A 3D Gesture Recognition Interface for Energy Monitoring and Control Applications

LUÍS SOUSA
ISE, University of the Algarve, Faro
PORTUGAL
lcsousa@ualg.pt

J.M.F. RODRIGUES
ISE, University of the Algarve, Faro
LARSyS, Lisbon
PORTUGAL
jrodrig@ualg.pt

JÂNIO MONTEIRO
ISE, University of the Algarve, Faro
INESC-ID, Lisbon
PORTUGAL
jmmontei@ualg.pt

PEDRO J. S. CARDOSO
ISE, University of the Algarve, Faro
LARSyS, Lisbon
PORTUGAL
pcardoso@ualg.pt

JORGE SEMIÃO
ISE, University of the Algarve, Faro
INESC-ID, Lisbon
PORTUGAL
jsemiao@ualg.pt

RICARDO ALVES
ISE, University of the Algarve, Faro
PORTUGAL
rmalves@ualg.pt

Abstract: Recently there has been an increase of interest in implementing a new set of home energy management devices that integrate Information Technologies, the Internet of Things and the ability to communicate between several appliances. Such energy management devices constitute an important milestone on the path to the Smart Grid, by being able to perform load-scheduling, taking into consideration different variables, such as the forecasted power generation from renewable sources, different tariffs’ rates, electric circuit constraints, user restrictions and correspondent comfort levels. To achieve it, they are required to maintain machine-to-machine communications with home appliances, while performing human-computer interaction through simple and intuitive solutions. In order to improve user experience when interacting with such devices, several depth and movement sensors are currently available that enable humans to interact with machines in innovative ways. In this paper, a 3D gesture recognition solution is explored, showing that when combined with 2D and 3D representations of building, objects and menus, allow humans to interact intuitively with energy management devices, in a way that cannot be achieved with other interface paradigms.

Key-Words: Human-Machine Interface, HCI, Leap Motion, Smart Grids
1 Introduction
Currently we are witnessing an increase in the energy produced by renewable sources, either motivated by the increase in the cost of oil exploitation or by an increment in environmental concerns, with its higher expression in the energy obtained from wind and photovoltaic sources. In this context, the traditional view of a distribution grid that uses centralized generators to provide power to consumers is being replaced by a “smart grid” solution where energy is produced based on a Distributed Generation (DG) schema [1]. In such solutions, the power generated by these energy sources varies according to environmental conditions, which are not controllable.

In this scenario, the traditional role of consumers is being replaced by a more proactive one. Not only in the sense that they should be able to produce energy locally, for self-consumption and/or feed it into the electrical grid, but also in the sense that they are expected to adjust their demand according to the power that is being produced.

By creating a new range of appliances that integrate Information Technologies (IT) and the Internet of Things (IoT) [2], users have now the possibility to get information, control or adjust automatically/semi-automatically or manually every equipment, machine or lamp in their building (e.g., house, hotel or factory). This allows an optimization of the consumption costs, based on the energy tariffs’ rates, the electric circuit constraints, the user restrictions and the correspondent comfort levels. One example of such a distributed load scheduling mechanism for micro grids can be found in [3].

In this context, in many buildings (home residences, hotels, factories, etc.), we are moving from dozens to hundreds, or even thousands of electric devices. The usability and ergonomics of the traditional selection menus and sub-menus using a keyboard, or a mouse (or even a touch screen) becomes unaffordable, for instance to give a command, or to query any information from device #1000. Many examples of possible different interactions can be mentioned, such as the request about the consumption of a lamp, or a group of lamps represented in a daily graph, or the changing of the washing cycle of a washing machine, or the selection of the home scenario that requires less consumption during a week day, or even to check all the consumption statistics of an air-conditioning machine in a specific location on a hotel.

Nowadays several types of three-dimensional (3D) sensors (see e.g. [4]) can be used to interpret specific human gestures, enabling a completely hands-free control of electronic devices, the manipulation of objects in a virtual world or the interaction with augmented reality applications.

These sensors, due to their reduced size and price, can be integrated in several different locations in the buildings, allowing the replacement of the traditional menu interface and peripherals devices. Furthermore, the new hands-free interfaces can provide easy navigation on the house/hotel/factory plant, with the real localization of the electric devices showed in any television or screen located on the building rooms. Within those plants, the user could navigate from device to device, individually selecting, turning them on/off or requesting information about them.

