



- (51) International Patent Classification:  
*G05B 11/58* (2006.01)    *G05B 15/00* (2006.01)
- (21) International Application Number:  
PCT/RU2013/000393
- (22) International Filing Date:  
13 May 2013 (13.05.2013)
- (25) Filing Language: English
- (26) Publication Language: English
- (71) Applicant: 3DIVI COMPANY [RU/RU]; 48, Makeeva pr., Miass, 456320 (RU).
- (72) Inventors: MOROZOV, Dmitry Aleksandrovich; 48, Makeeva pr., Miass, 456320 (RU). CHERNETSKAYA, Nina Vladimirovna; 64D, Lenina pr., Chelyabinsk, 454000 (RU).
- (74) Agent: SVETLOV, Ilya Aleksandrovich; a/ya 50, St.-Petersburg, 194156 (RU).

- (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).

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR CONTROLLING MULTIPLE ELECTRONIC DEVICES

(57) Abstract: Electronic devices can be controlled by a motion sensing device, such as a cellular phone, remote controller, gaming device, or any kind of wearable computers such as glasses, by tracking its location and orientation within a 3D environment and identifying user commands. For these ends, there is provided a computing device for obtaining depth maps of the 3D environment within which a user with the motion sensing device is present. The computing device further processes motion and orientation data received from the motion sensing device to associate the motion sensing device with a common coordinate system such as a coordinate system in relation of the computing device. Further steps include determining that the motion sensing device is oriented towards a particular electronic device to be controlled such as a home appliance. After receiving a user command, there is generated a corresponding control command for the electronic device.

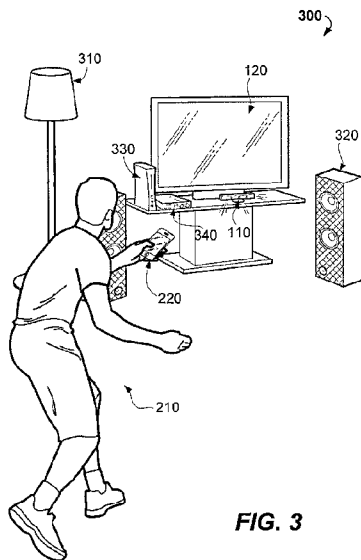


FIG. 3

WO 2014/185808 A1

**Published:**

— *with international search report (Art. 21(3))*

System and method for controlling multiple electronic devices

5

## TECHNICAL FIELD

[0001] This disclosure relates generally to human-computer interfaces and, more particularly, to the technology for controlling operation of one or more electronic devices, such as a home appliance, by determining that a user orients a user device having motion sensitivity (e.g., a smart  
10 phone, remote controller, gaming device, wearable computer, head-mounted glasses-like computing device, etc.) towards a predetermined location of one of said electronic devices and by identifying an activation event. The determination is made by real time processing of depth maps of a three-dimensional (3D) environment, within which the user is present, and of user  
15 device motion and orientation data.

## DESCRIPTION OF RELATED ART

[0002] The approaches described in this section could be pursued, but are not necessarily approaches that have previously been conceived or  
20 pursued. Therefore, unless otherwise indicated, it should not be assumed that any of the approaches described in this section qualify as prior art merely by virtue of their inclusion in this section.

[0003] Technologies associated with human-computer interaction have evolved over the last several decades. There are currently many various  
25 input devices and associated interfaces that enable users of various electronic devices including, for example, computers, game consoles, television devices, home appliances, lighting devices, audio systems, etc., to control them

remotely. Keyboards, keypads, pointing devices, joysticks, remote controllers, and touchscreens are just some examples of input devices that can be used to interact with various electronic devices. On the other hand, most of these electronic devices can be operated using dedicated input devices and every electronic device may have its own input device. In this regard, many users may find it annoying or difficult to utilize a plurality of input devices with no flexibility. There are known just a few examples when a traditional remote controller for a television device and a traditional remote controller for a corresponding audio system are combined into a single device. However, this approach does not solve the problem of flexibility to control multiple electronic devices of various types utilizing just a single remote control device or any suitable remote control device among a plurality of such devices.

[0004] One of the rapidly growing technologies in the field of human-computer interaction is the gesture recognition technology, which may enable the users to interact with multiple electronic devices naturally, using body language rather than mechanical devices. In particular, the users can make inputs or generate commands using gestures or motions made by hands, arms, fingers, legs, and so forth. For example, using the concept of gesture recognition, it is possible to point a finger at a computer screen and cause the cursor to move accordingly.

[0005] There currently exist various control systems having a 3D camera (which also referred herein as to a depth sensing camera), which captures scene images in real time, and a computing unit, which interprets captured scene images so as to generate various commands based on identification of user gestures. Typically, the control systems have very limited computation resources. Also, traditional small resolution of the 3D camera makes it difficult to identify and track motions of relatively small objects such as motion sensing devices.

[0006] On the other hand, the motion sensing devices may play an important role for human-computer interaction, especially, for gaming software applications. The motion sensing devices may refer to controller wands, remote control devices, or pointing devices which enable the users to  
5 generate specific commands by pressing dedicated buttons arranged thereon. Alternatively, commands may be generated when a user makes dedicated gestures using the motion sensing devices such that various sensors imbedded within the motion sensing devices may assist in determining and tracking user gestures. Accordingly, the computer or gaming console can be  
10 controlled via the gesture recognition technology, as well as by the receipt of specific commands originated by pressing particular buttons.

[0007] Typically, the control systems, when enabled, monitor and track all gestures performed by users. However, to enable the control systems to identify and track a location, motion and orientation of a motion sensing  
15 device having a relatively small size, a high resolution depth sensing camera and immoderate computational resources may be needed. It should be noted that state of the art 3D cameras, which capture depth maps, have a very limited resolution and high latency. This can make it difficult, or even impossible, for such systems to precisely locate the relatively small motion  
20 sensing device at the depth map and determine parameters such as its orientation. Today's motion sensing devices, on the other hand, may also include various inertial sensors which dynamically determine their motion and orientation. However, this information is insufficient to determine orientation of the motion sensing devices within the 3D environment within  
25 which they are used.

[0008] In view of the foregoing drawbacks, there is still a need for improvements of control systems that may be used in controlling multiple electronic devices with the help of a single motion sensing device.

## SUMMARY

[0009] This summary is provided to introduce a selection of concepts in a simplified form that are further described in the Detailed Description below. This summary is not intended to identify key or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

[0010] The present disclosure refers to “point-and-control” systems allowing a user to point or orient a user device towards a particular electronic device and control it (i.e., activate, deactivate, adjust operation) by providing a corresponding command. The user device may include any suitable electronic device allowing it to sense its motions and orientation. Some examples of user device may include devices that can be held in a user hand such as a cellular phone, game pad, remote controller, or other non-electronic devices, like sports implement, but provided with motion sensors. In certain aspects the user device may relate to wearable computing devices (e.g., glasses) having ability to sense its motions and orientation. Regardless of the user device type, it employs one or more sensors including, for example, accelerometers, gyroscopes, and magnetometers which generate user device motion data and user device orientation data in response to any movements of the user device. To control the particular electronic, the user may further need to make an input, which may refer to actuation of one or more buttons of the user device, interaction with a graphical interface of the user device, or predetermined gestures made by a hand or head of the user.

[0011] According to one or more embodiments of the present disclosure, provided is a control system including a depth sensing camera and/or video camera, which is/are used for obtaining depth maps of a 3D environment, within which at least one user holding or wearing the user

device is present. The control system may further include a communication module for receiving, from the user device, said user device motion data and user device orientation data associated with at least one motion of the user device. The control system may further include a computing unit, operatively  
5 coupled to the depth sensing camera (and/or video camera) and the communication unit. The communication unit may also be operatively coupled to a plurality of electronic devices, which can be controlled by the user. In general, the control system is configured to process the user device motion data, user device orientation data, depth maps, user inputs, and video  
10 (if captured) to identify a qualifying event, when the user orients the user device towards a particular electronic device and provides a corresponding input, and generate a corresponding control command for this electronic device.

[0012] In operation, according to one or more embodiments of the present disclosure, the user device may need to be properly identified and tracked within a coordinate system used by the control system (hereinafter referred to as "first coordinate system") before electronic devices can be controlled using the principles disclosed herein. For these ends, the user may need, optionally and not necessarily, to perform certain actions to assist the  
15 control system determine a location and orientation of the user device within the first coordinate system.

[0013] In one example, the user may be required to be present in front of the depth sensing device (and/or video camera), hold or wear the user device, and then make an input to the control system as to where specifically  
20 the user device is located (e.g., the user may specify that the user device is held in the right hand). In another example, the user may be required to be present in front of the depth sensing device (and/or video camera), and then place the user device in a predetermined location so as the control system

may determine the location of the user device within the first coordinate system and then easily track it within the first coordinate system.

[0014] In yet another example, the user may be required to be present in front of the depth sensing device (and/or video camera), hold or wear the user device, and make a predetermined gesture. For example, the user may need to make a gesture of moving the user device in parallel or perpendicular to the depth sensing device (and/or video camera), make a nodding motion (when a head-mountable user device is used), and so forth. When any of these predetermined motions are made by the user, the user device may generate corresponding user device motion data and user device orientation data using one or more motion sensors, and then wirelessly transmit this data to the control system. It should be noted, however, that the user device motion data and user device orientation data are associated with the second (internal) coordinate system (hereinafter referred to as "second coordinate system"), which differs from the first coordinate system used by the control system.

[0015] At the same time, the same motion of the user may be tracked by the depth sensing device (and/or video camera) which results in generation of a series of depth maps of a scene within which the user is present. The depth maps may include a plurality of coordinates associated with the user and possibly the user device on the first coordinate system. The depth maps may then be processed to identify the user and the at least one motion of the user hand or head. The computing unit may generate a virtual skeleton of the user, which skeleton may have multiple virtual joints having coordinates on the first coordinate system. Accordingly, once a motion of the at least one user hand or head is identified by processing of the depth maps, the computing unit obtains a corresponding set of coordinates on the first coordinate system. The set of these coordinates may constitute first motion

data.

[0016] According to one or more embodiments of the present disclosure, the user device motion data and user device orientation obtained from the user device may need to be transformed so as to relate to the coordinate system used by the control system before electronic devices can be controlled using the principles disclosed herein. For these ends, the data obtained from the user device may be optionally, and not necessarily, calibrated. The calibration, in general, means that the internal coordinate system used by the user device and corresponding coordinates are brought into accordance with the first coordinate system, or, in other words, a relation should be established between the first coordinate system and the second coordinate system. According to various embodiments of the present disclosure, the calibration may be performed in a number of various ways.

[0017] For example, in operation, the control system receives and processes the user device motion data and the user device orientation data from the user device. In particular, the user device motion data and the user device orientation data may be transformed by multiplying by a correlation matrix, transformation matrix or performing any other mathematical operation. The transformed user device motion data and user device orientation data now constitute second motion data. It should be noted however, that the user device motion data may not be necessarily transformed to constitute said second motion data.

[0018] At the succeeding steps, the computing unit compares (maps) the first motion data to the second motion data so as to find correlation between the motion of the at least one user hand or head identified on the depth maps and the motion of the user device as identified from processed user device motion and orientation data. Once such correlation is found, the computing unit may assign the set of coordinates

associated with the at least one user hand or head making said motion to the user device. Accordingly, the computing unit determines the location and orientation of the user device with the first coordinate system.

[0019] Further, the control system may track the location and orientation of the user device in real time. In certain embodiments, the control system may determine that the user device is oriented towards a predetermined location of a particular electronic device, which the user wants to control. In general, the electronic device may be selected from a group consisting of a home appliance, lighting device, audio or video system, computer, game console, etc. The control system may maintain a database storing locations of any used and controllable electronic devices.

[0020] Accordingly, if it is determined that the user orients the user device towards the particular electronic device, the user may be then prompted to make an input to control this electronic device. The input can be made by the user by actuating one or more physical or virtual buttons of the user device, by making a particular gesture, by providing a voice command, and so forth. In response to the user input and the determination, the computing unit generates a corresponding control command to the selected electronic device. The command may be then transmitted to this electronic device so as it changes its operation mode or certain operation function is adjusted.

[0021] Thus, the present technology allows for easy control of multiple electronic devices using a single user device, which technology does not require immoderate computational resources or high resolution depth sensing cameras. Other features, aspects, examples, and embodiments are described below.

## BRIEF DESCRIPTION OF THE DRAWINGS

- [0022] Embodiments are illustrated by way of example, and not by limitation in the figures of the accompanying drawings, in which like references indicate similar elements and in which:
- 5 [0023] FIG. 1 shows an example system environment for providing a real time human-computer interface.
- [0024] FIG. 2 is a general illustration of a scene suitable for controlling an electronic device.
- [0025] FIG. 3 is an illustration of a scene suitable for controlling  
10 multiple electronic devices using a user device.
- [0026] FIG. 4A shows a simplified view of an exemplary virtual skeleton as can be generated by a control system based upon a depth map.
- [0027] FIG. 4B shows a simplified view of exemplary virtual skeleton associated with the user a holding a user device.
- 15 [0028] FIG. 5 shows an environment suitable for implementing methods for controlling various electronic devices utilizing a user device.
- [0029] FIG. 6 shows a simplified diagram of user device according to an example embodiment.
- [0030] FIG. 7 shows an exemplary graphical user interface that may  
20 be displayed on a touchscreen of user device, when it is oriented towards a particular electronic device.
- [0031] FIG. 8 is a process flow diagram showing an example method for controlling one or more electronic devices using a user device.
- [0032] FIG. 9 is a process flow diagram showing another example  
25 method for controlling one or more electronic devices using a user device.
- [0033] FIG. 10 is a process flow diagram showing yet another example method for controlling one or more electronic devices using a user device.

[0034] FIG. 11 is a diagrammatic representation of an example machine in the form of a computer system within which a set of instructions for the machine to perform any one or more of the methodologies discussed herein is executed.

5

#### DETAILED DESCRIPTION

[0035] The following detailed description includes references to the accompanying drawings, which form a part of the detailed description. The drawings show illustrations in accordance with example embodiments. These example embodiments, which are also referred to herein as "examples," are described in enough detail to enable those skilled in the art to practice the present subject matter. The embodiments can be combined, other embodiments can be utilized, or structural, logical, and electrical changes can be made without departing from the scope of what is claimed. The following detailed description is therefore not to be taken in a limiting sense, and the scope is defined by the appended claims and their equivalents. In this document, the terms "a" and "an" are used, as is common in patent documents, to include one or more than one. In this document, the term "or" is used to refer to a nonexclusive "or," such that "A or B" includes "A but not B," "B but not A," and "A and B," unless otherwise indicated.

[0036] The techniques of the embodiments disclosed herein may be implemented using a variety of technologies. For example, the methods described herein may be implemented in software executing on a computer system or in hardware utilizing either a combination of microprocessors, controllers or other specially designed application-specific integrated circuits (ASICs), programmable logic devices, or various combinations thereof. In particular, the methods described herein may be implemented by a series of

computer-executable instructions residing on a storage medium such as a disk drive, solid-state drive or on a computer-readable medium.

### Introduction

5 [0037] The embodiments described herein relate to computer-implemented methods and corresponding systems for controlling multiple electronic devices by determining and tracking the current location and orientation of a user device through the use of depth maps.

[0038] In general, depth sensing cameras or 3D cameras can be used  
10 to generate depth maps of a scene within which a user may be located. Depth maps may be also generated by processing video data obtained using traditional video cameras. In certain embodiments however, a combination of both depth sensing cameras and video cameras may be utilized. In either case, the depth maps may be associated with a 3D coordinate system (e.g., a  
15 3D Cartesian coordinate system, also referred herein to as a “first coordinate system”) and include plurality of coordinates for objects present in the scene. The depth map analysis and interpretation can be performed by a computing unit operatively coupled to or embedding the depth sensing camera or video camera. Some examples of computing units may include a desktop computer,  
20 laptop computer, tablet computer, gaming console, audio system, video system, cellular phone, smart phone, personal digital assistant (PDA), set-top box (STB), television set, smart television system, in-vehicle computer, infotainment system, or any other wired or wireless electronic device. The computing unit may include, or be operatively coupled to, a communication  
25 unit which may communicate with various user devices and, in particular, receive motion and/or orientation data of user devices.

[0039] In general, the computing unit processes and interprets the depth maps of a scene in real time. By mere processing of depth maps, it may

identify at least a user, which present within the scene, identify user's limbs, head, and generate a corresponding virtual skeleton of the user. The skeleton may include multiple virtual "joints" related to certain body parts and which possess certain coordinates on the first coordinate system. By processing of the depth maps, the computing unit may further determine that the user makes at least one motion (gesture) using his hand, arm or head. The coordinates of every joint can be tracked by the computing unit, and thus corresponding "First" motion data can be generated in real time, which may include a velocity, acceleration, orientation, or other information related to the user's motion.

[0040] The term "user device," as used herein, refers to a motion sensing device including any suitable electronic device enabled to sense motions and orientation. Some examples of motion sensing devices include an electronic pointing device, remote controller, cellular phone, smart phone, video game console, handheld game console, game pad, computer (e.g., a tablet computer), a wand, and so forth. The motion sensing device may also relate to wearable computing device like a head-mounted computing devices implemented, for example, in the form of glasses. Some additional examples of motion sensing devices may include various non-electronic devices, such as sports implements, which may include, for example, a tennis racket, golf club, hockey or lacrosse stick, baseball bat, sport ball, etc. Regardless of what type of motion sensing device is used, it may include various removably attached motion (or inertial) sensors or imbedded motion (or inertial) sensors. The motion or inertial sensors may include, for example, acceleration sensors for measuring acceleration vectors in relation to an internal coordinate system, gyroscopes for measuring the orientation of the motion sensing device, and/or magnetometers for determining the direction of the motion sensing device with respect to a pole. In operation, the user device having said sensors

dynamically generates “user device motion data” (which include acceleration data) and “user device orientation data” (which include rotational data, e.g., an attitude quaternion), both associated with an internal coordinate system (referred herein to as “second coordinate system”). Further, this user device  
5 motion data and orientation data may be transmitted to the computing unit over a wired or wireless network for further processing. The user device may optionally include a user control system including, for example, a keypad, keyboard, one or more buttons, joystick, touchscreen, touch pad, and so forth. In case the touchscreen is utilized, it may be provided with a corresponding  
10 graphical user interface having one or more actionable buttons.

[0041] It should be noted, however, the user device may not be able to determine its exact location within the scene, for example within a 3D coordinate system associated with the control system (referred herein to as “first coordinate system”). Although various geo-positioning devices, such as  
15 Global Positioning System (GPS) receivers, may be used in the user devices, the accuracy and resolution for determining its location within the scene is very low. On the other hand, the depth sensing camera and/or video camera may be of low resolution to locate the user device on the depth maps and identify its exact location and orientation. In light of this, the present  
20 technology takes an advantage of combining data obtained from both the user device and the depths maps to precisely locate the user device and determine its orientation in real time using standard computational resources.

[0042] In operation, the user device should be properly identified on the first coordinate system used by the computing unit before various  
25 electronic devices can be controlled by the user using the principles disclosed herein. In other words, the location and orientation of the user device should be known in the first coordinate system prior to determination of where the user device is oriented at.

[0043] For these ends, the user device motion data and user device orientation obtained from the user device may need to be transformed/modified so as to relate to the first coordinate system and not the second coordinate system. This process is referred herein to as "calibration."

5 The calibration, in general, means that the internal coordinate system used by the user device and corresponding coordinates are brought into accordance with the first coordinate system, or, in other words, a relation should be established between the first coordinate system and the second coordinate system. According to various embodiments of the present disclosure, the

10 calibration may be performed in a number of various ways.

[0044] For example, in operation, the user may be required, first, to be present within the scene, i.e., in front of the depth sensing camera (and/or video camera), hold or wear the user device depending on its type, and then make an input or selection to the computing unit identifying where the user

15 device is located within the scene. For example, the user may give a voice command informing that the user device is held in a left hand. Thus, the computing unit may determine an exact location of the user device with the first coordinate system, provided the coordinates of the user hands or head are properly acquired.

20 [0045] In another embodiment, the user may be required, to be present within the scene and then place the user device to a predetermined location. For example, the user may be required to put the user device on a table in front of the depth sensing camera. Since the computing unit is aware of the coordinates of the predetermined location within the first coordinate

25 system, the computing unit may assign the same coordinates to the user device and then track its location.

[0046] In yet another embodiment, the user may be required, first, to be present within the scene, hold or wear the user device depending on its

type, and make a predetermined gesture. For example, if the user possesses a handheld user device, such as a cellular phone, the user may need to place the user device in a predetermined spot or make a gesture of moving the user device in parallel or perpendicular to the axis of the depth sensing camera (and/or video camera), etc. If the user possesses a wearable user device, such as a head-mounted, glasses-like computing device, the user may need to make a nodding motion or make a hand gesture by moving one of the hands from the head towards the depth sensing camera (and/or video camera). When any of these predetermined gestures are made by the user, the user device may be recognized by the computing unit and “bound” to the first coordinate system such that the user device’s location and orientation would be known. It may simplify further tracking of the user motions.

[0047] More specifically, when the user makes a gesture using the user device, it generates corresponding user device motion data and user device orientation data using one or more motion sensors embedded therein, which data is associated with the second coordinate system. The user device motion data and user device orientation data are then wirelessly transmitted to the computing unit.

[0048] When, the computing unit receives the user device motion data and user device orientation data associated with the second coordinate system, it may then optionally, and not necessarily, transform the received data so as to generate corresponding data, but in the first coordinate system. It may be done by transforming coordinates of the user device’s motion and orientation given in the second coordinate system to corresponding coordinates of the first coordinate system by the use of any suitable transitional function.

[0049] In general, this optional transformation process for the user device motion data may require utilization of the user device orientation data.

It may further include multiplying the user device motion data by user device orientation data (which may be optionally modified). For example, this process may include multiplying the user device motion data by a rotation matrix, an instantaneous rotation matrix or a calibrated instantaneous rotation matrix, all of which are based on the user device orientation data. In another example, this process may include multiplying the user device motion data by the calibrated instantaneous rotation matrix and by a predetermined calibration matrix. In yet another example, the user device orientation data may be modified by multiplying by a predetermined correlation matrix. The user device motion data, which is optionally transformed so as to fit the first coordinate system, is now referred to as "Second" motion data. In certain embodiments, the user device motion data and/or the user device orientation data may stay not transformed.

[0050] Further, the computing unit compares the First motion data retrieved from the processed depth maps to the Second motion data obtained from the user device motion data and user device orientation data. When it is determined that the first motion data and second motion data coincide, are similar or in any other way correspond to each other, the computing unit determines that the user device is held or worn by the user. Since coordinates of the user's arm, hand or head are known and tracked, the same coordinates are then assigned to the user device. Therefore, the user device can be associated with the virtual skeleton of the user so that the current location and orientation of the user device can be determined and further monitored on the first coordinate system in real time.

[0051] Further, the computing unit may then dynamically determine whether or not the user device is oriented by the user towards a predetermined location of a particular electronic device to be controlled. If it is ascertained by the computing unit that the user orients the user device

towards a predetermined location of a particular electronic device, the computing unit may then prompt the user to make an input, for example, by pressing a dedicated button on the user device or by making a predetermined gesture or by providing a voice command. For example, when the user  
5 possesses a handheld user device, such as a cellular phone, the user may point it towards a desired electronic device to be controlled, and then press a physical or virtual button to trigger a certain action or provide a voice command. In another example, when the user possesses a wearable user device, such as a head-mountable computing device, the user may orient the  
10 head towards a desired location of the electronic device (i.e., look at the electronic device to be controlled) and also point one of the user's hands towards the same electronic device to be controlled (such that a virtual line can be drawn connecting the user head, hand and the electronic device to be controlled).

15 [0052] Therefore, the computing unit may determine which electronic device the user points at using any of the above methods, and upon receipt of the user input (i.e., code generated by pressing a button or voice command), the computing unit generates a corresponding control command to the electronic device. The control command may be then transmitted to the  
20 electronic device using wired or wireless interface.

