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. 12,401,720

Inter Partes Review No. IPR2026-00280

DECLARATION OF HENRY H. HOUH

Amazon EX1002

TABLE OF CONTENTS

Page

I. INTRODUCTION 1
II. EXPERIENCE AND QUALIFICATIONS 1
III. MATERIALS CONSIDERED 7
IV. OVERVIEW OF THE '720 PATENT 7
A. Filing Date and Priority Dates 7
B. Specification 7
C. Claims 10
D. Prosecution History 14
V. RELEVANT LEGAL STANDARDS AND DEFINITIONS 17
A. Obviousness 17
B. Prior Art 20
VI. LEVEL OF ORDINARY SKILL IN THE ART FOR THE '720 PATENT 20
VII. CLAIM CONSTRUCTION 21
VIII. PRIOR ART 22
IX. ANALYSIS AND OPINIONS 24
A. Ground 1: Lindahl and Williams Render Obvious Claims 1, 2, 4-8, 11-26, and 29 24
B. Ground 2: Lindahl, Williams, and iPhone User Guide Render Obvious Claims 1-2, 23-26, 29 111
C. Ground 3: Lindahl, Williams, and Doherty Render Obvious Claims 4-5 133
D. Ground 4: Lindahl, Williams, and Rossing Render Obvious Claims 7-8,11-12 141
E. Ground 5: Lindahl, Williams, and Benesty Render Obvious Claims 9-10 154
F. Ground 6: Lindahl, Williams, and Goldsmith Render Obvious Claims 20-21 162

G. Ground 7: Lindahl, Williams, and Slaby Render Obvious Claims 14, 23, 28-29. 169

H. Ground 8: Williams, Lim, and Benesty Render Obvious Claims 1, 4-26, 29……...182

I. Ground 9: Williams, Lim, Benesty, and iPhone User Guide Render Obvious Claim 2 281

J. Ground 10: Williams, Lim, Benesty, and Doherty Render Obvious Claims 4-5……. 285

K. Ground 11: Williams, Lim, Benesty, and Rossing Render Obvious Claims 7-8, 11-12 290

L. Ground 12: Williams, Lim, Benesty, and Goldsmith Render Obvious Claims 20-21 297

M. Ground 13: Williams, Lim, Benesty, and Slaby Render Obvious Claims 14, 23, 28-29 301

N. Ground 14: Lindahl and Evermann Render Obvious Claims 1-2, 4-8, 11-26, 29.....305

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

I. INTRODUCTION

1. I have been retained by counsel for Petitioner Amazon.com Services LLC ("Amazon") as a technical expert in connection with the proceeding identified above. I submit this declaration in support of Amazon's Petition for *Inter Partes* Review of United States Patent No. 12,401,720 ("the '720 patent").

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

II. EXPERIENCE AND QUALIFICATIONS

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

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

5. I received a Ph.D. in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology ("MIT") in 1998. Beforehand, I received a Master of Science degree in Electrical Engineering and Computer Science in 1991, a Bachelor of Science degree in Electrical Engineering and Computer Science in 1989, and a Bachelor of Science degree in Physics in 1990, all from MIT.

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

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

8. I also helped to build the web pages that initiated the above-mentioned video sessions via a web interface. Vice President Al Gore visited our group in 1996 and received a demonstration of—and remotely

drove—a radio-controlled toy car with a wireless video camera mounted on it that was built by our group. This toy car device received commands transmitted over a network from a remote computer, and video data from the toy car was transmitted wirelessly then over a computer network back to the user controller. On occasion, we allowed users visiting our web site to drive the toy car from their remote computer while they watched the video on their computer. The video stream was encoded by TNS-designed hardware, streamed over the TNS-designed network, and displayed using TNS-designed software.

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

10. From 1997 to 1999, I was a Senior Scientist and Engineer at NBX Corporation, a start-up that made business telephone systems for streaming packetized audio over data networks instead of using traditional telephone lines. Our system incorporated an auto-attendant, which was a basic interactive voice response (IVR) application to route calls to an extension or voice mailbox. For example, the auto-attendant could be programmed to provide a prompt such as "Press 1 for sales, 2 for support, 3

for directions, or dial an extension" and route the call accordingly based on the caller's touch tone input. Our voicemail system also provided voice prompts and menu selections in which a caller or user may select various options in the voicemail system.

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

12. From 1999-2004, I was employed by Empirix or its predecessor company, Teradyne. Empirix was a leader in test tools for telecommunications protocols and systems, providing functional testing tools

as well as load testing tools. One of Empirix's core products was the Hammer, a telecommunications test platform that was used for functional and load testing. The Hammer was commonly used to test IVR systems, and Hammer also offered a testing service for IVRs. From my work with the testing service business, I received U.S. Patent 7,590,542, titled "Method of generating test scripts using a voice-capable markup language." The '542 patent disclosed a method "for generating test scripts for testing an IVR from an IVR's voice-capable markup language application." (EX1032, Abstract.) From 2000-2001, I conceived of and built a test platform for testing Voice-over-IP (VoIP) based on highly parallel network processors. The first application on this new test platform was a cloud emulator for simulating the effects of transmitting VoIP over a busy network. The platform was based on a multi-core high performance network processor chip.

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

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

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

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

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

III. MATERIALS CONSIDERED

17. In preparing this declaration, I have reviewed, among other things, the '720 patent and file history. I have also considered each of the documents cited in this Declaration and the documents specifically identified below and in the Exhibit List. In formulating my opinions, I have further relied upon my extensive experience in the relevant fields.

IV. OVERVIEW OF THE '720 PATENT

A. Filing Date and Priority Dates

18. The '720 patent was filed on May 8, 2025, and issued on August 26, 2025. The '720 patent claims earliest priority to a US provisional application which was filed on January 9, 2012, which I am informed that the Petitioner treats as the earliest effective filing date only for purposes of this proceeding.

B. Specification

19. The specification of US 12,401,720 is a vague aggregation of off-the-shelf technologies at the time of the invention. It recites long laundry lists of interchangeable components. "Actuators" could be speakers, "sounders," displays, coolers, heaters, lights, motors, pumps, etc. (*E.g.*, EX1001, 94:5-23; 96:12-97:31; 99:7-23, 100:10-101:8; 104:11-32; 105:51-106:55; 109:45-110:28.) "Sensors" could be smoke sensors, blood pressure

sensors, voltmeters, capacitance meters, magnetic sensors, strain gauge, angular sensors, etc. (*Id.*, 17:35-38; 17:39-45; 18:42-47; 18:65-66; 19:18-29; 19:30-35.) Various network technologies were described to be "known in the art." (*E.g.,* EX1001, 2:34-38; 6:42-43; 6:66-7:6; 9:35-38.) Each of these elements was well-known at the time, and the specification provides no technical details of how they are integrated. The result is a disjointed specification where components and processes are described vaguely and separately but never tied together in coherent and novel embodiments.

20. Although the claims of the '720 patent focus on "human voice data," the specification provides only sparse and scattered descriptions related to voice processing. (EX1001, claim 1.) A paragraph found in the Summary section describes voice processing over a network, and that paragraph's description is at a high level of generality. It describes "a router in the building (or vehicle), connected to the one or more in-building (or in-vehicle) networks and to the external network, and may be operative to pass digital data between the in-building (or in-vehicle) and external networks." (EX1001, 32:60-64.) Further, "a first device in the building (or vehicle)" can "transmit a digital data corresponding to the sensed voice to the router over the one or more in-building (or in vehicle) network." (EX1001, 32:64-33:10.) Further, a "voice processor" can "process[] the voice", and a

"control server may be operative to produce actuator commands in response to the voice processor output according to the control logic, and may be operative to transmit the actuator commands to the second device via the router." (EX1001, 33:9-14.)

21. FIG. 19 of the '720 specification is similarly high-level. (EX1001, FIG. 19, 134:38-135:30.) At step 191 ("Receive Voice Data"), "voice data sent from one or more microphones (which are part of one or more field units) is received at the controller location." (EX1001, 134:40-43.) At step 192 ("Voice Processing"), "[t]he voice data is checked, processed, conditioned, or otherwise manipulated." (EX1001, 134:43-45.) The specification then generally states that "[f]or, example, in the case of a voice processing that includes a voice recognition function, the detection of a specific human voice is an event that may trigger one or more actuators into action (or to stop an activity) by the control logic." (EX1001, 135:34-38.)

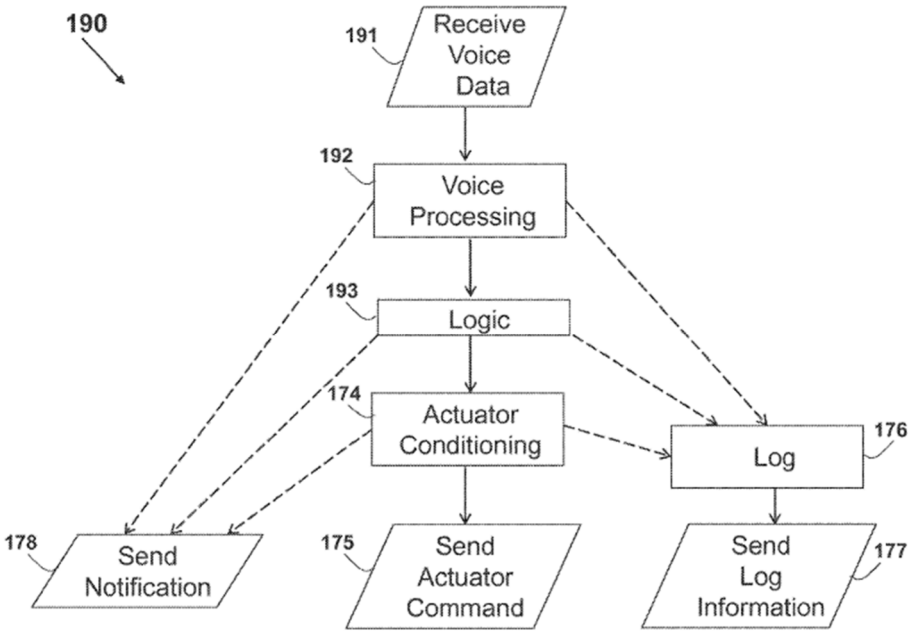


FIG. 19

C. Claims

22. I was asked to evaluate whether claims 1-2, 4-26, and 28-29 of the '720 patent are patentable.

23. There is one independent claim. Claim 1 is directed to a device. The first device comprises: (1) multiple microphones for capturing first human voice data; (2) a Wireless Local Area Network (WLAN) transceiver for transmitting digital data to, and for receiving digital data from, a WLAN; (3) a speaker for sounding audio data; (4) an electric light source for emitting light for emitting light; (5) an enclosure that houses the multiple microphones, the WLAN transceiver, the speaker, and the electric light source, and (6) one

or more processors programmed with computer program instructions. The computer program instructions, when executed, cause the device to (1) send, , to a server device over the Internet via the WLAN by the WLAN transceiver, at least part of the first human voice data captured by the multiple microphones; (2) play, by the speaker, an audio data received from the server device over the Internet via the WLAN by the WLAN transceiver, and (3) emit, by the electric light source, in response to data received from the server over the Internet via the WLAN by the WLAN transceiver. And the device is addressed in the WLAN using an Internet Protocol (IP) address that uniquely identifies the device in the WLAN or in the Internet.

24. There are 28 dependent claims. Claim 2 requires the device to further comprise, in the enclosure, a rechargeable battery for powering the device.

25. Claim 4 requires that the device be addressable using the IP address in the WLAN or in the Internet using distinct locally administered addresses or a universally administered digital addresses. Claim 5 requires the IP address be a layer 3 address and be static or dynamic IP address that is IPv4 or IPv6 type address.

26. Claim 6 requires the device to further comprise a volatile or non-volatile memory that stores the IP address.

27. Claim 7 requires the speaker to comprise an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker. Claim 8 requires that the speaker be operative for omnidirectional, unidirectional, or bidirectional pattern audio emitting.

28. Claim 9 requires the multiple microphones be configured to improve far-field audio reception. Claim 10 requires that the multiple microphones be arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array.

29. Claim 11 requires that each of the multiple microphones comprises, or consists of, an omnidirectional, unidirectional, or bidirectional microphone that is based on sensing of an incident sound-based motion of a diaphragm or a ribbon. Claim 12 requires that each of the multiple microphones comprises, or consists of, a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone.

30. Claim 13 requires that the light be a visible light for illumination or indication. Claim 14 requires that the light be a non-visible light that is infrared, ultraviolet, X-rays, or gamma rays. Claim 15 requires the electric light source to consist of, or comprise, a Solid-State Lighting (SSL), a Light

Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode.

31. Claim 16 requires the device to further comprise a memory that stores digital audio content, and the one or more processors are further programmed with computer program instructions that, when executed, cause the device to playing by the speaker the stored digital audio content. Claim 17 requires that the digital audio content be pre-recorded or synthesized audio content. Claim 18 requires that the device be further operative to play a simulating of a voice of a human being or to play music. Claim 19 requires that the digital audio content to comprise a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.

32. Claim 20 requires the WLAN be according to, or compatible with, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard. Claim 21 requires the WLAN to use a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band.

33. Claim 22 requires that the enclosure be a handheld enclosure.

34. Claim 23 requires that the device to further comprise in the enclosure an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, and that the actuator be

coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN. Claim 24 requires the actuator to consist of, or comprise, a motion actuator that causes linear or rotary motion. Claim 25 requires the actuator to consist of, or comprise, an electromagnetic coil or an electromagnet operative for generating a magnetic or electric field. Claim 26 requires the actuator to consist of, or comprise, an electrical signal generator. Claim 28 requires the actuator to consist of, or comprise, a thermoelectric actuator that is a heater or a cooler and is configured for affecting a temperature of a solid, a liquid, or a gas object, by conduction, convection, force convention, thermal radiation, or by transfer of energy by phase changes.

35. Claim 29 requires the physical phenomenon to comprise temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current.

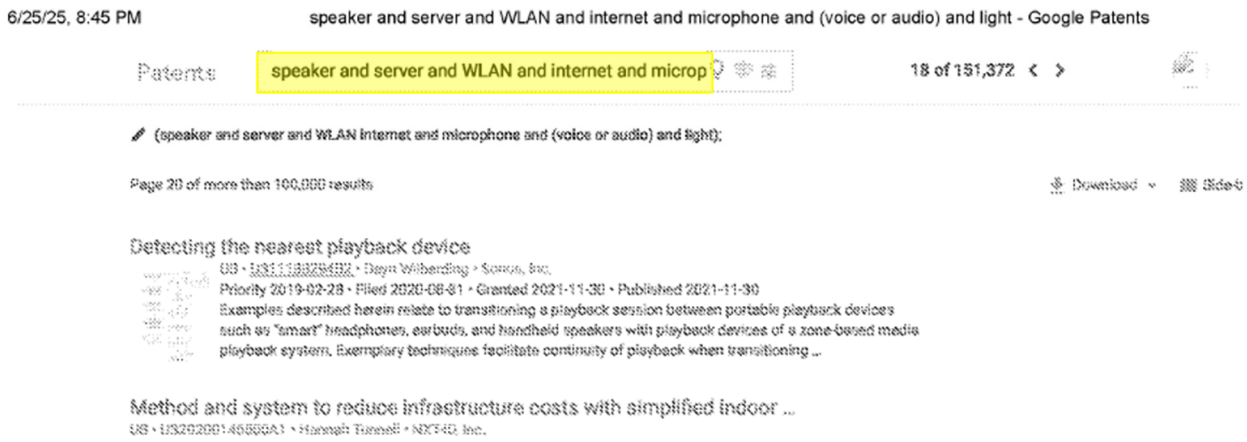
D. Prosecution History

36. I understand that during prosecution of the ’720 patent, the Examiner did not issue any Office Actions.

37. The application that matured into the ’720 Patent was filed on May 8, 2025. (EX1003, 470-475.) No Office Action was issued. The

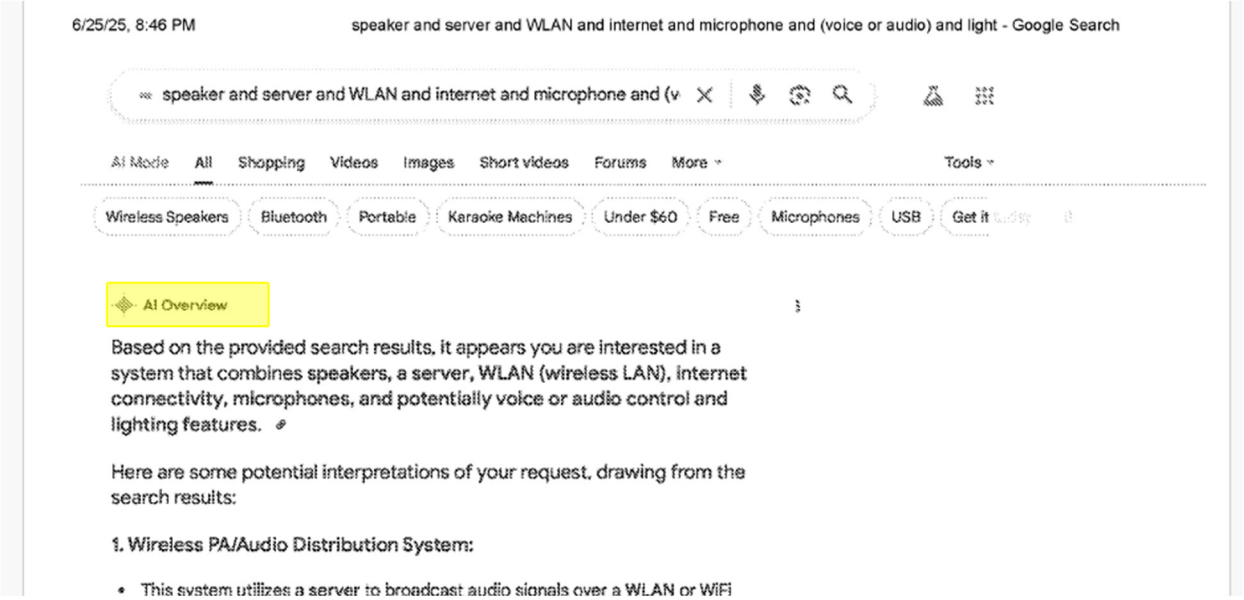
Examiner allowed the claims on July 9, 2025, within **2 months** of filing, citing the entirety of claim 1 as the reason for allowance. (*Id.*, 170-176.)

38. The Examiner performed a defective search on Google Patents, yielding over 100,000 results. His search string reads: "**speaker** and **server** and **WLAN** and **internet** and **microphone** and (**voice** or **audio**) and **light**." (EX1003, 212-213.) I note that this search string is deficient not only because it disregards numerous claim limitations, but also because it does not even appear to restrict the results by any priority date. The top 20 results returned by Google Patents were included as his search results.



(Excerpt of Examiner's Search Strategy and Results, EX1003, 212-213.)

39. The Examiner also relied on Google's "AI Overview," generated in response to the flawed search string, as part of his search strategy and results. Because the search string itself was overly broad and not limited by the application's alleged priority date, the resulting AI Overview is irrelevant.



40. The Examiner's PE2E searches were also at a high level of generality. His search strings are mostly focused on mere combinations of generic components, such as "light," "server," "microphones," and "Internet." (EX1003, 221-227.)

41. If the Examiner was relying on the parent prosecution for a thorough examination, this approach is doubly flawed. First, I note that the claims of the '720 patent is significantly broader than any of its parents. For example, the '710 patent (*i.e.*, the first issued patent in the family) requires "a controlled device in a building," "an Internet-connected server device external to the building," "capturing, by the microphone in the client device, a second human voice data; sending to the server, by the client device via the wireless network, the captured second human voice data; receiving, by the server over

the Internet, the captured second human voice data; processing, by the server, the captured second human voice data; responsive to the processing, sending a second message, by the server to the controlled device over the Internet; receiving, by the controlled device via the wireless network, the second message; and operating a second actuator in the controlled device in response to the received second message." ('710 patent, claim 1.) None of these is recited in claim 1 of the '720 patent. Second, the examination of the parent patents was just as deficient. For example, the parent '710 patent went through *ex parte* appeal, where the PTAB affirmed the rejections, finding that a client device that receives a voice command triggering operation of a controlled device to be obvious. (EX1004, 389-421.) On remand, the Applicant then added a trivial requirement that a voice command trigger operation of the client device itself. (EX1004, 344-352, 121-130.) Despite the minimal and well-known addition, the Examiner abruptly allowed the application without offering any substantive explanation. (*See* EX1004, 8-14.)

V. RELEVANT LEGAL STANDARDS AND DEFINITIONS

A. Obviousness

42. I have been informed that a claim is unpatentable if the differences between the claimed subject matter and the prior art are such that

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

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

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

- the claimed subject matter is simply a combination of prior art elements according to known methods to yield predictable results;
- the claimed subject matter is a simple substitution of one known element for another to obtain predictable results;
- the claimed subject matter uses known techniques to improve similar devices or methods in the same way;
- the claimed subject matter applies a known technique to a known device or method that is ready for improvement to yield predictable

results;

- the claimed subject matter would have been "obvious to try" choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success;
- there is known work in one field of endeavor that may prompt variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations would have been predictable to a POSITA;
- there existed at the time of conception and reduction to practice a known problem for which there was an obvious solution encompassed by the patent's claims; and
- there is some teaching, suggestion, or motivation in the prior art that would have led a POSITA to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed subject matter.

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

B. Prior Art

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

VI. LEVEL OF ORDINARY SKILL IN THE ART FOR THE '720 PATENT

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

48. A POSITA in the field of the '720 patent would have a bachelor's degree in electrical engineering, computer engineering, or equivalent field, and two years of experience working on remote device control and/or networked devices. An advanced graduate degree in the disciplines cited above may make up for the work experience indicated above.

49. I agree with this level of ordinary skill for the '720 patent. Based on my experience, as described above and in Appendix A, I meet the definition of the qualifications of a POSITA, now and as of the alleged priority date. For purposes of this declaration, in general, and unless otherwise noted, my statements and opinions, such as those regarding my own experience, what a POSITA would have understood or known generally (and specifically related to the references I consulted herein), and how a POSITA would have understood and used the terminology found in the '720 patent, reflect the knowledge that existed in the relevant field at the time the '720 patent priority application was filed, and with the state of the art at that time.

VII. CLAIM CONSTRUCTION

50. I have been informed that claim construction requires consideration of the words of the claims themselves, the specification, the prosecution history, and extrinsic evidence concerning relevant scientific principles, the meaning of technical terms, and the state of the art. I have

also been informed that the specification is usually definitive, as it was the single best guide to the meaning of a disputed term. I have been informed that claim terms are presumed to carry their ordinary and customary meaning as understood by a POSITA at the time of the purported invention. I have been informed that, in the context of this proceeding, which requires the Board to construe the terms only to the extent necessary to resolve the issue at hand—invalidity of the claims under 35 U.S.C. §103, Petitioner believes the Board need not construe any of the claims because, as set forth in the Declaration and Petition, the claims are invalid under 35 U.S.C. §103 under any reasonable interpretation, and thus, no claim construction is necessary to resolve any issue before the Board.

VIII. PRIOR ART

51. I understand that each of the references discussed herein qualifies as prior art against the '720 patent:

52. U.S. Publication No. 2010/0088100 ("Lindahl," EX1011) was filed on October 2, 2008, and published on April 8, 2010.

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

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

55. "iPhone User Guide For iOS 5.0 Software" ("iPhone User Guide," EX1008) was published by Apple Inc., by at least November 7, 2011.

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

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

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

59. "Wireless Communications" ("Goldsmith," EX1042) was written by Andrea Goldsmith, and published by Cambridge University Press by at least January 20, 2006.

60. U.S. Publication No. 2008/0154608 ("Evermann," EX1012) was filed on February 12, 2007, and published on June 26, 2008.

61. U.S. Patent No. 8,751,056 ("Slaby," EX1020) was filed on November 30, 2011.

62. The Examiner never found, and the Applicant never cited the references during prosecution.

IX. ANALYSIS AND OPINIONS

A. Ground 1: Lindahl and Williams Render Obvious Claims 1, 2, 4-8, 11-26, and 29

1. Summary of Lindahl

63. Lindahl was filed on October 2, 2008, and published on April 8, 2010, before the earliest provisional application filing date of the '720 patent, and thus qualifies as prior art under §102(a), (b), and (e). (EX1011, 1.) The Examiner never found, and Applicant never cited, Lindahl during prosecution.

64. Lindahl is assigned to Apple Inc. and relates to "electronic devices such as ***portable electronic devices*** that can capture ***voice commands*** and contextual information." (EX1011, [0001], *see also* Abstract.) FIG. 2 provides an illustrative portable electronic device. (EX1011, [0016].)

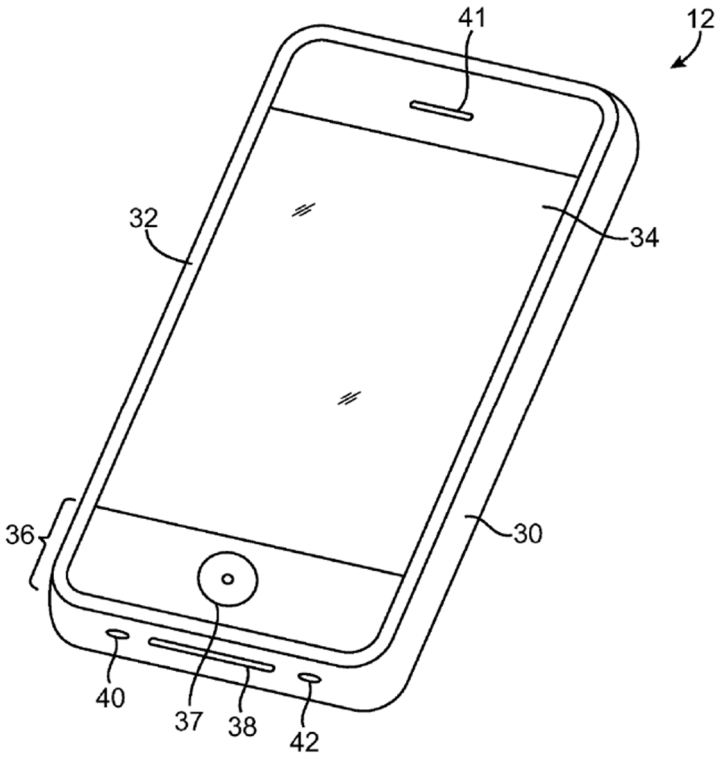


FIG. 2

65. Lindahl's user device can receive voice commands and provide a visual or auditory response. "[D]uring a media playback operation, device 12 may receive a voice command to '***find more music like this***.'" (EX1011, [0042].) The "device 12 may upload the voice command and relevant contextual information . . . to ***computing equipment 14***." *Id.* ***Computing equipment 14*** "may be a server (e.g., an internet server)" and "may receive and process the voice command." (EX1011, [0028], [0042].) "[A]fter processing the voice command and contextual information, ***computing***

equipment 14 may transmit results associated with the voice command to user device 12 which may then display the results." (EX1011, [0043].)

66. Lindahl's device includes all of the hardware components recited in claim 1 of the '720 patent. With reference to FIG. 3, Lindahl's device can include a "user input devices 50 such as . . . ***microphones***." (EX1011, [0048].) It further includes display device(s) such as "***liquid-crystal display (LCD) screens*** or other screens, ***light-emitting diodes (LEDs)***, and other components that present visual information and status data," as well as "audio equipment such as ***speakers and other devices for creating sound***." (EX1011, [0048].) Lindahl's device also includes "[w]ireless communications devices 54 [which] may be used to support local and ***remote wireless links***. Examples of local wireless links include infrared communications, ***Wi-Fi ® (IEEE 802.11)***." (EX1011, [0052].) "Wireless communications devices 54 [which] may include communications circuitry such as radio-frequency (RF) ***transceiver circuitry***." (EX1011, [0049].) As shown in FIG. 3, these components are all enclosed in a single housing.

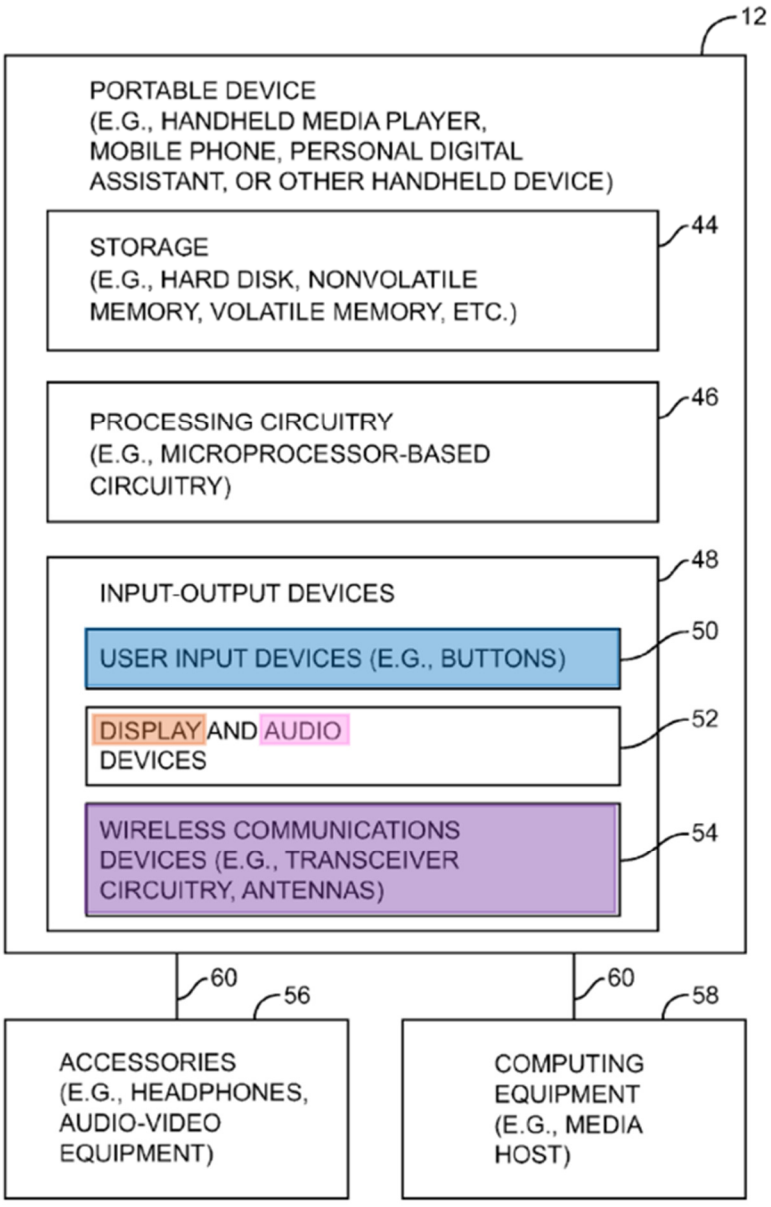


FIG. 3

67. Lindahl discloses several ***server devices***. "***Computing equipment 14*** may include any suitable computing equipment such as a personal desktop computer, a laptop computer, a ***server***, etc. . . . The computing equipment may be a ***server (e.g., an internet server)***[.]"

(EX1011, [0028].) As another example, "***[c]omputing equipment 18*** (e.g., ***one or more servers***) may be associated with one or more online services." (EX1011, [0029].)

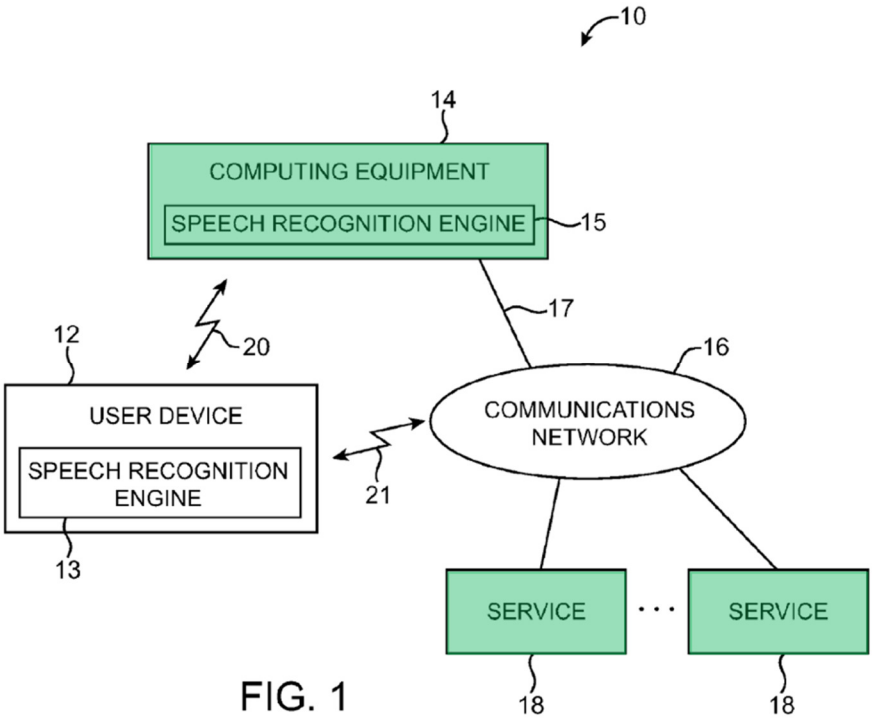


FIG. 1

68. Lindahl's user device communicates with the server over the Internet via a WLAN. Lindahl discloses that "voice commands may be transmitted by user device 12 to ***computing equipment 14*** over ***communications path 20***. Voice commands may also be conveyed to ***remote services 18*** over ***network 16*** (e.g., via ***path 21*** or via ***path 20***, equipment 14, and path 17)." (EX1011, [0025].) "Communications network 16 may be based on any suitable communications network or networks such as a radio-frequency network, the ***Internet***, an Ethernet network, a wireless network, ***a***

Wi-Fi network, a Bluetooth network, a cellular telephone network, or a combination of such networks." (Lindahl, [0032].)

69. Lindahl's user device sends speech commands to the server for processing. As shown in FIG. 6 below, "[a]t step 74, user device 12 may record a ***voice command***. The ***voice command*** may be recorded as an audio clip when a user pressed and releases a record button or supplies other user input directing device 12 to capture the ***voice command***. The voice command may be digitized by device 12 and stored in storage associated with user device 12 such as storage 44." (EX1011, [0065].) "At step 78, user device 12 may upload recorded ***voice commands*** and stored contextual information to computing equipment such as ***equipment 14 or equipment 18***. User device 12 may upload recorded ***voice commands*** and stored contextual information to ***computing equipment 14 or equipment 18*** using any suitable communications path." (EX1011, [0068], [0070], [0028]-[0029].)

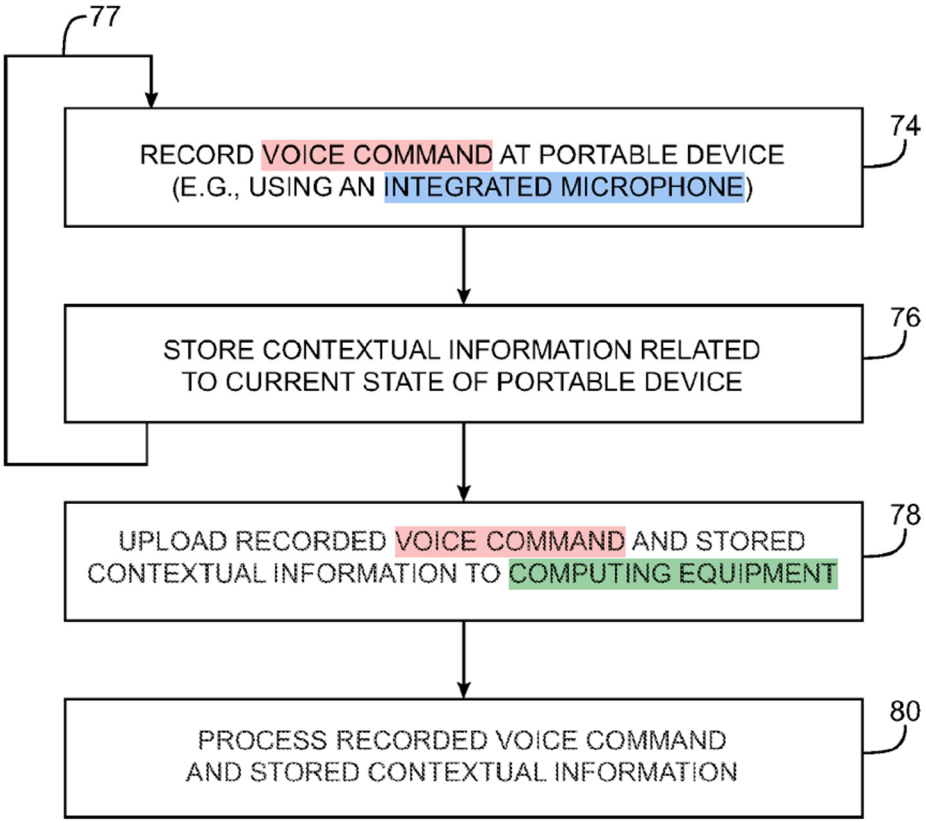


FIG. 6

70. Lindahl discloses that "***computing equipment 14 or 18*** may transmit ***information related to processing and responding*** to the voice command to user device 12. In response, user device 12 may then respond to the voice command." (EX1011, [0072].) Lindahl discloses that "[d]evice 12 may receive ***data signals*** from ***computing equipment 14*** in response to uploading voice commands and contextual information. The ***data*** received by device 12 from ***equipment 14*** in response to voice commands and contextual information may be used by device 12 to carry out requests associated with

the voice commands. For example, after processing the voice command and contextual information, ***computing equipment 14*** may transmit ***results associated with the voice command*** to user device 12 which may then ***display*** the ***results***." (EX1011, [0043].)

2. Summary of Williams

71. Williams was filed on January 28, 2011, and is prior art under §102(e). (EX1006, 1.) The Examiner never found, and Applicant never cited, Williams during prosecution. Williams is assigned to Amazon Technologies, Inc. Williams relates to "utilizing audio information that may already exist within a home or other location." (EX1006, [0007].) For example, "[a] user may ask a question, such as '***What's the recipe for deviled eggs?***'" and the system "may use the speech synthesis functionality of the system to provide ***audio answers*** to spoken questions." (EX1006, [0048].) Further, Williams' system "may accept spoken commands from a user regarding on-premises devices, and may respond by interacting with such devices through available network connections." (*Id.*, [0051]; *see also id.*, [0001], [0008], [0009], [0024].)

72. Various components of Williams are depicted in FIG. 1, which illustrates "an operational environment in which applications may utilize an application engine, and may respond to ***audio from user premises*** to provide

various services to users." (EX1006, [0003].) Williams discloses that "[a]n individual interface device 116 may include a ***microphone 202*** that is configured to detect ambient noise, sounds, and ***speech***." (*Id.*, [0023].) For network connection, "the devices within the home 102 are illustrated as connecting to the network 114 via the in-home router 112." (EX1006, [0013].)

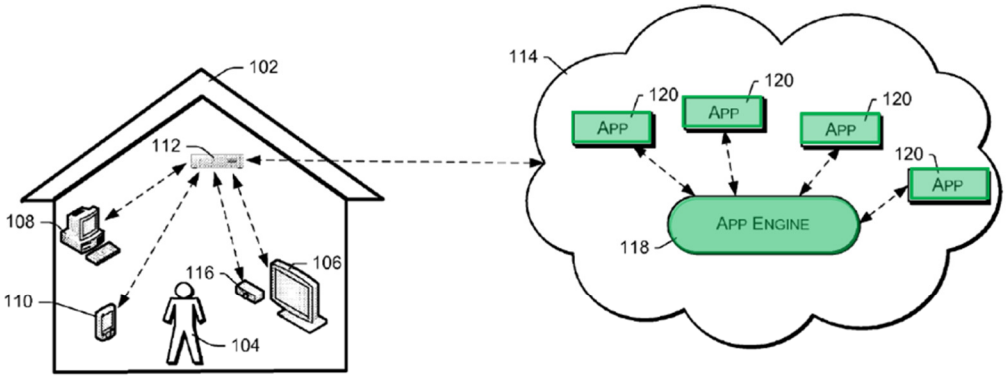


FIG. 1

73. As shown in FIG. 1, "[t]he network 114 represents what has come to be known as the 'cloud,' which may include a vast array of network-based entities, servers, clients, applications, and services." (EX1006, [0014].) Williams discloses that the network 114 includes "servers" including "a ***server*** 400 that may be used to implement the functionality of the application engine 118 and/or any of the applications 120." (*Id.*, [0052].)

74. Williams discloses that the interface device 116 receives ***speech commands***, which are processed by the ***server*** and cause actuation of a speaker of the interface device 116. "In operation, the application engine 118 [of the ***server 400***] receives ***audio-based information*** from the interface devices 116. The audio information may include audio streams, ***user commands or notifications derived from vocal utterances***, environmental information derived from on-premises audio, and so forth. The application engine 118 [of the ***server 400***] processes the audio-based information to determine various data about user activities, status, environment, commands, etc. This data can then be obtained by authorized applications 120, which can ***act on the information to provide or enhance services to the user 104***." (EX1006, [0022].) For example, the server can "provide ***audio 222*** for eventual transmission to the interface device 116, to be played on the ***speaker 206***." (EX1006, [0037].)

3. A POSITA would have been motivated to combine Lindahl and Williams

75. At the time of the invention, it would have been obvious for a POSITA to combine Lindahl and Williams. Lindahl and Williams both relate to processing and responding to human speech commands via a remote server. Lindahl discloses "using a portable electronic device such as user

device 12 to receive and upload voice commands and in using computing equipment such as computing equipment 14 to process the uploaded voice commands." (EX1011, [0073].) "[A]fter processing the voice command and contextual information, computing equipment 14 may transmit results associated with the voice command to user device 12 which may then display the results." (EX1011, [0043].) Similarly, Williams' server receives and processes audio information and provides an audio response for playback. (EX1006, [0027], [0036].) Thus, Lindahl and Williams are analogous to the '720 patent and are from the same field of endeavor. They relate to the same technical problems (*e.g.*, processing and responding to human speech commands).

76. In particular, Williams expressly teaches that the ***server*** transmits "audio" data back to the user device, which plays it back via a speaker. For example, "[t]he API 214 of the application engine 118 may also allow the application 120 to provide ***audio 222*** for eventual transmission to the interface device 116, to be played on the speaker 206." (EX1006, [0036].)

77. The Lindahl-Williams combination would result in a system where the server (*e.g.*, Lindahl's computing equipment 14 and/or 18) receives a voice command, processes the voice command, and transmits an audio response (*e.g.*, "audio 222") to the user device for playback. This way, the

device can respond to any voice command using an audio response (*e.g.*, speech, music via a speaker) as taught by Williams, a visual response (*e.g.*, via a display) as taught by Lindahl, or both depending on the context. Multi-modal feedback improves the usability of the device.

78. Although it is possible for a server to send ***text*** data to a user device, which in turn converts the text data into audio data for playback, there were numerous advantages to sending ***audio*** data directly as taught by Williams. For example, sending audio directly would reduce or eliminate the need for local text-to-speech processing at the device, thereby avoiding additional software or hardware resources and improving battery life. Lindahl expressly contemplates that the processing power on the user device 12 is limited and it is advantageous to offload the processing to the remote server. "[U]ser device 12 might not be able to execute the voice command immediately for reasons such as an inability to perform adequate speech processing due to a ***lack of available processing power***." (EX1011, [0035].) Doing so reduces the burden on the local computing resources (*e.g.*, to perform text-to-speech conversion) and results in enhanced and faster results because the server has more processing power. Thus, it would have been obvious to send audio directly to the Lindahl-Williams' device with limited proceeding powers.

79. While audio transmission may potentially require more bandwidth than text transmission, it would have been an acceptable tradeoff given its numerous advantages. Audio transmission does not demand significant bandwidth to begin with. Further, transmitting audio would improve user experience by ensuring the audio received is consistent (*e.g.*, in voice, tone, pronunciation, timing, effects) across every device even if they have different software and/or hardware configurations. Server-side audio generation would be more sophisticated, resulting in higher-quality audio for playback. Further, server-side text-to-speech processing can be faster than waiting for the device to synthesize speech. Moreover, sending audio directly provides better security, ensuring that the user hears exactly what was intended and not affected by user settings, malicious manipulation, or missing voices.

80. Further, it would have been obvious to try combining the teachings of Lindahl and Williams. There are only a finite number of identified and predictable solutions for transmitting information from a remote server to a device for audio output, either sending audio directly or sending other forms of data (*e.g.*, text) to be converted into audio on the device. Transmitting audio directly rather than text represents a routine design choice between known alternatives. The combination is straightforward and readily

achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Williams demonstrates the feasibility and benefits of direct audio transmission, and a POSITA would have had a reasonable expectation of success in applying that approach to Lindahl's architecture.

4. Claim 1

a. 1[PR]: "A device comprising:"[1]

81. Lindahl-Williams discloses a device. For example, Lindahl relates "generally to ***electronic devices***, and more particularly, to electronic devices such as ***portable electronic devices*** that can capture voice commands and contextual information." (EX1011, [0001], *see also* Abstract.) Lindahl provides "[a] portable ***electronic device*** such as a handheld ***electronic device***[.]" (*Id*., [0006].) "The electronic device may have a microphone that is used to receive voice commands. The electronic device may use the microphone to record a user's voice. The recording of the user's voice may be stored as a digital audio file in storage associated with the electronic device." *Id*.

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

82. Lindahl describes ***a user device 12***. Lindahl discloses that "[a] user in system 10 may have an electronic device such as ***user device 12***. User device 12 may be used to receive voice commands (e.g., to record a user's voice)." FIG. 2 below shows "an illustrative ***user device 12*** in accordance with an embodiment of the present invention." (EX1011, [0038].)

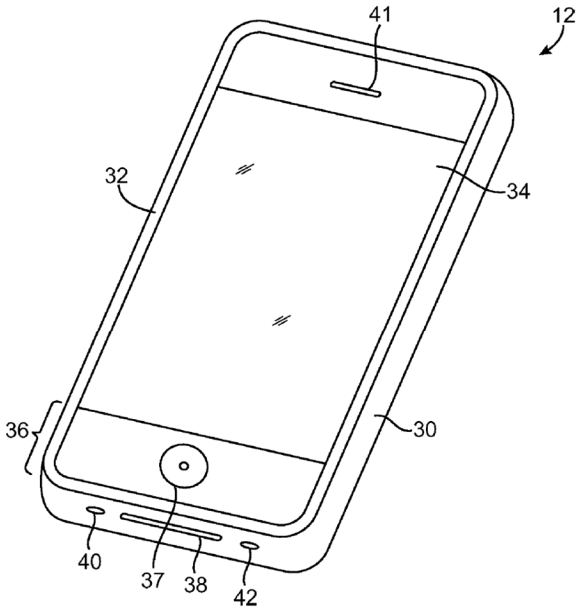


FIG. 2

83. Lindahl's user device 12 may take the form of various electronic devices. For example, "[u]ser device 12 may be a ***mobile telephone***, a mobile telephone with media player capabilities, a media player, a handheld computer, a game player, a global positioning system (GPS) device, a combination of such devices, or any other suitable electronic device such as a

portable device." (EX1011, [0044].) Lindahl also discloses that "the portable electronic devices may be wireless electronic devices . . . [that] may be, for example, handheld wireless devices such as cellular telephones, media players with wireless communications capabilities, handheld computers (also sometimes called personal digital assistants), global positioning system (GPS) devices, and handheld gaming devices." (EX1011, [0023]-[0024], [0027].) "Example of hybrid portable electronic devices include a cellular telephone that includes media player functionality, a gaming device that includes a wireless communications capability, a cellular telephone that includes game and email functions, and a portable device that receives email, supports mobile telephone calls, has music player functionality and supports web browsing. These are merely illustrative example." (EX1011, [0024], [0027].)

84. Further, Williams also discloses this element. (*Infra* §IX.H.3.a.)

b. 1[A]: "multiple microphones for capturing first human voice data;"

85. Lindahl-Williams discloses or suggests ***multiple microphones*** for capturing ***first human voice data***. First, Lindahl discloses its device comprises ***multiple microphones***. For example, with reference to FIG. 3, "[i]nput-output devices 48 can include ***user input devices 50*** such as buttons,

touch screens, joysticks, click wheels, scrolling wheels, touch pads, key pads, keyboards, ***microphones***, cameras, etc." (EX1011, [0048].)

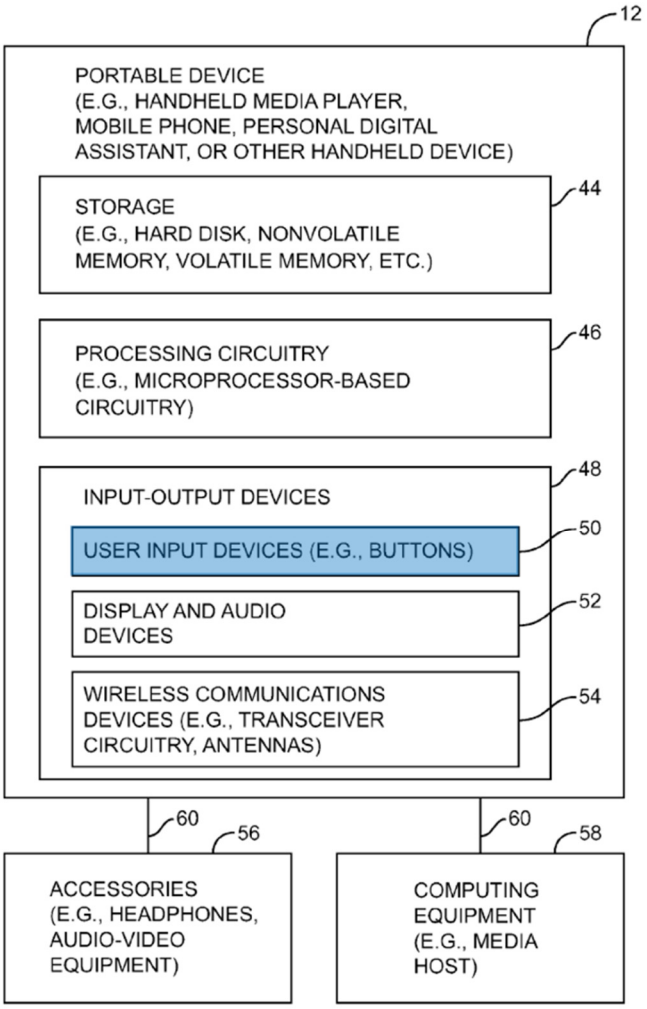


FIG. 3

86. I note that Lindahl is a patent assigned to Apple Inc. Consistent with Lindahl, many iPhones by 2012 were equipped with multiple microphones, as shown below. (EX1008, 9.) It would have been obvious to a POSITA that Lindahl's user device has multiple microphones.



87. Lindahl discloses multiple instances of user device 12 "capturing" human voice data. For example, "user device 12 may be used to ***capture voice commands*** from a user during the operation of user device 12." (EX1011, [0034].) "***Voice commands*** may be ***captured*** with an electronic device and uploaded to computing equipment for further processing. Electronic devices that may be used in this type of environment may be portable electronic devices such as laptop computers or small portable computers of the type that are sometimes referred to as ultraportables." (EX1011, [0023].)

88. Lindahl discloses several exemplary scenarios for capturing voice commands. For instance, "during a media playback operation, device 12

may receive a voice command to '***find more music like this***.'" (EX1011, [0042].) "As an example, a user may be using a mapping application on a handheld electronic device. The device may be presenting the user with a map that indicates the user's current position. The user may press a button or may otherwise instruct the handheld electronic device to record the phrase '***I like American restaurants in this neighborhood***.' In response, the electronic device may record the spoken phrase." (EX1011, [0013].) As a further example, device 12 may receive voice commands such as "***identify this car***" and "***create new event for Sunday, July 18th: James's Birthday***." (EX1011, [0063], [0067].) These voice commands are ***first human voice data***.

89. It would have been obvious that the capturing is done by Lindahl's "***microphones***". In fact, Lindahl expressly discloses that "[a] user can control the operation of user device 12 by supplying ***commands*** through ***user input devices 50***," which includes "***microphones***". (EX1011, [0048].)

90. While Lindahl describes receiving voice commands using "microphone" in certain places, its explicit disclosure of "microphone**s**" makes it clear that there are multiple microphones in Lindahl. (EX1011, [0006], [0040], [0048].) A POSITA would have understood that each of Lindahl's microphones would capture voice command uttered near the device.

It was common knowledge by 2012 that using multiple microphones improve signal clarity and reduce background noise, which is critical to enhance speech recognition (*e.g.*, by Siri). (EX1010, 1021.) Further, it was commercially implemented in iPhones at the time. (EX1008, 9.) Thus, a POSITA would have understood and find it obvious that Lindahl's "microphone**s**" are used to receive voice commands.

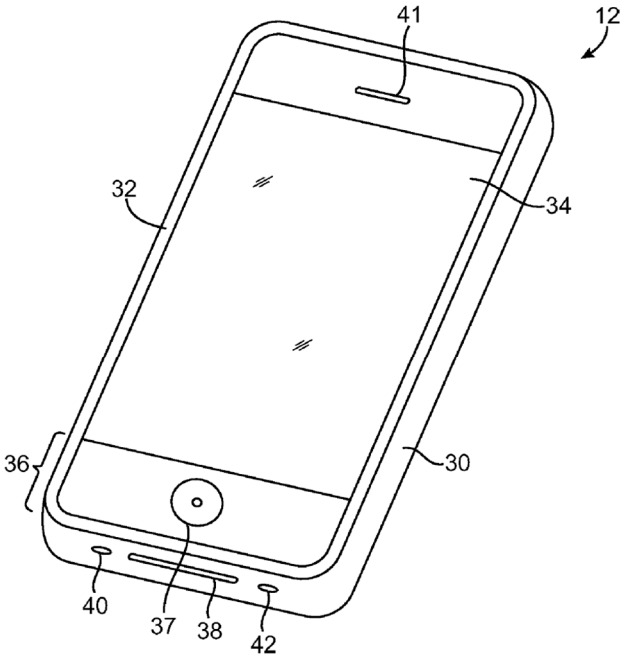


FIG. 2



91. Further, Williams also discloses this element. (*Infra* §IX.H.3.b.)

c. 1[B]: "a Wireless Local Area Network (WLAN) transceiver for transmitting digital data to, and for receiving digital data from, a WLAN;"

92. Lindahl-Williams discloses a ***WLAN transceiver*** for transmitting digital data to, and for receiving digital data from, a WLAN.

93. First, Lindahl discloses a device comprising a ***WLAN transceiver***. As shown in FIG. 3 below, Lindahl's device 12 comprises "[w]ireless communications devices 54 [which] may include communications circuitry such as radio-frequency (RF) ***transceiver circuitry*** formed from one or more integrated circuits, power amplifier circuitry, passive RF components, one or more antennas, and other circuitry for handling RF wireless signals." (EX1011, [0049].) "Wireless communications devices 54 may be used to

support local and ***remote wireless links***. Examples of local wireless links include infrared communications, ***Wi-Fi ® (IEEE 802.11)***, Bluetooth®, and wireless universal serial bus (USB) links." (EX1011, [0052].) In other words, Lindahl's device 12 has a "transceiver circuity" for supporting Wi-Fi. A POSITA would have understood that this "transceiver circuitry" discloses a Wi-Fi transceiver.

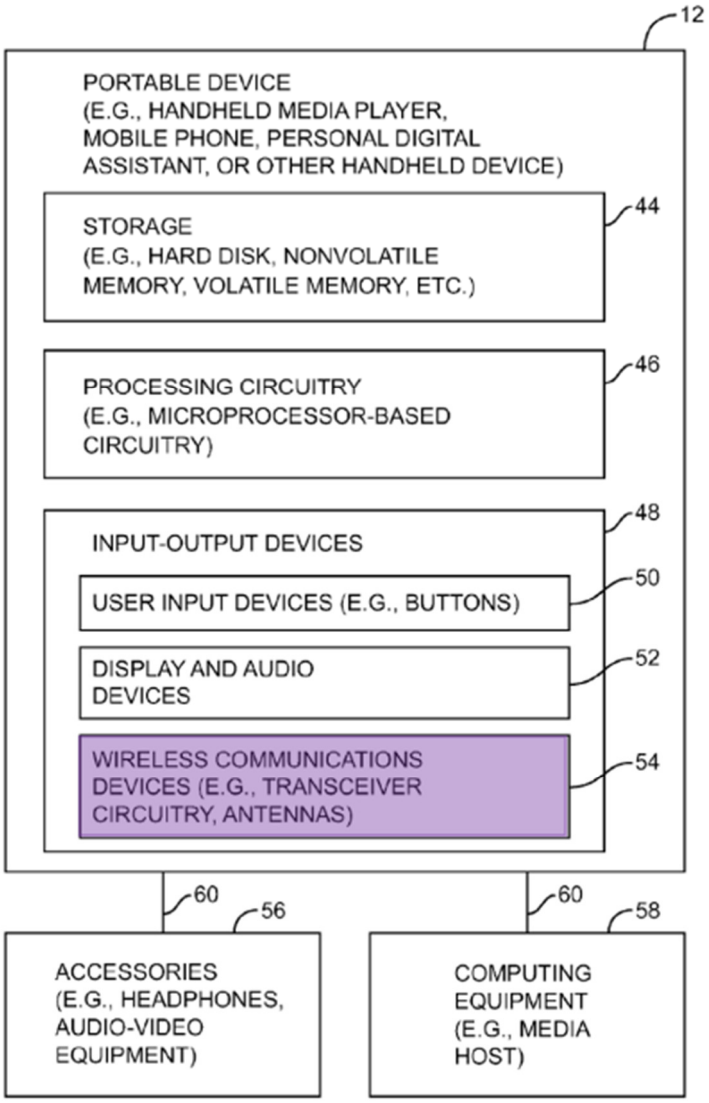


FIG. 3

94. Lindahl's Wi-Fi transceiver is a WLAN transceiver. At the time of the invention, it was common knowledge that WiFi implements a WLAN. "Wireless local area networks (***WLANs***) are built around the family of IEEE 802.11 standards . . . The 802.11a/b/g family of standards are collectively referred to as ***Wi-Fi***, for wireless fidelity." (EX1042, 600-601.) "***Wi-Fi*** is a registered trademark owned by the Wi-Fi Alliance, and it refers to ***wireless local area network (WLAN)*** products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards." (EX1013, 7:48-51.) This is consistent with the description of the '720 patent, which states that "[i]n order to allow interoperability between equipment manufactured by different vendors, few ***WLAN standards*** have evolved, as part of the ***IEEE 802.11 standard group***, ***branded as WiFi*** (www.wi-fi.org). IEEE 802.11b describes a communication using the 2.4 GHz frequency band and supporting communication rate of 11 Mb/s, IEEE 802.11a uses the 5 GHz frequency band to carry 54 MB/s and IEEE 802.11g uses the 2.4 GHz band to support 54 Mb/s." (EX1001, 8:25-32.) Thus, it would have been obvious to a POSITA that Lindahl's "transceiver circuitry" is a Wi-Fi transceiver, which is a type of WLAN transceiver. (EX1042, 600-601; EX1013, 7:38-51; EX1001, 8:25-32.)

95. Indeed, by 2012, handheld cellular telephones were commonly equipped with a WiFi transceiver to enable WiFi connectivity. (EX1037, 3; EX1035, 1:20-23; EX1011, [0024], [0027], FIG. 2.)

96. Second, Lindahl's WLAN transceiver is for transmitting digital data to, and for receiving digital data from, a WLAN. It was common knowledge that a "WLAN" transceiver would be for transmitting digital data to, and for receiving digital data from a "WLAN" as required by the claim. (EX1035, 3:17-20, 3:38-40; EX1036, [0013]-[0015]; *infra* IX.A.4.h-IX.A.4.j.) Indeed, Lindahl discloses "[u]ser device 12 may ***handle communications*** over one or more wireless communications bands such as local area network bands and cellular telephone network bands." (EX1011, [0038]; *see also id.*, [0027], [0044], [0031].) A POSITA would have understood that such communications involve both transmitting and receiving digital data via a WLAN.

97. Indeed, as discussed in detail below, in Lindahl, "voice commands" that compass digital data are transmitted to the WLAN. (*Infra* §IX.A.4.h.) For example, FIG. 1 shows "an illustrative environment in which a user may interface with system components using voice commands." (EX1011, [0025].) With reference to FIG. 1 below, Lindahl discloses that "***voice commands*** may be transmitted by user device 12 to computing

equipment 14 over communications path 20. Voice commands may also be conveyed to remote services 18 over network 16 (e.g., via path 21 or via path 20, equipment 14, and path 17)." (EX1011, [0025].) Lindahl expressly teaches that these communication paths and networks can all be based on WLAN. (EX1011, [0026], [0030], [0031], [0032].) "Communications paths in system 10 may, if desired, be based on ***wireless communications technology*** such as satellite technology, radio-frequency (RF) technology, wireless universal serial bus technology, and ***Wi-Fi***® or Bluetooth® 802.11 wireless link technologies." (EX1011, [0030].)

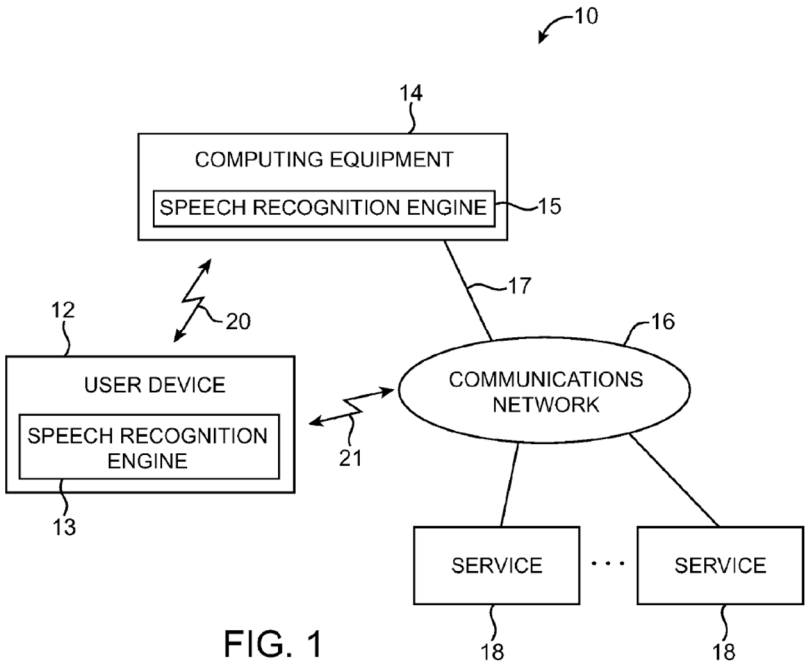


FIG. 1

98. Further, as discussed in detail below, after sending the audio data to the server device (i.e., computer equipment) for processing, "[t]he

computing equipment may also ***transmit information*** to the electronic device that allows the electronic device to respond to the voice command." (*Infra* § IX.A.4.i; EX1011, Abstract.) Such transmission of information is over the Internet via the WLAN.

99. Further, Williams also discloses this element. (*Infra* § IX.H.3.c.)

d. 1[C]: "a speaker for sounding audio data;"

100. Lindahl-Williams discloses a ***speaker*** for sounding audio data. For example, Lindahl discloses that "[d]evice 12 may also have audio capture and ***playback capabilities***." (EX1011, [0040].) Specifically, with reference to FIG. 2, Lindahl discloses its device 12 includes "[p]ort 41 [which] may be, for example, a ***speaker*** port." (EX1011, [0040], FIG. 2.) Indeed, in and before 2012, iPhones were equipped with an earpiece receiver speaker (*e.g.*, positioned toward the top part of the device). (EX1008, 929, 148.)

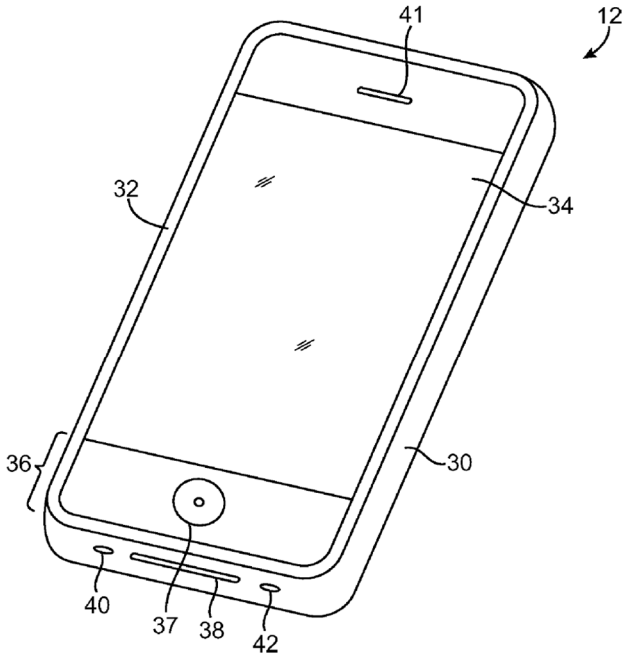


FIG. 2

101. Lindahl's speaker is for sounding audio. Further, with reference to FIG. 3, Lindahl discloses its device 12 comprises display and audio devices 52. Lindahl discloses that "[d]isplay and audio devices 52 may also include audio equipment such as ***speakers and other devices for creating sound***." (EX1011, [0048].)

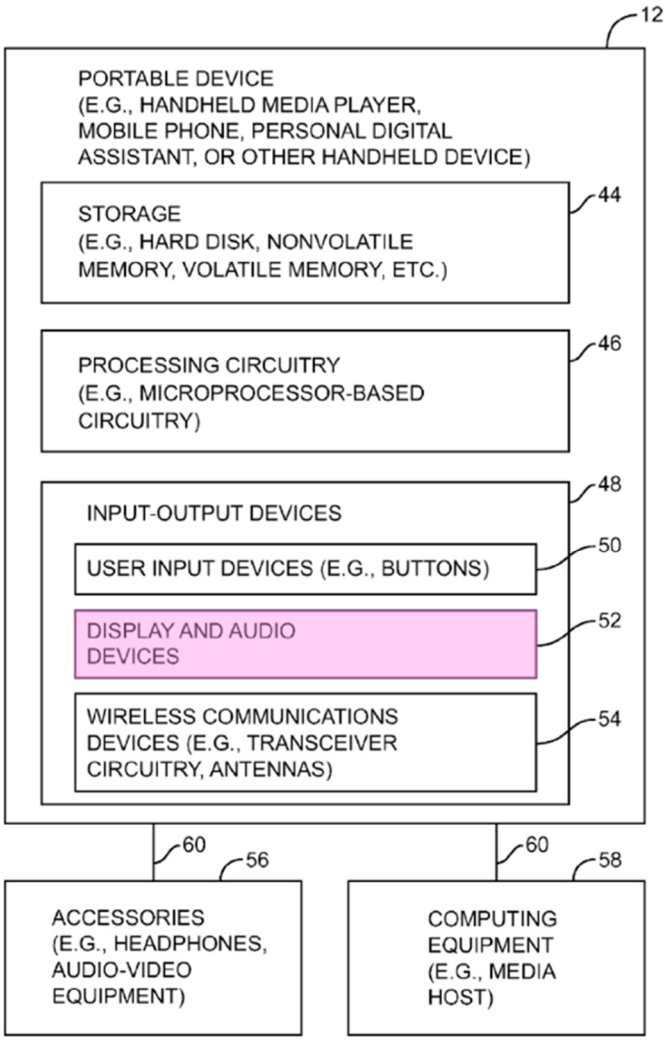


FIG. 3

102. The specification of the '720 patent confirms that all speakers are for sounding audio data. "A ***loudspeaker (a.k.a. speaker)*** is a ***sounder*** that ***produces sound*** in response to an ***electrical audio signal input***, typically ***audible sound***. The most common form of loudspeaker is the electromagnetic (or dynamic) type, uses a paper cone supporting a moving voice coil electromagnet acting on a permanent magnet. Where accurate reproduction of sound is required, multiple loudspeakers may be used, each reproducing a part

of the audible frequency range. A loudspeaker is commonly optimized for middle frequencies; tweeters for high frequencies; and sometimes supertweeter is used which is optimized for the highest audible frequencies." (EX1001, 105:64-106:7.)

103. Further, Williams also discloses this element. (*Infra* § IX.H.3.d.)

e. 1[D]: "an electric light source for emitting light for emitting light;"

104. Lindahl discloses an ***electric light source*** for emitting light for emitting light. I note that the '720 patent provides many examples of electric light source. The '720 patent provides that "[t]he light source may be an electric light source for converting electrical energy into light, and may consist of, or comprise, a lamp, such as an incandescent, a fluorescent, or a gas discharge lamp. The electric light source may be based on Solid-State Lighting (SSL) such as a Light Emitting Diode (LED) which may be Organic LED (OLED), a polymer LED (PLED), or a laser diode." (EX1001, 39:19-26.) "LEDs are efficient, and emit more light per watt than incandescent light bulbs. They can emit light of an intended color without using any color filters as traditional lighting methods need. " (EX1001, 95:64-67.)

105. According to the ’720 patent, a display is an electric light source. It states that “a display for presentation of visual data or information, commonly on a screen, [] may consist of ***an array (e.g., matrix) of light emitters*** or light reflectors, and may present text, graphics, image or video.” (EX1001, 109:45-50.) In particular, a liquid crystal display comprises backlit lights such as LEDs. “A display may be a Liquid Crystal Display (LCD) display, which utilize two sheets of polarizing material with a liquid crystal solution between them. An electric current passed through the liquid causes the crystals to align so that light cannot pass through them. Each crystal, therefore, is like a shutter, either allowing a ***backlit light*** to pass through or blocking the light. . . . An ***LED-backlit LCD*** is a flat panel display that uses ***LED backlighting*** . . . Three forms of LED may be used: ***White edge-LEDs*** around the rim of the screen, using a special diffusion panel to spread the light evenly behind the screen (the most usual form currently), an ***array of LEDs*** arranged behind the screen whose brightness are not controlled individually, and a dynamic “local dimming” ***array of LEDs*** that are controlled individually or in clusters to achieve a modulated backlight light pattern.” (EX1001, 110:64-111:3, 111:11-23.)

106. Consistent with the descriptions of the ’720 patent, Lindahl discloses an electric light source for emitting light. For example, Lindahl’s

"[d]isplay 34 may be a touch screen ***liquid crystal display*** (as an example)." (EX1011, [0039].) With reference to FIG. 3, Lindahl discloses its device 12 comprises display and audio devices 52, which "may include ***liquid crystal display (LCD) screens*** or other screens, ***light-emitting diodes (LEDs)***, and other components that present visual information and status data." (EX1011, [0048].)

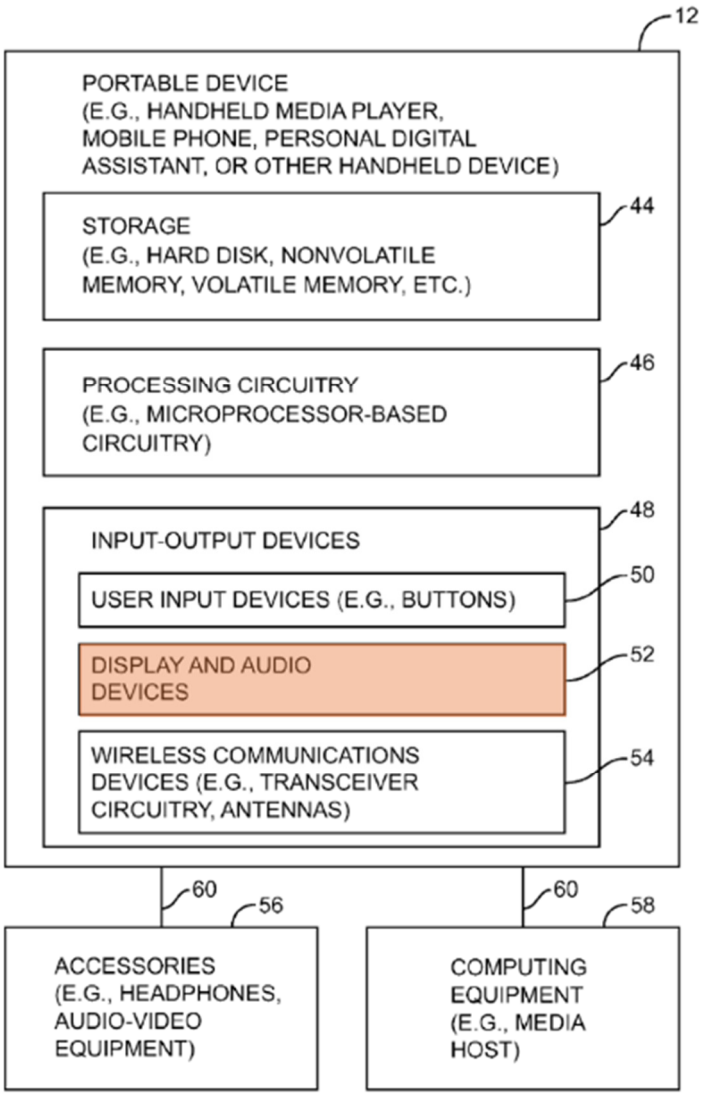


FIG. 3

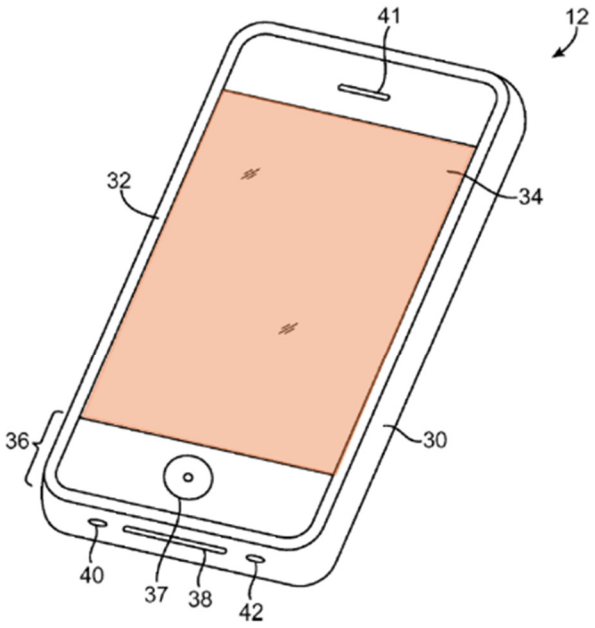


FIG. 2

f. 1[E]: "an enclosure that houses the multiple microphones, the WLAN transceiver, the speaker, and the electric light source;"

107. Lindahl discloses an enclosure that houses the ***multiple microphones***, the ***WLAN transceiver***, the ***speaker***, and the ***electric light source***. With reference to FIG. 2, Lindahl's device 12 has an enclosure formed by the outer shell and the front panel of the "***[d]isplay 34***." (EX1011, [0039].) The enclosure houses the light emitting components (*e.g.*, LEDs) of the display. Further, "[d]evice 12 may have a ***microphone*** for receiving voice commands. Openings 42 and 40 may, if desired, form ***microphone*** and ***speaker*** ports. . . . Device 12 may also have audio capture and playback capabilities. Device 12 may be able to receive voice commands from a user

and other audio though [sic] a ***microphone*** (e.g., formed as part of one or more ports such as openings 40 and 42). Port 41 may be, for example, a ***speaker*** sport [sic]." (EX1011, [0040].) These are all housed in the enclosure.

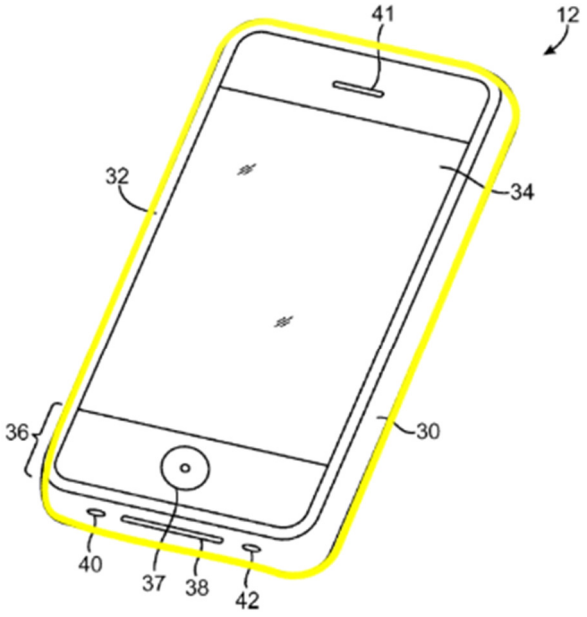


FIG. 2

108. With reference to FIG. 3 below, Lindahl again confirms that its device 12 comprises ***user input devices 50*** which includes ***multiple microphones***, display and audio devices 52 which include ***LCD screens or other screens***, ***LEDs***, ***speakers*** and other devices for creating sound, and wireless communications devices 54 which include a ***WLAN transceiver***. (EX1011, [0048], [0049].)

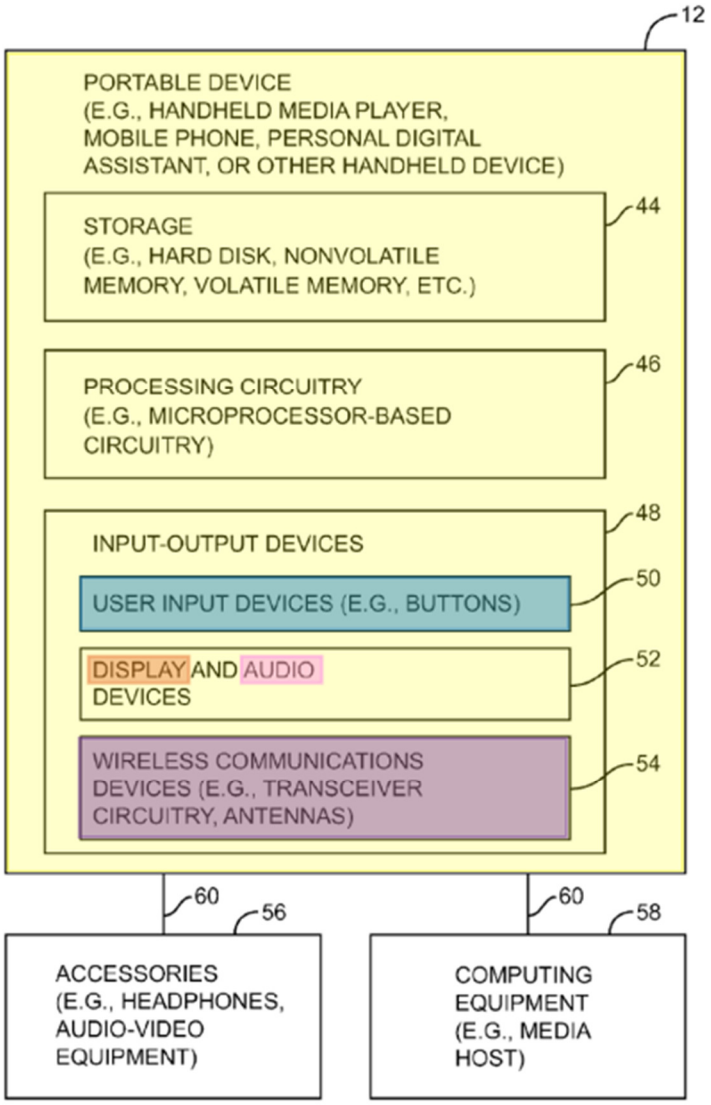


FIG. 3

109. Further, Williams also discloses this element. (*Infra* § IX.H.3.f.)

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

110. Lindahl-Williams discloses this element. For example, with reference to FIG. 3, Lindahl discloses "[p]rocessing circuitry 46 may be used

to control the operation of user device 12. Processing circuitry 46 may be based on a ***processor*** such as a ***microprocessor and other suitable integrated circuits***. With one suitable arrangement, processing circuitry 46 and storage 44 are used to run ***software*** on user device 12, such as speech recognition applications, internet browsing applications, voice-over-internet protocol (VOIP) telephone call applications, email applications, media playback applications, operating system functions (e.g., operating system functions supporting speech recognition capabilities), etc. Processing circuitry 46 and storage 44 may be used in implementing ***analog-to-digital conversion functions for capturing audio*** and may be used to implement speech recognition functions." (EX1011, [0046].) Thus, Lindahl discloses one or more processors programmed with computer program instructions. It would have been obvious to a POSITA that Lindahl's device comprises one or more processors programmed with computer program instructions that, when executed, cause the device to perform the recited operations in the following elements. (*Infra* §§IX.A.4.h-IX.A.4.j.)

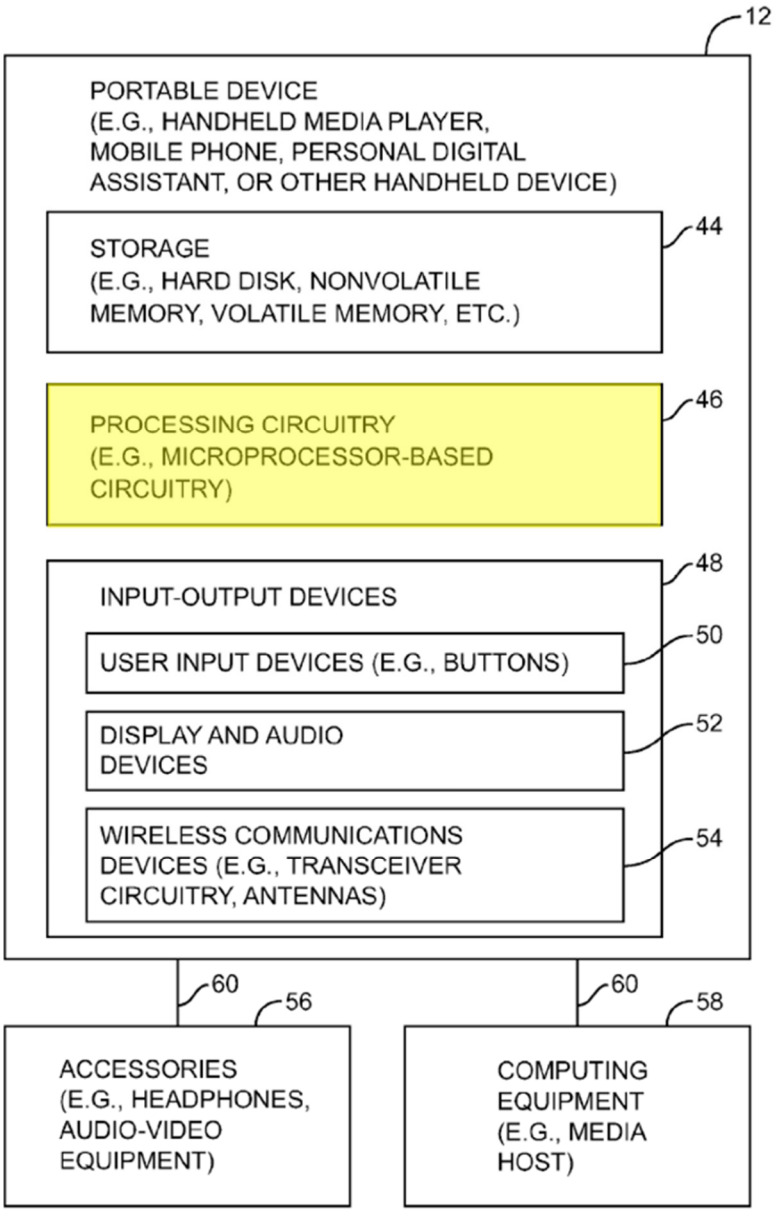


FIG. 3

111. Further, Williams also discloses this element. (*Infra* § IX.H.3.g.)

h. 1[G]: "send, to a server device over the Internet via the WLAN by the WLAN transceiver, at least part of the first human voice data captured by the multiple microphones;"

112. Lindahl-Williams discloses this element. First, Lindahl discloses a ***server device***. Lindahl discloses that "[***c***]***omputing equipment 14*** may include any suitable computing equipment such as a personal desktop computer, a laptop computer, a ***server***, etc. . . . The computing equipment may be a ***server (e.g., an internet server)***, a local area network computer with or without internet access, a user's own personal computer, a peer device (e.g., another user device 12), any other suitable computing equipment, and combinations of multiple pieces of computing equipment." (EX1011, [0028].) As another example, Lindahl discloses a computing equipment 18 and that, "***[c]omputing equipment 18*** (e.g., ***one or more servers***) may be associated with one or more online services." (EX1011, [0029].)

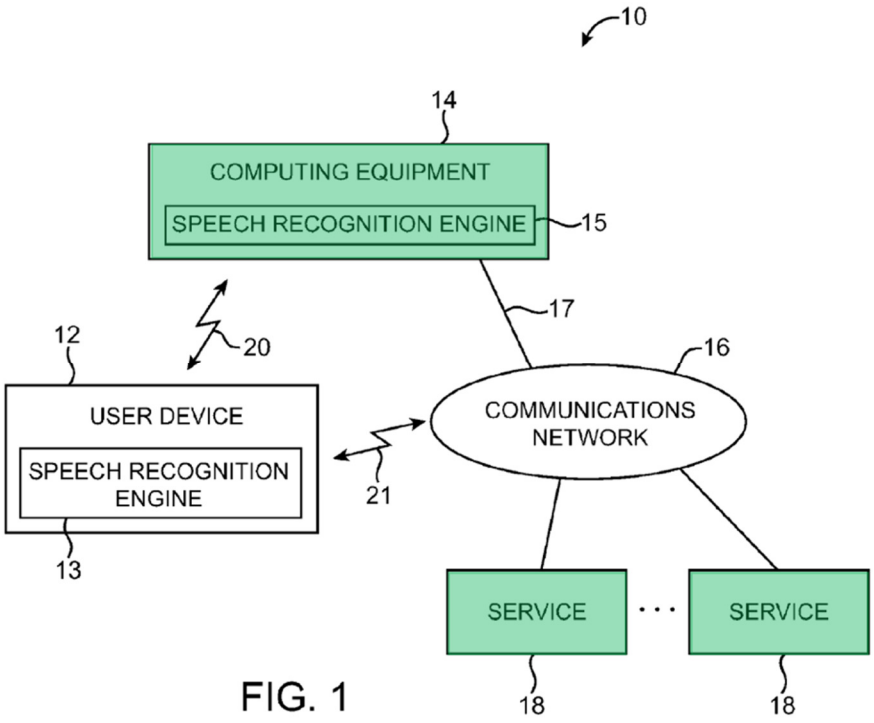


FIG. 1

113. Second, Lindahl discloses sending at least part of the captured ***first human voice data*** by the ***multiple microphones*** to the ***server***.

Referring to FIG. 1, Lindahl discloses that "[u]ser device 12 may be used to receive ***voice commands*** (e.g., to record a user's voice) . . . . If desired, the ***voice commands*** may be transmitted by user device 12 to ***computing equipment 14*** over communications path 20. Voice commands may also be conveyed to ***remote services 18*** over network 16 (e.g., via path 21 or via path 20, equipment 14, and path 17)." (EX1011, [0025], *see also*, Abstract ("[t]he electronic device may transmit the ***voice command*** and the contextual information to computing equipment such as a desktop computer or a ***remote***

server.").) As discussed above, Lindahl's "voice commands" are captured by its multiple microphones. (EX1011, [0048], [0034].)

114. Further, FIG. 6 shows the "[i]llustrative steps involved in using a portable electronic device such as user device 12 to receive and upload ***voice commands*** and in using computing equipment such as ***computing equipment 14*** to process the uploaded voice commands." (EX1011, [0073].)

115. "At step 74, user device 12 may record a ***voice command***. The ***voice command*** may be recorded as an audio clip when a user pressed and releases a record button or supplies other user input directing device 12 to capture the ***voice command***. The voice command may be digitized by device 12 and stored in storage associated with user device 12 such as storage 44." (EX1011, [0065], *see also*, Abstract.) As discussed above, it would have been obvious that the ***voice command*** is captured by the ***microphones*** in device 12. (*Supra* § IX.A.4.b.)

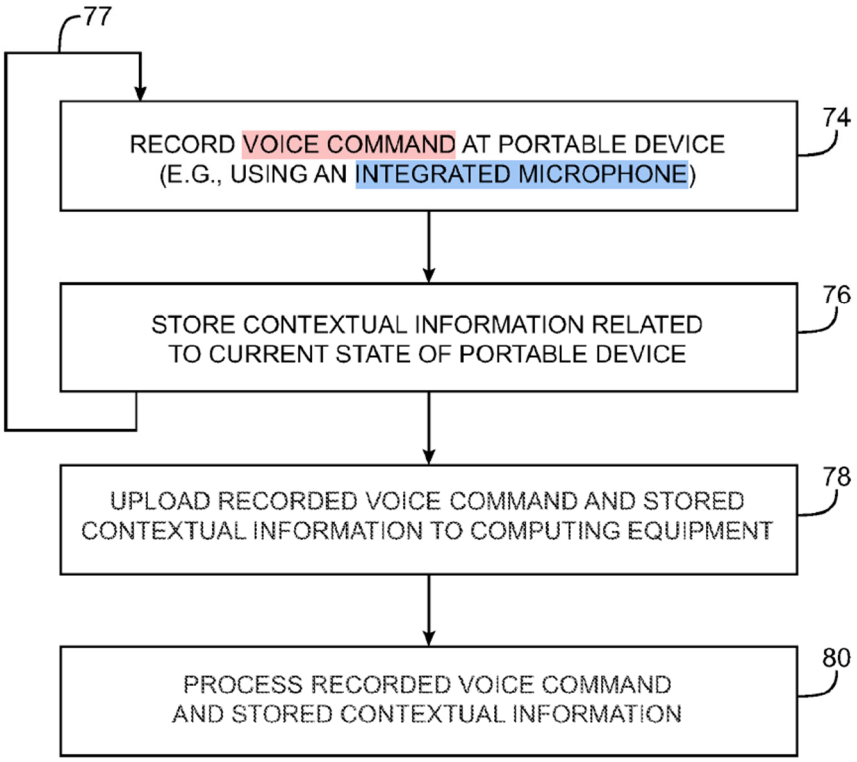


FIG. 6

116. "At step 78, user device 12 may upload recorded voice commands and stored contextual information to computing equipment such as ***equipment 14 or equipment 18***. User device 12 may upload recorded voice commands and stored contextual information to ***computing equipment 14 or equipment 18*** using any suitable communications path." (EX1011, [0068], [0070].)

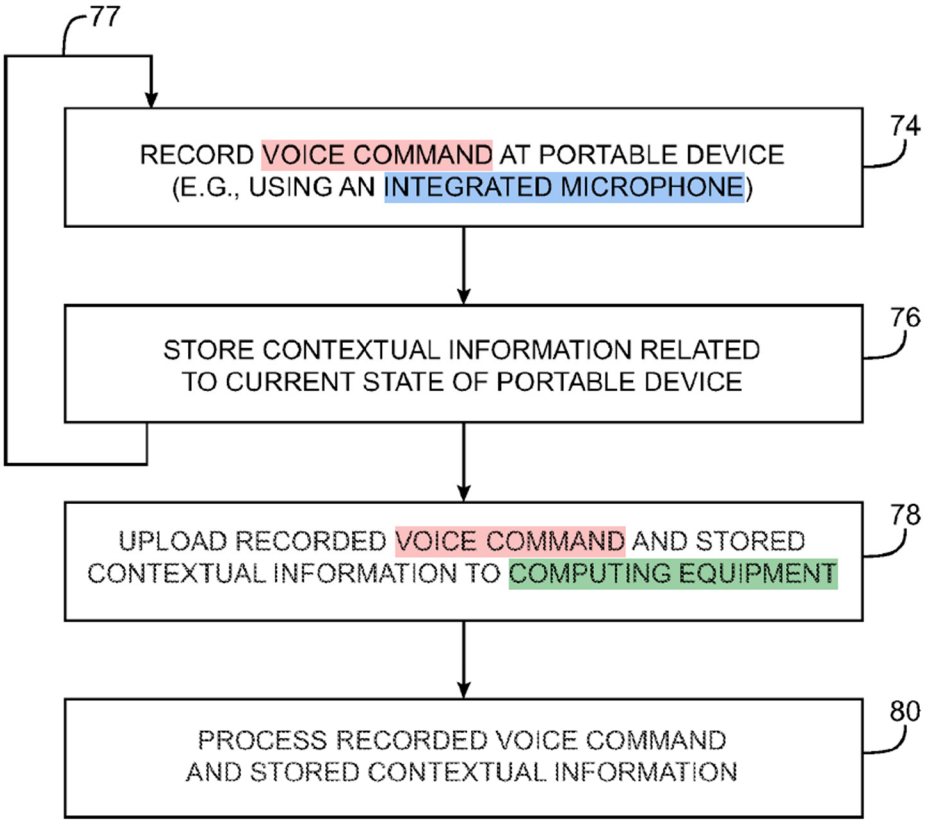


FIG. 6

117. It would have been obvious that the "***voice command***" would be at least part of the first human voice data captured by Lindhal's "***microphones***." (*Supra* § IX.A.4.b.) For example, Lindahl's multiple microphones would simultaneously capture the "***voice command***." Alternatively, at least one of Lindahl's multiple microphones captures the "***voice command***," which is part of all audio data captured by the multiple microphones ("first human voice data"). Alternatively, the "first human voice data" includes a longer audio stream or file, and only the segment

comprising the voice command (*i.e.*, "at least part of the first human voice data") is sent to the server for processing.

118. Lindahl explains the benefits for transmitting voice commands to a remote server for processing rather than processing it locally on the user device. For example, Lindahl explains that "[b]ecause ***computing equipment 14 or 18*** may include any type of computing equipment including desktop computers and ***computer servers*** which generally have relatively large amount of processing and storage capabilities compared to portable devices such as user device 12, ***computing equipment 14 or 18*** will generally have a larger dictionary of voice commands that the equipment can recognize using speech recognition operations. By uploading voice commands and contextual information from user device 12 to ***computing equipment 14 or 18***, the ***probability that a given voice command can be successfully processed and fulfilled will generally increase***." (EX1011, [0082], *see also*, [0012].)

119. Lindahl explains that "user device 12 may forward the captured ***voice command audio clip*** and, if desired, contextual information to ***computing equipment 14*** for processing. ***Computing equipment 14*** may use engine 15 to implement speech recognition capabilities that allow ***computing equipment 14*** to respond to voice commands that ***user device 12 might otherwise have difficulties in processing***. For example, if user device 12 were

to receive a voice command to '***find Italian restaurants near me***,' user device 12 might not be able to execute the voice command immediately for reasons such as an inability to perform adequate speech processing ***due to a lack of available processing power***, an inability to perform a search requested by a voice command due to a lack of network connectivity, etc. In this type of situation, device 12 may save the ***voice command*** (e.g., as a recorded audio file of a user's voice) and relevant contextual information (e.g., the current location of user device 12) for transmission to ***computing equipment 14*** for further processing of the voice command. Device 12 may transmit voice commands and contextual information to ***computing equipment 14*** at any suitable time (e.g., when device 12 is synched with ***computing equipment 14***, as the voice commands are received by device 12, whenever device 12 is connected to a communications network, etc.). These transmissions may take place simultaneously or as two separate but related transmissions." (EX1011, [0035].)

120. Further, "[b]ecause user device 12 can upload voice commands and contextual information to ***computing equipment 14 or 18***, user device 12 may be able to support an ***increased amount of voice commands and may be able to respond in a more complete manner*** than if user device 12 performed

speech recognition operations without the assistance of ***equipment 14 or 18***." (EX1011, [0084].)

121. Third, Lindahl discloses such sending of the captured first human voice data to the ***server*** is over the ***Internet via the WLAN*** by the WLAN transceiver. For example, Lindahl discloses that "***voice commands*** processing may be performed ***remotely***. For example, the electronic device may transmit one or more recorded ***voice commands*** and associated contextual information to computing equipment such as a desktop computer. Captured ***voice commands*** and contextual information may also be uploaded to ***server computing equipment*** ***over a network***." (EX1011, [0008], [0009].)

122. Referring to FIG. 1, Lindahl discloses that "the ***voice commands*** may be transmitted by user device 12 to ***computing equipment 14*** over ***communications path 20***. Voice commands may also be conveyed to ***remote services 18*** ***over network 16 (e.g., via path 21 or via path 20, equipment 14, and path 17)***." (EX1011, [0025].)

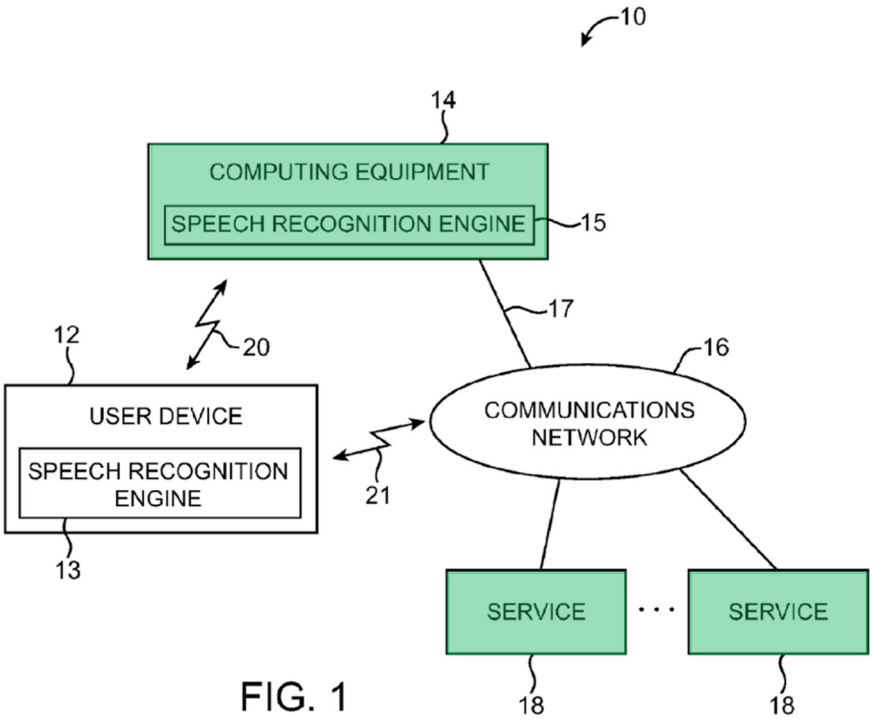


FIG. 1

123. Further, referring to FIG. 6, "[a]t step 78, user device 12 may upload recorded voice commands and stored contextual information to ***computing equipment such as equipment 14 or equipment 18***. User device 12 may upload recorded voice commands and stored contextual information to ***computing equipment 14 or equipment 18*** using ***any suitable communications path***. For example, user device 12 may transmit voice commands and contextual information to ***equipment 14*** directly over communications ***path 20***, indirectly through ***communications network 16 over paths 17 and 21***, or may upload them to ***equipment 18*** over ***network 16***." (EX1011, [0068].)

124. As discussed above, Lindahl discloses these communication paths (*e.g.*, path 20, over network 16, or via network 16 over path 17 and 21) comprise networks such as the Internet and a Wi-Fi network. (*Supra* § IX.A.4.c.) Lindahl's "[n]etwork 16 may be, for example, a local area network, a wide area network such as the ***Internet,*** a wired network, a ***wireless network***, or a network formed from multiple networks of these types." (EX1011, [0026].) "Communications network 16 may be based on any suitable communications network or networks such as a radio-frequency network, the ***Internet***, an Ethernet network, a wireless network, ***a Wi-Fi network***, a Bluetooth network, a cellular telephone network, or a ***combination of such networks***." (EX1011, [0032].) Further, Lindahl discloses that "[c]ommunications ***path 17*** and the other paths in system 10 such as ***path 20*** between device 12 and ***equipment 14***, ***path 21*** between device 12 and network 16, and the paths between network 16 and ***services 18*** may be based on any suitable wired or ***wireless communications technology***. . . . Communications paths in system 10 may, if desired, be based on ***wireless communications technology*** such as satellite technology, radio-frequency (RF) technology, wireless universal serial bus technology, and ***Wi-Fi***® or Bluetooth® 802.11 wireless link technologies." (EX1011, [0030].) Thus, Lindahl discloses

such transmission of voice commands to the computer equipment 14 or 18 (e.g., one or more Internet servers) are over the Internet via the WLAN.

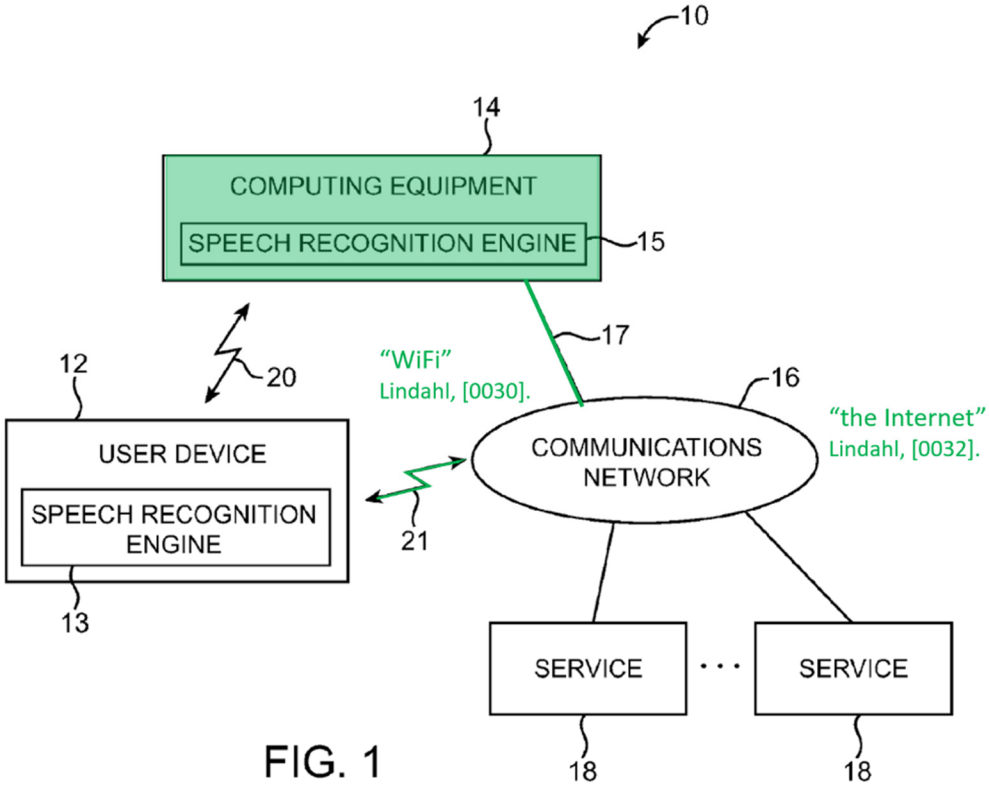


FIG. 1

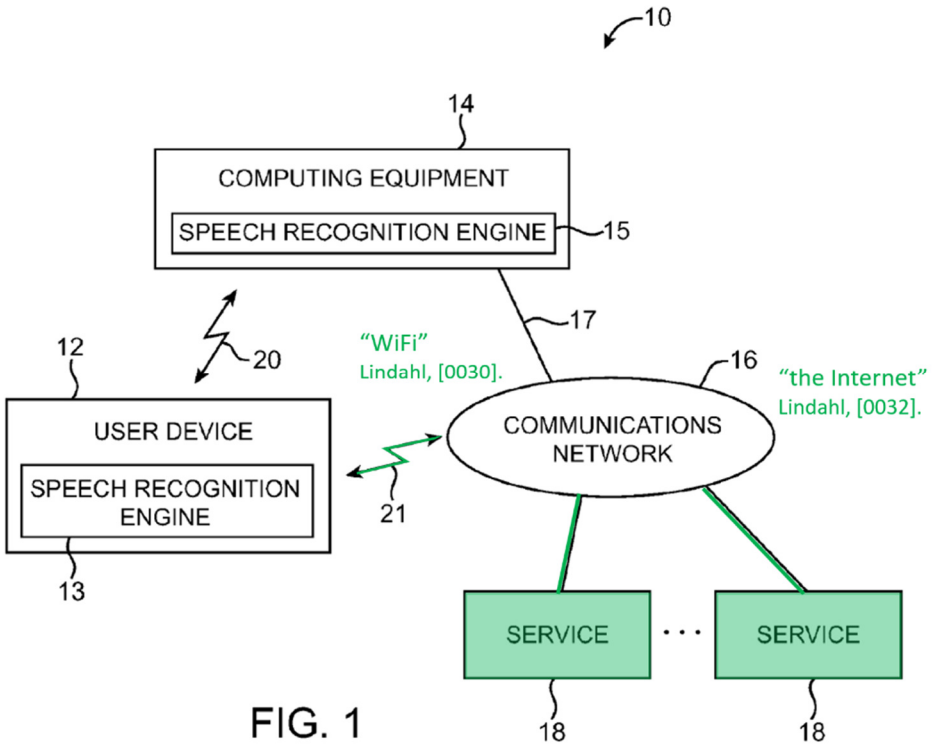


FIG. 1

125. As discussed above, Lindahl discloses a ***WLAN transceiver*** in device 12's wireless communication devices 54 "to support local and remote wireless links. Examples of local wireless links include infrared communications, ***Wi-Fi***® (IEEE 802.11)[.]" (*Supra* § IX.A.4.c; EX1011, [0052].) Thus, it would have been obvious to a POSITA that the transmission of voice commands over the Internet via the WLAN from the device to the server device is by the ***WLAN transceiver***.

126. Further, Williams also discloses this element. (*Infra* § IX.H.3.h.)

i. 1[H]: "play, by the speaker, an audio data received from the server device over the Internet via the WLAN by the WLAN transceiver;"

127. The Lindahl-Williams system discloses playing, by the ***speaker***, an ***audio data*** received from the ***server*** device ***over the Internet via the WLAN*** by the ***WLAN transceiver***.

128. First, Lindahl discloses or at least suggests receiving ***audio data*** from the ***server*** device ***over the Internet via the WLAN*** by the ***WLAN transceiver***. Lindahl discloses that after sending the audio data to the server device (i.e., computer equipment) for processing, "[t]he computing equipment may also transmit ***information*** to the electronic device that allows the electronic device to respond to the voice command." (EX1011, Abstract.)

129. For example, Lindahl discloses that "[a]fter a recorded voice command and associated contextual information have been transferred to a desktop computer, ***remote server***, or other computing equipment, the computing equipment may process the voice command using a speech recognition operation. The computing equipment may use the results of the speech recognition operation and any relevant contextual information together to respond to the voice command properly. For example, the computing equipment may respond to the voice command by displaying search results or performing other suitable actions. If desired, the computing equipment may

convey ***information*** back to the electronic device in response to the voice command." (EX1011, [0009].) It would have been obvious to a POSITA that such "information" may take the form of audio data.

130. Lindahl discloses that such transmission of ***audio data*** by the ***server device*** is ***over the Internet via the WLAN***. As discussed, the sending of voice commands to the server is over the Internet via the WLAN, it would be obvious that the receipt of audio data from the server is through the same communication path. Indeed, Lindahl discloses "[c]ommunications ***path 20*** may be used for one-way or ***two-way transmissions*** between user device 12 and ***computing equipment 14***. For example, user device 12 may transmit voice commands and contextual information to computing equipment 14. After receiving voice commands and contextual information from user device 12, ***computing equipment 14*** may process the voice commands and contextual information using a speech recognition engine such as speech recognition engine 15. Engine 15 may be provided as a standalone software component or may be integrated into a media playback application or other application. If desired, ***computing equipment 14*** may transmit ***data signals*** to user device 12. ***Equipment 14*** may, for example, transmit ***information*** to device 12 in response to voice commands transmitted by device 12 to system 14. For example, when a voice command transmitted by device 12 includes a

request to search for information, ***system 14*** may transmit ***search results*** back to device 12." (EX1011, [0031].) Lindahl discloses that communication path 20 (*i.e.*, one of the communication paths in system 10) may be based on wireless communications technology such as ***Wi-Fi***. (*Supra* §§ IX.A.4.c, IX.A.4.h; EX1011, [0030].)

131. Lindahl discloses that "***computing equipment 14 or 18*** may transmit ***information related to processing and responding*** to the voice command to user device 12. In response, user device 12 may then respond to the voice command." (EX1011, [0072].) Lindahl discloses that it is "particularly beneficial" for "***computing equipment 14 or 18*** [to] transmit ***information related to processing and responding*** to the voice command to user device 12" when the devices are connected "through long-range communications paths . . . such as the ***Internet***)." (EX1011, [0071], [0033].) It would have been obvious to a POSITA that the receipt of audio data from the server device ***over the Internet via the WLAN*** is by the ***WLAN transceiver***. (*Supra* §§IX.A.4.c, IX.A.4.h; EX1011, [0052].)

132. To the extent that Lindahl does not expressly disclose receiving audio data and playing it by the speaker (which is not conceded), Williams discloses this. A POSITA would have been motivated to combine Lindahl and Williams. (*Supra* § IX.A.3.)

133. For example, Williams discloses playing audio data received from the server device over the Internet via the WLAN by the ***speaker***. William discloses, with reference to FIG. 2, "API 214 of the application engine 118 [run on ***server 400***] may also allow the application 120 [run on ***server 400***] to provide ***audio 222*** for eventual transmission to the interface device 116, to be played on the ***speaker 206***." (EX1006, [0037].) "[T]he ***audio 222*** may be provided from the application 120 as a digital audio stream." *Id.*

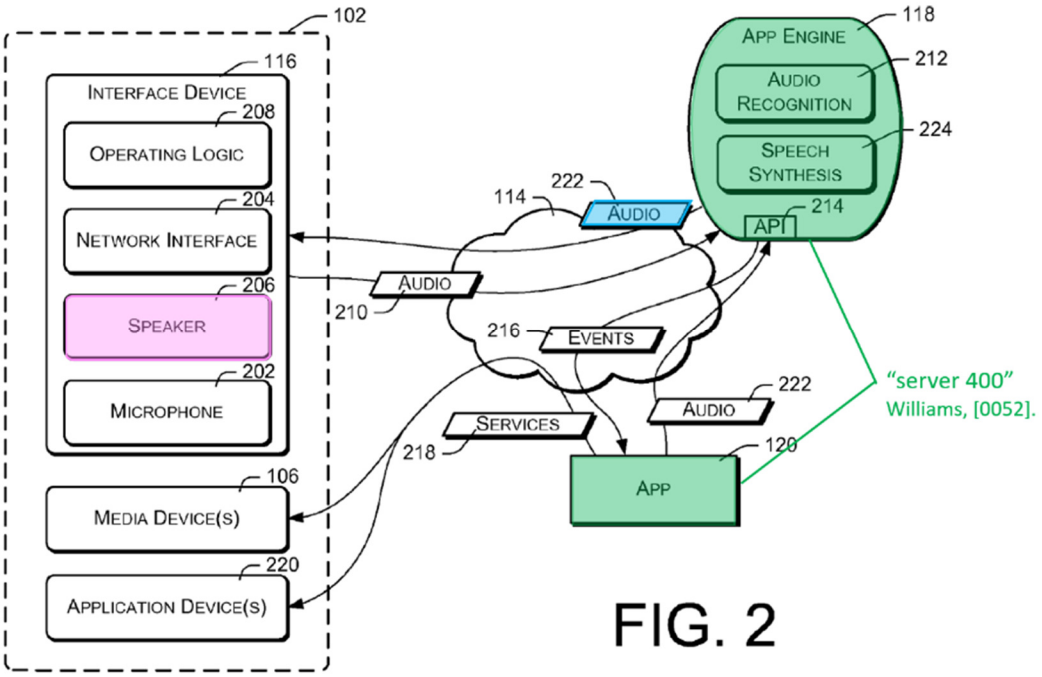


FIG. 2

134. Williams discloses several examples of providing to the user an ***audio response*** (i.e., "audio data") generated by the ***cloud server***. For example, Williams discloses that "[a]n application may provide general

information in response to users' uttered questions. A user may ask a question, such as 'What's the recipe for deviled eggs?' The application may respond by emailing or otherwise providing a recipe for deviled eggs. In some cases, the application may use the speech synthesis functionality of the system to provide ***audio answers to spoken questions***." (EX1006, [0048].) As another example, "such an application may allow the user to inquire about a flight or other event. The system might respond by sending information to the user via a device associated with the user, or by ***audibly responding*** using speech synthesis." (EX1006, [0050].) As another example, Williams discloses "a cloud-based application may allow users to order tickets without needing a computer. A user may tell the application what movie they want to see. The application may use pre-configured information about the user (such as location) to determine which theater is near the user. The application may respond to a spoken command to order movie tickets by using the speech synthesis functionality of the system to ***query the user regarding movie selections and times***." (EX1006, [0045].) A POSITA would have understood that the audio response or information associated with the audio response sent to the device disclose the "audio data".

135. Williams discloses that the audio data is received by the device from the ***server*** device ***over the Internet via the WLAN*** by the ***WLAN***

transceiver. With reference to FIG. 1, Williams discloses "the in-home devices connect to ***a router or access point 112***, which is in turn connected to ***a public network 114***. Connection between in-home devices and the router 112 may be through wired or ***wireless technologies***." (EX1006, [0012].) Thus, it would have been obvious that the audio 222 is received by the "***router or access point 112***" from the ***cloud server*** via ***the Internet*** or "***public network 114***" and then wirelessly transmitted to the interface device via the home wireless network (i.e., the WLAN). (*Id.*; *infra* §§ IX.H.3.c, IX.H.3.h.)

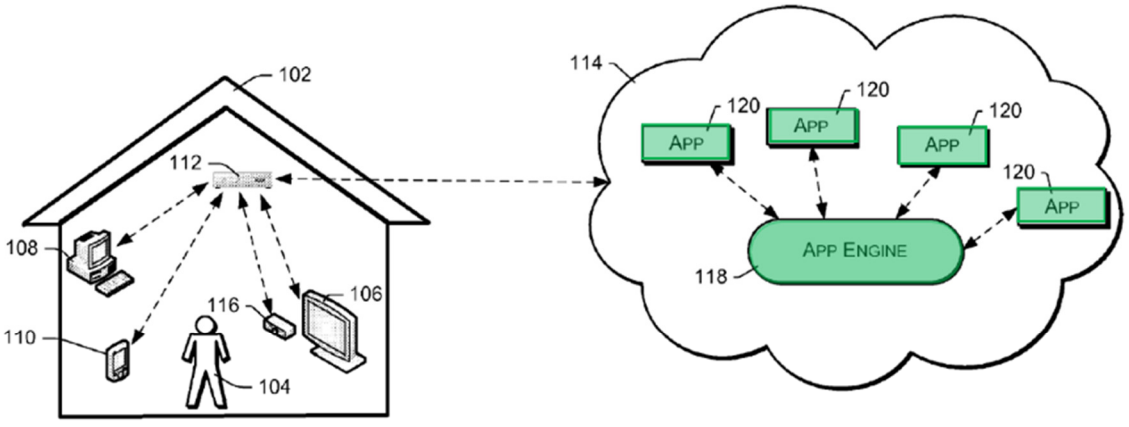


FIG. 1

136. Further, as shown in FIG. 2, audio 222 is sent by the ***cloud server*** via the ***wireless network 114*** and the interface device 116 receives the audio 222 via the ***wireless network 114***, which by itself includes both WLAN and

the Internet and thus discloses receiving audio data "via the WLAN over the Internet". (*See also Supra* §§ IX.H.3.c, IX.H.3.h; EX1006, [0015].)

137. As discussed, Williams' "network interface 204" "facilitates network communications with various entities, including the application engine 118 and one or more applications 120, via the network 114." (EX1006, [0023], [0007], [0012], [0016], [0020].) Thus, it would have been obvious that such communication over the Internet via the WLAN is via the network interface 204 which includes the ***WLAN transceiver***. (*Supra* §§IX.H.3.c, IX.H.3.h.)

138. A POSITA would have been motivated to combine Lindahl and Williams, resulting in a system where Lindahl's server (*e.g.*, computing equipment 14/18) transmits a response in the form of audio data (*e.g.*, "audio 222") to the user device for playback as taught by Williams. (*Supra* §IX.A.3.) Multi-modal feedback improves the usability of the device.

j. 1[I]: "and emit[2], by the electric light source, in response to data received from the server over the Internet via the WLAN by the WLAN transceiver,"

139. Lindahl-Williams discloses this element.

[2] I note that the limitation appears to be missing the word "light" after "emit[.]"

140. First, as discussed above, Lindahl discloses receiving ***data*** from the ***server*** device (*e.g.*, ***computing equipment 14 or 18***) over the Internet via the WLAN by the WLAN transceiver. (*Supra* § IX.A.4.i; EX1011, Abstract, [0009], [0072], [0031], [0030], [0052].) For example, Lindahl discloses that "***computing equipment 14 or 18*** may transmit ***information related to processing and responding*** to the voice command to user device 12. In response, user device 12 may then respond to the voice command. This type of arrangement may be particularly beneficial when user device 12 and the computing equipment are not physically located near each other (e.g., when user device 12 is only connected to ***computing equipment 14 or 18*** through long-range communications paths such as through a communications network such as the ***Internet***)." (EX1011, [0072].) As discussed above, Lindahl's device receives data from the server device ***over the Internet via the WLAN***. (*Supra* § IX.A.4.i; EX1011, Abstract, [0009], [0072], [0031], [0030].) It would be obvious that the receipt of data is by the ***WLAN transceiver***. (*Supra* § IX.A.4.c; EX1011, [0052].)

141. Second, Lindahl discloses emitting by the electric light source in response to the data received from the server over the Internet via the WLAN by the WLAN transceiver. As discussed above, Lindahl discloses ***electric light sources*** in user device 12. (*Supra* §IX.A.4.e; EX1011, [0040], [0048].)

Lindahl discloses that "[d]evice 12 may receive ***data signals*** from ***computing equipment 14*** in response to uploading voice commands and contextual information. The ***data*** received by device 12 from ***equipment 14*** in response to voice commands and contextual information may be used by device 12 to carry out requests associated with the voice commands. For example, after processing the voice command and contextual information, ***computing equipment 14*** may transmit ***results associated with the voice command*** to user device 12 which may then ***display*** the ***results***." (EX1011, [0043].) Thus, the displaying of the results is in response to Lindahl's user device receiving the results.

