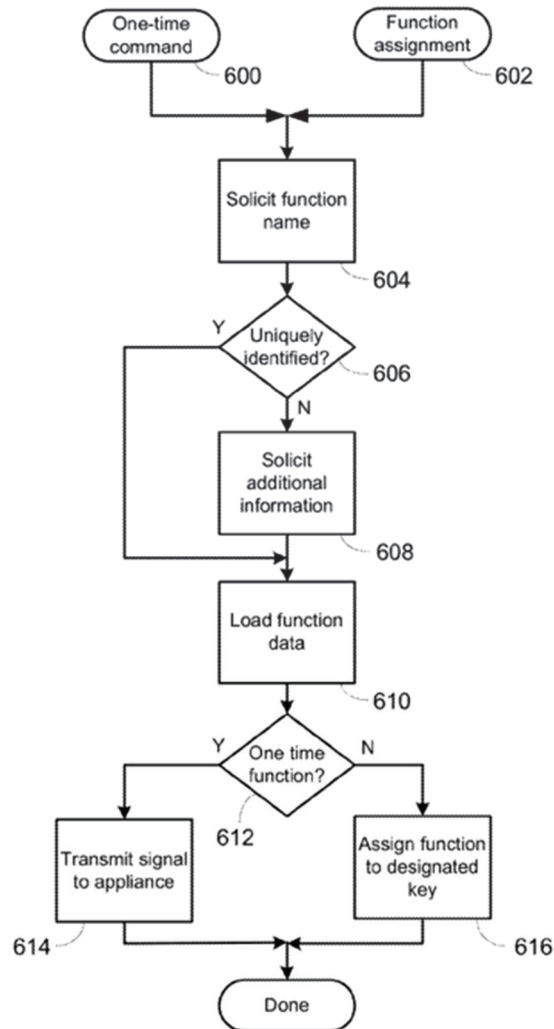


Figure 6 “Appliance command”
Lim, [0035].

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

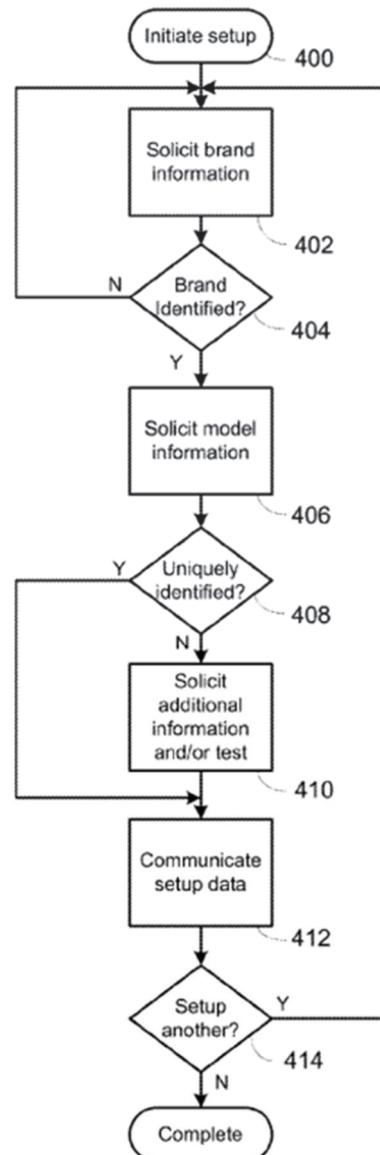
together to the server. (EX1002, ¶¶116-117.)

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

Lim discloses or at least suggests this element. As discussed, Lim’s *first/controlling device* sends the captured first and second human voice data to the *server* for speech processing. (*Supra* § IV.A.2.g.) It would have been obvious that *speech processing results* (i.e., *first message* and *second message*) are sent from the *server* back to the *client device*. (EX1002, ¶¶118-120.) Indeed, at the time of the invention, this was a routine operation by remote speech processing servers. (E.g., EX1009; EX1036, 307; EX1037, 2367.)

First, Lim discloses or at least suggests that the *first/controlling device* receives a *first message* from the *server* in response to the sending of the captured *first human voice data*. (EX1002, ¶119.) For example, at step 402, a user’s voice input specifying “*brand name of the appliance*” captured by the *first/controlling device* may be “presented to a speech recognition means [*remote server 500*] for processing.” (EX1005, [0027].) A POSITA would find it obvious that *speech processing result(s)* (i.e., first message) is sent back to the *first/controlling device* such that “[a]t step 404, *if the speech recognition means is unable to uniquely identify the desired brand* [based on the *speech processing result(s)*], step 402 may be repeated [by *first/controlling device*] in order to solicit additional

identifying information from the user.” (EX1005, [0028]; EX1002, ¶121.) A POSITA would have understood that 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, providing a buzz, or playing a sound). (*Id.*; EX1009; EX1036, 307; EX1037, 2367.)



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

Figure 4

In Lim, there are other instances of *first messages* received by the *first/controlling device* for it to determine the next prompt, for example, to determine “[i]f the initially provided model information is insufficient to uniquely identify the appliance [based on the *speech processing result(s)*]” in block 408 or

“if [the user] wish to configure additional appliances [based on the *speech processing result(s)*]” in block 414. (EX1005, [0029]-[0030].) Such *first message* is received “in response to the sending of the captured *first human voice data*” because the *speech processing result(s)* is generated based on the *server*’s processing of *first human voice data*. (EX1002, ¶¶122-124.)

Lim similarly discloses or at least suggests its *first/controlling device* receives a *second message* from the *server* in response to the sending of the captured *second human voice data*. For example, in FIG. 6 below, “at step 604 the user may be prompted to *speak the desired function name*.” (EX1005, [0035].) The voice data is sent to the “remote server” for processing. *Id.* Then, *speech processing result(s)* (*i.e.*, second message) is sent back to the *first/controlling device* such that “[o]nce the desired function has been uniquely identified [based on the *speech processing result(s)*], at step 610 the indicated command code data may be provisioned to the controlling device.” (EX1005, [0036]; EX1002, ¶¶125-126.) A POSITA would find it obvious that the second message would be the recognized speech (*e.g.* text strings), a result of further analyzing the recognized speech, etc. (*Id.*; EX1036, 307; EX1037, 2367.)

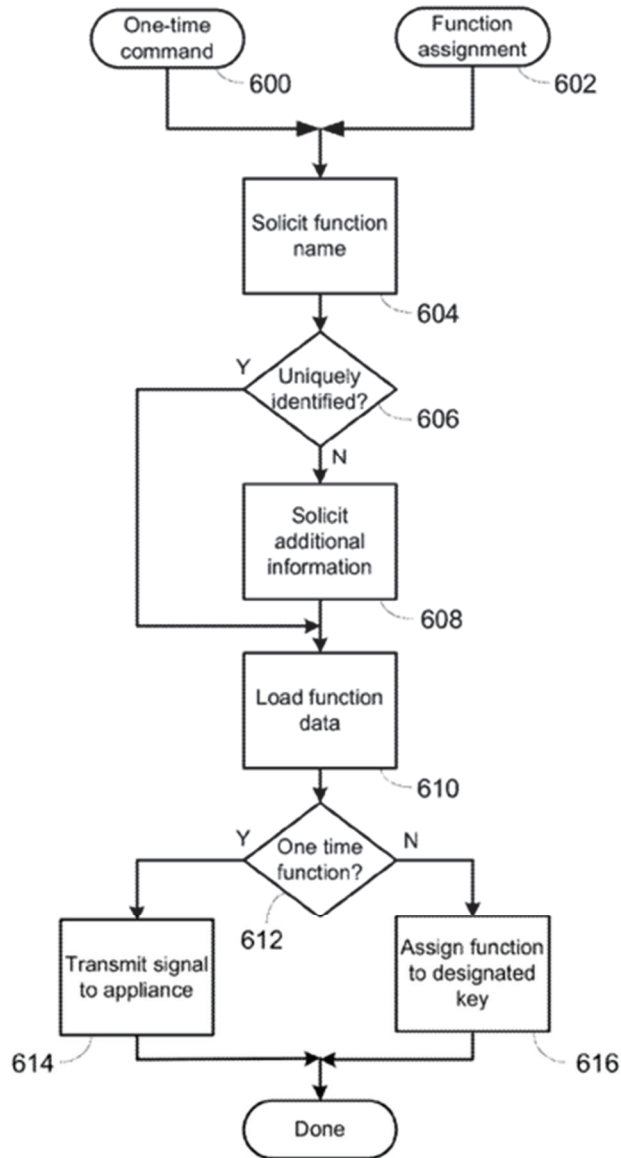


Figure 6 “Appliance command”
Lim, [0035].

Another example of the *second message* is the “*command code data*” downloaded from the *server 500* in block 610 in response to the sending of the captured *second human voice data*. “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 . . . *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, ¶¶127-128.)

The *second message* is sent from the *server* “in response to the sending” of the captured *second human voice data* because both the *speech processing result(s)* and the *command data/codesets* are identified and sent based on *server*’s processing of the *user command*. (EX1005, [0022], [0024]; EX1002, ¶127.)

As discussed above, Lim discloses a *first/controlling device* communicating with a *server device via the WLAN* over the *Internet*. (*Supra* §§IV.A.2.a, IV.A.2.g.) It thus follows that Lim’s *controlling device* similarly receives the recited “*first message*” and “*second message*” from the server *via the WLAN* over the *Internet*. (EX1002, ¶129.)

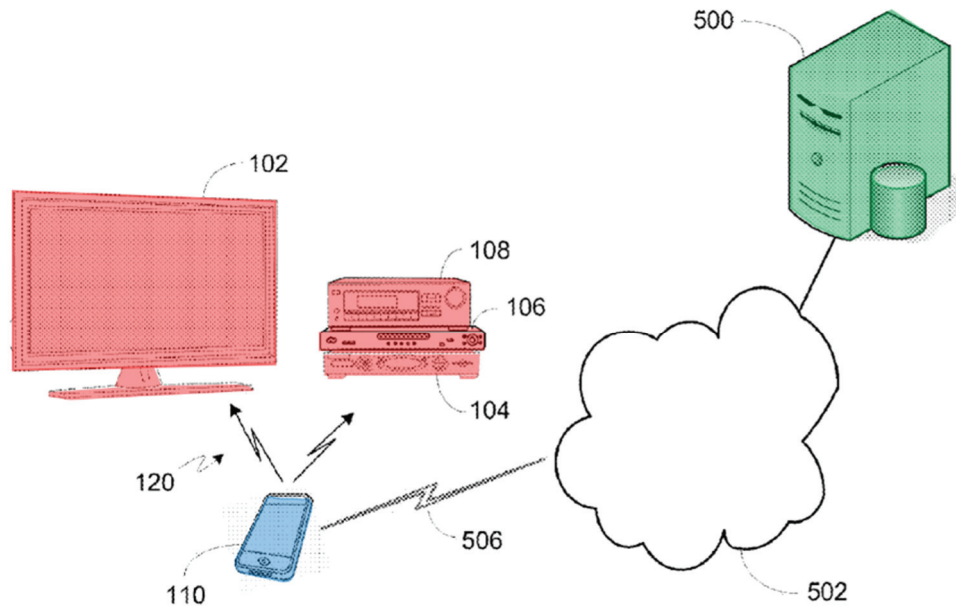


Figure 5b

Alternatively, as discussed, a POSITA would find it obvious that a user command to operate an appliance (*e.g.*, speaking the desired function name in FIG. 6) in Lim can include multiple parts, such as a trigger phrase (“*first human voice data*”) and the instruction (“*second human voice data*”), consistent with common implementations. (EX1002, ¶130.) It follows that the processing of the first human voice data would trigger sending of *corresponding speech processing results* (“*first message*”) to the controlling device to provide audio feedback as it does in FIG. 4. Further, the processing of the second human voice data would trigger sending of *corresponding speech processing results* of the instruction and *command data/codesets* (“second message”) as taught in FIG. 6. (EX1002, ¶¶130-131; *infra* §§IV.A.2.i-IV.A.2.l.)

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

As discussed, Lim’s *first/controlling device* has a *speaker* and a *buzzer*, both of which are sounders. (*Supra* §IV.A.2.d; EX1005, [0019], [0023], FIGs. 2-3; EX1002, ¶132.)

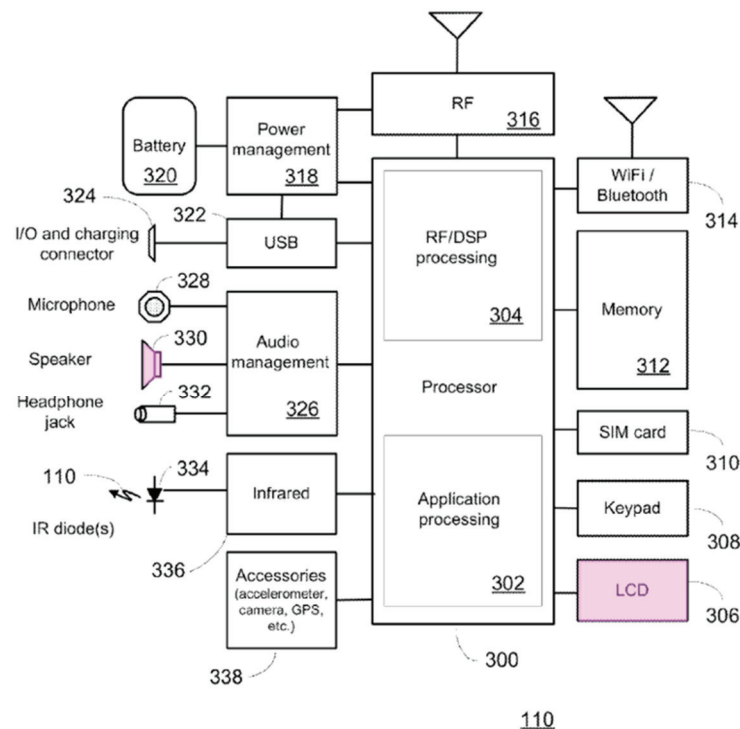
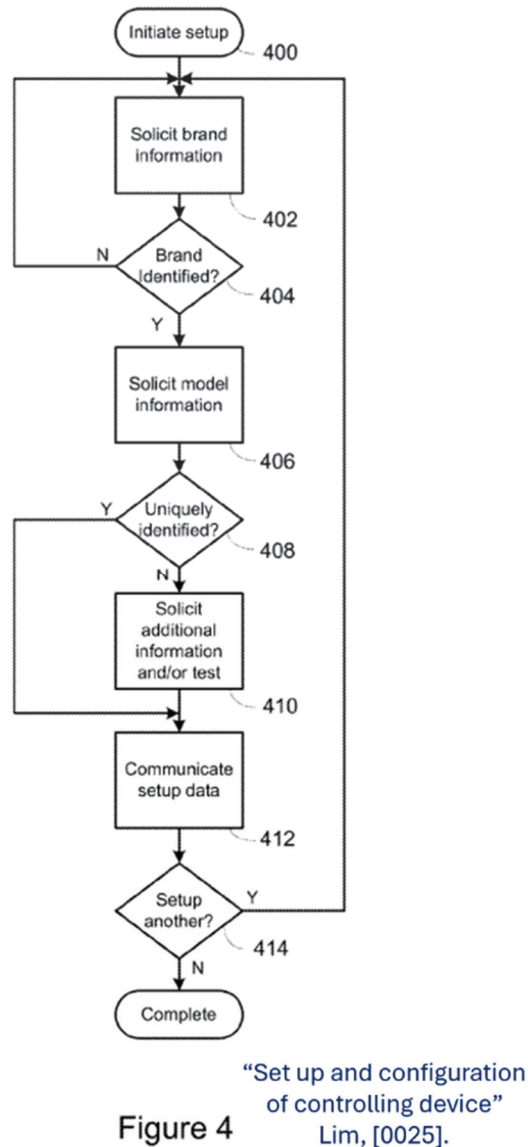


Figure 3

In response to *speech processing result(s)* (“first message”) from the *server* 500, the *first 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*.” (EX1005, [0027].) Lim suggests, and a POSITA would find it obvious, that the operation of sounder is based on *speech processing result(s)*. (EX1002, ¶133.) “[I]f the speech recognition means is unable to uniquely identify the desired brand

[based on *speech processing result(s)*]" in block 404, the controlling device may "*solicit* additional identifying information from the user," "*present*[]" a short list of possible brand identities," etc. (EX1005, [0028], [0036].) "Once a brand has been successfully identified [based on *speech processing result(s)*], at step 406 the user may be *prompted* to provide spoken appliance type and/or model number information." (*Id.*, [0029]; EX1002, ¶¶133-134.) Additional prompts are provided in blocks 410 and 414. (*Id.*; EX1005, [0029]-[0030].)