[0053] The term "electronic device," as used herein, may refer to a wide range of controllable electronic devices including, for example, a computer, computer periphery device, printing device, scanning device, gaming console, game pad, television system, set-top box, video system,  
25 audio system, mini system, speaker(s), microphone, router, modem, networking device, satellite receiver, lighting device, lighting system, heating device, electronic lock, louvers control system, gate opening/closing system, window hanging opening/closing system, home appliance, cleaning machine,

vacuum cleaner, oven, microwave oven, refrigerator, washing machine, drying machine, coffee machine, boiler, water heater, telephone, facsimile machine, entertainment system, infotainment system, in-vehicle computer, navigation device, security system, alarm system, air conditioning system, fan, ceiling ventilator, electronic toy, electrical vehicle, any element of “intelligent home” equipment, and so forth.

[0054] The technology described herein provides for easy and effective methods for controlling various electronic devices using any suitable user device through the use of traditional depth sensing camera/video camera. Below are provided description of various embodiments and example of the present technology with reference to the drawings.

#### Examples of Human-Computer Interface

[0055] With reference now to the drawings, FIG. 1 shows an example system environment 100 for providing a real time human-computer interface. The system environment 100 includes a control system 110, a display device 120, and an entertainment system 130.

[0056] The control system 110 is configured to capture various user gestures/motions and user inputs, interpret them, and generate corresponding control commands, which are further transmitted to the entertainment system 130. Once the entertainment system 130 receives commands generated by the control system 110, the entertainment system 130 performs certain actions depending on which software application is running. For example, the user may point a user device towards the entertainment system 130 and press a dedicated button to activate it or adjust its operation. This may be identified by the control system 110 and a corresponding command can be generated for the entertainment system 130. Similarly, the

user may point the user device towards the display device 120 to control its operation.

[0057] The entertainment system 130 may refer to any electronic device such as a computer (e.g., a laptop computer, desktop computer, tablet  
5 computer, workstation, server), game console, television (TV) set, TV adapter, smart television system, audio system, video system, cellular phone, smart phone, and so forth. Although the figure shows that the control system 110 and the entertainment system 130 are separate and stand-alone devices, in some alternative embodiments, these systems can be integrated within a  
10 single device.

[0058] FIG. 2 is a general illustration of a scene 200 suitable for controlling an electronic device. In particular, this figure shows a user 210 interacting with the control system 110 with the help of a user device 220. The control system 110 may include a depth sensing camera (and/or video  
15 camera), a computing unit, and a communication unit, which can be stand-alone devices or embedded within a single housing (as shown). Generally speaking, the user and a corresponding environment, such as a living room, are located, at least in part, within the field of view of the depth sensing camera.

[0059] The control system 110 may be configured to dynamically capture depth maps of the scene in real time and further process the depth maps to identify the user, its body parts/limbs, user head, determine one or more user gestures/motions, determine location and orientation of the user device 220 based on data received from the user device, etc. The control  
20 system 110 may also determine if the user holds the user device 220 in one of the hands, and determine the motions of the user device 220. The control system 110 may also determine specific motion data associated with user gestures/motions, wherein the motion data may include coordinates, velocity  
25

and acceleration of the user's hands or arms. The user gestures/motions may be represented as a set of coordinates on a 3D coordinate system (also referred herein to as the first coordinate system) which result from the processing of the depth map. For this purpose, the control system 110 may generate a  
5 virtual skeleton of the user as shown in FIG. 4A, 4B and described below in greater details.

[0060] The control system 110 may be also configured to receive user inputs made through the user device 220 and generate corresponding control commands based upon determination of user device orientation and  
10 the user inputs.

[0061] As discussed above, the user device 220 may refer to a pointing device, controller wand, remote control device, a gaming console remote controller, game pad, smart phone, cellular phone, PDA, tablet computer, head-mountable computing device, or alike. In certain  
15 embodiments, the user device 220 may also refer to non-electronic devices such as sports implements equipped with ad hoc motion and orientation sensors, key pad, communication device, etc.

[0062] In general, the user device 220 is configured to generate motion and orientation data, which may include acceleration data and  
20 rotational data associated with an internal coordinate system, with the help of embedded or removably attached acceleration sensors, gyroscopes, magnetometers, or other motion and orientation detectors. The user device 220, however, may not determine its exact location within the scene and the 3D coordinate system associated with the control system 110. The motion and  
25 orientation data of the user device 220 can be transmitted to the control system 110 over a wireless or wired network for further processing. In addition, the user may make an input via the user device which may be also

transmitted to the control system 110 over a wireless or wired network for further processing.

[0063] When the control system 110 receives the motion data and orientation data from the user device 220, it may calibrate the user device motion data and user device orientation data with the 3D coordinate system used in the control system 110 by transforming these data using calibration, applying correlation matrices, scaling, or other methods. The transformed user device motion data (which is also referred to as "second motion data") is then compared (mapped) to the motion data derived from the depth map (which is also referred to as "first motion data"). By the result of this comparison, the control system 110 may compare the motions of the user device 220 and the gestures/motions of a user's hands/arms. When these motions match each other or somehow correlate with or are similar to each other, the control system 110 acknowledges that the user device 220 is held in a particular hand of the user, and assigns coordinates of the user's hand to the user device 220. In addition to that, the control system 110 determines the orientation of user device 220 on the 3D coordinate system by processing the orientation data obtained from the user device 220 and optionally from the processed depth maps.

[0064] In various embodiments, this technology can be used for determining that the user device 220 is in "active use," which means that the user device 220 is held or worn by the user 210 who is located in the sensitive area of the depth sensing camera/video camera. In contrast, the technology can be used for determining that the user device 220 is in "inactive use," which means that the user device 220 is not held/worn by the user 210, or that it is held by a user 210 who is not located in the sensitive area of the depth sensing camera.

[0065] FIG. 3 is an illustration of a scene 300 suitable for controlling multiple electronic devices using a user device 220. As shown in this figure, the user 210 interacting with the control system 110 with the help of the user device 220, which is a smart phone in this example. The control system 110  
5 may be in communication with various controllable electronic devices present in the scene 300, which include a lamp 310, speakers 320 of an audio system (not shown), a game console 330, a video system 340, display device 120 (e.g., a television device). The control system 110 may be aware of location of each of these electronic devices by storing their coordinates associated with a 3D  
10 coordinate system used by the depth sensing camera and the control system 110. It should be noted that the electronic devices may be within the field of view of the depth sensing camera, or may be out of the field of view of the depth sensing camera.

[0066] According to various embodiments of the present disclosure,  
15 the user may control any of these electronic devices 120, 310-340, by merely orienting (pointing) the user device 220 towards a desired electronic device. When it's determined that the user orients the user device 220 towards a desired electronic device 120, 310-340, the control system 110 may cause the user device 220 to display a specific graphical user interface including one or  
20 more actionable buttons. By activating one of these actionable buttons, the user makes an input which is transmitted to the control system 110. Upon receipt of the user input, the control system 110 generates a corresponding command for the electronic device, which the user points at. In this regard, the user 210 may, for example, point the user device 220 towards the lamp  
25 310, press a dedicated button on the user device 220, causing thereby turning on or turning off the lamp 310. Similarly, the user may control any other electronic device 120, 310-340.

[0067] In certain embodiments, the user input may be generated by not prompting the user to make an input via the user device 220, but by tracking and identifying a specific user gesture. For example, to control a specific electronic device, the user may point the user device 220 towards the  
5 desired electronic device, and then make a hand or arm gesture, or head gesture (e.g., a nod motion).

[0068] In certain other embodiments, the user input may relate to a user voice command. For example, to control a specific electronic device, the user may point the user device 220 towards the desired electronic device 120,  
10 310-340, and then make say a voice command, e.g., "Turn on" or "Activate".

#### Virtual Skeleton Representation

[0069] FIG. 4A shows a simplified view of an exemplary virtual skeleton 400 as can be generated by the control system 110 based upon the  
15 depth map. As shown in the figure, the virtual skeleton 400 comprises a plurality of virtual "bones" and "joints" 410 interconnecting the bones. The bones and joints, in combination, represent the user 210 in real time so that every motion of the user's limbs is represented by corresponding motions of the bones and joints.

20 [0070] According to various embodiments, each of the joints 410 may be associated with certain coordinates in the 3D coordinate system defining its exact location. Hence, any motion of the user's limbs, such as an arm, may be interpreted by a plurality of coordinates or coordinate vectors related to the corresponding joint(s) 410. By tracking user motions via the  
25 virtual skeleton model, motion data can be generated for every limb movement. This motion data may include exact coordinates per period of time, velocity, direction, acceleration, and so forth.

[0071] FIG. 4B shows a simplified view of exemplary virtual

skeleton 400 associated with the user 210 holding the user device 220. In particular, when the control system 110 determines that the user 210 holds and the user device 220 and then determines the location (coordinates) of the user device 220, a corresponding mark or label can be associated with the  
5 virtual skeleton 400.

[0072] According to various embodiments, the control system 110 can determine an orientation of the user device 220. More specifically, the orientation of the user device 220 may be determined by one or more sensors of the user device 220 and then transmitted to the control system 110 for  
10 further processing and representation in the 3D coordinate system. In this case, the orientation of user device 220 may be represented as a vector 320 as shown in FIG. 3B.

#### Control System

15 [0073] FIG. 5 shows an environment 500 suitable for implementing methods for controlling various electronic devices 120, 310-340 utilizing a user device 220. As shown in this figure, there is provided the control system 110, which may comprise at least one depth sensing camera 510 configured to capture a depth map. The term "depth map," as used herein, refers to an  
20 image or image channel that contains information relating to the distance of the surfaces of scene objects from a depth sensing camera 510. In various embodiments, the depth sensing camera 510 may include an infrared (IR) projector to generate modulated light, and an IR camera to capture 3D images. Alternatively, the depth sensing camera 510 may include two digital  
25 stereo cameras enabling it to generate a depth map. In yet additional embodiments, the depth sensing camera 510 may include time-of-flight sensors or integrated digital video cameras together with depth sensors.

[0074] In some example embodiments, the control system 110 may

optionally include a color video camera 520 to capture a series of 2D images in addition to 3D imagery already created by the depth sensing camera 510. The series of 2D images captured by the color video camera 520 may be used to facilitate identification of the user, and/or various gestures of the user on the depth map. In yet more embodiments, the only color video camera 520 can be used, and not the depth sensing camera 510. It should also be noted that the depth sensing camera 510 and the color video camera 520 can be either stand alone devices or be encased within a single housing.

[0075] Furthermore, the control system 110 may also comprise a computing unit 530 for processing depth map data and generating control commands for one or more electronic devices 560 (e.g., the electronic devices 120, 310-340) as described herein. The computing unit 530 is also configured to implement steps of particular methods for determining a location and orientation of the user device 220 as described herein.

[0076] In certain embodiments, the control system 110 may also include at least one motion sensor 570 such as a movement detector, accelerometer, gyroscope, or alike. The motion sensor 570 may determine whether or not the control system 110 is moved or differently oriented by the user. If it is determined that the control system 110 or its elements are moved, then a new calibration or virtual binding process may be required. In other words, once the control system 110 or its elements are moved, the 3D coordinate system and the internal coordinate system of the user device are disoriented from each other and new calibration matrices are required for proper calibrating the user device motion data with the 3D coordinate system. In certain embodiments, when the depth sensing camera 510 and/or the color video camera 520 are separate devices not present in a single housing with other elements of the control system 110, the depth sensing camera 510 and/or the color video camera 520 may include internal motion sensors 570.

[0077] The control system 110 also includes a communication module 540 configured to communicate with the user device 220 and one or more electronic devices 560. More specifically, the communication module 540 may be configured to wirelessly receive motion and orientation data from the user device 220 and transmit control commands to one or more electronic devices 560 via a wired or wireless network. The control system 110 may also include a bus 550 interconnecting the depth sensing camera 510, color video camera 520, computing unit 530, communication module 540, and optional motion sensor 570. Those skilled in the art will understand that the control system 110 may include other modules or elements, such as a power module, user interface, housing, control key pad, memory, etc., but these nodules and elements are not shown not to burden the description of the present technology.

[0078] Any of the aforementioned electronic devices 560 can refer, in general, to any electronic device configured to trigger one or more predefined actions upon receipt of a certain control command. Some examples of electronic devices 560 include, but are not limited to, computers (e.g., laptop computers, tablet computers), displays, audio systems, video systems, gaming consoles, entertainment systems, lighting devices, cellular phones, smart phones, home appliances, and so forth.

[0079] The communication between the communication module 540 and the user device 220 and/or one or more electronic devices 560 can be performed via a network (not shown). The network can be a wireless or wired network, or a combination thereof. For example, the network may include the Internet, local intranet, PAN (Personal Area Network), LAN (Local Area Network), WAN (Wide Area Network), MAN (Metropolitan Area Network), virtual private network (VPN), storage area network (SAN), frame relay connection, Advanced Intelligent Network (AIN) connection,

synchronous optical network (SONET) connection, digital T1, T3, E1 or E3 line, Digital Data Service (DDS) connection, DSL (Digital Subscriber Line) connection, Ethernet connection, ISDN (Integrated Services Digital Network) line, dial-up port such as a V.90, V.34 or V.34bis analog modem connection, 5 cable modem, ATM (Asynchronous Transfer Mode) connection, or an FDDI (Fiber Distributed Data Interface) or CDDI (Copper Distributed Data Interface) connection. Furthermore, communications may also include links to any of a variety of wireless networks including WAP (Wireless Application Protocol), GPRS (General Packet Radio Service), GSM (Global System for 10 Mobile Communication), CDMA (Code Division Multiple Access) or TDMA (Time Division Multiple Access), cellular phone networks, Global Positioning System (GPS), CDPD (cellular digital packet data), RIM (Research in Motion, Limited) duplex paging network, Bluetooth radio, or an IEEE 802.11-based radio frequency network. The network can further include or interface with 15 any one or more of the following: RS-232 serial connection, IEEE-1394 (Firewire) connection, Fiber Channel connection, IrDA (infrared) port, SCSI (Small Computer Systems Interface) connection, USB (Universal Serial Bus) connection, or other wired or wireless, digital or analog interface or connection, mesh or Digi® networking.

20

### User Device

[0080] FIG. 6 shows a simplified diagram of the user device 220 according to an example embodiment. As shown in the figure, the user device 220 comprises one or more motion and orientation sensors 610, as well 25 as a wireless communication module 620. In various alternative embodiments, the user device 220 may include additional modules (not shown), such as an input module, a computing module, display, touchscreen, and/or any other modules, depending on the type of the user device 220

involved.

[0081] The motion and orientation sensors 610 may include gyroscopes, magnetometers, accelerometers, and so forth. In general, the motion and orientation sensors 610 are configured to determine motion and orientation data which may include acceleration data and rotational data (e.g., an attitude quaternion), both associated with an internal coordinate system. In operation, motion and orientation data is then transmitted to the control system 110 with the help of the communication module 620. The motion and orientation data can be transmitted via the network as described above.

10 [0082] FIG. 7 shows an exemplary graphical user interface (GUI) 700 that may be displayed on a touchscreen of the user device, when it is oriented towards a specific electronic device 560. In certain embodiments, the GUI 700 may be generated by a control command received from the control system 110.

15 [0083] In particular, the GUI 700 may include an information box 710 to display information regarding an electronic device 560 which the user device points at. For example, if the user device 220 is oriented by the user towards to a lamp 310, the information box 710 may display a corresponding message, such as one or more of a text message (e.g., "You point at the lamp"), an image message or an animated message.

[0084] The GUI 700 may also include one or more actionable buttons 720 for controlling the electronic device 560. For example, there may be actionable buttons to turn on or turn off the electronic device 560, or there may be actionable buttons to adjust its operation (e.g., lighting power). In certain embodiments, the GUI 700 may be generated by a control command received from the control system 110.

### Examples of Operation

[0085] FIG. 8 is a process flow diagram showing an example method 800 for controlling one or more electronic devices 560 using a user device 220. The method 800 may be performed by processing logic that may  
5 comprise hardware (e.g., dedicated logic, programmable logic, and microcode), software (such as software run on a general-purpose computer system or a dedicated machine), or a combination of both. In one example embodiment, the processing logic resides at the control system 110.

[0086] The method 800 can be performed by the units/devices  
10 discussed above with reference to FIG. 5. Each of these units or devices may comprise processing logic. It will be appreciated by one of ordinary skill in the art that examples of the foregoing units/devices may be virtual, and instructions said to be executed by a unit/device may in fact be retrieved and executed by a processor. The foregoing units/devices may also include  
15 memory cards, servers, and/or computer discs. Although various modules may be configured to perform some or all of the various steps described herein, fewer or more units may be provided and still fall within the scope of example embodiments.

[0087] As shown in FIG. 8, the method 800 may commence at  
20 operation 805, with the depth sensing camera 510 generating a depth map by capturing a plurality of depth values of scene in real time. The depth map may be associated with or include a 3D coordinate system (i.e., a first coordinate system) such that all identified objects within the scene may have particular coordinates. The depth map may then be transmitted to the  
25 computing unit 530.

[0088] At operation 810, the communication unit 540 may receive user device motion data and user device orientation data from the user device 220. At operation 815, the computing unit 530 may process the depth map(s)

and the user device motion data and user device orientation data from the user device 220. By the result of the processing, the computing unit 530 may determine a location and orientation of the device within the 3D coordinate system (and not internal coordinate system of the user device).

5 [0089] At operation 820, the computing unit 530 may determine that the user device 220 is oriented towards a predetermined direction, i.e. towards a predetermined location of one of a plurality of electronic devices 560. For this purpose, a database may be maintained storing locations of the electronic devices 560. Those skilled in the art should understand that the  
10 locations of the electronic devices 560 may be either manually input by user during setup of the control system 110 or the depth sensing camera 510 may automatically determine or assist in determining the locations of the electronic devices 560.

[0090] At operation 825, the communication unit 540 may receive a  
15 user input message from the user device 220. The user input message may include a user command based upon activation of one of the actionable buttons 720 or a user voice command or user control gesture/motion.

[0091] At operation 830, the computing unit 530 may generate a control command for the electronic device 560, which the user points at using  
20 the hand held device 220. The control command may be based on the determination made at the operation 820 and the user input message received at the operation 825. Further, the control command may be transmitted to the electronic device 560 utilizing the communication unit 540. At the following operations (not shown), the electronic device 560 makes a corresponding  
25 action or an operation.

[0092] FIG. 9 is a process flow diagram showing another example method 900 for controlling one or more electronic devices 560 using a user device 220. The method 900 may be performed by processing logic that may

comprise hardware (e.g., dedicated logic, programmable logic, and microcode), software (such as software run on a general-purpose computer system or a dedicated machine), or a combination of both. In one example embodiment, the processing logic resides at the control system 110. The method 900 can be performed by the units/devices discussed above with reference to FIG. 5.

[0093] As shown in FIG. 9, the method 900 may commence at operation 905, with the depth sensing camera 510 and/or color video camera 520 generating a depth map by capturing a plurality of depth values of scene in real time. The depth map may be associated with or include a first coordinate system such that all identified objects within the scene may have particular coordinates.

[0094] At operation 910, the depth map can be processed by the computing unit 530 to identify the user 210 on the depth map, user hands, user head, identify a motion of at least one user hand or head, and generate corresponding "first motion data" of the identified user motion. The first motion data may include a plurality coordinates associated with a virtual skeleton corresponded to the first coordinate system.

[0095] At operation 915, the computing unit 530 acquires user device motion data and user device orientation data from the user device 220 via the communication module 440. These user device motion data and user device orientation data are corresponded to a second coordinate system of the user device 220. In certain embodiments, the first coordinate system differs from the second coordinate system.

[0096] At operation 920, the computing unit 430 may optionally transform (calibrate) the user device motion data and/or the user device orientation data so as this data relates to the first coordinate system. In other words, any coordinates of the second coordinate system are modified into

corresponding coordinates of the first coordinate system. Either transformed or not, this data is now referred to as "second motion data."

[0097] According to certain embodiments, the transformation of the user device motion data, if performed, may be performed by the computing unit 530 using the user device orientation data and optionally correlation parameters/matrices and/or calibration parameters/matrices so that the user device motion data corresponds to the first coordinate system and not to the second coordinate system of the user device 220. In an example embodiment, the user device motion data is multiplied by a predetermined correlation (calibration) matrix and a current rotation matrix, where the current rotation matrix is defined by the user device orientation data, while the predetermined correlation (calibration) matrix may define correlation between two coordinate systems. As a result of multiplication, the transformed user device motion data (which is also referred herein to "second motion data") is meets the first coordinate system.

[0098] At operation 925, the computing unit 530 compares the second motion data to the first motion data. If the first and second motion data correspond (or match or they are relatively similar) to each other, the computing unit 530 selectively assigns the coordinates of the user's hand to the user device 220. Thus, the precise location and orientation of user device 220 is determined within the first coordinate system.

[0099] Once the orientation of the user device 220 is known, at operation 930, the computing unit 530 may determine that the user device 220 is oriented towards a predetermined direction, i.e. towards a predetermined location of one of a plurality of electronic devices 560. For this purpose, a database may be maintained storing locations of the electronic devices 560.

[00100] At operation 935, the communication unit 540 may receive a user input message from the user device 220. The user input message may

include a user command based upon activation of one of the actionable buttons 720 or a user voice command or user control gesture/motion.

[00101] At operation 940, the computing unit 530 may generate a control command for the electronic device 560, which the user points at using the hand held device 220. Further, the control command may be transmitted to the electronic device 560 utilizing the communication unit 540. At the following operations (not shown), the electronic device 560 makes a corresponding action or an operation.

[00102] In various embodiments, the described technology can be used for determining that the user device 220 is in active use by the user 210. As mentioned earlier, the term "active use" means that the user 210 is identified on the depth map or, in other words, is located within the viewing area of depth sensing camera 510 when the user device 220 is moved.

[00103] FIG. 10 is a process flow diagram showing yet another example method 1000 for controlling one or more electronic devices 560 using a user device 220. The method 1000 may be performed by processing logic that may comprise hardware (e.g., dedicated logic, programmable logic, and microcode), software (such as software run on a general-purpose computer system or a dedicated machine), or a combination of both. In one example embodiment, the processing logic resides at the control system 110.

[00104] As shown in FIG. 10, the method 1000 may commence at operation 1005, with the depth sensing camera 510 and/or color video camera 520 generating a depth map by capturing a plurality of depth values of scene in real time. The depth map may be associated with or include a first coordinate system such that all identified objects within the scene may have particular coordinates.

[00105] At operation 1010, the depth map can be processed by the computing unit 530 to identify the user 210 on the depth map, user hands,

identify a motion of at least one user hand or head, and generate corresponding "first motion data" of the identified user motion. The first motion data may include a plurality coordinates associated with a virtual skeleton corresponded to the first coordinate system.

5 [00106] At operation 1015, the computing unit 530 acquires user device motion data and user device orientation data from the user device 220 via the communication module 440. These user device motion data and user device orientation data are associated with a second coordinate system of the user device 220.

10 [00107] At operation 1020, the computing unit 430 optionally transforms (calibrates) the user device motion data so as to generate related motion data (i.e., series of coordinates) but associated with the first coordinate system. The transformed user device motion data or non-transformed user device motion data is now referred to as "second motion data". In other  
15 words, the user device motion data are recalculated such that it "fits" said first coordinate system. The transformation (calibration) process may be performed by the computing unit 530 using the user device orientation data and optionally correlation parameters/matrices and/or calibration parameters/matrices so that the user device motion data corresponds to the  
20 first coordinate system and not to the second coordinate system of the user device 220. In an example embodiment, the user device motion data is multiplied by a predetermined correlation (calibration) matrix and a current rotation matrix, where the current rotation matrix is defined by the user device orientation data, while the predetermined correlation (calibration)  
25 matrix may define correlation between two coordinate systems. As a result of multiplication, the transformed user device motion data is associated with the first coordinate system.

[00108] At operation 1025, the computing unit 530 compares the

second motion data to the first motion data. If the first and second motion data correspond (or match or are relatively similar) to each other, the computing unit 530 selectively assigns the coordinates of the user's hand to the user device 220. Thus, the precise location and orientation of user device  
5 220 is determined within the first coordinate system.