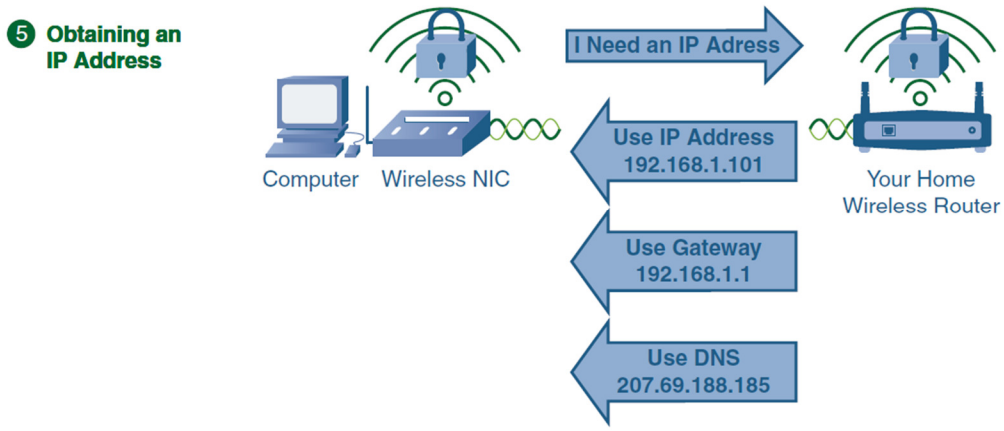
142. A POSITA would have understood that displaying the results, as taught by EX1011, involves emitting light by the electric light source. For example, in EX1011, "[d]isplay and audio devices 52 may include ***liquid crystal display (LCD) screens or other screens***, ***light-emitting diodes (LEDs)***, and other components that present visual information and status data." (EX1011, [0048].) Thus, displaying the "***results associated with the voice command***" would involve emitting light by the ***LCD screen*** or by the ***LEDs***.

k. 1[J]: "wherein the device is addressed in the WLAN using an Internet Protocol (IP) address that uniquely identifies the device in the WLAN or in the Internet."

143. Lindahl-Williams discloses or at least suggests that the device is addressed in the WLAN using an Internet Protocol (IP) address that uniquely identifies the device in the WLAN or in the Internet.

144. As discussed, Lindahl discloses that its device comprises a WLAN transceiver and operates in a WLAN, which is connected to the Internet. (*Supra* §IX.A.4.c.) A POSITA would have understood that this wireless network involves associating each device with a device identifier or address that uniquely identifies the device in the WLAN. Indeed, at the time of the invention, it was well-known that networked devices are associated with IP addresses. The IP address may be an assigned IP address for the device (*e.g.*, by the router), the router's IP address, a combination of the two, or a globally unique IP address. "***Each computer or device that you want to join your wireless network needs a wireless network interface card (NIC)***." (EX1007, 99.) "If the wireless router grants access, it then typically ***assigns the NIC an IP address*** to use on the wireless network, at which point we can begin exchanging data," "***each device (computer, laptop, and so on) in a home network needs to have an IP address***, which is how the device is known on the network . . . Most home networks use IP addresses in the range of 192.168.1.100 to 192.168.1.149, and the ***router assigns each device an address*** from this pool when it boots." (*Id.*, 108, 210, 340, *see also*, *id.*,

175.) These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to communicate with devices outside of the user's WLAN. (EX1007, 340.)



145. Further, Williams also discloses this element. (*Infra* § IX.H.3.k.)

5. Claim 2: "The device according to claim 1, further comprising, in the enclosure, a rechargeable battery for powering the device."

146. Lindahl-Williams discloses or at least suggests this claim. Referring to FIG. 3, Lindahl discloses that "user device 12 may include storage 44. Storage 44 may include one or more different types of storage such as . . . volatile memory (e.g., ***battery-based*** static or dynamic random-

access-memory)[.]" (EX1011, [0045].) Thus, Lindahl discloses its device comprises a battery.

147. As shown in FIG. 2, Lindahl discloses that "[u]ser device 12 may be a mobile telephone." (EX1011, [0044], [0027].) In light of this disclosure, a POSITA would have understood that Lindahl's handheld mobile phone comprises a rechargeable battery within its enclosure such that it can be repeatedly used without being tethered to a power source. Indeed, by 2012, mobile phones such as iPhones were commonly equipped with a rechargeable battery. (EX1008, 34.)

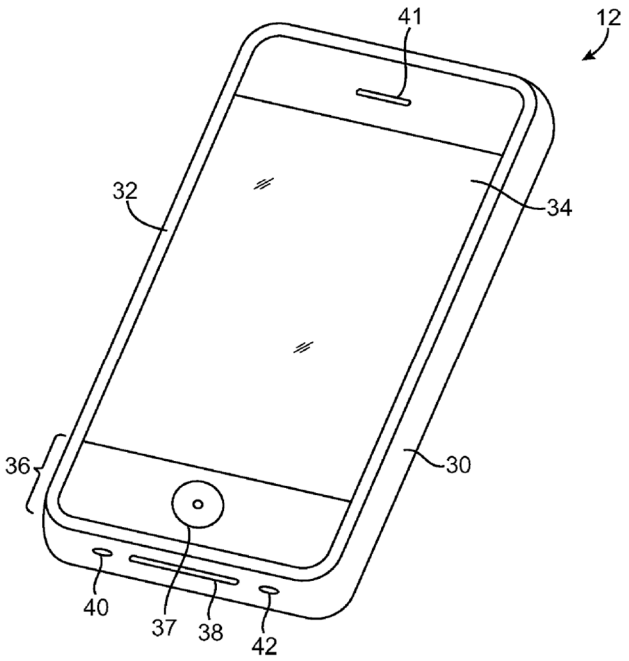


FIG. 2

6. Claim 4: "The device according to claim 1, further being addressable using the IP address in the WLAN or in the Internet using distinct locally administered addresses or a universally administered digital addresses."

148. Lindahl-Williams discloses or at least suggests this claim. As discussed, Lindahl discloses or at least suggests that the device is addressed in the WLAN using an IP address that uniquely identifies the device in the WLAN or in the Internet. (*Supra* § IX.A.4.k.) It was well-known that networked devices are associated with IP addresses. The IP address may be an assigned IP address for the device (*e.g.*, by the router), the router's IP address, a combination of the two, or a globally unique IP address. (EX1007, 99, 108, 175, 210, 340.) These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to communicate with devices outside the user's WLAN. *Id*. A POSITA would have understood the router assigned IP addresses to be locally administered addresses. "Most home networks use IP addresses in the range of 192.168.1.100 to 192.168.1.149, and the ***router assigns each device an address*** from this pool when it boots." (EX1007, 108, 210, 175.)

149. When communicating with Internet servers using NAT, the device may utilize the router's globally unique IP address. "Network Address Translation (NAT)[:] The technology by which a home router translates the private IP addresses used by computers on your home network into a single ***public IP address*** assigned to your broadband modem, providing the ability for multiple computers to share a connection, and also providing a degree of privacy because the computers on your home network are not able to be accessed directly from the Internet." (EX1007, 340). A POSITA would have understood that these globally unique IP addresses are universally administered.

150. Additionally or alternatively, Lindahl discloses or at least suggests that the device has a MAC address that is a universally administered digital address. It was common knowledge, and a POSITA would have understood that communicating via a wireless network and with a remote server over the Internet via a wireless router (as in Lindahl) involves associating the device with a device identifier or address such as a MAC address that uniquely identifies the device in the network. (*Supra* §§ IX.A.4.c, IX.A.4.h; EX1011, [0025], [0026], [0030], [0032].) "***Each computer or device that you want to join your wireless network needs a wireless network interface card (NIC)***," and "***each wireless card has a***

unique identifier called a MAC address." (EX1007, 99, 166.) Each networked device is thus associated with a globally unique MAC address for identification of the device in the network. This unique identifier is assigned to the NIC at its manufacture and is thus a universally administered address (locally stored in the non-volatile memory). "***[A] MAC address*** . . . is a 48-bit binary number generally written as 12 hexadecimal digits (six groups of two digits, the groups separated by dashes or colons). The address is ***set at the time of the NIC's manufacture*** . . . The second least significant bit of the most significant byte MAC address is known as the ***universal or locally administered address (U/L) bit*** . If this bit is set to 0, the address is administered universally. This means that the address was assigned by the Institute of Electrical and Electronics Engineers (IEEE) and is unique across the globe. If the bit is set to 1, it indicates that the address was selected locally by a network administrator (overriding the NIC's BIA)." (EX1023, 24.)

151. Furthermore, both the universally administered MAC address and the locally administered IP address are used in conjunction with each other. "On the DLC header, the data is functioning on Layer 2, so a MAC address is needed, as we saw. However, for the IP header, you have a set of IP addresses that should reinforce the fact you learned in Chapter 1—that IP functions on

Layer 3 (the network layer) of the OSI model." (EX1023, 218, *see also id.,* 312 ("ARP allows IP-enabled devices on your network to dynamically map physical (MAC) addresses to IP addresses"), 313, Figure 6.6.) Thus, a POSITA would have understood (and found it obvious) that the device may comprise a MAC address that uniquely identifies the device globally – both in the WLAN and in the Internet, and that the MAC address may be administered universally, and that the device also is addressable using an IP address in the WLAN.

7. Claim 5: "The device according to claim 4, wherein the IP address is a layer 3 address and is static or dynamic IP address that is IPv4 or IPv6 type address."

152. Lindahl-Williams discloses or at least suggests this claim. As discussed above, Lindahl discloses that the device (*e.g.*, user device 12 or interface device 116) is addressable in the WLAN or in the Internet using an IP address. (*Supra* §§ IX.A.4.c, IX.A.4.k; EX1011, [0025], [0026], [0030], [0032]; EX1006, [0012], [0020], [0023], FIG. 1; EX1007, 99, 108, 210, 175.)

153. This claim describes well-known knowledge about IP addresses at the time of the invention. At the time of the invention, it was well-known that Internet Protocol (IP) addresses are either static or dynamic, are Layer 3 addresses and are IPv4 addresses. "In the case of today's Internet, these addresses are Internet protocol version 4 (IPv4) addresses." (EX1029, 127.)

"Internet protocol (IP) addresses are the unique numbers used to identify individual connections to the Internet." *Id.* "Internet Protocol (IP) is a Layer 3 protocol that contains addressing and control information that allows packets to be routed." (EX1023, 11.) "The IP system in common use today is known as IPv4, for Internet Protocol version 4. A newer system, IPv6, or Internet Protocol version 6, has been developed and exists today in small deployments." (*Id*., 13.) "There are currently two main versions of IP and their associated IP address families; ***IPv4 and IPv6***. IP addresses can be ***either static or dynamic*** [.]" (EX1024, 2:60-62.) "The broadband networks typically assign a *dynamic IP address* to your broadband service. Dynamic means the IP address can change[.]" (EX1007, 277.) Further, an IP address may be a "*static IP address*, which means . . . your address never changes." *Id.*

8. Claim 6: "The device according to claim 1, further comprising a volatile or non-volatile memory that stores the IP address."

154. Lindahl-Williams discloses this claim.

155. For example, with reference to FIG. 3 below, Lindahl discloses that "use device 12 may include storage 44. Storage 44 may include one or more different types of storage such as hard disk drive storage, ***nonvolatile memory*** (e.g., flash memory or other electrically-programmable-read-only

memory), ***volatile memory*** (e.g., battery-based static or dynamic random-access-memory), etc. Storage 44 may be used to store voice commands and contextual information about the state of device 12 when voice commands are received." (EX1011, [0045].)

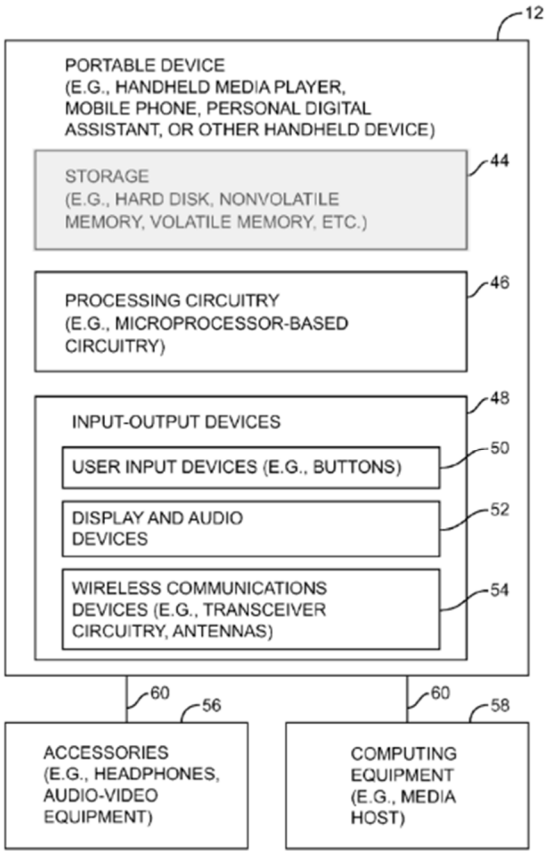
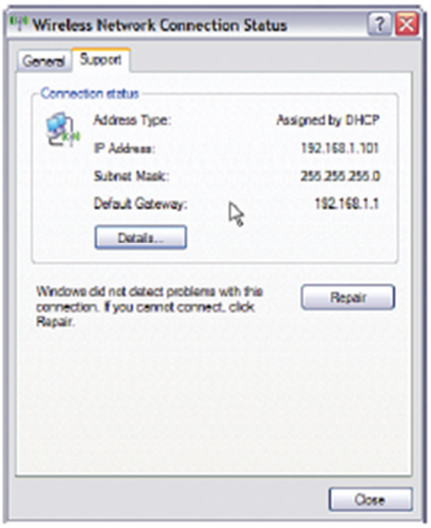


FIG. 3

156. As discussed above, Lindahl discloses that the user device 12 is addressable in the Internet or WLAN with an IP address and that it communicates with a server device over the Internet via the WLAN via a wireless router. (*Supra* §§ IX.A.4.c, IX.A.4.k.) A POSITA would have understood that the IP address would be locally stored in the disclosed

memory in the device to enable wireless connectivity. For example, as shown in the figure below in Doherty, a textbook titled "Wireless Home Networking Simplified", the networked devices can display the associated IP addresses, which means the IP addresses are locally stored in the device. (EX1007, 211, Figure 10-44, 265.) Thus, a POSITA would have understood that the "nonviolatile memory" and "volatile memory" in "storage 44" as disclosed in Lindahl's user device 12 locally stores the IP addresses.

Figure 10-44 Verifying IP Address Assignment with Windows XP



157. Moreover, Lindahl discloses that "[s]torage 44 may be used to store voice commands and ***contextual information about the state of device 12*** when voice commands are received." (EX1011, [0045].) Lindahl discloses that "[t]he contextual information may include any information that is available about the current state of the user device. For example, the

contextual information may include information related to . . . a current web-browsing operation (e.g., a current web-address), the geographic location of the user device (e.g., a location determined using a location sensor, a location derived from ***information associated with communications path 20 and 21*** such as which cellular telephone network or other network the device is connected to, or location data manually entered by a user)." (EX1011, [0060].) A POSITA would have understood, and it would have been obvious that the contextual information may include "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].)

9. Claim 7: "The device according to claim 1, wherein the speaker comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

158. A POSITA would have understood that this claim recites a standard speaker, and the patent confirms this understanding. (EX1001, 105:48-106:62.) Lindahl-Williams discloses this claim. For example, Lindahl discloses a speaker in the device for playing audio data. (*Supra* §§ IX.A.4.d, IX.A.4.i; EX1011, [0048], [0040], FIGs. 2, 3.) A POSITA would

have understood that an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker, a ribbon or planar magnetic loudspeaker are standard and common implementations of speakers. (EX1014, 321-322, 763.) Thus, a POSITA would have understood (or it would be obvious) that Lindahl's speakers are the recited speakers in claim 7.

10. Claim 8: "The device according to claim 1, wherein the speaker is operative for omnidirectional, unidirectional, or bidirectional pattern audio emitting."

159. A POSITA would have understood that this claim recites a standard speaker, and the patent confirms this understanding. (EX1001, 105:48-106:62.) Lindahl-Williams discloses this claim. For example, as discussed, Lindahl discloses a speaker in the device for playing audio data. (*Supra* §§ IX.A.4.d, IX.A.4.i; EX1011, [0048], [0040], FIGs. 2, 3.) A POSITA would have understood that a standard speaker is operative for omnidirectional pattern audio emitting. (EX1018, 122; EX1014, 763, 765.) Thus, a POSITA would understand (or it would be obvious) that Lindahl's speakers would operate as claimed.

11. Claim 11: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, an omnidirectional, unidirectional, or bidirectional microphone that is based on sensing of an incident sound-based motion of a diaphragm or a ribbon."

160. A POSITA would have understood that this claim recites standard microphones, and the patent confirms this understanding. (EX1001,

Claim 9, 17:47-54.) Lindahl-Williams discloses this claim. For example, as discussed, Lindahl discloses its user device 12 comprises user input devices 50 such as ***microphones***. (EX1011, [0048], *see also, id*., [0006], [0040].) A POSITA would have understood that a standard microphone is an omnidirectional microphone based on sensing an incident sound-based motion of a diaphragm. (EX1014, 757, 309.) Thus, a POSITA would understand (or it would be obvious) that each of Lindahl's standard microphones is the recited microphone in claim 11.

12. Claim 12: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

161. A POSITA would have understood that this claim recites standard microphones, and the patent confirms this understanding. (EX1001, Claim 9, 17:47-54.) Lindahl-Williams discloses this claim. For example, as discussed, Lindahl discloses its user device 12 comprises user input devices 50 such as ***microphones***. (EX1011, [0048], *see also, id*., [0006], [0040].) A POSITA would understand that dynamic microphone and piezoelectric microphone are standard and common implementations of microphones. (EX1014, 758.) Thus, a POSITA would understand (or it would be obvious) that each of Lindahl's standard microphones is/includes the recited microphone in claim 12.

13. Claim 13: "The device according to claim 1, wherein the light is a visible light for illumination or indication."

162. Lindahl-Williams discloses this claim. For example, as discussed, Lindahl discloses "***liquid crystal display (LCD) screens*** or other screens, ***light-emitting diodes (LEDs)***, and other components that present visual information and status data" in user device 12. (*Supra* § IX.A.4.e; EX1011, [0048].) As discussed, the '720 patent explains that a display may include an array of light emitters and that an LCD display may include LED backlit lights. (EX1001, 109:45-50; 110:64-111:3, 111:11-23.) Thus, a POSITA would understand, and the '720 patent confirms that display screens and LEDs are or include electric light sources.

163. Lindahl's light emitting diodes (LEDs) and light emitters/LED backlit lights in the LCD display screen emit visible light for illumination, consistent with the description of the '720 patent. "In one example, the device ***illuminates a visible light***, such as ***a Light-Emitting-Diode (LED)*** . . . Further, the illumination can be directed for lighting a surface, such as a surface including an image or a picture." (EX1001, 89:50-57.)

14. Claim 14: "The device according to claim 1, wherein the light is a non-visible light that is infrared, ultraviolet, X-rays, or gamma rays."

164. Lindahl-Williams discloses this claim. For example, as discussed, Lindahl discloses "***liquid crystal display (LCD) screens*** or other

screens, ***light-emitting diodes (LEDs)***, and other components that present visual information and status data" in user device 12. (*Supra* §§IX.A.4.e, IX.A.13; EX1011, [0048].) As discussed, LCD screens use LED backlighting. (*Supra* §IX.A.4.e; EX1001, 110:64-111:3, 111:11-23.) A POSITA would understand (and it would be obvious) that LCDs emit non-visible infrared radiation as ordinary thermal emission from the heated backlight and panel. Generally, objects that generate heat, including LEDs and LCDs, will emit thermal radiation, that is, some energy in the infrared spectrum. "It was mentioned above that in ***any object*** every atom and every molecule vibrate. The average kinetic energy of vibrating particles is represented by the temperature. . . . The electromagnetic radiation associated with heat is called ***thermal radiation***. Figure 3.41 shows the total electromagnetic radiation spectrum spreading from γ rays to radio waves. Thermal radiation is predominantly situated in the ***mid and far-infrared (IR)*** spectral ranges." (EX1030, 126.)

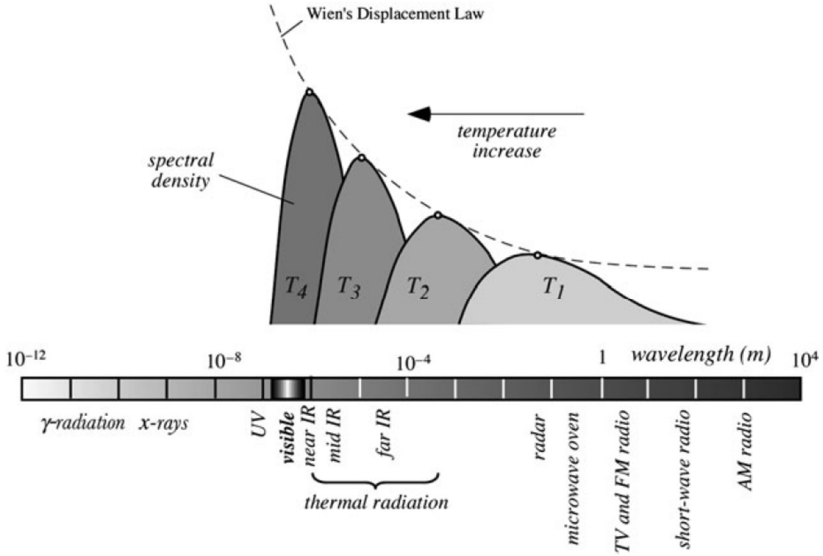


Fig. 3.41 Spectrum of electromagnetic radiation

15. Claim 15: "The device according to claim 1, wherein the electric light source consists of, or comprises, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode."

165. Lindahl-Williams discloses this claim. As discussed above, Lindahl's user device 12 comprises "***liquid crystal display (LCD) screens*** or other screens, ***light-emitting diodes (LEDs)***, and other components that present visual information and status data." (*Supra* §§IX.A.4.e, IX.A.13; EX1011, [0048].)

16. Claim 16: "The device according to claim 1, further comprising a memory that stores digital audio content, and the one or more processors are further programmed with computer program instructions that, when executed, cause the device to playing by the speaker the stored digital audio content."

166. Lindahl-Williams discloses this claim. Lindahl's device has "music player functionality" and may perform a "media playback operation"

to play "a music file or a video file". (EX1011, [0024], [0027], [0034], [0061], [0042].) A POSITA would understand (and it would be obvious) that Lindahl's device stores the media files in its local memory and Lindahl's "processing circuitry 42" are programmed with instructions that plays the media files on its speaker when executed. (EX1011, [0046]; EX1008, 60-61.)

167. Additionally or alternatively, Williams discloses that "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' speaker "us[es] the ***speech synthesis functionality*** of the system to ***query the user regarding movie selections and times***". (EX1006, [0037], [0045]-[0046] and [0050].) A POSITA would find it obvious that Lindahl-Williams, which receives audio 222, discloses this claim.

17. Claim 17: "The device according to claim 16, wherein the digital audio content is pre-recorded or synthesized audio content."

168. Lindahl-Williams discloses this claim. For example, Lindahl's device has "music player functionality" and may perform a "media playback operation" to play "a music file or a video file". (EX1011, [0024], [0034], [0061], [0042].) A POSITA would have understood, and it would be obvious that music files or video files are pre-recorded.

169. Additionally or alternatively, Williams discloses that "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' speaker "us[es] the ***speech synthesis functionality*** of the system to ***query the user regarding movie selections and times***". (EX1006, [0037], [0045]-[0046] and [0050].) A POSITA would find it obvious that Lindahl-Williams, which receives audio 222, discloses this claim.

18. Claim 18: "The device according to claim 16, further operative to play a simulating of a voice of a human being or to play music."

170. Lindahl-Williams discloses this claim. For example, as discussed above, Lindahl discloses or at least suggests storing music or video files and playing the media/music, thus, Lindahl's device is operative to play music. (EX1011, [0024], [0027], [0034], [0061], [0042].)

171. Additionally or alternatively, Williams discloses that "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' speaker "us[es] the ***speech synthesis functionality*** of the system to ***query the user regarding movie selections and times***". (EX1006, [0037], [0045]-[0046] and [0050].) A POSITA would find it obvious that Lindahl-Williams, which receives audio 222, discloses this claim.

19. Claim 19: "The device according to claim 18, wherein the digital audio content comprises a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

172. Lindahl-Williams discloses this claim. For example, as discussed above, Lindahl discloses or at least suggests storing music or video files and playing the media/music, thus, Lindahl's device is operative to play music. (EX1011, [0024], [0027], [0034], [0061], [0042].) A POSITA would have understood, and it would be obvious, that music and video files comprises a syllable, a word, a phrase, a sentence using male or female voice.

173. Additionally or alternatively, Williams discloses that the server may "***query*** the user regarding ***movie selections and time***" and "provide ***audio answers to spoken questions***" and "audibly respond[] using ***speech*** synthesis" in response to user queries. (*Supra* § IX.A.4.i; EX1006, [0046], [0048], [0050], [0037].) Thus, Williams discloses, and a POSITA would have found it obvious, that the digital audio content (*e.g.*, audio 222) comprises at least a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice.

20. Claim 20: "The device according to claim 1, wherein the WLAN is according to, or compatible with, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard."

174. Lindahl discloses or at least suggests this claim. As discussed, Lindahl discloses that device 12 communicates with ***computing equipment 14***

and or 18 (*i.e.*, uploading the voice commands and receiving information related to speech processing) over the Internet (*i.e.*, network 16), and via WiFi, a form of WLAN, directly or though a gateway device such as a router. (*Supra* §§ IX.A.4.c, IX.A.4.h; EX1011, [0025], [0032], [0026], [0030].) "Network 16 may be, for example, a local area network, a wide area network such as ***the Internet***, a wired network, a wireless network . . . User device 12 may transmit voice commands and contextual information to ***computing equipment 14*** through ***communications network 16***. User device 12 may also transmit voice commands and contextual information to ***computing equipment 14*** directly via ***communications path 20***." (EX1011, [0026].) Lindahl discloses that "[c]ommunications paths in system 10 [which includes, path 20, path 21, and path 17] may, if desired, be based on ***wireless communications technology*** such as satellite technology, radio-frequency (RF) technology, wireless universal serial bus technology, and ***Wi-Fi*** ® or Bluetooth® 802.11 wireless link technologies." (EX1011, [0030].)

175. A POSITA would have understood that WiFi implements a WLAN and is based on the IEEE 802.11 standards which include the 802.11a/b/g family. "Wireless local area networks (***WLANs***) are built around the family of IEEE 802.11 standards . . . The 802.11a/b/g family of standards are collectively referred to as ***Wi-Fi***, for wireless fidelity."

(EX1042, 600-601.) "***Wi-Fi*** is a registered trademark owned by the Wi-Fi Alliance, and it refers to ***wireless local area network (WLAN)*** products that are based on the Institute of Electrical and Electronics Engineers' (IEEE) 802.11 standards." (EX1013, 7:48-51.) This is consistent with the description of the '720 patent which states that "[i]n order to allow interoperability between equipment manufactured by different vendors, few ***WLAN standards*** have evolved, as part of the ***IEEE 802.11 standard group***, ***branded as WiFi*** (www.wi-fi.org). IEEE 802.11b describes a communication using the 2.4 GHz frequency band and supporting communication rate of 11 Mb/s, IEEE 802.11a uses the 5 GHz frequency band to carry 54 MB/s and IEEE 802.11g uses the 2.4 GHz band to support 54 Mb/s." (EX1001, 8:25-32.)

21. Claim 21: "The device according to claim 1, wherein the WLAN uses a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

176. As discussed, Lindahl-Williams discloses or at least suggests the WLAN is according to, or compatible with, IEEE 802.11b standard. (*Supra* §IX.A.20.) A POSITA would understand that WLAN over 802.11b is a form of ISM Band II frequency. (EX1042, 22-23, Table 1.2.)

22. Claim 22: "The device according to claim 1, wherein the enclosure is a handheld enclosure."

177. Lindahl-Williams discloses this claim. As discussed, Lindahl discloses a portable electronic device. (*Supra* §§IX.A.4.a, IX.A.5.) Lindahl discloses that the portable electronic device may be in the form of a handheld device such as a cellular telephone. For example, Lindahl discloses that "[u]ser device 12 may have any suitable form factor. For example, user device 12 may be provided in the form of a ***handheld device***, desktop device, or even integrated as part of a larger structure such as a table or wall. With one particularly suitable arrangement, which is sometimes described herein as an example, user device 12 may be provided with a ***handheld form factor***. For example, device 12 may be a ***handheld electronic device***. Illustrative handheld electronic devices that may be provided with voice command recording capabilities include ***cellular telephones***, media players, media players with wireless communications capabilities, handheld computers (also sometimes called personal digital assistants), global positioning system (GPS) devices, handheld gaming devices, and other handheld devices." (EX1011, [0027].)

23. Claim 23: "The device according to claim 1, further comprising in the enclosure an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, the actuator that is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN."

178. Lindahl discloses multiple speakers in the device. Lindahl discloses that "[d]isplay and audio devices 52 may also include audio equipment such as ***speakers and other devices for creating sound***." (EX1011, [0048].) Thus, Lindahl discloses comprising a speaker or speaker system (in addition to the speaker discussed with reference to claim 1) in the enclosure.

179. A POSITA would have understood that the speaker or speaker system is an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, namely, audio, or sound. For example, the actuator can convert electrical energy into sound waves transmitted through the air through the movement of a diaphragm, often cone-shaped or dome-shaped, which produces "sound" or acoustical energy. "Most loudspeakers use the moving-coil principle (Figure 10.8). The coil is wound on a former at the apex of a stiff diaphragm which is held in flexible supports so that the coil is the intense field of a permanent magnet. A signal current through the coil will then ***vibrate the diaphragm, producing sound waves*** at the same frequency and with the same waveshape as the signal." (EX1018, 122.)

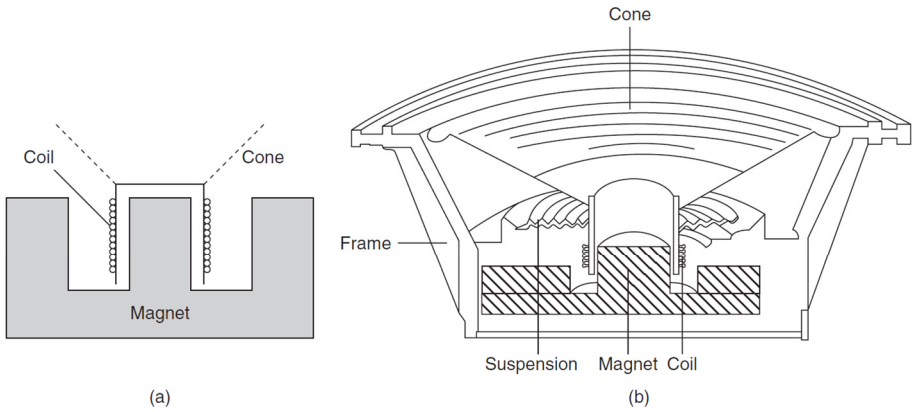


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

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

181. The specification of the '720 patent states that a speaker is an actuator for producing sound. "An actuator may be an ***electroacoustic actuator, such as a sounder*** which converts electrical energy to sound waves transmitted through the air, an elastic solid material, or a liquid, usually by means of a vibrating or moving ribbon or diaphragm." (EX1001, 105:51-55.) "A ***loudspeaker (a.k.a. speaker) is a sounder*** that produces sound in response to an electrical audio signal input, typically audible sound. The most common form of loudspeaker is the electromagnetic (or dynamic) type, uses a paper cone supporting a moving voice coil electromagnet acting on a permanent

magnet. Where accurate reproduction of sound is required, multiple loudspeakers may be used, each reproducing a part of the audible frequency range. A loudspeaker is commonly optimized for middle frequencies; tweeters for high frequencies; and sometimes supertweeter is used which is optimized for the highest audible frequencies." (EX1001, 105:64-106:7.)

182. As discussed below, a speaker or speaker system comprises a digital-to-analog converter ("DAC"), which is an actuator. (*Infra* §IX.A.26.) The speaker system is housed in the enclosure.

183. It would have been obvious that each speaker/speaker system is coupled to the WLAN transceiver and controlled in response to data received from the WLAN in Lindahl-Williams. (*Supra* §IX.A.4.i.)

24. Claim 24: "The device according to claim 23, wherein the actuator consists of, or comprises, a motion actuator that causes linear or rotary motion."

184. As discussed, Lindahl discloses multiple speakers in the device 12. (*Supra* § IX.A.23; EX1011, [0048].) A POSITA would have understood that a speaker is a motion actuator that causes "linear motion." First, a speaker converts electrical energy to mechanical energy. "In a manner reminiscent of microphones, few loudspeakers convert directly between electrical and acoustical energy. Instead, the ***electrical energy is converted to mechanical energy, specifically the movement of a diaphragm***,

often cone-shaped or dome-shaped, which then radiates acoustical energy, sometimes aided by an acoustical coupling horn or lens." (EX1014, 763.) As loudspeakers convert electrical energy to mechanical energy resulting in movement of a diaphragm, it is clear that loudspeakers are motion actuators.

185. Second, a speaker moves linearly. "[I]n the low-frequency range, the ***linear motion of an electrodynamical loudspeaker***, for example, can be described by a single-degree-of-freedom oscillator." (EX1014, 903.) Further, a dynamic or moving coil loudspeaker moves axially. "Ernst W. Siemens was the first to describe the ***dynamic or moving-coil loudspeaker***, with a circular coil of wire in a magnetic field and supported so that ***it could move axially***." (EX1014, 18.) Thus, an electrodynamical loudspeaker is a motion actuator that causes linear motion.

25. Claim 25: "The device according to claim 23, wherein the actuator consists of, or comprises, an electromagnetic coil or an electromagnet operative for generating a magnetic or electric field."

186. As discussed, Lindahl discloses multiple speakers in the device 12. (*Supra* § IX.A.23; EX1011, [0048].) A POSITA would have understood that a speaker commonly comprises, an electromagnetic coil operative for generating a magnetic or electric field, consistent with the specification of the '720 patent. The'720 patent states, "[a] loudspeaker (a.k.a. speaker) is a sounder that produces sound in response to an electrical

audio signal input, typically audible sound. The most common form of loudspeaker is the electromagnetic (or dynamic) type, uses a paper cone supporting a moving ***voice coil electromagnet acting on a permanent magnet***. Where accurate reproduction of sound is required, multiple loudspeakers may be used, each reproducing a part of the audible frequency range. A loudspeaker is commonly optimized for middle frequencies; tweeters for high frequencies; and sometimes supertweeter is used which is optimized for the highest audible frequencies." (EX1001, 105:64-106:7.)

187. "Most loudspeakers use the ***moving-coil*** principle (Figure 10.8). The coil is wound on a former at the apex of a stiff diaphragm which is held in flexible supports so that the ***coil is the intense field of a permanent magnet***. A signal current through the coil will then vibrate the diaphragm, producing sound waves at the same frequency and with the same waveshape as the signal." (EX1018, 122, 123.)

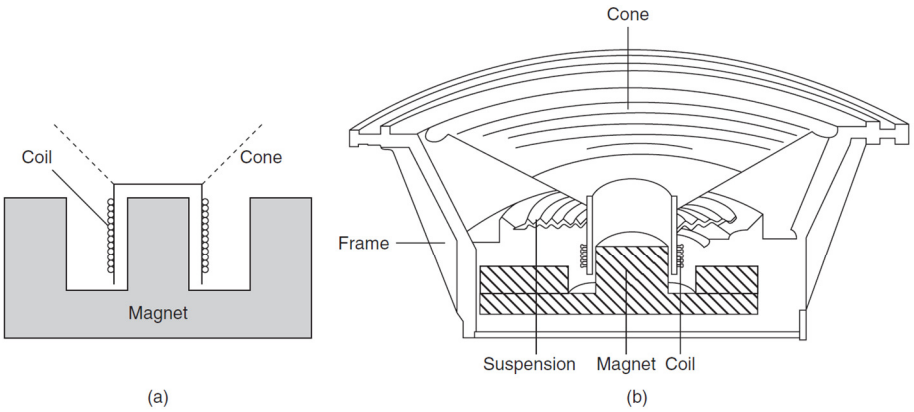


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

26. Claim 26: "The device according to claim 23, wherein the actuator consists of, or comprises, an electrical signal generator."

188. Lindahl-Williams discloses or at least suggests its device comprises a digital-to-analog converter, which is an actuator comprising an electrical signal generator. By 2012, it was well-known that a speaker system comprises a digital-to-analog converter ("DAC") for converting received digital audio into analog signals for playback. Further, Lindahl discloses analog-to-digital conversion functions. For example, Lindahl's "[p]rocessing circuitry 46 and storage 44 may be used in implementing analog-to-digital conversion functions for capturing audio." (EX1011, [0046].) "[D]evice 12 may have a microphone amplifier and associated analog-to-digital converter circuitry that digitizes audio clips." (EX1011, [0087].) Williams discloses that the audio received by the device may be in the form of a digital audio stream. "The API 214 of the application engine 118

may also allow the application 120 to provide audio 222 for eventual transmission to the interface device 116, to be played on the speaker 206. . . . [t]he audio 222 may be provided from the application 120 as a digital audio stream." (EX1006, [0037].) As Lindahl already supports analog-to-digital conversion and Williams discloses transmission and receipt of digital audio data, it would have been obvious to a POSITA that Lindahl-Williams's device may include a DAC for converting the received digital audio signal back into an analog signal for playback.

189. A POSITA would have understood that a DAC is an actuator consisting or comprising an electrical signal generator. This is consistent with the description of the '720 patent. The '720 patent discloses that "[t]he ***actuator*** may be controlled by an analog input, and the second device may comprise a ***digital to analog converter*** coupled to the analog input, for ***converting a digital data to an actuator input signal***." (EX1001, 34:25-28.) "In another example, the sounder simulates the voice of a human being or generates music, typically by using an electronic circuit having a memory for storing the sounds (e.g., music, song, voice message, etc.), a ***digital to analog converter*** to reconstruct the electrical representation of the sound, and a driver for driving a loudspeaker, which is ***an electro-acoustic transducer that converts an electrical signal to sound***." (EX1001, 107:13-20.) "An

analogue- to-digital converter (ADC) is a circuit that will convert a voltage level into a binary number. A varying voltage, such as an audio signal, or a slowly changing voltage, such as a temperature measurement voltage, will often need to be converted into a set of digital signals, so that microprocessor-based equipment can use the data . . . A digital-to-analogue converter (DAC) performs the opposite conversion, from a set of binary numbers into a ***voltage level***." (EX1018, 170.) Because the DAC outputs actuator input signal (*i.e.*, an electrical signal), the DAC is an actuator that consists of or comprises an electrical signal generator.

190. It would have been obvious that the digital-to-analog converter is communicatively coupled to the WLAN transceiver and activated in response to data (e.g., incoming digital audio signals) received from the WLAN in Lindahl-Williams. (Supra §IX.A.4.i.)

27. Claim 29: "The device according to claim 23, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

191. As discussed, Lindahl discloses multiple speakers in the device 12. (*Supra* § IX.A.23; EX1011, [0048].) A POSITA would have understood that a speaker is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, audio or sound. And a

POSITA would have understood that a DAC is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, electrical voltage, or electrical current.

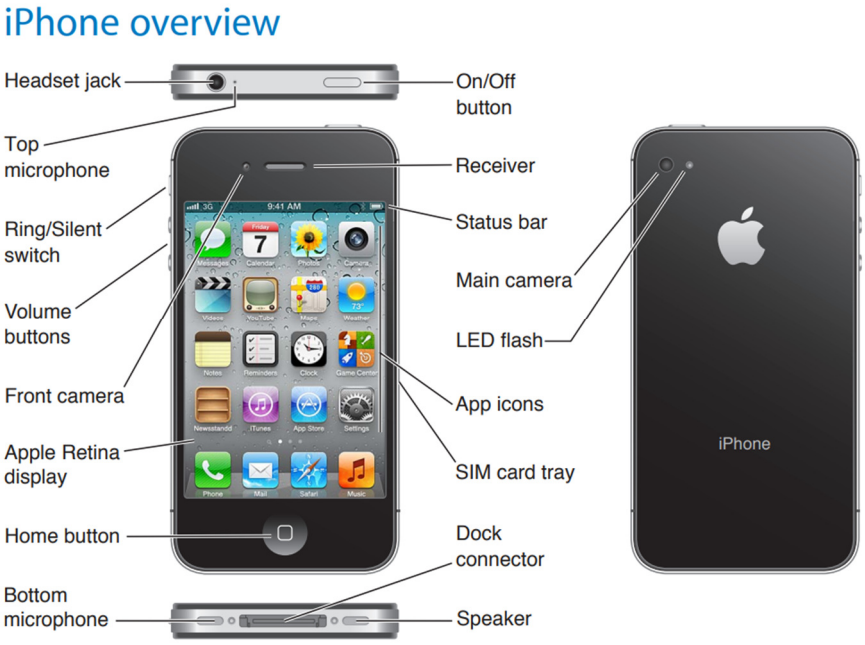
B. Ground 2: Lindahl, Williams, and iPhone User Guide Render Obvious Claims 1-2, 23-26, 29

1. Summary of iPhone User Guide

192. The iPhone User Guide was published/made publicly available at least by November 2011, before the earliest provisional application filing date of the '720 patent, and thus qualifies as prior art under §102(a). (EX1048.) The Examiner never found, and Applicant never cited, iPhone User Guide during prosecution.

193. Specifically, the iPhone User Guide is a publicly available document to purchasers, end-users, and service personnel using the iPhone, supplied to the purchasers and end-users along with the iPhone product or obtainable through their exercise of reasonable diligence (e.g., requesting a copy directly from Apple, obtaining a copy from the Apple service center, etc.).

194. iPhone User Guide provides an overview of the various components and features of an iPhone, as shown below:



195. The iPhone User Guide teaches a device with multiple microphones (e.g., top microphone and bottom microphone) and multiple speakers (e.g., receiver speaker and "Speaker"). (EX1008, 9.) Further, iPhone User Guide discloses that "iPhone has an internal rechargeable battery." (EX1008, 34.) "Charge the battery: Connect iPhone to a power outlet using the included Dock Connector to USB Cable and USB power adapter." *Id.*

196. iPhone User Guide discloses that the speaker is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN. For example, iPhone User Guide discloses FaceTime video calls. "FaceTime video calls (iPhone 4 or later) let you see as well as ***hear the person you're talking to***. You can make a video call to

someone with a Mac or iOS device that supports FaceTime. No setup is needed, but you must have a ***Wi-Fi connection to the Internet***." (EX1008, 45.)

2. Claim 1

a. 1[A]: "multiple microphones for capturing first human voice data;"

197. Claim 1 is obvious in further view of iPhone User Guide.

198. As discussed, with reference to FIG.2, Lindahl discloses a mobile phone as a device. (*Supra* §IX.A.4.aIX.A.22; EX1011, FIG. 2.) For example, Lindahl discloses that "[u]ser device 12 may be a ***mobile telephone***, a mobile telephone with media player capabilities, a media player, a handheld computer, a game player, a global positioning system (GPS) device, a combination of such devices, or any other suitable electronic device such as a portable device." (EX1011, [0044].) Lindahl also discloses that "the portable electronic devices may be wireless electronic devices . . . [that] may be, for example, handheld wireless devices such as ***cellular telephones***, media players with wireless communications capabilities, handheld computers (also sometimes called personal digital assistants), global positioning system (GPS) devices, and handheld gaming devices." (EX1011, [0023]-[0024].)

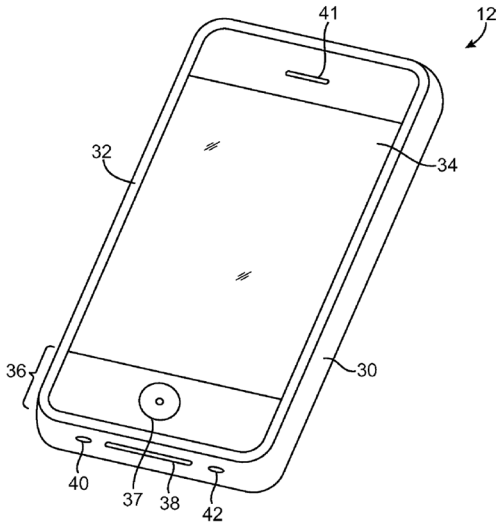
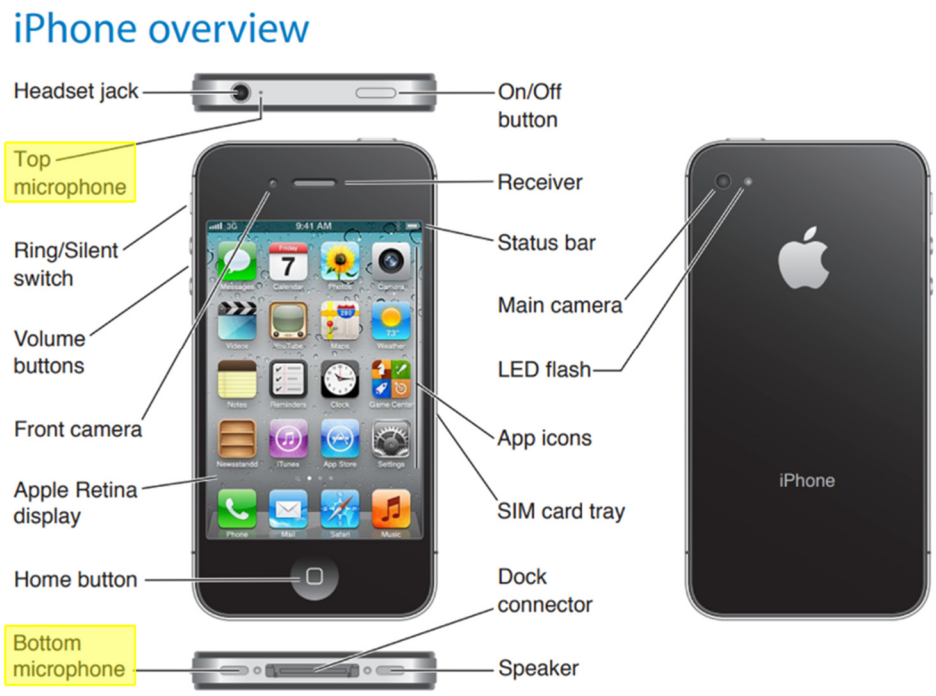


FIG. 2

199. iPhone User Guide discloses a smart phone with multiple microphones for capturing audio. For example, iPhone User Guide discloses the smart phone comprises multiple microphones, the "Top microphone" and "Bottom microphone." (EX1008, 9.) iPhone User Guide discloses using the "iPhone microphone" and "built-in microphone" to capture human voice data. (EX1008, 29, 119.)



200. It would have been obvious to combine Lindahl-Williams with iPhone User Guide. By 2012, the iPhone was a prevalent type of mobile phone. iPhone User Guide is a product manual providing an overview of the various components and features of an iPhone. Lindahl itself discloses a mobile telephone as a device, and thus suggests a motivation to combine with iPhone User Guide. (*Supra* §§IX.A.4.a-IX.A.22; EX1011, [0044], [0027], FIG. 2.) Both Lindahl and Williams are related to a device for capturing and processing audio data. (EX1011, [0048], [0040], [0006], FIGs. 2, 3; EX1006,[0016], [0019], [0023], [0037], FIG. 2, claims 1-3, 21.) iPhone User Guide is a product manual of a device (*i.e.*, an iPhone) that performs the same operations. (EX1008, 9, 29.) Thus, Lindahl, Williams, and iPhone

User Guide are from the same field of endeavor, and all relate to the same technical problems of capturing and processing audio data. (*Supra* § IX.A.3.) Thus, a POSITA would have been motivated to look to iPhone User Guide for techniques and details related to components for implementing Lindahl-Williams. Specifically, a POSITA would have been motivated to look to iPhone User Guide for implementing mobile phones. For example, the combination would result in a system where iPhone User Guide's multiple microphones (*e.g.*, top and bottom microphones) are incorporated in Lindahl-Williams to improve usability of the device and the quality of the captured audio. Mobile phones were known to include multiple microphones by 2012. (EX1008, 9.) The combination is straightforward and readily achievable by a POSITA without requiring undue effort and with a reasonable expectation of success, as demonstrated by successful implementations in the field. Further, the combination merely applies a known component/implementation method to Lindahl-Williams which already contemplated multiple microphones. (EX1011, [0048].) At the time of the invention, it was common for smart devices to have multiple microphones. Thus, application of multiple microphones as taught by iPhone User Guide in the same field would be motivated by customer expectations, design incentives, and other market forces.

b. 1[E]: "an enclosure that houses the multiple microphones, the WLAN transceiver, the speaker, and the electric light source;"

201. iPhone User Guide discloses an enclosure housing the recited elements such as the multiple microphones. For example, iPhone User Guide's multiple microphones are enclosed in the smart phone enclosure. (EX1008, 9.)



202. It would have been obvious to combine Lindahl-Williams with iPhone User Guide. (*Supra* §IX.B.2.a.)

c. 1[G]: "send, to a server device over the Internet via the WLAN by the WLAN transceiver, at least part of the first human voice data captured by the multiple microphones;"

203. As discussed, iPhone User Guide discloses a smart phone with multiple microphones for capturing audio data. (*Supra* §IX.B.2.a; EX1008, 9, 29, 117.)

204. It would have been obvious that at least part of the first human voice data captured by the multiple microphones is sent to Lindahl-Williams's server over the Internet via the WLAN by the WLAN transceiver. (*Supra* §IX.A.4.h.)

205. It would have been obvious to combine Lindahl-Williams with iPhone User Guide. (*Supra* §IX.B.2.a.)

3. Claim 2: "The device according to claim 1, further comprising, in the enclosure, a rechargeable battery for powering the device."

206. Claims 2, 23-26, 29 are obvious in further view of iPhone User Guide.

207. As discussed, with reference to FIG.2, Lindahl discloses a mobile phone as a device. (*Supra* §§IX.A.4.a-IX.A.22; EX1011, [0044], [0027]), [0023]-[0024], FIG. 2.)

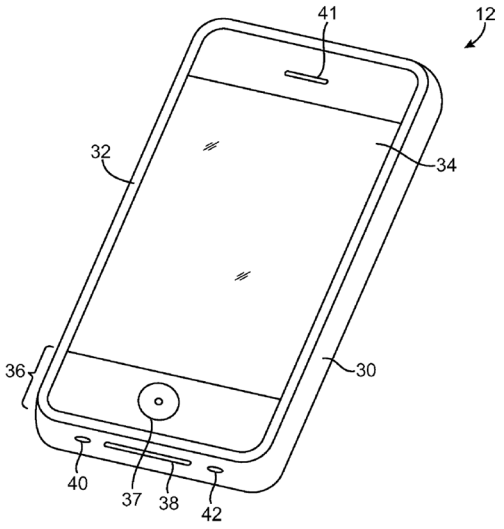


FIG. 2

208. Lindahl discloses, with reference to FIG. 3, that the user device 12 comprises storage 44 which includes "***battery-based*** static or dynamic random-access-memory." (EX1011, [0045].) In light of this disclosure, a POSITA would have understood that Lindahl's user device 12 comprises a rechargeable battery within its enclosure.

209. iPhone User Guide discloses a smart phone with a rechargeable battery. For example, iPhone User Guide discloses that "iPhone has an internal rechargeable battery." (EX1008, 34.) "Charge the battery: Connect iPhone to a power outlet using the included Dock Connector to USB Cable and USB power adapter." *Id*.

210. As discussed, Lindahl, Williams, and iPhone User Guide are from the same field of endeavor, and all relate to the same technical problems

of capturing and processing audio data. (*Supra* § IX.B.2.) Thus, a POSITA would have been motivated to look to iPhone User Guide for techniques and details related to components for implementing Lindahl-Williams.

211. It would have been obvious to a POSITA to apply iPhone User Guide's teachings related to mobile phones to Lindahl's user device 12, as various features of iPhone User Guide can improve it in the same way and yield predictable results. For example, the combination would result in a portable electronic device with a rechargeable battery as taught by iPhone User Guide. The known benefits of incorporating a rechargeable battery into a portable handheld device would prompt the application of iPhone User Guide in the same field based on design incentives, other market forces, and user expectations for portable systems, particularly where the device performs energy-intensive operations such as wireless communication, audio collection, and audio output. In such portable electronic devices, a rechargeable battery predictably improves usability and convenience by allowing the user to restore power through recharging the device. The benefits are predictable and directly aligned with the practical requirements of a portable device intended for frequent, repeated use.

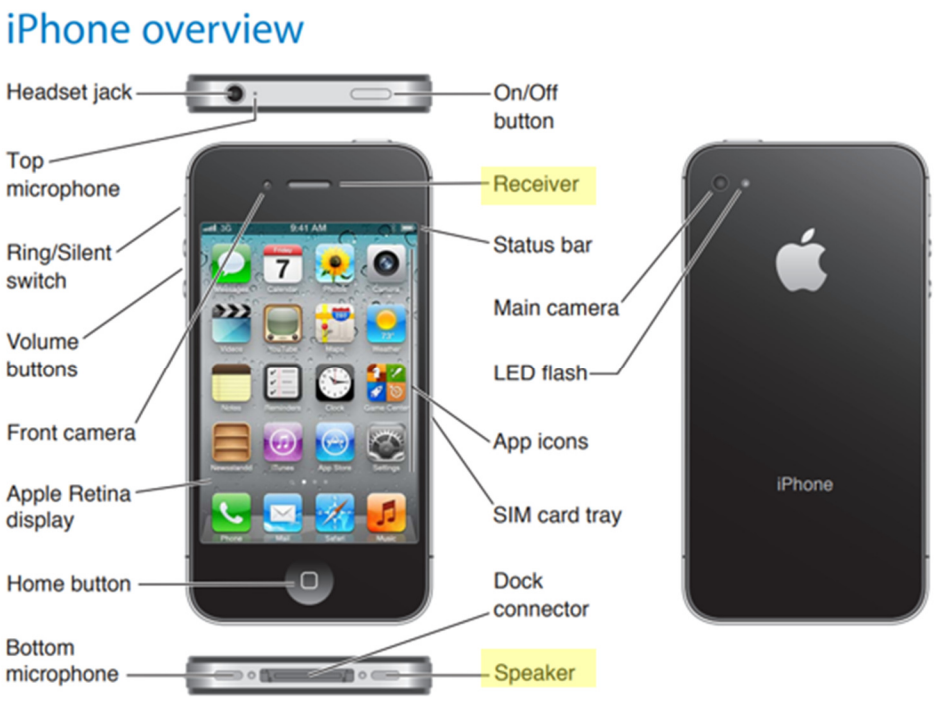
212. Further, it would have been obvious to try combining the technical features of iPhone User Guide with the teachings of Lindahl-

Williams, for example, by using a rechargeable battery to power the device, without requiring undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Further, there were a finite number of ways to power a device, such as via a rechargeable battery, a disposable battery, or external power supply, that were commonly seen in the industry and easy to manufacture and integrate into the Lindahl-Williams system. It would have been obvious to try, and a POSITA would have reasonably expected a rechargeable battery to work successfully as a straightforward substitution for disposable cells or as an alternative to external power supply. The selection of a rechargeable battery would therefore have been a routine engineering choice involving known components performing their expected function, yielding predictable results of improved portability, convenience, and reliable operation.

4. Claim 23: "The device according to claim 1, further comprising in the enclosure an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, the actuator that is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN."

213. iPhone User Guide discloses that a cellular phone may comprise multiple speakers. For example, iPhone User Guide discloses in the "iPhone Overview" that the iPhone includes both a "Receiver" and a "Speaker".

(EX1008, 9.) A POSITA would have understood (or it would have been obvious) that these are two speakers. "[T]he device 100 comprises at least a display 180 (which may or may not be touch-sensitive), a ***main speaker 104***, a ***telephone receiver/speaker 108***[.]" (EX1019, 3:20-22.) "***Different speakers***, such as a ***loudspeaker*** and a ***telephone receiver speaker***, may be positioned at different locations on the device (e.g., front, back, left side, right side, top, bottom, etc.)." (EX1019, 3:52-55.) One of the speakers can be mapped to the recited "speaker" in claim 1 and the other speaker can be mapped to the "actuator" in this claim.



214. Further, Lindahl also discloses multiple speakers in the device. Lindahl discloses that "[d]isplay and audio devices 52 may also include audio equipment such as ***speakers and other devices for creating sound***." (EX1011, [0048].) Thus, both iPhone User Guide and Lindahl disclose additionally comprising a speaker in the enclosure.

215. The specification of the '720 patent states that a speaker is an actuator. "An actuator may be an ***electroacoustic actuator, such as a sounder*** which converts electrical energy to sound waves transmitted through the air, an elastic solid material, or a liquid, usually by means of a vibrating or moving ribbon or diaphragm." (EX1001, 105:51-55.) "A ***loudspeaker (a.k.a. speaker) is a sounder*** that produces sound in response to an electrical audio signal input, typically audible sound. The most common form of loudspeaker is the electromagnetic (or dynamic) type, uses a paper cone supporting a moving voice coil electromagnet acting on a permanent magnet. Where accurate reproduction of sound is required, multiple loudspeakers may be used, each reproducing a part of the audible frequency range. A loudspeaker is commonly optimized for middle frequencies; tweeters for high frequencies; and sometimes supertweeter is used which is optimized for the highest audible frequencies." (EX1001, 105:64-106:7.)

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

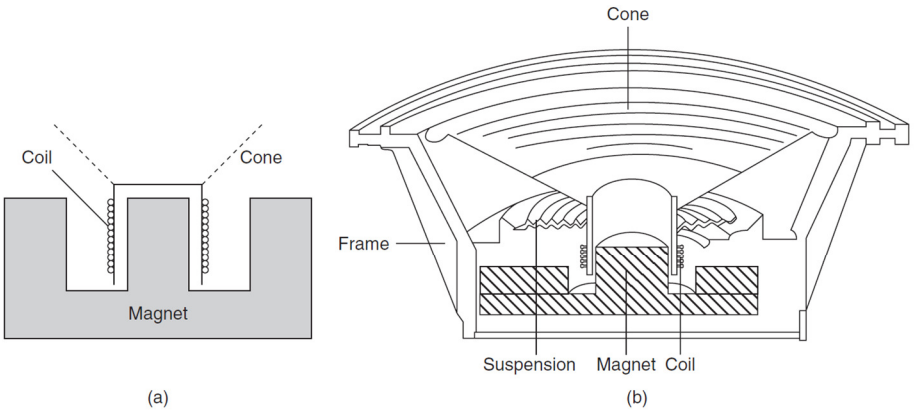


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

217. "[F]ew loudspeakers convert directly between electrical and acoustical energy. Instead, the electrical energy is converted to ***mechanical***

energy, specifically the ***movement of a diaphragm*** . . . which then radiates acoustical energy." (EX1014, 763.)

218. Thus, a speaker is or has an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, namely, audio, or sound.

219. iPhone User Guide discloses that the actuator is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN. iPhone User Guide discloses FaceTime video calls. "FaceTime video calls (iPhone 4 or later) let you see as well as ***hear the person you're talking to***. You can make a video call to someone with a Mac or iOS device that supports FaceTime. No setup is needed, but you must have a ***Wi-Fi connection to the Internet***." (EX1008, 45.) As discussed, Wi-Fi is a form of WLAN. (*Supra* § IX.A.4.c.) A POSITA would understand from this disclosure that, during a FaceTime call, the iPhone uses its WLAN (WiFi) transceiver to receive call data (*e.g.*, incoming network packets) over the Wi-Fi network connection. A POSITA would also understand that the iPhone would receive incoming audio data packets over the Wi-Fi connection (via the WLAN transceiver), and output the resulting audio through an audio output component such as the iPhone speaker. In this manner, the iPhone speaker actuator (*e.g.*, the loudspeaker at the bottom) is

activated or controlled in response to data received from the WLAN, for example, the incoming audio stream received through the Wi-Fi connection during the FaceTime call.

