

(43) International Publication Date
27 June 2013 (27.06.2013)

W I P O | P C T

(10) International Publication Number
WO 2013/096023 A1

- (51) **International Patent Classification:**
G06F 3/01 (2006.0 1) *G06F 3/16* (2006.0 1)
- (21) **International Application Number:**
PCT/US20 12/069030
- (22) **International Filing Date:**
12 December 2012 (12.12.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
13/332,348 20 December 2011 (20.12.2011) US
- (71) **Applicant (for all designated States except US):** **MICROSOFT CORPORATION** [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).
- (72) **Inventors:** **TAN, Desney S.**; Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **PATEL, Shwetak**; Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **MORRIS, Daniel S.**; Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US). **GUPTA, Sidhant**; Microsoft Corporation, LCA - International Patents, One Microsoft Way, Redmond, Washington 98052-6399 (US).
- (81) **Designated States (unless otherwise indicated, for every kind of national protection available):** AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

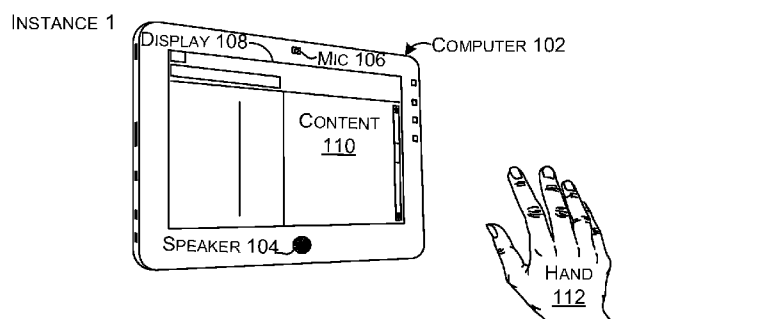
- (84) **Designated States (unless otherwise indicated, for every kind of regional protection available):** ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.1 7(H))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.1 7(in))

Published:

- with international search report (Art. 21(3))

(54) **Title:** USER CONTROL GESTURE DETECTION

(57) **Abstract:** The description relates to user control gestures. One example allows a speaker and a microphone to perform a first functionality. The example simultaneously utilizes the speaker and the microphone to perform a second functionality. The second functionality comprises capturing sound signals that originated from the speaker with the microphone and detecting Doppler shift in the sound signals. It correlates the Doppler shift with a user control gesture performed proximate to the computer and maps the user control gesture to a control function.



WO 2013/096023 A1

User Control Gesture Detection

BACKGROUND

[0001] Computers are ubiquitous in everyday life. However, instead of having 'a computer', users tend to have, or interact with, lots of different computers. For instance, a
5 typical user may have a desktop computer at work, a laptop at home, a Smartphone, a pad computer, a video game console, and/or an electronic book, among others. Further, the user tends to interact with other computers that are integrated into other devices. For example, the user's car likely has a computer that the user interacts with to get directions, listen to music, etc. Various user controls have been developed for these different types of
10 computers. For instance, a mouse and a physical keyboard tend to be the most common way of interacting with and controlling a desktop computer. Mice were not convenient with notebook computers which tend to use a trackpad-type device instead. Neither mice nor trackpads were feasible with Smartphones and pad type computers which instead tend to have touch sensitive screens (e.g. virtual keyboards) which the user can contact to
15 control the device.

[0002] Each of these control technologies has disadvantages. For instance, controlling devices by contacting the touch screen tends to make smudges all over the screen and/or is rendered inoperable if the user is wearing gloves. Potentially more importantly is that these control technologies do not lend themselves to use across all of
20 the device types. As a result the user has to learn different techniques for controlling his/her different devices. For example, assume that in the course of a day the user listens to music on each of his/her computing devices and wants to turn off the music. When listening on the desktop the user likely uses the mouse to turn a music playing application from 'play' to 'stop'. The user may use the notebook's trackpad in a similar manner. In the
25 car, the user may use a specific hardware switch to stop the music. On the Smartphone of the pad-type computer, the user may touch the 'stop' icon on the touch screen. This lack of continuity is inconvenient. Further, this inconvenience is magnified when the user tries to control a computer with which the user is unfamiliar. The present concepts relate to techniques that can be utilized to control a computer and can be readily applied to any (or
30 all) computers, such as those listed above.

SUMMARY

[0003] The described implementations relate to user control gestures. One example allows a speaker and a microphone to perform a first functionality. The example

simultaneously utilizes the speaker and the microphone to perform a second functionality. The second functionality comprises capturing sound signals that originated from the speaker with the microphone and detecting Doppler shift in the sound signals. It correlates the Doppler shift with a user control gesture performed proximate to the computer and maps the user control gesture to a control function.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] The accompanying drawings illustrate implementations of the concepts conveyed in the present application. Features of the illustrated implementations can be more readily understood by reference to the following description taken in conjunction with the accompanying drawings. Like reference numbers in the various drawings are used wherever feasible to indicate like elements. Further, the left-most numeral of each reference number conveys the Figure and associated discussion where the reference number is first introduced.

[0005] FIGS. 1-7 show examples of Doppler-based user control gesture detection scenarios in accordance with some implementations of the present concepts.

[0006] FIG. 8 shows an example system upon which Doppler-based user control gesture detection techniques can be employed in accordance with some implementations of the present concepts.

[0007] FIG. 9 is a flowchart of an example of a Doppler-based user control gesture detection method or technique in accordance with some implementations of the present concepts.

DETAILED DESCRIPTION

OVERVIEW

[0008] This patent relates to utilizing user control gestures (hereinafter, "user gestures") to control computers. More specifically, the user gestures can be detected by the computer by emitting an audio signal and detecting the emitted signal. Any differences between the emitted signal and the detected signal can be analyzed to detect the user gesture based upon the Doppler effect or Doppler shift. Individual user gestures can be identified from the analyzed audio signal data. Individual user gestures can be mapped to specific control functionalities of the computer. The specific control functionality can be implemented to control the computer.

USE CASE EXAMPLES

[0009] For purposes of explanation consider introductory FIG. 1, which shows a computer 102 manifest as a pad-type computer. Computer 102 includes a speaker 104, a

microphone 106, and a display 108. In this example, content 110, such as text, is displayed on a portion of the display 108.

[00010] FIG. 1 is divided into instance 1 through instance 3 to help illustrate how a user can control the computer by making user gestures. During instances 1-3, computer 102 can generate an audio signal via speaker 104. The audio signal can be captured by microphone 106. The captured signal can be compared to the emitted signal for differences. These differences can be produced by the user gesture effect on the signal (e.g., its Doppler effect). Different user gestures create different Doppler effects and thus individual user gestures can be distinguished and identified.

10 [00011] In the illustrated case, the user gesture is a hand gesture made with the user's right hand 112. (In some implementations, the gesture could alternatively be performed with the left hand, which is not illustrated here to avoid clutter on the drawing page). In this example the user starts with the hand 112 generally parallel to the display 108 as illustrated in instance 1. The user then makes the user gesture (arrow 114) by tilting the hand downward as illustrated in instance 2. The user gesture 114 is detected by the computer 102 as a difference in the received signal from the emitted signal. State another way, the user gesture can cause changes to a frequency profile of the captured signal from what would otherwise be expected in the absence of the user gesture. This difference or change can be caused by the user's hand creating a Doppler shift on the received or captured signal. In this example, the computer maps the detected gesture to a 'scroll down command'. The computer then scrolls down the content 110 (arrow 116).

[00012] Instance 3 shows the user's hand 112 at the end of the gesture now generally perpendicular to the display 108. Also, the content 110 has been scrolled downward when compared to instance 1. Note also, that in instances 1-3 the user did not have to physically contact the computer to accomplish the control functionality.