The main contribution of this paper is the presentation of this new human computer interaction (HCI) approach, to do energy monitoring and control applications, where the 3D hands-free interface combine 2D and 3D representations of objects, buildings and configuration options, allowing humans to interact intuitively with electric devices, in a way that cannot be achieved with other interface paradigms.

The remainder of the paper has the following structure. Section 2 presents the state of the art and a deeper contextualization of the problem in hand. Section 3 addresses the development of the interaction, and in Section 4 the resulting interfaces are presented, as well as the prove of concept. Finally, conclusions and future work are presented in Section 5.

2 Contextualization and State of the Art
In the context of the Energy Management in the Smart Grid, several scientific works [5][6][7] have analyzed the importance of an efficient management of home grids by consumers, or by an Energy Consumption Scheduling device (ECS). The role of an ECS is to optimally schedule loads in order to better harness the energy produced locally or shift them to work at periods of time when the tariff rates are lower, while reflecting user preferences. This also requires that loads, like HVAC (heating, ventilation and air conditioning)
and Home Appliances, should be able to communicate with the ECS device and shift their working periods, or adjust the power they consume, according to the power generated locally from renewable sources, or according to a supplier’s tariff, which in turn may change dynamically.

As mention in the Introduction (Section I), by creating a new range of appliances that integrate IT and IoT [2], with the ability of communicate between devices, we can respond dynamically to the varying tariffs, and a reduction on CO₂ emissions to the atmosphere is possible [8], while ensuring at the same time higher returns on investments made in renewable energy sources. Due to this, as described in [9], Smart Appliances are characterized as an important milestone on the path to the Smart Grid.

In this case, the ECS should decide which appliance is expected to work first and which ones should be scheduled to work later, reflecting a level of priority that should be commensurate with the user’s preferences.

In the search for a lower cost or tariff, one important feature associated with the ECS is to prevent electrical overloads that may happen when several appliances are scheduled to work at the same time period. In fact, while overloads may prevent the user from using the ECS, the quality of the scheduling algorithm from a user perspective will also determine the degree of freedom given by the user to it.

The user’s ability to manage their energy consumption according to the production is a critical feature of Smart Grids, and a base for innovation, new products and services. In order to support this capability, the communication between different devices such as meters, appliances, electric vehicles, energy management systems and distributed energy resources (including renewable energy and storage) must occur using secure, standard and open procedures. In this context, several protocols and architectures have been recently defined, like the Smart Energy Profile - Version 2 (SEP 2.0) [10], IEEE 1888 [11], and the OpenADR 2.0 [12], and can already be used by the ECS to communicate and control appliances.

One of these protocols, the Smart Energy Profile [6], SEP, results from the collaboration between the low-power ZigBee, the Wi-Fi and the HomePlug power-line technologies, building a power management architecture for Micro Grids, supported on IP networks.

In March 2011, the Institute of Electrical and Electronics Engineers (IEEE) announced the approval and publication of the Standard for Ubiquitous Green Community Control Network Protocol (IEEE 1888 TM) [11] within the Ubiquitous Green Community Control Network Protocol (UGCCNet). Originating in China, the IEEE 1888 standard defines itself as a global standard within the IoT, which aims at energy efficiency through the management of renewable energy, based on communication using Internet protocols and Information and Communication Technologies. Another communication protocol for Smart Grids is the Open Automated Demand Response (OpenADR) version 2.0 [12]. The OpenADR is an evolution and extension of the first version, developed by the Demand Response Research Center at Lawrence Berkeley National Laboratory, and is supported by the OpenADR industrial alliance, having been developed as part of the standard OASIS Energy Interoperation 1.0, published in February 2012.

If on one hand protocols that allow communication between different devices of a Micro Grid are being developed, a proper human-machine interface is needed to support the management of all these distributed resources. This is the purpose of the human machine interface (HMI) presented in this paper.

In terms of HMI, several solutions can nowadays be used by ECS devices to obtain and set user preferences. Traditionally, these interfaces have been based in touchscreens, or communicate with users using web pages that can be accessed from personal terminals like PCs, tablets or mobile phones. Nevertheless, the interaction between humans and machines can be made in many different ways, being crucial to obtain the expected level of control.