220. Second, a POSITA would understand that there is a communication/control path from the WLAN transceiver to the speaker in order for the speaker to output audio in response to call data during the FaceTime call. Thus, the actuator (*e.g.*, loudspeaker at the bottom) is communicatively coupled to the WLAN transceiver (*e.g.*, via processors).

221. As discussed, Lindahl, Williams, and iPhone User Guide are from the same field of endeavor, and all relate to the same technical problems (*e.g.*, capturing and responding to audio commands and wireless communication). (*Supra* § IX.B.2.) Further, a POSITA would have been motivated to look to iPhone User Guide, a product manual of a mobile phone, for components and implementation details of a conventional mobile phone for implementing Lindahl-Williams, which explicitly discloses a portable electrical device in the form of a phone that captures and uploads voice commands to a server for processing and receives audio data from the server for playback on a speaker. (EX1011, [0073], [0043]; EX1006,[0036].)

222. It would have been obvious to a POSITA to apply iPhone User Guide's teachings regarding conventional mobile phone audio components

and Wi-Fi call functionality to Lindahl-Williams, as various features of iPhone User Guide can improve Lindahl-Williams' mobile phone in the same way and yield predictable results. A POSITA would have been motivated to include an additional speaker (*e.g.*, the loudspeaker at the bottom) in Lindahl-Williams' phone as a common and well-known implementation feature of cellular phones to support multiple audio playback modes, such as ordinary handset/earpiece use and "speaker phone mode" (hands-free) operation, for example, allowing the user to "[l]isten through the built-in iPhone speaker instead of the receiver." (EX1008, 118.) Further, incorporating an additional speaker predictably improves usability by enabling clearer and louder playback, providing better audibility in noisy environments, and allowing flexibility in audio routing depending on the use case (*e.g.*, directing different audio streams or alerts to different outputs). These known benefits of incorporating an additional speaker into a mobile phone would prompt the application of iPhone User Guide in the same field based on design incentives and other market forces favoring improved multimedia/voice performance on mobile devices.

223. For example, the combination would result in a portable electronic device with an additional speaker, such as a loudspeaker at the bottom in addition to the receiver speaker, coupled to the WLAN transceiver

to be actuated by data received from the WiFi connection to enable audio or video calls, as taught by iPhone User Guide.

224. Further, it would have been obvious to try combining the technical features of iPhone User Guide with the teachings of Lindahl-Williams, by including an additional speaker in the phone and implementing WLAN-driven speaker actuation as taught by iPhone User Guide, without requiring undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, using routine smartphone architectures, as demonstrated by successful implementations in the field. There was a finite number of predictable audio output hardware configurations for phones, such as an integral earpiece speaker, an externally connected speaker, and/or an additional integral loudspeaker, that were commonly seen in the industry and easy to manufacture and integrate into the Lindahl-Williams system. Prior art such as iPhone User Guide demonstrated successful implementation of multiple speakers in a cellular phone. Moreover, Lindahl itself suggests multiple speakers (*e.g.*, multiple speaker ports such as port 41, and ports 40 or 42) in its device and thus suggests a motivation to combine. *E.g.*, "[D]isplay and audio devices 52 may also include audio equipment such as ***speakers.***" (EX1011, [0048], [0040], FIG. 2.) The known benefits and widespread

adoption of multi-speaker phone designs and WLAN-based voice/multimedia playback would therefore have motivated a POSITA to incorporate these conventional features into the Lindahl-Williams system with a reasonable expectation of success, yielding predictable improvements in usability, hands-free operation, and providing audible feedback based on WLAN-received data.

5. Claim 24: "The device according to claim 23, wherein the actuator consists of, or comprises, a motion actuator that causes linear or rotary motion."

225. As discussed, iPhone User Guide discloses that a mobile phone may comprise multiple speakers, for example, a cellular phone may comprise a receiver speaker and a general loudspeaker. (*Supra* § IX.B.4.)

226. This claim describes how a conventional speaker operates. A POSITA would have understood that a speaker is a motion actuator that causes linear motion. As discussed, a speaker converts electrical energy to mechanical energy. (*Supra* § IX.A.24.) "In a manner reminiscent of microphones, few loudspeakers convert directly between electrical and acoustical energy. Instead, the ***electrical energy is converted to mechanical energy, specifically the movement of a diaphragm***, often cone-shaped or dome-shaped, which then radiates acoustical energy, sometimes aided by an acoustical coupling horn or lens." (EX1014, 763.) As loudspeakers convert

electrical energy to mechanical energy resulting in movement of a diaphragm, it is clear that loudspeakers are motion actuators.

227. Further, a speaker moves linearly. (*Supra* § IX.A.24.) "[I]n the low-frequency range, the ***linear motion of an electrodynamical loudspeaker***, for example, can be described by a single-degree-of-freedom oscillator." (EX1014, 903.) Further, a dynamic or moving coil loudspeaker moves axially. "Ernst W. Siemens was the first to describe the ***dynamic or moving-coil loudspeaker***, with a circular coil of wire in a magnetic field and supported so that ***it could move axially***." (EX1014, 18.) Thus, an electrodynamical loudspeaker itself is a motion actuator that causes linear motion.

228. It would have been obvious to combine Lindahl-Williams and iPhone User Guide. (*Supra* § IX.B.4.)

6. Claim 25: "The device according to claim 23, wherein the actuator consists of, or comprises, an electromagnetic coil or an electromagnet operative for generating a magnetic or electric field."

229. As discussed, iPhone User Guide discloses that a mobile phone may comprise multiple speakers, such as a receiver speaker and a general loudspeaker. (*Supra* § IX.B.4.) As discussed, a speaker is an actuator. (*Supra* § IX.A.24.)

230. This claim describes how a conventional speaker operates. Consistent with the specification of the '720 patent, a POSITA would have understood that a speaker commonly comprises, an electromagnetic coil operative for generating a magnetic or electric field. The'720 patent states, "[a] loudspeaker (a.k.a. speaker) is a sounder that produces sound in response to an electrical audio signal input, typically audible sound. The most common form of loudspeaker is the electromagnetic (or dynamic) type, uses a paper cone supporting a moving ***voice coil electromagnet acting on a permanent magnet***. Where accurate reproduction of sound is required, multiple loudspeakers may be used, each reproducing a part of the audible frequency range. A loudspeaker is commonly optimized for middle frequencies; tweeters for high frequencies; and sometimes supertweeter is used which is optimized for the highest audible frequencies." (EX1001, 105:64-106:7.)

231. "Most loudspeakers use the ***moving-coil*** principle (Figure 10.8). The coil is wound on a former at the apex of a stiff diaphragm which is held in flexible supports so that the ***coil is the intense field of a permanent magnet***. A signal current through the coil will then vibrate the diaphragm, producing sound waves at the same frequency and with the same waveshape as the signal." (EX1018, 122.)

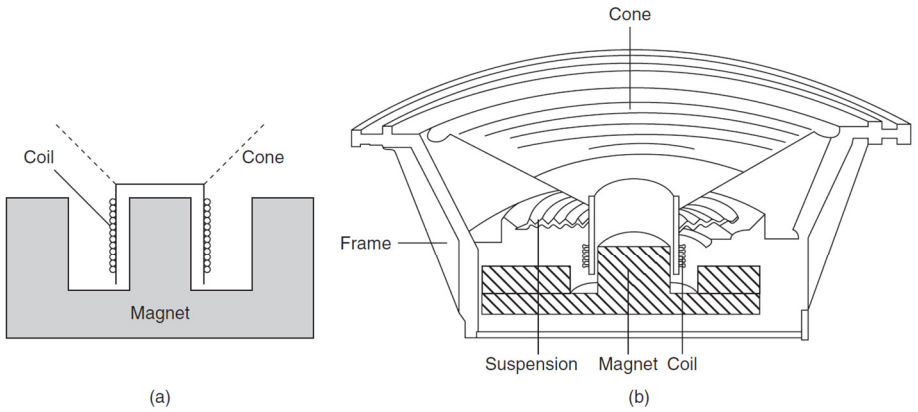


Figure 10.8 Moving-coil loudspeaker: (a) cross-section of coil and magnet, and (b) cutaway view of complete unit

232. It would have been obvious to combine Lindahl-Williams and iPhone User Guide. (*Supra* § IX.B.4.)

7. Claim 26: "The device according to claim 23, wherein the actuator consists of, or comprises, an electrical signal generator."

233. As discussed, it was well-known that a speaker system comprises a DAC for converting digital audio into analog signals for playback. (*Supra* § IX.A.26; EX1014, 714, 522; EX1018, 170.) A DAC is an actuator generating electrical signals, consistent with the '720 patent. (EX1001, 34:25-28, 107:13-20.) It would have been obvious that the Lindahl-Williams-iPhone User Guide system would include a digital-to-analog converter for converting received "digital audio stream" into analog signals for improved digital audio playback functions. (EX1006, [0037]; EX1011, [0087].) The inclusion of a digital-to-analog converter is straightforward and

readily achievable by a person of ordinary skill in the art, as demonstrated by successful and common implementations in the field.

8. Claim 29: "The device according to claim 23, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

234. iPhone User Guide discloses that a mobile phone may comprise multiple speakers, for example, a mobile phone may comprise both a receiver speaker and a general loudspeaker. (*Supra* § IX.B.4.)

235. This claim describes how a conventional speaker operates. A speaker is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, audio or sound. (*Supra* § IX.A.27.) Further, a DAC is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, electrical current or electrical voltage. *Id*.

236. It would have been obvious to combine Lindahl-Williams and iPhone User Guide. (*Supra* § IX.B.4.)

C. Ground 3: Lindahl, Williams, and Doherty Render Obvious Claims 4-5

1. Summary of Doherty

237. Doherty was published in 2007, before the earliest provisional application filing date of the '720 patent and thus qualifies as prior art under

§102(a) and (b). The Examiner never found, and Applicant never cited, Doherty during prosecution. Doherty is a textbook titled "Wireless Home Networking Simplified" published by Cisco. (EX1007, Title.) Doherty teaches readers to set up wireless networks at home using conventional products. (EX1007, Introduction, Chapter 6.) A POSITA would have looked to Doherty for wirelessly connecting in-home devices.

238. Doherty is related to wireless networking and communication among devices. (EX1007, Introduction, Chapter 6.)

2. Claim 4: "The device according to claim 1, further being addressable using the IP address in the WLAN or in the Internet using distinct locally administered addresses or a universally administered digital addresses."

239. Claim 1 is obvious in view of Lindahl and Williams. (*Supra* § IX.A.4.) Claims 4-5 are obvious in further view of Doherty. As discussed above, Lindahl discloses or at least suggests that the device (*e.g.*, user device 12 or interface device 116) is addressable using the IP address in the WLAN or in the Internet. (*Supra* §IX.A.4.k) 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. (EX1007, 99, 108, 175, 210, 340.) These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the

router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to communicate with devices outside of the user's WLAN. *Id*.

240. Doherty discloses (and a POSITA would understand) that the router assigned IP address is locally administered. "Most home networks use IP addresses in the range of 192.168.1.100 to 192.168.1.149, and the ***router assigns each device an address*** from this pool when it boots." (EX1007, 108, 210, 340, *see also*, *id*., 175.)

241. Further, Doherty discloses (and a POSITA would understand) that the router's IP address is globally unique and is a universally administered address. "Network Address Translation (NAT)[:] The technology by which a home router translates the private IP addresses used by computers on your home network into a single public IP address assigned to your broadband modem, providing the ability for multiple computers to share a connection, and also providing a degree of privacy because the computers on your home network are not able to be accessed directly from the Internet." (EX1007, 340).

242. Additionally or alternatively, Doherty discloses that each networked device has a NIC that has a unique MAC address. "***Each computer or device that you want to join your wireless network needs a***

wireless network interface card (NIC)," and "***each wireless card has a unique identifier called a MAC address***." (EX1007, 99, 166.) In light of this disclosure, a POSITA would understand that each networked device is associated with a globally unique MAC address for identification of the device in the wireless network. This unique identifier is assigned to the NIC at its manufacture and is thus a universally administered address (locally stored in the non-volatile memory). "***[A] MAC address*** . . . is a 48-bit binary number generally written as 12 hexadecimal digits (six groups of two digits, the groups separated by dashes or colons). The address is ***set at the time of the NIC's manufacture*** . . . The second least significant bit of the most significant byte MAC address is known as the ***universal or locally administered address (U/L) bit*** . If this bit is set to 0, the address is administered universally. This means that the address was assigned by the Institute of Electrical and Electronics Engineers (IEEE) and is unique across the globe. If the bit is set to 1, it indicates that the address was selected locally by a network administrator (overriding the NIC's BIA)." (EX1023, 24.) Thus, a POSITA would have understood (and found it obvious) that the device may comprise MAC addresses that uniquely identify the device in the WLAN or in the Internet, and that the MAC address may be administered universally.

243. Both the universally administered MAC address and the locally administered IP address are used in conjunction with each other when identifying the device. "On the DLC header, the data is functioning on Layer 2, so a MAC address is needed, as we saw. However, for the IP header, you have a set of IP addresses that should reinforce the fact you learned in Chapter 1—that IP functions on Layer 3 (the network layer) of the OSI model." (EX1023, 218, *see also id.,* 312 ("ARP allows IP-enabled devices on your network to dynamically map physical (MAC) addresses to IP addresses,"), 313, Figure 6.6.)

244. It would have been obvious for a POSITA to apply Doherty's teachings of conventional wireless networking components and addressing techniques to Lindahl-Williams. Lindahl, Williams, and Doherty are related to wireless communication technologies. As discussed, Lindahl discloses that device 12 communicates with ***computing equipment 14 and or 18*** (*i.e.*, uploading the voice commands and receiving information related to speech processing) over the Internet (*i.e.*, network 16), and via WiFi, a form of WLAN, directly or though a gateway device such as a router. (*Supra* §§ IX.A.4.c, IX.A.4.h; EX1011, [0025], [0032], [0026], [0030].) Williams discloses that the interface device 116 communicates with a wireless network (e.g., via a router). (*Infra* §§ IX.H.3.c, IX.H.3.h; EX1006,[0012] – [0013],

[0020], Figure 1.) Doherty, a textbook directed to "wireless home networking," explains that "[e]ach computer or device that you want to join your wireless network needs a wireless network interface card (NIC)," and "each wireless card has a unique identifier called a MAC address." (EX1007, 99, 166.) Thus, Lindahl, Williams, and Doherty are from the same field of endeavor, and all relate to the same technical problem of providing wireless communications between devices.

245. It would have been obvious for a POSITA to look to Doherty for well-known implementation details of conventional wireless networking to apply to Lindahl-Williams. Doherty teaches that each networked device includes a NIC and that the NIC includes a unique MAC address which is a universally administered address, used for device identification on a local network. Incorporating a NIC with a unique MAC address into Lindahl-Williams' device would have been a routine and predictable design choice, because MAC addressing is a fundamental and ubiquitous feature of IEEE 802 networking and is required for ordinary device participation in standard WLANs. Indeed, a POSITA would understand that for Lindahl's portable device to communicate with an Internet server over Wi-Fi, and for Williams' interface device to communicate with the cloud application engine using "existing networking protocols and techniques," the device necessarily

employs conventional addressing techniques and thus would include (or be implemented with) a NIC having a unique MAC address in the same way as taught by Doherty.

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

247. Further, there would have been a reasonable expectation of success in combining Doherty with Lindahl-Williams because the use of a NIC with a unique MAC address is a well-established and widely standardized networking feature implemented in essentially all WLAN-

capable devices. The combination is straightforward and readily achievable by a POSITA using routine engineering practices, and it merely applies a known, fundamental networking mechanism (unique MAC addressing) to the wireless communications already contemplated by Lindahl and Williams.

3. Claim 5: "The device according to claim 4, wherein the IP address is a layer 3 address and is static or dynamic IP address that is IPv4 or IPv6 type address."

248. As discussed, Lindahl-Williams-Doherty discloses claim 4. Doherty discloses, and a POSITA would understand that Internet Protocol (IP) addresses are either static or dynamic, are Layer 3 addresses and are IPv4 addresses. "The broadband networks typically assign a *dynamic IP address* to your broadband service. Dynamic means the IP address can change[.]" (EX1007, 277.) Further, an IP address may be a "*static IP address*, which means . . . your address never changes." *Id.* "In the case of today's Internet, these addresses are Internet protocol version 4 (IPv4) addresses." (EX1029, 127.) "Internet protocol (IP) addresses are the unique numbers used to identify individual connections to the Internet." *Id.* "There are currently two main versions of IP and their associated IP address families; ***IPv4 and IPv6***. IP addresses can be ***either static or dynamic*** [.]" (EX1024, 2:60-62.) "Internet Protocol (IP) is a Layer 3 protocol that contains addressing and control information that allows packets to be routed."

(EX1023, 11.) "The IP system in common use today is known as IPv4, for Internet Protocol version 4. A newer system, IPv6, or Internet Protocol version 6, has been developed and exists today in small deployments." (*Id*., 13.)

D. Ground 4: Lindahl, Williams, and Rossing Render Obvious Claims 7-8, 11-12

1. Summary of Rossing

249. Rossing was published by at least July 23, 2007, and is prior art under §102(a), (b). (EX1041, 1.) The Examiner never found, and Applicant never cited, Rossing during prosecution. Rossing is a textbook titled "Springer Handbook of Acoustics." (EX1014, Title.) Rossing discloses how conventional microphones and speakers operate.

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

251. Rossing also discloses different types of microphones such as a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone. "Two types of microphones avoid any mechanical attachment

to the diaphragm: the ***ribbon microphone*** and the ***condenser microphone***. A popular variant of the ***dynamic microphone***, the ribbon microphone, uses a thin metal ribbon diaphragm suspended in a magnetic field." (EX1014, 758.) "A popular modern variant is the electret condenser microphone, perfected in 1962 by Sessler and West of Bell Laboratories." (EX1014, 759.)

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

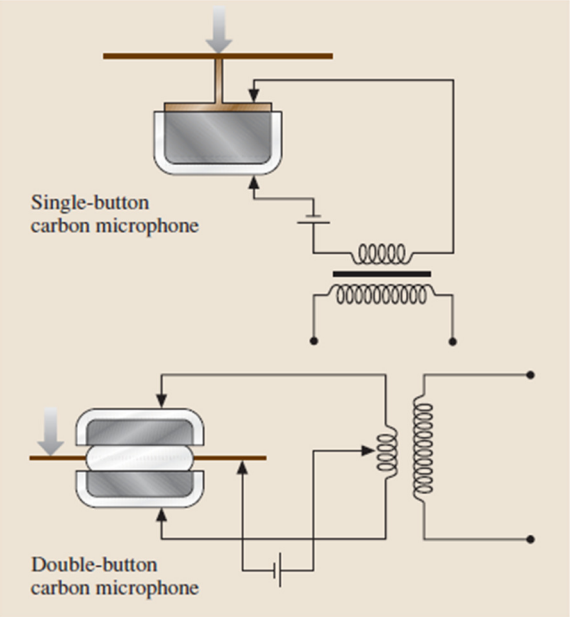


Fig. 18.11 A carbon microphone (After [18.25])

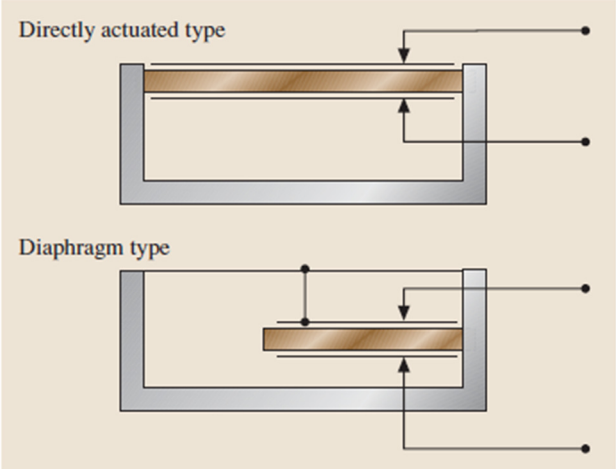


Fig. 18.12 Piezoelectric microphones (After [18.29])

253. Rossing also discloses how a conventional speaker operates. Rossing discloses that "in a manner reminiscent of microphones, few loudspeakers convert directly between electrical and acoustical energy. Instead, the ***electrical energy is converted to mechanical energy, specifically the movement of a diaphragm***, often cone-shaped or dome-shaped, which

then radiates acoustical energy, sometimes aided by an acoustical coupling horn or lens." (EX1014, 763.) "In the low-frequency range, the ***linear motion of an electrodynamical loudspeaker***, for example, can be described by a single-degree-of-freedom oscillator." (EX1014, 903.)

254. Rossing discloses that the soundwaves outputted by speakers are omnidirectional. "***Sound radiates from a loudspeaker in all directions***, depending on angle and frequency," and that "it is rare for a speaker to have truly constant directionality across its entire pass-band, in part from that fact that most are at least somewhat directional at mid and high frequencies, and, because of the long wavelengths involved, almost unavoidably ***omni-directional at low frequencies***." (EX1014, 765.)

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

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

2. Claim 7: "The device according to claim 1, wherein the speaker comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

257. Claim 1 is obvious in view of Lindahl and Williams. (*Supra* § IX.A.4.) Claims 7-8 and 11-12 are obvious in further view of Rossing.

258. A POSITA would understand that this claim refers to standard speakers. Lindahl and Williams disclose playing audio data by a speaker in the device (e.g., ***speakers and other devices for creating sound or speaker 206***). (*Supra* §§ IX.A.4.d, IX.A.4.i; *Infra* §§ IX.H.3.d, IX.H.3.i; EX1011, [0048], [0040], FIGs. 2, 3; EX1006,[0016], [0023], [0037], FIG. 2.)

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

260. Both Lindahl and Williams are related to operating microphones and speaker(s). (EX1011, [0048], [0040], [0006], FIGs. 2, 3; EX1006,[0016], [0019], [0023], [0037], FIG. 2, claims 1-3, 21.) Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1014, 757-758, 763, 765, 903.) Thus, Lindahl, Williams, and Rossing are from the same field of endeavor, and all relate to the same technical problems of operating microphone(s) and speaker(s). (*Supra* § IX.A.3.) Thus, a POSITA would have been motivated to look to Rossing for techniques and details related to speakers for implementing Lindahl-Williams.

261. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional speakers to Lindahl' speaker, as various speaker features can improve Lindahl's speaker in the same way and yield predictable results. Lindahl itself discloses standard speakers and thus suggests a motivation to combine. *Id.* The known benefits of various speaker features would prompt the application of Rossing in the same field based on design incentives and other market forces. Further, it would have been obvious to try combining the technical features of Rossing with the teachings of Lindahl-Williams, for example, by using an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker in the

device to play digital audio content, without requiring undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of types of speakers, such as an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, that were commonly seen in the industry and easy to manufacture and integrate into the Lindahl-Williams system. (EX1014, 321-322, 763.)

3. Claim 8: "The device according to claim 1, wherein the speaker is operative for omnidirectional, unidirectional, or bidirectional pattern audio emitting."

262. A POSITA would have understood that this claim refers to standard properties of a speaker. Lindahl and Williams disclose playing audio data by a speaker in the device (e.g., ***speakers and other devices for creating sound or speaker 206***). (*Supra* §§ IX.A.4.d, IX.A.4.i; *infra* §§ IX.H.3.d, IX.H.3.i; EX1011, [0048], [0040], FIGs. 2, 3; EX1006,[0016], [0023], [0037], FIG. 2.)

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

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

4. Claim 11: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, an omnidirectional, unidirectional, or bidirectional microphone that is based on sensing of an incident sound-based motion of a diaphragm or a ribbon."

264. A POSITA would understand that the claim refers to properties and components of a standard microphone.

265. Both Lindahl and Williams disclose microphone(s) in the device. For example, Lindahl discloses its user device 12 comprises user input devices 50 such as ***microphones***. (EX1011, [0048], *see also, id.*, [0006], [0040].) Williams discloses "[a]n audio interface device 116 … having a ***microphone*** and a network interface." (EX1006, [0016]; *see also*, *id.*, [0007], [0019], claims 1-3 and 21.)

266. Rossing discloses that a microphone includes a diaphragm and operates by sensing movement of the diaphragm. Rossing states that "[c]ommon to ***virtually all microphones*** ever made is a ***diaphragm***. Most

microphones do not directly sense the instantaneous pressure deviations of an airborne acoustic wave. Instead, the sound waves impinging on a diaphragm cause movement of the diaphragm, and it is this movement that is sensed and converted to an electrical audio signal." (EX1014, 757.)

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

268. As discussed, Lindahl, Williams, and Rossing are from the same field of endeavor and all relate to the same technical problems of operating microphone(s) and speaker(s). (*Supra* § IX.D.2.) A POSITA would have been motivated to look to Rossing for techniques and details regarding these components for implementing the system described in Lindahl-Williams.

269. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Lindahl-Williams, for example, as various microphone features can improve Lindahl-Williams' microphone in the same way and yield predictable results. Lindahl discloses

standard microphones and thus suggests a motivation to combine. *Id*. The known benefits of adopting various standard microphone components (for example, due to availability and popularity) would prompt the application of Rossing in the same field based on design incentives and market forces. Further, it would have been obvious to try combining Rossing with the teachings of Lindahl-Williams, for example by using an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon, without requiring undue effort and with a reasonable expectation of success. There was a finite number of types of microphones, such as an omnidirectional, unidirectional, or bidirectional microphone, that were commonly known in the industry and easy to manufacture and integrate into Lindahl-Williams. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field.

5. Claim 12: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

270. A POSITA would understand that the claim refers to standard types of microphones. As discussed, both Lindahl and Williams disclose microphone(s) in the device. For example, Lindahl discloses its user device 12 comprises user input devices 50 such as ***microphones***. (EX1011, [0048],

see also, id., [0006], [0040].) Williams discloses "[a]n audio interface device 116 … having a ***microphone*** and a network interface." (EX1006, [0016]; *see also*, *id.*, [0007], [0019], claims 1-3 and 21.)

271. Besides diaphragm-based microphone as disclosed above, Rossing discloses a variety of other microphones such as the ribbon microphone or the condenser microphone. (EX1014, 757.) "Two types of microphones avoid any mechanical attachment to the diaphragm: the ribbon microphone and the condenser microphone. A popular variant of the dynamic microphone, the ribbon microphone, uses a thin metal ribbon diaphragm suspended in a magnetic field. Movement of the diaphragm induces a current in the diaphragm, which can then be amplified. In effect, the diaphragm acts as a single-turn moving coil." (EX1014, 758.) "The condenser microphone is based on the notion of electrical capacitance. A metal, or other conducting, diaphragm is placed close to a second, fixed conducting plate, forming a capacitor, and a polarizing voltage is applied between the plates. Movement of the diaphragm causes alterations in the capacitance, in turn causing corresponding alteration of the charge on the plates, which is converted to a dynamic voltage by a resistor, and amplified as necessary." *Id.*

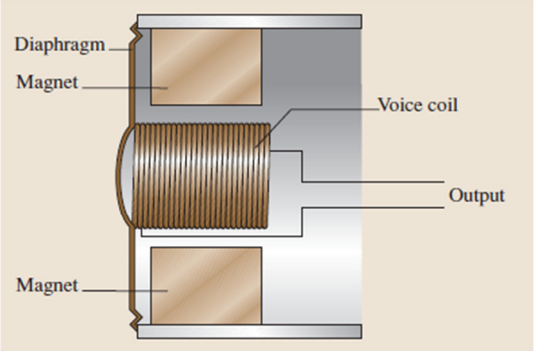


Fig. 18.10 A dynamic microphone uses a coil of wire attached to a diaphragm, suspended in a magnetic field (After [18.24])

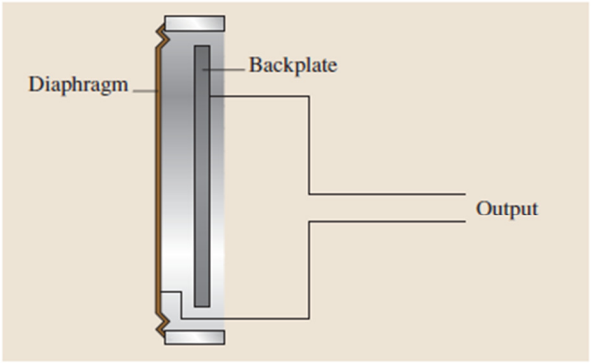


Fig. 18.9 A condenser microphone consists of a very thin diaphragm suspended parallel to a backplate (After [18.24])

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

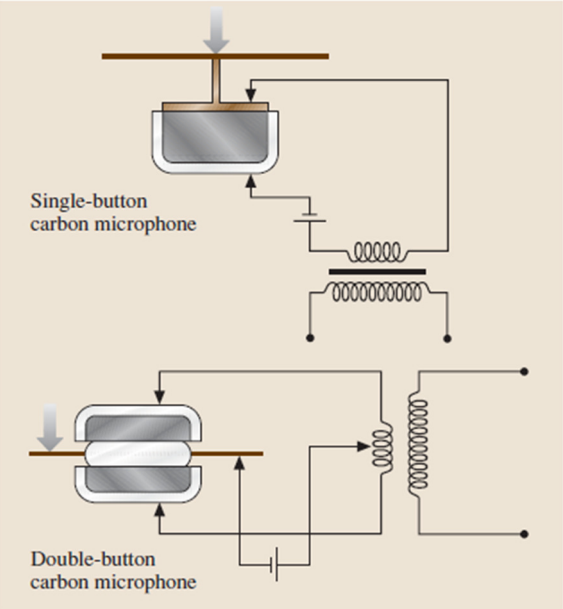


Fig. 18.11 A carbon microphone (After [18.25])

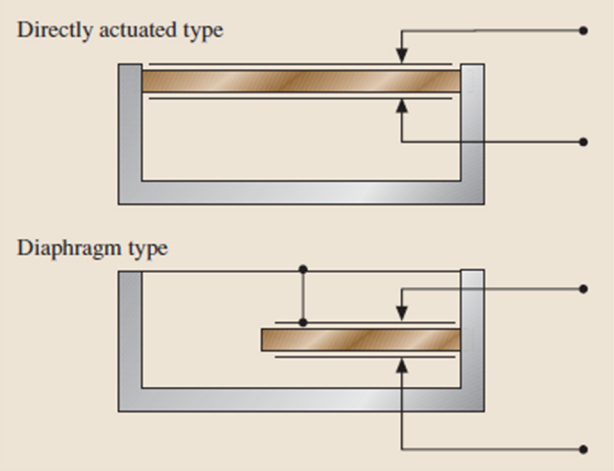


Fig. 18.12 Piezoelectric microphones (After [18.29])

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

exhibit good sample-to-sample consistency, making them ***practical for multi-microphone arrays***." (EX1014, 759.)

274. As discussed, Lindahl, Williams, and Rossing are from the same field of endeavor and all relate to the same technical problems of operating microphone(s) and speaker(s). (*Supra* § IX.D.2.) A POSITA would have been motivated to look to Rossing for techniques and details regarding these components for implementing the system described in Lindahl-Williams.

275. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Lindahl-Williams, for example, as various microphone features can improve Lindahl-Williams' microphone in the same way and yield predictable results. The known benefits of adopting various standard microphone components (for example, due to popularity and practical benefits of "good sample-to-sample consistency") would prompt the application of Rossing in the same field based on design incentives and market forces. Further, it would have been obvious to try combining Rossing with the teachings of Lindahl-Williams, for example by using a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone, without requiring undue effort and with a reasonable expectation of success. There was a finite number of types of microphones, such as a condenser, an electret, a dynamic, a ribbon, a carbon,

or a piezoelectric microphone, that were commonly known in the industry and easy to manufacture and integrate into Lindahl-Williams. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field.

E. Ground 5: Lindahl, Williams, and Benesty Render Obvious Claims 9-10

1. Summary of Benesty

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

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

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

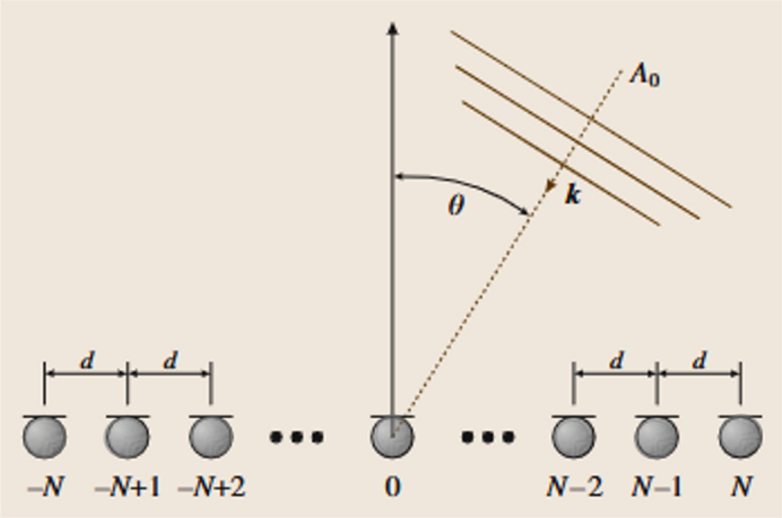


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

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

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

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

2. Claim 9: "The device according to claim 1, wherein the multiple microphones are configured to improve far-field audio reception."

281. Claim 1 is obvious in view of Lindahl and Williams. (*Supra* § IX.A.4.) Claims 9-10 are obvious in further view of Benesty.

282. Benesty discloses the use of microphone arrays, which comprises multiple microphones, to improve far-field audio reception. Benesty discloses 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.) A microphone array includes multiple microphones arranged in an array, as shown below. The schematic diagram FIG. 50.1 of Benesty illustrates a uniformly spaced linear microphone array by ***microphones*** spaced by a distance *d*.

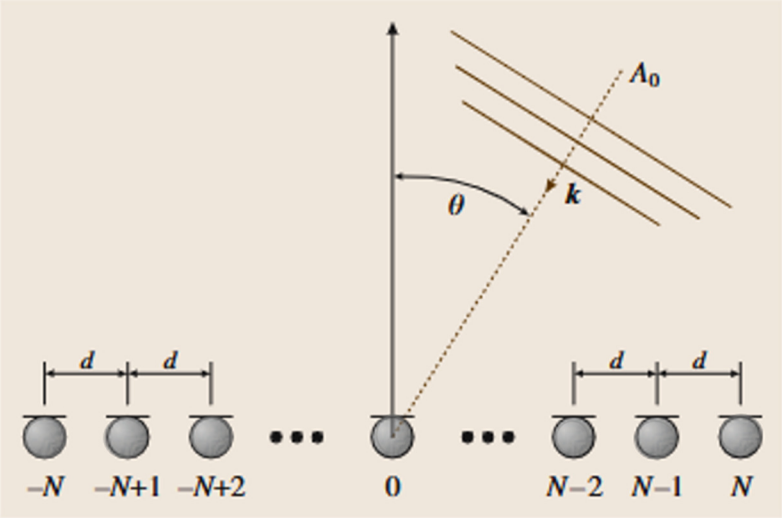


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

283. Benesty first discloses a problem with quality of audio data capturing and processing. "Hands-free audio communication is now a major feature in mobile communication systems as well as audio and video conferencing systems. One problem that becomes evident to users of these systems is the decrease in communication quality due to the pickup of room reverberation and background noise." (EX1010, 1021.)

284. Benesty then discloses the limitations with previous solutions to "us[e] microphones placed close to the desired talker or source." (EX1010, 1021.) "First, ***it is not always possible or desirable to place the microphone very close to the talker's mouth***. Second, by placing the microphone close to the talker's mouth, one has to deal with rapid level variation as the talker moves his or her mouth relative to the microphone." *Id.*

285. Benesty then introduces the solution of using microphone arrays which can offer similar audio performance as that of closely placed microphones, thereby improving audio collection where "it is not [] possible or desirable to place the microphone very close to the talker's mouth." (EX1010, 1021.) "One solution is to use beamforming microphone arrays, which can offer significant directional gain so as to result in similar audio performance to that of closely placed microphones." *Id*. Thus, Benesty teaches improving far-field audio reception (*e.g*., where the talker is at a distance from the microphone) using microphone arrays.

286. It would have been obvious to a POSITA to apply Benesty's techniques of arranging the multiple microphones into a microphone array to Lindahl-Williams's device to improve far-field audio reception. Lindahl, Williams, and Benesty are related to capturing and processing audio data. Lindahl-Williams contemplates capturing audio commands. For example, Lindahl discloses "electronic devices such as ***portable electronic devices*** that can capture ***voice commands*** and contextual information." (EX1011, [0001], *see also* Abstract.) Such devices comprise "input-output devices 48 [that] can include . . . ***user input devices 50*** such as . . . ***microphones***, cameras, etc." (EX1011, [0048].) FIG. 1 of Williams illustrates "an operational environment in which applications may utilize an application engine, and may

respond to ***audio from user premises*** to provide various services to users." (EX1006, [0003].) Benesty discloses hands-free audio communication and speech processing techniques. Benesty states that "hands-free audio communication is now a major feature in mobile communication systems as well as audio and video conferencing systems. One problem that becomes evident to users of these systems is the decrease in communication quality due to the pickup of room reverberation and background noise." (EX1010, 1021.) Thus, Lindahl, Williams, and Benesty are from the same field of endeavor and all relate to the same technical problem of providing hands-free audio communication between users and devices.

287. The Lindahl-Williams-Benesty combination would result in a system where multiple microphones or the microphone-array technique as taught by Benesty are incorporated in Lim-Williams' device to improve the quality of captured audio. For example, Benesty's microphone array(s) may be included in Lindahl-Williams' device 12 or interface device 116 to capture the first human voice data. Lindahl itself already discloses multiple microphones and thus suggests a motivation to combine. *Id*. Because Lindhal, Williams, and Benesty relate to audio capturing and processing, a POSITA would be motivated to use the known techniques of Benesty to improve Lindahl-Williams' ability to capture and process audio. Even in use

cases where background noise is desired, a microphone array can capture richer audio information, enabling more efficient and accurate processing, for example, separating or balancing speech and background noise when desired.

288. Combining the teachings of Lindahl-Williams and Benesty as such would yield predictable results of effectively improving audio capturing and processing. Further, it would have been obvious to try combining the technical features of Benesty with the teachings of Lindahl-Williams, as there is a finite number of identified, predictable resolutions, without requiring undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of potential solutions to improve audio signal capturing, such as additional software processing or incorporating additional hardware components. The known benefits of Benesty would prompt the application of Benesty in the same field based on design incentives and other market forces to produce a system architecture that responds to captured human voice data accurately, having improved audio communication signal pickup.

3. Claim 10: "The device according to claim 9, wherein the multiple microphones are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array."

289. 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.) "One solution is to use beamforming microphone arrays, which can offer significant directional gain so as to result in similar audio performance to that of closely placed microphones." *Id.*

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

291. Benesty also discloses that "[a]lternatively, a TDE-based source localization algorithm involves a two-step procedure. In the first step, a set of relative time differences of arrival (TDOAs) for different microphone pairs are calculated. In the second step, the acoustic source location is estimated

from the TDOAs and using the a priori knowledge about the locations of the microphones. TDE-based source localization algorithms have many merits: they cope well with both narrowband and broadband source signals; their localization resolution can be flexibly adjusted by varying the sampling rate and microphone array size; the effect of room reverberation is only of concern in the first step TDE and several recently developed robust TDE algorithms are promising for future practical use; and in general they are computationally efficient." (EX1010, 1044.)

292. It would have been obvious to a POSITA to combine the teachings of Lindahl and Benesty. (*Supra* § IX.E.2.)

F. Ground 6: Lindahl, Williams, and Goldsmith Render Obvious Claims 20-21

1. Summary of Goldsmith

293. Goldsmith was published by at least January 20, 2006, and is prior art under §102(a) and (b). (EX1042, 2.) The Examiner never found, and Applicant never cited, Goldsmith during prosecution. Goldsmith is a textbook titled "Wireless Communications." (EX1042, Title.) Goldsmith provides an overview of wireless communication techniques, including the wireless systems available at the time of publication in 2005.

294. Goldsmith discloses that the wireless systems available in 2005 includes, cellular telephone systems, cordless phones, wireless LANs, wide

area wireless data services, broadband wireless access, paging systems, satellite networks, low-cost low-power radios such as Bluetooth and Zigbee, as well as ultrawideband radios. (EX1042, 8-21.)

295. Goldsmith discloses the wireless LAN standard, IEEE 802.11b standard, the standard which many companies developed their products on, and that "[t]wo additional standards in the 802.11 family were developed to provide higher data rates than 802.11b. The ***IEEE 802.11a wireless LAN*** standard . . . the other standard, ***802.11g***." (*Id*., 15-16.)

296. Goldsmith discloses that 802.11b WLAN is one of the ISM Band II frequencies. (EX1042, 22-23, Table 1.2.)

Table 1.2: Unlicensed U.S. spectrum allocations

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

297. Goldsmith is a textbook published by Cambridge University Press and authored by Andrea Goldsmith, a professor of Electrical Engineering at the Stanford University. Goldsmith was well-known among

POSITAs. A POSITA would have looked to Goldsmith for wireless communication systems, protocols, and techniques of applying them.

2. Claim 20: "The device according to claim 1, wherein the WLAN is according to, or compatible with, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard."

298. Claim 1 is obvious in view of Lindahl and Williams. (*Supra* § IX.A.4) Claims 20-21 are obvious in further view of Goldsmith.

299. As discussed, Lindahl discloses or at least suggests this claim. (*Supra* §§ IX.A.4.c, IX.A.4.h, IX.A.20; EX1011, [0025], [0032], [0026], [0030].)

300. Goldsmith discloses Wireless Local Area Network (WLAN) based on IEEE 802.11b as one of the common wireless network communication protocols well before the priority date .

301. Goldsmith discloses that wireless LAN was well-known and widely adopted. "Wireless LANs provide high-speed data within a small region, e.g. a campus or small building, as users move from place to place." (EX1042, 15.) "The second generation of wireless LANs in the U.S. operate with 80 MHz of spectrum in the 2.4 GHz ISM band. A wireless LAN standard for this frequency band, the ***IEEE 802.11b standard***, was developed to avoid some of the problems with the proprietary first generation systems. The

standard specifies direct sequence spread spectrum with data rates of around 1.6 Mbps (raw data rates of 11 Mbps) and a range of approximately 150m. The network architecture can be either star or peer-to-peer, although the peer-to-peer feature is rarely used. ***Many companies developed products based on the 802.11b standard, and after slow initial growth the popularity of 802.11b wireless LANs has expanded considerably***. ***Many laptops come with integrated 802.11b wireless LAN cards. Companies and universities have installed 802.11b base stations throughout their locations, and many coffee houses, airports, and hotels offer wireless access***, often for free, to increase their appeal." (EX1042, 16.)

302. Goldsmith discloses a WLAN based on IEEE 802.11a and 802.11g. "Two additional standards in the 802.11 family were developed to provide higher data rates than 802.11b. The ***IEEE 802.11a wireless LAN*** standard operates with 300 MHz of spectrum in the 5 GHz U-NII band. The 802.11a standard is based on multicarrier modulation and provides 20-70 Mbps data rates. Since 802.11a has much more bandwidth and consequently many more channels than 802.11b, it can support more users at higher data rates. There was some initial concern that 802.11a systems would be significantly more expensive than 802.11b systems, but in fact they quickly became quite competitive in price." (EX1042, 16.) In addition, "the other

standard, ***802.11g***, also uses multicarrier modulation and can be used in either the 2.4 GHz and 5 GHz bands with speeds of up to 54 Mbps. ***Many wireless LAN cards and access points support all three standards to avoid incompatibilities***." (EX1042, 16.)

303. It would have been obvious for a POSITA to apply Goldsmith's disclosure of various types of wireless networking techniques and protocols to Lindahl-Williams. Lindahl, Williams, and Goldsmith are related to wireless communication technologies. As discussed, Lindahl discloses that device 12 communicates with ***computing equipment 14 and or 18*** (*i.e.*, uploading the voice commands and receiving information related to speech processing) over the Internet (*i.e.*, network 16), and via WiFi, a form of WLAN, directly or though a gateway device such as a router. (*Supra* §§ IX.A.4.c, IX.A.4.h; EX1011, [0025], [0032], [0026], [0030].) Williams discloses that the interface device 116 communicates with a wireless network (e.g., via a router). (*Infra* §§ IX.H.3.c, IX.H.3.h; EX1006,[0012]-[0013], [0020], Figure 1.) Goldsmith discloses various existing wireless connection systems in a chapter titled "1.4 Current Wireless Systems" (EX1042, 8-21.) Thus, Lindahl, Williams, and Goldsmith are from the same field of endeavor and all relate to the same technical problem of providing wireless communications between devices.

304. A POSITA would be motivated to use the known wireless networking techniques of WLAN over IEEE 802.11b, 802.11a, and/or 802.11g to improve the usability of Lindahl-Williams in the same way. Combining the teachings of Lindahl-Williams and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

305. A POSITA would have looked to Goldsmith's disclosure of WLAN based on IEEE 802.11b, IEEE 802.11a, and IEEE 802.11g, and understood that the WLAN communications between device 12 and the remote server as disclosed in Lindahl-Williams can be WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g. Lindahl expressly discloses WiFi based on IEEE 802.11, and thus suggests a motivation to combine. (EX1011, [0030], [0052]) Further, Goldsmith explains the popularity of 802.11b wireless LANs and the wide coverage of locations with 802.11b base stations installed. (EX1042, 16.) Goldsmith discloses that 802.11a can support more users at higher data rates, and 802.11g can be used in either the 2.4 GHz and 5 GHz bands. Goldsmith discloses that "many wireless LAN cards and access points support all three standards to avoid incompatibilities." *Id.* Customer expectations, product design incentives and requirements, and

strong market forces would have prompted the application of known and popular WLAN standards, such as IEEE 802.11a/b/g in Lindahl-Williams.

306. Further, the combination is straightforward and readily achievable by a person of ordinary skill in the art without requiring undue effort and with a reasonable expectation of success, as demonstrated by successful implementations in the field. There was a finite number of potential wireless WLAN network communications, such as WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g, that were commonly known in the industry in 2012 and were easy to apply to Lindahl-Williams. The known benefits and feasibility of including a well-known wireless networking technique such as WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g as taught by Goldsmith to connect in-home devices with a remote server would prompt the application of Goldsmith in the same field based on design incentives and other market forces.

3. Claim 21: "The device according to claim 1, wherein the WLAN uses a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

307. Goldsmith summarizes in the table below the unlicensed radio frequency spectrum allocations in the U.S., which expressly shows that

802.11b and 802.11g WLANs are ISM Band II frequencies. (EX1042, 22-23, Table 1.2.)

Table 1.2: Unlicensed U.S. spectrum allocations

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

308. As discussed above, a POSITA would be motivated to combine the teachings of Goldsmith and Lindahl-Williams and apply the well-known wireless communication techniques such as WLAN over 802.11b or 802.11g, which are ISM Band II frequencies, as taught in Goldsmith to Lindahl-Williams. (*Supra* §§ IX.F.2, IX.A.21, IX.A.4.c.)

G. Ground 7: Lindahl, Williams, and Slaby Render Obvious Claims 14, 23, 28-29.

1. Summary of Slaby

309. Slaby was filed on November 30, 2011, and is prior art under §102(e). (EX1020, 1.) The Examiner never found, and Applicant never cited, Slaby during prosecution.

310. Slaby discloses that "a user computer device, such as a ***mobile device***, is provided that comprises a temperature sensitive touchscreen having

a temperature sensitive user interface comprising ***multiple thermal energy emitter***/detector devices, such as ***thermocouples***. The multiple thermal energy emitter/detector devices are capable both of detecting thermal energy and ***generating thermal energy***." (EX1020, 3:36-45.) Slaby discloses that it is advantageous to include multiple thermal energy emitter/detector devices in a mobile device for "improved capabilities in terms of detecting one or more mobile device operational conditions and providing support for improved detection capabilities." (EX1020, 3:36-39.)

2. Claim 14: "The device according to claim 1, wherein the light is a non-visible light that is infrared, ultraviolet, X-rays, or gamma rays."

311. Claim 1 is obvious in view of Lindahl-Williams. (*Supra* § IX.A.4.) Claims 14, 23, 28-29 are obvious in further view of Slaby. Referring to FIG. 1, Slaby discloses a device 102 (*e.g.*, a "wireless mobile device"). The device 102 includes a "touchscreen 104" that is "temperature sensitive." (EX1020, 5:47-50, 5:65-6:2.) Specifically, the touchscreen 104 includes "multiple thermal energy emitter/detector devices," which are "capable of generating and emitting thermal energy." (EX1020, 6:2-21, *see also*, 32:36-45.) These thermal energy emitter/detector devices would emit non-visible light such as infrared through thermal radiation. As discussed, objects that generate heat will emit thermal radiation, that is, some energy in

the infrared spectrum. (*Supra* § IX.A.13.) "It was mentioned above that in ***any object*** every atom and every molecule vibrate. The average kinetic energy of vibrating particles is represented by the temperature. . . . The electromagnetic radiation associated with heat is called ***thermal radiation***. Figure 3.41 shows the total electromagnetic radiation spectrum spreading from γ rays to radio waves. Thermal radiation is predominantly situated in the ***mid and far-infrared (IR)*** spectral ranges." (EX1030, 126.)

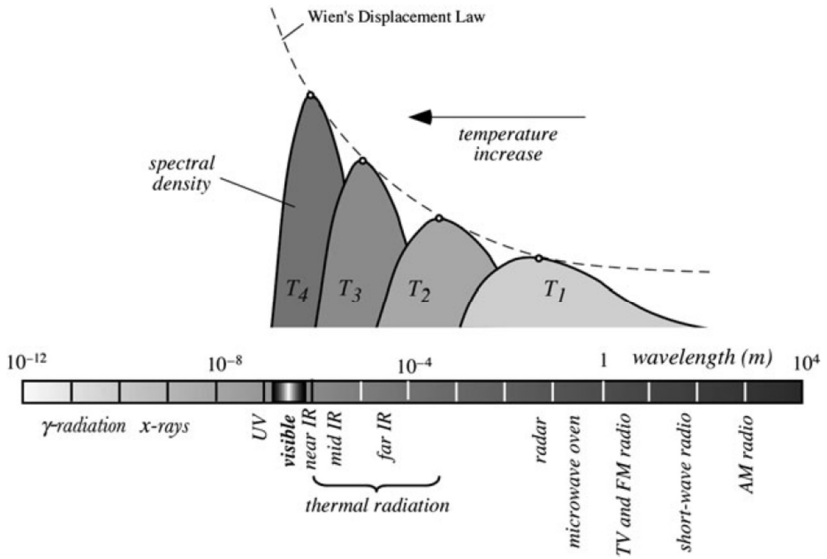


Fig. 3.41 Spectrum of electromagnetic radiation

312. It would have been obvious to a POSITA to combine the teachings of Lindahl-Williams and Slaby. (*Infra* § IX.G.3)

3. Claim 23: "The device according to claim 1, further comprising in the enclosure an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, the actuator that is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN."

313. Referring to FIG. 1, Slaby discloses a device 102 (*e.g.*, a "wireless mobile device"). The device 102 includes a "touchscreen 104" that is "temperature sensitive." (EX1020, 5:47-50, 5:65-6:2.) Specifically, the touchscreen 104 includes "multiple thermal energy emitter/detector devices," which are "capable of generating and emitting thermal energy." (EX1020, 6:2-21, *see also*, 32:36-45.) Slaby's thermal energy emitter/detector devices are in the phone enclosure because they are a part of the "temperature sensitive touchscreen." (EX1020, 6:2-21, 3:36-45, *see also*, 32:36-45.) These thermal energy emitter/detector devices generate thermal energy "in response to the application of the current" and thus are actuators that are directly or indirectly affecting, changing, producing, or creating heat/temperature. *Id.*

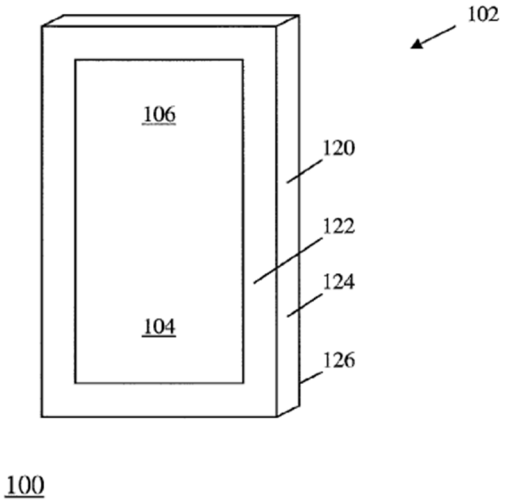


FIG. 1

314. Slaby's device 102 has a WLAN transceiver. Slaby's device can be "wireless mobile device" that is "capable of sending and receiving communication signals on a wireless network." (EX1020, 5:47-55.) It has a "[t]ransceiver 310 include[ing] at least one wireless receiver . . . and at least one wireless transmitter . . . for receiving and transmitting wireless signals" and supports "WiFi connectivity." (EX1020, 7:58-66, 27:55-59.) As discussed, WiFi is a form of WLAN. (EX1042, 600-601; EX1013, 7:48-51.) Thus, in other words, Slaby's device has a WLAN transceiver.

315. Slaby's thermal energy emitter/detector devices are coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN. First, they can be activated to display an image in response to the image being wirelessly downloaded to the device. An image "may be downloaded, wirelessly or over a wired connection, by the user computer device." (EX1020, 17:11-17.) In response, the device "activate, that is, heat up, thereby producing (908) a corresponding thermal image in temperature sensitive user interface 108." (EX1020, 18:3-15.) "[A]n activating of thermal energy emitter/detector devices 110 causes the devices to generate thermal energy, in turn causing a heating up of the thermally sensitive film or ink 112 proximate to the heated up thermal energy emitter/detector devices, thereby ***producing an image and/or color change*** in

the film or ink corresponding to the heated up devices, which image may be displayed to a user of the user computer device." (EX1020, 6:46-52.)

316. Second, Slaby's thermal energy emitter/detector devices are activated to perform authentication in response to an authentication request, which is wirelessly received by the device. The device 102 "may self-determine to thermally authenticate itself ***based on a short-range*** (for example, Bluetooth, infra-red, near field communication (NFC), or thermally-generated) ***authentication request*** received from another electronic device." (EX1020, 20:9-14.) Slaby discloses that "a ***short-range*** wireless protocol" includes "the Bluetooth protocol or a Wireless Local Area Network (WLAN) protocol that operates in accordance with the IEEE (Institute of Electrical and Electronics Engineers) 802.xx standards, for example, the 802.11 or 802.16 standards." (EX1020, 23:42-46.) In response to receiving the authentication request, the device "selectively activat[es] (1708), in temperature sensitive user interface 108." (EX1020, 20:19-28.)

317. Third, Slaby's thermal energy emitter/detector devices are activated to pre-tune the temperature in response to weather information, which is wirelessly received by the device. Specifically, "[i]n order to facilitate an operation of temperature sensitive touchscreen 104 in all environmental and operating conditions, user computer device 102 further

provides for an ***auto-biasing, that is, a pre-tuning, of a temperature of the thermal energy emitter/detector devices 110*** of temperature sensitive user interface 108." (EX1020, 29:50-56.) Slaby teaches that the mobile device performs the auto-bias function after receiving weather data. For example, "user computer device 102 may . . . use[] a ***received broadcast of weather data*** to estimate the temperature corresponding to your GPS position, such as the 'Thermometer' . . . or weather applications." (EX1020, 30:21-29.) A POSITA would have understood that the device 102 receives the weather data via WLAN.

318. "[T]he auto-biasing could be enabled by placing a ***heating element***, such as a resistive element, for example, a resistor, near the thermocouple junction areas. Power (for example, an applied Voltage or current) then could be continuously applied to the heating elements or could be applied in bursts in time (averaging effects) until the ***thermocouple junction temperatures are elevated by few degrees, for example, 5-10 degrees***." (EX1020, 31:51-67.)

319. Further, Slaby's thermal energy emitter/detector devices are activated in response to "being within range of a WiFi node." "[U]ser computer device 102 may store, in at least one memory device 304, a queue of processor intensive (heat generating) tasks that processor 302 of the user

computer device needs to perform but is waiting for some condition to occur, such as being plugged into a power outlet or ***being within range of a WiFi node***. If the auto-biasing needs to be performed, the processor could decide to perform one or all of the intensive tasks now, rather than waiting for the occurrence of the condition, so that the internally generated heat can be used to bias the thermal energy emitter/detector devices 110." In other words, Slaby performs auto-biasing (*i.e.*, activating its thermal energy emitter/detector devices) in response to data received via the WLAN, which indicates whether the device is within range of a WiFi node. (EX1020, 32:25-31.) A POSITA would understand, and it would be obvious, that Slaby's thermal energy emitters, or heating elements such as resistors are coupled to the WLAN transceiver and are activated or controlled in response to data received from the WLAN in order to carry out the heat generating tasks.

320. It would have been obvious to a POSITA to combine the teachings of Lindahl-Williams and Slaby. Lindahl and Slaby both relate to providing a touchscreen for receiving touch inputs. Lindahl's user device 12 comprises input-output devices 48 which includes "***touch screens***," and device 12's "[d]isplay 34 may be a ***touch screen*** liquid crystal display." (EX1011, [0048], [0039].) Slaby's device 102 includes a "touchscreen 104" that is "temperature sensitive." (EX1020, 5:47-50, 5:65-6:2.) Thus,

Lindahl and Slaby are from the same field of endeavor and all relate to the same technical problems of providing a touchscreen for receiving touch inputs.

321. Slaby provides a temperature-sensitive touchscreen with multiple thermal energy emitter/detector devices to improve detection. Slaby discloses that it is advantageous to include multiple thermal energy emitter/detector devices in a mobile device for "improved capabilities in terms of detecting one or more mobile device operational conditions and providing support for improved detection capabilities [of the touchscreen]." (EX1020, 3:36-39.) Further, Slaby's temperature-sensitive touchscreen provides various functionalities such as to "produc[] a corresponding thermal image," to "authenticate the user computer device" by generating thermal patterns/images, and to "auto-bias the temperature of the temperature sensitive user interface in order to better assure proper operation of the interface in all operating conditions." (EX1020, 3:45-48, 3:52-56, 20:19-39.)

322. A POSITA would have been motivated to combine Lindahl, Williams, and Slaby to include Slaby's thermal energy emitter/detector devices in Lindahl-Williams' touch screen or include Slaby's temperature-sensitive touchscreen in Lindahl-Williams' device, thereby resulting in a system with a temperature-sensitive touchscreen to provide the same types of

enhancements described by Slaby, (*e.g.*, device authentication, display of thermal images, auto-bias of use interface temperature to improve detection capabilities). Further, Lindahl itself discloses a touch screen and thus suggests a motivation to combine. (*Id.*; EX1011, [0048], [0039]; EX1020, 5:47-50, 5:65-6:2.)

323. Incorporating Slaby's thermal energy emitter/detector devices into Lindahl-Williams yields the predictable result of improving the usability of the mobile device. Slaby's "auto-bias" functionality would assure proper operation of the temperature sensitive user interface in all operating conditions and improve detection on the touchscreen, and Slaby further teaches using thermal patterns/images to authenticate the user computer device. Further, it would have been obvious to try combining the teachings of Lindahl-Williams and Slaby. There are only a finite number of potential solutions to improve touchscreen detection and assure proper touchscreen operation across operating conditions such as improving touch-sensing algorithms, applying environmental/temperature compensation and recalibration, adding active heating/auto-biasing, and/or incorporating multiple sensing elements. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Slaby has successfully demonstrated that a mobile device

may use thermal energy emitting/detector devices to "display thermal image," "authenticate the user computer device," or to "assure proper operation" of the user interface. Thus, there would have been a reasonable expectation of success for combining Lindahl-Williams and Slaby. The above discussed benefits of Slaby would prompt the application of Slaby in the same field based on design incentives and other market forces.