25 [00013] As mentioned above, computer 102 includes a speaker 104 and a microphone 106. The speaker and microphone could be dedicated hardware devices that are utilized only for the user gesture detection functionality. However, such need not be the case. In fact most computers already include at least one microphone and at least one speaker. These existing hardware devices can be utilized to reduce costs and complexity associated with adding additional hardware devices. Further, the present gesture detection functionality can be achieved with the existing speaker and microphone without interrupting or diminishing existing uses of these devices. For instance, assume that in instance 1, the user is engaged in an ongoing telephone call (audio only) or web-chat

(audio and video (while not shown for sake of brevity, computer 102 could include a camera)). Thus, during the telephone call audio signals are being played by speaker 104 from the other party and audio signals, such as the user speaking, are being captured by microphone 106. Simultaneously, the speaker can be emitting one or more audio signals
5 for use in the gesture detection functionality. In some implementations, these audio signals can be a steady tone emitted at a frequency that the user cannot hear. For instance, most speakers can emit signals up to 22k-24k Hertz and most microphones can capture these signals. However, humans generally cannot hear anything above about 18k-20k Hertz. Thus, speaker 104 and microphone 106 can be simultaneously used for the
10 telephone call and for user gesture detection, yet the user only hears a normal telephone call. When the user is not making a gesture the emitted gesture detection signal and the captured gesture detection signal are generally identical. For instance, a steady 20k Hertz signal is emitted and received.

[00014] While the user is engaged in the telephone call the user can be looking at
15 content 110. The user may want to see the content above content 110 and can make the user gesture 114 to scroll down. This movement changes the audio signal so that the captured signal is slightly different from the emitted gesture detection signal. This difference can be analyzed to determine the user's gesture (e.g., to distinguish between different gestures). The difference between the emitted gesture detection signal and the
20 detected gesture detection signal diminishes once the gesture is completed. Despite the dual functionality performed by the speaker and the microphone, the user's telephone call progresses in a normal fashion (e.g., the user does not even need to be aware of the underlying gesture detection functionality that is simultaneously leveraging the speaker 104 and the microphone 106). In other examples, the user gesture may be utilized to
25 effect and/or operate cooperatively with the first function. For example, the user gesture could be used to mute the telephone call or control the volume of the telephone call, among others.

[00015] Further, this user gesture detection implementation is not dependent on the condition of the user's hand. For instance, the user may have gloves on or the user's hand
30 may be dirty or oily. Such conditions may hinder the ability of a touch-screen to detect user touches. However, the present implementations do not require user contact and can be unaffected by gloves, dirt, etc.

[00016] Further, the present user gesture recognition implementation can be readily applied to each of the user's computers so that the user can use the same user gestures to

control his/her various computers. For instance, the user could use the illustrated user gesture to control his/her desktop computer, in-dash car computer, smartphone, and notebook computer (among others) in the same way that the user controlled pad type computer 102 in FIG. 1.

5 [00017] FIGS. 2-7 collectively show another Doppler-based user gesture detection implementation that relates to a computer 202. In this case, computer 202 is manifest as a notebook-type computer. Computer 202 includes two speakers 204(1) and 204(2), two microphones 206(1) and 206(2), a display 208, and a user input device manifest as a keyboard/trackpad 210. In this example, a music player application 212 is running on the
10 computer. As shown in FIG. 2, the music player is in a "stop" condition (e.g., not playing) as indicated at 214. Assume that a Doppler-based user gesture detection feature is operating on the computer. For instance, the Doppler-based user gesture detection feature could be causing one or more signals to be emitted from one or both of the speakers 204(1) and 204(2) and for the signals to be captured by one or both of microphones 206(1)
15 and 206(2). As described above relative to FIG. 1, the emitted and captured signals can be analyzed for differences and the differences can be identified as user gestures for controlling the computer.

[00018] FIG. 3 shows a user engaging the keyboard/trackpad 210 with his/her left and right hands 302 and 304, respectively. In this implementation, engaging the
20 keyboard/trackpad 210 inactivates the Doppler-based user gesture detection feature to avoid interpreting typing movements as user gestures. For instance, the Doppler-based user gesture detection feature may stop causing signals to be emitted from the speaker(s) or it may simply not analyze the captured signals when the user is engaging the keyboard/trackpad.

25 [00019] In other examples, the Doppler-based user gesture detection feature may operate cooperatively with other input mechanisms to provide control functionalities that are otherwise unavailable to the user or to make such control functionalities more user-friendly. For instance, in an un-illustrated example, in a 3-D application, the user may select an object with the trackpad. The user may then be able to 'drag' the selected object
30 with the trackpad. Further, the user may be able to 'turn' or 'rotate' the selected object by performing the side-to-side control gesture described and illustrated above relative to FIG. 1. Of course this is just one example, and many other 'combination' control options can be mapped by the user to control the computer as desired with traditional input mechanisms in combination with Doppler-based user gesture detection.

[00020] Continuing with the illustrated example, FIG. 4 shows a subsequent time where the user has disengaged from the keyboard/trackpad 210 (e.g., is no longer physically making key inputs or trackpad inputs). As such, the Doppler-based user gesture detection feature is once again activated. FIG. 5 shows the user moving his/her left hand 5 302 forward (toward the display 208) as indicated by arrow 502 and simultaneously moving his/her right hand 304 backward (away from the display 208) as indicated by arrow 504. This user movement effects the sound captured by speakers 204(1) and/or 204(2). This effect can be detected and identified by the Doppler-based user gesture detection feature as a user gesture. In this example, the user gesture is mapped to 10 changing the state of the music player or music player application 212. As such, the Doppler-based user gesture detection feature causes the music player to go from the stop condition 214 of FIG. 4 to a "play" condition 506 of FIG. 5. As a result musical sounds are emitted from the speakers as indicated at 508.

[00021] FIGS. 6-7 show a scenario where the user can control the computer 202 15 with a user gesture even though the user is not in front of the computer and is not in line of sight of the screen. In FIG. 6 as the user approaches computer 202, the computer is still outputting musical sounds as indicated at 508. The user can change the music playing status of the computer by repeating the user gesture discussed relative to FIGS. 4-5 that entails simultaneously moving one hand toward the computer and one hand away from the 20 computer. In FIGS. 4-5 the user gesture caused the computer's music playing application to switch from a stopped state to a playing state. In FIG. 6, the computer is still in the playing state and repeating the user gesture switches the computer back to the stopped state. This is evidenced by the absence of musical notes emitted from the computer in FIG. 7. The user gesture can be recognized by the computer, and used to control the 25 computer, without the user being in front of the computer or in line-of-sight of certain devices on the computer, such as camera devices. In fact, in some implementations, the user can control the computer from any direction or dimension as long as the user is within a defined range extending radially from the computer.

[00022] To summarize the discussion relative to FIGS. 1-7, the Doppler-based user 30 gesture detection implementations detect motion around a computer and use properties of the detected motion, such as speed, direction, and/or amplitude, to recognize a rich set of user gestures. For instance, by tracking the direction and speed, a user gesture of moving the hand up or down can scroll a webpage in real-time. Doppler-based user gesture detection implementations can also detect two hands moving in opposite directions,

among other detectable user gestures. The concepts can be implemented on any type of computer that has a microphone and a speaker. The microphone and the speaker need not be dedicated to gesture detection and can simultaneously perform other functions. The described implementations are robust and can perform satisfactorily in both busy and
5 noisy environments and quieter environments. The concepts can be implemented across multiple different types of devices so that the user can control any of the devices with the same user gestures.

SYSTEM EXAMPLES

[00023] FIG. 8 shows an example of a system 800 that can accomplish the Doppler-based user gesture detection concepts introduced above. System 800 includes a computer
10 802, that in this case is manifested as a smartphone. In this example, the computer 802 includes a speaker 804, a microphone or mic 806, a Doppler-based user gesture identification component 808, an operating system 810, a processor 812, and storage 814. Doppler-based user gesture identification component 808 can be an application, an
15 application part, or a part of the operating system 810.

[00024] Computer 802 can be configured to communicate with other computers and/or computing resources, such as computing resources located in the cloud 816 via network 818.

[00025] The term "computer" or "computing device" as used herein can mean any
20 type of device that has some amount of processing capability and/or storage capability. Examples of computers can include traditional computing devices, such as personal computers, cell phones, smart phones, personal digital assistants, cameras, video cameras, or any of a myriad of ever-evolving or yet to be developed types of computing devices. Computers can also be components of a larger device. For instance, a dashboard computer
25 is included in many newer cars.