Desktop or mobile applications, including Internet browsing, make use of Graphical User Interfaces (GUI), being currently prevalent among the available solutions. Other solutions like Voice User Interfaces (VUI) use speech recognition to control devices. More recently, multimodal interfaces [13] allow humans to interact with machines in a way that cannot be achieved with other interface paradigms.

One of the new paradigms for human computer interaction are the three-dimensional (3D) sensors, such as Kinect [4], Leap Motion [14], Structure Sensor [15], or Asus Xtion [16]. Those sensors can be used to interpret specific human gestures, enabling a completely hands-free control of electronic devices, the manipulation of objects in a virtual world or the interaction with augmented reality applications. Many of these tracking and gesture recognition sensors have a huge importance.
in the videogames industries, once with the appropriate software, they have also the capability to detect the user skeleton and/or tracking a single or several users, while replicating with accuracy the user movements in a 3D mesh.

There is in the literature an enormous amount of applications, where gesture recognition and tracking is referenced, e.g., interactive art installations [17], applications to help disable or old people [18], air painting application [19], robotic arm manipulation [20] or applications in sign language [21]. One of the sensors mentioned in the above publications, that, due to its size, price and specific range of applications, has many potentials of use, is the Leap Motion sensor [14][19][20][21]. It is a recent, but widely known sensor for hand, fingers and gesture recognition, with a very high accuracy and speed.

The Leap Motion uses two monochromatic infrared (IR) cameras and three infrared LEDs (for more details see [22]). A smaller observation area and a higher resolution differentiate it from the Kinect and Asus Xtion sensors, which are more suitable for body tracking. To use it, users place their hands in front of the device, not too close, nor too far (in height and width) and execute predefined movements.

In the context of this work, in other to enhance the user’s experience, the sensors are combined with a 3D representation of options concerning energy management of home appliances.

Currently several solutions can be used to make such 3D representation. One of these solutions is the Unity cross-platform game engine [23]. Unity is a game creation system developed by Unity Technologies that includes a game engine and integrated development environment (IDE). It is currently used to develop video games for web sites, desktop platforms, consoles, and mobile devices. It is now the default software development kit (SDK) for the Nintendo Wii and has been extended to target more than fifteen platforms [23]. This platform is also the ideal one to create the graphics interfaces for our application.

3 Gesture Recognition with Leap Motion

Leap Motion [14] has an Application Programming Interface (API) capable of detecting multiple hand gestures, such as straight line movement by hand with fingers extended, a circle movement by a finger, a forward tapping movement by a finger and a downward tapping movement by a finger. These gestures (swipe gesture, circle gesture, screen tap gesture and key tap gesture) can be seen in [14] or in [24].

For each of the gestures there are a few optional configurations properties, to improve gesture detection. For example in a circle gesture, where the user can do a circle with a finger, there are two selectable properties: minimum radius and minimum arc (by default, minimum radius was set to 5mm and minimum arc was set to 1.5π radians). For a swipe gesture, there are also two properties selectable, minimum length, set by default to 150mm, and minimum velocity, set to 1000mm/s. These are examples of the “out of the box” gestures, possible to retrieve with Leap Motion API.

It is also possible to join some of these gestures, like the positions and rotation of each finger and the use of both hands at the same time, to recognize other gestures, such as open or close hand (used for example for zoom in or out) which can be done with one or both hands.

For the development of the interface we tried to define a short number of movements, and at the same time select the most intuitive ones - the swipe, was considered as the best approach. Thus, swipe gesture was the gesture mainly used in the interface developed in this paper, e.g., for the interaction with the energy monitoring and control applications.

After regulating the minimum length swipe and velocity to 100mm and 400mm/s, respectively, it is possible to detect swipe gestures at any direction. Leap Motion API has a direction vector for the swipe gesture, i.e., a gesture completely recognized is associated with a 3D direction vector. This vector has values ranging from -1.0 to +1.0.