4. Claim 28: "The device according to claim 23, wherein the actuator consists of, or comprises, a thermoelectric actuator that is a heater or a cooler and is configured for affecting a temperature of a solid, a liquid, or a gas object, by conduction, convection, force convention, thermal radiation, or by transfer of energy by phase changes."

324. Slaby discloses thermoelectric actuators. First, Slaby discloses thermoelectric actuators that are heaters. As discussed, Slaby discloses that the mobile device may comprise multiple thermal energy emitter/detector devices such as ***thermocouples*** that can emit thermal energy. (EX1020, 3:36-45.) A POSITA would have understood that a thermocouple that emits thermal energy is a thermoelectric actuator. The '720 patent specification confirms that "[a] thermoelectric actuator may be an electric heater, converting electrical energy into heat, using resistance, or a dielectric heater." (EX1001, 26:25-27.) Indeed, Slaby discloses that "the thermal energy emitter/detector devices 110" "***activate, that is heat up***" "in response to the application of the ***current or voltage***." (EX1020, 18:9-13, 20:31-33.)

Slaby's thermal energy emitter/detector devices convert electrical energy into heat, and thus is a thermoelectric actuator.

325. Slaby further discloses that the mobile device may include a ***heating element***, such as a resistor, placed near the thermocouple junction areas to increase or decrease the ***temperature*** of the temperature sensitive user interface. (EX1020, 31:6-12, *see also*, 3:52-56; 31:51-67.) Slaby discloses that "the auto-biasing could be enabled by placing a ***heating element***, such as a resistive element, for example, a resistor, near the thermocouple junction areas. ***Power (for example, an applied voltage or current)*** then could be continuously applied to the heating elements or could be applied in bursts in time (averaging effects) until the ***thermocouple junction temperatures are elevated by few degrees, for example, 5-10 degrees***." (EX1020, 31:51-67.) Slaby's resistor, which converts electrical energy into heat using resistance and raises the temperature of the thermocouple junction, is a thermoelectric actuator, consistent with the description of the '720 patent. (EX1001, 26:25-27.)

326. A POSITA would have understood that Slaby's thermal energy emitter/detector devices and heating elements (*e.g.*, resistor) are heaters that are configured for affecting a temperature of a solid, a liquid, or a gas object, by conduction, convection, force convection, thermal radiation, or by transfer

of energy by phase changes. A POSITA would understand that there are three ways to transfer thermal energy from one object to another, such as conduction, convection, and radiation, and the implementation of any one of them to heat up another object or itself would be predictable and straightforward. "Thermal energy may be transferred from one object to another by three ways: conduction, convection, and radiation." (EX1030, 121.) For example, Slaby's resistor are heaters that are placed near thermocouple junctions to affect the temperature of the junctions and elevate their temperatures. (EX1020, 31:51-67.)

327. Second, Slaby discloses thermoelectric actuators that are coolers. Slaby discloses that "user computer device 102 may ***decrease a temperature*** of the temperature sensitive user interface 108, and more particularly of thermal energy emitter/detector devices 110, through use of a ***thermoelectric cooling system*** or a liquid coolant system. For example, user computer device 102 may activate a ***fan*** (not shown) included in the user computer device or may ***decrease a temperature*** of the temperature sensitive user interface through use of a thin-film thermoelectric material[.]" (EX1020, 32:36-44.) This thermoelectric cooling system/liquid coolant system decreases the temperature of the temperature sensitive user interface 108, and thus, is configured for affecting a temperature of a solid, a liquid, or a gas

object, by conduction, convection, force convention, thermal radiation, or by transfer of energy by phase changes. For example, a POSITA would understand that "***[f]orced convection*** of air is produced by a ***fan*** or blower." (EX1030, 125.)

328. It would have been obvious to a POSITA to combine the teachings of Lindahl-Williams and Slaby. (*Supra* § IX.G.3.)

5. Claim 29: "The device according to claim 23, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

329. As discussed, Slaby discloses thermoelectric actuators (*e.g.*, thermal energy emitter, heating elements, thermoelectric cooling system) in the device. (*Supra* §§ IX.G.3, IX.G.4.) A POSITA would understand (and it would be obvious) that these thermoelectric actuators directly or indirectly affect, change, produce, or create a physical phenomenon, namely, temperature as claimed. *Id.*

330. It would have been obvious to a POSITA to combine the teachings of Lindahl-Williams and Slaby. (*Supra* § IX.G.3.)

H. Ground 8: Williams, Lim, and Benesty Render Obvious Claims 1, 4-26, 29

1. Summary of Lim

331. 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***." (*Id.*) Lim discloses that the controlling device may be a ***smart device 110*** in the form of a smart phone, as shown in FIGs. 1b and 5b. (*Id.*, [0023].)

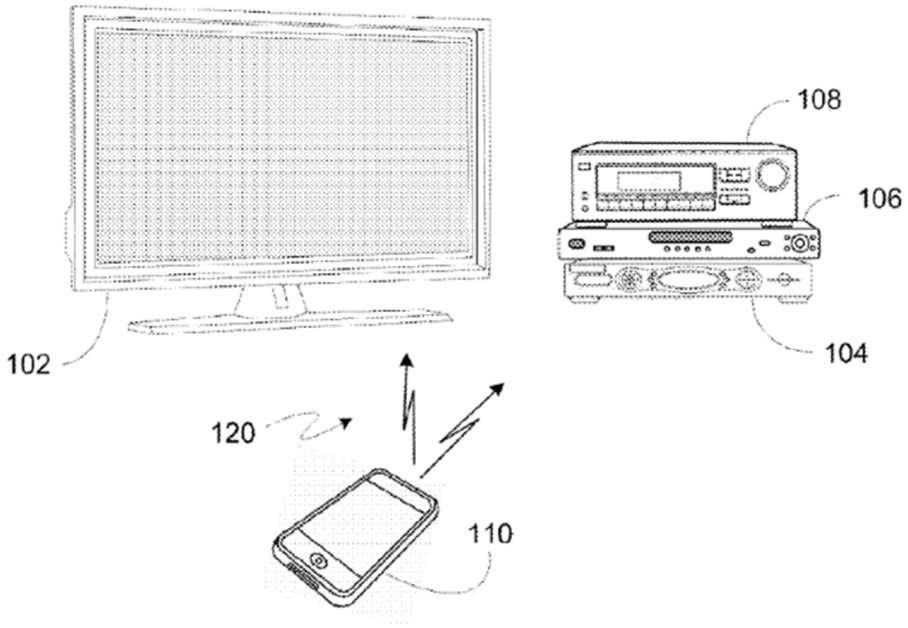


Figure 1b

332. Lim discloses that the ***device*** receives ***speech commands*** and responds with a ***visual or audio output*** from the device itself. For example, Lim discloses a "setup and configuration" stage for the ***controlling device***. (EX1005, [0025].) During this stage, the ***controlling device*** can "prompt the

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

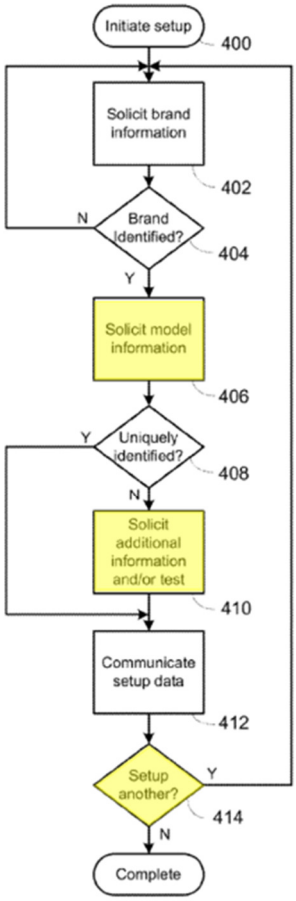


Figure 4 "Set up and configuration of controlling device" Lim, [0025].

333. In Lim, the speech commands are transmitted to a ***server*** for processing. Further, Lim discloses (1) the ***server*** is accessible via the ***Internet*** and (2) the ***device*** connects to the ***Internet*** via ***WiFi*** through a gateway device such as a router. For example, Lim discloses that "at step 402 the ***controlling device*** may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. . . . The resulting spoken syllables may be digitized and presented to a speech recognition means for processing . . . all or part of said speech recognition

processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the ***Internet*** or PSTN." (EX1005, [0027], [0034]-[0035].)

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

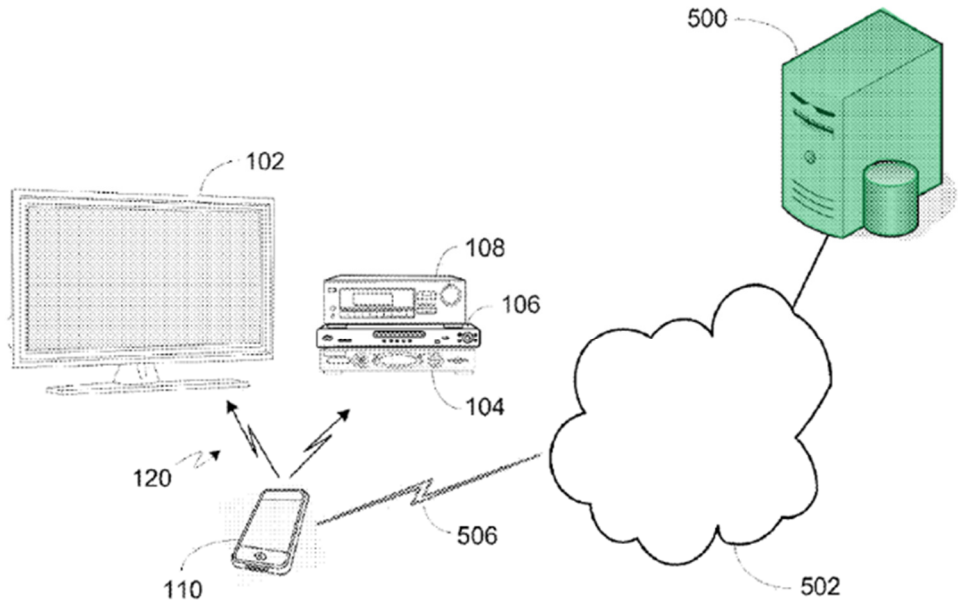


Figure 5b

335. Lim discloses that the ***device*** has a microphone, a display, a speaker, various sensors, and WiFi/Bluetooth capabilities. "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include .

. . processing means 300 which may comprise both an application processing section 302 and an RF/DSP processing section 304; an ***LCD display 306***; a keypad 308 which may comprise physical keys, touch keys overlaying LCD 306, or a combination thereof; a subscriber identification module (SIM) card 310; memory means 312 which may comprise ROM, RAM, Flash, or any combination thereof; ***WiFi and/or Bluetooth wireless interface(s) 314***; a wireless telephony interface 316; power management circuitry 318 with associated battery 320; a USB interface 322 and connector 324; an audio management system 326 with associated ***microphone 328***, ***speaker 330***, and headphone jack 332; built-in or add on IR communication means comprising a transmitter and/or receiver 336 with associated IR output diode(s) 334; and various optional accessory features 338 such as a digital camera, GPS, accelerometer, etc." (EX1005, [0023].)

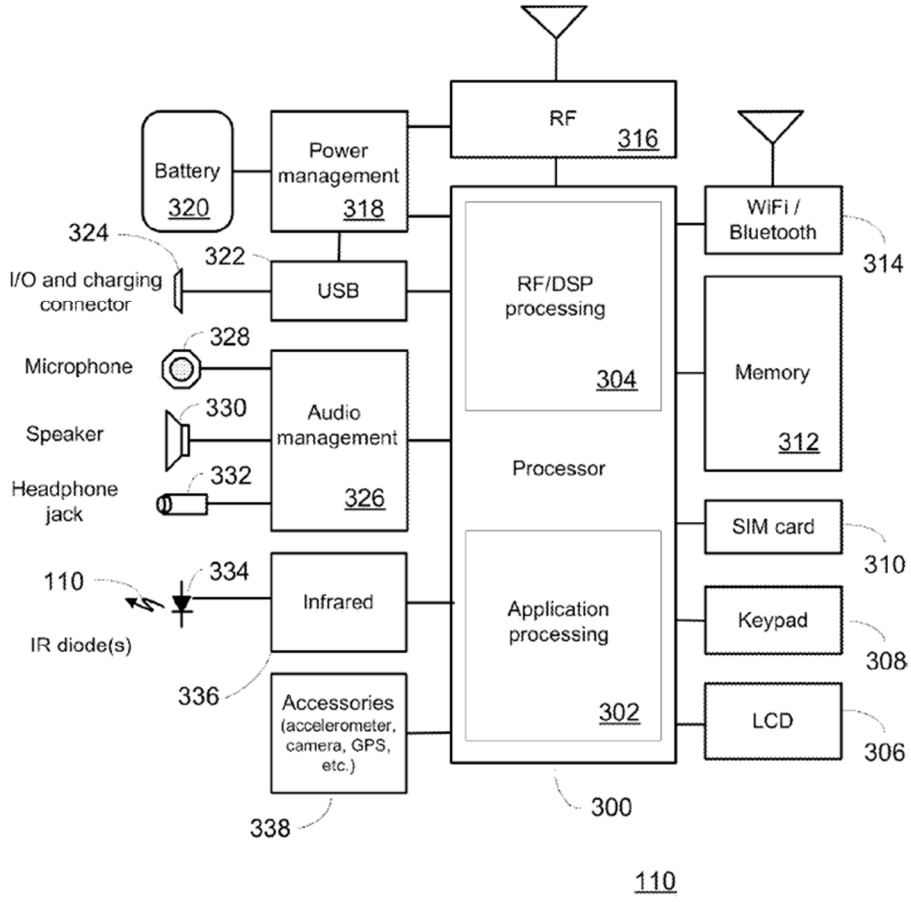


Figure 3

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

336. At the time of the invention, it would have been obvious for a POSITA to combine Williams, Lim, and Benesty.

337. First, at the time of the invention, it would have been obvious for a POSITA to combine Williams and Lim. Williams and Lim both relate to server processing and providing feedback to users in response to server communications. (*See* EX1006, [0027], [0036]; EX1005, [0007].) Thus,

both Williams and Lim are analogous art from the same field of endeavor and they both relate to the same technical problems (*e.g.*, processing and responding to speech commands).

338. Lim teaches providing visual feedback in addition to audio feedback to users. For example, the visual output may "take the form of an illuminated ***LED*** and/or '***speak' symbol***, a ***displayed text or image***[.]" (EX1005, [0027], [0019].) Thus, the Williams-Lim combination would result in a system that provides visual feedback in response to a voice command. For example, Williams' audio interface device would include a display and/or LEDs to provide visual feedback to the user. The display would provide visual feedback (*e.g.*, texts, images) in response to data (*e.g.*, processing results) received from the server. The LEDs would serve as one or more status lights, which turn on/off and/or change color to indicate different device states and modes (*e.g.*, based on processing results received from the server), or to prompt user to provide voice input as taught by Lim, in response to data received from the server. In Williams-Lim, the visual feedback can be provided in addition to certain audio feedback (*e.g.*, before, after, or concurrently) or provided instead of certain audio feedback sometimes. For example, Williams' server can receive a voice command and, in addition to sending audio 222, sends some processing results (e.g.,

text, image) as taught by Lim to the interface device for display or operating LEDs.

339. Incorporating a display and/or LEDs in Williams' audio interface device yields the predictable result of improving usability of the device by enabling multimodal feedback. Visual feedback would allow the device to present information (*e.g.*, device status, network connectivity, authorization state, prompts, confirmation messages, error codes) that may be difficult to convey by speech alone or that may be missed in noisy environments. A POSITA would also have understood that visual indicators such as LEDs provide low-power, immediate notifications (e.g., "listening," "processing," "message received," "microphone muted," or "network disconnected") that can alert the user before or without playing audio. Incorporating these known visual feedback mechanisms would predictably improve the user experience of Williams' system by reducing ambiguity and increasing the likelihood that the user notices and correctly interprets the outputs.

340. Further, incorporating an indicator light in Williams' audio interface device yields the predictable result of improving usability of the device. Changes in the light can indicate the device's current state or mode to a user, *e.g.*, whether the device is waiting for the server, whether the device has received data from the server, whether the device is playing the response.

The indicator light can also serve as an additional prompt for the user to provide voice data. These benefits align with common design incentives in consumer electronics toward multimodal user interfaces, and would have prompted application of Lim's display/LEDs to Williams' interface devices. Because both Williams and Lim relate to capturing and processing human voice data and providing user feedback in response to data received from server, a POSITA would be motivated to use the known techniques of Lim to improve the usability of Williams' interface device in the same way by enabling multimodal feedback in the form of visual cues, allowing users to understand the shifts in device status and modes through the changes in the LED light, and to understand when the device requires additional voice input based on processing results received from the server.

341. The known benefits of Lim would prompt the application of Lim in the same field based on design incentives and market forces and yield the predictable result of enhancing the usability of the system. Further, it would have been obvious to try combining the teachings of Williams and Lim. There are only a finite number of identified and predictable solutions for providing feedback to users, indicating to a user the device state or mode, or prompt user to provide speech input, such as visual cues, auditory cues, tactile cues, physical cues, etc. The combination is straightforward and readily

achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Lim has successfully demonstrated that visual user feedback means improves usability, and thus there would have been a reasonable expectation of success for combining Williams and Lim.

342. Further, it would have been obvious to a POSITA to apply Benesty's techniques of including additional microphones to Williams-Lim for capturing human voice data. Williams, Lim, and Benesty are all related to capturing and processing human audio data. (*e.g.*, EX1006,[0003], [0016], [0023], [0037]; EX1005, [0019], [0023]; EX1010, 1021, 980.) Thus, Williams, Lim, and Benesty are analogous art to the '720 patent from the same field of endeavor and all relate to the same technical problems (*e.g.*, capturing and processing human speech data). (*Id.*; EX1001, 105:48-52, 105:61-106:4.)

343. Benesty discloses hands-free audio communication and speech processing techniques. Benesty states that "hands-free audio communication is now a major feature in mobile communication systems as well as audio and video conferencing systems. One problem that becomes evident to users of these systems is the decrease in communication quality due to the pickup of room reverberation and background noise." (EX1010, 1021.) The

Williams-Lim-Benesty combination would result in a system where multiple microphones as taught by Benesty are incorporated in Williams-Lim to capture human voice data to improve the quality of captured audio. For example, Benesty's microphone array(s) would be included in Williams' audio interface device 116 to capture the first human voice data. Because Williams, Lim, and Benesty relate to audio capturing and processing, a POSITA would be motivated to use the known techniques of Benesty to improve Williams-Lim's ability to capture and process audio. Even in use cases where background audio is desired, a microphone array can capture richer audio information, enabling more efficient and accurate processing, for example, separating or balancing speech and background noise when desired.

344. Combining the teachings of Williams, Lim, and Benesty as such would yield predictable results of effectively improving audio capturing and processing. Further, it would have been obvious to try combining the technical features of Benesty with the teachings of Williams-Lim, as there is a finite number of identified, predictable resolutions, without requiring undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of potential solutions to improve audio signal collection, such as

additional software processing or incorporating additional hardware components. The known benefits of Benesty would prompt the application of Benesty in the same field based on design incentives and other market forces to produce a system architecture that allows accurate capturing and processing of human voice data.

3. Claim 1

a. 1[PR]: "A device comprising:"

345. Williams-Lim-Benesty discloses a device. For example, Williams discloses an audio interface device, *e.g.*, a smart-home hub. Williams states that "the user's home 102 may be equipped with one or more on-premises ***audio monitoring devices 116*** . . . having a microphone and a network interface". (EX1006, [0016].) "The ***interface device*** may also have a speaker or other form of audio output." *Id.*

346. FIG. 2 provides additional details on Williams' device. FIG. 2 "is a block diagram illustrating additional details regarding the system of FIG. 1." (EX1006, [0004].) With reference to FIG. 2, Williams discloses that "[a]n individual ***interface device 116*** may include a microphone 202 that is configured to detect ambient noise, sounds, and speech. The ***interface device*** 102 [*sic*] may also include a network interface 204 that facilitates network communications with various entities, including the application engine 118

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

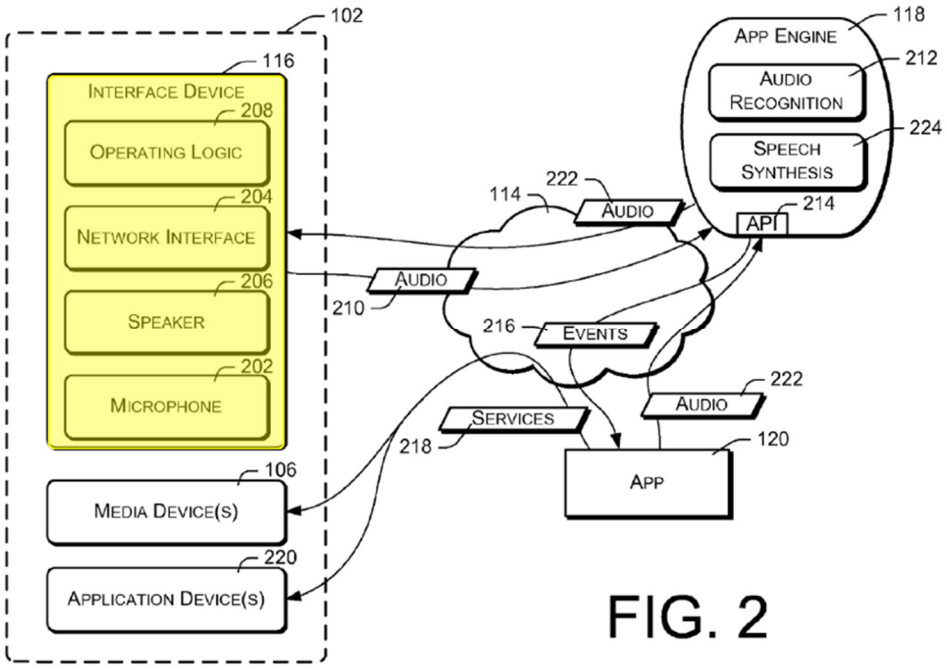


FIG. 2

347. Additionally or alternatively, Lim also discloses a device. Lim discloses that "a ***controlling device*** may be provisioned with means to receive voice input from a user, e.g., a microphone". (EX1005, [0007].) Lim discloses that the ***controlling device*** may be "a ***smart device 110***, which device may comprise any of a smart phone, a tablet computer, a portable email device, an e-book, a hand-held games and/or game controller, a portable media player, etc." as shown in FIGs. 1b and 5b. (*Id*., [0023].)

b. 1[A]: "multiple microphones for capturing first human voice data;"

348. Williams-Lim-Benesty discloses ***multiple microphones*** for capturing ***first human voice data***. Both Williams and Lim disclose a device comprising at least one microphone for capturing human voice data. For example, Williams discloses or suggests ***microphones***. Williams discloses that "audio interface devices 116 can be placed within the home 102 so that their ***microphones*** detect ambient sound or noise within the home 102". (EX1006, [0007]; *see also* [0019], claims 1-3 and 21.) To the extent that Williams does not expressly disclose that a single audio interface device 116 has multiple microphones, the configuration would have been obvious to a POSITA, given common knowledge that multiple microphones can enhance audio collection and processing. (EX1010, 1021.)

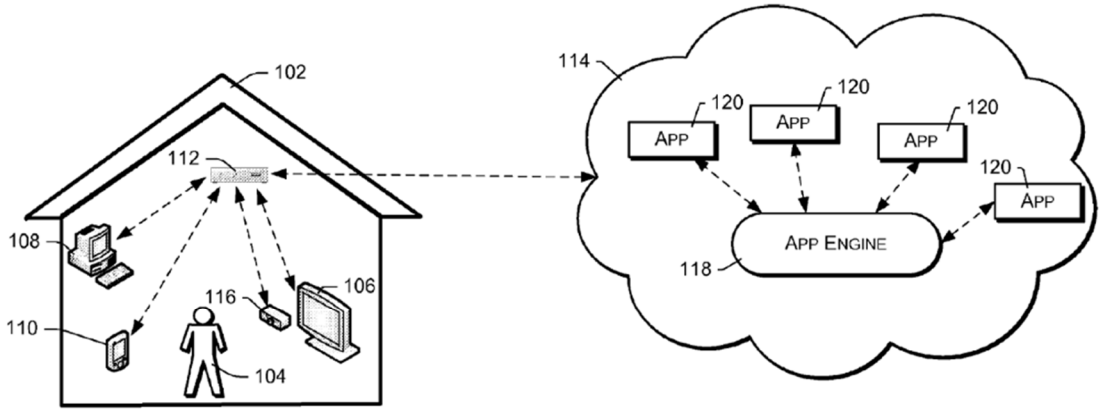


FIG. 1

349. Williams discloses capturing ***first human voice data*** by the ***microphones*** in the device. Williams's in-home interface device has "***microphone*** configured to receive ***audio from a nearby user***[.]" (EX1006, Claim 21; *see also id.*, claims 1-3.) Williams also discloses a system comprising "a cloud-based service that is accessible by on-premises audio monitoring devices to receive ***audio-based information***", where "the ***audio-based information*** comprises ***audio streams***" and/or "***user commands derived from vocal utterances of the users.***" (EX1006, claims 1-3.) Williams further discloses that "[a]udio and audio-based information may be formatted as files, ***audio streams,*** data objects, notifications, text streams, and so forth", and that "[t]he audio-based information may comprise [of] an ***audio stream*** or ***data*** derived from an ***audio stream.***" (EX1006, [0038]–[0039].) Williams

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

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

351. As another example, Williams states that "a cloud-based application may allow users to order tickets without needing a computer. ***A user may tell the application what movie they want to see***. The application may use pre-configured information about the user (such as location) to determine which theater is near the user. The application may respond to a spoken command to order movie tickets by using the speech synthesis

functionality of the system to query the user regarding movie selections and times." (EX1006, [0045].) "As another example, such an application may allow the user to ***inquire about a flight or other event***. The system might respond by sending information to the user via a device associated with the user, or by audibly responding using speech synthesis." (EX1006, [0050].)

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

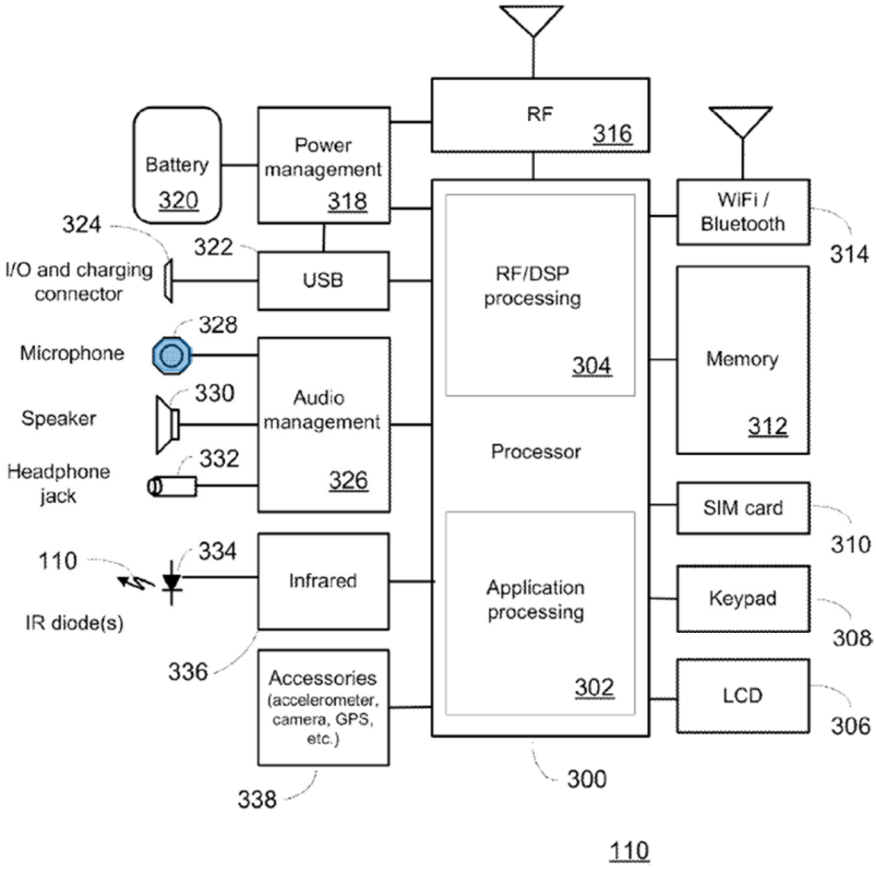


Figure 3

353. Lim discloses capturing ***first human voice data*** by the ***microphone*** in the ***device***. Lim discloses many instances where the controlling device receives speech commands ("***first human voice data***") and provides a visual or audio response on the device itself. For example, Lim discloses a "setup and configuration" stage for the ***device***. (EX1005, [0025], [0013].) "[D]uring setup and configuration of the ***controlling device***, a user may provide appliance identification via ***spoken input***. . . ." (EX1005, [0025].)

354. FIG. 4 discloses numerous examples of "***first human voice data***". Referring to FIG. 4 below, "[o]nce the setup mode has been initiated, at step 402 the controlling device may prompt the user to ***speak the brand name of the appliance*** to be added to the configuration of the controlling device." (EX1005, [0027].) Further, "[a]t step 404, if the speech recognition means is unable to uniquely identify the desired brand, step 402 may be repeated in order to solicit additional identifying information from the user. For example, a user may be presented with a short list of possible brand identities from which to select; may be prompted to ***spell the first few characters of the brand name***; may be requested to ***speak the name of a geographic region or a service provider***; etc.; as appropriate to facilitate the identification process." (*Id.*, [0028].) "Once a brand has been successfully identified, at step 406 the user may be prompted to ***provide spoken appliance type and/or model number information*** for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, e.g., to further narrow down the codeset choices." *Id.* "If the ***initially provided model information*** [*e.g.*, provided previously in step 406] is insufficient to uniquely identify the appliance, at step 410, additional user input may be solicited. For example, a user may be requested to ***indicate a device type*** if this was not already provided; to ***provide geographic region***

or service provider; to select from a series of displayed images of appliances and/or OEM remote controls; etc." (EX1005, [0029].) "Thereafter, at step 414 the user may be prompted to ***indicate if they wish to configure additional appliances***. If so, the process may be repeated starting at step 402 [where the controlling device prompts the user to speak the brand name of the appliance to be added to the configuration]." (*Id.*, [0030].) Lim discloses that each such voice data is sent to the ***remote server 500*** for "speech recognition processing" and based on the processing, the device provides a next prompt via "a displayed text or image", "a synthesized spoken message", etc. (EX1005, [0027]-[0030].) The user's voice inputs sent to the ***server*** device are all "***first human voice data***." It would have been obvious to a POSITA that the controlling device's ***microphone*** captures these voice inputs.

355. Benesty is a textbook providing an overview of well-known speech processing techniques, including speech recognition, speaker recognition, speech enhancement via noise reduction, and microphone arrays. (EX1010, table of contents.)

356. Benesty discloses that "***microphone arrays*** have the ability to offer directional gains that can ***significantly improve the quality of signal pickup in reverberant and noisy environments***." (EX1010, 1021.) "One solution is to use beamforming microphone arrays, which can offer significant

directional gain so as to result in similar audio performance to that of closely placed microphones." *Id.*

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

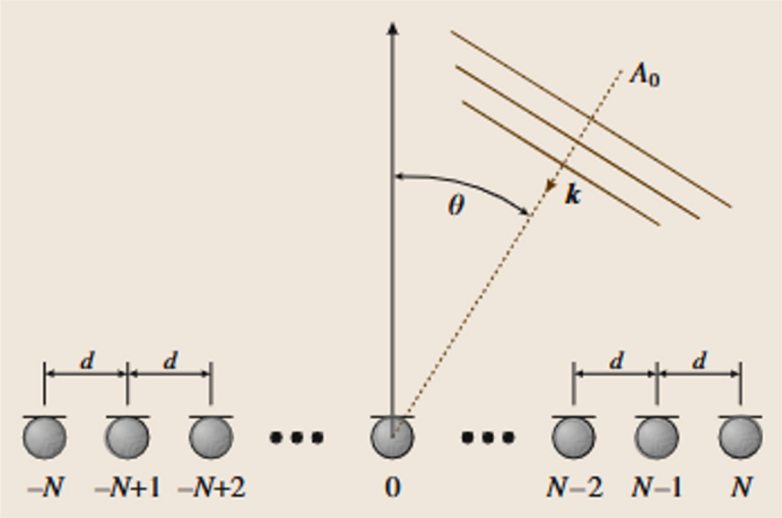


Fig. 50.1 Schematic of uniformly spaced linear array containing $2N + 1$ microphones spaced by d

358. It would have been obvious to a POSITA that incorporating additional microphones into Williams-Lim as taught by Benesty would improve the quality of the captured audio data for further processing. (*Supra* § IX.H.2.)

c. 1[B]: "a Wireless Local Area Network (WLAN) transceiver for transmitting digital data to, and for receiving digital data from, a WLAN;"

359. Williams-Lim-Benesty discloses this element. Both Williams and Lim disclose a device with a ***WLAN transceiver*** for transmitting digital data to, and for receiving digital data from, a WLAN. For example, Williams discloses that the device (i.e., ***audio monitoring devices 116***) can connect to a local ***router or access point*** that in turn connects to the Internet. With reference to FIG. 1, Williams discloses "the ***devices shown within the home 102*** are merely examples of a wide variety of devices that may be present within a user premises or associated with a user. Many such devices may have some sort of ***network connectivity***. In this case, the ***in-home devices*** connect to ***a router or access point 112***, which is in turn connected to ***a public network 114. Connection between in-home devices and the router 112 may be through wired or wireless technologies***." (EX1006, [0012].) In other words, Williams expressly discloses the ***interface device 116*** transmits and receives digital data over a wireless network through wireless router/access point 112. "[T]he ***interface devices 116*** communicate with and transmit environmental information to a cloud-based application service or engine 118, ***via the router 112*** and the network 114." (EX1006, [0020], [0023].)

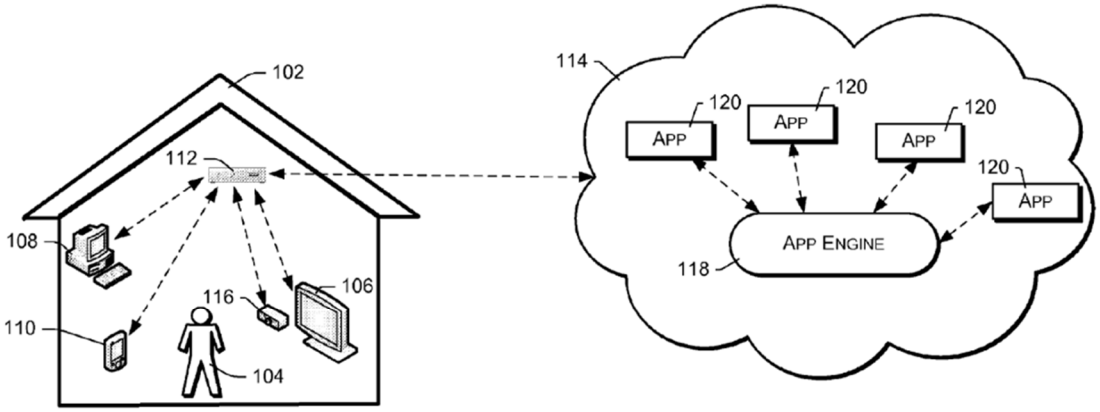


FIG. 1

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

(EX1007, 11, 100.) Such a local wireless network through a local router is exactly what is shown in FIG. 1 of Williams, consistent with the description of the '720 patent.

361. The '720 patent explains that a device can "communicate (such as send sensor info or receive actuator commands) with the router (or gateway) or the control server using the same or different ***WANs used by the router***[.]" (EX1001, 30:13-18.) The '720 patent also explains that "[t]he external network or one or more of the in-building (or in-vehicle) networks may be a Local Area Network (LAN), and the router, the first device, or the second device may comprise a LAN interface, and the LAN interface may include a LAN port and a LAN transceiver." (*Id*., 46:5-9.)

362. Alternatively, I note that Williams also discloses network 114 may be a WLAN. (EX1006, [0015].) Williams states, "[t]he network 114 may thus include various types of communication networks, including ***local-area networks***, wide-area networks, the Internet, wireless networks, cable television networks, telephone networks, cellular communications networks, etc. The techniques discussed herein may also be practiced in smaller networks, including ***localized*** and/or private networks." *Id.*

363. As discussed, Williams' device is part of a WLAN network and wirelessly communicates over the WLAN via a router 112. In light of this

disclosure, a POSITA would have understood that Williams' device has a ***WLAN transceiver***. A WLAN transceiver is the hardware component that receives wireless signals and is necessary for WLAN communication. "The mobile WLAN device 12a includes a WLAN transceiver 20 operatively coupled to antenna 22 and processing circuitry 24 operatively coupled to the WLAN transceiver to ***send and receive suitable information via a WLAN protocol***." (EX1035, 3:17-20.) "The mobile WLAN devices 12a-12n send messages 34a-34n wirelessly using their respective WLAN transceivers, ***as known in the art***." (EX1035, 3:38-40.) "'[W]ireless terminal' represents any wireless device capable of wireless communication . . . Wireless terminal 100 includes a user interface 102, memory 104, antenna 106, transceiver 110 . . . Antenna 106, coupled to transceiver 110, receives and transmits wireless signals[.]" (EX1036, [0013]-[0015].) Thus, it would have been obvious to a POSITA that Williams' device would have the necessary hardware and software components, including a WLAN transceiver, to communicate in a WLAN network as disclosed.

364. Indeed, Williams' device has a "network interface", which requires a transceiver. Williams discloses, with reference to Figure 2, that "[t]he interface device 102 may also include a ***network interface 204*** that facilitates network communications with various entities, including the

application engine 118 and one or more applications 120, via the network 114." (EX1006, [0023], [0007], [0012], [0016], [0020].) The '720 patent explains that "a network interface may include a ***transceiver*** or modem and a communication port for coupling to the network[.]" (EX1001, 28:63-67.) Considering this disclosure, a POSITA would have understood that the "***network interface 204***" contains a WLAN transceiver for coupling the interface device to the WLAN network e.g., via a WLAN interface card. (*Id*., EX1007, 99-100.) "Each computer or device that you want to join your wireless network needs a ***wireless network interface card (NIC)***. A wireless NIC provides a number of functions, including: [e]nabling a computer (or device) to connect to your home network through a wireless signal; [s]ecuring the communications path between the ***wireless router and your computers and devices*** connected over wireless; Providing a way for you to connect your computer to other wireless networks outside your home." (EX1007, 99-100).

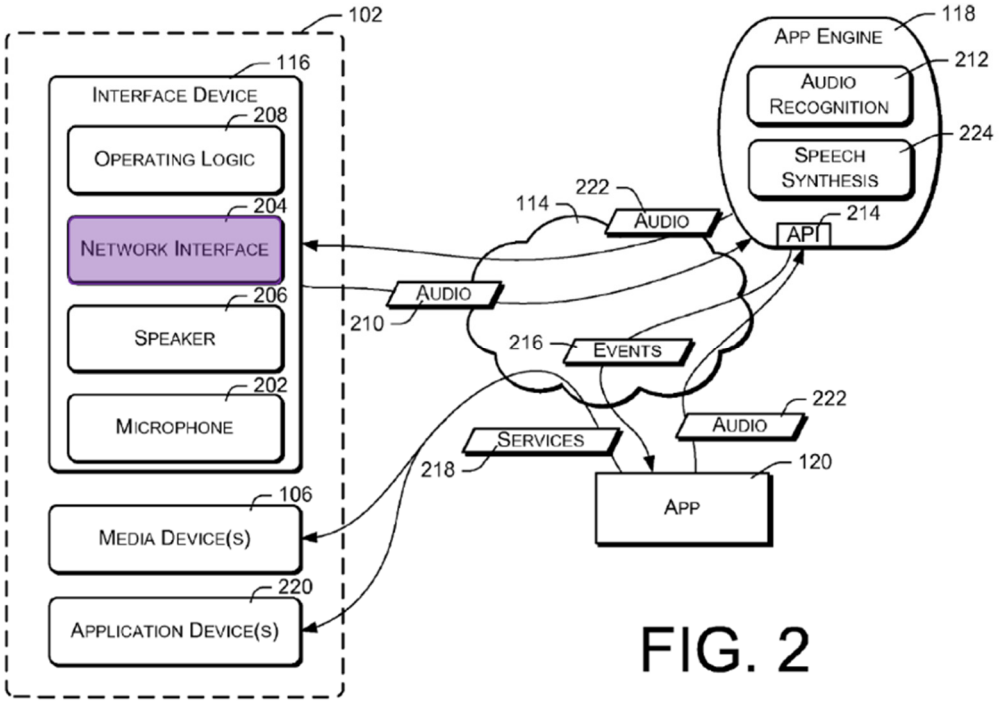


FIG. 2

365. And as discussed in detail below, Wiliams' device transmits digital data (*e.g.*, voice commands, audio 210) to a remote server and receives digital data (*e.g.*, audio 222) from the remote server over a WLAN. (*Infra* §§IX.H.3.h-IX.H.3.j.)

366. Additionally or alternatively, Lim discloses that the "controlling device 100 for use in commanding the functional operations of one or more appliances may include, . . . , Zigbee, ***WiFi*** , or Bluetooth ***transceiver***[.]" (EX1005, [0019], [0027].) A POSITA would have understood that a "WiFi transceiver" is a type of WLAN transceiver. (*Supra* § IX.A.4.c; EX1042, 600-601; EX1013,, 7:38-51; EX1001, 8:25-32.) While this statement refers

to controlling device 100, it would have been obvious to a POSITA that the same can be applied to controlling device 110.

367. With respect to the controlling device 110, Lim discloses that "[a] representative smart device . . . in the form of a smart phone, may include . . . ***WiFi and/or Bluetooth wireless interface(s) 314***; a wireless telephony interface 316[.]" (EX1005, [0023].) A POSITA would have understood that a "WiFi wireless interface" refers to the hardware and software that enable WiFi communication, which necessarily include a WLAN transceiver. "The LAN/WAN network 100d may comprise suitable logic, circuitry, ***interfaces***, and/or code that may be ***operable to enable implementation*** of one or more wired and/or ***wireless LAN*** or WAN standards and/or protocols . . . Exemplary LAN technologies may comprise, for example, those based on IEEE 802.1 standards, including, for example, ***WiFi-based networks***." (EX1021, [0067].) Indeed, in 2012, mobile phones were commonly equipped with a WiFi transceiver to enable WiFi connectivity. "Each smartphone [smartphones that can communicate with existing WiFi infrastructure] nowadays is equipped with a WiFi transceiver." (EX1037, 3.) "[W]ireless mobile devices are being made with integrated WLAN transceivers[.]" (EX1035, 1:20-23.) Thus, a POSITA would have understood that Lim's ***controlling device 110*** has a ***WLAN transceiver***.

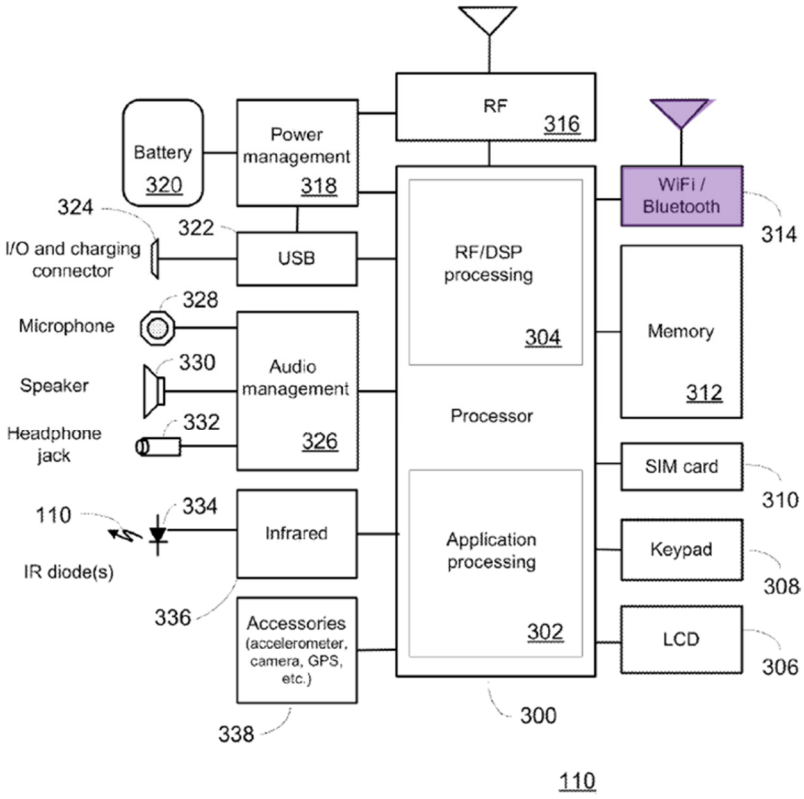


Figure 3

368. And as discussed in detail below, Lim's device transmits digital data (*e.g.*, digitized spoken syllables) to a remote server and receives digital data (*e.g.*, speech processing results) from the remote server over a WLAN. (*Infra* §§IX.H.3.h-IX.H.3.j.)

d. 1[C]: "a speaker for sounding audio data;"

369. Williams-Lim-Benesty discloses this element. Both Williams and Lim disclose a speaker for sounding audio data in the devices. (*Supra* § IX.A.4.d.) For example, Williams also discloses a ***speaker*** in the ***interface device***. Williams discloses that "the ***interface device*** may also have a

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

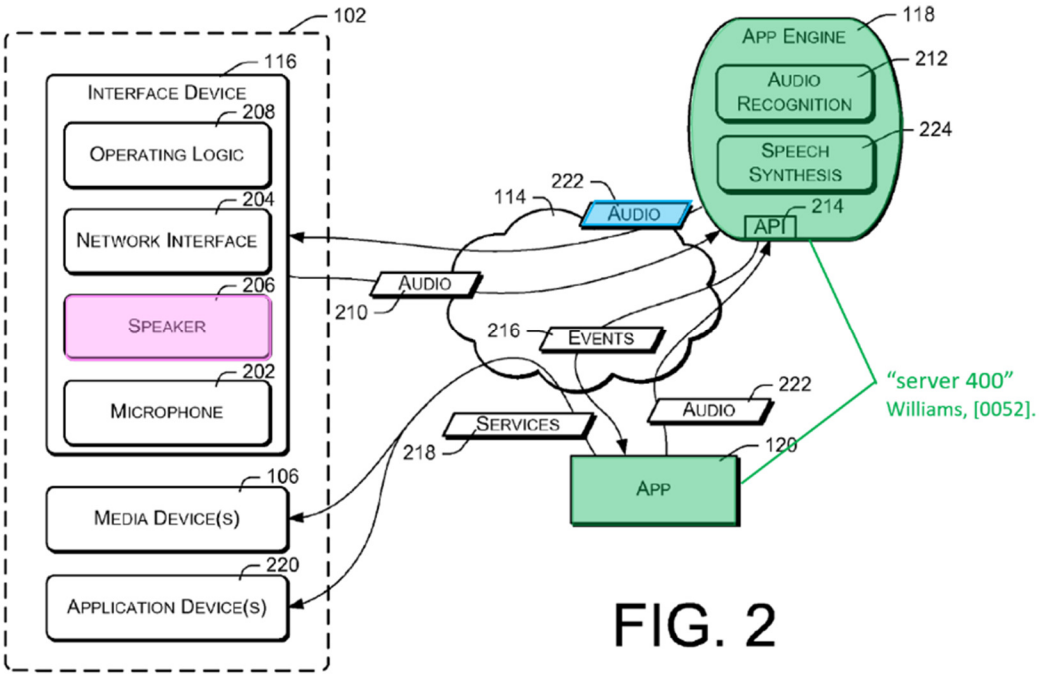


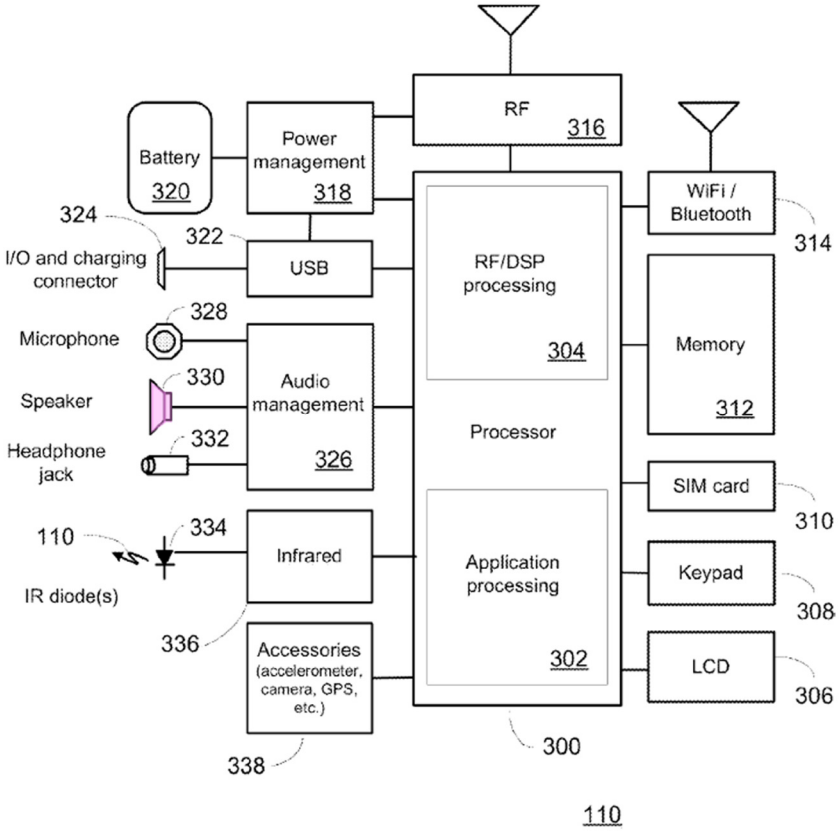
FIG. 2

370. Second, Williams' speaker is for sounding audio data, consistent with the '720 patent. (EX1001, 105:64-106:7.) For example, Williams discloses that audio may be played on the ***speaker 206***. Williams discloses that, "[t]he API 214 of the application engine 118 may also allow the application 120 to provide ***audio 222*** for eventual transmission to the ***interface device 116***, ***to be played on the speaker 206***." (EX1006, [0037].)

371. Additionally or alternatively, Lim's device has a ***speaker***. "[A] controlling device 100 for use in commanding the functional operations of one

or more appliances may include, . . . , a means 202 to provide feedback to the user (e.g., one or more LEDs, illuminable keys, display, buzzer, voice synthesizer with associated ***speaker***, and/or the like)[.]" (EX1005, [0019].)

"A representative smart device, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated microphone 328, ***speaker 330***[.]" (EX1005, [0023].)



372. Further, Lim's ***speaker*** is for sounding audio data. Lim operates its ***speaker*** in the device to provide an auditory prompt. For example, "at step 402 the controlling device may prompt the user to speak the brand name of the appliance to be added to the configuration of the

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

e. 1[D]: "an electric light source for emitting light for emitting light;"

373. Williams-Lim-Benesty discloses this element. Lim discloses that the controlling device has a ***display***. For example, "a controlling device 100 for use in commanding the functional operations of one or more appliances may include, . . . , a key matrix 216 (e.g., hard keys, soft keys such as a touch sensitive surface overlaid on a liquid crystal (LCD), . . . , a means 202 to provide feedback to the user (e.g., one or more LEDs, illuminable keys, ***display***[.])" (EX1005, [0019].) "A representative smart device, illustrated in FIG. 3 in the form of a smart phone, may include . . . an ***LCD display 306***; a keypad 308 which may comprise physical keys, touch keys overlaying LCD 306, or a combination thereof[.]" (EX1005, [0023].) As discussed, a display comprises electric light source for emitting light.

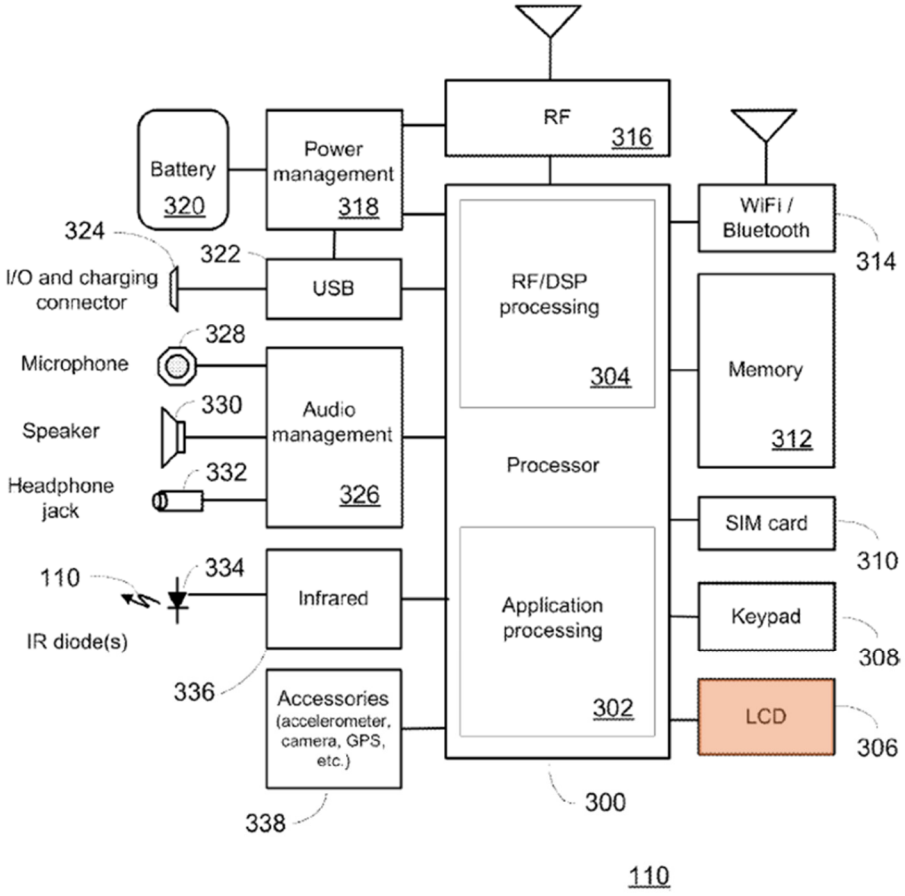


Figure 3

374. Moreover, Lim discloses that the ***device*** has lights such as LEDs. As discussed, the '720 patent explains that an "electric light source may be based on Solid-State Lighting (SSL) such as a Light Emitting Diode (LED)[.]" (EX1001, 39:10-21.) Consistent with the '720 patent, Lim discloses that "a controlling device 100 for use in commanding the functional operations of one or more appliances may include, . . . , a key matrix 216 (e.g., hard keys, soft keys such as a touch sensitive surface overlaid on a liquid crystal (LCD), . . . , a means 202 to provide feedback to the user (e.g., ***one or more LEDs***,

illuminable keys, display[.])" (EX1005, [0019].) It would have been obvious that smart device 110 is similarly equipped with ***LEDs***.

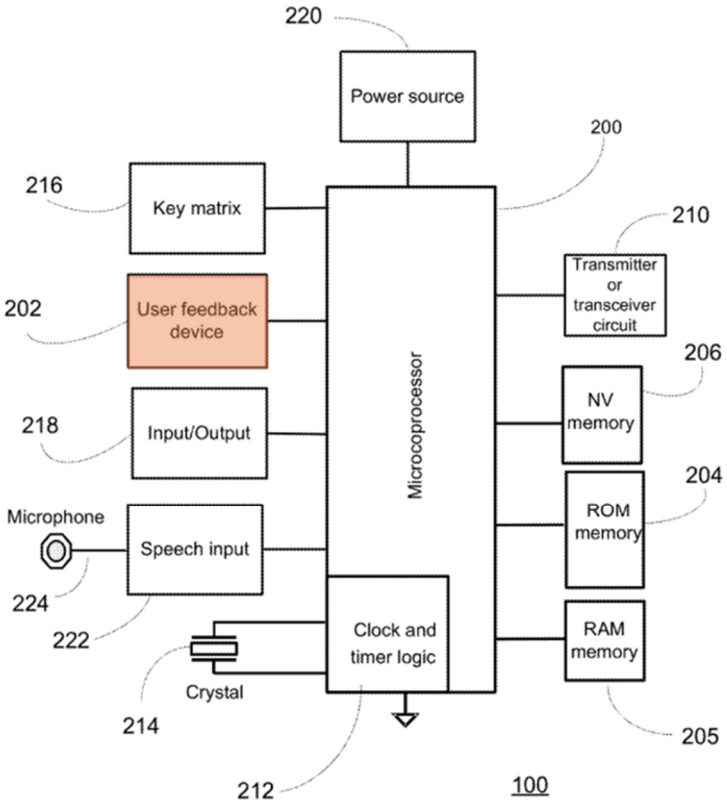


Figure 2

375. Lim's electric light sources are for emitting light. For example, Lim operates its ***display*** and ***LEDs*** in the controlling device to provide a visual prompt. For example, "at step 402 the controlling device may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. This and subsequent prompts for voice input may take the form of an ***illuminated LED*** and/or '***speak' symbol***, a

displayed text or image, a synthesized spoken message (e.g., where controlling device user feedback means 202 comprises speech synthesis means and a speaker, or where a smart device app has access to speaker 330), etc." (EX1005, [0027].)

376. It would have been obvious to a POSITA that the Williams-Lim-Benesty combination discloses this element. (*Supra* § IX.H.2.)

f. 1[E]: "an enclosure that houses the multiple microphones, the WLAN transceiver, the speaker, and the electric light source;"

377. Williams-Lim-Benesty discloses this element. Williams discloses an ***enclosure*** that houses the ***microphone(s)***, the ***WLAN transceiver***, and the ***speaker***. As discussed, Williams's device comprises ***microphone(s)***, a ***WLAN transceiver***, and a ***speaker***. (*Supra* §§ IX.H.3.b, IX.H.3.c, IX.H.3.d.) For example, with reference to FIG. 2, "[a]n individual interface device 116 may include a ***microphone 202*** that is configured to detect ambient noise, sounds, and speech. The interface device 102 may also include a ***network interface 204*** that facilitates network communications with various entities, including the application engine 118 and one or more applications 120, via the network 114. The interface device may also include a ***speaker 206***, and operating logic 208." (EX1006, [0023].) As shown in FIGs. 1

and 2 below, Williams' interface device 102 houses the WLAN transceiver, speaker, and microphone within one enclosure.