[00026] Processing capability can be provided by one or more processors (such as processor 812) that can execute data in the form of computer-readable instructions to provide a functionality. Data, such as computer-readable instructions, can be stored on storage, such as storage 814. The storage can be internal and/or external to the computing
30 device. For instance, memory card 820 can be thought of as storage that is removably received by the computer. The storage can include any one or more of volatile or non-volatile memory, hard drives, flash storage devices, and/or optical storage devices (e.g., CDs, DVDs etc.), among others. As used herein, the term "computer-readable media" can include transitory and non-transitory instructions. In contrast, the term "computer-

readable storage media" excludes transitory instances. Computer-readable storage media can include "computer-readable storage devices". Examples of computer-readable storage devices include volatile storage media, such as RAM, and non-volatile storage media, such as hard drives, optical discs, and flash memory, among others.

5 [00027] Doppler-based user gesture identification component 808 can include and/or access a signal control module 822, a signal analysis module 824, a user interaction module 826, and a datatable 828.

[00028] Briefly, the signal control module 822 can control what signals are emitted for use by the Doppler-based user gesture identification component 808. The signal
10 control module 822 can also control capturing the signals. The signal analysis module 824 can analyze the emitted and captured signals to identify user gesture induced Doppler shifts. The user interaction module 826 can provide interfaces, such as GUIs that explain to a user how to properly make pre-installed gestures and/or can allow the user to define individual gestures and to map the gestures to individual computer control commands.
15 These mappings can be stored in the datatable 828.

[00029] Recall that Doppler-based user gesture identification component 808 can use the phenomenon known as the Doppler shift to sense motion. The motion can be detected as the shift in frequency of a sound wave in response to a moving object, an effect called the Doppler effect. This frequency shift is proportional to the velocity with
20 which the object moves and the source frequency. In the present implementations, the original source (the speakers) and listener (the microphone) are stationary, thus in absence of any motion, there is no frequency change. When a user moves his/her hand, however, it reflects the sound waves, causing a shift in frequency. This frequency is measured by the microphone (f_r) and can be described by the following equation, which can be employed
25 by the signal analysis module 824.

$$f_r = f_t \cdot \left(\frac{c + v}{c - v} \right)$$

where, f_r = *perceived frequency at microphone*;
 f_t = *original frequency from speaker*;
 c = *speed of sound in air*;
 v = *velocity of target /hand*

[00030] Signal control module 822 can generate a continuous pilot tone (or multiple pilot tones), played through the computer's speaker(s) 804. Generally high frequencies (typically in the range of 18-22 kHz on commodity audio systems) are employed so that the user does not hear the pilot tone. However, lower frequencies such as down to and
5 below 6 kHz can be employed. Additionally, the higher the frequency, the greater the shift for a given velocity, which makes it computationally easier to estimate motion at a given resolution. The upper bound is largely a function of most laptop and phone speaker systems only being capable of producing audio at up to 22 kHz. Note also, that in some implementations, the signal control module may be configured to use signals associated
10 with another functionality rather than generating a dedicated pilot tone. For instance, if the speakers are playing music for the user, frequency elements of the music can be utilized as the pilot tone. The emitted musical frequencies can then be compared to the captured frequencies to identify differences caused by user gestures.

[00031] Due to variations in hardware as well as filtering in speaker and
15 microphone systems, signal control module 822 can, in some implementations, perform an initial calibration to find one or more performant (and potentially optimal) tone frequency(s) (no user intervention is required). In one implementation, it performs a 500 millisecond frequency sweep, and keeps track of peak amplitude measurements as well as the number of candidate motion events detected (i.e., potential false positives). Signal
20 control module 822 can select the highest frequency at which minimum false events are detected and the peak is most isolated (i.e., the amplitude is at least 3 dB greater than next highest peak in the sweep range). Note that the frequency sweep can be repeated periodically and/or upon some other occurrence or according to various parameters. For instance, if at some point excessive noise (e.g., interference) is detected in the captured
25 signal the frequency sweep may be repeated to look for a 'cleaner' frequency. For example, the noise can be compared to a predefined threshold. If the noise level exceeds the predefined threshold then the testing can be repeated. This testing and retesting of potential pilot tones can allow the Doppler-based user gesture detection techniques to perform satisfactorily in different and/or changing environments.

[00032] Further, the signal control module 822 may simultaneously emit multiple
30 pilot tones, each at a different but steady frequency. The corresponding captured signals may all be analyzed for redundancy or to obtain more accurate results. Alternatively, one of the pilot tones can be characterized as 'better' or 'best' (e.g., cleaner) and may be selected for analysis.

[00033] With the high frequency pilot tone being emitted, any motion in proximity (around a one meter range depending on speed) of the computer 802 will cause Doppler shifted reflections to be picked up by the microphone, which is continuously sampled at 44,100 Hz, in one implementation.

5 [00034] In some implementations, the signal analysis module 824 can buffer the incoming time domain signal from the microphone 806 and compute Fast Fourier Transform (FFT) with 2048-point Hamming window vectors. This yields 1024-point magnitude vectors that are spread equally over the spectral width of 22,050 Hz. After each FFT vector is computed, it can be further processed by the signal analysis module
10 824. For instance, the signal analysis module can perform signal conditioning, bandwidth extraction, followed by motion detection, and finally feature extraction.

[00035] In some cases, signal conditioning can be thought of as utilizing testing data relating to how fast users can move their hands. Some implementations can operate on the assumption that the fastest speed at which users can move their hands in front of a
15 computer is about 3.9 m/sec. Hence, such implementations can conservatively bound signals of interest at 6 m/sec. Given the described sampling rate and FFT size, this yields about 33 frequency bins on either side of the emitted peak. (Some implementations may operate on a presumption that faster movement is created by a user employing some type of pointing device to make the user gesture. For instance, a user can hold and move a
20 ruler or pencil faster than the user can move his/her hand. This fast movement can be mapped to a separate control command).

[00036] The signal analysis module 824 can interpret shifted frequencies as user gestures. For instance, the signal analysis module can detect user gestures that effectively increase (or decrease) the bandwidth of the pilot tone (i.e., window averaging and spectral
25 leakage blur the movement of the peak). To detect this, signal analysis module 824 can compute the bandwidth of the pilot tone by scanning the frequency bins on both sides independently until the amplitude drops below 10% of pilot tone peak. Stated another way, the present techniques do not need to analyze the entire frequency spectrum and can instead focus the analysis around the pilot tone (for instance 200 HZ on either side).
30 Further, using a relative amplitude drop can allow the system to respond dynamically, such as when the user changes the volume of the speakers.

[00037] For most cases, this bin analysis is sufficient for inferring motion. However, if the shift is large enough, the reflected signal can separate from the pilot tone's peak rather than blurring the peak. To address this issue the signal analysis module 824

can perform a second scan, looking beyond the stopping point of the first scan. If a second peak with at least 30% energy of the primary tone is found, the first scan can be repeated looking for amplitude drops calculated from the second peak.

[00038] Motion Detection and Feature Extraction: In some implementations, the frequency vectors have a per-bin resolution of 21.5 Hz. With a pilot tone of 20 kHz this translates to detecting movements as slow as 18.5 cm/sec. The signal analysis module 824 can consider the bandwidth of the pilot tone itself, which with no motion can be ~80 Hz, for example. The pilot tone bandwidth can vary from 60-120 Hz (1-3 bins on either side of the tone) depending on the quality of the commodity hardware's sound system. Thus, signal analysis module 824 can consider a "motion event" to occur when there is a frequency shift bandwidth of 4 or more bins, for example. This threshold allows sufficiently slow movements of the hand to be detected while ignoring false positives due to variations in the bandwidth.

[00039] Measurable Properties

[00040] In addition to the fundamental frequency shift, signal analysis module 824 can also compute other useful features for inferring gestures.

[00041] Velocity: The measured frequency change is proportional to the absolute speed of the target. Signal analysis module 824 can measure the difference between the original (e.g., pilot tone) and reflected frequency to differentiate between slow, medium and fast gestures.

[00042] Direction: Determining whether the hand is moving towards or away from the computer can be made from the sign of shift in frequency. A positive shift indicates towards and negative shift indicates away. In multiple microphones scenarios the signal analysis module 824 can compare the detected Doppler shift differences for the multiple microphones to identify further information about the user control gesture.

[00043] Proximity and Size of Target: The amplitude of the observed signal increases as the target moves closer to the computer, and it also increases with size and reflectivity of the target. For example, a larger hand or open palm can manifest as larger amplitude than a smaller or fist hand.

[00044] Time Variation: Measuring the variation of the above three properties over time allows signal analysis module 824 to observe both the rate of change and use it for filtering spurious signals. For example, any motion that lasts for a very short period of time can be reliably filtered out, while longer lasting motion events can be used to identify activities like walking towards or away from the computer.

GESTURES AND USE CASES

[00045] The features described above can be combined to form complex gestures. For instance, as described above relative to FIG. 1, scrolling can be effectively detected and used to control a web-browser or word processing application, among others. Some
5 implementations can employ a clutching mechanism to reduce and/or prevent inadvertent scrolling as the hand returns to a particular position. For example, using the velocity feature and scrolling only when it meets a certain speed criteria can make this possible.

[00046] Double Tap or Tapping Gesture: By observing the change in direction over time, the 'frequency' at which the direction is changing can be computed. The value of
10 this direction frequency can be used for detecting tap gestures, which can be further used to distinguish quick taps, much like a mouse double-click, from slower taps.

[00047] The two handed seesaw gesture was illustrated relative to FIGS. 2-7 and entails moving both hands in opposite directions at the same time to activate the gesture. It is detected by the presence of both shifted up and down frequency components in the
15 same FFT vector.

[00048] Proximity and Flick: This gesture is activated when at least N-consecutive motion events in the same direction are detected. A large N can signify whether a person is walking (as opposed to moving the hand which has a predictably smaller N consecutive motions). The walking gesture can be used to automatically put a computer to sleep or
20 wake it up as a user walks away or towards it. The above listing of detectable gestures is provided for purposes of explanation and many other user gestures can be detected.

[00049] The user interaction module 826 can set up the user gestures on the computer. For instance, some implementations could come with one or more predefined user gestures. In such a case, the user interaction module may cause an interface, such as
25 a graphical user interface (GUI) to be generated that explains and/or shows the user how to make individual user gestures. The user interaction module may also allow the user to define what computer control to map to the user gesture. For instance, the GUI may show how to make the scroll gesture shown in FIG. 1.

[00050] The GUI can then allow the user to map the user gesture to a computer
30 control function of his/her choosing. One such example is evidenced in horizontal row 832 of datatable 828. The datatable includes a gesture name column 834, a gesture profile column 836, and a control function column 838. In row 832, the gesture name is 'scrolling' and the gesture profile is 'A'. The A represents the graphed frequency profile recorded for the gesture, which is not illustrated for sake of brevity. In this case, the

scrolling gesture is mapped to the control function of 'scroll downward', which as mentioned above and is user adjustable. For instance, the user might change the corresponding control function to 'decrease volume' on the music player application described relative to FIG. 1.

5 [00051] In another example, assume that the user wants to 'teach' the user a new gesture. The GUI can offer this feature. For instance, the GUI may instruct the user to perform the user gesture one or more times. The user interaction and/or the signal analysis modules 826 and 824 can analyze the signals during the user gesture performance until a satisfactory gesture profile is obtained. For instance, assume that the user performs the
10 two handed seesaw gesture illustrated relative to FIGS. 2-7 above. The signal analysis module 824 can identify the gesture profile as 'B' as indicated in row 840 of the datatable. The user interaction module 826 can allow the user to define the corresponding control function, such as by providing a listing of control functions for the user. In this case, assume that the user selected 'change play status' of the music player.

15 [00052] At a subsequent point in time, the user may want to control the computer 802 and may perform a user gesture. The user gesture can be captured and analyzed by the signal analysis module 824 to determine if the captured gesture profile of the performed gesture matches (or approximately matches as defined by a similarity threshold) any of the gesture profiles in data table 828. If there is an approximate match,
20 the corresponding control function (e.g., from the horizontal row of the matching gesture profile) can be performed.

[00053] In some implementations, if excessive amounts of incorrect approximate matches are made, the user can adjust the similarity threshold upward. In the opposite case where the gestures are not identified as matches, the user can adjust the similarity
25 threshold downward via the user interaction module 826.

[00054] Additionally, the user interaction module 826 can allow the user to set and/or adjust a range of detection. For instance, the user can set a threshold amplitude value, such that only occurrences in a frequency profile that exceed the amplitude threshold are mapped to the datatable 828. Alternatively, the user may adjust the volume
30 of the speakers, which in turn controls the loudness (e.g. amplitude) of the emitted pilot tone. The volume control can be used to regulate the effective detection range, which could be useful in crowded situations in which the user may explicitly not want to pick up ambient movement.

[00055] In some implementations, the Doppler-based user gesture identification component 808 can be pre-installed on the computer 802 by a manufacturer or an intermediary that prepares and/or sells the computer to an end user (e.g., customer). In other implementations, the Doppler-based user gesture identification component 808 can be installed by the user. For instance, the user may download the Doppler-based user gesture identification component over network 818 as a set of computer-readable instructions that are then stored on the computer in a tangible form, such as on storage 814. In another implementation, the user may obtain the Doppler-based user gesture identification component on a storage device such as memory card 820 to install the Doppler-based user gesture identification component on his/her computer.

[00056] Further, in the illustrated implementation, computer 802 is configured with a general purpose processor 812 and storage 814. In some alternative configurations, the computer can include a system on a chip (SOC) type design. In such a case, functionality provided by various components of the computer can be integrated on a single SOC or multiple coupled SOC's. In one such example, the computer can include shared resources and dedicated resources. An interface(s) can facilitate communication between the shared resources and the dedicated resources. As the name implies, dedicated resources can be thought of as including individual portions that are dedicated to achieving specific functionalities.

[00057] Shared resources can be storage, processing units, etc. that can be used by multiple functionalities. In this example, the shared resources can include the processor.

METHOD EXAMPLES

[00058] FIG. 9 illustrates a flowchart of a technique or method 900 that is consistent with at least some implementations of the present concepts.

[00059] At block 902, the method can emit a sound signal from a computer. The sound signal can be emitted from speaker(s) on the computer. In some implementations, these speaker(s) can be existing multipurpose speakers that can be utilized to achieve other functionalities on the computer.

[00060] At block 904, the method can capture the sound signal at the computer. The sound signals can be captured with microphone(s). In some implementations, the microphones can be existing multipurpose microphones that can be utilized to achieve other functionalities on the computer.

[00061] At block 906, the method can compare a frequency range of the emitted sound signal to the same frequency range of the captured sound signal to detect Doppler shift created by a user control gesture.

5 [00062] At block 908, the method can map the user control gesture to a control function.

[00063] At block 910, the method can execute the control function on the computing device.

10 [00064] The order in which the above-mentioned methods are described is not intended to be construed as a limitation, and any number of the described blocks can be combined in any order to implement the method, or an alternate method. Furthermore, the method can be implemented in any suitable hardware, software, firmware, or combination thereof, such that a computing device can implement the method. In one case, the method is stored on a computer-readable storage media as a set of instructions such that execution by a computer causes the computer to perform the method.

15 CONCLUSION

[00065] Although techniques, methods, devices, systems, etc., pertaining to Doppler-based user gesture detection are described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or acts described. Rather, the specific features and acts are disclosed as exemplary forms of implementing the claimed methods, devices, systems, etc. Also, as used in this document the terms "a" and "an" are intended to introduce an element, which may be singular or plural. Therefore, "a" or "an" can be used to cover both the singular sense (e.g., one) and the plural sense (e.g., multiple).