[00109]           Once the orientation of the user device 220 is known, at operation 1030, the computing unit 530 may determine that the user device 220 is oriented towards a predetermined direction, i.e. towards a predetermined location of one of a plurality of electronic devices 560. For this  
10 purpose, a database may be maintained storing locations of the electronic devices 560.

[00110]           At operation 1035, the communication unit 540 may track user gestures and identify that the user 210 makes a predetermined gesture with a hand, arm, body, or hand.

15 [00111]           At operation 1040, the computing unit 530 may generate a control command for the electronic device 560, which the user points at using the hand held device 220, based on the identified user gesture. Further, the control command may be transmitted to the electronic device 560 utilizing the communication unit 540. At the following operations (not shown), the  
20 electronic device 560 makes a corresponding action or an operation.

### Example of Computing Device

[00112]           FIG. 11 shows a diagrammatic representation of a computing device for a machine in the example electronic form of a computer system  
25 1100, within which a set of instructions for causing the machine to perform any one or more of the methodologies discussed herein can be executed. In example embodiments, the machine operates as a standalone device, or can be connected (e.g., networked) to other machines. In a networked deployment,

the machine can operate in the capacity of a server, a client machine in a server-client network environment, or as a peer machine in a peer-to-peer (or distributed) network environment. The machine can be a personal computer (PC), tablet PC, STB, PDA, cellular telephone, portable music player (e.g., a portable hard drive audio device, such as a Moving Picture Experts Group Audio Layer 3 (MP3) player), web appliance, network router, switch, bridge, or any machine capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that separately or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies discussed herein.

[00113] The example computer system 1100 includes one or more processors 1102 (e.g., a central processing unit (CPU), graphics processing unit (GPU), or both), main memory 1104, and static memory 1106, which communicate with each other via a bus 1108. The computer system 1100 can further include a video display unit 1110 (e.g., a liquid crystal display (LCD) or cathode ray tube (CRT)). The computer system 1100 also includes at least one input device 1112, such as an alphanumeric input device (e.g., a keyboard), cursor control device (e.g., a mouse), microphone, digital camera, video camera, and so forth. The computer system 1100 also includes a disk drive unit 1114, signal generation device 1116 (e.g., a speaker), and network interface device 1118.

[00114] The disk drive unit 1114 includes a computer-readable medium 1120 that stores one or more sets of instructions and data structures (e.g., instructions 1122) embodying or utilized by any one or more of the methodologies or functions described herein. The instructions 1122 can also reside, completely or at least partially, within the main memory 1104 and/or

within the processors 1102 during execution by the computer system 1100. The main memory 1104 and the processors 1102 also constitute machine-readable media.

[00115] The instructions 1122 can further be transmitted or received  
5 over the network 1124 via the network interface device 1118 utilizing any one of a number of well-known transfer protocols (e.g., Hyper Text Transfer Protocol (HTTP), CAN, Serial, and Modbus).

[00116] While the computer-readable medium 1120 is shown in an example embodiment to be a single medium, the term "computer-readable  
10 medium" should be understood to include either a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers), either of which store the one or more sets of instructions. The term "computer-readable medium" shall also be understood to include any medium that is capable of storing, encoding, or carrying a set of  
15 instructions for execution by the machine, and that causes the machine to perform any one or more of the methodologies of the present application. The "computer-readable medium may also be capable of storing, encoding, or carrying data structures utilized by or associated with such a set of instructions. The term "computer-readable medium" shall accordingly be  
20 understood to include, but not be limited to, solid-state memories, and optical and magnetic media. Such media may also include, without limitation, hard disks, floppy disks, flash memory cards, digital video disks, random access memory (RAM), read only memory (ROM), and the like.

[00117] The example embodiments described herein may be  
25 implemented in an operating environment comprising computer-executable instructions (e.g., software) installed on a computer, in hardware, or in a combination of software and hardware. The computer-executable instructions may be written in a computer programming language or may be

embodied in firmware logic. If written in a programming language conforming to a recognized standard, such instructions may be executed on a variety of hardware platforms and for interfaces associated with a variety of operating systems. Although not limited thereto, computer software  
5 programs for implementing the present method may be written in any number of suitable programming languages such as, for example, C, C++, C#, .NET, Cobol, Eiffel, Haskell, Visual Basic, Java, JavaScript, or Python, as well as with any other compilers, assemblers, interpreters, or other computer languages or platforms.

10

### Conclusion

[00118] Thus, methods and systems for controlling one or more electronic devices through real time tracking of location and orientation of a user device have been described. Although embodiments have been  
15 described with reference to specific example embodiments, it will be evident that various modifications and changes can be made to these example embodiments without departing from the broader spirit and scope of the present application. Accordingly, the specification and drawings are to be regarded in an illustrative rather than a restrictive sense.

20

## CLAIMS

1. A method for controlling one or more electronic devices using a user device, the method comprising:
- 5           acquiring, by a processor communicatively coupled with a memory, a depth map, wherein the depth map is associated with a first coordinate system;
- processing, by the processor, the depth map to generate first motion data associated with at least one motion of at least one user hand or head of a
- 10          user;
- acquiring, by the processor, user device motion data and user device orientation data associated with at least one motion of the user device;
- generating, by the processor, second motion data based at least in part on the user device motion data and the user device orientation data;
- 15          comparing, by the processor, the first motion data to the second motion data to set coordinates and orientation of the user device within the first coordinate system;
- determining, by the processor, that the user device is oriented towards an electronic device; and
- 20          based at least in part on the determination, generating, by the processor, a control command for the electronic device.
2. The method of claim 1, further comprising:
- receiving, by the processor, a user input message from the user device;
- 25          and
- wherein the generating of the control command for the electronic device is based at least in part on the user input message.

3. The method of claim 2, further comprising:  
based at least in part on the determination, generating, by the processor, a prompting command for prompting the user to use the user device to generate the user input message; and  
5 sending, by the processor, the prompting command to the user device.
4. The method of claim 3, wherein the prompting command is configured to cause the user device to provide a graphical user interface including one or more actionable buttons.  
10
5. The method of claim 4, wherein the one or more actionable buttons, when activated by the user, are configured to cause the user device to generate the user input message.
- 15 6. The method of claim 1, further comprising sending, by the processor, the control command to the electronic device.
7. The method of claim 1, further comprising:  
determining, by the processor, that the user makes one or more  
20 predetermined gestures; and  
wherein the generating of the control command for the electronic device is based at least in part on the determination.
8. The method of claim 7, wherein the one or more predetermined gestures  
25 include at least one of the following: pointing a hand towards the electronic device, orienting a head towards the electronic device, and nodding.

9. The method of claim 1, wherein the comparing of the first motion data to the second motion data is further used to determine that the at least one motion of the user device is correlated with the at least one motion of the at least one user hand or head of the user.

5

10. The method of claim 1, wherein the first motion data includes a set of coordinates associated with the at least one user hand or head.

11. The method of claim 10, wherein the setting of coordinates and orientation of the user device within the first coordinate system includes assigning, by the processor, the set of coordinates associated with the at least one user hand or head to the user device.

12. The method of claim 1, wherein the second motion data is associated with the first coordinate system.

13. The method of claim 1, wherein the user device motion data and the user device orientation data are associated with a second coordinate system, wherein the second coordinate system differs from the first coordinate system.

20

14. The method of claim 13, wherein the generating of the second motion data comprises multiplying, by the processor, the user device motion data by a correlation matrix and a rotation matrix, wherein the rotation matrix is

25 associated with the user device orientation data.

15. The method of claim 1, further comprising determining, by the processor, one or more orientation vectors of the user device within the first coordinate system based at least in part on the user device orientation data.
- 5 16. The method of claim 1, further comprising generating, by the processor, a virtual skeleton of the user, the virtual skeleton comprises at least one virtual joint of the user; wherein the at least one virtual joint of the user is associated with the first coordinate system.
- 10 17. The method of claim 16, wherein the processing of the depth map further comprises determining, by the processor, coordinates of the at least one user hand or head on the first coordinate system, wherein the coordinates of the at least one user hand or head are associated with the virtual skeleton.
- 15 18. The method of claim 16, wherein the setting of the coordinates and orientation of the user device within the first coordinate system includes assigning the coordinates of the at least one user hand or head associated with the virtual skeleton to the user device.
- 20 19. The method of claim 1, wherein the second motion data includes at least acceleration data.
20. The method of claim 1, wherein the user device orientation data includes at least one of: rotational data, calibrated rotational data or an attitude  
25 quaternion associated with the user device.
21. The method of claim 1, wherein the depth map is acquired from at least one depth sensing device.

22. The method of claim 1, wherein the depth map is acquired from at least one video camera.
- 5 23. The method of claim 1, wherein the user device includes a handheld or wearable motion sensing device.
24. The method of claim 1, wherein the user device includes a cellular phone.
- 10 25. The method of claim 1, wherein the user device includes a head-mountable computing device.
26. A method for controlling one or more electronic devices using a user device, the method comprising:
- 15       acquiring, by a processor communicatively coupled with a memory, a depth map, wherein the depth map is associated with a first coordinate system;
- acquiring, by the processor, user device orientation data;
- processing, by the processor, the depth map and the user device
- 20 orientation data to determine the orientation of the user device within the first coordinate system;
- determining, by the processor, that the user device is oriented towards an electronic device;
- receiving, by the processor, a user input; and
- 25       based at least in part on the determination and on the user input, generating, by the processor, a control command and transmitting the same to the electronic device.

27. The method of claim 26, wherein the control command is configured to turn on the electronic device.
28. The method of claim 26, wherein the control command is configured to  
5 turn off the electronic device.
29. The method of claim 26, wherein the control command is configured to adjust operation of the electronic device.
- 10 30. The method of claim 26, wherein the user input message includes instructions on the way how the electronic device is operated.
31. The method of claim 26, further comprising:  
acquiring, by the processor, user device motion data associated with at  
15 least one motion of the user device;  
generating, by the processor, second motion data based at least in part on the user device motion data and the user device orientation data; and  
comparing, by the processor, the first motion data to the second motion data to determine coordinates and orientation of the user device within the  
20 first coordinate system.
32. The method of claim 26, further comprising processing, by the processor, the depth map to determine coordinates of the one or more electronic devices within the first coordinate system.  
25
33. The method of claim 26, further comprising maintaining a database including location(s) of the one or more electronic devices within the first coordinate system.

34. The method of claim 26, wherein the determination of the orientation of the user device within the first coordinate system includes associating coordinates of the at least one user hand or head to the user device.

5

35. The method of claim 26, further comprising prompting, by the processor, the user to identify a location of the user device so as to determine a location of the user device within the first coordinate system.

10 36. The method of claim 26, further comprising prompting, by the processor, the user to place the user device in a predetermined location so as to determine a location of the user device within the first coordinate system.

37. A system for controlling one or more electronic devices using a user  
15 device, the system comprising:  
    a depth sensing device configured to obtain a depth map of a three-dimensional environment within which a user is present;  
    a wireless communication module configured to receive, from the user device, user device motion data and user device orientation data associated  
20 with at least one motion of the user device, and receive a user input; and  
    a computing unit communicatively coupled to the depth sensing device and the wireless communication unit, the computing unit is configured to:  
        identify, on the depth map, a motion of at least one user hand or  
25 head;  
        determine, by processing the depth map, coordinates of the at least one user hand or head on a first coordinate system;

generate first motion data associated with the at least one motion of the user hand or head;

5 generate second motion data based at least on the user device motion data, wherein the second motion data is associated with the first coordinate system;

compare the first motion data and the second motion data so as to determine correlation therebetween;

based on the correlation, determine a location and orientation of the user device on the first coordinate system; and

10 determine that the user device is oriented towards an electronic device; and

based at least in part on the determination that the user device is oriented towards an electronic device and based on the user input, generate a control command for the electronic device.

15

38. The system of claim 37, wherein the user device is selected from a group comprising: an electronic pointing device, a motion sensing device, a cellular phone, a smart phone, a remote controller, a video game console, a video game pad, a handheld game device, a computer, a tablet computer, a sports  
20 implement, a head-mounted computing device, and a head-mounted communication device.