Additionally or alternatively, Lim’s *buzzer* meets this element. Lim’s *first/controlling device* “provide[s] feedback to the user (e.g., one or more LEDs, illuminable keys, display, *buzzer*, voice synthesizer with associated speaker, and/or the like).” (EX1005, [0019].) Thus, Lim certainly suggests that providing the prompts discussed above would include sounding a buzzer, for example, to get the user’s attention to start listening to the questions. (EX1002, ¶135.)

It would have been obvious that Lim sounds its *speaker* and/or *buzzer* in the *first device* in response to *first message* because the *speech processing result(s)* is needed to determine if, when, and/or which next prompt should be sounded.

(EX1002, ¶¶136-137.)

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

Lim’s “*controlling device 100* is configured to control *various controllable appliances*”, each of which is a *second device*. They 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].)

Lim’s *second devices* have actuators (“first actuator”). For example, Lim’s appliances (e.g., *television 102, DVD players*) have displays, which are actuators. (EX1018, 175; EX1004, 396; EX1025, FIG. 15; EX1002, ¶¶138-139.) The ’721 patent confirms that displays are actuators. (*Id.*; EX1001, 25:33-38, 40:50-41:6, 109:42-47, 110:61-67.) This is further consistent with the Board’s finding during prosecution. (EX1004, 396 (finding that “image display device 51 and air conditioner 52 [] are controlled devices with actuators”); EX1002, ¶140.)

Lim’s appliances (e.g., *television 102, DVD players*) also have speakers, which are actuators, consistent with the ’721 patent. (EX1002, ¶141; e.g.,

EX1018, 122-123; EX1014, 763; EX1001, 105:48-63.)

Lim's appliances (*e.g.*, **drapery, fans, HVAC systems**) have motion actuators. (EX1002, ¶¶142-143.) For example, HVAC has motion actuators, as shown below. (*Id.*; EX1020, 401.) The Board has found that an air conditioner has actuators during prosecution. (*See* EX1004, 396.) Further, "**VCRs**", "**CD players**", and "**DVD players**" have motion actuators (*e.g.*, for spinning the disk). (EX1018, 226-227, 280, Figure 23.7; EX1002, ¶¶144-145.)

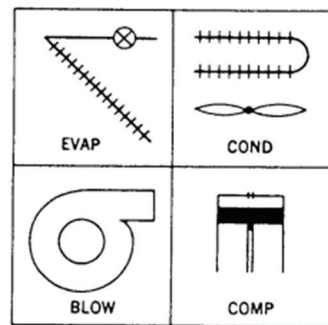


Fig 13-1. Self contained air conditioner

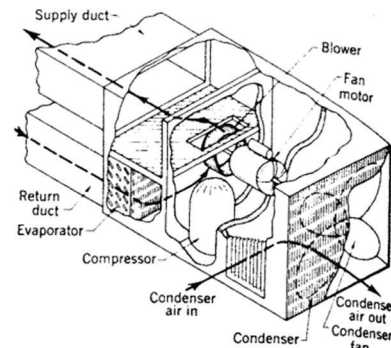


Fig 13-3. Horizontal self contained air conditioned unit

(EX1020, 401.)

Lim's **first/controlling device** sends a **control message** that comprises a control data for controlling a first actuator in the second device. "[T]he **controlling device 100** may be capable of transmitting **commands** to the **appliances** . . . using . . . **networked protocol 120**, to cause the **appliances** to perform operational functions." (EX1005, [0017].) Similarly, an "app may serve to allow use of the **smart phone** to transmit **commands** suitable for operation of **entertainment and other appliances**." (*Id.*, [0024].) As shown in FIGs. 1b and

5b, the *controlling device 110* transmits commands to the *appliances*. (EX1002, ¶146.) It would have been obvious that Lim's *commands* are messages. The *commands* are transmitted using "networked protocol 120" which, as discussed, includes WiFi (*e.g.*, via the WiFi transceiver). (*Supra* §IV.A.2.b.) A POSITA would have found it obvious that data transmitted over WiFi (such as these *commands*) are structured as messages, consistent with the '721 patent. (EX1001, 148:37-42.)

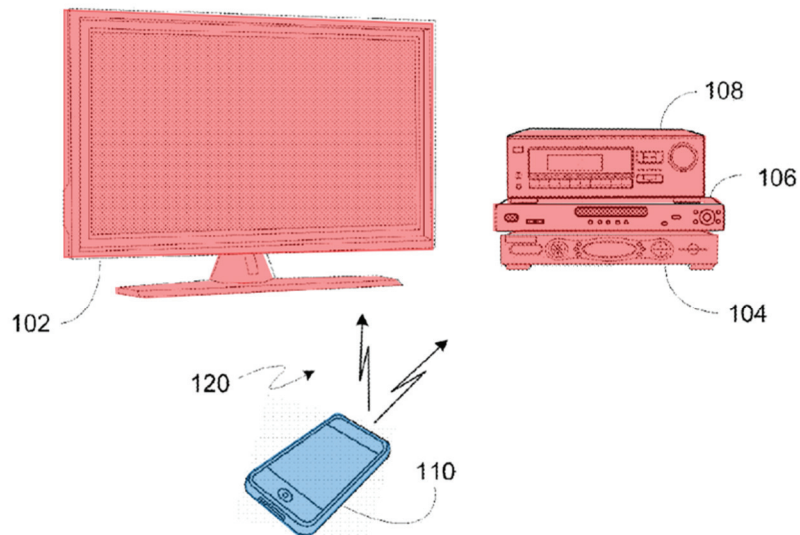


Figure 1b

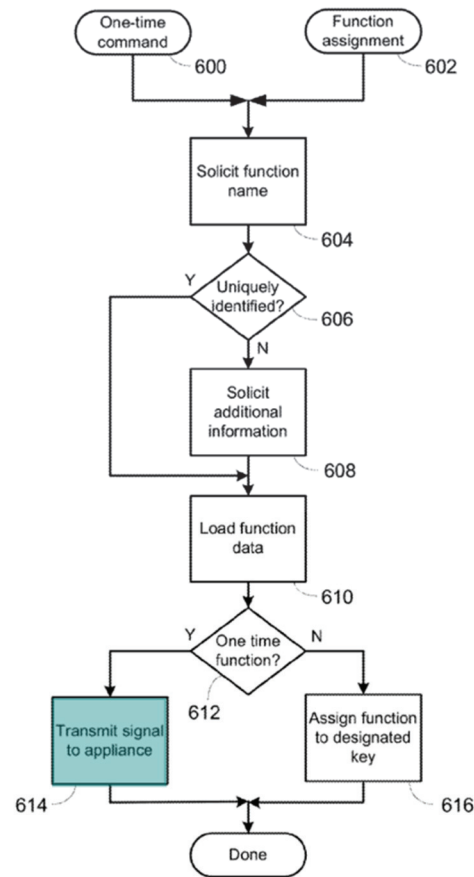


Figure 6 “Appliance command”
Lim, [0035].

Lim’s *control message* is or comprises *control data* for controlling actuators in the *second device*. For example, Lim’s controlling device stores “data which serves to define the necessary control protocols and command values for use in transmitting *command signals* to controllable *appliances* (collectively, the *command data*).” (EX1005, [0020], [0036]; *supra* §IV.A.2.h.) Thus, Lim discloses that its “*command signals*” comprise control data (*e.g.*, Lim’s “command values”). (EX1002, ¶147.)

The actuators in Lim's *second devices* are controlled accordingly. The command can control a *DVD player* to “[p]ower on,” “[f]ast forward,” “[p]ause,” and “[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 controlling actuators such as displays, speakers, and motion actuators. (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 similarly controlled by the control message. (EX1005, [0017]; EX1002, ¶¶148-149.)

Lim's *control message* is sent to the *second device* via the WLAN. “[T]he *controlling device 100* may be capable of transmitting *commands* to the *appliances*, using . . . *networked protocol 120*,” which is also shown in FIG. 5b. (EX1005, [0017].) This “networked protocol 120” includes WiFi because Lim discloses that the first device (*e.g., smart device 110*) includes a *WiFi transceiver/interface*. (EX1005, [0023], [0019].) Further, in 2012, home appliances that are WiFi-enabled were well-known. (EX1012, [0018]; EX1007, 321.) Thus, Lim suggests and a POSITA would find it obvious, that Lim's *controlling device* would send control messages via WiFi to the *controllable appliances 102/104/106/108* connected to the same home WLAN. (*Supra*

§IV.A.2.b; EX1002, ¶¶150-151.)

Alternatively, to the extent that the first and second human voice data are sent together (*e.g.*, as parts of one query), it would have been obvious that Lim, in response to receiving a multi-part user command related to an appliance (*e.g.*, in FIG. 6), would provide visual/audio feedback on the controlling device as it does in the configuration stage (FIG. 4) (responsive to the first message) and cause sending of a control message (responsive to the second message). (EX1002, ¶152.)

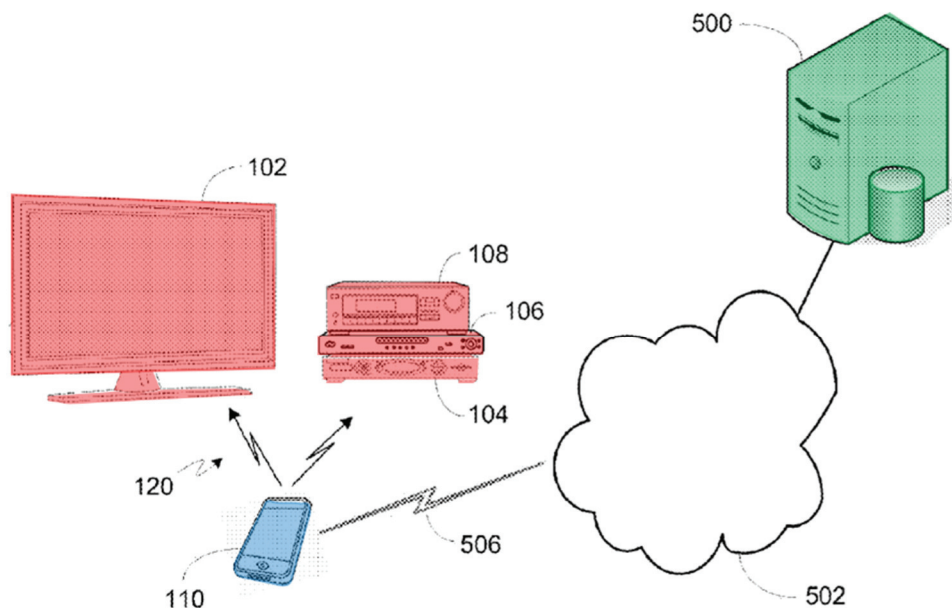


Figure 5b

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

The '721 patent states that a physical phenomenon comprises “temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current.” (EX1001, 16:45-50, claims 5-6.) A

POSITA would understand that Lim's actuator(s) (*e.g.*, speakers, displays, HVAC-related motors, VCR-related motors) are controlled to directly or indirectly affect, change, or produce at least audio, light, motion, temperature, or humidity.

(EX1002, ¶¶153-154.)

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

Lim discloses this element. As discussed, the second message can be the *speech processing result(s)* that allows the *first/controlling device* to determine whether to proceed to block 610. (*Supra* §IV.A.2.h.) Responsive to the *speech processing result(s)*, the *first device* obtains function data (block 610) and sends (block 614) the *command* to the *appliance*. (EX1002, ¶¶155-156.)

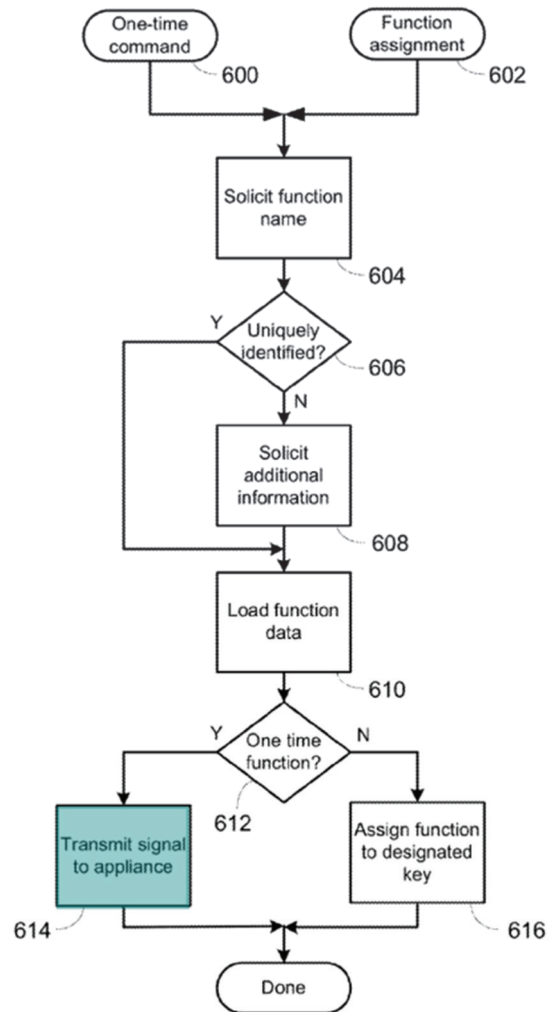


Figure 6 “Appliance command”
Lim, [0035].

As discussed, another example of the *second message* is the “*command code data*”. At block 610, the *first/controlling device* can download the *indicated command data/codesets* from the remote server. (*Supra* §IV.A.2.h; EX1005, [0024], [0030].) Then, “at step 614, the *provisioned data* may be utilized [for the *first/controlling device*] to transmit the desired *command* to the *target appliance*.” (EX1005, [0036], [0017].) Thus, the *command* is sent by the *first/controlling*

device to the *appliance* in response to the received *second message* (e.g., *command code data*). (EX1002, ¶¶157-158.)

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

Lim discloses this element. As discussed, each of Lim’s *first device* and *second device* is WiFi-enabled and connects to the Internet via the home WLAN. (*Supra* §§IV.A.2.a, IV.A.2.j.) A POSITA would have understood that implementing this network involves associating each device with an IP address. (EX1002, ¶159.) Not only was this well-known, Lim confirms this by expressly stating that its *first/controlling device* has an “*IP address associated with* [it]”. (EX1005, [0038]; EX1007, 99.)

A POSITA would have understood that the IP address may be an assigned IP address for the device (e.g., by the router), the router’s IP address, a combination of the two, or a globally unique IP address. (EX1002, ¶160.) “[E]ach device (computer, laptop, and so on) in a home network needs to have an IP address.” (EX1007, 99, 108, 210, 175.)

3. Claims 5/6

Lim’s *first/controlling device* further comprises a second actuator. For example, Lim’s *controlling device 100* “may include . . . a means 202 to provide feedback to the user (e.g., *one or more LEDs*, illuminable keys, *display* [.])”

(EX1005, [0019], [0023].) Lim’s *smart device* includes “an *LCD display 306*” and it also would have been obvious that it would include LEDs. (*Id.*; EX1002, ¶161) The ’721 patent confirms that displays and LEDs are all actuators. (*Supra* §§IV.A.2.j-IV.A.2.k; EX1001, 39:9-20; EX1002, ¶¶162-164.) Displays and LEDs are configured to directly or indirectly affect, change, produce, or create light. *Id.*

4. Claim 9/29/30/33

A POSITA would have understood that these claims recite standard acoustic devices (*e.g.*, microphone, speaker, buzzer), and the patent confirms this understanding. (EX1001, Claim 9, 17:47-54, 105:48-106:62; EX1002, ¶¶165-166, 191-193, 198.) Lim’s controlling device comprises a standard microphone. (EX1005, [0019], [0023].) Lim’s sounder comprises a speaker or a buzzer. (*Supra* §§IV.A.2.d, IV.A.2.i; EX1005, [0019], [0027]-[0029], [0037]; EX1002, ¶¶192-193.) A POSITA would understand that Lim’s microphone and speaker operate as claimed. (EX1014, 757-758, 309, 763, 765; EX1018, 122-123; EX1002, ¶¶165-166, 191-193, 198.)

5. Claims 10/11/12

Lim discloses that the *controlling device 100* “may include, . . . a *ROM memory 204*, a *RAM memory 205*, . . . , a *non-volatile read/write memory 206*[.]” (EX1005, [0019].) The *smart device 110*, “illustrated in FIG. 3 in the form of a smart phone, may include . . . *memory means 312 which may comprise ROM*,

RAM, Flash[.]” (EX1005, [0023]; EX1002, ¶¶167-171.)

As discussed above, Lim’s **controlling device** has an IP address. (*Supra* §IV.A.2.m; EX1005, [0027], [0038], FIG. 5b.) It was common knowledge that a networked device locally stores its IP address. (EX1007, 211, Figure 10-44, 265.) At the time of the invention, it was well-known that Internet Protocol (IP) addresses are either static or dynamic, and are IPv4 addresses. (EX1029, 127; EX1007, 277; EX1002, ¶¶168-171.)

Further, Lim discloses that the first/controlling device is addressable in the Internet and WLAN and communicates via WLAN (*e.g.* a router) with remote server over the Internet. (*Supra* §§IV.A.2.a, IV.A.2.m; EX1005, [0027], [0038], FIG. 5b.) At the time of the invention, it was well-known that networked devices are associated with MAC addresses that are MAC-48, EUI-48 or EUI-64 type. (EX1007, 99, 166; EX1026, 2:50-59; EX1002, ¶169.) Thus, Lim suggests, and it would be obvious in view of POSITA knowledge, that Lim’s networked devices are associated with MAC-48, EUI-48 or EUI-64 type MAC addresses. *Id.*

6. Claims 13/14/15

Lim discloses that the first/controlling device communicates with the **remote server 500** over the **Internet** (*i.e.*, network 502), and **via WiFi**. (*Supra* §§IV.A.2.a, IV.A.2.g; EX1005, [0027], Figure 5b.) Consistent with the ’721 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. (*Supra* §IV.A.2.b; EX1040, 600-601; EX1013, 7:48-51; EX1001, 8:25-32; EX1002, ¶¶171-173.) Further, 802.11b WLANs are ISM Band II frequencies. (EX1040, 22-23; EX1002, ¶¶174-176.)

Lim’s *first/controlling device* can be a wireless smartphone. (*Supra* §IV.A.2.a; EX1005, [0027]; EX1002, ¶177.) Thus, Lim suggests and renders obvious that the first device is configured to communicate over a cellular telephone network (*e.g.*, 3G/GSM). (EX1040, 12-13; EX1053; EX1002, ¶178.)

7. Claims 17/21/22/24

Lim’s “*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.j; EX1005, [0017].)

These disclose the recited “household appliance” in claim 17 and the recited “HVAC system,” “air conditioner or a heater,” “home cinema system,” “CD or DVD player,” and “light fixture” recited in claims 21-22 and 24. (EX1002, ¶¶179-182.)

8. Claim 25

Lim discloses and suggests that the *first/controlling device* (*e.g.*, a smart phone or other portable devices) is operated within a *building* such as a home. (*Supra* §IV.A.2.a; EX1005, [0018]; EX1021, 1:53-58, 2:19-26.)

Lim's *server device* is external to the *home*. A POSITA would understand, and it would be obvious that Lim's "*remote server 500*" is external to the user's *home* because it is "remote." (EX1002, ¶¶183-184; EX1005, [0027], [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. (EX1002, ¶184; EX1008, 8:26-32; EX1006, [0022], Figure 2; EX1009, [0080], [0100]-[0102], [0110]-[0117].)

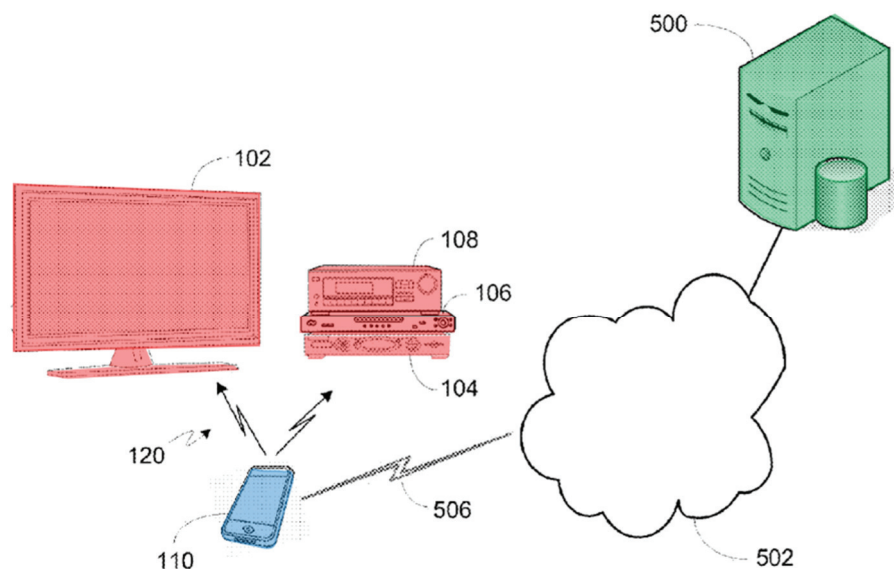


Figure 5b

9. Claims 26/27

Lim discloses or suggests claims 26-27. Lim's *controlling device* includes displays and LEDs. (*Supra* §IV.A.3.) A POSITA would have understood that Lim's displays and LEDs are electric light sources for converting electrical energy

into light that emits a visible light, consistent with the description of the '721 patent. (EX1001, 39:9-20, 40:50-52; EX1018, 174; EX1002, ¶¶185-186.)

10. Claim 28

Claim 28 depends on claim 27. Lim's sounder plays synthesized digital audio content, simulates the voice of a human being, and sounds syllables/words/phrases using human voice. (*Supra* §§IV.A.2.i, IV.A.9; EX1002, ¶¶187-190.) Lim's prompts for voice input "may take the form of . . . a *synthesized spoken message . . .*" (*Id.*; EX1005, [0027].)

11. Claim 31

Lim's *first/controlling device* can be "in the form of a *smart phone*" and "may include . . . *processing means 300 . . . ; WiFi and/or Bluetooth wireless interface(s) 314; . . . ; microphone 328, speaker 330[.]*" (EX1005, [0023]; EX1002, ¶¶194-197.) All the claimed components are housed in a smart phone housing/enclosure as shown in FIG. 1b below.

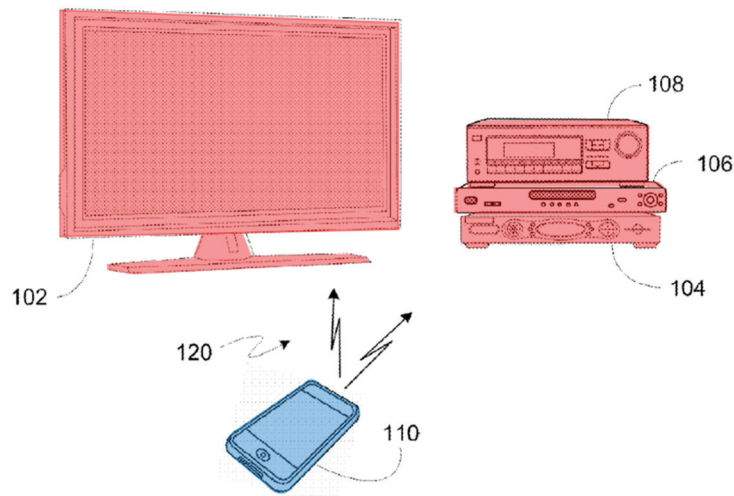


Figure 1b

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 “remotely processing a local audio command to control a local device.” (EX1008, Abstract, Claim 1; EX1002, ¶199.)

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, ¶200.) 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 . . . The processing involves comparing the characteristics . . . of the

audio signal to stored voice information to *determine whether the audio signal voice matches that of an authorized user.*” (*Id.*; EX1008, 5:61-6:6.)

It would have been obvious to combine Wittenberg with Lim. They are analogous art to the ’721 patent. Both relate to remotely processing speech for controlling devices or providing services. (EX1005, [0007], [0027]; EX1008, Claim 1; EX1002, ¶201.)

For example, the Lim-Wittenberg combination would result in a system where voice recognition technique as taught by Wittenberg is incorporated in Lim’s server to determine if the person issuing the voice data is authorized to make the command (*e.g.*, the homeowner). (EX1002, ¶205.) 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 as Lim already performs server-side speech processing, making the addition a straightforward software modification. (EX1002, ¶¶202-206.) For example, Wittenberg expressly teaches “[v]oice recognition processing provides an *additional degree of security.*” (EX1008, 5:61-6:6.) Wittenberg recognizes the benefits of applying its techniques to home appliances. (EX1008, 6:7-13; EX1002, ¶203.)

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

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.” (EX1009, [0007]; EX1002, ¶¶207-208.)

2. Claim 3

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2) Claims 3-4, 11 are obvious in further view of Gruber. ((EX1002, ¶¶209-223.) Gruber discloses that “*a physical device* running virtual assistant 1002 *may have one or more sensors*” such as “an accelerometer, compass, GPS unit, altitude detector, light sensor, thermometer,” and “[s]uch sensors can provide sources of contextual information. Example of such information include . . . the current light level, temperature and other environmental measures.” (*Id.*; EX1009, [0218]-[0225].)

Gruber further discloses that the contextual information provided by sensors can be transmitted to a server (*e.g.*, server 1340) for “interpreting the received user input”. (EX1009, [0366]–[0367], [0179], [0168], [0080], [0100]–[0104], [0110]–[0111], claims 43, 49; EX1002, ¶¶210-212.)

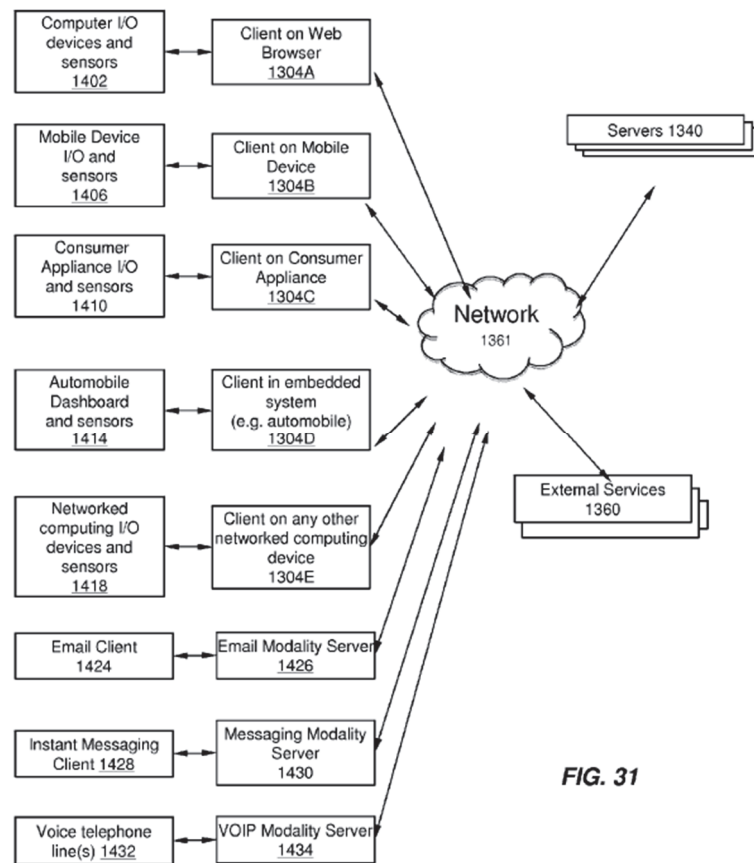


FIG. 31

For example, if a user asks, “what is the weather like?”, the transmitted sensor data can be used to formulate an answer without the user needing to specify the location/environmental parameter. (EX1009, [0354].)

Lim and Gruber are analogous art to the '721 patent. Both relate to processing and responding to human speech commands. (EX1005, [0007], [0027]; EX1009, [0007]; EX1002, ¶213.)

A POSITA would have been motivated to combine Lim with Gruber because Gruber teaches sending contextual information (*e.g.*, temperature, location), along with speech commands, to the server to improve server-side

speech processing. Lim's controlling device already (1) has built-in sensors (*e.g.*, GPS, accelerometer) and (2) sends voice commands to the server for processing. Thus, adding more sensors and sending sensor data to the server, as taught by Gruber, would be a straightforward, predictable improvement with a reasonable expectation of success. (EX1002, ¶¶214-218; EX1009, Abstract; EX1005, [0023].)

For example, the Lim-Gruber combination would result in a system where sensors (as taught by Gruber and/or Lim) are incorporated in Lim's first device to capture sensor data (*e.g.* temperature/location) to be transmitted to Lim's server (*e.g.* **remote server 500**) via the WLAN over the Internet. (EX1002, ¶216.) Lim's server would then use the sensor data to improve its speech processing, as taught by Gruber, and transmit the speech processing result(s) (first message) back to the device. *Id.*

3. Claim 4

Claim 4 depends from claim 3. Gruber discloses various sensors including “an accelerometer, compass, GPS unit, . . . **light sensor, thermometer**, . . . and the like.” (*Supra* §IV.C.2; EX1009, [0225].) The contextual information from these sensors can include “the **current light level, temperature, and other environmental measures**.” (EX1009, [0222].) A POSITA would have understood that photoelectric sensors and thermoelectric sensors are common light sensors and

temperature sensors. (EX1030, 116, 466; EX1002, ¶¶219-220.)