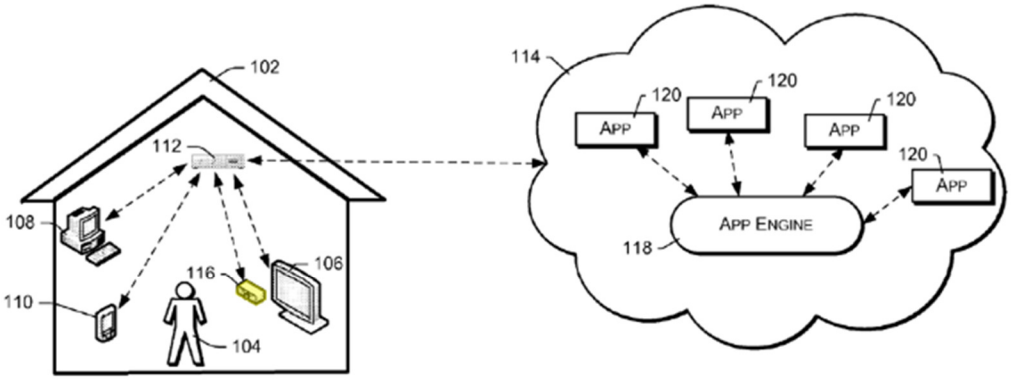


FIG. 1

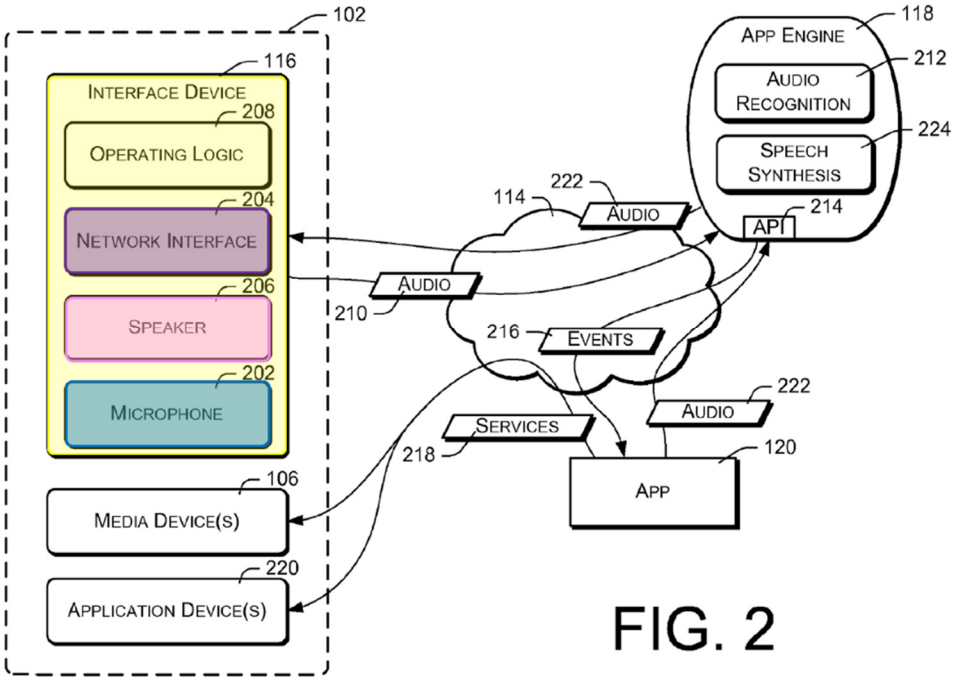


FIG. 2

378. As discussed, it would have been obvious for a POSITA to combine Williams, Lim, and Benesty. For example, it would have been obvious for a POSITA to combine Williams and Lim by incorporating electric light source(s) as taught by Lim in Williams. Further, it would have been obvious for a POSITA to combine Williams-Lim and Benesty by incorporating multiple microphones as taught by Benesty in Williams-Lim. (*Supra* § IX.H.2.) It would have been obvious that these components would all be housed in an enclosure (*i.e.*, Williams' interface device).

379. I note that Lim discloses an ***enclosure*** that houses the ***microphone***, the ***WLAN transceiver***, the ***speaker***, and the ***electric light source***. Lim's device comprises a microphone, a WLAN transceiver, a speaker, and an electric light source. (*Supra* §§ IX.H.3.b, IX.H.3.c, IX.H.3.d, IX.H.3.e.) For example, Lim's "controlling device 100, for use in commanding the functional operations of one or more appliances may include, . . . a means 202 to provide feedback to the user (e.g., one or more ***LEDs***, illuminable keys, ***display***, buzzer, voice synthesizer with ***associated speaker***, and/or the like), an input/output port 218 such as a serial interface, USB port, modem, Zigbee, ***WiFi***, or Bluetooth ***transceiver***, etc., . . . and speech input means comprising a ***microphone 224*** [.]" (EX1005, [0019].)

380. Lim's smart device 110, "illustrated in FIG. 3 in the form of a smart phone, may include as needed for a particular application, . . . an ***LCD display 306***; . . . ; ***WiFi and/or Bluetooth wireless interface(s) 314***; . . . ; an audio management system 326 with associated ***microphone 328***, ***speaker 330***, and headphone jack 332[.]" (EX1005, [0023].)

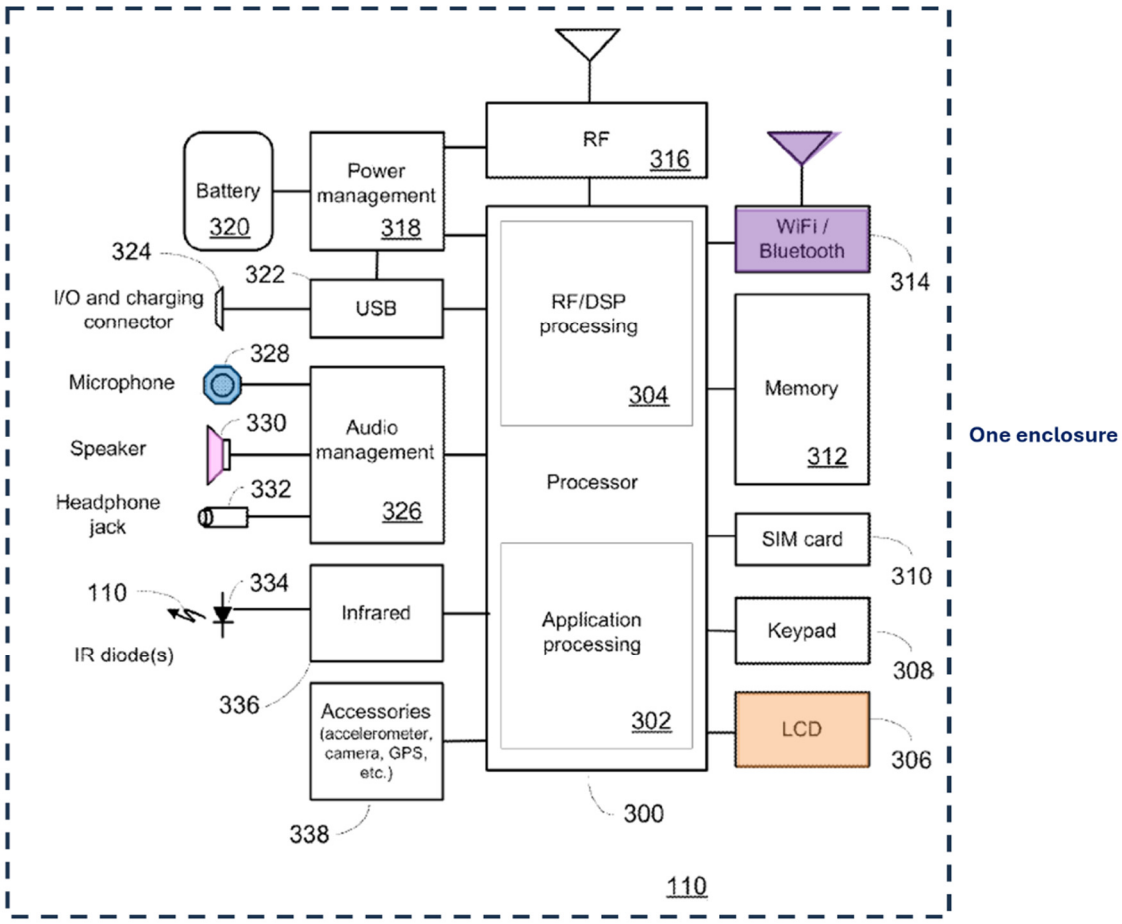


Figure 3

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

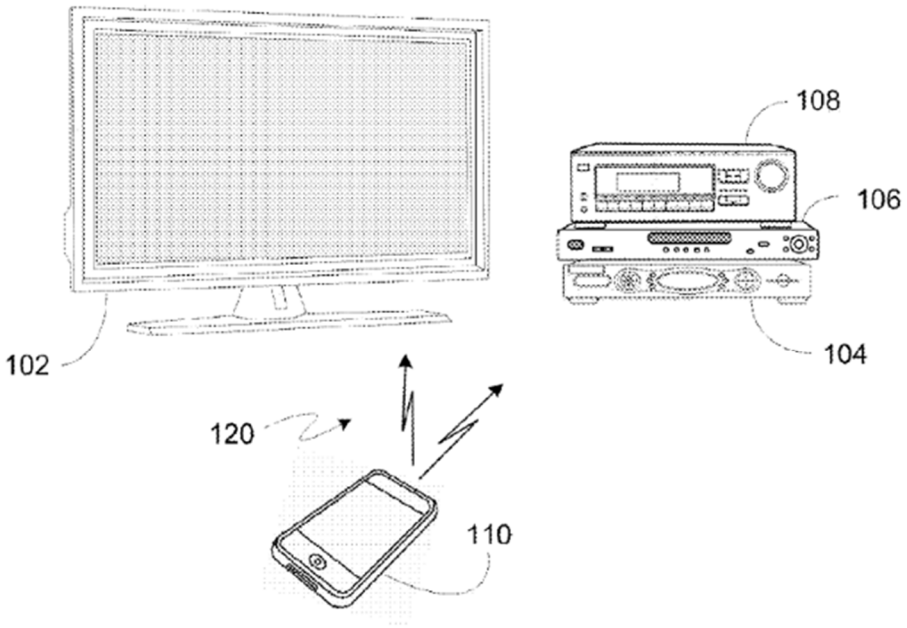


Figure 1b

382. Lim discloses that "FIG. 3 illustrates a block diagram of exemplary components of the controlling device of FIG. 1b." (EX1005, [0012].) All the exemplary components of the smart device 110,are enclosed within an enclosure of the smart device.

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

383. Williams-Lim-Benesty discloses this element. Both Williams and Lim disclose this element. For example, Williams discloses one or more processors in the device programmed with computer program instructions. Williams discloses that "[t]he interface device may also include a speaker 206, and operating logic 308. The operating logic 208 may be implemented as ***one***

or more programmable processing units, associated memory, and appropriate instructions stored in the memory and executed by the processing unit. Other types of logic may also be used to configure the interface device 116 to ***perform the functions described here***." (EX1006, [0023].) It would have been obvious to a POSITA that Williams' device comprises one or more processors programmed with computer program instructions that, when executed, cause the device to perform the following recited operations. (*Infra* §§IX.H.3.h-IX.H.3.j.)

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

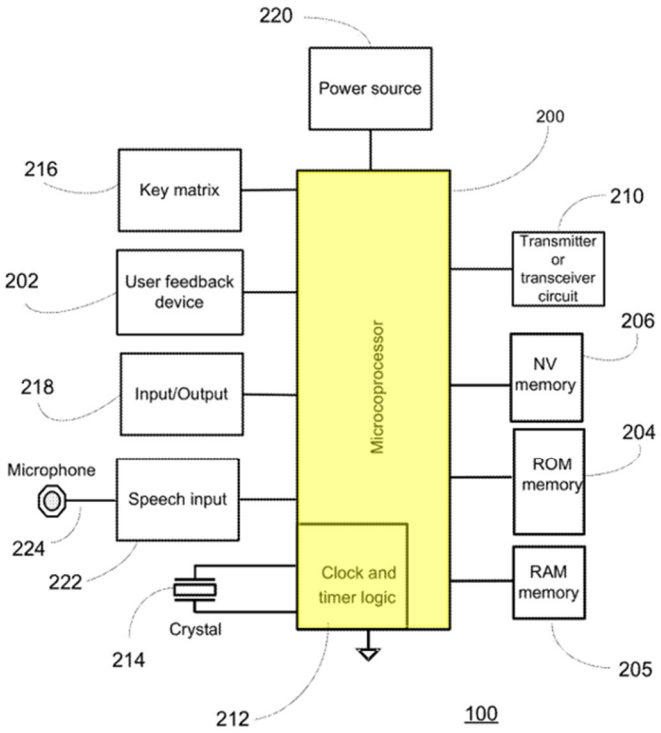


Figure 2

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

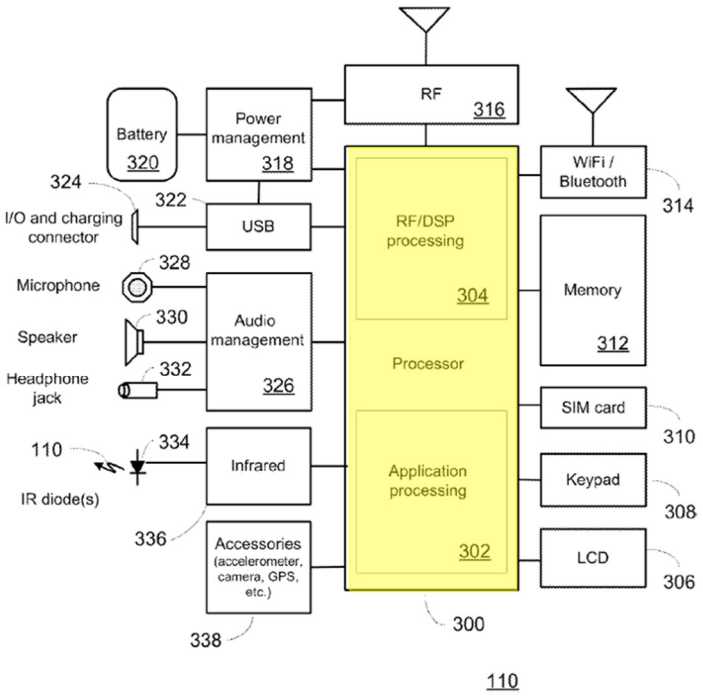


Figure 3

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

operations that may be performed by the controlling device 100 may include ***displaying information/data***[.]" (EX1005, [0021]; *infra* §§IX.H.3.h-IX.H.3.j.)

h. 1[G]: "send, to a server device over the Internet via the WLAN by the WLAN transceiver, at least part of the first human voice data captured by the multiple microphones;"

387. Williams-Lim-Benesty discloses this element. Both Williams and Lim disclose sending, to a server device over the Internet via the WLAN by the WLAN transceiver, at least part of the first human voice data captured by the multiple microphones. For example, Williams discloses a ***server device***. Williams discloses that "a ***server 400*** that may be used to implement the functionality of the application engine 118 and/or any of the applications 120." (EX1006, [0052].)

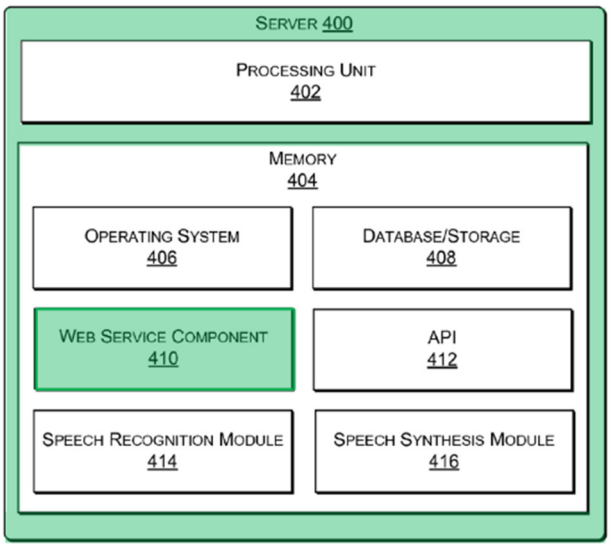


FIG. 4

388. Williams' server is over the Internet. Williams uses "cloud" and "Internet" interchangeably. (*See, e.g.,* EX1006, [0001] ("Homes and other user premises are increasingly equipped with always-on ***Internet or 'cloud'*** connectivity.").) As shown in FIG. 1, the "***server*** . . . that may be used to implement the functionality of the application engine 118 and/or any of the applications 120" is in the cloud network 114. Thus, 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).

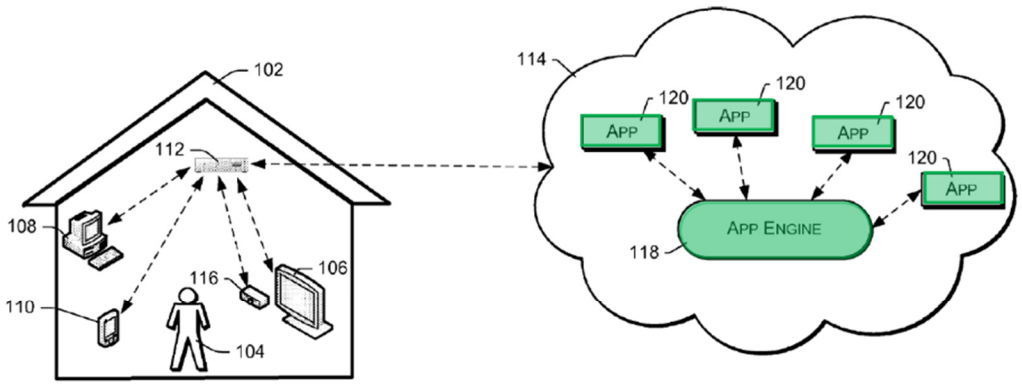


FIG. 1

389. Indeed, Williams discloses that "the network 114 represents what has come to be known as ***the 'cloud,' which may include a vast array of network-based entities, servers, clients, applications, and services.*** Blocks

within the network 114 represent cloud-based applications and/or services that are relevant to this discussion." (EX1006, [0014].) "The network 114 may be a conglomeration of various discrete networks and systems, interconnected in various ways so that for functional purposes they appear more or less as a unified network. The network 114 may thus include various types of communication networks, including local-area networks, wide-area networks, ***the Internet***, wireless networks, cable television networks, telephone networks, cellular communications networks, etc. The techniques discussed herein may also be practiced in smaller networks, including localized and/or private networks." (EX1006, [0015].) Further, Williams discloses that the "applications 120 may be ***web-based applications*** that are accessible to users via a network browser, operating in conjunction with the computer 108 or another Internet terminal or access device." (EX1006, [0021]; *see also id.*, [0055], [0007]-[0009], Abstract.)

390. Second, Williams's device sends at least part of the ***first human voice data*** captured by the ***multiple microphones*** to the ***server***. Williams discloses that "[i]n operation, the application engine 118 [of ***server 400***] receives ***audio-based information*** from the interface devices 116. The audio information may include ***audio streams, user commands or notifications***

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

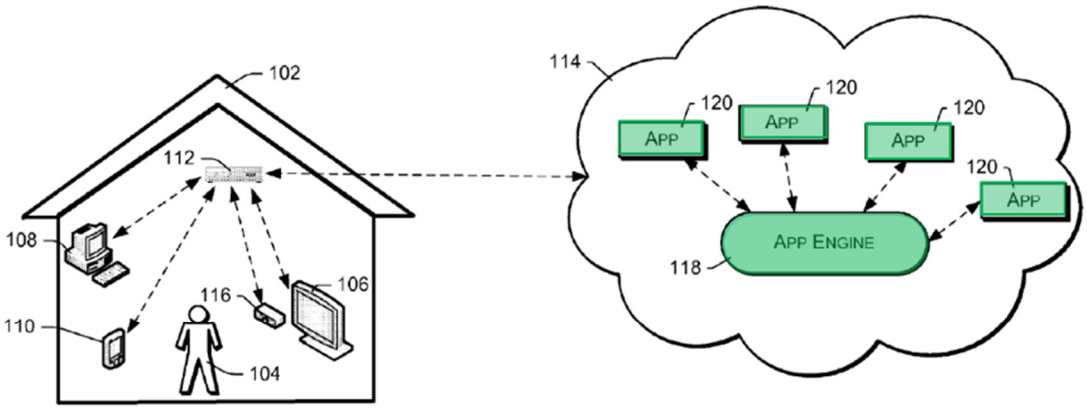


FIG. 1

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

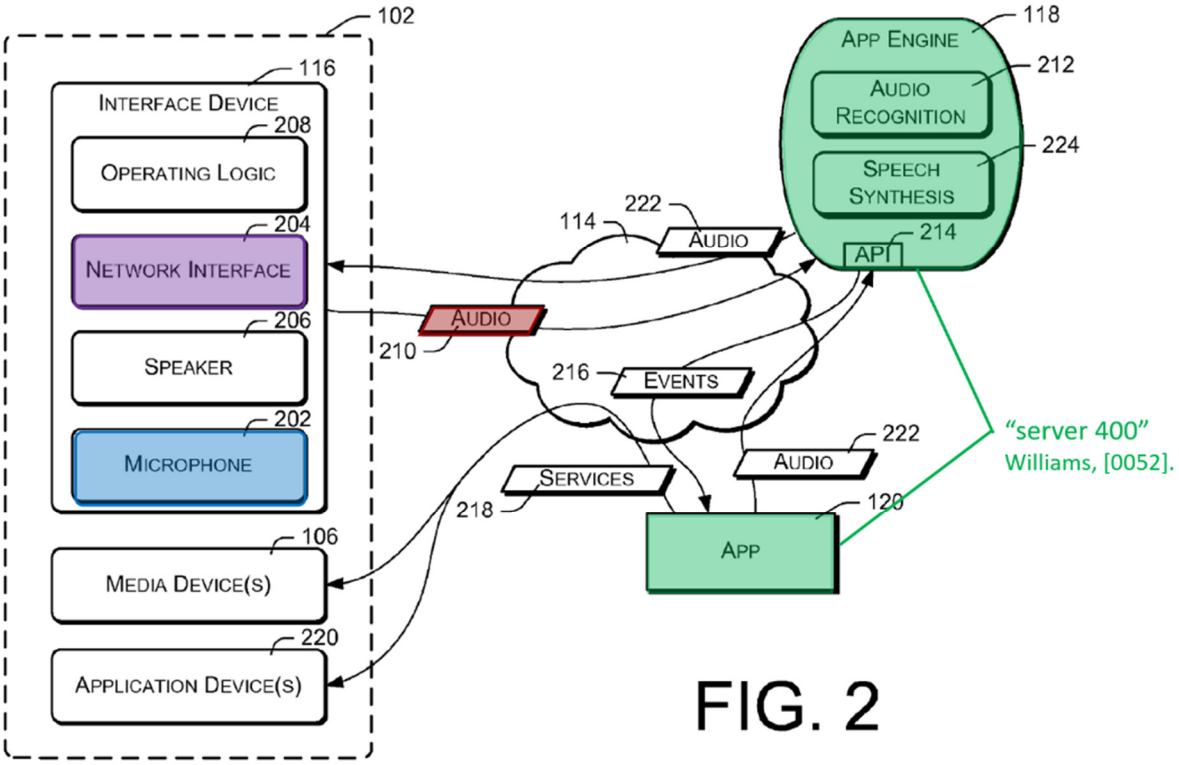


FIG. 2

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

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

393. Williams discloses that the device sends at least part of the ***first human voice data*** captured by the ***multiple microphones*** to the ***server*** device over the ***Internet via the WLAN*** by the WLAN transceiver. As shown in FIG. 1, the cloud-based ***server 400*** is over the Internet (*e.g.*, network 114) and is connected to the in-home router 112. As discussed, the in-home router is part of the WLAN network. (*Supra* §IX.H.3.c; EX1007, 11, 100.) Thus, a POSITA would have recognized that Williams discloses a network architecture where the ***audio data*** would be sent by interface device 116 to the cloud-based server device over the Internet via the WLAN (*e.g.*, via the router 112). Indeed, Williams discloses "the interface devices 116 communicate with and transmit environmental information to a cloud-based application service or engine 118 [of ***server 400***], ***via the router 112 and the network 114***." (EX1006, [0020]; *see also Supra* § IX.H.3.c.)

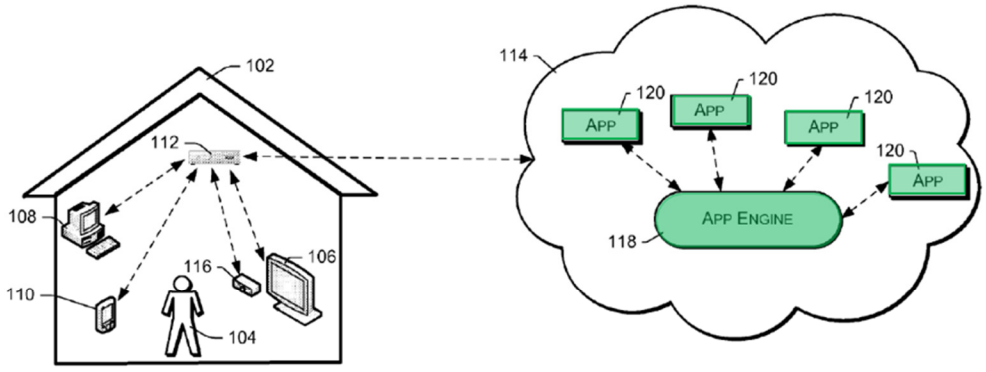
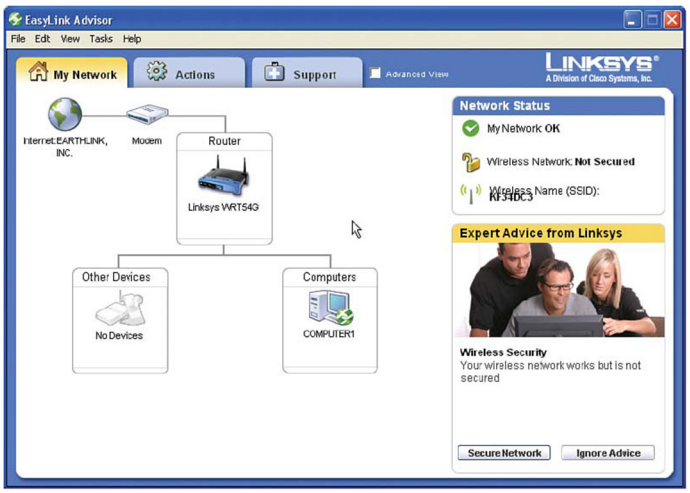


FIG. 1

394. Williams' network architecture in FIG. 1 is consistent with the descriptions of the '720 patent. The '720 patent explains that "[t]he system may comprise a router in the building, . . . an Internet-connected control server (referred herein also as 'cloud server' and 'gateway server') external to the building (or vehicle) storing the control logic, and communicatively coupled to the router over the Internet via the external network." (EX1001, 32:28-45.)

395. Indeed, this was the common configuration for a home network. (*Id.*; EX1007, 11, see also, *Id.*, 100.) For example, "[f]igure 7-28 shows what the initial ***network*** should look like if things are working properly, including ***connectivity to the Internet***." (*Id.*, 96; *see also id.*, Figure 7-28.)

Figure 7-28 EasyLink Graphical View of Your Network



396. It would have been obvious that such communication by Williams' interface device is via its network interface 204, which includes the ***WLAN transceiver***. (*Supra* § IX.H.3.c.) Indeed, Williams' "network interface 204" "facilitates network communications with various entities, including the application engine 118 and one or more applications 120, via the network 114." (EX1006, [0023], [0007], [0012], [0016], [0020].)

397. I note that network 114 by itself also discloses both the WLAN and the Internet. "***The network 114*** may thus include various types of communication networks, including ***local-area networks***, wide-area networks, ***the Internet***, ***wireless networks***, cable television networks, telephone networks, cellular communications networks, etc." (*Supra* § IX.H.3.c.) In FIG. 2, ***audio 210*** is sent via the wireless network 114. Williams discloses that the interface device 116, the application engine 118, and the application

120 [of ***server 400***] all communicates over the same wireless network (e.g., network 114). (*Supra* § IX.H.3.c.)

398. I note that Williams mentions "[s]ome of the applications 120 may serve as resources for programs that execute on local devices such as the computer 108 or the personal device 110. Such local programs may operate in conjunction with the cloud-based applications 120 to provide various different services and functionality, ranging from email services to games." (EX1006, [0021].) A POSITA would have understood that this does not mean that the server itself is local; rather, the server works with local devices. Williams refers to a local program that would provide a graphical user interface (e.g., a graphical user interface for an email inbox or a game), which can update based on messages received from an application 120 that is hosted on a remote server device. Indeed, "most Internet-based email accounts use POP3 (Post Office Protocol v.3). POP3 allows an email client to ***collect the email on a remote server (POP server)*** . . . When an email client wishes to make use of the service, it establishes a TCP connection with the server host. When the connection is established, the POP3 server sends a greeting. The client and POP3 server then exchange commands and responses." (EX1017, 664.)

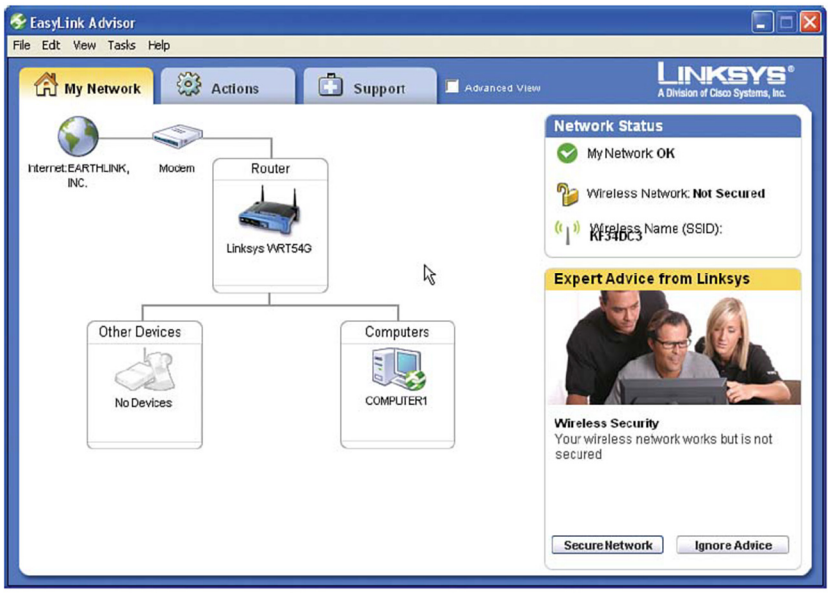
399. Additionally or alternatively, Lim also discloses this element. For example, Lim discloses a ***server device***. Lim discloses that "all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as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

400. Further, Lim's device communicates with the ***remote server 500*** over the Internet via the WLAN by the ***WLAN transceiver***. Lim discloses that "access to such network [i.e., 'a network 502 such as the Internet or PSTN'] may be . . . via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].) In other words, Lim's device is connected to a WiFi network (i.e., WLAN) via a router, which in turn accesses the ***remote server 500*** over the Internet. Indeed, this was the common configuration for a home network. As discussed in a textbook titled "Wireless Home Networking Simplified", "with a wireless LAN, the computers have to find a way to connect, or associate, themselves to the LAN, because they are not physically connected. The way they do this is by ***tuning themselves to the wireless router*** using the name of the router . . . Once a computer has

connected, or associated, itself to the access point, it becomes part of the LAN . . . ***Now the computer can send and receive information*** to and from other computers on the LAN or ***to and from the Internet by transmitting over wireless to the wireless router (or access point)***. The wireless router acts as the central communication routing point." (EX1007, 11, 100.)

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

Figure 7-28 EasyLink Graphical View of Your Network



402. As discussed, Lim's device includes a ***WLAN transceiver*** which is the hardware component that receives wireless signals and is necessary for WLAN communication. (*Supra* § IX.H.3.c; EX1021, [0067]; EX1035, 3:17-20, 3:38-40; EX1036, [0013]-[0015].) Thus, it would have been obvious that

any communication between the device and a server device over the Internet via the WLAN (*e.g.*, WiFi) is by the ***WLAN transceiver***. Thus, Lim discloses a device communicating with a ***server device*** over the ***Internet via the WLAN*** by the ***WLAN transceiver***.

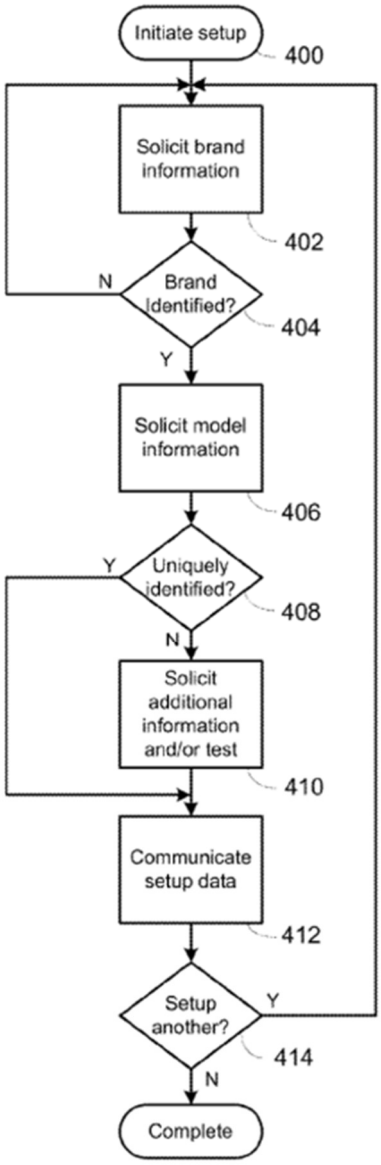
403. Second, Lim discloses that the device sends at least part of the ***first human voice data*** captured by the ***multiple microphones*** to the ***server*** over the ***Internet via the WLAN*** by the WLAN transceiver.

404. Lim discloses that "at step 402 the controlling device may prompt the user to speak the ***brand name of the appliance*** to be added to the configuration of the controlling device. . . . The resulting ***spoken syllables*** may be digitized and presented to a speech recognition means for processing and comparison to a phonetic database of brand names which have been cross referenced to appropriate codesets within the library of codesets. As appropriate for a particular embodiment, all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the Internet or PSTN. In the latter case, access to such network . . . may be via a local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router (not illustrated)." (EX1005, [0027].) In other words, the "***first human voice***

data" is sent by the controlling device over the Internet via the WLAN to a ***server device***.

405. It would have been obvious to a POSITA that other instances of "***first human voice data***" are similarly sent to the server for processing. Some of these are shown in FIG. 4. "At step 404, if the speech recognition means is unable to uniquely identify the desired brand, step 402 may be repeated in order to solicit additional identifying information from the user. For example, a user may be presented with a short list of ***possible brand identities*** from which to select; may be prompted to ***spell the first few characters of the brand name***; may be requested to speak the ***name of a geographic region or a service provider***; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].) Lim discloses that "[o]nce a brand has been successfully identified, at step 406 the user may be prompted to provide ***spoken appliance type and/or model number information*** for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, e.g., to further narrow down the codeset choices . . . At step 408 the ***spoken syllables comprising the model identity*** may be processed and compared to a phonetic database of model numbers by the speech recognition means which have also been cross referenced to codesets within a library of codesets. If the initially provided

model information is insufficient to uniquely identify the appliance, at step 410 additional user input may be solicited. For example, a user may be requested to ***indicate a device type*** if this was not already provided; to ***provide geographic region or service provider***; to select from a series of displayed images of appliances and/or OEM remote controls; etc." (*Id*., [0029].)



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

Figure 4

i. 1[H]: "play, by the speaker, an audio data received from the server device over the Internet via the WLAN by the WLAN transceiver;"

406. Williams-Lim-Benesty discloses the device plays, by the speaker, an audio data received from the ***server*** device over the Internet via the WLAN

by the WLAN transceiver. For example, as discussed, Williams discloses this element. (*Supra* § IX.A.4.i.) William discloses, with reference to FIG. 2, "API 214 of the application engine 118 [run on ***server 400***] may also allow the application 120 [run on ***server 400***] to provide ***audio 222*** for eventual transmission to the interface device 116, to be played on the ***speaker 206***." (EX1006, [0037].) "[T]he ***audio 222*** may be provided from the application 120 as a digital audio stream." (*Id*.)

j. 1[I]: "and emit, by the electric light source, in response to data received from the server over the Internet via the WLAN by the WLAN transceiver,"

407. Williams-Lim-Benesty discloses or suggests this element. For example, as discussed, Lim discloses an electric light source for emitting light. (*Supra* § IX.H.3.e.) Further, Lim discloses or suggests receiving data from the server over the Internet via the WLAN by the WLAN transceiver. As discussed, Lim discloses that the device sends at least part of the captured ***first human voice data*** to the ***server***. (*Supra* § IX.H.3.h.) It would have been obvious to a POSITA that, after sending ***first human voice data*** to the ***remote server*** for speech recognition processing, the controlling device would receive some ***speech processing result(s)*** associated with the ***first human voice data*** from the ***remote server*** in order to provide a responsive next prompt. Indeed, Lim discloses providing responsive prompts (via a display or LEDs)

following the user's voice inputs, and it would have been obvious that these responsive prompts are or derived from ***speech processing result(s)*** (i.e., data) received from the ***remote server***. A POSITA would have understood that the ***speech processing result(s)*** may be the recognized speech (e.g., text strings), a result of processing the recognized speech (e.g., semantic processing to determine meaning and adequacy of the speech), and/or information associated with the next prompt (e.g., text or audio, information to display, information triggering a visual or auditory output such as flashing a light indicator or playing a sound). (*E.g.*, EX1009; EX1038, 307 ("[t]he client side has to transmit the speech signal to the server. The server returns the recognition result for the client after processing."); EX1039, 2367 ("the server computer recognizes the speech and displays the recognized text back to the user."))

408. For example, "at step 402 the controlling device may prompt the user to ***speak the brand name of the appliance*** to be added to the configuration of the controlling device . . . The ***resulting spoken syllables*** may be digitized and presented to a speech recognition means for processing . . . all or part of said speech recognition processing . . . may be performed by programming resident on a ***remote server 500*** accessible via a network 502 such as the Internet or PSTN." (EX1005, [0027].) "At step 404, ***if the***

speech recognition means is unable to uniquely identify the desired brand, step 402 may be repeated in order to solicit additional identifying information from the user. For example, a user may be presented with a short list of possible brand identities from which to select; may be prompted to spell the first few characters of the brand name; may be requested to speak the name of a geographic region or a service provider; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].) In this example, the ***speech processing result(s)*** can be the recognized speech (e.g., text strings), a result of processing the recognized speech (e.g., determining "***if the speech recognition means is unable to uniquely identify the desired brand***"), information associated with the responsive next prompt (e.g., text or audio), and/or information to display, information triggering a visual or auditory output such as flashing a light indicator or playing a sound, depending on the distribution of speech processing between the server and the controlling device.

409. Similarly, "at step 406 the user may be prompted to provide spoken appliance type and/or model number information for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, e.g., to further narrow down the codeset choices. . . At step 408 the ***spoken syllables comprising the model identity***

may be processed and compared to a phonetic database of model numbers by the speech recognition means which have also been cross referenced to codesets within a library of codesets. ***If the initially provided model information is insufficient to uniquely identify the appliance***, at step 410 additional user input may be solicited." (EX1005, [0029].) In this example, the ***speech processing result(s)*** can be the recognized speech (e.g., text strings), a result of processing the recognized speech (e.g., determining "***[i]f the initially provided model information is insufficient to uniquely identify the appliance***"), and/or information associated with the responsive next prompt (e.g., text or audio, information to display, information triggering a visual or auditory output such as flashing a light indicator or playing a sound), depending on the distribution of speech processing between the server and the controlling device.

410. "Thereafter, at step 414 the user may be prompted to ***indicate if they wish to configure additional appliances***. ***If [the user indicates that they wish to configure additional appliances]***, the process may be repeated starting at step 402." (EX1005, [0030].) In this example, the ***speech processing result(s)*** can be the recognized speech (e.g., text strings), a result of processing the recognized speech (e.g., determining "***[i]f [the user indicates that they wish to configure additional appliances]***"), and/or

information associated with the responsive next prompt (e.g., text or audio, information to display, information triggering a visual or auditory output such as flashing a light indicator or playing a sound), depending on the distribution of speech processing between the server and the controlling device.

411. At the time of the invention, remote speech recognition servers routinely transmit results of speech recognition back to the requesting device. (*E.g.*, EX1009; EX1038, 307; EX1039, 2367.) Further, in light of Lim's disclosure of responsive prompts following the user's voice inputs, it would have been obvious to a POSITA that ***speech processing result(s)*** (*e.g.*, recognized speech, analysis of or information associated with the recognized speech, the next prompt, information to display, information triggering a visual or auditory output such as flashing a light indicator or playing a sound) is sent from the ***remote server 500*** back to the device for display. For example, "[t]he client side has to transmit the speech signal to the server. The server returns the recognition result for the client after processing." (EX1038, 307.) "The server computer recognizes the speech and displays the recognized text back to the user." (EX1039, 2367.) Further, it would have been obvious that such data is received by the device from the ***server*** "in response to the sending of the captured ***first human voice data***" because the recognized speech, analysis of or information associated with the recognized

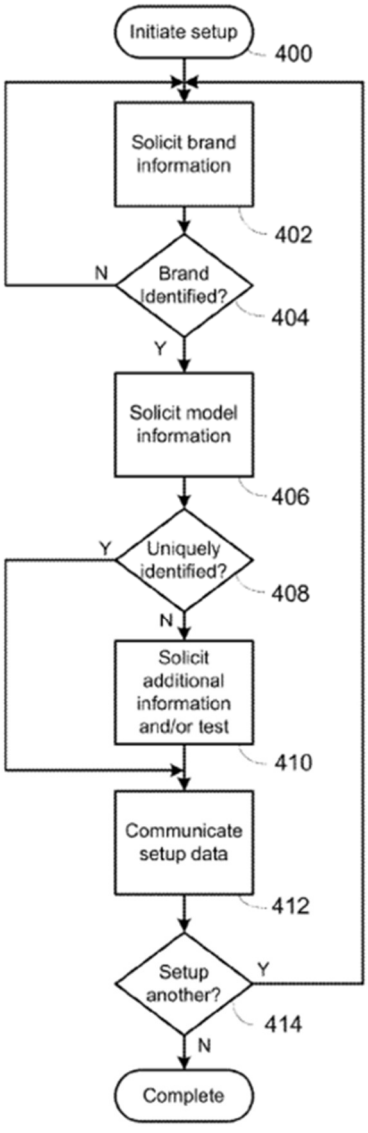
speech, the next prompt (*e.g.*, text or audio or information to display), and/or information triggering a visual or auditory output such as flashing a light indicator or playing a sound is generated based on the ***server***'s processing of ***first human voice data***.

412. Lim discloses that the device receives data from the server over the ***Internet via the WLAN***. As discussed above, Lim discloses a controlling device communicating with a ***server device*** (e.g., sending at least part of the captured ***first human voice data***) ***via the WLAN*** over the ***Internet***. (*Supra* §IX.H.3.h.) It thus follows that Lim's controlling device receives data from the server in response to the sending of the captured ***first human voice data*** over the ***Internet via the WLAN***. As discussed, Lim's device includes a ***WLAN transceiver*** which is the hardware component that receives wireless signals and is necessary for WLAN communication. (*Supra* § IX.H.3.c; EX1021, [0067]; EX1035, 3:17-20, 3:38-40; EX1036, [0013]-[0015].) Thus, it would have been obvious that any communication between the device and a server device (*e.g.*, receiving data by the device from the server) over the Internet via the WLAN (*e.g.*, WiFi) is by the ***WLAN transceiver***.

413. Lim discloses emitting, by the electric light source, in response to data received from the server over the Internet via the WLAN by the WLAN transceiver. Specifically, Lim's ***display*** and ***LEDs*** emit in the device to

provide a visual prompt. The prompt is provided in response to a user voice input and, more specifically, in response to data received from the speech recognition ***server 500*** (*i.e.*, speech processing result(s)), such as, recognized speech, analysis of or information associated with the recognized speech, next prompts, information to display, and/or information triggering a visual output such as flashing a light indicator. For example, "at step 402 the controlling device may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. This and subsequent prompts for voice input may take the form of an ***illuminated LED*** and/or '***speak' symbol***, a ***displayed text or image***, … etc." (EX1005, [0027].) Lim discloses that the prompts for voice input are based on the recognized speech. For example, Lim discloses that after step 402, "[a]t step 404, ***if the speech recognition means is unable to uniquely identify the desired brand***, step 402 may be repeated in order to ***solicit*** additional identifying information from the user. For example, a user may be ***presented*** with a short list of possible brand identities from which to select; may be ***prompted*** to spell the first few characters of the brand name; may be ***requested*** to speak the name of a geographic region or a service provider; etc.; as appropriate to facilitate the identification process." (Id., [0028].) In addition, "***[o]nce a brand has been successfully identified***, at step 406 the user may be ***prompted*** to provide

spoken appliance type and/or model number information for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, e.g., to further narrow down the codeset choices." (*Id.*, [0029].)



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

Figure 4

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

415. It would have been obvious that Lim emits its ***display*** and ***LEDs*** in the device in response to the data received from the remote speech recognition server (i.e., ***remote server 500***) (*i.e.*, speech processing result(s)), such as, recognized speech, analysis of or information associated with the recognized speech, the next prompt, information to display, and/or information triggering a visual output such as flashing a light indicator to display the prompt or alert the user.

k. 1[J]: "wherein the device is addressed in the WLAN using an Internet Protocol (IP) address that uniquely identifies the device in the WLAN or in the Internet."

416. Williams-Lim-Benesty discloses this element. Both Williams and Lim disclose or at least suggest this element. As discussed, Williams discloses that its device operates in a WLAN (e.g., home network formed by router 112), which is connected to the Internet. (*Supra* § IX.H.3.c.) A POSITA would have understood that this wireless network involves associating each device with a device identifier or address (such as an IP address) that uniquely identifies the device in the WLAN. (EX1007, 99, 108, 210, *see also*, *id*., 175.) 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. These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to communicate with devices outside of the user's WLAN. (EX1007, 340.)

417. Further, Lim also discloses this element. As discussed, Lim discloses that its device is WiFi-enabled and connects to the Internet via WLAN. (*Supra* §§ IX.H.3.c, IX.H.3.h) Thus, a POSITA would have

understood (and it would be obvious) that this wireless network similarly involves associating Lim's device with a device identifier or address (such as an IP address) that uniquely identifies the device in the WLAN. 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. (EX1007, 99, 108, 175, 210, 340.) These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to communicate with devices outside of the user's WLAN. *Id*.

418. Indeed, Lim expressly confirms this. Lim discloses with reference to FIG. 4, that "at step 404, the controlling device may request the user to "speak the name of a geographic region." (EX1005, [0028].) Lim discloses that "region information may be requested from a user if a spoken brand name is not immediately identified at step 404 of FIG. 4, and used to filter search results when multiple possible matches were initially identified." (EX1005, [0038].) Lim discloses that "in embodiments which are ***Internet server based*** region information may be derivable from the ***IP address associated with the*** requesting device[.]" *Id*. Thus, Lim expressly discloses that the requesting device (e.g., the smart device 110) is addressable

in the Internet using an IP address. It would have been obvious to incorporate Lim's networking feature into the combination to enable wireless connectivity.

4. Claim 4: "The device according to claim 1, further being addressable using the IP address in the WLAN or in the Internet using distinct locally administered addresses or a universally administered digital addresses."

419. Williams-Lim-Benesty discloses this claim. As discussed above, both Williams and Lim disclose or at least suggest that the device is addressed in the WLAN using an IP address that uniquely identifies the device in the WLAN or in the Internet. (*Supra* §§ IX.H.3.c, IX.H.3.k; EX1005, [0027], [0038], FIG. 5b; EX1006, [0012], [0020], [0023], FIG. 1; EX1007, 99, 108, 210, 175, 340.) It was well-known that networked devices are associated with IP addresses. The IP address may be an assigned IP address for the device (*e.g.*, by the router), the router's IP address, a combination of the two, or a globally unique IP address. (EX1007, 99, 108, 175, 210, 340.) These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to communicate with devices outside of the user's WLAN. *Id.* A POSITA would have

understood the router assigned IP addresses to be locally administered addresses. "Most home networks use IP addresses in the range of 192.168.1.100 to 192.168.1.149, and the ***router assigns each device an address*** from this pool when it boots." (EX1007, 108, 210, 175.)

420. When communicating with Internet servers using NAT, the device may utilize the router's globally unique IP address. "Network Address Translation (NAT)[:] The technology by which a home router translates the private IP addresses used by computers on your home network into a single ***public IP address*** assigned to your broadband modem, providing the ability for multiple computers to share a connection, and also providing a degree of privacy because the computers on your home network are not able to be accessed directly from the Internet." (EX1007, 340). A POSITA would have understood that these globally unique IP addresses are universally administered.

421. Additionally or alternatively, both Williams and Lim disclose or at least suggest that the device has a MAC address that is a universally administered digital address. It was common knowledge, and a POSITA would have understood that communicating via a wireless network and with a remote server over the Internet via a wireless router (as in Williams and Lim) involves associating the device with a device identifier or address such as a

MAC address that uniquely identifies the device in the network. (*Supra* §§ IX.H.3.c, IX.H.3.h; EX1006, [0012], [0020], [0023], FIG. 1; EX1005, [0007], [0023], [0027]-[0029], [0038], FIG. 5b.) "***Each computer or device that you want to join your wireless network needs a wireless network interface card (NIC)***," and "***each wireless card has a unique identifier called a MAC address***." (EX1007, 99, 166.) Each networked device is thus associated with a globally unique MAC address for identification of the device in the network.

422. This unique identifier is assigned to the NIC at its manufacture and is thus a universally administered address (locally stored in the non-volatile memory). "***[A] MAC address*** . . . is a 48-bit binary number generally written as 12 hexadecimal digits (six groups of two digits, the groups separated by dashes or colons). The address is ***set at the time of the NIC's manufacture*** . . . The second least significant bit of the most significant byte MAC address is known as the ***universal or locally administered address (U/L) bit*** . If this bit is set to 0, the address is administered universally. This means that the address was assigned by the Institute of Electrical and Electronics Engineers (IEEE) and is unique across the globe. If the bit is set to 1, it indicates that the address was selected locally by a network administrator (overriding the NIC's BIA)." (EX1023, 24.)

423. Furthermore, both the universally administered MAC address and the locally administered IP address are used in conjunction with each other. "On the DLC header, the data is functioning on Layer 2, so a MAC address is needed, as we saw. However, for the IP header, you have a set of IP addresses that should reinforce the fact you learned in Chapter 1—that IP functions on Layer 3 (the network layer) of the OSI model." (EX1023, 218, *see also id.,* 312 ("ARP allows IP-enabled devices on your network to dynamically map physical (MAC) addresses to IP addresses"), 313, Figure 6.6.) Thus, a POSITA would have understood (and found it obvious) that the device may comprise a MAC address that uniquely identifies the device globally – both in the WLAN and in the Internet, and that the MAC address may be administered universally, and that the device also is addressable using an IP address in the WLAN.

5. Claim 5: "The device according to claim 4, wherein the IP address is a layer 3 address and is static or dynamic IP address that is IPv4 or IPv6 type address."

424. Williams-Lim-Benesty discloses this claim. As discussed above, both Williams and Lim disclose that the device is addressable in the Internet and WLAN with an IP address. (*Supra* §§ IX.H.3.c, IX.H.3.k; EX1005, [0027], [0038], FIG. 5b; EX1006, [0012], [0020], [0023], FIG. 1; EX1007, 99, 108, 210, 175, 340.)

425. At the time of the invention, it was well-known that Internet Protocol (IP) addresses are either static or dynamic, are Layer 3 addresses and are IPv4 addresses. "In the case of today's Internet, these addresses are Internet protocol version 4 (IPv4) addresses." (EX1029, 127.) "Internet protocol (IP) addresses are the unique numbers used to identify individual connections to the Internet." *Id.* "Internet Protocol (IP) is a Layer 3 protocol that contains addressing and control information that allows packets to be routed." (EX1023, 11.) "The IP system in common use today is known as IPv4, for Internet Protocol version 4. A newer system, IPv6, or Internet Protocol version 6, has been developed and exists today in small deployments." (*Id*., 13.) "There are currently two main versions of IP and their associated IP address families; ***IPv4 and IPv6***. IP addresses can be ***either static or dynamic*** [.]" (EX1024, 2:60-62.) "The broadband networks typically assign a *dynamic IP address* to your broadband service. Dynamic means the IP address can change[.]" (EX1007, 277.) Further, an IP address may be a "*static IP address*, which means . . . your address never changes." *Id.*

6. Claim 6: "The device according to claim 1, further comprising a volatile or non-volatile memory that stores the IP address."

426. Williams-Lim-Benesty discloses this claim. For example, Williams discloses that "[t]he ***interface device*** may also include . . . operating logic 208. The operating logic 208 may be implemented as one or more ***programmable processing units, associated memory***, and appropriate instructions stored in the memory and executed by the progressing unit." (EX1006, [0023].)

427. As discussed above, Williams discloses that the interface device is addressable in the Internet and WLAN with an IP address and that it communicates with a remote server over the Internet via WLAN via a wireless router. (*Supra* §§ IX.H.3.c, IX.H.3.k; EX1006, [0012]–[0013], [0020], Figure 1.)

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

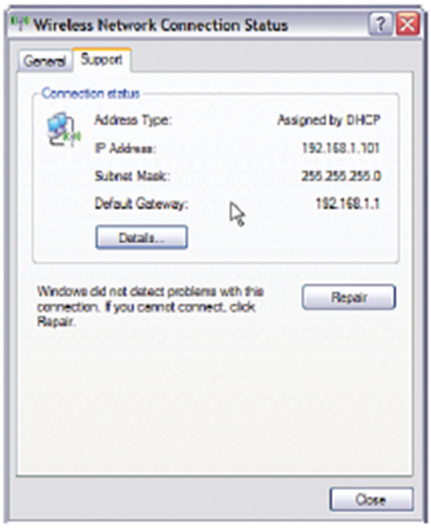
429. Additionally or alternatively, Lim discloses that the "controlling device 100 for use in commanding the functional operations of one or more appliances may include, as needed for a particular application, a processor 200 coupled to a ***ROM memory 204, a RAM memory 205***, . . . , a ***non-volatile***

read/write memory 206[.]" (EX1005, [0019].) The smart device 110, "illustrated in FIG. 3 in the form of a smart phone, may include . . . ***memory means 312 which may comprise ROM, RAM, Flash***, or any combination thereof[.] " (EX1005, [0023].)

430. As discussed above, Lim discloses that the controlling device is addressable in the Internet and WLAN with an IP address and that it communicates via WLAN and with a remote server over the Internet via a wireless router. (*Supra* §§ IX.H.3.c, IX.H.3.k; EX1005, [0027], [0038], FIG. 5b.)

431. A POSITA would have understood that the "ROM memory 204, RAM memory 205, non-volatile read/write memory 206, and memory means 312" as disclosed in Lim's device (e.g., smart device 110) would be used to store the associated IP addresses. For example, as discussed above and shown below, networked devices can display the IP addresses, which means the IP addresses are locally stored in the device. (*Supra* § IX.A.8; EX1007, 211, Figure 10-44, *see also, Id*., 265.)

Figure 10-44 Verifying IP Address Assignment with Windows XP



7. Claim 7: "The device according to claim 1, wherein the speaker comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

432. Williams-Lim-Benesty discloses this claim. As discussed, both Williams and Lim disclose a speaker for playing audio data (e.g., ***voice synthesizer with associated speaker, speaker 330, or speaker 206***), consistent with the descriptions of the '720 patent, in the device (e.g., interface device 116, smart device 110). (*Supra* §§ IX.H.3.d, IX.H.3.i; EX1005, [0019], [0023], FIGs. 2, 3; EX1006, [0016], [0023], [0037], FIG. 2; EX1001, 105:48-52, 105:61-106:4.) As discussed, a POSITA would have understood that an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker, a ribbon or planar magnetic loudspeaker are standard and common implementations of speakers. (*Supra* §IX.A.9.) Thus, a POSITA

would have understood (or it would be obvious) that Williams-Lim-Benesty's speaker is the recited speaker in claim 7.

8. Claim 8: "The device according to claim 1, wherein the speaker is operative for omnidirectional, unidirectional, or bidirectional pattern audio emitting."

433. Williams-Lim-Benesty discloses this claim. As discussed, both Williams and Lim disclose a standard speaker in the device (e.g., ***voice synthesizer with associated speaker, speaker 330, or speaker 206***). (*Supra* §§ IX.H.3.d, IX.H.3.i; EX1005, [0019], [0023], FIGs. 2, 3; EX1006, [0016], [0023], [0037], FIG. 2; EX1001, 105:48-52, 105:61-106:4.) As discussed, a POSITA would have understood that a standard speaker is operative for omnidirectional pattern audio emitting. (*Supra* §IX.A.10; EX1018, 122; EX1014, 763, 765.) Thus, a POSITA would understand (or it would be obvious) that Williams-Lim-Benesty's speaker is a speaker as recited in claim 8.

9. Claim 9: "The device according to claim 1, wherein the multiple microphones are configured to improve far-field audio reception."

434. Williams-Lim-Benesty discloses this claim. As discussed above, Benesty discloses this claim. (*Supra* §IX.E.2.)

435. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Benesty. (*Supra* § IX.H.2.)

10. Claim 10: "The device according to claim 9, wherein the multiple microphones are arranged as a directional microphones array operative to estimate a number, magnitude, frequency, Direction-Of-Arrival (DOA), distance, or speed of a phenomenon impinging the microphones array."

436. Williams-Lim-Benesty discloses this claim. As discussed above, Benesty discloses this claim. (*Supra* §§ IX.E.3.)

437. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Benesty. (*Supra* § IX.H.2.)

11. Claim 11: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, an omnidirectional, unidirectional, or bidirectional microphone that is based on sensing of an incident sound-based motion of a diaphragm or a ribbon."

438. Williams-Lim-Benesty discloses this claim. As discussed above, Williams-Lim-Benesty discloses multiple microphones in the device. (*Supra* §IX.H.3.) As discussed, a POSITA would have understood that a standard microphone is an omnidirectional microphone based on sensing an incident sound-based motion of a diaphragm. (EX1014, 757, 309.) Thus, a POSITA would understand (or it would be obvious) that each of Williams-Lim-Benesty's standard multiple microphones is the recited microphone in claim 11.

12. Claim 12: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

439. Williams-Lim-Benesty discloses this claim. As discussed above, Williams-Lim-Benesty discloses multiple microphones in the device. (*Supra* §IX.H.3.) As discussed, a POSITA would understand that dynamic microphone and piezoelectric microphone are standard and common implementations of microphones. (EX1014, 758.) Thus, a POSITA would understand (or it would be obvious) that each of Williams-Lim-Benesty's standard multiple microphones is the recited microphone in claim 12.

13. Claim 13: "The device according to claim 1, wherein the light is a visible light for illumination or indication."

440. Williams-Lim-Benesty discloses this claim. For example, Lim discloses this claim. As discussed, Lim's device includes ***LEDs*** and ***displays***. Lim discloses "a controlling device 100 for use in commanding the functional operations of one or more appliances may include, . . . , a means 202 to provide feedback to the user (e.g., ***one or more LEDs***, illuminable keys, ***display***, buzzer, voice synthesizer with associated speaker, and/or the like)[.]" (EX1005, [0019]; *Supra* §§IX.H.3.e, IX.H.3.j.) "A representative ***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an ***LCD display 306***; a keypad 308 which may comprise physical keys, touch keys overlaying LCD 306, or a combination thereof[.]" (EX1005, [0023].) As discussed, a display comprises an electric light source for

emitting light, for example, a LCD display may comprise LED backlighting which comprises "an array of LEDs arranged behind the screen." (EX1001, 109:45-50, 110:64-111:3, 111:11-23.)

441. A POSITA would have understood that LEDs are electric light sources that emit a visible light for illumination or indication. "Light-emitting diodes (LEDs) indicators are diodes that emit light when current passes in the forward direction." (EX1018, 174.) This is consistent with the '720 patent, which explains that "an actuator may be a light source that emits visible or non-visible light (infrared, ultraviolet, X-rays, or gamma rays) such as for illumination or indication . . . The light source may be an electric light source for converting electrical energy into light . . . The electric light source may be based on Solid-State Lighting (SSL) such as a ***Light Emitting Diode (LED)***[.]" (EX1001, 39:15-26.) "In one example, the device ***illuminates a visible light***, such as ***a Light-Emitting-Diode (LED)*** . . . Further, the illumination can be directed for lighting a surface, such as a surface including an image or a picture." (EX1001, 89:50-57.) It would have been obvious to a POSITA that Lim's LCD display and LEDs would be incorporated in the combination to enhance functionality and improve usability. (*Supra* §IX.H.2.)

14. Claim 14: "The device according to claim 1, wherein the light is a non-visible light that is infrared, ultraviolet, X-rays, or gamma rays."

442. Williams-Lim-Benesty discloses this claim. For example, Lim discloses this claim. As discussed, 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***[.]" (*Supra* §§IX.H.3.c, IX.H.3.k; EX1005, [0019].)

443. Further, Lim's "***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an ***LCD display 306***[.]" (EX1005, [0023].) As supported by the '720 patent, LCD screens may use LED backlighting comprising of LED-backlit lights such as in the form of "an array of LEDs arranged behind the screen". (EX1001, 110:64-111:3, 111:11-23.) As discussed, LCD displays and LEDs emit non-visible light. (*Supra* § IX.A.13.)

444. In addition, Lim discloses its "smart device . . . may include . . . IR communication means comprising a transmitter and/or receiver 336 with associated ***IR output diode(s) 334***[.]" (EX1005, [0023].) Indeed, Lim supports transmission of commands to appliance "using any convenient ***IR***" protocols. (EX1005, [0017].) Lim's IR output diode(s) 334 emit non-visible light that is infrared. It would have been obvious to a POSITA that

Lim's LEDs and other light outputting features would be incorporated in the combination to enhance functionality and improve usability.

15. Claim 15: "The device according to claim 1, wherein the electric light source consists of, or comprises, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode."

445. Williams-Lim-Benesty discloses this claim. For example, Lim discloses this claim. As discussed, 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***[.]" (*Supra* §§IX.H.3.c, IX.H.3.k; EX1005, [0019].)

446. Further, Lim's "***smart device***, illustrated in FIG. 3 in the form of a smart phone, may include . . . an ***LCD display 306***[.]" (EX1005, [0023].) As supported by the '720 patent, LCD screens may use LED backlighting comprising of LED-backlit lights such as in the form of "an array of LEDs arranged behind the screen". (EX1001, 110:64-111:3, 111:11-23.) Thus, Lim discloses electric light sources that comprise of at least a light emitting diode (LED). It would have been obvious to a POSITA that Lim's LCD display and LEDs would be incorporated in the combination to enhance functionality and improve usability. (*Supra* §IX.H.2.)

16. Claim 16: "The device according to claim 1, further comprising a memory that stores digital audio content, and the one or more

processors are further programmed with computer program instructions that, when executed, cause the device to playing by the speaker the stored digital audio content."

447. Williams-Lim-Benesty discloses this claim. For example, Williams discloses this claim. William discloses, with reference to FIG. 2, "API 214 of the application engine 118 [run on ***server 400***] may also allow the application 120 [run on ***server 400***] to provide ***audio 222*** for eventual transmission to the interface device 116, to be played on the ***speaker 206***." (EX1006, [0037].) "[T]he ***audio 222*** may be provided from the application 120 as a digital audio stream." *Id*.

448. Williams discloses that "[t]he interface device may also include a speaker 206, and operating logic 208. The operating logic 208 may be implemented as one or more programmable processing units, ***associated memory***, and ***appropriate instructions stored in the memory and executed by the processing unit***. Other types of logic may also be used to configure the interface device 116 to ***perform the functions described here***." (EX1006, [0023].) First, a POSITA would understand (and it would be obvious) that Williams' device stores the audio 222 in its "associated memory" upon receipt from the server to be played by the speaker. Second, Williams discloses that the operating logic 208 and other types of logic are used to "configure the device to perform the functions described", which necessarily includes

receiving audio data from the server "to be played on the speaker 206." (EX1006, [0023], [0037].) Thus, a POSITA would understand (and it would be obvious) that Williams' "programmable processing units" are programmed with instructions that, when executed, cause the device to play the stored digital audio content by the speaker.

449. Further, Lim's device plays a digital audio content that is synthesized human speech, which comprises words, phrases, and sentences. "This and subsequent prompts [by the device] for voice input may take the form of . . . a ***synthesized spoken message*** (*e.g.*, where controlling device user feedback means 202 comprises ***speech synthesis means*** and ***a speaker***, or where a smart device app has access to ***speaker 330***), etc." (EX1005, [0027].) A POSITA would understand (and it would be obvious) that Lim's device stores the synthesized spoken messages in its "memory means 312 [] compris[ing] ROM, RAM" and its "processing means 300" is programmed with instructions that plays the messages when executed. (EX1005, [0023]-[0024], [0019]-[0020].) It would have been obvious that, in Williams-Lim-Benesty, Williams' audio 222 would include synthesized spoken message as taught by Lim to provide speech output.

17. Claim 17: "The device according to claim 16, wherein the digital audio content is pre-recorded or synthesized audio content."

450. Williams-Lim-Benesty discloses this claim. For example, Williams discloses this claim. Williams teaches "provid[ing] audio 222 for eventual transmission to the interface device 116, to be played on the ***speaker 206***." (*Supra* §§ IX.H.3.d, IX.A.4.i; EX1006, [0037].) Further, Williams discloses that the server may perform speech synthesis to response to the user commands. "The audio may be in the form of text, in which ***speech synthesis or text-to-speech conversion*** may be ***performed by the application engine 118*** or by the operating logic 108 of the interface device 116." (EX1006, [0037], *see also id.*, [0045] and [0050].) As another example, Williams discloses that "the application may respond to a spoken command to order movie tickets by using the ***speech synthesis functionality*** of the system to ***query the user regarding movie selections and time***." (*Id.*, [0046].)

451. Thus, Williams discloses that the digital audio content (*e.g.*, 222) is synthesized audio content.

452. As another example, Lim discloses "a controlling device 100 for use in commanding the functional operations of one or more appliances may include, . . . , a means 202 to provide feedback to the user (e.g., one or more LEDs), illuminable keys, display, buzzer, ***voice synthesizer with associated speaker***, and/or the like[.]" (EX1005, [0019].) "A representative smart device, illustrated in FIG. 3 in the form of a smart phone, may include . . . an

LCD display 306; . . . an audio management system 326 with associated microphone 328, ***speaker 330***[.]" (EX1005, [0023]; *Supra* §IX.H.3.d.)

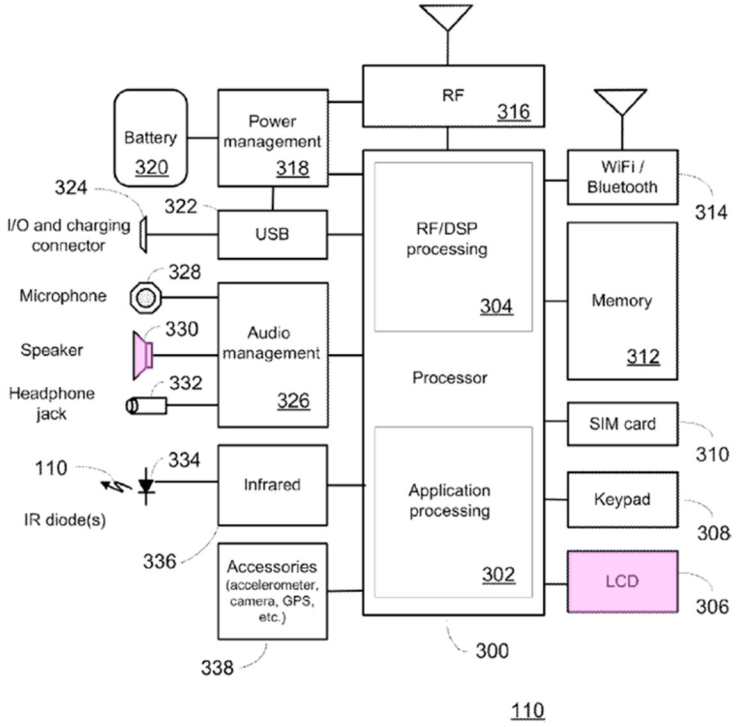


Figure 3

453. Lim discloses playing audio output by the speaker in the device. For example, "at step 402 the controlling device may prompt the user to speak the brand name of the appliance to be added to the configuration of the controlling device. This and subsequent prompts for voice input may take the form of an illuminated LED and/or 'speak' symbol, a displayed text or image, a ***synthesized spoken message*** (e.g., where controlling device ***user feedback means 202*** comprises speech synthesis means and ***a speaker***, or where a smart device app has access to ***speaker 330***), etc." (EX1005, [0027].) It would

have been obvious that, in Williams-Lim-Benesty, Williams' audio 222 would include synthesized spoken message as taught by Lim to provide speech output.

18. Claim 18: "The device according to claim 16, further operative to play a simulating of a voice of a human being or to play music."

454. Williams-Lim-Benesty discloses this claim. For example, Williams discloses this claim. Williams discloses that "the application may respond to a spoken command to order movie tickets by using the ***speech synthesis functionality*** of the system to ***query the user regarding movie selections and time***." (*Supra* §IX.A.18; EX1006, [0046].) Williams discloses several exemplary use cases of operating the speaker to play a simulating of a voice of a human being. For example, Williams discloses that "[a]n application may provide general information in response to users' uttered questions. A user may ask a question, such as 'What's the recipe for deviled eggs?' The application may respond by emailing or otherwise providing a recipe for deviled eggs. In some cases, the application may use the ***speech synthesis*** functionality of the system to provide ***audio answers to spoken questions***." (EX1006, [0048].) A POSITA would have understood that this involves simulating a voice of a human being.

455. As another example, Williams states that "such an application may allow the user to inquire about a flight or other event. The system might respond by sending information to the user via a device associated with the user, or by ***audibly responding*** using ***speech synthesis***." (EX1006, [0050].)

456. Additionally or alternatively, Lim discloses this claim. Lim discloses "prompt[ing] the user to speak the brand name of the appliance to be added to the configuration of the controlling device. This and subsequent prompts for voice input may take the form of an illuminated LED and/or 'speak' symbol, a displayed text or image, a ***synthesized spoken message*** (e.g., where controlling device ***user feedback means 202*** comprises speech synthesis means and ***a speaker***, or where a smart device app has access to ***speaker 330***), etc." (EX1005, [0027].)

457. Lim discloses several examples of such prompts. For example, Lim discloses that after step 402, "[a]t step 404, ***if the speech recognition means is unable to uniquely identify the desired brand*** [based on ***speech processing result(s)***], step 402 may be repeated in order to ***solicit*** additional identifying information from the user. For example, a user may be ***presented*** with a short list of possible brand identities from which to select; may be ***prompted*** to spell the first few characters of the brand name; may be ***requested*** to speak the name of a geographic region or a service provider; etc.;

as appropriate to facilitate the identification process." (EX1005, [0028].) In addition, "***[o]nce a brand has been successfully identified*** [based on ***speech processing result(s)***], at step 406 the user may be ***prompted*** to provide spoken appliance type and/or model number information for processing and comparison to a phonetic database of model number data which is applicable to the identified brand if necessary, e.g., to further narrow down the codeset choices." (*Id*., [0029].) A POSITA would have understood that these voice prompts are digital audio data played by the speaker in the controlling device.

458. Thus, Lim discloses that the device is operative to play a simulating of a voice of a human being (*i.e*., synthesized spoken message). It would have been obvious that, in Williams-Lim-Benesty, Williams' audio 222 would include synthesized spoken message as taught by Lim to provide speech output.

19. Claim 19: "The device according to claim 18, wherein the digital audio content comprises a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

459. Williams-Lim-Benesty discloses this claim. For example, Williams discloses this claim. (*Supra* §IX.A.19.) Additionally or alternatively, as discussed, Lim discloses providing audio prompts for voice input in the form of "a ***synthesized spoken message*** (e.g., where controlling device ***user feedback means 202*** comprises speech synthesis means and ***a***

speaker, or where a smart device app has access to ***speaker 330***), etc." (EX1005, [0027].)

460. Further, Lim discloses examples of such prompts for voice input. "For example, a user . . . may be ***prompted*** to spell the first few characters of the brand name; may be ***requested*** to speak the name of a geographic region or a service provider; etc.; as appropriate to facilitate the identification process." (EX1005, [0028].)

461. Thus, Lim discloses that the digital audio content (*i.e.*, synthesized spoken message) comprises a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice. It would have been obvious that, in Williams-Lim-Benesty, Williams' audio 222 would include synthesized spoken message as taught by Lim to provide speech output.

20. Claim 20: "The device according to claim 1, wherein the WLAN is according to, or compatible with, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard."

462. Williams-Lim-Benesty discloses this claim. Both Williams and Lim disclose this claim. For example, as discussed, Williams discloses that the interface device 116 communicates over a WLAN (e.g., via a router). (*Supra* §§ IX.H.3.c, IX.H.3.h, IX.A.4.i; EX1006, [0012]–[0013], [0020],

Figure 1.) As discussed, WiFi implements WLAN and is based on the IEEE802.11 standards. (EX1040, 600-601; EX1013, 7:48-51; EX1001, 8:25-32.) Williams expressly discloses usage of "other wireless data communications technologies" between interface device and application engine and "various network communications between the components of FIG. 2 (which includes the client device, the controlled device, and the server device) may be performed using existing networking protocols and techniques." (EX1006, [0020], [0012], [0013], [0038].) A POSITA would understand that WLAN over IEEE 802.11a/b/g are common wireless technologies in or before 2012. (EX1042, 600-601; EX1013, 7:48-51; EX1001, 8:25-32.) Thus, a POSITA would understand that Williams discloses or suggests the WLAN is according to, or base on, Institute of Electrical and Electronics Engineers (IEEE) 802.11a, IEEE 802.11b, IEEE 802.11g.

463. Additionally or alternatively, Lim discloses this claim. Lim discloses that the controlling device communicates with the ***remote server 500*** (*i.e.*, provisioning the speech recognition result and/or additional prompts for voice input and/or indicated command code data to the first device) over the ***Internet*** (*i.e.*, network 502), and ***via WiFi***, a form of WLAN, through a gateway device such as a ***router***. (*Supra* §§ IX.H.3.c, IX.H.3.h; EX1005,

[0027], Figure 5b; EX1040, 600-601; EX1013, 7:48-51; EX1001, 8:25-32.) Lim discloses that access to "a remote server 500 accessible via a network 502 such as the Internet" "may be via an intermediary appliance such as STB 104, illustrated at 506, or may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].) Thus, a POSITA would understand that Lim discloses or suggests the WLAN is according to, or base on, Institute of Electrical and Electronics Engineers (IEEE) 802.11a, IEEE 802.11b, IEEE 802.11g. It would have been obvious to incorporate Lim's networking features into the combination to improve networking capabilities.

21. Claim 21: "The device according to claim 1, wherein the WLAN uses a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

464. Williams-Lim-Benesty discloses this claim. For example, both Williams and Lim disclose or suggest WLAN that is according to, or based on, Institute of Electrical and Electronics Engineers (IEEE) 802.11b and a POSITA would understand that WLAN over 802.11b is a form of ISM Band II frequency. (*Supra* §§IX.H.20, IX.A.21; EX1042, 22-23, Table 1.2.) It would have been obvious to incorporate Lim's networking features into the combination to improve networking capabilities.

22. Claim 22: "The device according to claim 1, wherein the enclosure is a handheld enclosure."

465. Williams-Lim-Benesty discloses this claim. For example, Lim discloses that its controlling device can be a smart device 110 in the form of a smart phone. (*Supra* § IX.H.3.a; EX1005, [0018], [0023].) "A representative smart device [is] illustrated in FIG. 3 in the form of a smart phone." (EX1005, [0023].) A POSITA would understand that a smart phone enclosure is a handheld device. Further, Lim discloses that the ***smart device 110*** "may comprise any of a ***smart phone, a tablet computer***, a portable email device, an e-book, a ***hand-held games and/or game controller***, a portable media player." (EX1005, [0018].) A POSITA would find it obvious that the smart device may take the form of other portable devices such as that of other handheld devices.

466. It would have been obvious to make Williams' device handheld, like Lim, such that the device is near the user and "user speech can be clearly received." (EX1006, [0019].) It would have been obvious that the Williams-Lim-Benesty device can be implemented in the form of Lim's smart phone to improve portability and usability.

23. Claim 23: "The device according to claim 1, further comprising in the enclosure an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, the actuator that is

coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN."

467. Williams-Lim-Benesty discloses claim 23.

468. Benesty discloses "one-or two dimensional loudspeaker array" and "stereophonic sound reproduction" which uses multiple speakers. (EX1010, 1100, 1103.) Benesty discloses that "[s]tereophony comprises a number of techniques for reproduction with ***two or more loudspeakers***. They describe how to feed the loudspeakers with appropriate signals to create suitable ***spatial cues*** at the ears of a listener." (EX1010, 1100.)

469. Benesty also discloses vector-based amplitude panning where "that panning is not only applied to two loudspeakers, but to two or three adjacent loudspeakers out of an arbitrary number distributed on a circle or a sphere around the listener. In this way, virtual sources can appear on a full horizontal circle or anywhere in space." (EX1010, 1103.) "Vector-based amplitude panning extends the tangent panning law for two loudspeakers to panning between ***adjacent speakers of a one- or two-dimensional loudspeaker array***." (EX1010, 1103.)

470. Further, as discussed below, a speaker system comprises a digital-to-analog converter ("DAC"), which is an actuator. (*Infra* §IX.H.26.) The speaker system is comprised in the enclosure.

471. It would have been obvious to a POSITA to combine the teachings of Williams-Lim and Benesty. As discussed, Williams-Lim and Benesty are from the same field of endeavor, and all relate to the same technical problems of audio collection and feedback. (*Supra* § IX.H.2.)

472. It would have been obvious to a POSITA to apply Benesty's teachings regarding conventional audio components and techniques to improve audio collection and playback to Williams-Lim's device, as various features of Benesty can improve Williams-Lim's device in the same way and yield predictable results. A POSITA would have been motivated to include a loudspeaker array (*e.g.*, additional speakers) in Williams-Lim's device as a conventional, known, and well-established design choice in a consumer device to improve audio playback quality. Incorporating an additional speaker predictably improves usability of the device by enabling clearer and louder playback, providing better audibility in noisy environments, and allowing stereo playback or spatial audio effects. Further, adding an additional speaker would increase maximum output volume while reducing distortion by distributing acoustic load across multiple transducers. Multiple speakers also predictably improve sound dispersion and coverage, allowing users located at different areas in a room to more reliably hear audio output without requiring excessive volume. These known benefits of incorporating

an additional speaker into a consumer device would prompt the application of Benesty in the same field based on design incentives and other market forces favoring improved multimedia/voice performance with clearer, louder, and more uniform audio output on consumer audio devices.

473. A POSITA would have been motivated to couple the additional speaker in Williams-Lim's device with the WLAN transceiver and operate it to play audio data in the same way.

474. Further, it would have been obvious to try combining the technical features of Benesty with the teachings of Williams-Lim, for example, by including an additional speaker in the device as taught by Benesty and implementing the same WLAN-driven speaker actuation, without requiring undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, using routine networked device architectures, as demonstrated by successful implementations in the field. There was a finite number of predictable ways to improve audio output, such as using a loudspeaker array or adopting active noise cancelling, that were used in the industry and could be integrated into the Williams-Lim system. Prior art such as Benesty demonstrated successful implementation of multiple speakers in a consumer device. The known benefits and widespread adoption of multi-speaker

designs and WLAN-based voice/multimedia playback would therefore have motivated a POSITA to incorporate these conventional features into the Williams-Lim system with a reasonable expectation of success, yielding predictable improvements in usability, hands-free operation, and improved audible playback based on WLAN-received data.

24. Claim 24: "The device according to claim 23, wherein the actuator consists of, or comprises, a motion actuator that causes linear or rotary motion."

475. As discussed, Williams-Lim-Benesty's device comprises additional speaker(s). (*Supra* § IX.H.23.)

476. As discussed, a POSITA would have understood that a speaker is a motion actuator that causes linear motion. (*Supra* § IX.A.24; EX1014, 763, 903, 18.)

477. It would have been obvious to a POSITA to combine Williams, Lim, and Benesty. (*Supra* §§ IX.H.2, IX.H.23.)

25. Claim 25: "The device according to claim 23, wherein the actuator consists of, or comprises, an electromagnetic coil or an electromagnet operative for generating a magnetic or electric field."

478. As discussed, Williams-Lim-Benesty's device comprises additionally, a speaker (*e.g.*, an actuator). (*Supra* § IX.H.23.)

479. As discussed, a POSITA would have understood that a speaker commonly comprises an electromagnetic coil operative for generating a

magnetic or electric field, consistent with the specification of the '720 patent. (*Supra* § IX.A.25; EX1001, 105:64-106:7; EX1018, 122.)

480. It would have been obvious to a POSITA to combine Williams, Lim, and Benesty. (*Supra* §§ IX.H.2, IX.H.23.)

26. Claim 26: "The device according to claim 23, wherein the actuator consists of, or comprises, an electrical signal generator."

481. Williams-Lim-Benesty discloses or at least suggests its device comprises a digital-to-analog converter, which is an actuator comprising an electrical signal generator. As discussed, it was well-known that a speaker system comprises a DAC, an electrical signal generator actuator, for converting received digital audio into analog signals for playback. (*Supra* § IX.A.26; EX1014, 714, 522; EX1018, 170; EX1001, 34:25-28, 107:13-20.) Further, it would have been obvious to a POSITA that Williams-Lim-Benesty's device would include a digital-to-analog converter for converting the received digital audio signal back into an analog signal for playback (*e.g.*, of audio 222). Lim already supports analog-to-digital conversion and Williams discloses transmission and receipt of digital audio data. Lim's "controlling device 100 . . . may include, as needed for a particular application, speech input means comprising a microphone 224 and associated circuitry 222 for example an analog to digital converter, speech digitizer, etc.,

as known in the art." (EX1005, [0019].) Williams discloses that the audio received by the device may be in the form of a digital audio stream. "The API 214 of the application engine 118 may also allow the application 120 to provide audio 222 for eventual transmission to the interface device 116, to be played on the speaker 206. . . . [t]he audio 222 may be provided from the application 120 as a digital audio stream." (EX1006, [0037].) It would have been obvious to a POSITA that the combination would include a DAC.

27. Claim 29: "The device according to claim 23, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

482. Williams-Lim-Benesty discloses claim 29.

483. As discussed, Williams-Lim-Benesty's device comprises additionally, a speaker. (*Supra* § IX.H.23.) A speaker is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, audio or sound. (*Supra* § IX.A.26.) Further, a DAC is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, electrical current or electrical voltage. *Id*.

484. It would have been obvious to a POSITA to combine Williams, Lim, and Benesty. (*Supra* §§ IX.H.2, IX.H.23.)

I. Ground 9: Williams, Lim, Benesty, and iPhone User Guide Render Obvious Claim 2

1. Claim 2: "The device according to claim 1, further comprising, in the enclosure, a rechargeable battery for powering the device."

485. Claim 1 is obvious in view of Williams-Lim-Benesty. (*Supra* §IX.H.3.) Claim 2 is obvious in further view of iPhone User Guide.

486. iPhone User Guide discloses claim 2. (*Supra* §IX.B.3.)

487. Further, Lim discloses a mobile phone with internal battery. "A representative smart device, illustrative in FIG. 3 in the form of a smart phone, may include . . . power management circuitry 318 with associated ***battery 320***[.]" (EX1005, [0023].) Lim discloses that the smart device "may comprise any of a smart phone, a tablet computer, a portable email device, an e-book, a hand-held games and/or game controller, a portable media player." (EX1005, [0018].) A POSITA would understand and find it obvious that a smart phone, tablet computer, or any other forms of portable devices comprise a rechargeable battery for powering the device. Indeed, in 2012, handheld mobile phones are commonly equipped with a rechargeable battery. (EX1008, 34.)

488. It would have been obvious to a POSITA to combine the teachings of Williams-Lim-Benesty and iPhone User Guide. Both iPhone User Guide and Williams-Lim-Benesty relate to audio capturing. For example, Williams discloses that the "in-home interface device [] compris[es]:

a ***microphone*** configured to receive ***audio from a nearby user***[.]" (EX1006, Claim 21; *see also id.*, claims 1-3.) Lim discloses that "a controlling device may be provisioned with means to receive voice input from a user, e.g., a ***microphone***, such that when placed in a setup mode the controlling device may solicit spoken appliance descriptive information, for example brand and model number, to be processed by a speech recognition system in order to identify the appliance to be controlled and the command data to which it is responsive." (EX1005, [0007].) Benesty discloses hands-free audio communication, audio collection via microphones, and speech processing techniques. (EX1010, 1021.) iPhone User Guide is a product manual of a device (*i.e.*, an iphone) comprising of multiple microphones to receive user audio. (EX1008, 9, 29.) Thus, Williams-Lim-Benesty and iPhone User Guide are from the same field of endeavor, and all relate to the same technical problems of audio collection. Thus, a POSITA would have been motivated to look to iPhone User Guide for techniques and details related to components of the device for implementing Williams-Lim-Benesty.

489. It would have been obvious to a POSITA to apply iPhone User Guide's teachings related to device components to Williams-Lim-Benesty's interface device, as various features of iPhone User Guide can improve Williams-Lim-Benesty's device in the same way and yield predictable results.

For example, the combination would result in a device with a rechargeable battery as taught by iPhone User Guide. The known benefits of incorporating a rechargeable battery into a consumer device would prompt the application of iPhone User Guide in the same field based on design incentives, other market forces, and user expectations for consumer devices, particularly where the device performs power consuming operations such as wireless communications, audio capture, audio output. In such electronic devices, a rechargeable battery predictably improves usability and convenience by allowing the user to recharge the device, rather than replacing disposable cells. The benefits are predictable and directly aligned with the practical requirements of a portable device intended for frequent, repeated use.

490. Further, it would have been obvious to try combining the technical features of iPhone User Guide with the teachings of Williams-Lim-Benesty, for example, by using a rechargeable battery to power up the device, without requiring undue effort and with a reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Further, there were a finite number of ways to power up a device, such as via a rechargeable battery, a disposable battery, or external power supply, that were commonly seen in the industry and easy to manufacture and

integrate into the Williams-Lim-Benesty system. It would have been obvious to try, and a POSITA would have reasonably expected a rechargeable battery to work successfully as a straightforward substitution for disposable cells or as an alternative to external power supply. The selection of a rechargeable battery would therefore have been a routine engineering choice involving known components performing their expected function, yielding predictable results of improved portability, convenience, and reliable operation.

J. Ground 10: Williams, Lim, Benesty, and Doherty Render Obvious Claims 4-5

1. Claim 4: "The device according to claim 1, further being addressable using the IP address in the WLAN or in the Internet using distinct locally administered addresses or a universally administered digital addresses."

491. Claim 1 is obvious in view of Williams, Lim, and Benesty. (*Supra* § IX.H.3.) Claims 4-5 are obvious in further view of Doherty.

492. Doherty discloses claim 4. (*Supra* § IX.C.2.)

493. As discussed above, both Williams and Lim disclose that the device is addressable using the IP address in the WLAN. (*Supra* §IX.H.3.k.) 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. (EX1007, 99, 108, 175, 210, 340.) These IP addresses uniquely

identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to communicate with devices outside of the user's WLAN. *Id*.

494. Additionally or alternatively, each of Williams and Lim's device has a NIC that has a unique MAC address. (*Supra* §IX.C.2.) As discussed, Lim's device is WiFi-enabled and communicates via WLAN and with a remote server over the Internet via a wireless router. (*Supra* §§ IX.H.3.c, IX.H.3.h; EX1005, [0027], [0038], FIG. 5b; EX1006, [0012], [0020], [0023], FIG. 1; EX1007, 99, 108, 210, 175.) In light of Doherty, a POSITA would have understood (and find it obvious) that each of Williams and Lim's device is associated with a globally unique MAC address that uniquely identifies the device in the WLAN or in the Internet. (EX1007, 99, 166; EX1023, 24.) This unique identifier is assigned to the NIC at its manufacture and is thus a universally administered address (locally stored in the non-volatile memory). (EX1023, 24.) Both the universally administered MAC address and the locally administered IP address are used in conjunction with each other when identifying the device. (EX1023, 218, *see also id.,* 312, 313, Figure 6.6.)

495. It would have been obvious for a POSITA to apply Doherty's teachings of conventional wireless networking components and addressing techniques to Williams-Lim-Benesty. Williams, Lim, Benesty, and Doherty are related to wireless communication technologies. Williams discloses that the interface device 116 communicates with a wireless network (e.g., via a router). (*Supra* §§ IX.H.3.c, IX.H.3.h, IX.A.4.i; EX1006, [0012] – [0013], [0020], Figure 1.) Lim discloses that access to "a remote server 500 accessible via a network 502 such as the Internet" "may be via an intermediary appliance such as STB 104, illustrated at 506, or may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].) Benesty discloses speech data transmission over a wireless (WiFi) network. (EX1010, 309.) Doherty, a textbook directed to "wireless home networking," explains that "[e]ach computer or device that you want to join your wireless network needs a wireless network interface card (NIC)," and "each wireless card has a unique identifier called a MAC address." (EX1007, 99, 166.) Thus, Williams, Lim, Benesty, and Doherty are from the same field of endeavor, and all relate to the same technical problem of providing wireless communications between devices.

496. It would have been obvious for a POSITA to look to Doherty for well-known implementation details of conventional wireless networking to apply to Williams-Lim-Benesty. Doherty teaches that each networked device includes a NIC and that the NIC includes a unique MAC address which is a universally administered address, used for device identification on a local network. Incorporating a NIC with a unique MAC address into Williams-Lim-Benesty's device would have been a routine and predictable design choice, because MAC addressing is a fundamental and ubiquitous feature of IEEE 802 networking and is required for ordinary device participation in standard WLANs. Indeed, a POSITA would understand that for Williams-Lim-Benesty's device to communicate with an Internet server over Wi-Fi, the device necessarily employs conventional addressing techniques and thus would include (or be implemented with) a NIC having a unique MAC address in the same way as taught by Doherty.

497. A POSITA also would have been motivated by strong market incentives and product design requirements to include a NIC with a unique MAC address to ensure compatibility with the vast installed base of existing home routers, access points, and network infrastructure. Wireless consumer devices are expected to interoperate with standard Wi-Fi networks and MAC addressing is a universally adopted mechanism for network identification,

association, filtering, and access control in such environments. Incorporating a unique MAC address would therefore predictably improve interoperability, simplify deployment, and allow the Williams-Lim-Benesty device to function seamlessly with existing networks. These design incentives and market forces would have prompted a POSITA to apply Doherty's conventional NIC/MAC teachings to Williams-Lim-Benesty to meet consumer expectations for wireless connectivity and broad network compatibility.

498. Further, there would have been a reasonable expectation of success in combining Doherty with Williams-Lim-Benesty because the use of a NIC with a unique MAC address is a well-established and widely standardized networking feature implemented in essentially all WLAN-capable devices. The combination is straightforward and readily achievable by a POSITA using routine engineering practices, and it merely applies a known, fundamental networking mechanism (unique MAC addressing) to the wireless communications already contemplated by Williams-Lim-Benesty.

2. Claim 5: "The device according to claim 4, wherein the IP address is a layer 3 address and is static or dynamic IP address that is IPv4 or IPv6 type address."

499. Doherty discloses claim 5. (*Supra* § IX.C.3.) Doherty discloses, and a POSITA would understand that Internet Protocol (IP) addresses are either static or dynamic, are Layer 3 addresses and are IPv4

addresses. (EX1007, 277; EX1029, 127; EX1023, 11, 13; EX1024, 2:60-62.)

500. It would have been obvious to combine Williams-Lim-Benesty and Doherty. (*Supra* § IX.J.1.)

K. Ground 11: Williams, Lim, Benesty, and Rossing Render Obvious Claims 7-8, 11-12

1. Claim 7: "The device according to claim 1, wherein the speaker comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

501. Claim 1 is obvious in view of Williams, Lim, and Benesty. (*Supra* § IX.H.3.) Claims 7-8, 11-12 are obvious in further view of Rossing.

502. Rossing discloses claim 7. (*Supra* § IX.D.2.)

503. A POSITA would understand that this claim refers to standard types of speakers. As discussed, Rossing discloses that there are different types of speaker at the time of the invention, such as an electromagnetic speaker, a piezoelectric speaker, an electrostatic loudspeaker, or a magnetic loudspeaker. (EX1014, 763, 321-322.)

504. Williams, Lim, Benesty, and Rossing are from the same field of endeavor and reasonably pertinent to the same technical problems of operating microphone(s) and speaker(s). Both Williams and Lim disclose a speaker in for playing audio data (e.g., ***voice synthesizer with associated speaker,***

speaker 330, or speaker 206), consistent with the descriptions of the '720 patent, in the device (e.g., interface device 116, smart device 110). (*Supra* §§ IX.H.3.d, IX.H.3.i; EX1005, [0019], [0023], FIGs. 2, 3; EX1006, [0016], [0023], [0037], FIG. 2; EX1001, 105:48-52, 105:61-106:4.) Benesty discloses methods for speech processing, audio collection, and acoustic feedback. (EX1010, 1021, 980.) Rossing is a textbook providing basic information on how microphones and speakers operate. (EX1014, 757-758, 763, 765, 903.) Thus, a POSITA would have been motivated to look to Rossing for techniques and details related to speakers for implementing Williams-Lim-Benesty, as various speaker features can improve William-Lim-Benesty's speaker in the same way and yield predictable results. The known benefits of various features would prompt the application of Rossing in the same field based on design incentives and other market forces.

505. Further, it would have been obvious to try combining the technical features of Rossing with the teachings of Williams-Lim-Benesty, for example, by using an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker in the client device to output audio reply, without requiring undue effort and with a reasonable expectation of success. (*Supra* § IX.D.2.) Both Williams and Lim disclose standard speakers and thus

suggests a motivation to combine. (EX1005, [0019], [0023], FIGs. 2, 3; EX1006, [0016], [0023], [0037], FIG. 2.) There was a finite number of types of speakers, such as an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, that were commonly seen in the industry and easy to manufacture and integrate into the Williams-Lim-Benesty system. (EX1014, 321-322, 763.) The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field.

2. Claim 8: "The device according to claim 1, wherein the speaker is operative for omnidirectional, unidirectional, or bidirectional pattern audio emitting."

506. Rossing discloses claim 8. (*Supra* § IX.D.3.)

507. A POSITA would have understood that this claim refers to standard properties of a speaker.

508. As discussed above, both Williams and Lim disclose a speaker in the device for playing audio data (e.g., ***voice synthesizer with associated speaker, speaker 330, or speaker 206***). (*Supra* §§ IX.H.3.d, IX.H.3.i; EX1005, [0019], [0023], FIGs. 2, 3; EX1006, [0016], [0023], [0037], FIG. 2.)

509. Further, in light of Rossing, a POSITA would have understood that the speaker operates by emitting audible sound waves that are of an

omnidirectional, unidirectional, or bidirectional pattern. For example, "***[s]ound radiates from a loudspeaker in all directions***, depending on angle and frequency … It is rare for a speaker to have truly constant directionality across its entire pass-band, in part from that fact that most are at least somewhat directional at mid and high frequencies, and, because of the long wavelengths involved, almost unavoidably ***omni-directional at low frequencies***." (EX1014, 765.) It would have been obvious to combine Williams-Lim-Benesty and Rossing. (*Supra* § IX.K.1.)

3. Claim 11: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, an omnidirectional, unidirectional, or bidirectional microphone that is based on sensing of an incident sound-based motion of a diaphragm or a ribbon."

510. Rossing discloses claim 11. (*Supra* § IX.D.4.)

511. A POSITA would understand that the claim refers to properties and components of a standard microphone.

512. As discussed, both Williams and Lim disclose microphone(s) in the device. For example, Lim discloses that the controlling device has a ***microphone***. For example, "a controlling device 100 for use in commanding the functional operations of one or more appliances may include . . . speech input means comprising ***a microphone 224*** and associated circuitry 222 for example an analog to digital converter, speech digitizer, etc., as known in the

art." (EX1005, [0019].) "A representative smart device, illustrated in FIG. 3 in the form of a smart phone, may include . . . an audio management system 326 with associated ***microphone 328***, speaker 330, and headphone jack 332 [.]" (EX1005, [0023].) Williams discloses "[a]n audio interface device 116 … having a ***microphone*** and a network interface." (EX1006, [0016]; *see also*, *id.*, [0007], [0019], claims 1-3 and 21.)

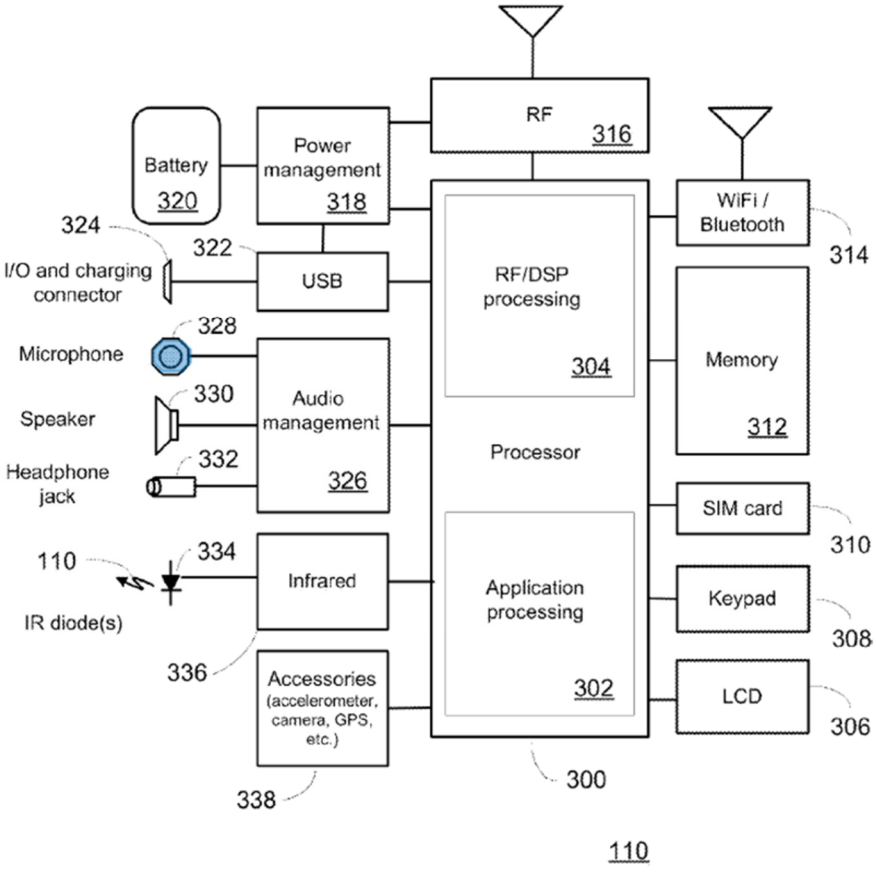


Figure 3

513. As discussed, Williams-Lim-Benesty and Rossing are from the same field of endeavor, and all relate to the same technical problems of

operating microphone(s). (*Supra* § IX.K.1; EX1006, [0016]; *see also*, *id.*, [0007], [0019], claims 1-3 and 21; EX1005, [0019], [0023]; EX1010, 1021; EX1014, 757-758, 763, 765, 903.) Thus, a POSITA would have been motivated to look to Rossing for techniques and details regarding these components for implementing the system described in Williams-Lim-Benesty.

514. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional microphone components to Williams-Lim-Benesty, which expressly disclose microphones, as various microphone features can improve Williams-Lim-Benesty's microphones in the same way and yield predictable results. Both Williams and Lim disclose standard microphones and thus suggests a motivation to combine. (EX1006, [0016]; *see also*, *id.*, [0007], [0019], claims 1-3 and 21; EX1005, [0019], [0023].) The known benefits of adopting various standard microphone components (for example, due to availability and popularity) would prompt the application of Rossing in the same field based on design incentives and market forces. Further, it would have been obvious to try combining Rossing with the teachings of Williams-Lim-Benesty, for example by using an omnidirectional, unidirectional, or bidirectional microphone that is based on the sensing an incident sound-based motion of a diaphragm or a ribbon in the device to capture the first human data, without requiring undue effort and with a

reasonable expectation of success. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. There was a finite number of types of microphones, such as an omnidirectional, unidirectional, or bidirectional microphone, that were commonly known in the industry and easy to manufacture and integrate into Williams-Lim-Benesty.

4. Claim 12: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

515. Rossing discloses claim 12. (*Supra* § IX.D.5.)

516. A POSITA would understand that the claim refers to standard types of microphones.

517. As discussed, Rossing discloses that there are different types of microphones at the time of the invention, such as a condenser, an electret, a ribbon, a carbon, and a piezoelectric microphone. (EX1014, 757-759.)

518. As discussed, Williams, Lim, Benesty, and Rossing are from the same field of endeavor, and all relate to the same technical problems of operating microphone(s). (*Supra* § IX.K.1.) A POSITA would have been motivated to look to Rossing for techniques and details regarding these components for implementing the system described in Williams-Lim-Benesty.

519. It would have been obvious to a POSITA to apply Rossing's teachings related to conventional components to Williams-Lim-Benesty, for example, as various microphone features can improve Williams-Lim-Benesty' microphone in the same way and yield predictable results. The known benefits of adopting various standard microphone components (for example, due to popularity and practical benefits of "good sample-to-sample consistency" which supports usage with microphone arrays) would prompt the application of Rossing in the same field based on design incentives and market forces. Further, it would have been obvious to try combining Rossing with the teachings of Williams-Lim-Benesty, for example by using a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone, without requiring undue effort and with a reasonable expectation of success. There was a finite number of types of microphones, such as a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone, that were commonly known in the industry and easy to manufacture and integrate into Williams-Lim-Benesty. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field.

L. Ground 12: Williams, Lim, Benesty, and Goldsmith Render Obvious Claims 20-21

1. Claim 20: "The device according to claim 1, wherein the WLAN is according to, or compatible with, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard."

520. Claim 1 is obvious in view of Williams, Lim, and Benesty. (*Supra* § IX.H.3.) Claims 20-21 are obvious in further view of Goldsmith.

521. As discussed, Lim discloses or at least suggests this claim. (*Supra* §§ IX.H.3.c, IX.H.20; EX1005, [0027]; EX1042, 600-601; EX1013, 7:48-51; EX1001, 8:25-32.)

522. Further, Goldsmith discloses claim 20. (*Supra* § IX.F.2.)

523. It would have been obvious for a POSITA to apply Goldsmith's disclosure of various types of wireless networking techniques and protocols to Williams-Lim-Benesty. Both Williams-Lim-Benesty and Goldsmith are related to wireless communication technologies. As discussed, Williams discloses that the interface device 116 and home devices (media devices 106 and application devices 220) communicate with a wireless network (e.g., via a router). (*Supra* §§ IX.H.3.c, IX.H.3.h, IX.A.4.i; EX1006, [0012] – [0013], [0020], Figure 1.) Williams expressly discloses usage of "other wireless data communications technologies" between interface device and application engine and "various network communications between the components of FIG. 2 (which includes the client device, the controlled device, and the server

device) may be performed using existing networking protocols and techniques." (EX1006, [0020], [0012], [0013], [0038].) Lim discloses that access to "a remote server 500 accessible via a network 502 such as the Internet" "may be via an intermediary appliance such as STB 104, illustrated at 506, or may be via a ***local network connection 508, such as WiFi or Bluetooth, through a gateway device such as a router*** (not illustrated)." (EX1005, [0027].) Benesty discloses speech data transmission over a wireless (WiFi) network. (EX1010, 309.) Goldsmith discloses various existing wireless connection systems in a chapter titled "1.4 Current Wireless Systems," which includes WiFi. (EX1042, 8-21.)

524. A POSITA would be motivated to use the known wireless networking techniques of WLAN over IEEE 802.11b, 802.11a, and/or 802.11g to improve the usability of Williams-Lim-Benesty in the same way. Combining the teachings of Williams-Lim-Benesty and Goldsmith as such would yield predictable results of effectively improving usability and convenience of the control system.

525. A POSITA would have looked to Goldsmith's disclosure of WLAN based on IEEE 802.11b, IEEE 802.11a, and IEEE 802.11g, and understood that the WLAN (WiFi) communications between devices as disclosed in Williams-Lim-Benesty can be WLAN based on IEEE 802.11b,

802.11a, and/or 802.11g, especially since (1) Williams-Lim-Benesty expressly provided for WiFi communication (a form of WLAN) and (2) Goldsmith explains the popularity of 802.11b wireless LANs and the wide coverage of locations with 802.11b base stations installed. (EX1042, 15-16.) Lim itself discloses WiFi and thus suggests a motivation to combine. *Id*. Further, Goldsmith discloses that 802.11a can support more users at higher data rates, and 802.11g can be used in either the 2.4 GHz and 5 GHz bands. Goldsmith discloses that "many wireless LAN cards and access points support all three standards to avoid incompatibilities." *Id*. Customer expectations, strong market forces, and product design requirements and incentives would prompt the application of a known and popular standard (*e.g*., WLAN over 802/11a/b/g) in Williams-Lim-Benesty.

526. Further, the combination is straightforward and readily achievable by a person of ordinary skill in the art without requiring undue effort and with a reasonable expectation of success, as demonstrated by successful implementations in the field. There was a finite number of potential wireless network communications, such as infrared, BlueTooth, Zigbee, 3G, 4G, UWB, and GSM, that were commonly known in the industry and were easy to apply to Williams-Lim-Benesty. The known benefits and feasibility of including a well-known wireless networking technique such as

WLAN based on IEEE 802.11b, 802.11a, and/or 802.11g as taught by Goldsmith to connect in-home devices with a remote server would prompt the application of Goldsmith in the same field based on design incentives and other market forces.

2. Claim 21: "The device according to claim 1, wherein the WLAN uses a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

527. Goldsmith discloses claim 21. (*Supra* § IX.F.3.)

528. As discussed above, a POSITA would be motivated to combine the teachings of Goldsmith and Williams-Lim-Benesty and apply the well-known wireless communication techniques such as WLAN over 802.11b or WLAN over 802.11g, which are ISM radio band frequencies, as taught in Goldsmith to Williams-Lim-Benesty. (*Supra* §IX.L.1.)

M. Ground 13: Williams, Lim, Benesty, and Slaby Render Obvious Claims 14, 23, 28-29

1. Claim 14: "The device according to claim 1, wherein the light is a non-visible light that is infrared, ultraviolet, X-rays, or gamma rays."

529. Claim 1 is obvious in view of Williams, Lim, and Benesty. (*Supra* § IX.H.3.) Claims 14, 23, 28-29 are obvious in further view of Slaby.

530. Slaby discloses claim 14. (*Supra* § IX.G.2.)

531. It would have been obvious to a POSITA to combine the teachings of Williams-Lim-Benesty and Slaby. (*Infra* § IX.M.2.)

2. Claim 23: "The device according to claim 1, further comprising in the enclosure an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, the actuator that is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN."

532. Slaby discloses claim 23. (*Supra* § IX.G.3.)

533. It would have been obvious to a POSITA to combine the teachings of Williams-Lim-Benesty and Slaby. Lim and Slaby both relate to providing a touchscreen for receiving touch inputs. Lim's controlling device 100 includes "soft keys such as a touch sensitive surface overlaid on a liquid crystal (LCD), and/or an EL display," and similarly, Lim's smart device 110, includes "touch keys over laying LCD 306[.]" (EX1005, [0019], [0023].) Slaby's device 102 includes a "touchscreen 104" that is "temperature sensitive." (EX1020, 5:47-50, 5:65-6:2.) Thus, Lim and Slaby are from the same field of endeavor, and both relate to the same technical problems of providing a touchscreen for receiving touch inputs.

534. Slaby incorporates temperature-sensitive touchscreen with multiple thermal energy emitter/detector devices to improve detection. Slaby discloses that it is advantageous to include multiple thermal energy emitter/detector devices in a mobile device for "improved capabilities in terms

of detecting one or more mobile device operational conditions and providing support for improved detection capabilities [of the touchscreen]." (EX1020, 3:36-39.) Further, Slaby's temperature-sensitive touchscreen provides various functionalities such as to "produce[] a corresponding thermal image," "authenticate the user computer device" by generating thermal patterns/images and to "auto-bias the temperature of the temperature sensitive user interface in order to better assure proper operation of the interface in all operating conditions." (EX1020, 3:45-48, 3:52-56, 20:19-39.)

535. A POSITA would have been motivated to combine Williams-Lim-Benesty and Slaby to include Slaby's thermal energy emitter/detector devices in Williams-Lim-Benesty's touch screen or include Slaby's touchscreen in Williams-Lim-Benesty's device, thereby resulting in a system with temperature-sensitive touchscreen to provide the same types of enhancements described by Slaby, (*e.g.*, device authentication, display of thermal images, auto-bias of use interface temperature to improve detection).

536. Incorporating Slaby's thermal energy emitter/detector devices in Williams-Lim-Benesty yields the predictable result of improving the usability of the device. Slaby's "auto-bias" functionality would assure proper operation of the temperature sensitive user interface in all operating conditions and improve detection on the touchscreen, and Slaby further teaches using

thermal patterns/images to authenticate the user computer device. Lim itself discloses a touch screen and thus suggests a motivation to combine. (EX1005, [0019], [0023]; EX1020, 5:47-50, 5:65-6:2.) Further, it would have been obvious to try combining the teachings of Williams-Lim-Benesty and Slaby. There are only a finite number of potential solutions to improve touchscreen detection and assure proper touchscreen operation across operating conditions such as improving touch-sensing algorithms, applying environmental/temperature compensation and recalibration, adding active heating/auto-biasing, and incorporating multiple sensing elements. The combination is straightforward and readily achievable by a person of ordinary skill in the art, as demonstrated by successful implementations in the field. Prior art such as Slaby has successfully demonstrated that a device may use the thermal energy emitting/detector devices to "display thermal image," "authenticate the user computer device," or to "assure proper operation" of the user interface. Thus, there would have been a reasonable expectation of success for combining Williams-Lim-Benesty and Slaby. The above discussed benefits of Slaby would prompt the application of Slaby in the same field based on design incentives and other market forces.

3. Claim 28: "The device according to claim 23, wherein the actuator consists of, or comprises, a thermoelectric actuator that is a heater or a cooler and is configured for affecting a temperature of a solid, a

liquid, or a gas object, by conduction, convection, force convention, thermal radiation, or by transfer of energy by phase changes."

537. Slaby discloses claim 28. (*Supra* § IX.G.4.)

538. It would have been obvious to a POSITA to combine the teachings of Williams-Lim-Benesty and Slaby. (*Supra* § IX.M.2.)

4. Claim 29: "The device according to claim 23, wherein the physical phenomenon comprises temperature, humidity, pressure, audio, vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

539. Slaby discloses claim 29. (*Supra* § IX.G.5.)

540. It would have been obvious to a POSITA to combine the teachings of Williams-Lim-Benesty and Slaby. (*Supra* § IX.M.2.)

N. Ground 14: Lindahl and Evermann Render Obvious Claims 1-2, 4-8, 11-26, 29

1. Summary of Evermann

541. Evermann was filed on February 12, 2007, and published on June 26, 2008, and is prior art under §102(a), (b), and (e). (EX1012, 1.) The Examiner never found, and Applicant never cited, Evermann during prosecution.

542. Evermann discloses a "mobile device 102" as shown in FIG. 1b. (EX1012, [0020].) Such mobile device "receiv[es] utterance that includes a search request from a user of the device" send the search request to a "remote server over a wireless data connection." (EX1012, abstract.)

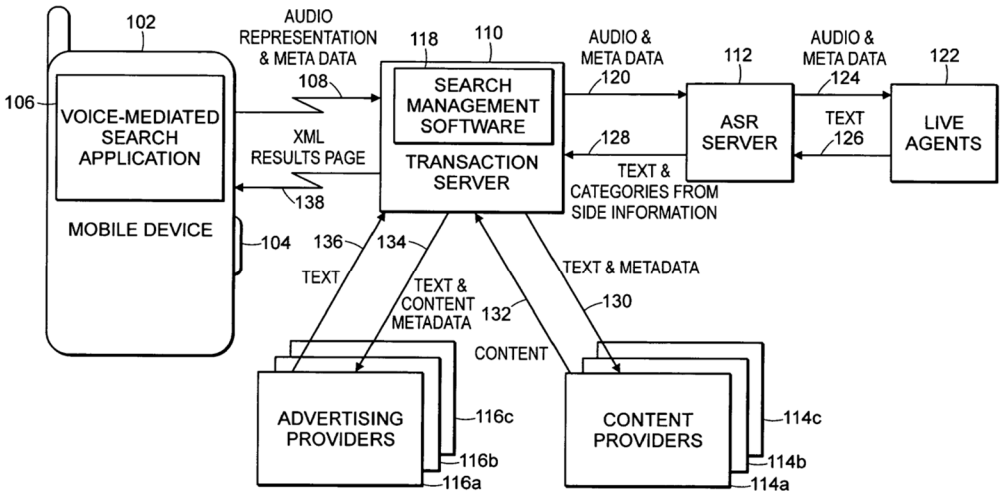


FIG. 1

543. Evermann's mobile device has a microphone, a display, a speaker, WLAN transceiver, battery, processor and associated memory. (EX, [0085]-[0086], [0020], FIG. 7.)

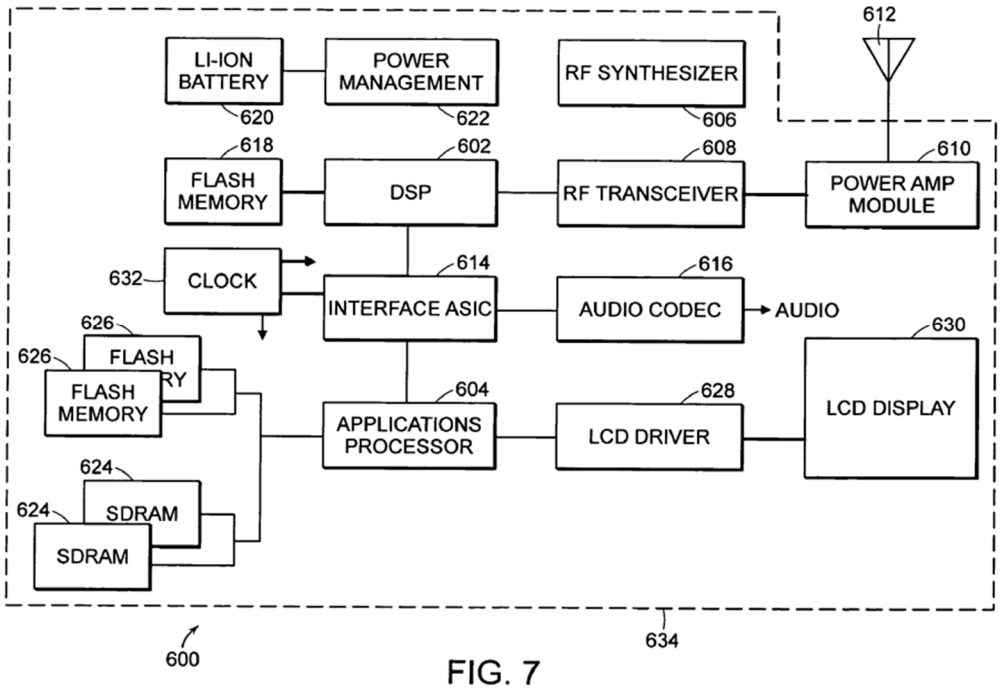


FIG. 7

544. Evermann's mobile device sends human voice data to a server device and receives multimodal results and advertisements from the server for playback and display. Evermann's device "recognizes the word 'search' and sends the ***utterance that follows*** to one or more ***remote servers*** for speech recognition and further handing[.]" (EX1012, [0030]-[0031], [0045].) "***Transaction server 110*** transmits the combined ***search results and advertisements*** to the phone over ***connection 138***," and such "results can be multimodal, *i.e.*, include text, ***graphics audio***, and ***video***." (EX1012, [0069], [0038].) Evermann discloses that "the voice search results and associated advertising content" may be presented to users "as an ***audio*** or a ***video message***." (EX1012, [0038].) For example, a user can say "search ringtones" and "the ***servers*** return the results to device 102, which then ***displays one or more screens of ringtone choices***." (EX1012, [0026].)

2. A POSITA would have been motivated to combine Lindahl and Evermann

545. It would have been obvious for a POSITA to combine Lindahl and Evermann. Lindahl teaches using a portable electronic device to receive and upload voice commands to computing equipment (*e.g.*, computing equipment 14/18) for processing and, after processing the voice command, transmitting information associated with the voice command back to the user

device for presentation. (EX1011, [0073], [0043].) Evermann teaches server-side processing of voice commands and delivery of search results and associated advertising content to a mobile device, including results that may include audio and other rich media formats. (EX1012, abstract, [0018], [0026], [0038].) They are analogous to the '720 patent and both relate to processing/responding to human speech commands via a remote server. (EX1001, 32:60-64.)

546. A POSITA would have been motivated to combine Lindahl and Evermann, resulting in a system where Lindahl's server (*e.g.*, computing equipment 14/18) or Evermann's server transmits search results and associated advertising content, both of which include audio data, to the user device as taught by Evermann. This is an advantageous way to surface relevant advertisement because audio- and video-inclusive advertising content can be presented to the user in connection with the search results returned in response to the user's speech command. Further, transmitting audio or video data from the server (*e.g.*, rather than pure text data) allows rich, multimodal feedback and ensures that the data is consistent, high-quality, and free from manipulation, degradation, and variation that could otherwise occur with device-side generation or rendering. (*Supra* §IX.A.3.) These benefits align with common design incentives and market forces in consumer electronics,

including improving user-facing presentation of results and advertisements and providing consistent multimedia experiences across different user devices.

3. Claim 1

a. 1[PR]: "A device comprising:"

547. Both Lindahl and Williams disclose this element. (*Supra* §IX.A.4.a.) Referring to FIGs. 1 and 2, Evermann discloses a "***mobile device 102*** [which] is a personal wireless communication device. Such as a cellular (cell) phone, that can receive audio input from a user." (EX1012, [0020].)