39. The system of claim 37, wherein the depth map is associated with the first coordinate system, and wherein the user device motion data and the user  
25 device orientation data are associated with a second coordinate system, wherein the second coordinate system differs from the first coordinate system.

40. The system of claim 37, wherein the computing unit is further configured to generate a prompting command for prompting the user to use the user device to generate the user input.
- 5 41. The system of claim 37, wherein the computing unit is further configured to cause the user device to display a graphical user interface having one or more actionable buttons, which, when activated by the user, causes generation of the user input.
- 10 42. The system of claim 37, wherein the one or more electronic devices include one or more of the following: a computer, a display, a television device, a gaming console, a printing device, a computer peripheral device, a telephone.
43. The system of claim 37, wherein the one or more electronic devices include  
15 one or more of the following: an audio device, a mini system, a speaker.
44. The system of claim 37, wherein the one or more electronic devices include one or more lighting devices.
- 20 45. The system of claim 37, wherein the one or more electronic devices include one or more home appliances.
46. The system of claim 37, wherein the depth sensing device further includes a video camera.  
25
47. The system of claim 37, further comprising a motion sensor configured to determine movement or disposition of the system.

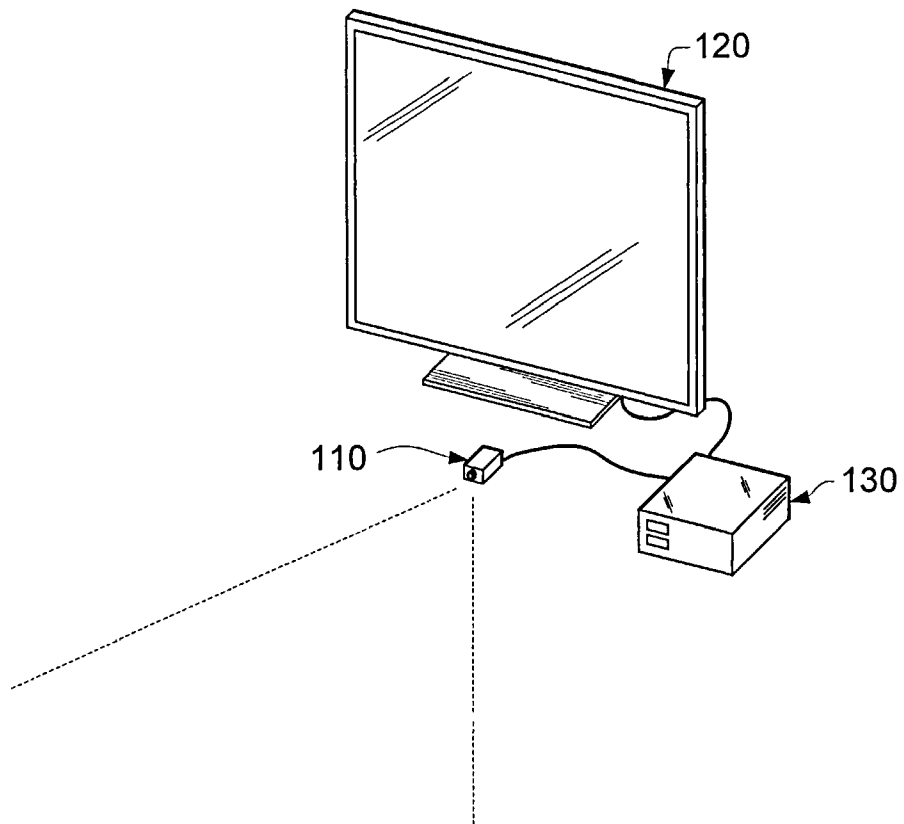
48. A non-transitory processor-readable medium having instructions stored thereon, which when executed by one or more processors, cause the one or more processors to implement a method for controlling one or more electronic devices using a user device, the method comprising:

- 5           acquiring a depth map, wherein the depth map is associated with a first coordinate system;
- processing the depth map to generate first motion data associated with at least one motion of at least one user hand or head of a user;
- acquiring user device motion data and user device orientation data
- 10       associated with at least one motion of the user device;
- generating second motion data based at least in part on the user device motion data and the user device orientation data;
- comparing the first motion data to the second motion data to set coordinates and orientation of the user device within the first coordinate
- 15       system;
- determining that the user device is oriented towards an electronic device; and
- based at least in part on the determination, generating a control command for the electronic device.

20

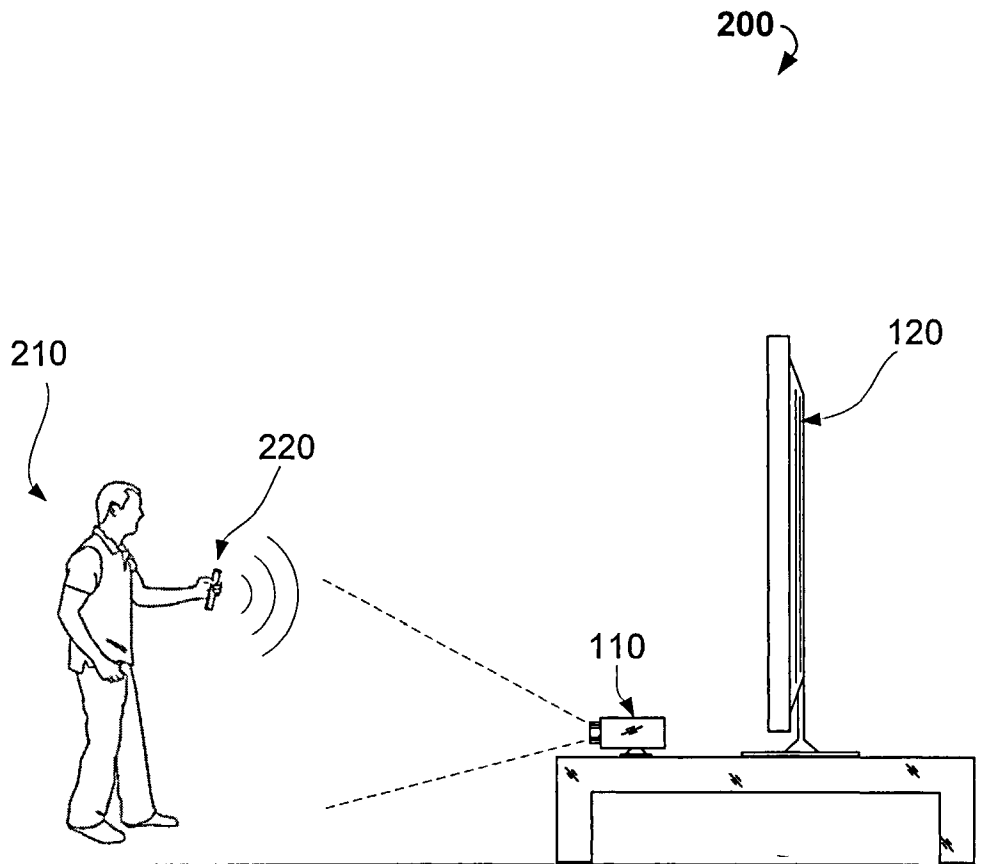
1/11

100



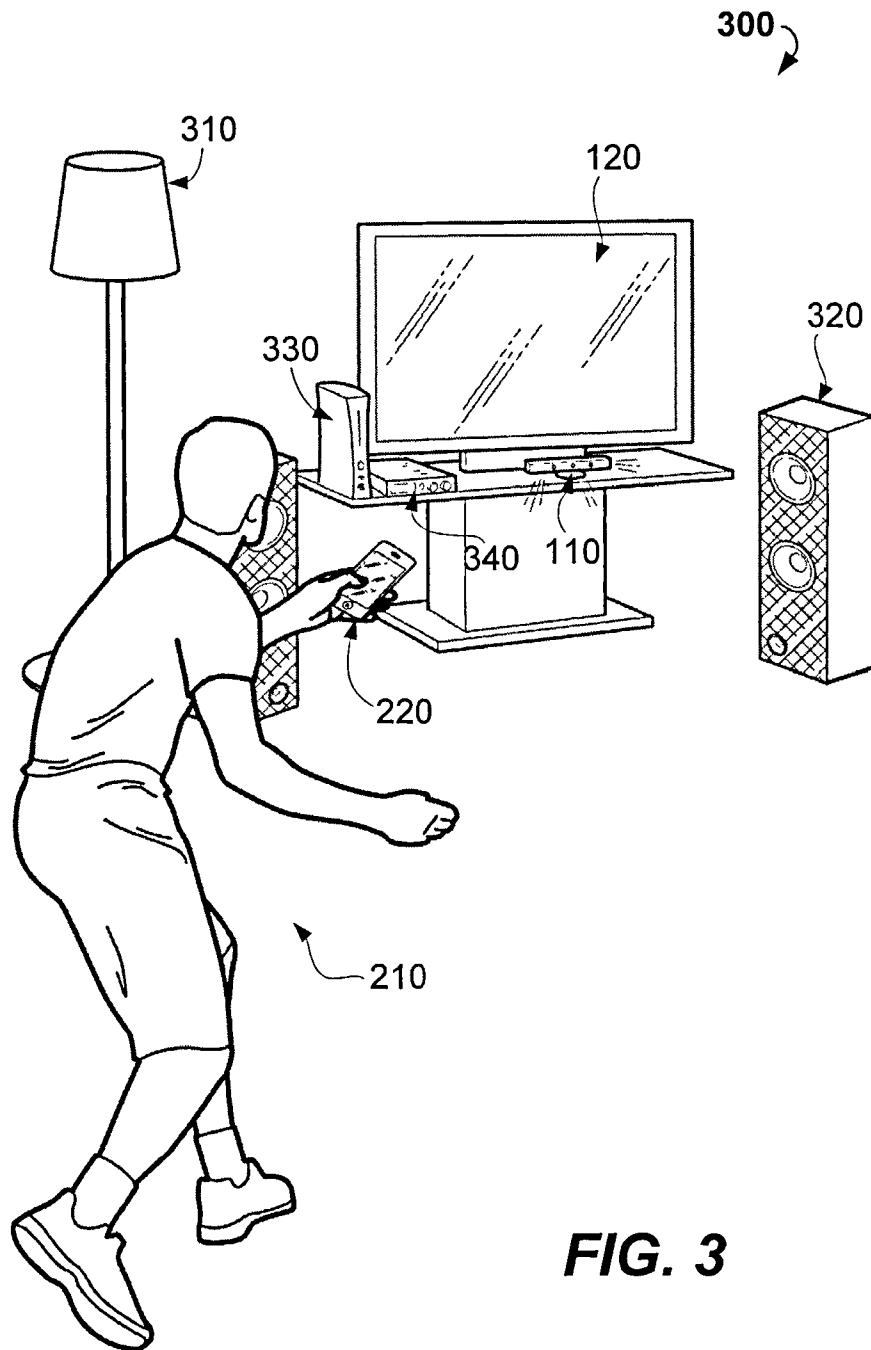
**FIG. 1**

2/11



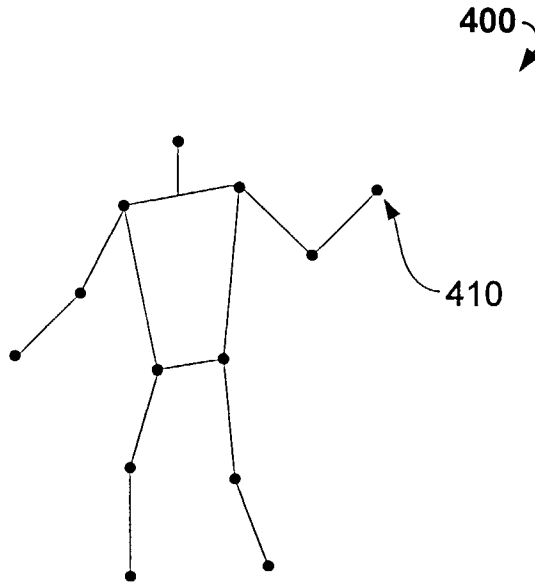
**FIG. 2**

3/11

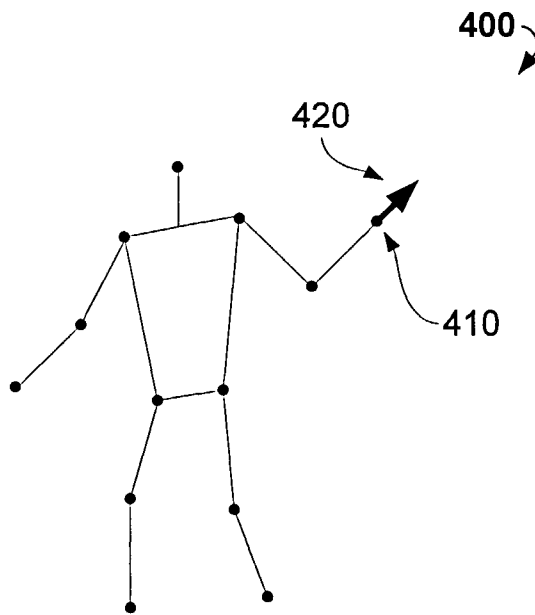


**FIG. 3**

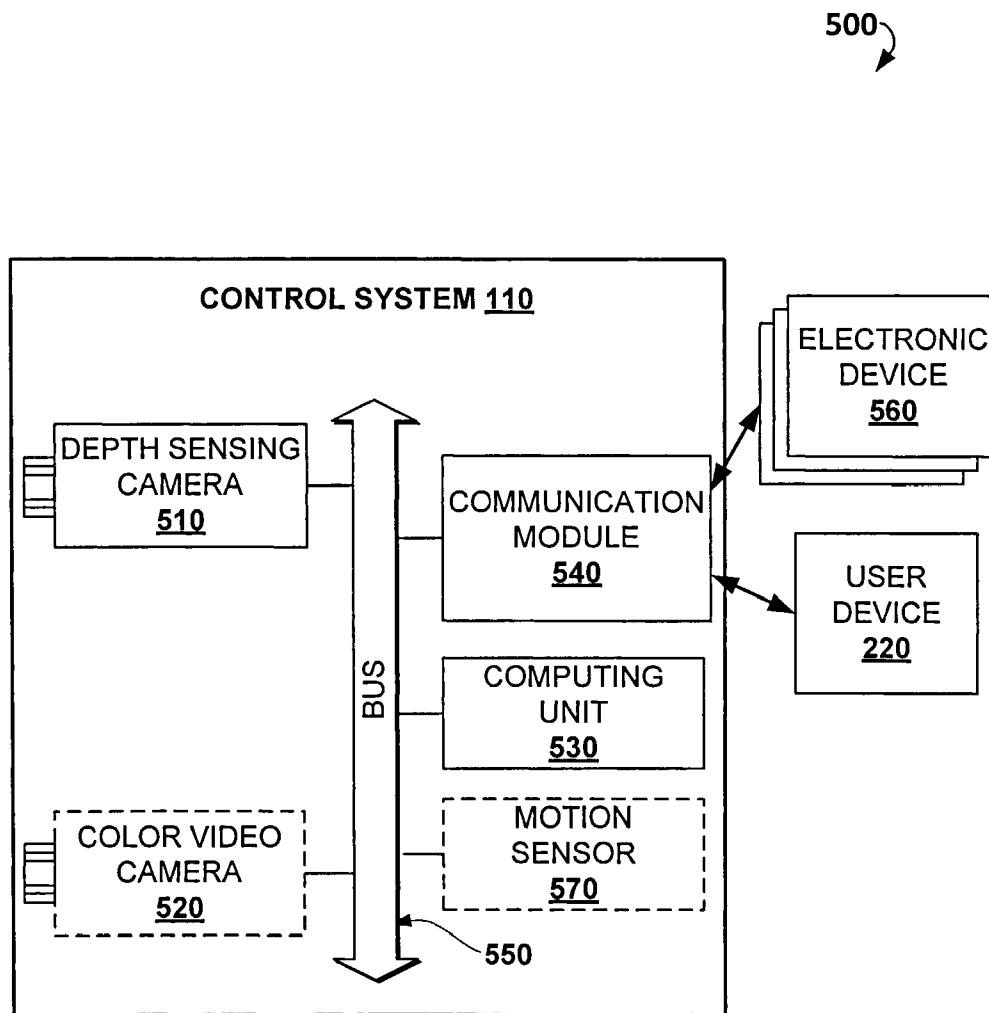
4/11



**FIG. 4A**

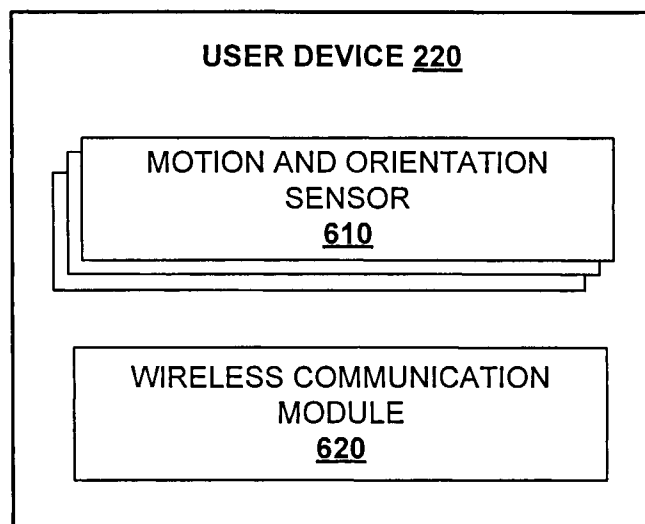


**FIG. 4B**



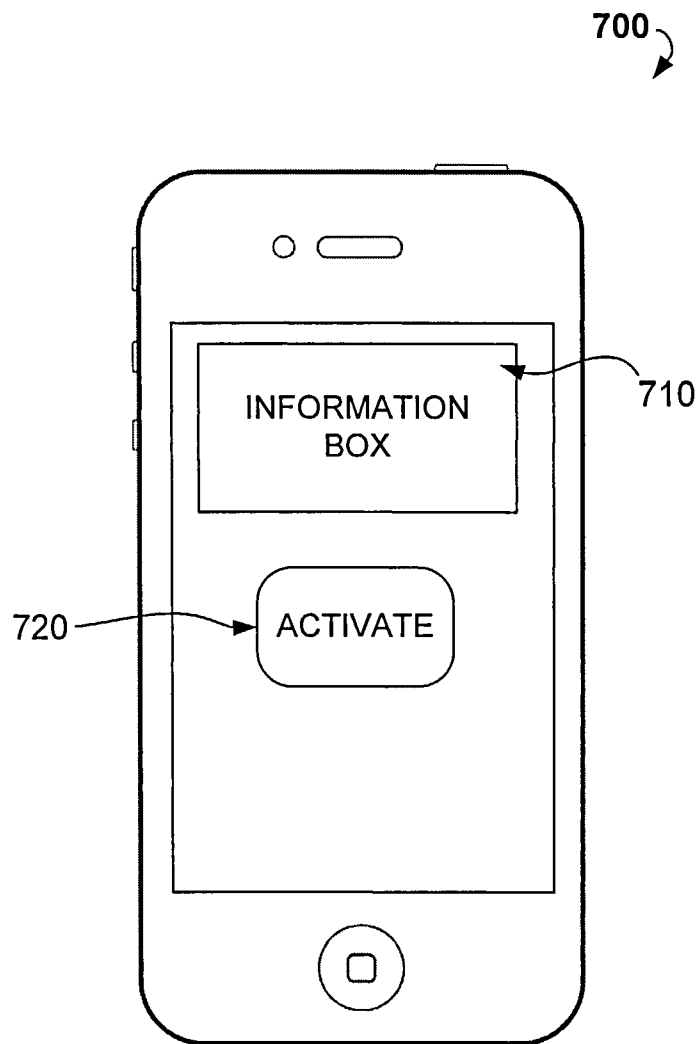
**FIG. 5**

6/11



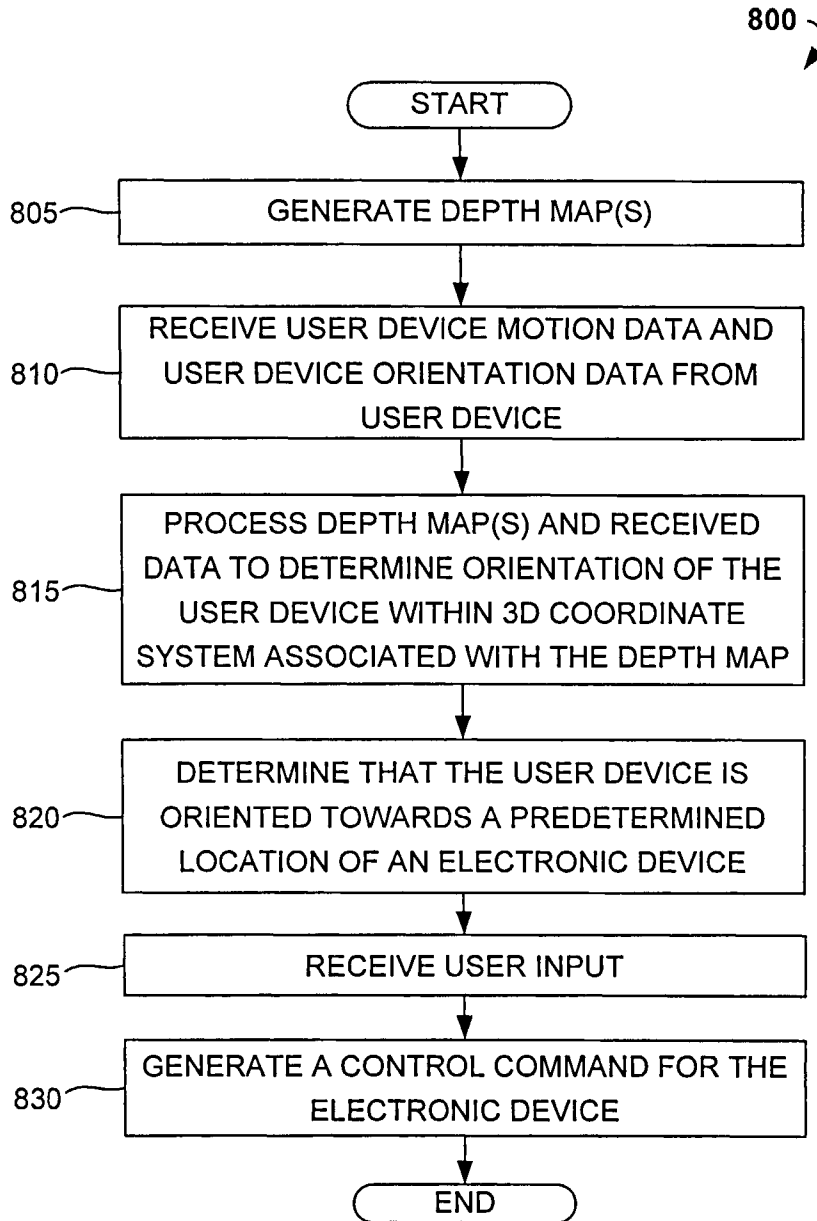
**FIG. 6**

7/11

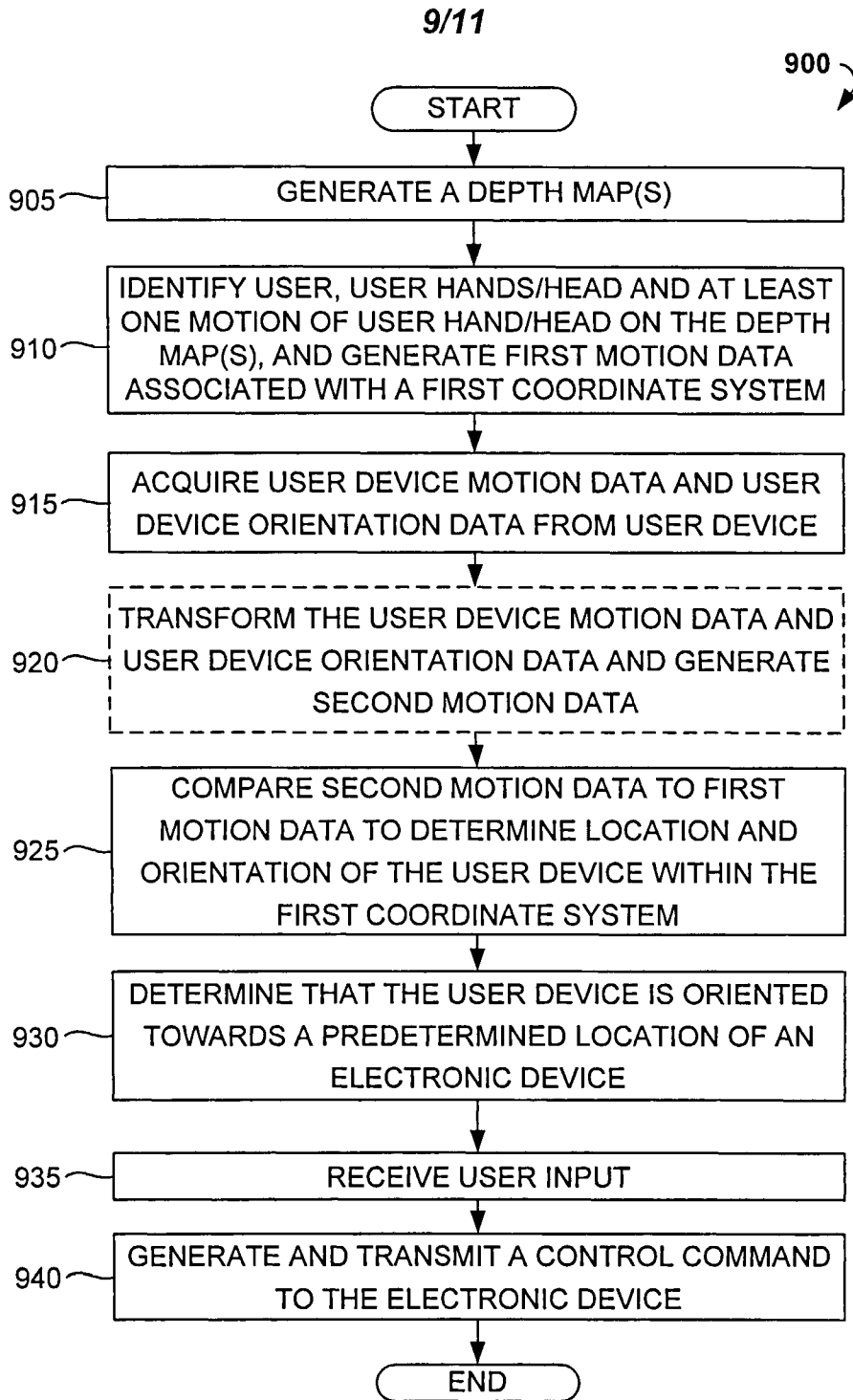


**FIG. 7**

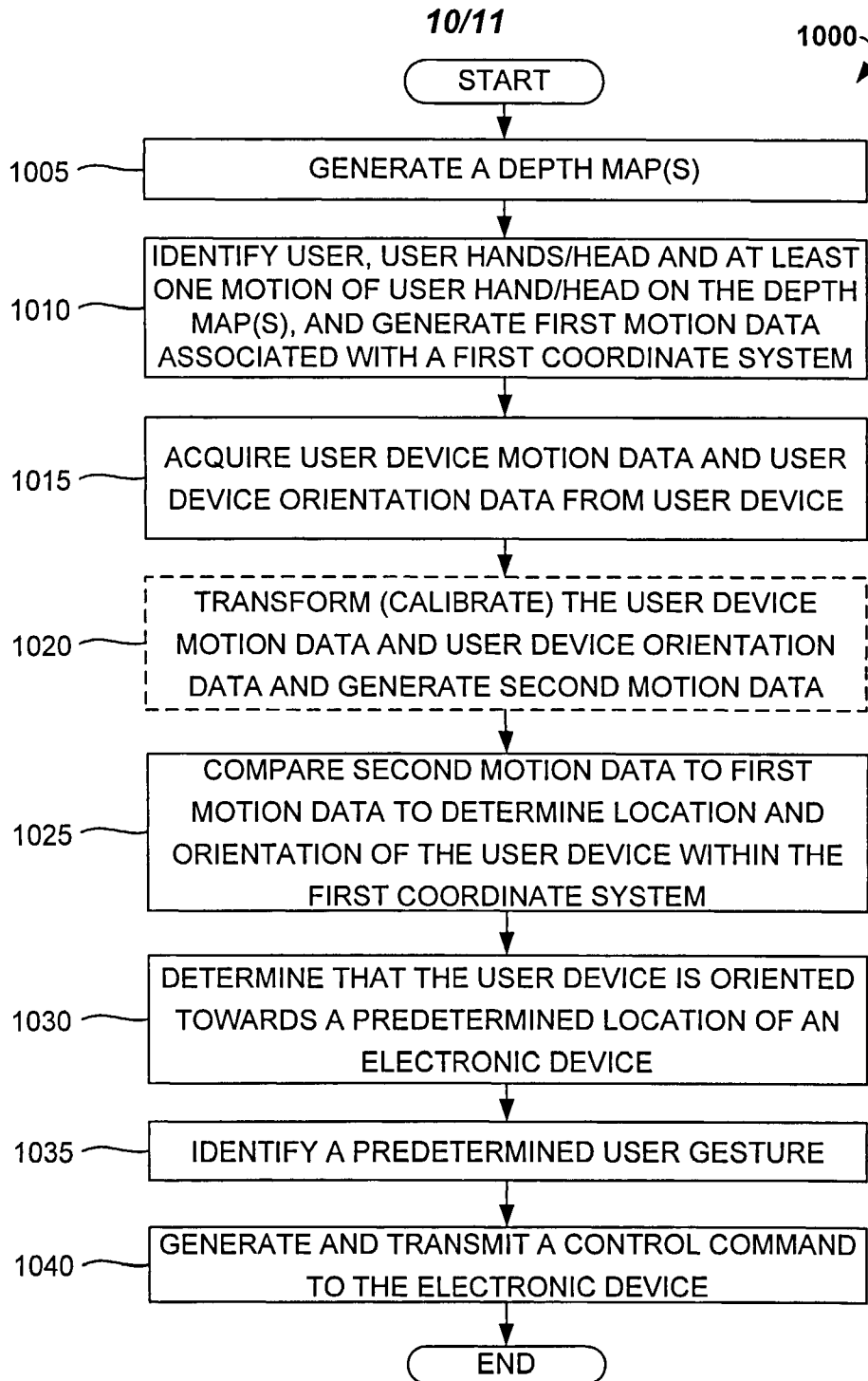
8/11



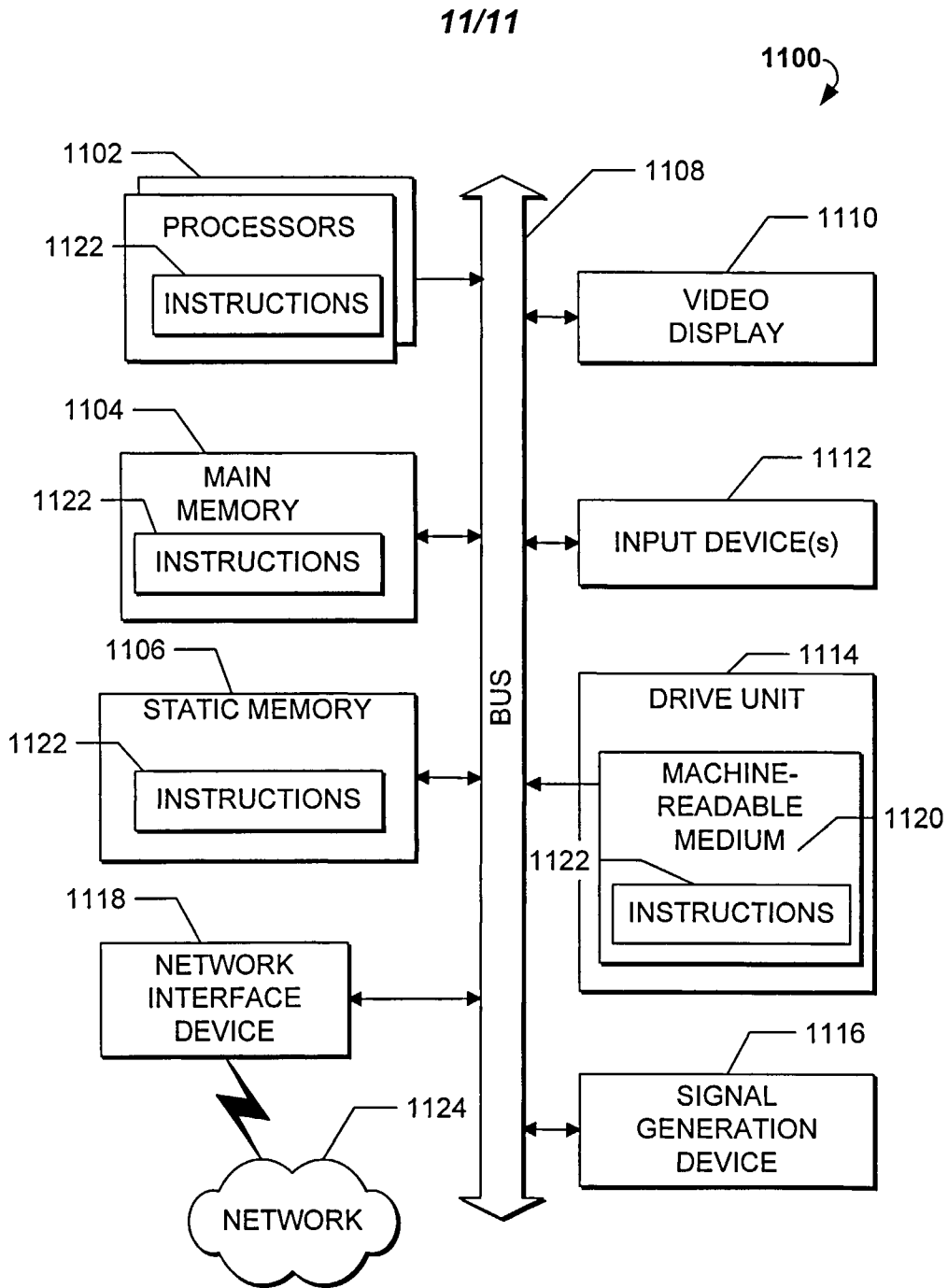
**FIG. 8**



**FIG. 9**



**FIG. 10**



**FIG. 11**

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/RU 2013/000393

A. CLASSIFICATION OF SUBJECT MATTER		
<i>G05B 11/58 (2006.01)</i> <i>G05B 15/00 (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)		
G06K 9/00-9/82, A63F 13/00, G06T 7/00, 7/20, G06F 3/00-3/0488, G06F 17/00, H04N 13/02 G09G 5/00, 5/08		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
PatSearch (RUPTO internal), USPTO, PAJ, Esp@cenet, Information Retrieval System of FIPS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2013/0010071 A1 (3DIVI) 10.01.2013, paragraphs [0009], [0010], [0013], [0014], [0016], [0017], [0020], [0031], [0033], [0034]-[0037], [0039]-[0042], [0045], [0046], [0048]-[0053], [0057], [0058], [0064], [0065]	1-48
A	US 2013/009861 A1 (3DIVI) 10.01.2013	1-48
A	WO 2011/011029 A1 (HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.) 27.01.2011	1-48
A	US 2009/0268945 A1 (MICROSOFT CORPORATION) 29.10.2009	1-48
A	US 2011/0154266 A1 (MICROSOFT CORPORATION) 23.06.2011	1-48
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents:	“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	
“A” document defining the general state of the art which is not considered to be of particular relevance	“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	
“E” earlier document but published on or after the international filing date	“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	
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family	
“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		
Date of the actual completion of the international search	Date of mailing of the international search report	
12 March 2014 (12.03.2014)	20 March 2014 (20.03.2014)	
Name and mailing address of the ISA/ FIPS Russia, 123995, Moscow, G-59, GSP-5, Berezhkovskaya nab., 30-1	Authorized officer	S. Finoshin
Facsimile No. +7 (499) 243-33-37	Telephone No. (499) 240-25-91	