4. Claim 11

Gruber discloses that its devices provide “signatures of connected networks, including Ethernet, Wi-Fi and Bluetooth. *Signatures include MAC addresses of network access points, IP addresses assigned, device identifiers such as Bluetooth names, frequency channels and other properties of wireless networks.*” (EX1009, [0218]-[0225].)

It would have been obvious to a POSITA to look to Gruber for implementation details for Lim’s networked device and addressing schemes (*e.g.*, MAC addresses) with a reasonable expectation of success as Lim’s first/controlling device is addressable in the Internet and WLAN. (*Supra* §§IV.A.2.a, IV.A.2.m, IV.C.2; EX1002, ¶¶221-223.)

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). (EX1039, 1.) The Examiner never found, and Applicant never cited, Benesty during prosecution. Benesty is a textbook titled “Springer Handbook of Speech Processing”. (EX1010, Title.) Benesty provides an overview of well-known speech processing techniques. (EX1010, 1021; EX1002, ¶¶224-227.)

Benesty was published by Springer and was well-known among POSITAs,

who would have looked to it for implementing/improving audio capturing/processing in Lim. (EX1002, ¶228.)

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, ¶¶229-236.) 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.)

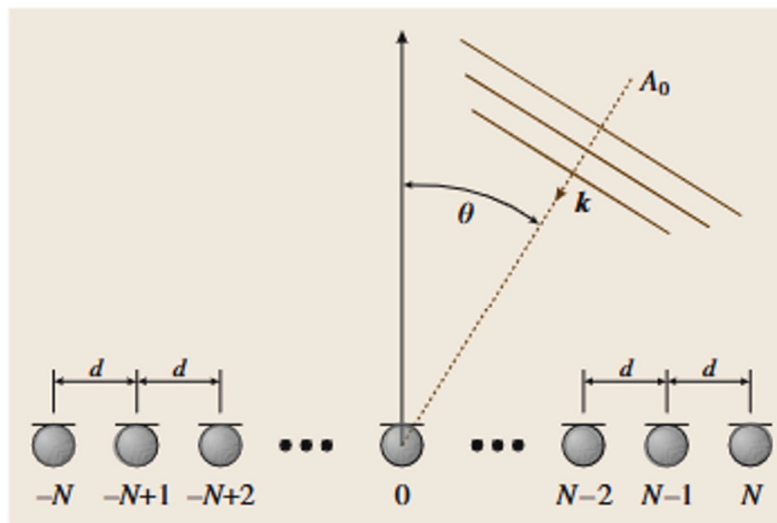


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

Benesty discloses that “microphone array beam scanning and **time delay estimation** (TDE)-based localization methods,” which can involve “**estimate of the direction** of an active speech source.” (EX1010, 1044.) Further, “a set of relative time differences of arrival (TDOAs) for different microphone pairs are

calculated.” (*Id.*; EX1002, ¶¶234-235.)

Lim and Benesty are analogous art to the '721 patent. Both relate to capturing/processing human audio data. (EX1005, [0007]; EX1010, 1021.) A POSITA would have been motivated to combine the two to include Benesty's microphone array techniques to improve audio capturing and processing and with a reasonable expectation of success. (EX1002, ¶¶231-233.) For example, the Lim-Benesty combination would result in a system where multiple microphones as taught by Benesty are incorporated in Lim to capture and process human voice data, yielding the predictable results of improved audio capturing and processing. (EX1002, ¶232.)

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

1. Summary of Rossing

Rossing was published by at least July 23, 2007, and is prior art under §102(a), (b). (EX1041, 1.) The Examiner never found, and Applicant never cited, Rossing during prosecution. Rossing is a textbook titled “Springer Handbook of Acoustics” and was well-known among POSITAs. (EX1014, Title; EX1002, ¶¶237-244.) A POSITA would have looked to Rossing for details and techniques of using microphones/speakers in Lim. *Id.*

2. Claim 9

Rossing discloses that a microphone includes a diaphragm and operates by

sensing movement of the diaphragm. “Common to virtually all microphones ever made is a diaphragm” and “the sound waves impinging on a diaphragm cause movement of the diaphragm[.]” (EX1014, 757; EX1002, ¶¶245-247.)

Rossing also discloses other types of microphones such as “the ribbon microphone and the condenser microphone[.]” piezoelectric microphones, and carbon microphones. (EX1014, 758; EX1002, ¶¶248-250.)

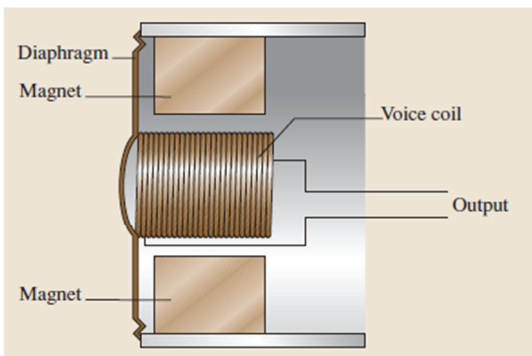


Fig. 18.10 A dynamic microphone uses a coil of wire attached to a diaphragm, suspended in a magnetic field (After [18.24])

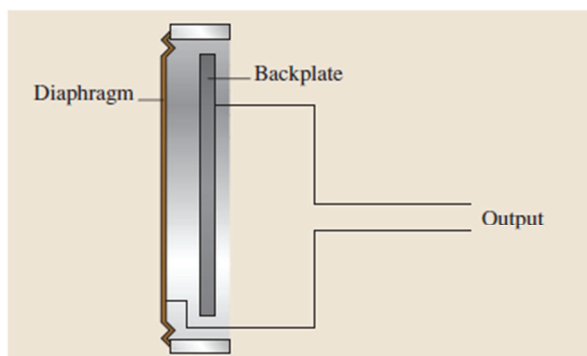


Fig. 18.9 A condenser microphone consists of a very thin diaphragm suspended parallel to a backplate (After [18.24])

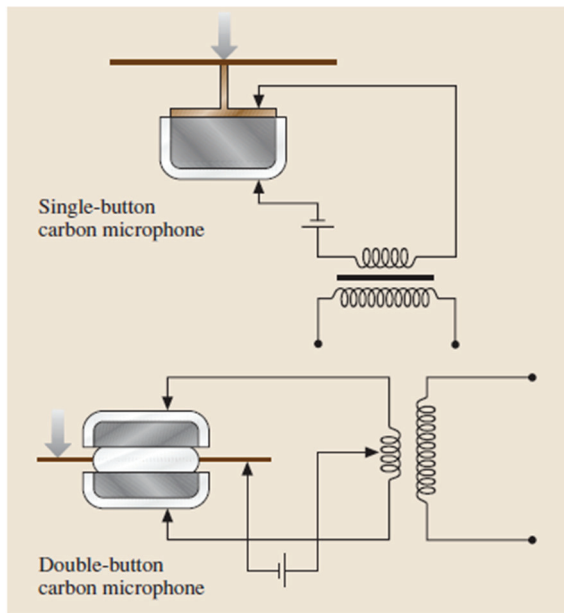


Fig. 18.11 A carbon microphone (After [18.25])

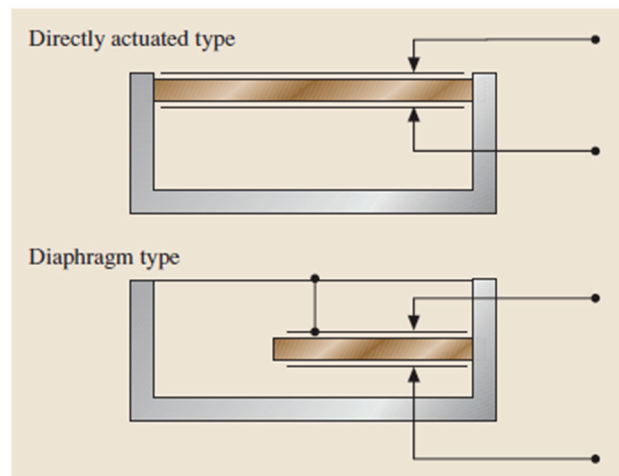


Fig. 18.12 Piezoelectric microphones (After [18.29])

Further, Rossing discloses that “[a] popular modern variant is the *electret condenser microphone*”, which are “*practical for multi-microphone arrays.*” (EX1014, 759.) Rossing also discloses “an ordinary omnidirectional microphone.” (EX1014, 309; EX1002, ¶¶250-251.)

It would have been obvious to a POSITA to apply Rossing, a textbook providing basic information on conventional microphones/speakers, to Lim and with a reasonable expectation of success. (EX1014, 757-758, 763, 765, 903; EX1002, ¶¶252-253.)

3. Claims 29/33

Regarding claim 29, Rossing explicitly discloses “electrostatic loudspeaker” and 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.) It would have been obvious to combine Lim and Rossing, as there was a finite number of types of conventional speakers as taught by Rossing to try applying to Lim. (EX1002, ¶¶254-257.)

Regarding claim 33, given 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, ¶¶258-259.) It would have been obvious to combine Lim and Rossing. (*Id.*; *supra*

§IV.E.2.)

F. Ground 6: 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). (EX1040, 2.) The Examiner never found, and Applicant never cited, Goldsmith during prosecution. Goldsmith, a textbook titled “Wireless Communications,” was published by Cambridge University Press and was well-known among POSITAs, who would have looked to it to implement Lim’s wireless capabilities. (EX1040, Title; EX1002, ¶¶260-265.)

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, ¶¶266-284.) Goldsmith introduces various IEEE802.11 standards, including “IEEE 802.11a wireless LAN standard” and “802.11g.” (EX1040, 16; EX1002, ¶¶268-270.)

Lim and Goldsmith are analogous art to the ’721 patent. Both relate to wireless communications. It would have been obvious for a POSITA to look to Goldsmith, a textbook on conventional wireless networking techniques, to implement the same in Lim, because Lim discloses using WiFi, a form of WLAN and thus suggests a motivation to combine. (EX1040, 8-21; *supra* §§IV.A.2.a, IV.A.2.b, IV.A.2.g; EX1005, [0027]; EX1002, ¶¶271-274.) For example, it would

have been obvious to try combining WLAN over IEEE 802.11a/b/g with Lim without undue effort and with a reasonable expectation of success. (EX1002, ¶274.)

3. Claim 14

Goldsmith expressly shows that 802.11b WLANs are ISM Band II frequencies. (EX1040, 22-23; EX1002, ¶275.)

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 Goldsmith and Lim. (*Supra* §IV.F.2; EX1002, ¶276.)

4. Claim 15

Goldsmith discloses the recited conventional wireless network/cellular communication protocols, for example, that “[t]he *third generation (3G) cellular systems are based on a wideband CDMA standard*”, which is also referred to as “UMTS” (EX1040, 12-13; 599-600.) It further discloses the conventional use of “*GSM*” and “*EDGE.*” (*Id.*; EX1002, ¶¶277-281.)

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.

A POSITA would have been motivated to look to Goldsmith for cellular protocols to implement Lim, as Lim’s device comprises a cellular phone. (*Supra* §§IV.A.2.a, IV.A.2.b, §IV.F.2; EX1005, [0027]; EX1002, ¶¶283-284.) In 2012, home appliances with cellular connectivity were well-known. (EX1031, Claims 14-16, 18; EX1002, ¶282.) Further, wireless technology commonly included cellular connectivity (*e.g.*, 3G, GSM) in 2012. (*Id.*; EX1040, 12-13.)

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

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 relates to “a system and method for providing home automation using a mobile device.” (EX1012, Abstract; EX1002, ¶¶285-288.)

2. Claims 18-20 and 23

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

obvious in further view of Chatterjee. (EX1002, ¶¶289-305.)

Lim discloses controlling various home appliances. (EX1005, [0017].)

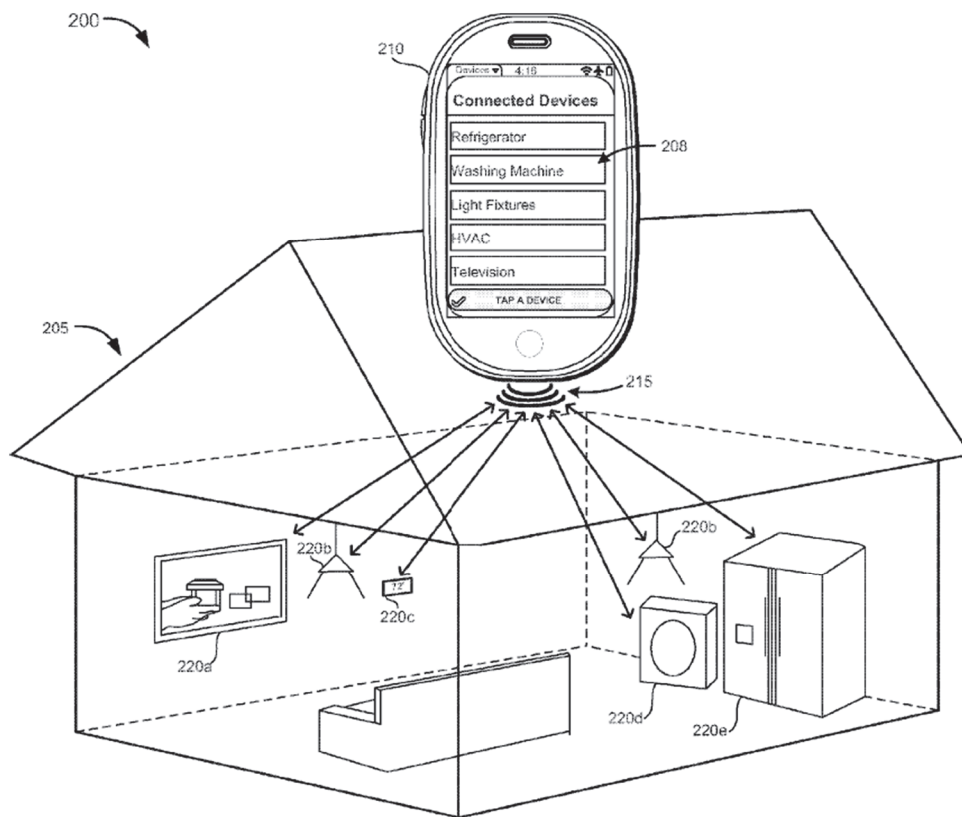
Similarly, Chatterjee describes that a variety of appliances can be controlled,

“such as controlling the temperature of a HVAC or toggling the light system on and off.” (EX1012, [0012]; EX1002, ¶¶290-291.)

Chatterjee discloses that *“home automation systems* provide centralized control for *lighting, HVAC (heating, ventilation and air conditioning), appliances (such as refrigerators and dishwashers)*, and similar systems.”

(EX1012, [0001].) Further, “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, ¶291.)

Further, FIG. 2 depicts a “smart refrigerator 220e”, “HVAC 120c”, and “a smart dish washer or washing machine 220d.” (EX1012, [0014]; EX1002, ¶292.)

**FIG. 2**

Lim and Chatterjee are analogous art to the '721 patent. (EX1001, Claim 1, 135:59-63; EX1005, [0017], [0007]; EX1012, Abstract.) A POSITA would understand Lim suggests other home appliances could be similarly controlled. (EX1002, ¶¶293-295.) A POSITA would have been motivated to combine Lim with Chatterjee and looked to Chatterjee for additional home automation devices that can be controlled. For example, the Lim-Chatterjee combination would result in a system where Lim's first device controls appliances such as oven, refrigerator, HVAC, and washing machine as taught by Chatterjee, thereby expanding the

capabilities of Lim with a reasonable expectation of success. (EX1002, ¶¶294-295.)

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

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.” (EX1013, 1:7-13, 8:8-16; EX1002, ¶¶306-308.)

2. Claims 26/27

Claim 1 is obvious in view of Lim. (*Supra* §IV.A.2.) Claims 26-27 are obvious in further view of Rosenberger. (EX1002, ¶¶309-318.) Rosenberger’s device includes a multi-colored LED status light that visually indicates 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.)

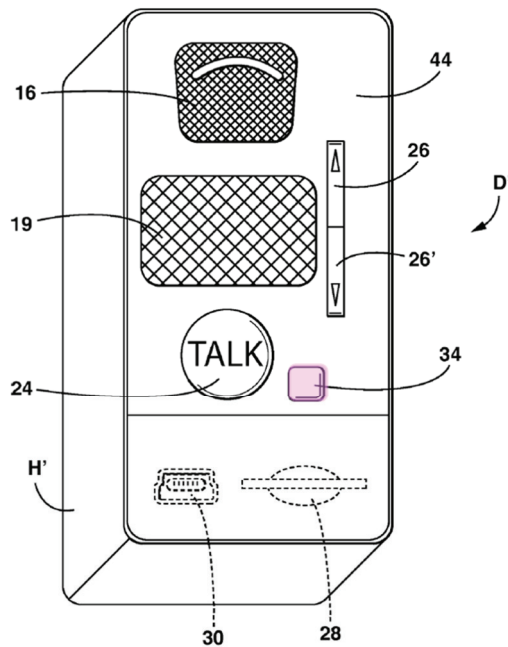


FIG. 2B

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, ¶¶310-311.)

Lim and Rosenberger are analogous art to the ’721 patent. (EX1002, ¶312.) Both relate to processing/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].) A POSITA would have

been motivated to combine Lim with Rosenberger to include Rosenberger's status light in the controlling device, which turns on/off and/or changes color to indicate device states and modes to improve the usability of the device, as taught by Lim/Rosenberger, and with a reasonable expectation of success. (EX1002, ¶¶312-316.)

3. Claim 28

Claim 28 depends on claim 27. Lim-Rosenberger's sounder (*e.g.*, *speaker* and *speaker 330*) plays digital audio content in the form of "*synthesized spoken message*" as recited in claim 28. (*Supra* §§IV.A.2.i, IV.A.2.h, IV.A.10; EX1005, [0027]; EX1002, ¶319.)

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

1. Summary of Williams

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, the system can "provide audio answers to spoken questions." (EX1006, [0048].) Further, the system "may accept spoken commands from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections." (*Id.*, [0051], [0001], [0008]-[0009], [0024]; EX1002, ¶320.)

Referring to FIG. 1, "[a]n individual *interface device 116* may include

a microphone 202 that is configured to detect ambient noise, sounds, and *speech*.”

(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, ¶321.)

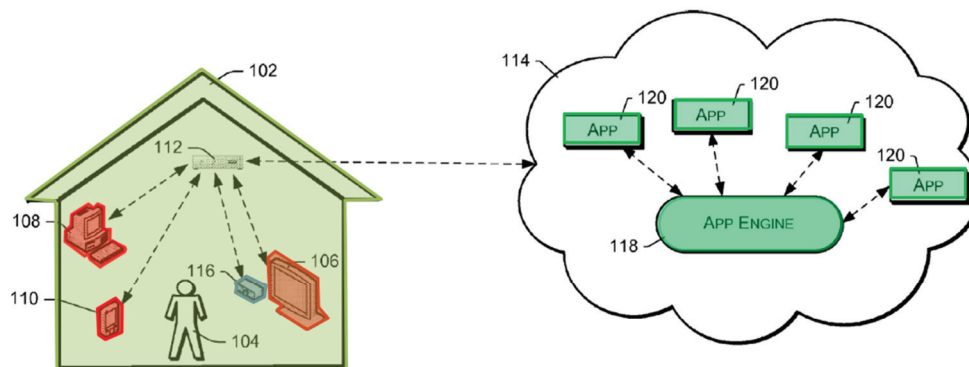


FIG. 1

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, ¶¶322-323.)

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

It would have been obvious for a POSITA to combine Williams and Lim. (EX1002, ¶¶324-328.) Williams and Lim are analogous art to the '721 patent and relate to processing/responding to human speech commands. (EX1006, [0027], [0036]; EX1005, [0007]; EX1002, ¶¶324-325.)

A POSITA would have been motivated to combine Williams with Lim, resulting in a system where Williams' *server 400* sends information for controlling appliances to Williams' *first device* (rather than to Williams' *second devices/appliances*). Williams' *first device* then in turn forwards such information (or command derived from the information) to the *second devices/appliances*. (EX1002, ¶¶326-327; *infra* §§IV.I.3.h, IV.I.3.j-IV.I.3.l.) By adopting Lim's teaching, the server only needs to establish and maintain connection with Williams' *first device*, rather than maintaining numerous individual connections with the *second devices/appliances*. *Id.* 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 security,

reliability, and scalability of the system by having a centralized, local hub (*i.e.*, the first device) that receives control information from the remote server and transmits commands to the second device. (EX1002, ¶328.) It would have been obvious to try combining communication techniques in Lim with Williams, as there is a finite number of identified, predictable solutions for transmitting information from a remote server to home appliances, without undue effort and with a reasonable expectation of success. *Id.* The above discussed benefits would prompt Lim’s application in the same field based on design incentives and other market forces. *Id.*

3. Claim 1

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.a; EX1005, [0019].)

First, Williams discloses a *first device* in a *building*. For example, “[*t*]he *user’s home 102* may be equipped with one or more on-premises *audio monitoring devices 116* . . . having a microphone and a network interface”. (EX1006, [0016]; EX1002, ¶¶329-332, 338-339.)

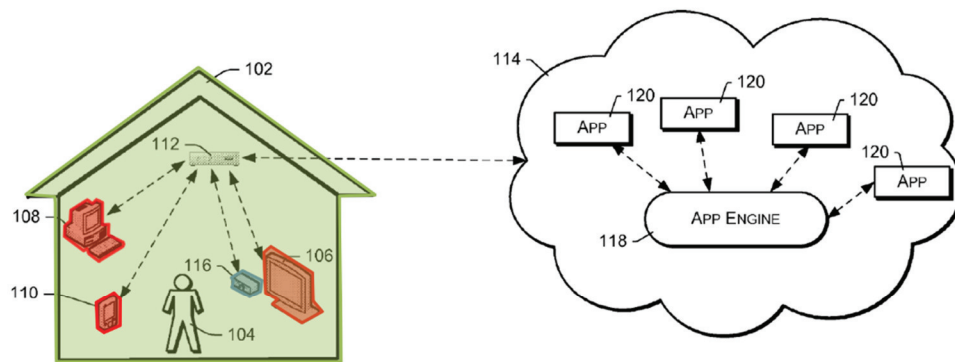


FIG. 1

Williams discloses that the *first device* is for use with *WLAN*. A WLAN network is characterized by a local access point (*e.g.* router) that broadcast wireless signals to networked devices. (EX1007, 11, 100; EX1002, ¶¶333-334.) This is consistent with the '721 patent, which explains that a device can “communicate . . . with the *router* (or gateway) or the control server using the same or different *WANs used by the router*[.]” (EX1001, 28:7-10; EX1002, ¶335.)

A POSITA would have understood that Williams’ in-home wireless network is a WLAN. (EX1007, 11, 100; EX1002, ¶¶333-334.) Referring to FIG. 1, “the in-home devices connect to *a router or access point 112*, which is in turn connected to a public network 114. Connection *between in-home devices and the router 112 may be through wired or wireless technologies.*” (EX1006, [0011]-[0013].) Thus, *interface device 116* and *media devices 106/108/110*

communicate over a wireless network through wireless router/access point 112, which is a WLAN. (EX1002, ¶¶333-334.)

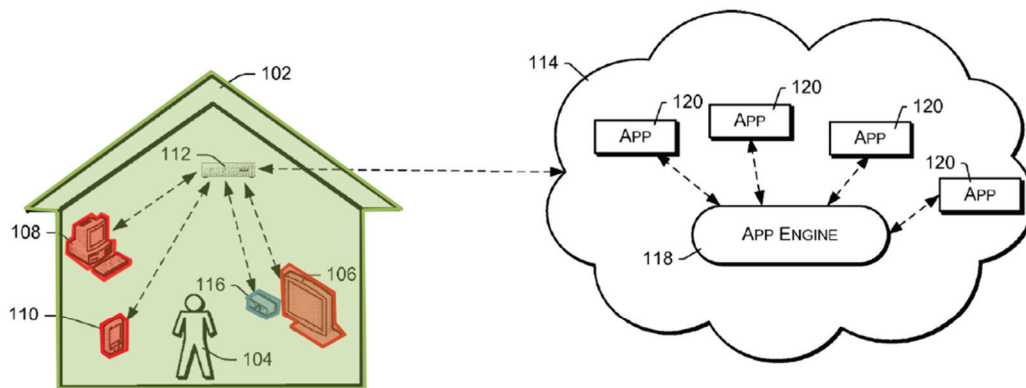


FIG. 1

Further, Williams’ *first device* is wirelessly connected to the cloud network 114, which itself includes “*local-area networks*.” (EX1006, [0013]-[0015]; EX1002, ¶¶336-337.) Again, Williams’ *first device* connects to a WLAN network and thus is for use with a WLAN. (*Id.*; EX1006, [0012].)

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.b; EX1002, ¶¶340-341.) As discussed, Williams discloses all devices in the home (including *audio interface device 116*) communicate with the public network 114 via WLAN (*e.g.*, via a router or access point 112). (*Supra* §IV.I.3.a; EX1006, [0007], [0012],

[0016], [0020], [0023].) A POSITA would have understood that Williams’ *first device* would have the necessary components to allow WLAN connection including a WLAN transceiver. (EX1027, 3:17-20, 3:38-40; EX1028, [0013]-[0015]; EX1002, ¶341.)

Indeed, as shown in FIG. 2 below, which “is a block diagram illustrating additional details regarding the system of FIG. 1,” Williams’ *first device* has a “network interface”, which a POSITA would understand to require a transceiver, consistent with the description of the ’721 patent. (EX1001, 28:62-64; EX1006, [0004], [0023], [0007], [0012], [0016], [0020]; EX1007, 99-100; EX1002, ¶342.)

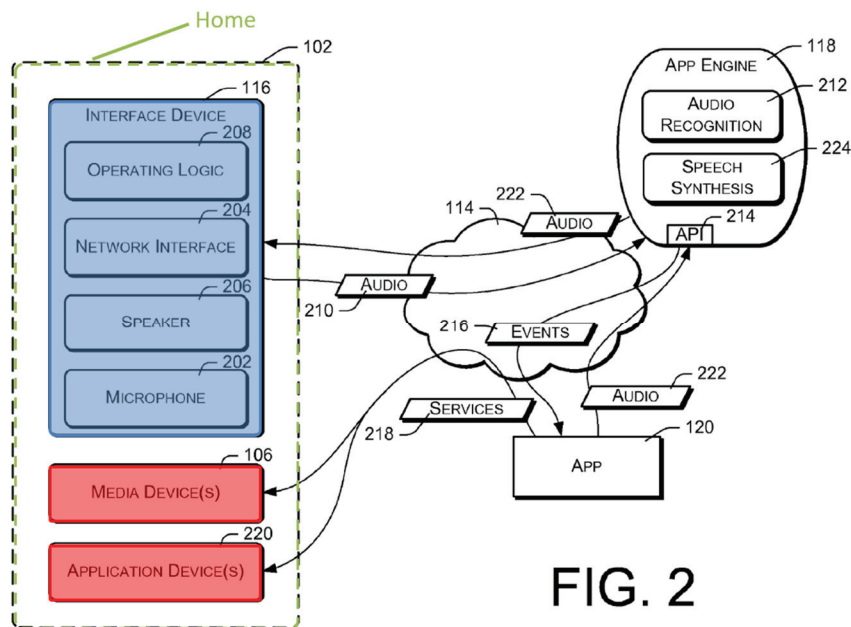


FIG. 2

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

Both Williams and Lim disclose this element. (*Supra* §IV.A.2.c.) 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], [0038]-[0039], Claim 1-3 and 21.) The speech data is captured by the microphone and sent to the cloud for processing. (*Id.*, [0022], [0024]-[0025]; EX1002, ¶¶343-345.)

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

Both Williams and Lim disclose this element. (*Supra* §IV.A.2.d; EX1005, [0019].) Williams discloses that “the *interface device* may also have a *speaker* or *other form of audio output*” and “the *interface device* may also include a *speaker 206*[.]” (EX1006, [0016], [0023], FIG. 2; EX1002, ¶¶346-347.) The speaker is a sounder. (EX1002, ¶348; *e.g.*, EX1018, 122, EX1014, 763; EX1001, 105:61-106:4.)

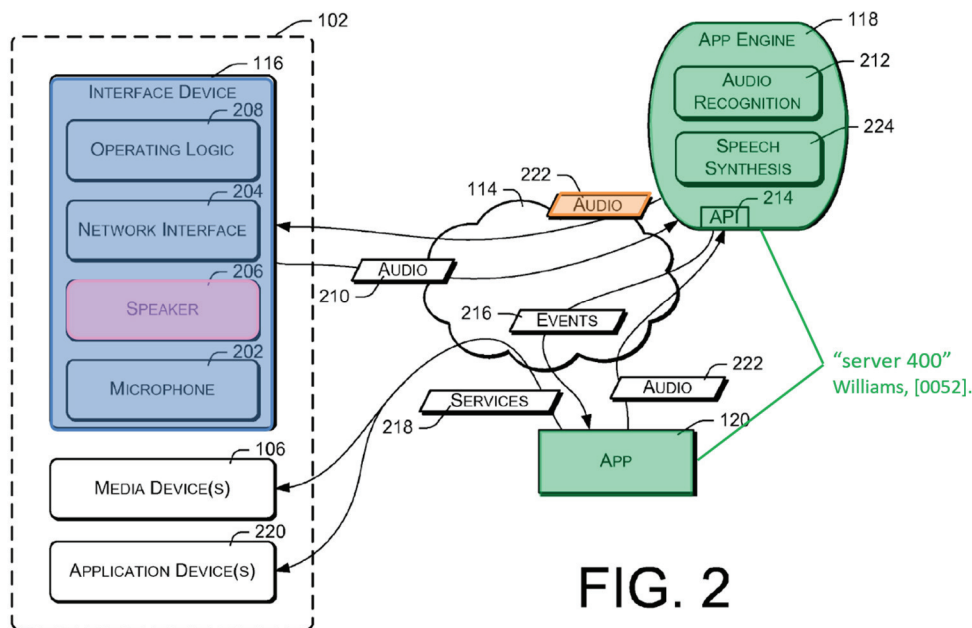


FIG. 2

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.e.)

For example, as shown in FIG. 2, Williams’ *interface device* includes operating logic 208. (EX1006, [0023].) “The operating logic 208 may be implemented as *one or more programmable processing units*, associated memory, and appropriate *instructions* stored in the memory and *executed by the processing unit*.” *Id.* It would have been obvious to a POSITA that execution of the computer program instructions by the processors would cause the first device to perform the recited operations. (*Infra* §§IV.I.3.f-IV.I.3.l; EX1002, ¶¶349-350.)

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

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

In Williams, the voice commands that trigger playing of audio 222 by the interface device are the recited “*first human voice data*”, and the voice commands that cause control of media equipment and home-automation devices by the interface device are the recited “*second human voice data*,” as discussed below. They can be parts of the same command. (EX1002, ¶¶352-353.)

Regarding “*first human voice data*,” Williams discloses many voice commands that cause a sounder (e.g. *speaker 206*) in the *interface device 116* to sound, including: “*What’s the recipe for deviled eggs?*”, an *inquiry “about a flight or other event”*, etc. (EX1006, [0048], [0050]; *infra* §IV.I.3.i; EX1002, ¶*.) Indeed, Williams’ system “provide[s] audio answers to *spoken questions*.” (EX1006, [0048].) It would have been obvious that these voice commands are captured using the microphone. (EX1002, ¶¶354-356.)

Regarding “*second human voice data*,” Williams “allow[s] a user to control various *on-premises devices and equipment, such as media equipment and home-automation devices*.” (EX1006, [0051]; *infra* §§IV.I.3.j-IV.I.3.l; EX1002, ¶359.) 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.” (*Id.*) Such spoken commands are “*second human voice data*.” (EX1002, ¶¶358-359.)

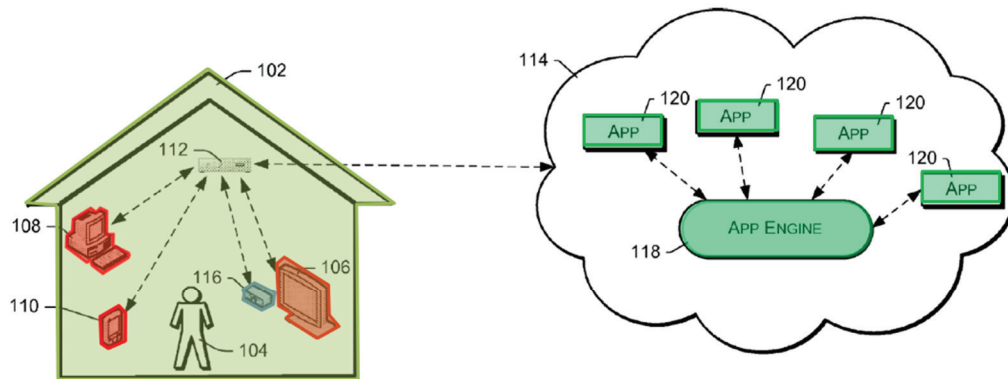


FIG. 1

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, ¶357.) A POSITA would have understood that the “*second human voice data*” (to the extent it is different from the first human voice data) is also captured using the microphone. (EX1002, ¶360.) Thus, two instances of “audio 210” can be mapped to the first and second human voice data, respectively.

To the extent that the first and second human voice data are captured/sent together (*e.g.*, as parts of one query) (which is not conceded), a POSITA would find it obvious that “audio 210” (FIG. 2) can include multiple parts, such as a trigger phrase (“*first human voice data*”) and the instruction (“*second human voice data*”), consistent with common implementations. (EX1023, 2:12-16, 34:63-

67; EX1013, 8:20-32; EX1024, [0045], [0047]; EX1002, ¶¶360.) Thus, the “audio 210” can include both the first and second human voice data and is captured/sent together. *Id.*

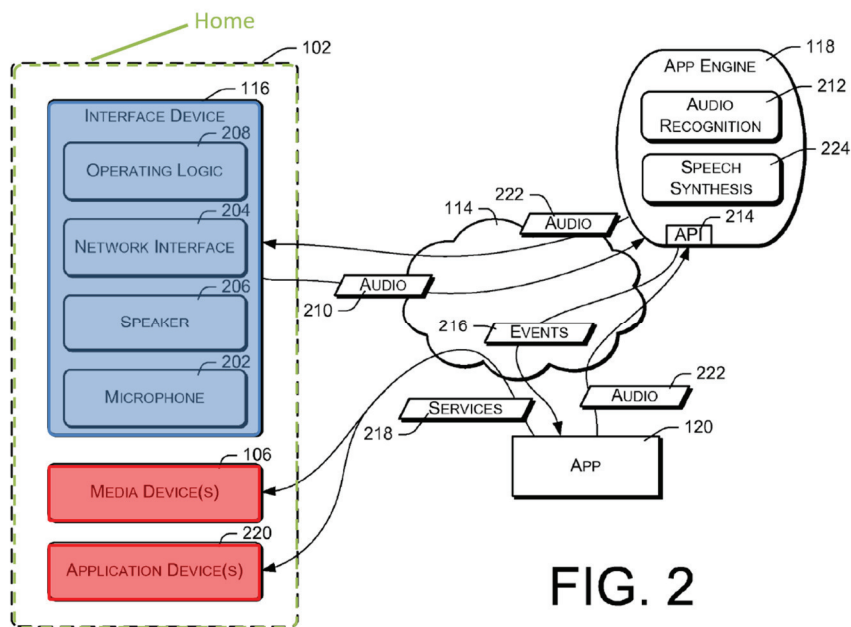


FIG. 2

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.g.) First, Williams discloses a *server device*. 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, ¶¶361-362.)

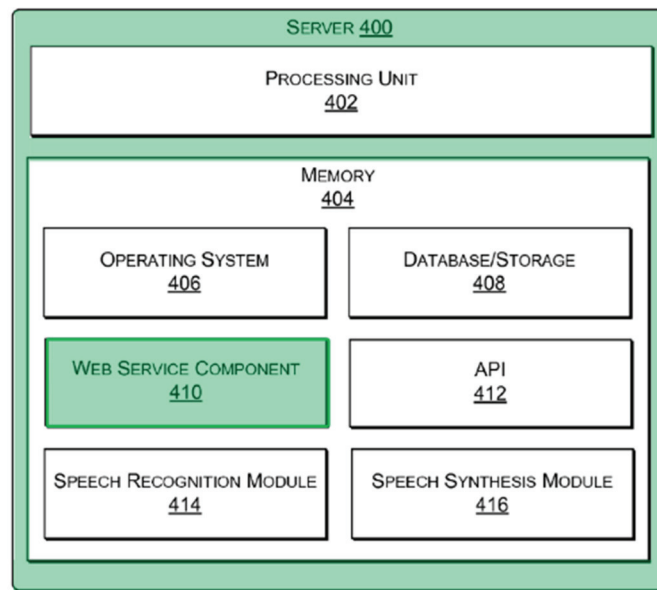


FIG. 4

Given FIG. 1, a POSITA would have understood that *server 400* (running 118 and 120) is an Internet-connected *server device* in the cloud, which is connected to the in-home WLAN (*e.g.* via router 112). (EX1002, ¶¶363-366.) This is consistent with the '721 patent, which explains that “[t]he system may comprise . . . an Internet-connected control server (referred herein also as ‘*cloud server*’ and ‘gateway server’) external to the building . . . and communicatively coupled to the router over the Internet via the external network.” (EX1001, 32:25-42.)

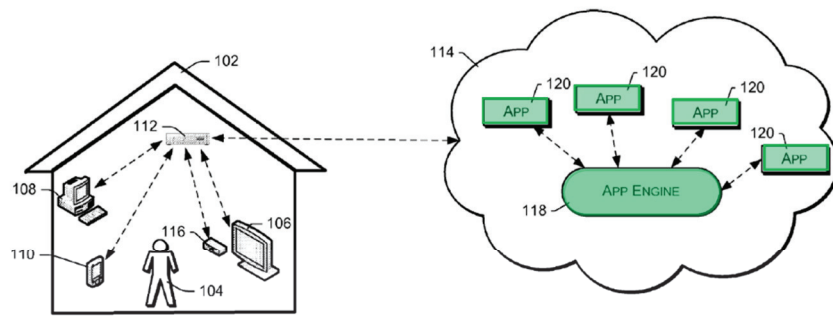


FIG. 1

Williams discloses that the *first device* sends the captured human voice data (including the *first human voice data* and *second human voice data*) to the *server* over the *Internet via the WLAN*. As shown in FIG. 2, *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, ¶¶367-370.) 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,” “audio-based information”) is similarly sent to the *server*. (EX1006, [0028], [0007], [0009], [0051]; EX1002, ¶371.) Alternatively, the “audio 210” sent to *server 400* can include or be both the first and second human voice data as discussed above. (*Id.*; *supra* §IV.A.2.f.)

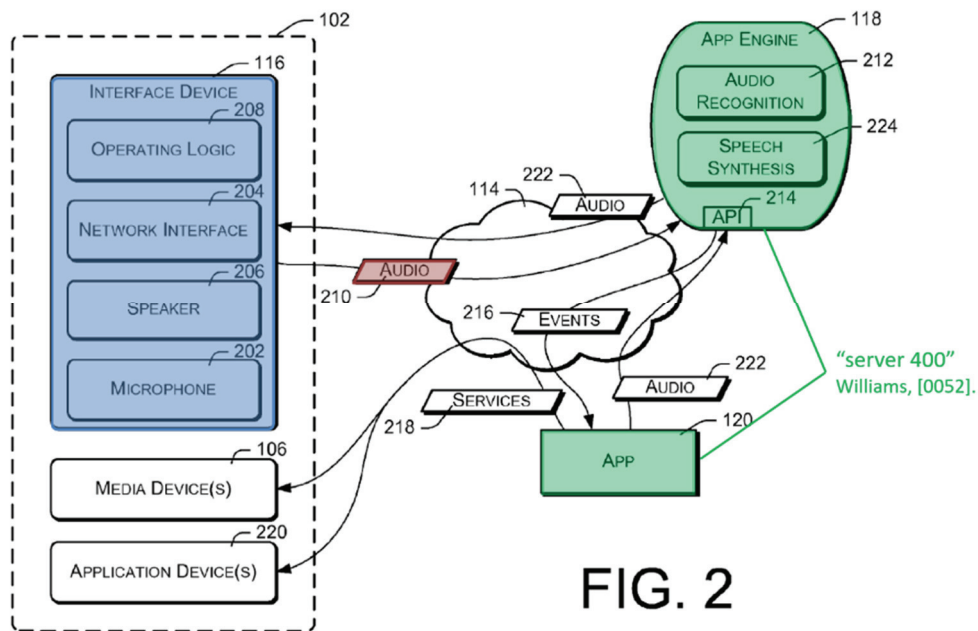


FIG. 2

Referring to FIG. 1, it would have been obvious that the *first human voice data* and *second human voice data* are sent by *interface device 116* to router 112 over the in-home WLAN 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.a; EX1002, ¶¶372-376.)

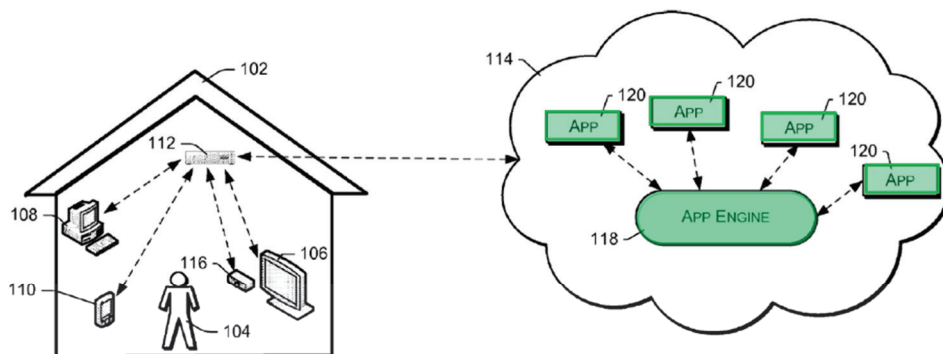
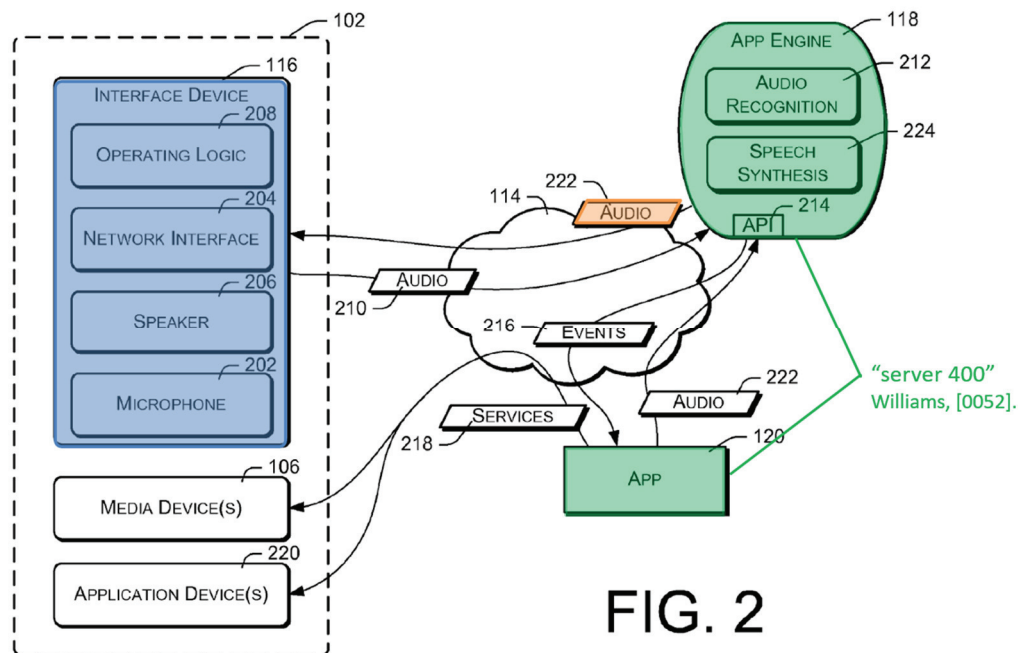


FIG. 1

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.h; EX1002, ¶377.)

For example, Williams’ *first device* receives, from the *server* device via the WLAN over the Internet, *first message* in response to the sending of the captured *first human voice data*. 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 *interface device 116* providing an *audio response* generated by the *cloud server*. (EX1006, [0045], [0048], [0050]; EX1002, ¶¶378-381.)



It would have been obvious that the *first message* is received “in response to the sending of the captured *first . . . human voice data*” as recited because the *first message* is generated as the result of speech processing of the *first human voice data*. (EX1002, ¶379.)

As shown in FIGs. 1 and 2, the *audio 222* is sent by the *cloud server* over the network 114 (i.e., “the Internet”) via the home WLAN (via router 112). (EX1006, FIG. 1, [0015], [0037], [0052]; *supra* §§IV.I.3.a, IV.I.3.g.)

Alternatively, the *wireless network 114* itself includes both WLAN and the Internet and thus discloses *first device* receiving data “via the WLAN over the Internet”. (*Id.*; EX1006, [0015]; EX1002, ¶382-383.)

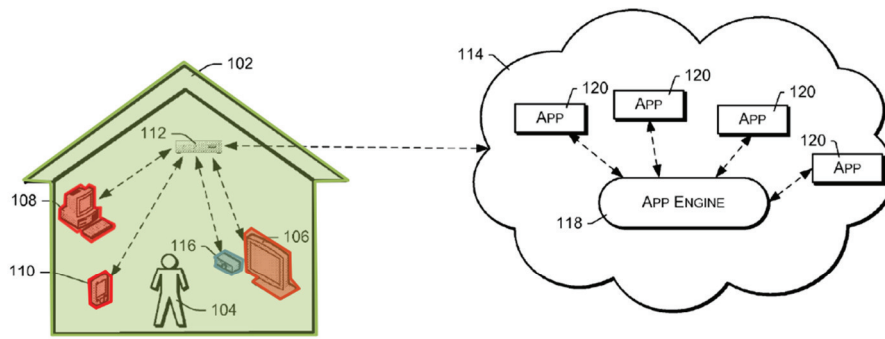


FIG. 1

Williams discloses receiving, from the *server* device via the WLAN over the Internet, a *second message* in response to the sending of the captured *second human voice data*. The *second human voice data* (e.g., “spoken commands from a user regarding on-premises devices”) would be processed in the same way as shown in FIG. 2. (EX1006, [0051]; EX1002, ¶384.) Williams discloses that the *server* “processes” it and as a result “may . . . initiate or customize *services 218*”. (EX1006, [0027], [0036].) A POSITA would have understood that providing *services 218* involves sending information related to the services 218 (e.g., a control message). (E.g., EX1019, 3:19-22; EX1008, Claim 1; EX1012, [0012]; EX1002, ¶¶385-386.) “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; EX1002, ¶384.)

As shown in FIGs. 1 and 2, the second message (*services 218*) is received from *server 400* over the wireless network 114 (e.g. the Internet) and over the in-home WLAN (via router 112 in FIG. 1). (EX1006, [0012]-[0013], [0052]-[0056]; EX1002, ¶387.)

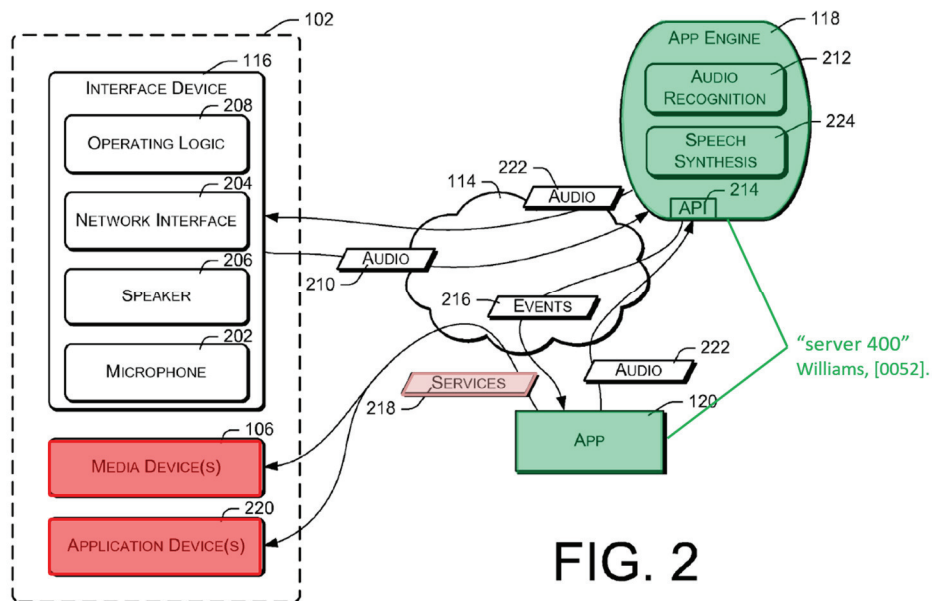


FIG. 2

To the extent that Williams does not disclose that the *first device* receives its *second message*, Lim discloses this element. (*Supra* §IV.A.2.h.) In the Williams-Lim system, the *server 400* would send information (e.g., related to *services 218*) to the *interface device 116* (as supposed to e.g., the *second devices*), and the *interface device 116* would then transmit commands to the *second devices* as

taught by Lim, resulting in more simplified, reliable, and secure system that is easier to scale, troubleshoot, and manage. (EX1002, ¶388; *Supra* §IV.I.2.)

Again, to the extent that the first and second human voice data are captured/processed together (*e.g.*, as parts of one query) (which is not conceded), a POSITA would find it obvious that “audio 210” can include multiple parts, such as a trigger phrase (“*first human voice data*”) and the instruction (“*second human voice data*”), consistent with common implementations. (EX1002, ¶389.) As shown in FIG. 2, processing of audio 210 triggers both first and second messages (*audio 222* and *services 218*). *Id.* For example, the first human voice data triggers sending of “*audio 222*” (“*first message*”) and the second human voice data triggers sending of information related to *services 218* (“*second message*”). (*Id.*; *infra* §§IV.I.3.i-IV.I.3.l.)

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.i.) For example, Williams discloses sounding digital audio content using the *sounder/speaker* in the *first device* in response to the received *first message*. Specifically, the cloud server “provide[s] *audio 222* for eventual transmission to the *interface device 116*, to be played on the *speaker 206*.” (EX1006, [0037], [0045], [0048]; EX1002, ¶¶390-395.) A speaker is a sounder. (*Supra* §§IV.A.2.d, IV.I.3.d; EX1002, ¶393; *e.g.*,

EX1018, 122, EX1014, 763; EX1001, 105:48-52, 105:61-106:4.).

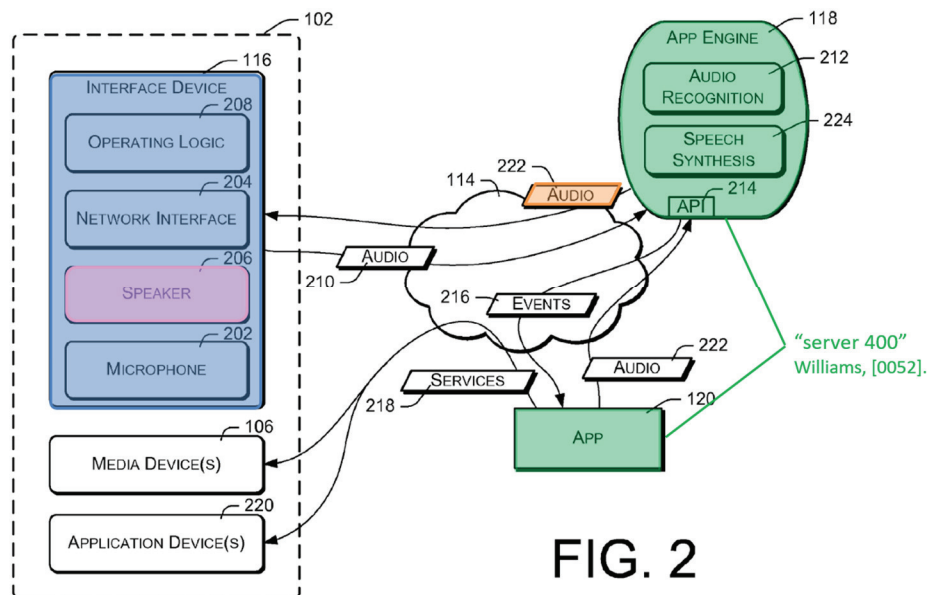


FIG. 2

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

Williams-Lim discloses this. (*Supra* §IV.A.2.j.)

First, Williams discloses *a second device*. Williams discloses “*control[ling]* various *on-premises devices and equipment*, such as *media equipment* and *home-automation devices*.” (EX1006, [0051], [0021], [0036], claim 20.) These are examples of the *second device*. (EX1002, ¶¶396-397.)

Williams’ *second devices* have actuators. For example, Williams’ “*televisions, audio players, video players*” contain displays and/or speakers, which are actuators. (*Supra* §IV.A.2.j; EX1002, ¶¶398-399; EX1006, [0011], [0009]-[0010], [0021], [0036], [0054]; EX1004, 396; EX1025, FIG. 15.) 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/windows, lights, locks, cameras, HVAC systems) have motion actuators. (E.g., EX1020, 401; EX1004, 396; EX1002, ¶¶400-401.)

Williams discloses sending a *control message*. For example, as shown in FIG. 2, Williams depicts sending information related to *services 218* by the application 120 [of the *server*] to the *media devices 106 and/or application devices 220*. (EX1006, [0042].) A POSITA would have understood that the information comprises a *control message* comprising control data to control the *second devices*, consistent with common implementations. (EX1019, 3:19-22, 6:23-27; EX1008, Claim 1; EX1012, [0012]; EX1002, ¶¶402-403.) Further, a POSITA would have understood that “control[ling] various *on-premises devices and equipment*” involves controlling “actuators” in them, for example, to playback a media item. (EX1002, ¶404.)

To the extent that Williams does not disclose that the *control message* is sent from the *first device*, it would have been obvious to a POSITA that the Williams-Lim combination discloses this element. (*Supra* §§IV.I.2; IV.A.2.j; EX1002, ¶405.)

As discussed, Williams' *first device* and *second devices* are connected to a WLAN network via a router. (*Supra* §IV.I.3.a.) A POSITA would have understood that by connecting to the same WLAN network, the *first device* and *second devices* would communicate with each other. (EX1007, 11, 100; EX1002, ¶406.) Thus, in Williams-Lim, Williams' first device would send the control message to a second device over the WLAN.

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.k.) For example, a POSITA would find it obvious that controlling Williams' *speakers/home automation devices/media devices* involves controlling actuators affecting, changing, producing, or creating sound or audio (*e.g.*, via entertainment devices), light, motion, temperature, humidity, etc. (*Supra* §§IV.I.3.j, IV.A.2.j-IV.A.2.k; EX1002, ¶¶407-408.)

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

In Williams-Lim, information related to *services 218* (second message) is sent to the first device. (*Supra* §IV.I.3.h.) Upon receiving the information, the first device then transmits the *control message* to the *appliances*. (*Supra* §IV.I.3.j.) A POSITA would find it obvious that the *control message* is or is derived from the information related to *services 218* (second message). (EX1019,

3:19-22, 6:23-27; EX1008, Claim 1; EX1012, [0012].) Thus, it would have been obvious that the *control message* is sent in response to the received second message. (EX1002, ¶409.)

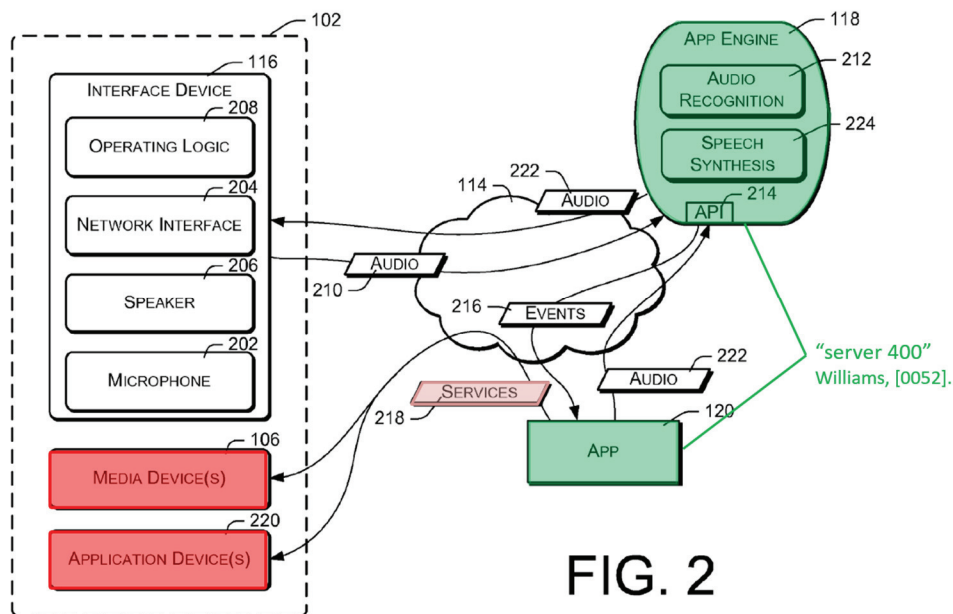


FIG. 2

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

Both Williams and Lim disclose this. (*Supra* §IV.A.2.m.) For example, as discussed, Williams discloses that each of the *first* and *second device* communicates with the remote server in the cloud via WLAN (i.e., over the Internet via a wireless router). (*Supra* §§IV.I.3.a, IV.I.3.j.) A POSITA would have understood that this involves associating each device with an IP address. (EX1007, 99, 108, 175, 210, EX1002, ¶¶410-411.) The IP address may be an assigned IP address for the device (*e.g.*, by the router), the router’s IP address, a

combination of the two, or a globally unique IP address. *Id.*

4. Claims 5/6/26/27

Lim discloses claims 5-6 and 26-27. (*Supra* §§IV.A.3, IV.A.9; EX1002, ¶¶412-417, 439-440.)

5. Claims 9/29/33

A POSITA would have understood that these claims recite standard acoustic devices (e.g., microphone, speaker). (EX1002, ¶¶418, 443, 446; *supra* §IV.A.4.) Both Williams and Lim disclose these claims. *Id.* Williams’ device comprises a standard microphone and speaker. (EX1006, [0016], [0007], [0019], [0037], [0048], claims 1-3 and 21.) A POSITA would understand that Williams’ microphone and speaker operate as claimed. (EX1014, 309, 757-758, 763, 765; EX1018, 122-123; EX1002, ¶¶418, 443, 446.)

6. Claims 10/11/12

Both Williams and Lim disclose these claims. (*Supra* §IV.A.5.) For example, Williams discloses that the *interface device* may include “one or more programmable processing units, *associated memory*[.]” (EX1006, [0023].) In Williams, the first device is addressable with an IP address and that it communicates via WLAN (e.g. a router) with remote server over the Internet. (*Supra* §§IV.I.3.m, IV.I.3.a; EX1006, [0012]–[0013], [0020], Figure 1.) At the time of the invention, it was well-known that IP addresses are either static or

dynamic, and are IPv4 addresses. (*Supra* §IV.A.5.) Further, a POSITA would have understood that the “associated memory” in Williams’ first device locally stores the IP addresses. (EX1002, ¶¶419-425; *e.g.*, EX1007, 211, Figure 10-44, 265.)

Further, Williams suggests, and it would be obvious in view of POSITA knowledge, that Williams’ networked device is associated with an MAC address that is MAC-48, EUI-48 or EUI-64 type as it is addressable in the Internet and WLAN and communicates via WLAN (*e.g.* a router) with a remote server over the Internet. (*Supra* §§IV.A.2.a, IV.A.2.g, IV.A.2.m, IV.I.3.a, IV.I.3.g, IV.I.3.m, IV.A.5; EX1006, [0012]–[0013], [0020], Figure 1; EX1005, [0027], [0038], FIG. 5b; EX1007, 99, 166; EX1026, 2:50-59; EX1002, ¶423.)

7. Claims 13/14/15

Both Williams and Lim disclose or suggest these claims. (*Supra* §§IV.A.6.) For example, Williams discloses usage of “cellular or other wireless data communications technologies,” thus suggesting communications via known 802.11 WLAN (ISM Band II frequencies) and 3G/GSM technologies. (*Id.*; EX1006, Figs. 1-2, [0012]–[0013], [0020], [0038]; EX1040, 8-21, 22-23; EX1002, ¶¶426-430.)

8. Claims 17/21/22/24

Both Williams and Lim disclose these claims. (*Supra* §IV.A.7; EX1005, [0017].) The second device in Williams includes devices such as “*televisions*,

audio players, video players” and “*home-automation devices*”. (*Supra* §IV.I.3.j; EX1006, [0051], [0011]; EX1002, ¶¶431-434.) These are “appliances” in claim 17. *Id.* These also disclose a “home cinema system” recited in claim 24. (EX1002, ¶434.) Further, it would have been obvious to a POSITA that Williams’ “video players” and “audio players” include DVD players and CD players at the time of the invention. (*Id.*; EX1014, 746.)

Lim discloses controlling appliances such as “*fans, HVAC systems, thermostats*”. (*Supra* §IV.A.7; EX1005, [0017].) It would have been obvious to a POSITA that the Williams-Lim combination discloses claims 21/22. (*Supra* §IV.I.2; EX1002, ¶¶432-433.)

9. Claim 25

Both Williams and Lim disclose this claim. (*Supra* §IV.A.8.) For example, as discussed, Williams’ *first device* is operated within a *building* (e.g., *user’s home 102*). (EX1006, [0016].) Further, Williams’ *server 400* is external to the *building* (e.g. the user premise 102) as shown in FIG. 1 below. Williams’ *server 400* “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, ¶¶435-438.)

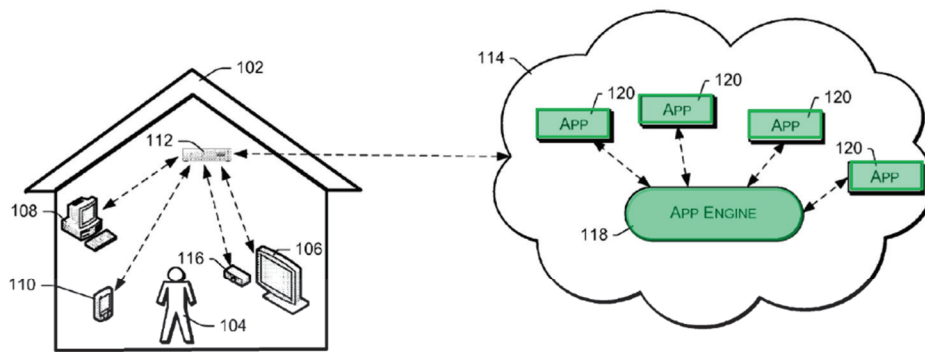


FIG. 1

10. Claim 28

Both Williams and Lim disclose this claim. (*Supra* §IV.A.10.) For example, Williams teaches “playing audio 222 on the *speaker 206*.” (*Supra* §§IV.I.3.d, IV.I.3.i; EX1006, [0037]; EX1002, ¶441.) “[T]he audio [222] may be in the form of text, in which *speech synthesis or text-to-speech conversion* may be performed by the application engine 118.” (EX1006, [0037].) Williams teaches that “[an] application may respond ... by using the *speech synthesis functionality* of the system to *query the user regarding movie selections and times*”. (EX1002, ¶442; EX1006, [0037], [0045]-[0046] and [0050].)

11. Claim 31

Both Williams and Lim disclose claim 31. (*Supra* §IV.A.11.) For example, as shown in FIGs. 1 and 2 below, Williams’ *interface device 116* houses the WLAN transceiver (*e.g., network interface 204*), sounder (*e.g., speaker 206*),

microphone (e.g., *microphone 202*), and processors (e.g., *operating logic 208* implemented as one or more programmable processing units) within one enclosure. (EX1006, [0023], FIG. 2; EX1002, ¶¶444-445.)

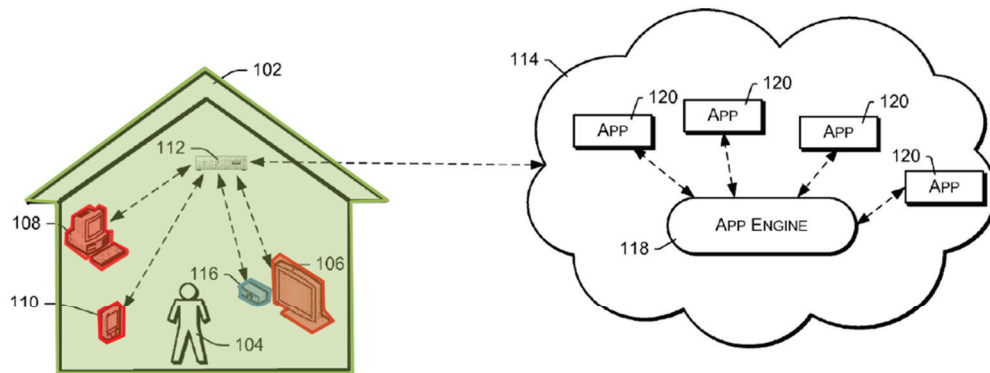


FIG. 1

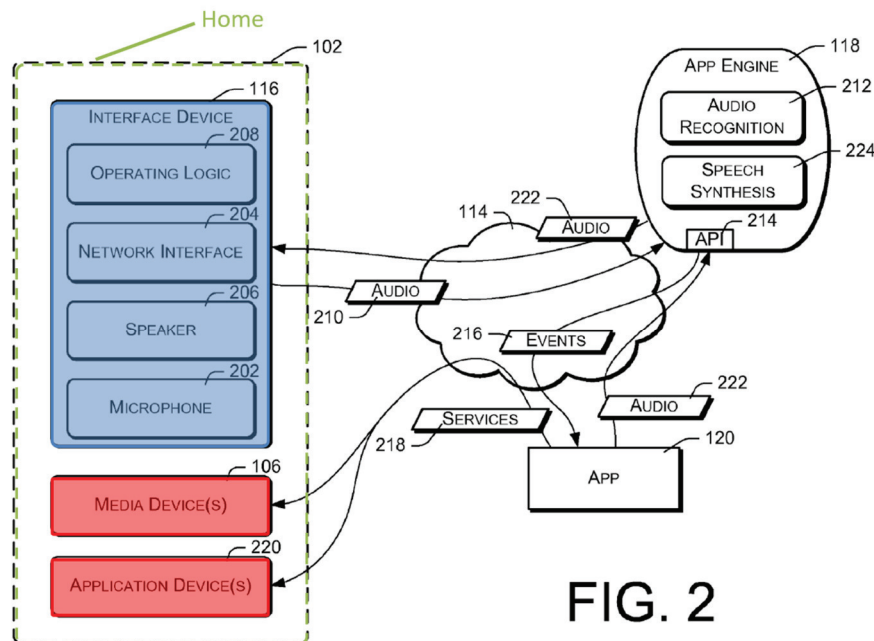


FIG. 2

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

Wittenberg discloses claim 2. (*Supra* §IV.B.2; EX1002, ¶¶447-448.)

It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Wittenberg. (EX1002, ¶¶449-454.) For example, all are analogous art to the '721 patent and relate to remotely processing audio information. (EX1006, [0036]; EX1005, [0007]; EX1008, Claim 1; EX1002, ¶449.)

The Williams-Lim-Wittenberg combination would result in a system where voice recognition technique as taught by Wittenberg is incorporated in Williams-Lim's server to determine if the person issuing the voice data is authorized to make the command. A POSITA would have been motivated to make the combination with a reasonable expectation of success as Lim already performs server-side speech recognition, making the addition a straightforward software modification. Combining them as such would yield predictable results of enhancing security of the system. (EX1002, ¶¶453-454.)

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

1. Claims 3-4

Gruber discloses claims 3-4. (*Supra* §§IV.C.2-IV.C.3; EX1002, ¶¶455-464.)

It would have been obvious to a POSITA to combine the teachings of Williams-

Lim and Gruber. (EX1002, ¶¶457-462.) For example, Williams, Lim, and Gruber are analogous art to the '721 patent and all relate to processing and responding to speech commands. (EX1006, [0027], [0036]; EX1005, [0007]; EX1009, [0007]; EX1002, ¶457.)

The Williams-Lim-Gruber combination would result in a system where sensors as taught by Lim/Gruber are incorporated in Williams-Lim's first device to capture sensor data to be transmitted to Williams-Lim's server to improve audio processing as taught by Gruber. (EX1006, [0013]; EX1009, [0013]; EX1002, ¶460.) For example, combining them as such would yield predictable results of improving audio processing as the sensor-derived information can be used to "interpret received user data." (EX1002, ¶461.)

2. Claim 11

Gruber discloses claim 11. (*Supra* §IV.C.4.) It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Gruber and look to Gruber for implementation details for Williams-Lim's network capabilities as both Williams and Lim's devices are network enabled and addressable in the Internet and WLAN. (*Supra* §§IV.A.2.a, IV.A.2.g, IV.A.2.m, IV.I.3.a, IV.I.3.g, IV.I.3.m, IV.I.6, IV.K.1; EX1002, ¶¶465-468.)

L. Ground 12: Williams, Lim, and Rosenberger Render Obvious Claims 5-6, 26-27, 28

1. Claims 5/6/26/27

Lim and Rosenberger each discloses claims 5-6 and 26-27. (*Supra* §IV.H.2; EX1002, ¶¶469-483.) Rosenberger’s device comprises “a *multi-colored LED status light 34*.” (EX1013, 8:8-16.) The ’721 patent confirms that LEDs are actuators affecting light. (EX1001, 39:9-20; EX1002, ¶477.)

It would have been obvious to combine the teachings of Rosenberger with Williams-Lim. (EX1002, ¶¶472-476.) They are analogous art to the ’721 patent and 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 include Rosenberger/Lim’s LED light indicator (*e.g.*, an additional actuator) in Williams-Lim’s first device, which can be operated (*e.g.*, turns on/off and/or changes color) to indicate device states and modes in response to information received from the server, to improve the usability of the device, as taught by Lim/Rosenberger, and with a reasonable expectation of success. (EX1002, ¶¶474-475.)

2. Claim 28

Williams-Lim-Rosenberger’s sounder plays digital audio content as recited in claim 28. (*Supra* §IV.H.3.) Williams teaches playing “audio 222,” which may be “speech synthesi[zed]” “on the *speaker 206*.” (*Supra* §§IV.I.3.d, IV.I.3.i; EX1006, [0037]; EX1002, ¶¶484-485.) Further, Williams-Lim-Rosenberger’s sounder “us[es] the *speech synthesis functionality* of the system to *query the user*

regarding movie selections and time.” (*Id.*, EX1006, [0037], [0045]-[0046] and [0050].) Thus, Williams plays synthesized digital audio content, simulates the voice of a human being, and sounds syllables/words/phrases using human voice. (EX1002, ¶486.)

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

Benesty discloses claims 7-8. (*Supra* §IV.D.2; EX1002, ¶¶487-493.) It would have been obvious to a POSITA to combine Benesty with Williams-Lim. (EX1002, ¶¶489-491.) All are analogous art and relate to providing hands-free audio communication. (EX1006, [0003], Figure 1; EX1005, [0007]; EX1010, 1021.) Williams-Lim-Benesty would result in a system where multiple microphones as taught by Benesty are incorporated in Williams-Lim’s client device, yielding the predictable results of improving audio capturing. (EX1002, ¶490.)

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

Rossing discloses these claims. (*Supra* §§IV.E.2-IV.E.3; EX1002, ¶¶494-507.) Williams and Lim both disclose use of such a *microphone* and *speaker* in the *first device*, consistent with the ’721 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.) Thus, a POSITA would have been

motivated to look to Rossing, a textbook on acoustic devices, for techniques for implementing Williams-Lim. (EX1002, ¶¶496-507.)

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

Goldsmith discloses claim 13-15. (*Supra* §§IV.F.2-IV.F.4; EX1002, ¶¶508-518.) It would have been obvious for a POSITA to look to Goldsmith for wireless networking techniques to apply to Williams-Lim, as Williams discloses using “cellular or other wireless data communications technologies”, and Lim discloses usage of “local network connection” such as WiFi. (EX1002, ¶¶509-518; EX1006, FIGs. 1-2, [0012]-[0013], [0020], [0038]; EX1005, [0027], Figure 5b; EX1040, 599-600; EX1013, 7:48-51.) For example, it would have been obvious to try combining Goldsmith’s wireless techniques with Williams-Lim without undue effort and with a reasonable expectation of success. (EX1002, ¶512.)

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

Chatterjee discloses claims 18-20 and 23. (*Supra* §IV.G.2; EX1002, ¶¶519-529.) It would have been obvious to a POSITA to combine Williams-Lim and Chatterjee. (EX1002, ¶¶521-523.) They are analogous art to the ’721 patent and all relate to controlling home appliances. (*Supra* §IV.G.2; EX1006, [0051], [0012], [0007]; EX1005, [0007], [0017]; EX1012, Abstract; EX1001, Claim 1, 135:59-63.) Combining them such that Williams-Lim can be used to control

Chatterjee's home devices would improve/expand the capabilities of Williams-Lim in the same way. (EX1002, ¶¶522-523.)

V. NOTICES AND STATEMENTS

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

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

- *Smart Speaker LLC v. Amazon.com Services LLC*, No. 2-25-cv-00707 (E.D. Tex.)
- IPR2026-00145, *Amazon.com Services LLC v. Smart Speaker LLC*, Patent No. 11,128,710
- IPR2026-00146, *Amazon.com Services LLC v. Smart Speaker LLC*, Patent No. 12,316,706
- IPR2026-00147, *Amazon.com Services LLC v. Smart Speaker LLC*, Patent No. 12,010,174
- IPR2026-00148 *Amazon.com Services LLC v. Smart Speaker LLC*, Patent No. 11,190,590

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

Lead Counsel	First Backup Counsel
Alex S. Yap Registration No.: 60609 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Tel: (213) 892-5200 Fax: (213) 892-5454 E-Mail: AYap@mofo.com	Mehran Arjomand Registration No.: 48231 MORRISON & FOERSTER LLP 707 Wilshire Blvd., Suite 6000 Los Angeles, CA 90017 Tel: (213) 892-5200 Fax: (213) 892-5454 E-Mail: MArjomand@mofo.com
	Second Backup Counsel
	Yunan Yuan Registration No.: 73637 MORRISON & FOERSTER LLP 755 Page Mill Road Palo Alto, CA 94304-1018 Tel: (650) 813-5600 Fax: (650) 494-0792 E-Mail: AYuan@mofo.com

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

Pursuant to 37 C.F.R. § 42.104(a), Petitioner certifies that the '721 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: February 27, 2026

Respectfully submitted,

By /s/ Alex S. Yap
Alex S. Yap
Registration No.: 60,609
MORRISON & FOERSTER LLP
707 Wilshire Blvd. Suite 6000
Los Angeles, CA 90017
Tel: (213) 892-5200
Fax: (213) 892-5454
E-Mail: AYap@mofo.com

Counsel for Petitioner

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

I hereby certify that this Petition for *Inter Partes* Review has 13,973 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: February 27, 2026

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 February 27, 2026, and served on February 28, 2026, by Priority Mail Express®, on the Patent Owner at the correspondence address indicated for U.S. Patent No. 12,401,721.

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

Dated: February 27, 2026

/s/ Alex S. Yap
Alex S. Yap