25

CLAIMS

1. A computer, comprising:
a speaker and a microphone; and,
a gesture identification component configured to cause sounds of a given
5 frequency to be emitted from the speaker and captured by the microphone, the gesture
identification component further configured to evaluate the captured sounds around the
given frequency to detect Doppler effect shifts caused by user gestures, the gesture
identification component further configured to map a Doppler effect shift to an individual
control function and to cause the control function to be executed by the computer, and
10 wherein the gesture identification component is further configured to allow the speaker
and the microphone to perform other functions during the emitting and capturing.
2. The computer of claim 1, further comprising a user input device and
wherein the gesture identification component is further configured not to evaluate the
captured sounds when a user is engaging the user input device.
- 15 3. The computer of claim 2, wherein the user input device comprises a
physical keyboard, a virtual keyboard, a mouse, a track pad, or a touch screen.
4. The computer of claim 1, wherein the speaker comprises a single speaker or
multiple speakers and the microphone comprises a single microphone or multiple
microphones.
- 20 5. The computer of claim 1, wherein the gesture identification component is
further configured to allow a user to define a range from the computer within which the
control function is to be executed and outside of which the control function is not to be
executed.
6. The computer of claim 5, wherein the gesture identification component is
25 further configured to evaluate an amplitude of Doppler effect shift to determine whether
the user movement occurred within the range.
7. The computer of claim 5, wherein the range extends radially in three
dimensions from the computer.
8. The computer of claim 1, wherein the other function of the speaker is
30 playing music and wherein the sounds of the given frequency comprise a portion of the
music.
9. The computer of claim 1, wherein the gesture identification component is
configured to test a range of frequencies and evaluate the corresponding captured sounds

and to select the given frequency when interference associated with the given frequency is relatively less than other frequencies of the range of frequencies.

10. The computer of claim 9, wherein the interference comprises other sounds generated by the speaker or background noise.

5 11. The computer of claim 9, wherein the gesture identification component is configured to retest the range of frequencies according to one or more parameters.

12. The computer of claim 11, wherein the one or more parameters relate to a predefined period of time or relate to detected interference at the given frequency exceeding a predefined threshold.

10 13. A computer-readable storage media having instructions stored thereon that when executed by a computing device cause the computing device to perform acts, comprising:

emitting a sound signal from a computer;

capturing sound at the computer;

15 comparing a frequency range of the emitted sound signal to the same frequency range of the captured sound to detect Doppler shift created by a user control gesture;

mapping the user control gesture to a control function; and,

executing the control function on the computing device.

20 14. The computer-readable storage media of claim 13, wherein the emitting comprises simultaneously emitting the sound signal at a constant frequency and other sound signals from a single speaker set.

15 15. The computer-readable storage media of claim 13, wherein the capturing is performed by multiple microphones and further comprising comparing the detected Doppler shift differences for the multiple microphones to identify further information about the user control gesture.

1/6

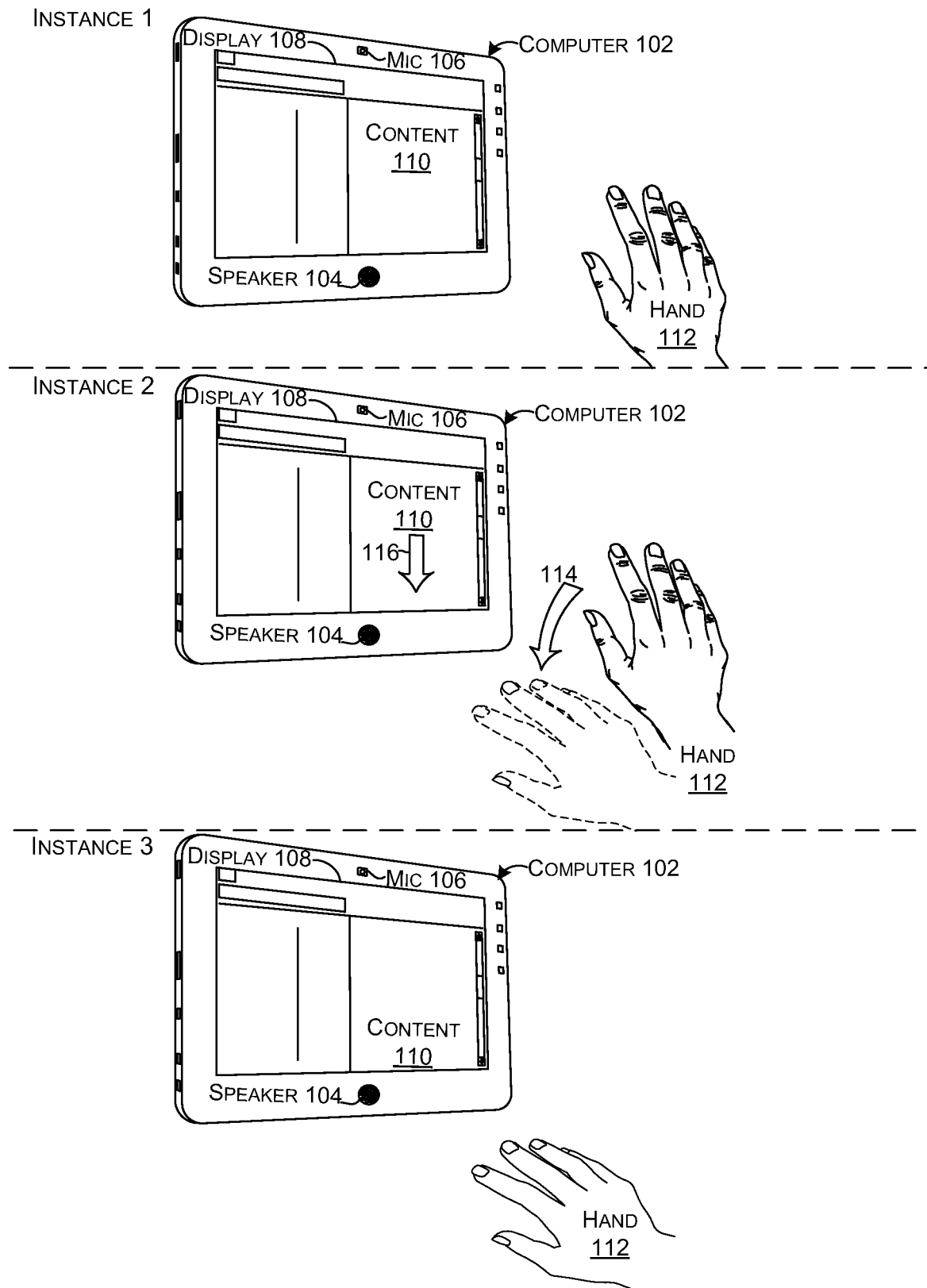
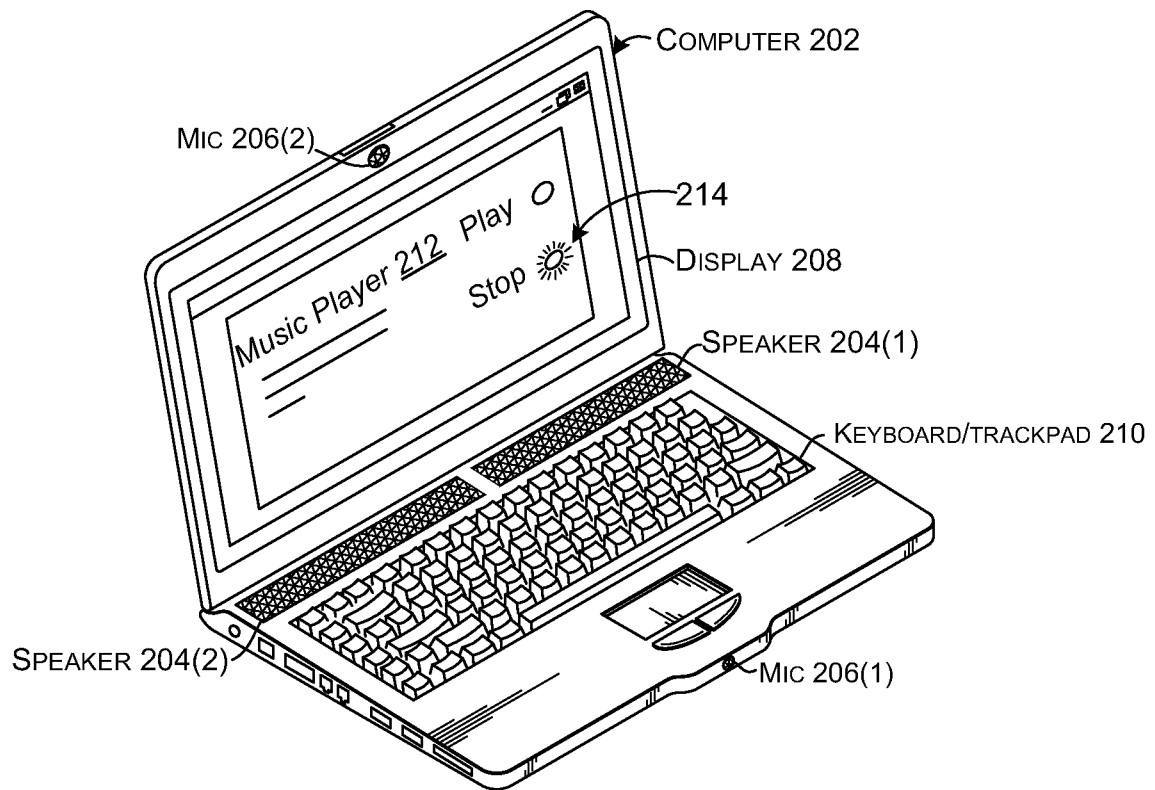
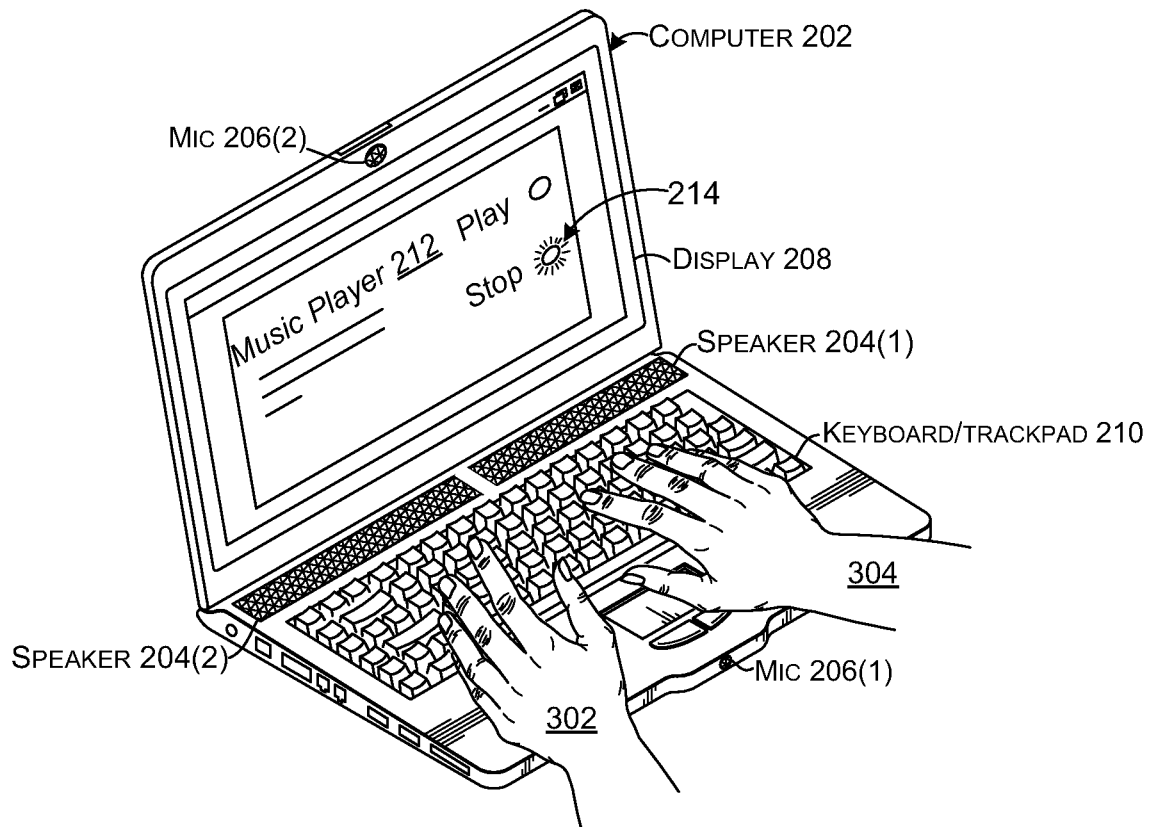
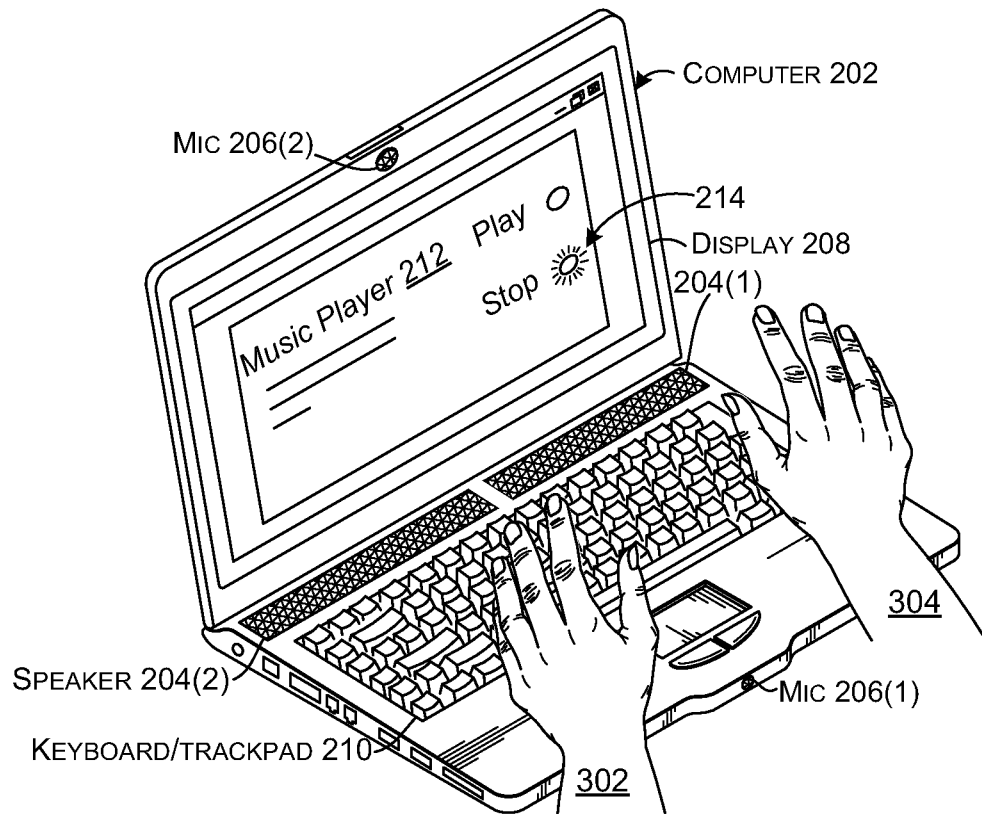
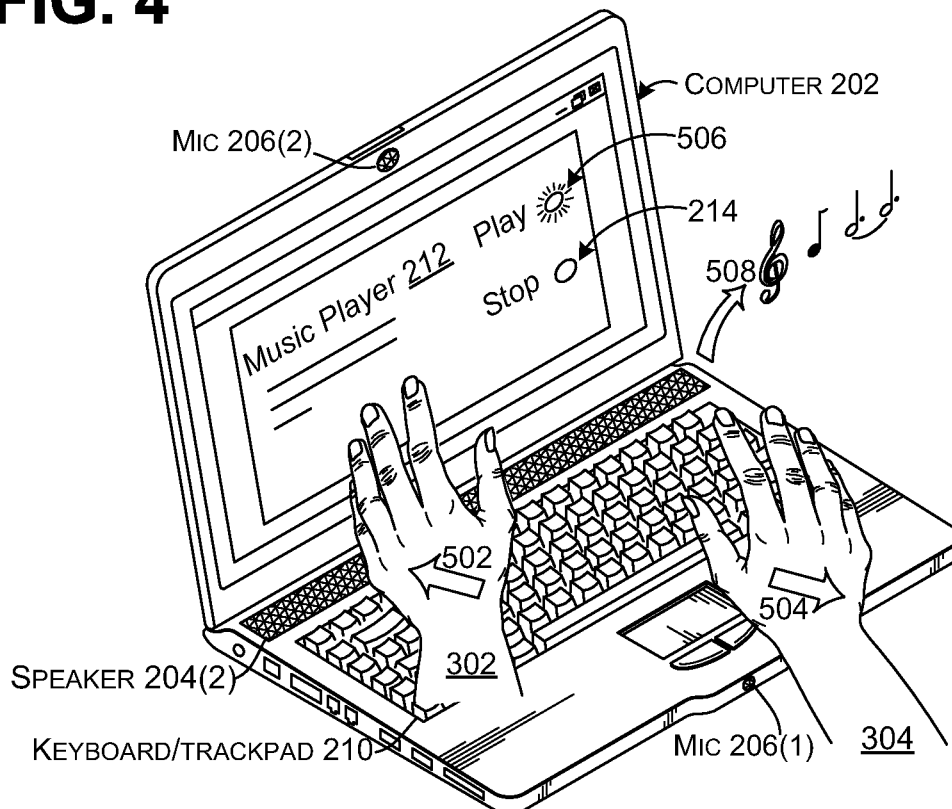


FIG. 1

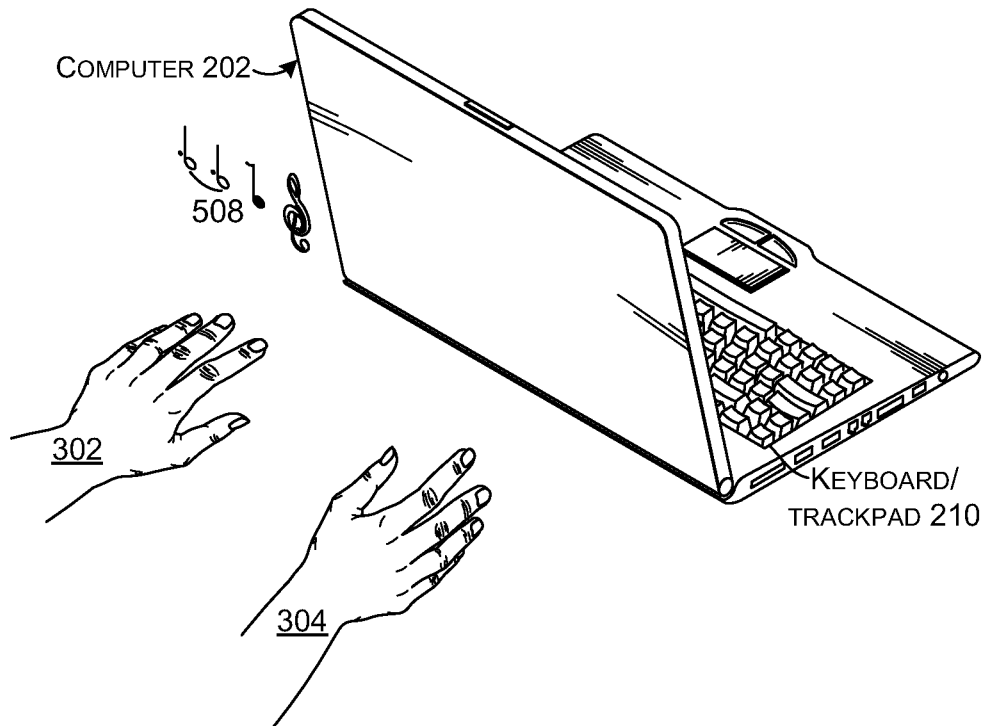
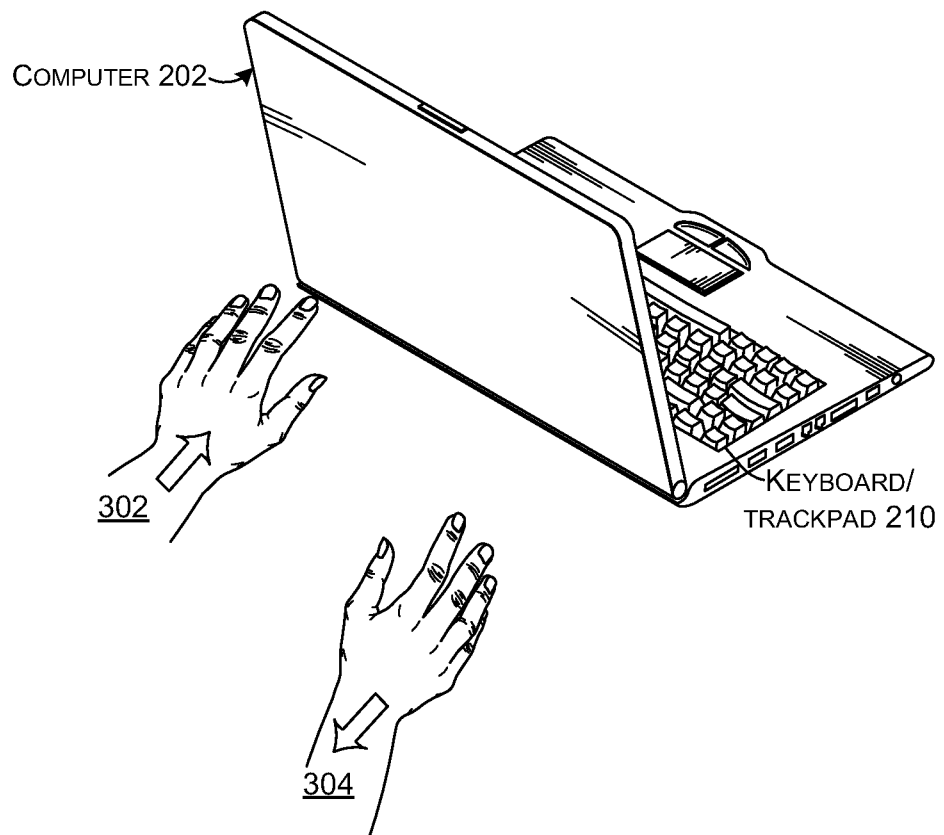
2/6

**FIG. 2****FIG. 3**

3/6

**FIG. 4****FIG. 5**

4/6

**FIG. 6****FIG. 7**

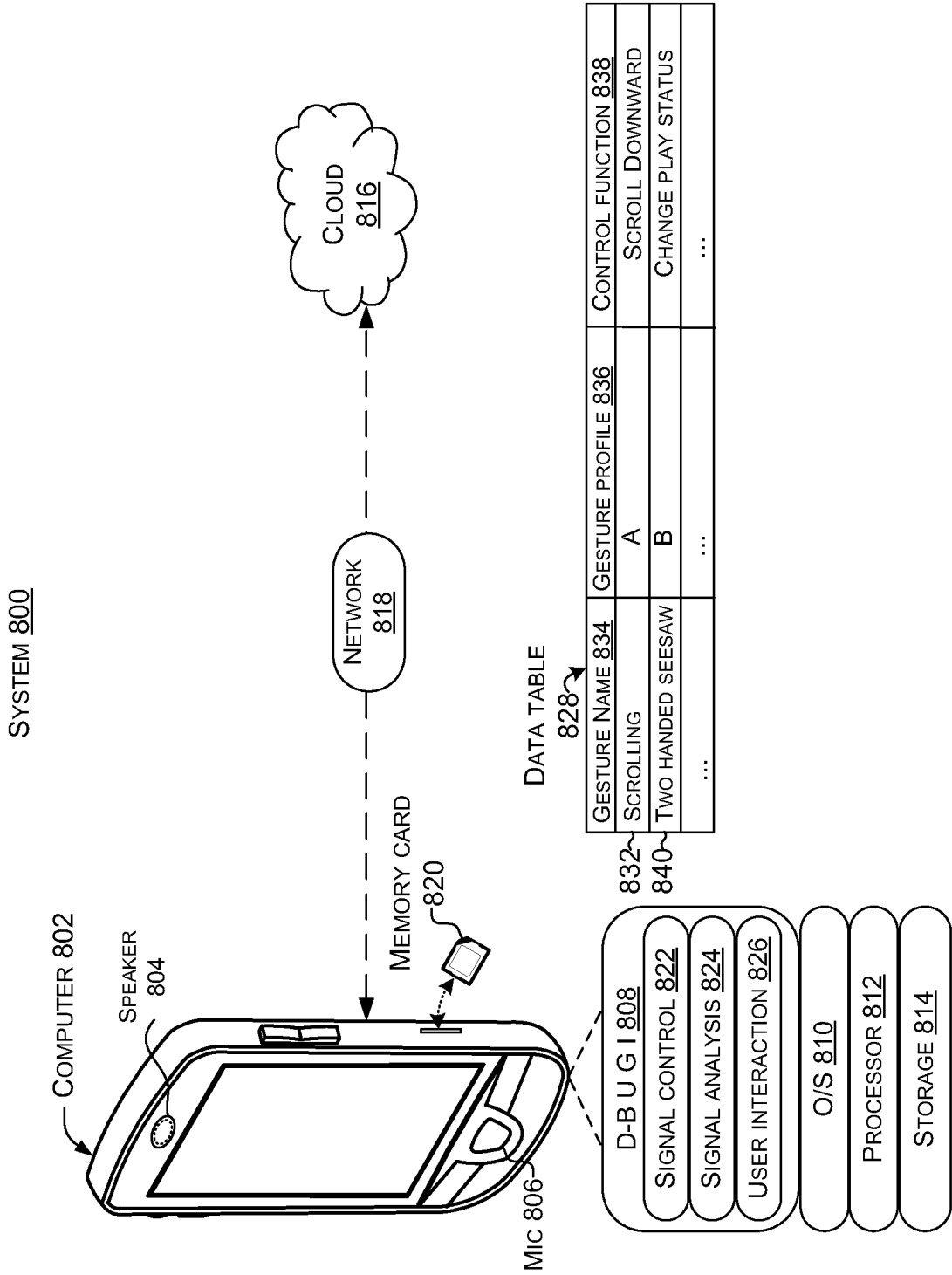
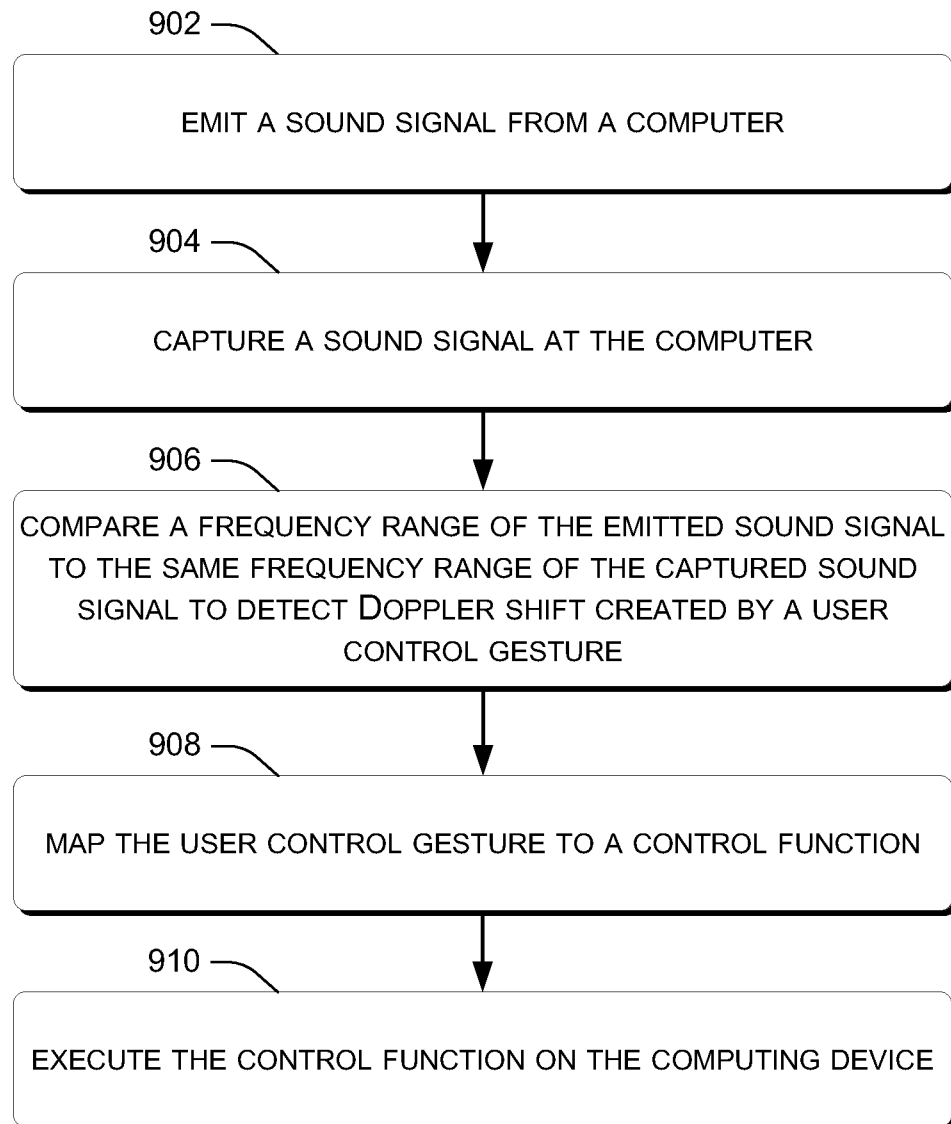


FIG. 8

6/6

METHOD 900**FIG. 9**

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2012/069030**A. CLASSIFICATION OF SUBJECT MATTER****G06F 3/01(2006.01)i, G06F 3/16(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F 3/01; G09G 5/08; H04B 17/00; G01P 3/36; G06K 9/00; GOIC 3/08; G09G 5/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS(KIPO internal) & Keywords: gesture, identification, sound, speaker, microphone, doppler effect shift, control, and similar terms.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2011-0103448 A1 (DAHL, T. et al.) 5 May 2011 See paragraphs [0075], [0082]-[0085], and [0095H0101];	13, 15
A	claims 48, 54, 63, and 75; and figure 1.	1-12, 14
X	KALGAONKAR, K. et al. 'ONE-HANDED GESTURE RECOGNITION USING ULTRASONIC DOPPLER SONAR.' In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2009). Taipei, Taiwan. 19-24 April, 2009. pp. 1889-1892.	13, 15
A	See pages 1889-1891 and figures 1-3.	1-12, 14
A	TARZIA, S. P. et al. 'Sonar-based Measurement of User Presence and Attention.' In: The 11th International Conference on Ubiquitous Computing (Ubicomp 09). Orlando, Florida, USA. Sep. 30 - Oct. 3, 2009. pp. 89-92.	1-15
A	See pages 90-91.	
A	US 2011-0181509 A1 (RAUTIAINEN, T. T. et al.) 28 July 2011 see paragraphs [0014H0039] and [0061H0088]; claims 1-2, 5, and 20; and figures 1 and 6.	1-15

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

21 MARCH 2013 (21.03.2013)

Date of mailing of the international search report

22 MARCH 2013 (22.03.2013)

Name and mailing address of the ISA/KR

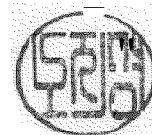
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan
City 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

NHO, Ji Myong

Telephone No. 82-42-481-8528



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2012/069030

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2010-0202656 A1 (RAMAKRISHNAN, B. R. et al.) 12 August 2010 See paragraphs [0017]-[0025] and [0035]-[0055] ; claim 1; and figures 1-5.	1-15
A	US 2007-0211022 A1 (BOILLOT, M.) 13 September 2007 See paragraphs [0021]-[0022] and [0032]-[0043] ; claims 1, 5-6, and 9 ; and figures 4 and 7-8.	1-15
A	US 2007-0121097 A1 (BOILLOT, M.) 31 May 2007 See paragraphs [0029]-[0039] , [0042]-[0043] , and [0069]; claims 1-4; and figures 1-5 and 7 .	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2012/069030

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2011-0103448 A 1	05,,05,,2011	EP 2271951 A 1 GB 0806196 DO JP 2011-516848 A WO 2009-122193 A 1	12. 01.,2011 14. 05.,2008 26. 05.,2011 08. 10.,2009
US 2011-0181509 A 1	28,,07,,2011	WO 2011-092628 A 1	04. 08.,2011
us 2010-0202656 A 1	12,,08,,2010	None	
us 2007-0211022 A 1	13,,09,,2007	US 8139029 B2	20. 03.,2012
us 2007-0121097 A 1	31.,05,,2007	US 2010-0204575 A 1 US 7414705 B2 US 7724355 B 1 US 8098544 B2	12. 08.,2010 19. 08.,2008 25. 05.,2010 17. 01.,2012