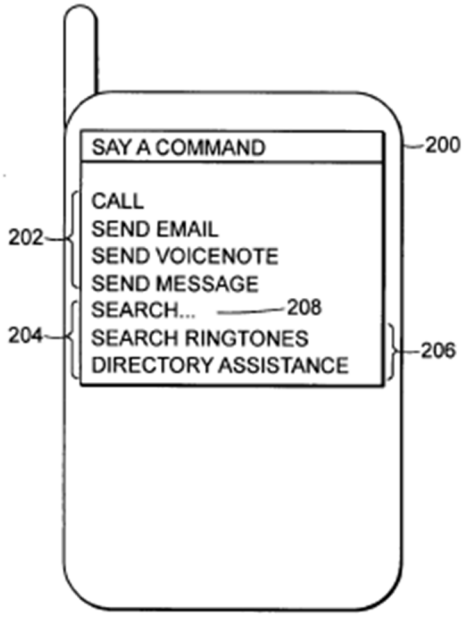


FIG. 2

b. 1[A]: "multiple microphones for capturing first human voice data;"

548. Lindahl-Evermann discloses this. (*Supra* §IX.A.4.b.) Evermann also discloses a microphone for capturing ***first human voice data***. Evermann discloses "a ***microphone***, and other input/output devices provided in the phone." (EX1012, [0086].) Evermann's mobile device 102 "receive[s] ***audio input from a user***" and "the received input includes an ***utterance by the user*** of the mobile device." (EX1012, [0007], [0020], [0025], abstract, claim 7.) Evermann discloses many examples of ***first human voice data***. For example, Evermann discloses that "[t]he device captures the user's ***spoken answer***, transmits it to the remote servers that recognize the speech and retrieve the available ring tones that correspond to the user's selected artist." (EX1012, [0026].)

c. 1[B]: "a Wireless Local Area Network (WLAN) transceiver for transmitting digital data to, and for receiving digital data from, a WLAN;"

549. Both Lindahl and Evermann disclose this element. (*Supra* §IX.A.4.c.) Evermann's mobile device connects with a server through "Wi-Fi connections" via a "***RF radio transceiver 608***." (EX1012, [0072], [0086], [0020].) "Among the items logged are all user actions that involve initiating a connection between mobile device 102 and a remote source . . . Such connections can be achieved via wireless data connection 108, or over other wireless or fixed connections, such as ***Wi-Fi connections***[.]" (EX1012,

[0072].) It was common knowledge that WiFi is a type of WLAN. (EX1042, 600-601; EX1013, 7:38-51; EX1001, 8:25-32.) Thus, Evermann's ***RF radio transceiver 608***, which implements "[t]he transmit and receive functions," is a WLAN transceiver. (EX1012, [0086].) As discussed, "the transmit and receive functions are implemented by . . . an ***RF radio transceiver 608***" which supports Wi-Fi connections, thus, Evermann's WLAN transceiver is for transmitting digital data to, and for receiving digital data from, a WLAN. (EX1012, [0086]; *infra* §§IX.N.3.h-IX.N.3.j.)

d. 1[C]: "a speaker for sounding audio data;"

550. Both Lindahl and Evermann disclose this element. (*Supra* §IX.A.4.d.) Evermann's mobile device includes a "***speaker***" for sounding audio. (EX1012, [0086].) With reference to FIG. 7, which illustrates "a typical platform on which mobile communications device 102 can be implemented," the device includes "an audio CODEC 616 provide interfaces to a ***speaker***, a microphone, and other input/output devices provided in the phone[.]" (EX1012, [0085], [0086].)

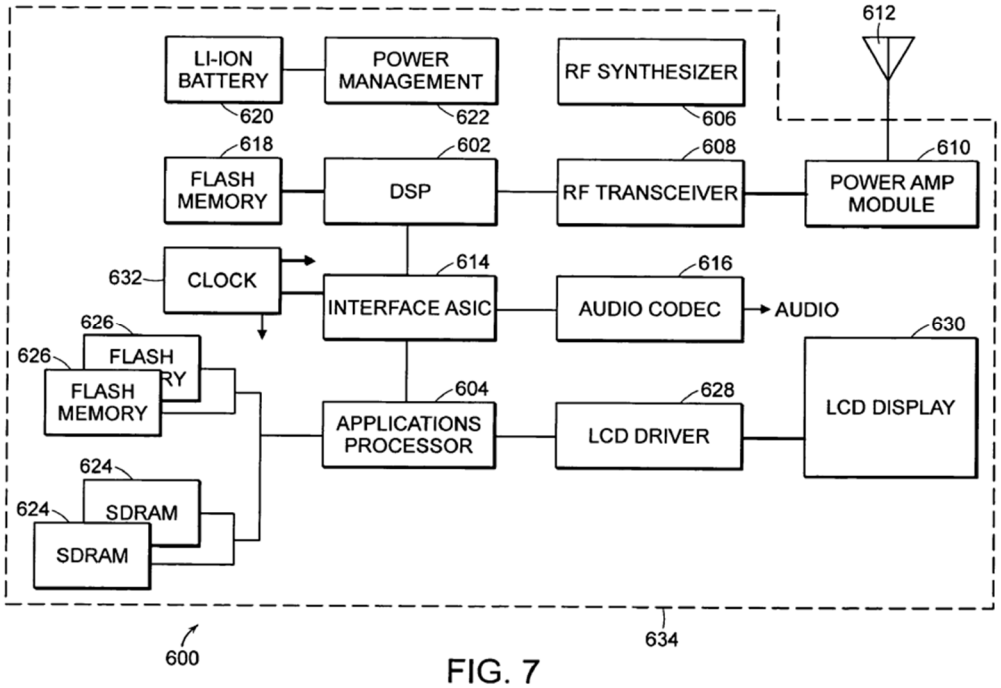


FIG. 7

e. 1[D]: "an electric light source for emitting light for emitting light;"

551. Both Lindahl and Evermann disclose this element. (*Supra* §IX.A.4.e.) Evermann's mobile device 102 includes a "***display for displaying text and graphics***." (EX1012, [0020], [0086], [0004], [0008].) Referring to FIG. 7, Evermann's mobile device 102 includes "a ***display screen***, such as a ***liquid crystal (LCD) display*** or a plasma display." (EX1012, [0086].) "The visual display device for the device includes an LCD driver chi 628 that drives an LCD display 630." (EX1012, [0089].) As discussed above, displays are or comprise electric light sources. (*Supra* §IX.A.4.e; EX1001, 39:10-21.)

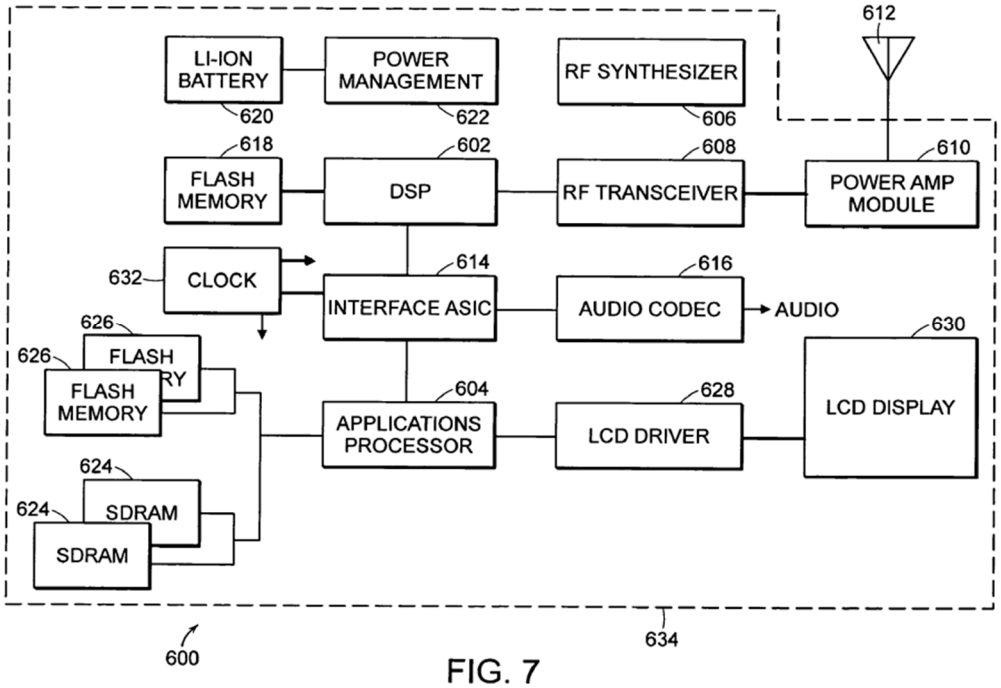


FIG. 7

f. 1[E]: "an enclosure that houses the multiple microphones, the WLAN transceiver, the speaker, and the electric light source;"

552. Lindahl-Evermann discloses this. (*Supra* §IX.A.4.f.) As shown in FIG. 7 below, Evermann's mobile device has a "***designed housing 634***" that comprises all claimed components, "an RF radio transceiver 608," "a speaker, a microphone," "a display screen, such as a liquid crystal (LCD) display" (EX1012, [0086], [0089], FIG. 7.)

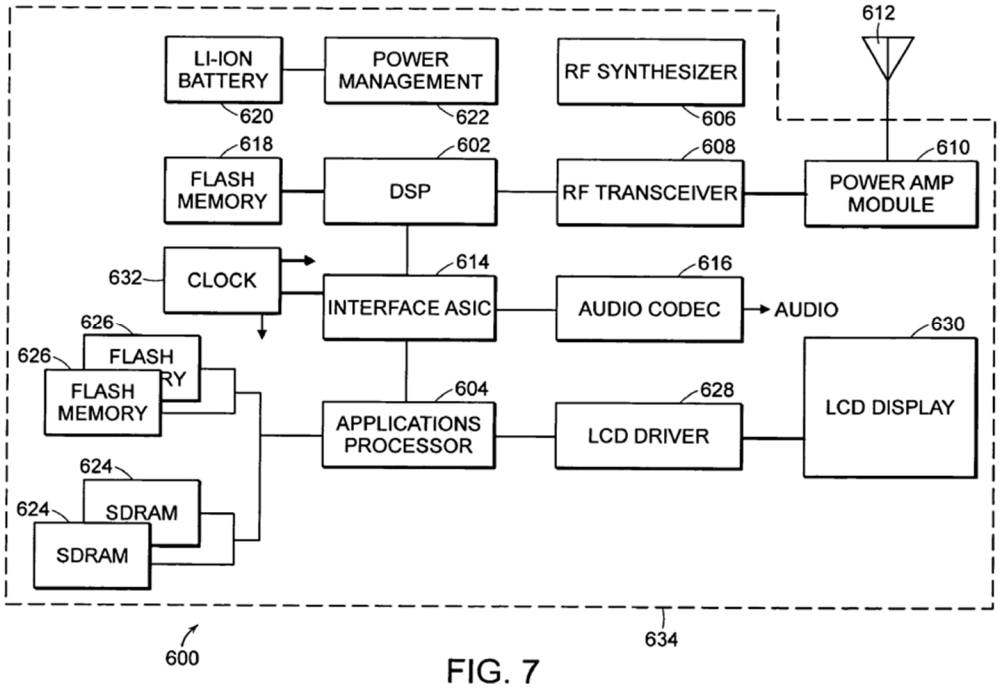


FIG. 7

553. As shown in FIG. 2, Evermann's mobile device has an enclosure formed by the outer shell and the front panel of the display.

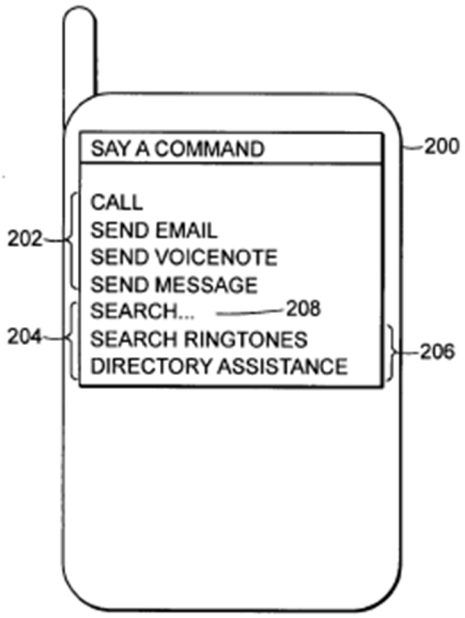


FIG. 2

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

554. Both Lindahl and Evermann disclose this element. (*Supra* §IX.A.4.g.) Evermann's mobile device includes "a ***processor system*** and ***memory storing code***, which when executed by the processor system, cause the mobile device to perform the functions of: receiving an utterance from a user of the mobile device, the utterance including a search request; presenting the received information on the mobile device[.]" (EX1012, [0009]-[0010], [0085], [0087].) Thus, Evermann's device comprises one or more processors programmed with computer program instructions that, when executed, cause the device to perform the recited operations. (*Infra* §§IX.N.3.h-IX.N.3.j.)

h. 1[G]: "send, to a server device over the Internet via the WLAN by the WLAN transceiver, at least part of the first human voice data captured by the multiple microphones;"

555. Lindahl-Evermann discloses this. (*Supra* §IX.A.4.h.) Evermann also discloses ***server devices***, such as "transaction server 110" and "speech recognition servers 112." (EX1012, [0039]-[0040]. [0091].) Evermann's device sends at least part of the first human voice data captured by the multiple microphones to the server device over the Internet via the WLAN by the WLAN transceiver. For example, Evermann's device "recognizes the word 'search' and sends the ***utterance that follows*** to one or

more ***remote servers*** for speech recognition and further handing[.]" (EX1012, [0030]-[0031], [0045].)

556. Evermann disclose that "when the user says '***search ring-tones***.' The device responds with a spoken and displayed prompt 'what artist?' The user then speaks the name of the artist. The device captures the user's ***spoken answer***, transmits it to the ***remote servers*** that recognize the speech and retrieve the available ring tones that correspond to the user's selected artist." (EX1012, [0026].)

557. As discussed above, the connections between the mobile device and the remote servers are through wireless connections. "Such connections [between the mobile device and remote source] can be achieved via wireless data connection 108, or over other wireless or fixed connections, such as ***Wi-Fi connections***[.]" (EX1012, [0072].) It was common knowledge that WiFi is a type of WLAN. (EX1042, 600-601; EX1013, 7:38-51; EX1001, 8:25-32.) Further, Evermann's mobile device accesses information stored in the server via the Internet. "The described embodiment is a mobile device and ***server system*** that provides a user of the mobile device with voice-mediated access to a wide range of information, such as directory assistance, financial data, or to search the Web. In general, this information is not stored on the device itself, but is stored on any ***server*** or other device to which the

mobile device has access either via predetermined relationship, or via a public access network, such as the ***Internet***." (EX1012, [0018].)

558. As discussed, Evermann's ***RF radio transceiver 608*** implements "the transmit and receive functions." (*Supra* §IX.N.3.c; EX1012, [0086].) A POSITA would understand (and found it obvious) that such transmission of human voice data to the server is via the WLAN transceiver.

i. 1[H]: "play, by the speaker, an audio data received from the server device over the Internet via the WLAN by the WLAN transceiver;"

559. Lindahl-Evermann discloses this. (*Supra* §IX.A.4.i.) Evermann's mobile device "receiv[es] ***information*** over the wireless data connection that is responsive to the search request." (EX1012, abstract.) For example, "***[t]ransaction server 110*** transmits the ***combined search results and advertisements*** to the phone over ***connection 138***," and such "results can be multimodal, *i.e.*, include text, ***graphics audio***, and ***video***." (EX1012, [0069], [0038], [0018].) Evermann discloses that "the voice search results and associated advertising content" may be presented to users "as an ***audio*** or a ***video message***." (EX1012, [0038].) For example, "[t]he device's voice mediated search software displays the results to the user as text, graphics, and ***video*** and, optionally as ***audio output*** of synthesized speech, ***sounds, or music***." (EX1012, [0018].) Thus, it would be obvious that such "***graphics***

audio and ***video***" received from the ***server*** are played on Evermann's ***speaker***. (EX1012, [0086].)

560. Further, in or before 2012, advertisements accompanied with music were known and studies have shown that it is beneficial to include audio such as popular music in advertisements as "song vocals, either original or altered, are more effective stimuli of advertising effects than instrumentals or no popular music." (EX1022, 434.) Thus, a POSITA would have it obvious that such "***advertisements***" transmitted by the ***transaction server 110*** to the phone would be accompanied with audio such as popular music or instrumentals.

561. As discussed, Evermann discloses sending at least part of the first human voice data captured to the server over the Internet via the WLAN, so a POSITA would understand and found it obvious that the audio data transmitted from server back to the mobile device is through the same communication path. Indeed, Evermann discloses that such connection 138 "is a ***wireless connection***," that the "connection between mobile device 10 and a remote source" can be via "Wi-Fi connections," and that the mobile device has access to information stored on any server via "the Internet." (EX1012, [0041], [0072], [0018].) As discussed, Evermann's ***RF radio transceiver 608*** implements "the transmit and receive functions." (*Supra*

§IX.N.3.c; EX1012, [0086].) It would have been obvious that the receipt of audio data is by the WLAN transceiver. (EX1012, [0072], [0086], [0020].)

562. A POSITA would have been motivated to combine Lindahl and Evermann, resulting in a system where Lindahl's server (*e.g.*, computing equipment 14/18) or Evermann's server transmits results/advertisement in the form of audio data (*e.g.*, audio or video) to the user device for playback as taught by Evermann. (*Supra* §IX.N.2.) Multi-modal feedback improves the usability of the device. Especially as Evermann teaches multi-modal feedback through both "spoken and displayed prompt." (EX1012, [0026].)

j. 1[I]: "and emit, by the electric light source, in response to data received from the server over the Internet via the WLAN by the WLAN transceiver,"

563. Lindahl-Evermann discloses this. (*Supra* §IX.A.4.j.)

564. Evermann discloses that "[w]hen ***transaction server 110*** returns the ***voice search results and associated advertising content*** to mobile device 102, VMSA 106 . . . ***presents*** [the information] to the user as ***text and graphics*** on the device's ***display***." (EX1012, [0038], [0018], abstract.) For example, a user can say "search ringtones" and "the ***servers*** return the results to device 102, which then ***displays one or more screens of ringtone choices***." (EX1012, [0026].) Thus, Evermann's display emits in the device to present

the server-received results/information to user. (EX1001, 39:19-26, 109:45-50, 110:64-111:3, 111:11-23.)

565. As discussed above, the "voice search results and associated advertising content" are sent to the mobile device 102 from the transaction server 110 over the Internet via the WLAN and by the WLAN transceiver. (*Supra* §§IX.N.3.c, IX.N.3.i.)

k. 1[J]: "wherein the device is addressed in the WLAN using an Internet Protocol (IP) address that uniquely identifies the device in the WLAN or in the Internet."

566. Lindahl-Evermann discloses this. (*Supra* §IX.A.4.k.) As discussed, at the time of the invention, it was well-known that networked devices are associated with IP addresses that uniquely identifies the device in the WLAN and in the Internet. (*Id*.; EX1007, 99.) It would have been obvious Evermann's networked device is addressed with an associated IP address. (*Supra* §§IX.N.3.c, IX.N.3.h.) 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. (EX1007, 99, 108, 175, 210, 340.) These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers

using the Network Address Translation ("NAT") feature of the router to communicate with devices outside of the user's WLAN. *Id*.

4. Claim 2: "The device according to claim 1, further comprising, in the enclosure, a rechargeable battery for powering the device."

567. Both Lindahl and Evermann disclose or at least suggest this claim. (*Supra* §IX.A.5.) Evermann's mobile device comprises "a Li-Ion battery 620 [that] powers the phone." (EX1012, [0087].) It would be obvious that Evermann's lithium battery is rechargeable to improve usability and portability of portable mobile phone intended for frequent, repeated use.

5. Claim 4: "The device according to claim 1, further being addressable using the IP address in the WLAN or in the Internet using distinct locally administered addresses or a universally administered digital addresses."

568. Both Lindahl and Evermann disclose or at least suggest this claim. (*Supra* §IX.A.6.) As discussed above, Evermann discloses or at least suggests that the device may be associated with 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. (*Supra* §IX.N.3.k; EX1007, 99, 108, 175, 210, 340.) These IP addresses uniquely identify the device within the user's WLAN, and further these devices would also be addressed using the router's globally unique IP address when communicating with Internet servers using the Network Address Translation ("NAT") feature of the router to

communicate with devices outside of the user's WLAN. *Id*. A POSITA would have understood the router assigned IP addresses to be locally administered addresses and the router's globally unique IP address to be a universally administered address. (EX1007, 108, 210, 175, 340.)

569. Further, Evermann discloses or at least suggests that its device has a MAC address that is a universally administered digital address. It was common knowledge, and a POSITA would have understood that communicating via a wireless network and with a remote server over the Internet via a wireless network (as in Evermann) involves associating the device with a device identifier or address such as a MAC address that uniquely identifies the device in the network. (*Supra* §IX.N.3.h, IX.A.4.h; EX1007, 99, 166; EX1023, 24.) Furthermore, both the universally administered MAC address and the locally administered IP address are used in conjunction with each other. (EX1023, 218, *see also id.,* 312, 313, Figure 6.6.)

6. Claim 5: "The device according to claim 4, wherein the IP address is a layer 3 address and is static or dynamic IP address that is IPv4 or IPv6 type address."

570. Both Lindahl and Evermann disclose or at least suggest this claim. (*Supra* §IX.A.7.) As discussed above, Evermann discloses that the

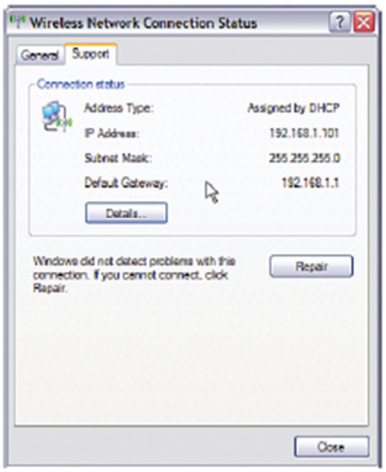
device is addressable in the Internet and WLAN with an IP address. (*Supra* §IX.N.3.h.)

571. As discussed, at the time of the invention, it was well-known that Internet Protocol (IP) addresses are either static or dynamic, are Layer 3 addresses and are IPv4 addresses. (EX1029, 127; EX1023, 11; EX1007, 277; EX1024, 2:60-62.)

7. Claim 6: "The device according to claim 1, further comprising a volatile or non-volatile memory that stores the IP address."

572. Both Lindahl and Evermann disclose this claim. (*Supra* §IX.A.8.) Evermann's mobile device includes "flash memory 618," "SDRAM 624 and flash memory 626." (EX1012, [0087], [0088].) A POSITA would have understood that Evermann's memory such as the "flash memory 618," "SDRAM 624 and flash memory 626" would be used to store the associated IP addresses. For example, as discussed above and shown below, networked devices can display the IP addresses which means the IP addresses are locally stored in the device. (*Supra* § IX.A.8; EX1007, 211, Figure 10-44, *see also, Id*., 265.)

Figure 10-44 Verifying IP Address Assignment with Windows XP



8. Claim 7: "The device according to claim 1, wherein the speaker comprises an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker (ESL), a ribbon or planar magnetic loudspeaker, or a bending wave loudspeaker."

573. Both Lindahl and Evermann disclose this claim. As discussed, Everman discloses a ***speaker*** in the device for playing audio data, consistent with the descriptions of the '720 patent. (*Supra* §§ IX.N.3.d, IX.A.4.d.) As discussed, a POSITA would have understood that an electromagnetic loudspeaker, a piezoelectric speaker, an electrostatic loudspeaker, a ribbon or planar magnetic loudspeaker are standard and common implementations of speakers. (*Supra* §IX.A.9.) Thus, a POSITA would have understood (or it would be obvious) that Lindahl-Evermann's speaker is the recited speaker in claim 7.

9. Claim 8: "The device according to claim 1, wherein the speaker is operative for omnidirectional, unidirectional, or bidirectional pattern audio emitting."

574. Both Lindahl and Evermann disclose this claim. As discussed, Everman discloses a standard speaker in the device. (*Supra* §§ IX.N.3.d, IX.A.4.d.) As discussed, a POSITA would have understood that a standard speaker is operative for omnidirectional pattern audio emitting. (*Supra* §IX.A.10; EX1018, 122; EX1014, 763, 765.) Thus, a POSITA would understand (or it would be obvious) that Williams-Lim-Benesty's speaker is a speaker as recited in claim 8.

10. Claim 11: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, an omnidirectional, unidirectional, or bidirectional microphone that is based on sensing of an incident sound-based motion of a diaphragm or a ribbon."

575. Both Lindahl and Evermann disclose this claim. (*Supra* §IX.A.11.) As discussed above, Lindahl-Evermann discloses multiple microphones in the device. (*Supra* §IX.N.3.b.) As discussed, a POSITA would have understood that a standard microphone is an omnidirectional microphone based on sensing an incident sound-based motion of a diaphragm. (EX1014, 757, 309.) Thus, a POSITA would understand (or it would be obvious) that each of Lindahl-Evermann's standard multiple microphones is the recited microphone in claim 11.

11. Claim 12: "The device according to claim 1, wherein each of the multiple microphones comprises, or consists of, a condenser, an electret, a dynamic, a ribbon, a carbon, or a piezoelectric microphone."

576. Both Lindahl and Evermann disclose this claim. (*Supra* §IX.A.12.) As discussed above, Lindahl-Evermann discloses multiple microphones in the device. (*Supra* §IX.N.3.b.) As discussed, a POSITA would understand that dynamic microphone and piezoelectric microphone are standard and common implementations of microphones. (EX1014, 758.)

12. Claim 13: "The device according to claim 1, wherein the light is a visible light for illumination or indication."

577. Both Lindahl and Evermann disclose this claim. As discussed, Evermann's mobile device includes a "***display for displaying text and graphics***." (EX1012, [0020], [0086], [0004], [0008].) Such display may be "a ***liquid crystal (LCD) display***" and it would be obvious that it would include LEDs. (EX1012, [0086]; *supra* §§IX.N.3.e, IX.A.13.)

578. As discussed, a POSITA would have understood that LEDs are electric light sources that emit a visible light for illumination or indication, consistent with the '720 patent. (*Supra* §IX.A.4.e; EX1018, 174; EX1001, 39:15-26, 89:50-57.)

13. Claim 14: "The device according to claim 1, wherein the light is a non-visible light that is infrared, ultraviolet, X-rays, or gamma rays."

579. Both Lindahl and Evermann disclose this claim. (*Supra* §IX.A.14.) As discussed, Evermann's mobile device includes a "***display for displaying text and graphics***." (EX1012, [0020], [0086], [0004], [0008].)

Such display may be "a ***liquid crystal (LCD) display***" and it would be obvious that it would include LEDs. (EX1012, [0086]; *supra* §§IX.N.3.e, IX.A.13.) As discussed, LCD displays and LEDs emit non-visible light. (*Supra* § IX.A.13.)

14. Claim 15: "The device according to claim 1, wherein the electric light source consists of, or comprises, a Solid-State Lighting (SSL), a Light Emitting Diode (LED), an Organic LED (OLED), a polymer LED (PLED), or a laser diode."

580. Both Lindahl and Evermann disclose this claim. (*Supra* §IX.A.15.) As discussed, Evermann's mobile device includes a "***display for displaying text and graphics***." (EX1012, [0020], [0086], [0004], [0008].) Such display may be "a ***liquid crystal (LCD) display***" and it would be obvious that it would include LEDs. (EX1012, [0086]; *supra* §§IX.N.3.e, IX.A.14.)

15. Claim 16: "The device according to claim 1, further comprising a memory that stores digital audio content, and the one or more processors are further programmed with computer program instructions that, when executed, cause the device to playing by the speaker the stored digital audio content."

581. Lindahl-Evermann discloses this claim. (*Supra* §IX.A.16.) Evermann's mobile device "displays the results to the user as text, graphics, and video and, optionally as audio output of ***synthesized speech, sounds, or music***." (EX1012, [0018].) A POSITA would understand (and it would be obvious) that Evermann's device stores the synthesized speech in its "SDRMA 624 or flash memory 626" and its "processor system and memory

storing code" are programmed with instructions that plays the speech when executed. (EX1012, [0009], [0088].) Evermann expressly provides that its "SDRMA 624 or flash memory 626" stores "advertisements" to be played. *Id*.

582. Further, Evermann discloses that the user may be provided with "available ring tones that correspond to the user's selected artist . . . The user can select a ringtone, and the device then downloads his selection to the device." (EX1012, [0026].) A POSITA would understand (and it would be obvious) that the ringtones (e.g., digital audio content) are stored in the device memory and Evermann's processors are further programmed with computer program instructions that, when executed, cause the device to play by the speaker, the stored digital audio content (*e.g*., downloaded ringtones). Indeed, Evermann's "processor can also run additional applications, such as a digital music player." (EX1012, [0085].)

16. Claim 17: "The device according to claim 16, wherein the digital audio content is pre-recorded or synthesized audio content."

583. Lindahl-Evermann discloses this claim. Evermann's mobile device "displays the results to the user as text, graphics, and video and, optionally as audio output of ***synthesized speech, sounds, or music***." (EX1012, [0018].)

17. Claim 18: "The device according to claim 16, further operative to play a simulating of a voice of a human being or to play music."

584. Lindahl-Evermann discloses this claim. Evermann's mobile device "displays the results to the user as text, graphics, and video and, optionally as audio output of ***synthesized speech, sounds, or music***." (EX1012, [0018].) Thus, Evermann's device is operative to play music or a speech which is simulating of a voice of a human being.

18. Claim 19: "The device according to claim 18, wherein the digital audio content comprises a syllable, a word, a phrase, a sentence, a short story, or a long story, using male or female voice."

585. Lindahl-Evermann discloses this claim. Evermann's mobile device "displays the results to the user as text, graphics, and video and, optionally as audio output of ***synthesized speech, sounds, or music***." (EX1012, [0018].) For example, Evermann discloses that "when the user says 'search ring-tones.' The device responds with a ***spoken*** and displayed ***prompt*** '***what artist***?' The user then speaks the name of the artist. The device captures the user's spoken answer, transmits it to the remote servers that recognize the speech and retrieve the available ring tones that correspond to the user's selected artist." (EX1012, [0026].) Thus, Evermann's device stores and plays digital audio content that comprises a syllable, a word, a phrase, a sentence etc., using male or female voice.

19. Claim 20: "The device according to claim 1, wherein the WLAN is according to, or compatible with, IEEE 802.11-2012, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, or IEEE 802.11ac standard."

586. Both Lindahl and Evermann disclose or at least suggest this claim. (*Supra* § IX.A.20.) As discussed, Evermann's mobile device communicates with a remote server via a "WiFi connection." (*Supra* §§IX.N.3.c, IX.N.3.h.) As discussed, WiFi is a form of WLAN based on the IEE 802.11 standard group. (EX1042, 600-601; EX1013, 7:48-51.)

20. Claim 21: "The device according to claim 1, wherein the WLAN uses a wireless communication over a licensed or an unlicensed radio frequency band, that is an Industrial, Scientific and Medical (ISM) radio band."

587. Both Lindahl and Evermann disclose or at least suggest this claim. (*Supra* § IX.A.21.) As discussed, Evermann's mobile device communicates with a remote server via a "WiFi connection." (*Supra* §§IX.N.3.c, IX.N.3.h.) As discussed, WiFi is a form of WLAN based on the IEE 802.11 standard group. (EX1042, 600-601; EX1013, 7:48-51.) IEE802.11b are ISM band II frequencies. (EX1042, 600-601, 22-23.)

21. Claim 22: "The device according to claim 1, wherein the enclosure is a handheld enclosure."

588. Both Lindahl and Evermann disclose this claim. (*Supra* § IX.A.22.) As discussed and shown in FIGs 1 and 2, Evermann's device is a

mobile device. Thus, Evermann's mobile device 12's "designed housing 634" is a handheld mobile device enclosure.

22. Claim 23: "The device according to claim 1, further comprising in the enclosure an actuator that is directly or indirectly affecting, changing, producing, or creating a physical phenomenon, the actuator that is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN."

589. Lindahl-Evermann discloses this claim. (*Supra* § IX.A.23.) As discussed, Lindahl discloses multiple speakers in the device, each of which is an actuator that is directly or indirectly affecting, changing producing, or creating a physical phenomenon, namely sound or audio. *Id*. Lindahl discloses that "[d]isplay and audio devices 52 may also include audio equipment such as ***speakers and other devices for creating sound***." (EX1011, [0048].) As discussed below, a speaker system comprises a digital-to-analog converter ("DAC"), which is an actuator. (*Infra* §IX.N.25.) The speaker system is comprised in the enclosure. A POSITA would understand that the speaker/speaker system is coupled to the WLAN transceiver for being activated or controlled in response to data received from the WLAN.

23. Claim 24: "The device according to claim 23, wherein the actuator consists of, or comprises, a motion actuator that causes linear or rotary motion."

590. Lindahl-Evermann discloses this claim. (*Supra* § IX.A.24.) As discussed, Lindahl discloses multiple speakers in the device. (*Id.*; EX1011, [0048].) A POSITA would have understood that a speaker is a motion actuator that causes "linear motion." (EX1014, 763, 903, 18.)

24. Claim 25: "The device according to claim 23, wherein the actuator consists of, or comprises, an electromagnetic coil or an electromagnet operative for generating a magnetic or electric field."

591. Lindahl-Evermann discloses this claim. (*Supra* § IX.A.25.) As discussed, Lindahl discloses multiple speakers in the device. (EX1011, [0048].) A POSITA would have understood that a speaker commonly comprises an electromagnetic coil operative for generating a magnetic or electric field, consistent with the specification of the '720 patent. (EX1001, 105:64-106:7; EX1018, 122, 123.)

25. Claim 26: "The device according to claim 23, wherein the actuator consists of, or comprises, an electrical signal generator."

592. As discussed, it was well-known that a speaker system comprises a DAC, an electrical signal generator actuator, for converting received digital audio into analog signals for playback. (*Supra* § IX.A.26; EX1014, 714, 522; EX1018, 170; EX1001, 34:25-28, 107:13-20.) It would have been obvious to a POSITA that Lindahl-Evermann's device would include a DAC.

26. Claim 29: "The device according to claim 23, wherein the physical phenomenon comprises temperature, humidity, pressure, audio,

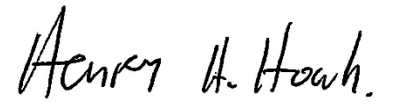
vibration, light, motion, sound, proximity, flow rate, electrical voltage, or electrical current."

593. Both Lindahl and Evermann disclose this claim. (*Supra* §IX.A.27.) As discussed, Lindahl discloses multiple speakers in its device. (*Supra* §IX.A.23; EX1011, [0048].) A POSITA would have understood that a speaker is configured to directly or indirectly affect, change, produce, or create a physical phenomenon, namely, audio or sound. Further, Lindahl-Evermann's device comprises a digital-to-analog converter, a POSITA would have understood the DAC is configured to directly or indirectly affect, change, or produce a physical phenomenon, namely, electrical voltage, or electrical current.

* * *

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and that these statements were made with knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code.

Dated: March 23, 2026



Henry H. Houh

Henry H. Houh

15 Ballard Ter.
Lexington, MA 02420
henry.houh@gmail.com
cell: 781-325-8273

Education

Massachusetts Institute of Technology, Cambridge, MA

- PhD in Electrical Engineering and Computer Science, February 1998. "Designing Networks for Tomorrow's Traffic," thesis supervised by Professor David Tennenhouse and Professor John Guttag. GPA 4.7/5.0
- Master of Science in Electrical Engineering and Computer Science, February 1991. "Demonstration of a laser repetition rate multiplier," thesis. GPA 4.5/5.0
- Bachelor of Science in Electrical Engineering and Computer Science, June 1989. "Boundary element analysis of arbitrarily shaped dielectric structures," thesis. GPA 4.7/5.0
- Bachelor of Science in Physics, February 1990. GPA 4.7/5.0

MITx, Cambridge, MA

- 2020 – 2021: Completed four courses in the Statistics and Data Science MicroMasters program, a sequence of MIT-offered graduate-level courses in statistics and machine learning.

Experience

Houh Consulting Inc. / Independent Consultant

- 2009 - present: Technical consultant specializing in Social Networking, Web 2.0, Web Site Development, Data Networking, Optical Networking, Telecommunications, Media Streaming and Voice Over IP. Clients include: Akin Gump Strauss Hauer & Feld LLP, Arnold & Porter Kaye Scholer LLP, BBN, Baker Botts LLP, Barnes & Thornburg LLP, BBN Technologies, Beus Gilbert McGroder PLLC, Bracewell LLP, Christensen O'Connor Johnson Kindness PLLC, Cooley LLP, Covington & Burling LLP, Cozen O'Connor P.C., DLA Piper, Davis Polk & Wardwell LLP, Desmarais LLP, Duane Morris LLP, Erise IP, P.A., Faegre Baker Daniels LLP, Finnegan, Henderson, Farabow, Garrett & Dunner, LLP, Fish & Richardson P.C., Goldman Ismail Tomaselli Brennan & Baum LLP, Haynes and Boone, LLP, Kaye Scholer, Kellogg Huber Hanson Todd Evans & Figel PLLC, Kirkland & Ellis LLP, Klarquist Sparkman, LLP, McCarthy Tetrault LLP, McGuireWoods LLP, Morgan, Lewis & Bockius LLP, Munger Tolles & Olson LLP, Orrick, Herrington & Sutfliffe LLP, O'Melveny & Myers LLP, Orrick, Herrington & Sutcliffe LLP, Paul Hastings LLP, Paul, Weiss, Rifkind, Wharton & Garrison LLP, Perkins Coie LLP, Quinn Emanuel Urquhart & Sullivan, LLP, Sidley Austin LLP, Smith Baluch LLP, Sterne, Kessler, Goldstein & Fox PLLC, Troutman Sanders LLP, Venable LLP, Weil, Gotshal & Manges LLP, Wiley Rein LLP, Wilmer Cutler Pickering Hale and Dorr LLP, and Winston & Strawn LLP.

BlocksCAD Inc. (H3XL spin-out)

- 2014 – 2025: Founder and member of Board of Directors. BlocksCAD is a visual Computer-Aided Design (CAD) program, designed to teach children mathematics and geometric visualization. BlocksCAD is being used in U.S. schools and worldwide in maker spaces, with 350,000+ users. BlocksCAD was selected to be part of the LearnLaunch 2017 Boost Accelerator, the MIT Play Labs 2018 Summer Accelerator, and the MassChallenge 2018 Boston Cohort. Recipient of DARPA DSO grant in 2014, and SBIR grant from the USDA in 2019. BlocksCAD was acquired by zSpace in 2025.

H3XL Inc. d/b/a Einstein's Workshop (formerly Lexington Robotics)

- 2009 - 2022: Founder and President. Started a science, technology, engineering and math enrichment program and creative/maker space, in 7,000 square feet of space. Serve 2,000+ kids and families annually. As of end of 2022, have delivered an estimated 250,000 student-instruction-hours of STEM courses. Principle Investigator for 2-year DARPA grant to improve 3D Computer Aided Design tool developed in-house for the purposes of teaching 8 year olds and up 3D CAD.

Eons

- 2008 - 2009: Chief Technology Officer. Created product that Eons acquired from BBN Technologies. Integrated BBN product with Eons social networking platform and significantly increased the Eons group creation rate. Helped evaluate advertisement platform offerings and rolled out the "Boom Network" advertising network. Eons raised $32 million from General Catalyst Partners, Charles River Ventures, Sequoia Capital, and Intel Capital.

BBN Technologies

- 2007 - 2008: Delta Division, Vice President of Technology. Grew "Boomerang" counter-sniper project engineering team and significantly de-risked $10 million worth of product deliveries. Identified sales team leading to $10 million dollars of new sales within 6 months and $100+ million in additional orders in the following two years. Boomerang was a significant asset leading to the acquisition of BBN by Raytheon in 2009. Created new business plan and grew team; launched new fully-featured social networking web site in 5 months. Served as lead expert witness in patent infringement lawsuit, resulting in $58 million jury award to client; verdict for patents I testified on were upheld on appeal which resulted in a $120 million settlement.
- 2004 - 2007: Delta Division, Director of Technology, responsible for commercializing IP and creating new businesses. Hired and grew division's initial engineering team. Wrote three business plans, two of which were funded. For call center business plan, acted as general manager, hiring and managing engineering team, inside sales team, and identifying and recruiting a new general manager. Identified and recruited other key employees to Delta Division, including senior members of team leading to successful internal sales growth and spin-outs of projects. Contracted by BBN to BBN spin-out PodZinger as VP of Operations and Technology. Additional areas studied include mesh networking, network security, satellite communications, and semantic web.

PodZinger Inc. (BBN spin-out, also known as EveryZing and now RAMP)

Amazon EX1002 Appendix A Page 2

- 2006: Vice President of Operations and Technology. Significantly upgraded capability of consumer-facing search site and redeployed web site from company to co-location facility. Identified key portions of infrastructure for upgrading and cost reduction. Helped write business plan, evaluating advertisement models of revenue. Hired in operations replacement and phased back to BBN.

Commonwealth Capital

- 2004: Entrepreneur-in-residence (informally), performed technical due diligence on business plans, brainstormed ideas for new businesses with venture partner. With venture partner, left for portfolio company BBN to form core of commercialization team.

Empirix, Inc./Teradyne, Inc.

- 2001 - 2004: Chief Technologist, Engineering Manager, Web Application Test Group. Researched potential new product areas; developed product plan and prototype. Responsible for three new and existing products. Managed off-shore development team. Chief architect for all web testing products. Re-architected core testing product, helped write javascript interpreter. Provided technical vision for core product.
- 2000 - 2001: Chief Technologist, Communication Infrastructure Test Group. Responsible for incorporating new technology internally, tracking new technologies, technical evaluation of partnerships and potential acquisitions. Helped develop division strategy. Developed plans which formed core capabilities for successful new products introduced in 2004-5.
- 2000 - 2001: Engineering Manager, Communications Infrastructure Test Group. Execution of new product plan developed in prior role. Grew team from four existing engineers to team of over 30 on immediate team and over 40 on project. Delivered new platform in one year. Platform and derivatives accounted for large portion of booked products for the division within 2 years and is currently (2008) a key portion of new product offerings.
- 2000: Empirix was formed as a spin-out of Teradyne in January 2000. Reported to CEO in carve-out of Empirix from Teradyne.
- 1999: (Teradyne) Director of Business Development, Software Test Units. Reported directly to Chairman of the Board/Founder and then to general manager of software test unit (6 divisions of Teradyne). Evaluated and researched acquisition and partnership candidates. Internally assessed technology position in market and gaps in product lines. Worked with senior division staff to develop new product strategies. Attended internal Teradyne divisional board meetings. Chairman served as my mentor.

3Com Corporation/NBX Corporation

- 1999: Software Engineer 5. Continued work after 3Com acquisition of NBX. Built cross-division relationships for new products and research directions. NBX was acquired by 3Com in March 1999.
- 1997 - 1999: (NBX) Senior Scientist and Engineer. Work in IP Telephony. Architected next-generation product. Protocol design and validation for core protocol now used tens of millions of times daily. Led team in integration of IP protocols into current product. Designed audio reconstruction algorithms. Developed applications for bug analysis and diagnosis of system problems. Implementation of network simulator. Work on collaborative projects with external

partners. Worked to identify gaps in product. Representative at numerous trade shows. Innovated novel methods of using product.

MIT Department of Electrical Engineering and Computer Science, Cambridge, MA

- 1991 - 1998: Research assistant, Telemedia Network Systems Group. Design, development, and implementation of Gigabit ATM network for distributed multimedia system. Studied host interface design issues. ATM network simulation.
- Spring 1989, Fall 1990, Spring 1995: Teaching assistant, Computation Structures digital systems course. (Spring 1995 Head TA)
- 1988 - 1989: Head laboratory teaching assistant for Computation Structures. Responsibilities included writing and revising lab assignments, and maintaining the lab.
- 1987: Laboratory teaching assistant for Computation Structures.
- 1987: Design, construction, and programming of 16-bit computer.

Agora Technology Group, Incorporated

- 1994 - 1996: Founder and CEO. Conceived and oversaw development of targeted advertising-supported Web sites. Responsible for company's vision and direction. Attended the first two WWW Conferences, presenting a workshop and paper at the first, and appearing on the "Commercialization and Economics of the Web" panel and chairing the "Where Commercial Services and the Web Are Headed" panel at the second. Sold company intact; is currently an operating stand-alone company.

AT&T Bell Laboratories, Holmdel, NJ

- 1989 - 1990: Implementation of cascadable all-optical fiber logic gate. Modelocking of all-fiber erbium laser. Construction of modelocked laser repetition rate booster. Strong optics laboratory and fiber optic experience.
- Summer 1988, 1987: Research in integrated optics. Analysis of rectangular waveguides using microwave modeling. Fabrication of integrated optical components.

Honors

- MIT Alumni Association Great Dome Award, 2010, Baker 60th Anniversary Reunion Co-Chair (highest group award given by MIT Alumni Association)
- MIT Alumni Association Presidential Citation Award (now known as Great Dome), 2008, Member of MIT Chairman's Salon committee
- MIT Alumni Association Bronze Beaver Award, 2007 (highest individual award given by MIT Alumni Association)
- MIT Alumni Association Volunteer Honor Roll, February 2004
- MIT Alumni Association Lobdell Award, 1999
- Boston Museum of Science Gold Pin for 1000 hours of Volunteer Service, April 1999
- MIT Alumni Association Presidential Citation Award (now known as Great Dome), 1997, Member of Alumni Online Communications Committee

Patents and Patent Publications

- U.S. Patent #9,697,231, H. Houh, J. Stern, R. Spina, M. Meteer, "Methods and apparatus for providing virtual media channels based on media search," July 4, 2017.
- U.S. Patent #9,697,230, H. Houh, J. Stern, "Methods and apparatus for dynamic presentation of advertising, factual, and informational content using enhanced metadata in search-driven media applications," July 4, 2017. See also WO2007056485.
- U.S. Patent #7,975,296, L. Apfelbaum, H. Houh, T. Mayberry and G. Friedman, "Automated security threat testing of web pages," July 5, 2011. See also US20030159063, WO2003067405.
- U.S. Patent #7,877,736, H. Houh and J. N. Stern, "Computer language interpretation and optimization for server testing," January 25, 2011. See also US20050138104, WO2005043300.
- U.S. Patent #7,801,910, H. Houh and J. N. Stern, "Method and apparatus for timed tagging of media content," September 21, 2010. See also US20070112837, US20090222442, WO2007056535.
- U.S. Patent #7,590,542, D. C. Williams, W. C. Hand, H. Houh, A. R. Seeley, "Method of Generating Test Scripts Using a Voice-Capable Markup Language," September 15, 2009. See also EP1530869, US20030212561, WO2003096663.
- U.S. Patent #6,967,963, H. H. Houh, P. Anderson, C. Gadda, "Telecommunication method for ensuring on-time delivery of packets containing time-sensitive data," November 22, 2005. See also EP1060400, WO2000033092, CA2318774.
- U.S. Patent #5,144,375, M. C. Gabriel, H. H. Houh, N. A. Whitaker, "Sagnac Optical Logic Gate," September 1, 1992. Also issued as European Patent # EP0456422, July 23, 1997, German Patent #DE69126913, August 28, 1997
- U.S. Patent Application Publication No. 20020015387, "Voice Traffic Packet Capture and Analysis Tool for a Data Network" (Abandoned in 2007)
- U.S. Patent Application Publication No. 20020016708, "Method and Apparatus for Utilizing a Network Processor as Part of a Test System" (Abandoned in 2007)
- U.S. Patent Application Publication No. 20020016937, "Method and Apparatus for Utilizing a Network Processor as Part of a Test System." (Abandoned in 2007) See also WO2002011413.
- U.S. Patent Application Publication No. 20070106646, "User-directed navigation of multimedia search results" (Abandoned in 2009)
- U.S. Patent Application Publication No. 20070106660, "Method and apparatus for using confidence scores of enhanced metadata in search-driven media applications" (Abandoned in 2009)
- U.S. Patent Application Publication No. 20070106685, "Method and apparatus for updating speech recognition databases and reindexing audio and video content using the same" (Abandoned in 2015)
- U.S. Patent Application Publication No. 20070118873, "Methods and apparatus for merging media content" (Abandoned in 2015)
- U.S. Patent Application Publication No. 20090222442, "User-directed navigation of multimedia search results" (Abandoned in 2016)
- U.S. Patent Application 11/395,732, "Search snippet creation for audio and video data" (Abandoned in 2009)
- U.S. Patent Application 11/774,931, "Methods and apparatus for managing a social networking web site"
- U.S. Patent Application 11/774,947, "Methods and apparatus for organizing media files"
- U.S. Patent Application 11/774,956, "Methods and apparatus for managing an online event"

- U.S. Provisional Patent Application 61/086,909, “Measuring and ranking relationship activity”
- U.S. Provisional Patent Application 61/086,914, “Detecting media object commonality”
- U.S. Provisional Patent Application 61/086,904, “Message categorization based on message characteristics”
- U.S. Provisional Patent Application 61/086,905, “Photo tagging to request action”

Publications

- “IP switching: server driven flow classification,” H. H. Houh, *Proceedings of the Washington University Workshop on Integration of IP and ATM* , November 1996.
- “Aurora at MIT,” D. D. Clark, H. H. Houh, and D. L. Tennenhouse, Editors, *MIT Laboratory for Computer Science Technical Report 673*, December 1995.
- “ViewStation Applications: Implications for Network Traffic,” C. J. Lindblad, D. Wetherall, W. Stasior, J. F. Adam, H. H. Houh, M. Ismert, D. Bacher, B. Phillips, and D. L. Tennenhouse, *IEEE Journal of Selected Areas in Communications*, 1995.
- “The VuNet Desk Area Network: Architecture, Implementation, and Experience,” H. H. Houh, J. F. Adam, M. Ismert, C. J. Lindblad, and D. L. Tennenhouse, *IEEE Journal of Selected Areas in Communications*, 13 (4), May, 1995.
- “Reducing the Complexity of ATM Host Interfaces,” H. H. Houh and D. L. Tennenhouse, Hot Interconnects II Symposium Proceedings, Stanford, August 11-12, 1994.
- “Media-intensive data communications in a `desk-area' network,” J. F. Adam, H. H. Houh, M. Ismert, and D. L. Tennenhouse, IEEE Communications, August 1994.
- “ViewStation Applications: Intelligent Video Processing Over A Broadband Local Area Network,” C. J. Lindblad, D. J. Wetherall, W. Stasior, B. Phillips, D. Bacher, J. Adam, H. Houh, M. Ismert, and D. L. Tennenhouse, Proceedings of the 1994 USENIX Symposium on High-Speed Networking, Oakland, CA, August 1994.
- “The Media Gateway: Live Video on the World Wide Web,” H. H. Houh, C. J. Lindblad, J. Soo, D. L. Tennenhouse, and D. J. Wetherall, Workshop at the 1994 World Wide Web Conference, Geneva, Switzerland, May 1994.
- “Active Pages: Intelligent Nodes on the World Wide Web ,” H. H. Houh, C. J. Lindblad, and D. J. Wetherall, Proceedings of the 1994 World Wide Web Conference, Geneva, Switzerland, May 1994.
- “Wavelength Division vs. Code Division Access Methods for Optical Networks,” H. H. Houh, Area Exam Paper, May 1993.
- “Experience with the VuNet: A Network Architecture for a Distributed Multimedia System,” J. F. Adam, H. H. Houh, D. L. Tennenhouse, Proceedings of the 18th Conference on Local Computer Networks, pp. 70-76, September 1993
- “The VudBoard: A Simple DMA Interface,” H. H. Houh, Proceedings of the 4th Gigabit Minijam, January 1994.
- “A Software-Oriented Approach to the Design of Media Processing Environments,” D. L. Tennenhouse, J. Adam, D. Carver, H. Houh, M. Ismert, C. Lindblad, W. Stasior, D. Weatherall, D. Bacher, and T. Chang., submitted to the International Conference on Multimedia Computing and Systems, May 1994.
- “A Network Architecture for Distributed Multimedia Systems,” J. F. Adam, H. H. Houh, M. Ismert, and D. L. Tennenhouse, submitted to the International Conference on Multimedia Computing and Systems, May 1994.

- "The Viewstation Collected Papers," D. L. Tennenhouse, J. Adam, C. Compton, A. Duda, D. Gifford, H. Houh, M. Ismert, C. Lindblad, W. Stasior, R. Weiss, D. Wetherall, D. Bacher, D. Carver, and T. Chang, MIT Laboratory for Computer Science Technical Report, MIT/LCS/TR-590, November 1993.
- "The Viewstation Collected Papers II," W. F. Stasior, D. L. Tennenhouse, J. F. Adam, D. R. Bacher, H. H. Houh, M. Ismert, C. J. Lindblad, B. M. Phillips, D. J. Wetherall, MIT Laboratory for Computer Science Technical Report, MIT/LCS/TR-696, April 1996.
- "A System's Perspective of the Sagnac Fiber Logic Gates and Their Possible Applications," A. Huang, N. Whitaker, C. Gabriel, H. Avramopoulos, P. M. W. French, H. H. Houh, and I. Chuang, Applied Optics, September 10, 1994
- "Complete Switching in a Three-Terminal Sagnac Switch," H. Avramopoulos, P. M. W. French, M. C. Gabriel, H. H. Houh, N. A. Whitaker, T. Morse, IEEE Phot. Tech. Lett. **3** (3), 235
- "Complete Switching in a Three-Terminal Sagnac Switch," H. Avramopoulos, P. M. W. French, M. C. Gabriel, H. H. Houh, N. A. Whitaker, IEEE/LEOS Annual Meeting, Paper PDP-13, November 1990
- "All-optical phase-locked oscillator," N. A.Whitaker, Jr., H. H. Houh, H. Avramopoulos, T. F. Morse, IEEE/LEOS Annual Meeting, Paper ELT2.4/MOO3, November 1990
- "Passive modelocking of an all-fiber erbium laser," H. Avramopoulos, H. H. Houh, N. A. Whitaker, M. C. Gabriel, T. F. Morse, IEEE/LEOS Conference on Optical Amplifiers and their Applications, Paper PDP-8, August 1990
- "Transverse modes, waveguide dispersion, and 30ps recovery in submicron GaAs/AlAs microresonators," J. L. Jewell, S. L. McCall, A. Scherer, H. H. Houh, N. A. Whitaker, A. C. Gossard, and J. H. English, Appl. Phys. Lett. **55** (1), July 3, 1989

Leadership, Activities and Interests

- Leadership
  - Reading Symphony Orchestra
    - Board of Directors (Treasurer), 2022 – present
  - Discovery Museum (Acton, MA)
    - Board of Directors, 2023 – present
    - Board of Directors, Development Committee, 2023 – present
    - Science and Technology Advisory Council, 2012 – present
    - The Discover Museum received the 2024 National Medal for Museum and Library Service, the nation's highest honor given to museums and libraries that make significant and exceptional contributions to their communities.
  - Tau Beta Pi National Engineering Honor Society
    - Executive Councilor and Secretary, 2024 – 2025, Treasurer 2026
    - Advisor, MA B Chapter at MIT, 2003 – present
    - District Director (National Officer), Tau Beta Pi, New England Area, 1991 – 2003
    - President, MA B Chapter at MIT, Fall 1988 – Spring 1989
    - Laureate award, 1989
  - MIT Alumni Association Board of Directors
    - MIT Annual Giving Board, 3 year term, 2024 – 2027
    - National Selection Committee, 3 year term, 2020 – 2023
    - K-12 STEM Initiatives Co-chair, 2013 – 2017
    - Awards Committee Chair, 2012 – 2014

        - Awards Committee, 3 year term, 2011 – 2014
        - Vice President, 2 year term, 2004 – 2006
        - Alumni Ad-Hoc Committee, Chair, 2005-2006
        - Board Member, 2 year term, 1997 – 1999
    - MIT Club of Boston
        - Board of Directors, 2006 – 2011
        - K-12 Initiatives Chair, 2009 – 2012
        - VP of Communications, MIT Club of Boston, 2003 – 2006
        - Past-President, MIT Club of Boston, 2002 – 2003
        - President, MIT Club of Boston, 2001 – 2002
        - President-Elect, MIT Club of Boston, 2000 – 2001
        - VP of Programs, MIT Club of Boston, 1999 – 2000
        - Activities Super-Chair, MIT Club of Boston, 1998 – 1999
    - MIT Enterprise Forum of Cambridge, Inc.
        - Past Chair, 2009 – 2011
        - In-NOW-vation Co-chair, 2010
        - Chair, 2007 – 2009
        - Vice Chair, 2005 – 2007
        - Executive Board Member, 2002 – 2011
        - Winter Workshop Co-Chair, February 2007 – conceived idea for conference, which sold-out and produced largest attendance numbers in recent memory
        - Spring Workshop Co-Chair, Spring 2004
        - Membership Committee Chair, Fall 2003 - 2006
        - 25th Anniversary Dinner Chair, Fall 2003
        - As Membership Chair and Board Member, started Special Interest Groups in 2004; a SIG won the MIT Presidential Citation award, the MIT Alumni Association's highest award for organizations, in 2006
    - Estabrook Elementary School PTA
        - Advisory committee to the superintendent on PCB issue, 2010-2011
        - 4th Grade after-school science program co-organizer, 2010-2012
        - 4th and 5th Grade before-school Math Olympiad co-organizer, 2009-2013
        - 5th Grade BBQ and Yearbook Committee, 2011, 2013
        - Family Math Night volunteer, 2008-2012
    - MIT Class of 1989
        - Class Agent, 2024 – present
        - 35-year Reunion Gift Committee, Co-chair, 2024
        - Secretary, five consecutive 5 year terms, 1989 – 2014
        - 30-year Reunion Committee, 2019
        - 25-year Reunion Committee, 2014
        - 20-year Reunion Committee and Gift Committee, 2009
        - 15-year Reunion Committee and Gift Committee, 2004
        - 10-year Reunion Committee and Gift Committee, 1999
        - 5-year Reunion Committee, 1994
        - Interim Treasurer, 1993 – 1994
        - Instituted annual senior class career fair, now raising over $100,000 annually for senior class activities, Fall 1988
    - Strong, consistent record of leadership dating to high school
- Acting

Amazon EX1002 Appendix A Page 8

  - ‘21’ (Sony Pictures), credited as “Chinatown Dealer,” 2007, Kevin Spacey's movie about the MIT Blackjack Team inspired by “Bringing Down the House” by Ben Mezrich, opened nationwide on March 28, 2008. 21 was the number one movie in US for two weeks and number one globally for one week. 21 also topped the DVD sales, Blu-ray sales and DVD rental charts.
  - Spring Lake Theater Company, first New York-area off-broadway production of “A Chorus Line,” played role of Mark, Summer 1990
- Former member of the MIT Blackjack team
- Producer for 10,000 Maniacs’ 2013 album “Music from the Motion Picture”
- Executive Producer for 10,000 Maniacs’ 2015 album “Twice Told Tales”
- Producer for 10,000 Maniacs’ 2016 DVD “10,000 Maniacs Live”
- Executive Producer for OK Go’s 2025 album “And The Adjacent Possible”
- 2024-25 production of “Romeo & Juliet” at the Circle in the Square Theater, nominated for 2025 Tony for Best Revival of a Play, Investor
- “The Interestings,” a new musical in progress by Sara Bareilles and Sarah Ruhl, Investor
- Violist, violinist, cellist, harpist, guitarist, singer, actor: played in many amateur/semi-professional groups including Firebird Pops Orchestra, Boston Civic Symphony Orchestra, Reading Symphony Orchestra, MIT Summer Philharmonic Orchestra, Carlisle Chamber Orchestra, Horizon Ensemble, Rivers Symphony Orchestra, Opera51 Orchestra, American Music Abroad Summer 2025 Tour, Horizon Ensemble, New England Film Orchestra, Freisinger Chamber Orchestra, Merrimack Valley Philharmonic, Longwood Symphony, MIT Symphony, and Somerville Community Chorus
- Violist, “Coming Home” album by Karen Phillips, 2014
- Google Scholar citations: 2224, h-index: 22, i10-index: 27 (since 2019: 539/12/12)
- Erdos-Bacon number of 7
- Erdos-Bacon-Sabbath number of 11