As shown on Fig. 1, the Leap Motion “sees” the 3D space as a standard Cartesian coordinate system, also known as right-handed orientation coordinate system. The origin of the coordinate system is centered at the top of the device, being the front of the device the side with the green light (see Fig. 1). The x-axis is placed horizontally along the device, with positive values increasing from left to right. The z-axis is placed also on the horizontal plane, perpendicular with x-axis and with values increasing towards the user (the front side of the device). The y-axis is placed in the vertical, with positive values increasing upwards (see Fig. 1).

As different types of swipe gesture exist we needed to detect and differentiate them.
Fig. 1 - Leap Motion coordinate system [24].

The interface was designed to react to six different independent types of swipe gestures as shown by Fig. 2, three of them, are the opposite of the other three:

(i) Select and deselect is realized by a top to bottom swipe or a bottom to top swipe respectively ($y$-axis).

(ii) Select different floors (see Fig. 4) of a house is realized by a front to back and back to front swipes, to select lower and upper floors respectively ($z$-axis).

(iii) Selecting the next and previous item/object is done by a left to right or right to left swipe respectively ($x$-axis).

In the first case (i), of a top to bottom and bottom to top swipe, the movement depends mainly on the $y$-axis. If the direction vector has an upward direction ($y \approx +1$) then a “deselect” action has occurred. Otherwise, if the vector has a downward direction ($y \approx -1$), then it is considered a “select” action. Since it is almost impossible to do a swipe gesture with a vector direction component of $x = 0 \land y = \pm 1 \land z = 0$, (1)

the algorithm should instead select a range of values to detect and differentiate between swipes types.

So, any swipe direction that agrees with the condition $y \leq -0.5 \land |x| \leq 0.5 \land |z| \leq 0.5$, (2)

is considered a downward swipe. Contrariwise, if a swipe direction agrees with the condition $y \geq 0.5 \land |x| \leq 0.5 \land |z| \leq 0.5$, (3)

then it is considered an upward swipe.

In the second case, (ii), we needed to analyze mainly the $z$-axis. As shown on Fig. 1, a vector with direction in $z$, who has a value approximately 1, is considered to be a front to front swipe. Otherwise, it is considered as a front to back swipe. Similar to (i), any swipe direction that agrees with the condition $z \leq -0.5 \land |x| \leq 0.5 \land |y| \leq 0.5$, (4)

is considered as a back to front swipe. Contrariwise, if a swipe direction agrees with the condition $z \geq 0.5 \land |x| \leq 0.5 \land |y| \leq 0.5$, (5)

then it is considered a back to front swipe.

The last case, (iii), where $x$-axis is the main axis, is again similar to (i) and (ii). Any swipe direction that agrees with the condition $x \leq -0.5 \land |z| \leq 0.5 \land |y| \leq 0.5$, (6)

is considered as a right to left swipe. Contrariwise, if a swipe direction agrees with the condition $x \geq 0.5 \land |z| \leq 0.5 \land |y| \leq 0.5$, (7)

then it is considered as a right to left swipe.

These swipes are mutually independent, i.e., for every type of swipes there is only one possible choice, see Fig. 2.

![Fig. 2 - Six types of swipes for the application.](image)

### 4 Interface and Interaction

In order to build a prove of concept, a house similar with a real one, was created using Sweet Home 3D [25], using data from different devices. This home includes three floors and a garden with an outside pool, creating 4 different areas, as represented in Fig. 3. The control interface was built using the same number of floors and rooms.

Each of those areas has a set of different electrical devices (appliances/equipment). For example in the ground floor (second from the bottom), there are various types of equipment such as a fridge, a stove, a washing machine, a drying machine, an air conditioning, and various types of lightings.

The interface was made in Unity 3D [23]. It was divided in 3 different levels. In the first level, (i) the user can select between different floors/zones. In the second level, (ii) the user can select the different machines in the selected floor. In the final level and after selecting a machine, (iii) a picture of the selected machine is shown together with some...
information regarding statistics, scheduled programs and the option to add new programs.

In the first level of the application, the four different zones were positioned one above the other, as shown in Fig. 3. On the second level, each electric device was considered an individual object. The selection of that appliance could be represented through an image scaling or by changing its color (see Fig. 7 top row; the device is marked with the red color). After performing some tests with users, it was decided to mark the selected object with an arrow (see Fig. 4).

Figure 4 shows the interface being operated. At the top the ground floor is shown, which is the default zone/floor selected. Right above and below are the first floor and basement, respectively. As explained previously (see section III), there are six different swipes that the user can do to navigate in the interface with the Leap Motion. To move from different floors/zones, the user has to do swipe movements along the z-axis of the Leap Motion. A back to front swipe moves the interface to the basement and a front to back swipe moves the interface to the first floor and, if repeated, to the outside zone.
Fig. 4 – The system in operation. In the top, the ground floor is presented and in the bottom the ground floor is selected.

A downward swipe selects the zone in the middle of the screen and the second application level appears. Simultaneously, the zone plan is zoomed in, as Fig. 4 bottom shows. When doing the opposite swipe (i.e., upwards swipe) the zone is deselected and changed to the first level of the application.

A default appliance is selected when one zone/floor is selected. For example, the default machine in the ground floor is the fridge in the kitchen, and is marked with a green arrow (see Fig. 4 bottom image). The swipes along the x-axis are used to select between different appliances (e.g., machines, lights, oven, and air conditioner). Doing a right to left swipe changes the selection for the next machine, while a left to right swipe changes to the previous selection.

The green arrow only “pre-selects” the machine. Doing a downward swipe selects the machine pointed by the green arrow, and changes to the third and last application level. In this last level, the 3D model of the selected machine (appliance/device) is shown together with a new menu showing information about the associated machine. As shown in Fig. 5, after selecting, e.g., the washing machine (from the ground floor), all the available information about the respective device is shown in a menu.

The last menu is composed by four different selectable options: New program (Fig. 5, top row), Scheduled Programs (Fig. 5, 2nd row), Statistics (Fig. 5, 3rd row) and Information (Fig. 5, last row).

To change between those options the user can do a swipe around the x-axis. If a left to right swipe is done the option from the right menu is selected, otherwise, in case a right to left swipe is done, the option from the left menu is selected (see Fig. 5).

Fig. 6 represents an example of the scheduling of several loads, after running the optimization algorithm and taking into consideration the user preferences. This graphic is one of the many options available in the General Configuration/Information menu.

Another option is the view representation of the plan, that can be “top view”, as in Fig. 3 or Fig. 4, or can be seen as “3D view” as can be seen in Fig. 7, bottom two rows. This menu is available in every plan (floor/house/hotel/factory/etc.), and is selected as any normal device. The enumeration of all the options available is out of the focus of the paper, nevertheless it is important to mention that the navigation on those options are done only with the 6 swipes explained in Section 3.

4 Discussion

In this paper a combined 3D gesture recognition solution and 2D/3D representation of buildings, objects and configuration options is explored, that allows humans to interact intuitively with electronic devices, in a way that cannot be achieved with other interface paradigms.

While the proposed system is still under development, requiring more tests to prove the final concept, the initial tests with a small group of users have shown very promising results, as all the opinions were favorable to the interface.
Nevertheless, some usability problems still exist, as well as design flaws. Some small problems were also noticed with the Leap Motion sensor, since it doesn’t work properly in some lighting conditions. Sometimes, gestures can be trigger/detected by moving the hand above the device, when the user don’t want to do a swipe gesture (to minimize this, tests with several people were done, and the configurations properties were changed accordingly; see Section 3).

It was also noticed that some users do the swipe movement in a diagonal (not in the direction that agrees totally with the conditions predefined in Section 3), in this case, as expected the system doesn’t respond, nevertheless, this aspect as to be resolved in the future, once it was notice that this is a natural movement that users do.

In terms of future work, a bigger case study is being prepared with more electric devices and rooms. Other 3D sensors, such as the Structure Sensor are also going to be explored, in order to create even a more intuitive interface, as well as minimize some false positives and negatives occurred in the swipe movement.

Fig. 5 – Example of a complementary information shown by the interface about each individual device.

Fig. 6 – Power consumption/production (in kW), during a 24 hours period.
Acknowledgments: This work was supported by FCT project PEst-OE/EEI/LA0009/2013 and projects QREN I&DT n." 33845, “PRHOLO: The realistic holographic public relations” (project leader SPIC – Creative Solutions [http://www.spic.pt/]) and QREN I&DT n." 30260 “MTI: Managing the intelligence” (project leader Certigarve [http://www.certigarve.pt/]).

References:


