

# A Kinect-based 3D hand-gesture interface for 3D databases

**Abstract.** The use of natural interfaces improves significantly aspects related to human-computer interaction and consequently the productivity and the overall performance. In this paper we present a novel framework to interact with data elements presented in a 3D space. The system provides two mechanisms to interact using 2D and 3D gestures based on data provided by Kinect and on hand detection and gesture interpretation algorithms. The proposed architecture is analysed indicating that 3D interaction with information is possible, and provides advantages over a 2D interaction over the same problem. Finally, two sets of experiments were performed to evaluate 2D and 3D interaction styles based on natural interfaces focusing on traditional interaction with 3D databases.

*Keywords:* Human-computer interaction, hand gesture, programming and developing interfaces, 3D data representation.

## 1. Introduction

Research in natural interfaces has increased significantly in the last years. For the most part, this is due to the emergence of acquisition devices, which can easily be built at low cost. Consequently, the amount of interaction mechanisms and the related interfaces is growing fast, including software and libraries written to support these systems. Nowadays, the advances in graphical interfaces have reached a breakpoint where simple interaction devices are not enough to provide adequate manipulation of 3D elements on a display. Also, 2D representations of the information are unable to provide an appropriate interaction experience [1]. Even if contemporary graphical interfaces have evolved from the typical writing code to visual programming environments [2], there are still some non-graphical components, which reduce the user's understanding and productivity [3]. Therefore, a better understanding of the working environment and the required tools could be a solution to these problems [4].

Recently, significant research was conducted to overcome the issues between real environments and computer interfaces, focusing on human computer interaction. For example, in [5] researchers addressed the importance of

introducing new communication between humans and computers, replacing the traditional methods and devices. This new communication has to be based on systems capable to use not only one type of interaction mechanism, but to integrate more than one natural interface, such as verbal or gesture communication. As a result, the human would be able to overcome all the inefficiencies and limitations related to element manipulation and environment understanding during the interaction with a software system. Additionally, better hardware integration with the working environment can improve the user's experience (e.g. interactive data processing).

New ways to interact with machines have been created based on computer vision and image understanding [6]. These advances aim to improve human-computer interfaces and their main objective is to provide natural interaction mechanisms based on body motion understanding, gestures analysis and sensor integration [7]. This effort is leading to the design and development of hardware interfaces appropriate to operate in specific applications like geographic data management, education, industrial design, architecture and web interfaces [8]. Additionally, in the video game industry, the need to provide new levels of experiences and much higher interaction between the users and the systems results in significant novel contributions in the area of user interfaces, [9].

The level of detail and accuracy in an interaction environment are also crucial parameters. For example, the connection between the graphic metaphor and the data to be manipulated [10] could be problematic, due to the spatial representation of the related elements. Therefore, in the case of a 3D framework, it usually provides tools to create software layers for specific components, such as graphic elements and the interaction language [11]. Also, since the developers have to create the objects that are going to be used as metaphors, the required level of detail is significant high [12]. Another issue is related with the flexibility of these metaphors since they should be flexible enough to be used in any framework including 3D environments for any kind of interface.

Gesture based systems that are focused on hand gesture controls, have become popular, especially in hand-held portable systems such as laptops, mobile phones and gaming devices [13, 14]. Since they support only two dimensional interactions, they are not able to perform naturally tasks and activities that are performed in three dimensions in real environments [2, 15]. Depth capturing

devices have provided new mechanisms for interface systems, as it was shown by Microsoft Kinect [16, 17]. An issue related with these interaction mechanisms is the accuracy in performance of gesture based systems and how this may affect the overall interaction [18, 19].

All the above interaction mechanisms and mainly the gesture based approaches can be utilized also in databases. Particularly, data selection and manipulation with hand-gestures can be applied on 3D datasets using 3D graphical representations. In that way, more natural and intuitive mechanisms are provided to interact with information and especially data elements in 3D datasets.

This paper presents the results of our studies in the field of 3D hand gesture interaction focusing on 3D datasets providing novel ways to improve the user experience. The advantages of using a natural human computer interface without external devices, such as data gloves, solves some of the most problematic issues in this area such as finger tracking, gesture identification and recognition in real time. The proposed novel framework provides mechanisms to interact with 3D databases based on 2D and 3D hand gestures. These interaction mechanisms are more intuitive than the classical two dimensional environments allowing flexibility and increased productivity.

This paper is organized as follows: the following section describes previous works in this area and the advances in human-computer interaction mechanisms. In section 3, the concept of 3D data representation is analyzed and two different 3D interaction approaches are proposed, based on modern acquisition devices. In section 4 and 5, experimental results are shown providing a comparative study with the traditional 2D data representation systems. Finally, the results are discussed and the conclusions are addressed.

## **2. Previous Work**

Structures to represent information in an intuitive way to understand and explore datasets have become a challenge for researchers. There is a fundamental need to create interactive tools that provide access to large amount of data, however, in order to handle modern database systems, advanced knowledge and training is essential. A new generation of databases aim to change their typical text-based representation to visual formats, where interaction can be achieved by using

natural interfaces, such as gesture based commands or multi-touch interactions, instead of complex sequences of commands [20].

Technologies aimed at improving interaction with databases led to the creation of new paradigms to visualise information. The graph databases provide mechanisms that improve the classical relational model. Graph databases represent information in the shape of graphs where each node corresponds to a specific data type with specific attributes (i.e. address, date, user id, etc.), while the nodes represent connections between the data elements (i.e. source, sink). The graph databases present advantages in the retrieval of information compared to the conventional relational models, offering a new way to store and retrieve data [21]. This type of representation also provides another advantage over relational methods. The internal representation of information is based on a graph model; the creation of interfaces capable of dealing with information under a graphic interface is intuitive and allows the use of both traditional and modern input devices, such as multi-touch and gesture based systems. The representation of information using this graph model has the problem of supporting only 2D interfaces to interact with information.

Other models of databases providing data modelling in multiple visual dimensions have been introduced. These models rely on a multi-dimensional representation of information, where the data can be perceived as a cube, where each “cell” of information contains a set of measures of interest, related with three information sources. This model is the one used by the paradigm known as On-Line Analytical Processing (OLAP), and graphical 3D interfaces, based on this method, can be created, focused on geographic and spatiotemporal data management systems [22]. Also this type of data modelling (OLAP data cube) has been successfully used to store and query real event data from sensors in smart buildings [23], where parameters such as temperature, humidity, luminescence and related events can be stored in a cubic cell that registers date, device and value related with the considered parameter. Even if OLAP data cubes provide a powerful tool to interact with information, their interfaces rely on 2D representation and traditional input devices, which reduce the level of effective interaction. The use of natural interface paradigms to interact with data presents an interesting alternative over traditional methods. User interfaces that support the use of gesture or touchless 3D interaction allow the better understanding and

manipulation of 3D graphic contents and are applicable over different interaction scenarios, where direct touch is not possible. These frameworks are able to utilise Kinect device, providing an alternative to traditional methods. Also they allow the design and development of a gesture vocabulary that can be used in “traditional” software interface, but with a separate module capable of connecting 3D gestures to complex browsing actions [24]. Even when these frameworks provide an interesting approach, the connection between gesture based interaction interfaces and databases is still required.

Approaches that combine the previous methods (touchless interaction and multi-dimensional databases) appear to be the next step in data interaction development. Natural user interfaces for OLAP cube based systems are possible to be implemented. A clear example is Data3 [25], that introduces a new approach to interact with multidimensional databases, where the dataset itself is modelled as a 3D cube interface (following the logical data representation of OLAP structure). In this interface, the interaction with the data cube is done by using gesture detection based on body motion capturing, provided by skeleton tracking and the OpenNI framework, without using direct hand and finger based interaction. The supported gestures are basically swipes (for rotation), pushes (for selection), and combinations of them using both hands. The initial definition of gestures is done under a declarative environment (text-based programming) using the AnduIN data stream engine to process the events coming from Kinect translating them in command gestures. This approach makes use of a 3D interaction and 3D data representation, which allows better understanding and faster user task performance. The main issue with this approach lies in the use of full body motion to generate the gestures that actually can be a problem in desk-based applications, where a direct hand gesture based interface would be more appropriate, allowing the users to perform the same tasks in 3D in a more efficient way and with less effort.

In the following section, a finger based two handed gesture interaction method is presented, in order to overcome the issues of the methods previously presented.

### 3. Proposed Methodology

The proposed methodology defines a common development framework for two hands-gesture interaction in 3D environments. The application is divided into layers, each of them containing specific tasks. The layer architecture and the connection between their components can be seen in Figure 1.

The architecture presented has similarities with a multi-touch architecture, such as the one presented by Echter and Klinker [26] but with several important improvements and modifications. The Hardware Data Acquisition layer takes the information directly from the device previously identified by the associated API, which in our case is Microsoft Kinect [27]. In multi-touch architectures, this task is performed by two different layers, following the touch user input/output (TUIO) protocol [28]. Also, depth detection is crucial during the performance of 3D activities, necessary for natural 3D interactions in indoor environments [29]. The Hand Gesture Acquisition layer is responsible for defining and identifying hand's shape, fingers and interpretations, which are related physically (fingers and hand relative relation). This process is not performed in multi-touch architectures, because all the interactions over the surface correspond to fingers and the correlation with the hand is not necessary. Furthermore, in this layer, the system selects the features that define the palm and then the ones that represent the fingers. The Gesture Interpretation layer works in a similar way to a multi-touch's interpretation layer, but in our approach, the hand position is also used to define the gestures, "translating" a gesture to a specific command, according to the interacting environment and the fingers' identification in the previous layer. The Command Graphic Association Layer makes visible the action, the association between the "logic" object and their graphic representation has to be performed in this layer. The information is passed immediately to the Graphic Interface layer, which takes control of the actions and changes in the environment after the performance of a predefined gesture, triggering a subsequent action. Finally, the graphic interface displays the outcome of the interaction. This layered architecture provides flexibility to define several combinations for 3D interactions for different applications. Also, it can provide a high degree of hardware independence, since the modular definition of the architecture allows replacing components in any layer.

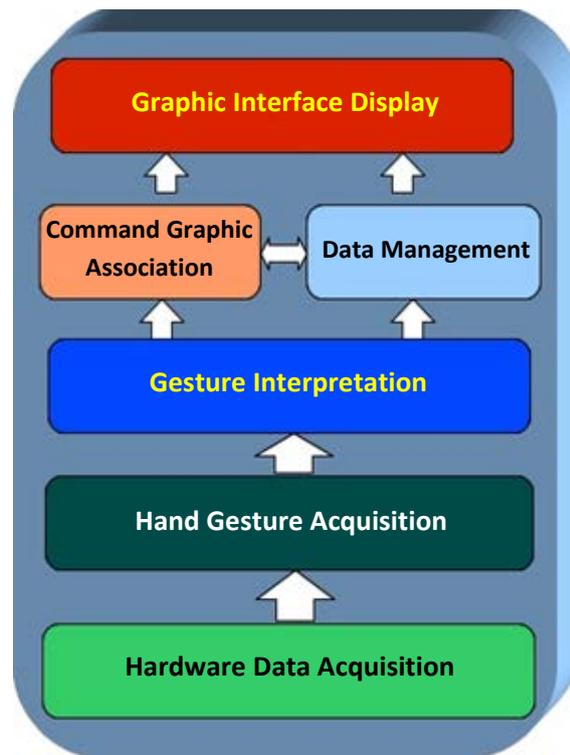


Figure 1: Layer architecture for two-handed gesture based systems.

### 3.1. Finger tracking

Finger tracking is related to a hand-gesture based interaction, since the proposed gesture is based on hand and finger correlation; and hands relative position. The detection of the hand is based on the depth map provided by Kinect, using an approach similar to the one presented by Xia [30], focusing on the extraction of the palm of each hand and the detection of the hands' contour using the segmented depth map. In order to identify and label each hand, the relative position in the detection space is considered. The hand in the right portion of the detection space will be the right and equivalently for the left hand. The positions are defined this way to provide the user direct feedback over the actions and the areas for each hand are defined from the centre of the detection space. If the hands cross over, since the hand detection is not connected with a skeleton tracking, the hand identification is switched, making it wrong. The system uses an approach similar to the one presented by Frati [31] to perform the detection of the fingers, where the hand needs to be held facing the device. The detected features can be seen in Figure 2 and the algorithm provides the fingertips of the index and the thumb,

which correspond to the start and end points of the convexity (that corresponds to the area where the shape of the hand has gaps, allowing the separation between fingers, identifying them individually).

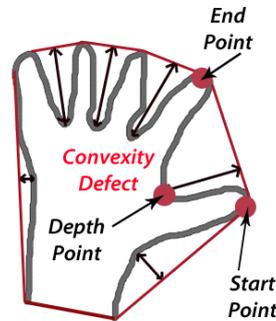


Figure 2: Fingers' detection algorithm using convexity, used by Fratti [31].

Each finger is recognised by the detection of end points and overall convexity. Each end point represents a finger, and the palm point will provide information about the relative 3D position of each finger on the hand, allowing the performance of 3D gestures improving the interaction. However, since the detection of the hand is provided by the segmented depth map, the problem of hand gesture detection is highly dependent on the distance from the sensor, as was also discussed by Tang in [32], mainly due to the occlusions and possible reflection problems. The interaction space for our approach is between 0.6m and 0.9m from the camera and the width of the interaction space is about 0.6m (the hands should face towards the camera in that space).

It's necessary to mention that hands are detected as independent elements, not connected to the whole body, which further limits the detection range, but provides more degrees of freedom and improved gesture recognition speed, since there is no need to calculate the rest of the articulations of the body. As a result, the detection of the hands and the fingers associated to each of them provides enough functionality for 3D hand-gesture interaction in real time [33]. The main advantages of this approach are related with the absence of a training period of the system. Also, the performance of the system is not affected by hand size changes, given the model to detect the fingers.

Regarding the performance, compared to the entire human body, the hand is a smaller object with more complex articulations and more easily affected by segmentation errors. Nevertheless novel algorithms [34] match the finger parts

and not the whole hand, distinguishing better, hand gestures. The accuracy of state of the art methods [35, 36] for hand and finger gesture recognition is more than 93% with an average latency of 0.075 seconds per frame, which makes this technology suitable enough for real-life HCI applications.

### 3.2. The interaction

The interaction in our approach, during the experiments is based on hand-gestures, and as a result the interface operates using only the hands, using a depth and video capturing device. During the gestures' definition, two sets were developed: a set of gestures that combine both 2D and 3D movements and a second that utilises only 2D gestures, which were performed during the experiments. The choice of the two sets of gestures was set to test whether the 3D gestures can improve user interaction times and performance compared with 2D gestures (which are commonly seen on traditional touch surfaces). The proposed hand-gesture interactions are based on the number and the position of the fingers. Changes in their position trigger different actions and responses at the system. For both sets of gestures, one of the hands indicates the function (or mode) and the other performs the action, which allowed avoiding gestures' mistakes according to previous experiments performed. Considering that, the interactions are divided into three types:

- **Movements:** These actions correspond to changes in the position and/or the orientation of 3D graphic elements in 2D or 3D space. The action is performed only if an object is already selected and is related to the actual location of the hand in the 3D space.
- **Selections:** The selections are applicable only to specific 3D elements in the environment and when successfully performed, some components or parts of them are highlighted. The selection process is based on two actions: locate the element that will be selected and the selection process itself.
- **Executions:** Interaction related to performing a particular action not defined as a previous one. These actions could be the result of a combination of the previous ones or just a single hand gesture.

For our experiments, these interactions will be further analysed regarding their implementation in the results section. Also, it should be mentioned that these

interactions are enough to perform the required tasks in a database system, since the actions that a standard 2D mouse can perform are a subset of them. Consequently, in the case of 3D databases the proposed interaction mechanisms are enough, supporting similar actions with a standard mouse but in a three dimensional space.

### **3.3. Three Dimensional Databases**

3D databases are a derivation of multidimensional databases and in our case, the type of a cubic database is considered. These kind of databases provides an interesting field of development and research due to the data mining features offered by this model (i.e. find correlations between data elements invisible in a 2D relational model, such as the relation between items sold, stores and dates over a multi-store company database) [37]. Cubic databases are useful in cases where relationships between different pieces of data are not totally clear and the connection of several information sources is required, which cannot be performed easily by the traditional 2D databases. An example is related to medical information, and the need to find associations between not obviously related features improving the diagnosis process and the patient's healthcare. As a result, important parameters for the diagnosis of diseases can be estimated, allowing a more efficient control of the demanding health services, especially for primary care [38]. For example a 3D dataset could store personal information of patients in the first table/dimension. A set of measurements with related information for each one could be available in the second table/dimension, and finally the actual measurements over a certain period of time could be part of the third table/dimension. However, due to the complexity of the traditional interaction models, we defined a novel approach that resembles the functionality, where the cube is formed by multiple tables linked together.

### **3.4. Suggested interface model**

The main interface used to analyse our proposed methodology was a 3D data interaction model. In more detail, a simplified model of a cube database with multiple faces was introduced, representing information about a group of patients. The interface can also manipulate larger and more complex databases by adding sliders that could be manipulated by a combination of a selection and moving gestures.

In order to provide a more intuitive interface, only two successive sides of the cube are displayed at any time instance (the front face with the patients' details and the right side with their measurements e.g. weight per month).

In more detail, one face of the cube contains basic personal information of the patients. The other faces have information about the weight or other measurements of the patients, for a period of several months, (e.g. July, August and September). The top face of the cube provides the option to terminate the application by performing the related gesture over the 'close' button.

The user is able to interact with this cube using both hands. The left hand is the function 'indicator', while the right hand actually performs the action on the screen. This configuration was selected in order to limit possible confusions between functionalities and also it can be reversed to facilitate both left and right hand users. The hand of the user and particularly the index finger is followed by a screen indicator to allow the users to have a visual representation of their exact position on the screen and on the cube.

To improve the feedback to the user, a visual text chart has been added to indicate the current function performed which changes according to the detection of the hand-finger gestures. This text indicator will show the function mode (e.g. movement or selection according the indicator hand), if the action is being performed (selecting or moving) and finally, in which column the action is being performed.

In the following section the experiments related to the cube database model are presented, the set of gestures and the experimental procedure are described focusing on the users performing basic information tasks with their hands.

## **4. Experiments**

In our experiments, two set of gestures were designed and developed mainly for interaction with 3D databases. The suggested sets of gestures were tested in order to demonstrate and evaluate the users experience and indicate the need for 3D interaction in these applications. The evaluation process was focused on an example of a simplified version of a 3D cube database containing information about patients and their measurements over a period of time. This simplified

model considers several patients providing personal details (ID Code, Name and Gender); a list of measurement features (e.g. weight, height, heart rate, etc.) including their importance (e.g. weight) and finally, the actual values for each feature of each patient over a period of few months (e.g. July, August and September).

With these experiments we wanted to investigate whether the 3D hand gesture interaction provided a better experience in comparison to 2D gesture based interfaces and also traditional mechanisms using the keyboard and the mouse based on Structured Query Language (SQL), where the main difference between 3D and 2D gestures lies in the use of depth of the fingers to perform gestures. Therefore, during our evaluation hand-gesture interactions were used to interact with the database and these gesture-based interfaces were compared with traditional SQL queries. Also, since we are focusing on a comparison of 2D and 3D hand gesture interaction mechanisms, an analysis of 2D mouse pointing interfaces with hand movement systems is not part of this work and an extended analysis is available in [19].

The experiments are divided into three stages, which are as follows:

- the *presentation stage* (where the interface is presented and explained)
- the *practice stage* (where the users can interact with the interface and use the available features), and
- the *task execution stage* (where the users perform the task and quantitative factors are recorded, such as task execution time, the number of users and system errors). The defined task for the evaluation was the columns' selection from the cube database simulating a querying procedure.

In order to provide better feedback to the users, another graphic element was added to the interface: a dialog box that indicates the columns correctly selected (according to the given task) by the users (on the bottom of the screen, as it can be seen in Figure 4). The main elements (according to the stages' definition) of our experiments are analysed in the following section.

#### **4.1. Set of experimental Gestures**

Two sets of gestures were used during the experiments, as mentioned in the previous section. The fingers combination was selected after several tests with the graphic interface.

The set of the selected gestures are enough to perform all the interactions analysed in section 3.2 and are further divided in 3D and 2D gestures, which correspond to the first and second interaction experiments, respectively. The amount of hand-finger combinations used in each gesture was determined experimentally to avoid the confusion between gestures, allowing the correct identification of the performed action, following further experimental results about finger detection reliability [39]. Also it was observed from the initial experiments that using numbers pointed by the fingers was more intuitive at this stage helping to memorize the available interaction mechanisms. For the 3D set of gestures, the related actions are defined below:

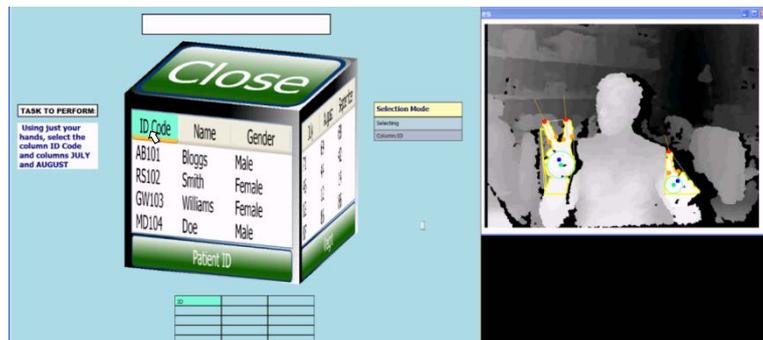
- **Rotation:** The cube can be rotated from left to right and vice versa around the vertical axis. The rotation action is performed by keeping the left hand totally open (all the five fingers, indicating the “Rotation Mode”) and simultaneously moving one finger (any, but the index finger is preferred) of the right hand from left to right or vice versa, depending on the face of the cube that the user wants to see. During this action the cube rotates smoothly from one side to the other according to the finger’s position. Other combinations of movements to provide rotation around other axis was considered, but to simplify the interaction process and to not confuse the user, the rotation feature was limited to rotations about the Y axis.
- **Selection:** The selection is considered more as a mode than an action allowing the identification of graphic elements to be selected. In order to enter in this mode, the users must show two fingers of their left hand and place the cursor over the selected element, using the indicator finger.
- **Clicking:** The clicking action works as the execution phase of the selection mode of an identified element, and because of that during the clicking process, the user must remain in selection mode (two fingers of the left hand have to be visible). The clicking is performed by placing the indicator finger over a selectable element and “pushing” (moving forward, towards the screen). The clickable elements on the cube are the “close” button, the column headers and the data rows. In order to choose a full column (to perform a specific task), it is just necessary to click on the column header.

The set of 2D gestures are defined below:

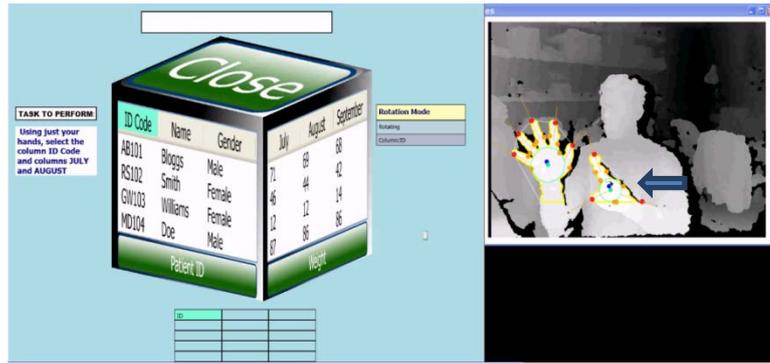
- **Rotation:** The rotation gestures are the same as the one for the 3D approach, because it does not have any 3D interaction itself. The combination of fingers and sequence of movements are the same for this experiment.
- **Swiping:** To perform a selection in the 2D cube interface it is necessary to swipe over the column or row. This process requires first the indicator to be positioned on the top or bottom of the column and then swipe along the column to be selected successfully. To avoid a wrong selections, the column is divided into several equivalent surface sections that must be swiped sequentially using a vertical (or horizontal) movement to perform a column (or row) selection that can be also performed from bottom to top or vice versa.

#### 4.2. Execution Task Description

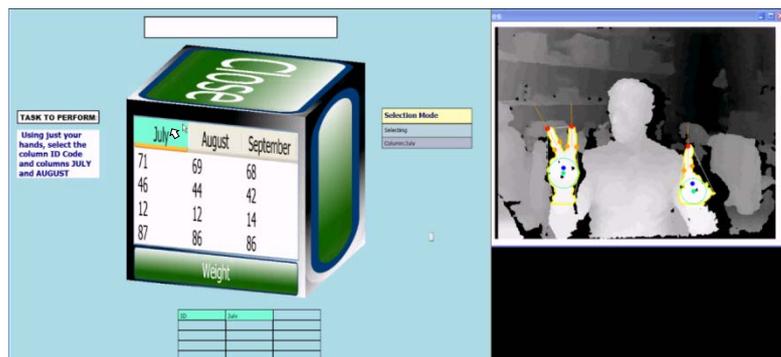
This task consists of a sequence of column selections, where a simple information selection query is performed on the 3D database tables containing patients' details. In this case, the user is asked to select the name and the weight information for the months July and August by selecting the corresponding columns. There is no specific order in the selection, but the combination of these data columns is needed for each face of the cube to complete the task. In the case of selecting the wrong data, the user is requested to repeat the task. This aspect reduces the possibility of a random selection allowing the user to focus on the required tasks. The ideal interaction sequence to perform the task for our 3D approach can be seen in Figure 3, to clarify how the interface works.



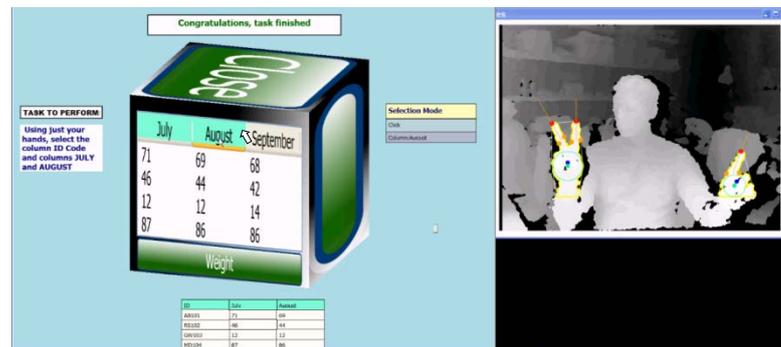
(a)



(b)



(c)



(d)

**Figure 3: Interaction sequence for the task in 3D gestures case (ideal scenario): a) First step: in selection mode, click on ID code b) Second step: in rotation mode, rotate the cube (moving the indicator finger from right to left, inverted in the captured image) to access the next table of the cube (months) c) Third step: in selection mode, click on July d) Fourth step: in selection mode, click on August and the task is completed.**

In more detail, the expected sequence of steps for this experiment is: first enter selection mode and perform the click over the ID Code column (the ID code is shown by default when the system starts), then it is necessary to rotate the cube to expose the face that contains the months' table. Once the rotation is completed, the user must select the months July and August. During the selection process, the user has constant feedback about the performed actions and when the columns selection is correctly completed; in the lower part of the screen a highlighted table

with the name of the selected column is shown. Also, the header of the selected column changes colour in the case of a correct clicking. Once the task is successfully completed, the selected data are displayed on the lower part of the interface (see Figure 4). However, the user can start rotating the cube, selecting the months and then rotating back and selecting ID Code or following any other sequence of actions to achieve the expected. If the user selects a wrong column during the process, the task must be restarted again.

The 2D interface works in a similar way but the main difference is that all the interaction is performed like on a touch device, which means there are no depth related movements. The selection instead is made by swiping over the data (e.g. columns) as was described previously without any 3D interactions. This swiping must start in a defined initial place of the data to be selected (in the case of a column, on selection mode, the finger must swipe over the top or bottom sections of the column and then move along the column and reach the opposite extreme). Also, to improve the feedback provided to the user, another feature was added: semi-transparent cells to indicate when the user passes over them were implemented changing colour (e.g. to red).

The general concept is to investigate the advantages of a 3D graphical based query using hand movements over an equivalent 2D interface and also the traditional SQL approaches using keyboard or mouse. In this analysis, both qualitative and quantitative information is obtained over the usability of the interfaces and the general user satisfaction. During this process the 2D and 3D interfaces are compared also with an SQL query using a cube dataset configuration based on the concept of information manipulation, access and retrieval (involving a process of multiple selections and join operations over different related tables). A cube interface configuration is suitable for a hand-gesture based interface, since interfaces of this type resemble aspects of real world interactions. The full interface for these experiments when the task is achieved can be seen in Figure 4.

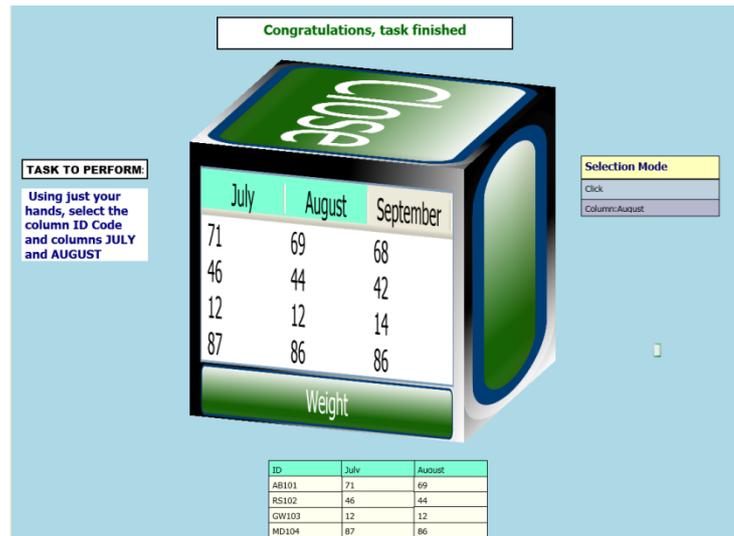


Figure 4: Displayed interface when the task is successfully completed.

### 4.3. Evaluation Procedure

During the evaluation process the experiments with the users were divided into of the following steps.

#### 4.3.1 Present and explain the experiment and its objectives

The objective of this section is to provide information about the possibility of using 3D hand gesture interfaces instead of the traditional 2D and SQL code based interaction to perform queries on a 3D database. As a result, the proposed experiments perform a comparative study evaluating the users' performance on the proposed prototype interface versus the traditional approaches.

#### 4.3.2 Demonstration

The aim of the demonstration section is to present to the users the interface and its elements, answering any related questions. Also, if the user is not familiar with SQL, the basics of the language are explained.

#### 4.3.3 Familiarise the subject with the interface

During the familiarisation with the interface stage the interaction mechanisms are presented to the users allowing them to practice with the basic movements and the on screen features.

#### 4.3.4 Subject performs the supported actions

The supported functions (e.g. rotate, click, etc.) are explained to the users and demonstrated in real time during this step. Furthermore, they are encouraged to

practice and perform these functions by themselves. The total time of training is less than ten minutes.

#### **4.3.5 Quantitative data collection**

Once the users are familiar with the environment and with the mechanisms to perform the available functions, the full task that was initially introduced is performed measuring the required time to successfully complete it, the number of errors during the users' interaction and also the errors due to the system inaccuracies. Therefore, the quantitative metrics used to evaluate the interfaces for each set of gestures were:

- *the required time to complete the task*
- *the number of errors caused by the users and*
- *the number of errors caused by the system.*

The error metrics (users and system errors) were needed to determinate the influence of them in the general performance; the user errors also provide information about the understanding of the gestures required to work with the interface. The system errors provide information about the correct identification of gestures and the possible improvements required. For each set of gestures, two types of errors were considered during the following actions: clicking in the case of the 3D approach and swiping in the 2D approach.

#### **4.3.6 Qualitative evaluation approach**

After the interaction task, a questionnaire is completed by the users, evaluating and comparing the available interfaces (i.e. visual 2D/3D and SQL). The questionnaire was the tool to collect users' feedback and to provide a qualitative analysis.

The model of questionnaire that was used was based on the questionnaires provided by IBM in their research about new interfaces on usability tests [40]. In this case the questionnaire is separated into three main sections: the first two sections evaluate the user's experience with the interface (where section 1 considers aspects related with the interaction process and section 2 aims to evaluate the interface itself) and the third section compares the interfaces and the interactions with a traditional text based SQL approach, as it is shown in Table 1. In all questions the user provides a response from 1 to 5 to evaluate the interface, where 1 is the lower score (extremely negative evaluation) and 5 is the maximum

score (extremely positive evaluation).

Table 1: Usability questionnaire questions.

Section 1	Section 2 - How would you rate:	Section 3- SQL interface compared with the proposed visual approach
Q1: Was the interaction easy to understand?	Q4: The Interface?	Q10: The selection is easier than SQL?
Q2: Was it easy to manipulate?	Q5: The Performance?	Q11: The task is more intuitive than SQL sentences?
Q3: Is the navigation system intuitive?	Q6: The functionality?	Q12: Is it easier to learn the proposed visual approach than SQL?
	Q7: The objective achieved?	Q13: Is the task faster to perform than with SQL?
	Q8: The user experience?	
	Q9: The hand gestures selected?	

Also, at the end of the questionnaire, a last question is asked about the preference of the users over the three approaches (3D hand gesture, 2D hand gesture and typical text based SQL interaction). The scale of evaluation in this case is from 1 to 3, with 1 corresponding to the most preferable approach and 3 to the less preferable one, where the users have to rank the interfaces according their preferences.

## 5. Experimental Results

In order to evaluate the proposed interfaces, experiments were conducted using 29 subjects aged between 20 and 50 years old. Regarding the subjects, 63% were males and 37% were female; and 59% had knowledge of SQL and databases. Also the level of programming knowledge and experience was well distributed among all the subjects from novice to expert. The statistical validation of the data was performed using the Wilcoxon signed-rank test [41]; given the non-normally distributed nature of our data (the level of skewness is too high to be considered normally distributed).

### 5.1. Qualitative results

In this section, the results obtained correspond to the answers given by the users in the questionnaire presented in section 3.6. The evaluation of the interface from the users is summarised in the following tables and graphs. Table 2 shows the median values of the scores for the answered questions in both approaches (with

the median absolute deviation value shown in the brackets), in order to avoid the influence of outliers.

Table 2: Median values and median absolute deviation in each question for 3D and 2D interaction approaches.

<u>S 1</u>	<u>Q1</u>		<u>Q2</u>		<u>Q3</u>	
<b>Median 3D</b>	5.0 (0.1)		4.0 (1.0)		5.0 (0.9)	
<b>Median 2D</b>	4.5 (0.8)		3.5 (1.1)		4.5 (0.8)	
<u>S 2</u>	<u>Q4</u>	<u>Q5</u>	<u>Q6</u>	<u>Q7</u>	<u>Q8</u>	<u>Q9</u>
<b>Median 3D</b>	4.0 (0.8)	4.0 (1.0)	4.0 (0.8)	4.0 (1.0)	4.0 (1.0)	4.0 (0.9)
<b>Median 2D</b>	4.0 (0.8)	4.0 (1.0)	4.0 (0.9)	4.0 (1.0)	4.0 (1.1)	4.0 (0.8)
<u>S 3</u>	<u>Q10</u>		<u>Q11</u>	<u>Q12</u>	<u>Q13</u>	
<b>Median 3D</b>	4.0 (1.0)		4.0 (1.0)	5.0 (0.9)	4.0 (1.2)	
<b>Median 2D</b>	4.0 (1.0)		4.0 (1.0)	5.0 (0.8)	4.0 (1.1)	

As we can observe in the first section, the main positive point for the users is related to the first question, which is correlated to the complexity of learning the interaction in both approaches, indicating that the whole mechanism is intuitive and no significant prior knowledge or training is required. The other question with high positive evaluation is related to how much intuitive the system is (question 3), showing that the 3D approach is slightly superior to the 2D one. That can be related to the click movement, which presents a more natural process to select items (as it can be done in the real world). Also, both methods are almost equally evaluated in relation to the manipulation mechanisms (question 2). Furthermore, it can be seen, the interaction mechanisms are considered highly intuitive.

All the questions assessed in section 1 of the questionnaire have values over 3 that indicate a positive evaluation of the interaction process. Also, the 3D gesture based approach has an advantage over the 2D one, indicating that users prefer to interact using 3D hand gestures in a 3D interface. The median absolute deviation over the questions in this section indicates the responses given by the users are in general alike.

In the second section, it can be seen that both approaches have equal evaluation, all over 3.0, indicating both interfaces are well accepted by users.

Section 3 shows a clear advantage for both hand gesture interaction methods over the traditional SQL approach for the defined task (with almost all the questions ranked over 3.5 points), where the aspect with higher score is related to the learning process of the proposed visual approaches over traditional

interfaces (question 12).

In general, the results for both approaches are similar. Regarding the obtained values during the evaluation all of them are above 3.5, indicating the general acceptance of the new approaches based on hand movement both for 2D and 3D interaction environments.

Table 3: Median values and median absolute deviation (comparison between the 3 approaches), in a scale of 1 (most desirable) to 3 (less desirable).

	<i><u>Approach 3D</u></i>	<i><u>Approach 2D</u></i>	<i><u>SQL</u></i>
<b>Median</b>	1.7	2.0	2.4
<b>Median absolute deviation</b>	0.7	0.8	0.6

The answers to the last question regarding the users' preferences, comparing the three methods, show the superiority of the hand gesture based methods and particularly the 3D interface over the traditional approaches. Table 3 summarises the qualitative preferences of all the subjects (last comparative question) and as it can be seen, the 3D approach is preferred by the users over the 2D approach. This can be explained by the fact that the clicking (3D feature) provides a more stable and intuitive selection mechanism in 3D interfaces.

Regarding the qualitative analysis, the median score values given to all the sections versus different age ranges are shown in Figures 5, 6 and 7. The median absolute deviation is added to the plots.

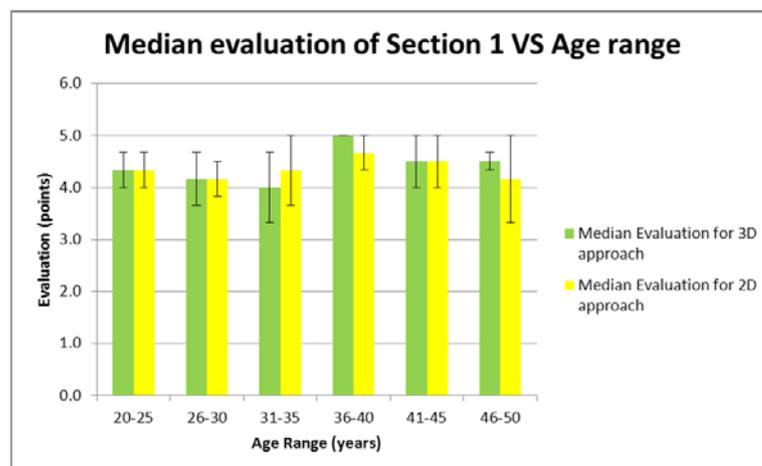
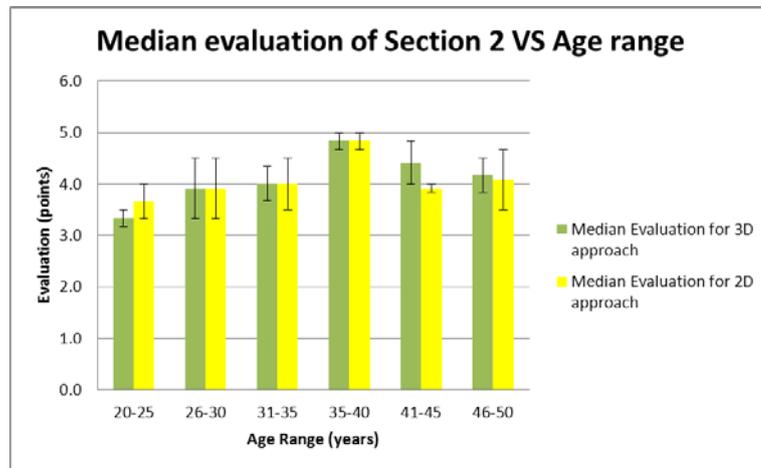


Figure 5: Median evaluation values of the section 1 of the questionnaire vs. age range (with median absolute deviation). The green bars show the scores given to the 3D approach; meanwhile the yellow bars show the score given to the 2D approach by different clusters of users' ages.

Figure 5 shows the median evaluation scores for all the questions on

section 1 (questions one to three) for the different age range of users. The users between 35 and 40 age range gave the highest scores to both hand gesture approaches, but with a clear advantage for the 3D ones. In general, except by the users between 31 to 35 years of age, both interaction approaches received scores over 4 (very positive) indicating that the visual gesture based interfaces provide a desirable way to interact with data.

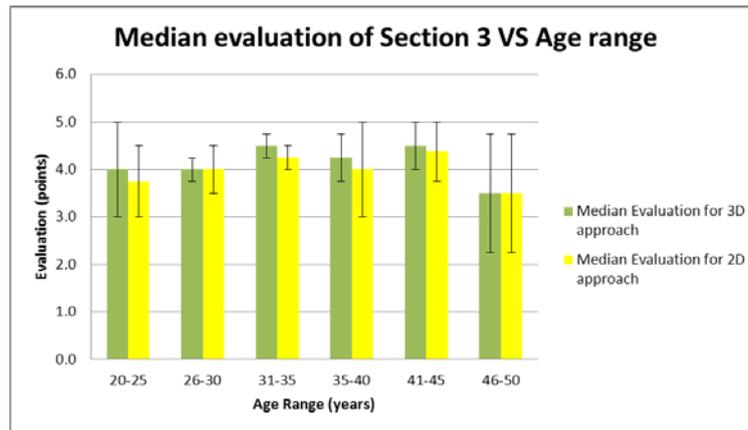
In the statistical evaluation the first section, the values obtained were *Wilcoxon Statistic* = 133,  $p < 0.05$  (one tailed), indicating that the median for the 3D results is significantly greater.



**Figure 6: Median evaluation values of the entire section 2 of the questionnaire vs. age range (with median absolute deviation). The green bars show the scores given to the 3D approach; meanwhile the yellow bars show the score given to the 2D approach by different clusters of users' ages.**

Figure 6 shows the median evaluation scores for all the questions on section 2 (questions 4 to 9), and again it can be seen that the group of users between 35 and 40 years of age gave the higher scores for both approaches, but in this case, both of them received the same median scores. In general, an advantage can be seen for the 3D gestures, but the evaluation for both approaches is a bit lower than in section 1. However, it is still higher than 3 indicating a good overall evaluation of the 3D interface.

The values obtained for the statistical evaluation were *Wilcoxon Statistic* = 173,  $p > 0.05$  (one tailed), indicating that the 3D results are not statistically significantly better than the 2D results for this section. This indicates the users consider equally acceptable both user graphic interfaces.



**Figure 7: Median evaluation values of the entire section 3 of the questionnaire vs. age range (with median absolute deviation). The green bars show the scores given to the 3D approach; meanwhile the yellow bars show the score given to the 2D approach by different clusters of users' ages.**

Figure 7 shows the results for the section 3 of the questionnaire (e.g. comparison between the hand gesture interfaces and traditional SQL). As it can be seen, the best evaluation was given by users in the range between 41 to 45 years of age, with an advantage for the 3D hand gesture approach. In general, again, the evaluation score is over 3, which indicates a preference of the users for the hand gesture interactions over the traditional text based ones. The lowest evaluation was given by users between 46 and 50 age range, which can be associated to the fact these users have been working generally with text- based interfaces and that would make hand gesture interfaces not that friendly for them, but still desirable. However the high median absolute deviation indicates different opinions between the users of this age range.

In the case of the third section, the values obtained were *Wilcoxon Statistic* = 139,  $p > 0.05$  (one tailed), indicating that the 3D results are not statistically significantly better than the 2D results for this section. This indicates the users consider both interfaces preferable to the traditional SQL text based interface similarly.

According to the results presented in the three plots, the significance tests in all sections, regardless of the age range, show the evaluation of the users is highly positive. This means that the hand gesture interaction in a 3D database representation presents advantageous features, especially in the area of understanding, learning and performance over traditional interfaces, especially when the hand gestures provide 3D features.

## 5.2. Quantitative results

In this section the results obtained by measuring time to complete the task, and the amount of errors occurred for each interface are analysed.

### 5.2.1 Time performance based evaluation

The overall median times for both approaches are shown in Table 4.

Table 4: Overall median time in seconds for 3D and 2D approach.

	3D Approach	2D Approach
<b>Median Time (seconds)</b>	22.7	23.8

As seen above, the users in general perform the task faster using the 3D hand gestures. This can be explained by the use of the “3D click” feature, which provides a faster selection of the columns than the swiping, since the rotation feature is the same in both.

Table 5 shows the average times for each gender and for users with and without SQL knowledge, highlighting the best time results.

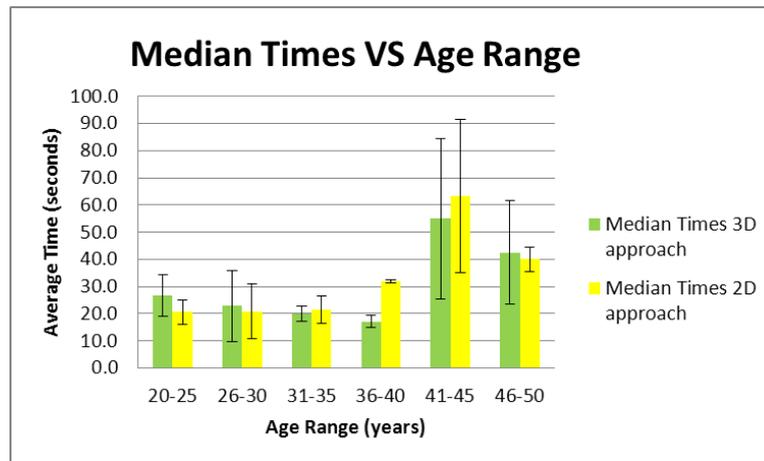
Table 5: Median time in seconds and age (in years) for males, females, people who knew SQL and people who did not know it.

	<u>Males</u>	<u>Females</u>	<u>SQL</u>	<u>No SQL</u>
<b>Median Age (years)</b>	32.0	32.5	34.0	29.5
<b>Median TimeAp 3D (seconds)</b>	22.1	<b><u>24.2</u></b>	<b><u>22.1</u></b>	23.1
<b>Median TimeAp 2D (seconds)</b>	<b><u>21.5</u></b>	43.9	29.0	<b><u>22.2</u></b>

It can be observed, the 3D gesture based interface outperforms the equivalent 2D one requiring less amount of time to complete the tasks in the females and SQL knowers groups. Furthermore, analysing the results over different sub-categories, it can be observed that men perform faster in 2D tasks than females, and also the people who know SQL perform tasks faster in the 3D approach than the people that do not have database programming knowledge. Additionally, there is no significant average age difference over all the available sub groups of users.

The time required to complete the tasks is further analysed providing a more accurate quantitative evaluation. Figures 8 to 12, presented below, show

how different aspects are related with the speed and the time required to accomplishing the tasks.

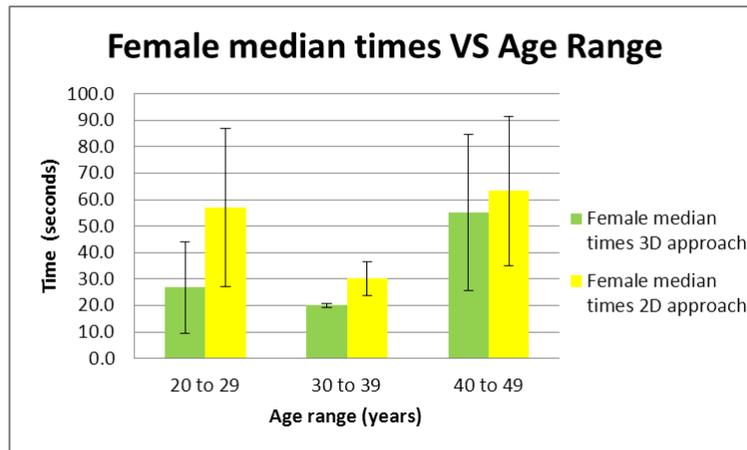


**Figure 8: Performance median times of all subjects based on their ages for both approaches (with median absolute deviation). The green bars show the performance times on the 3D approach; meanwhile the yellow bars show the performance times on the 2D approach by different clusters of users' ages.**

In figure 8 it can be seen that subjects between 26 and 40 age range have the best performance in time, especially in the 3D approach, while the slowest performance is for the subjects aged 41 to 45 range. Also, the 3D approach in general has better time performance than the 2D one. This can be explained by the previous use of graphical interfaces and other more interactive technologies, which would allow a better understanding of 3D interfaces and gesture interaction.

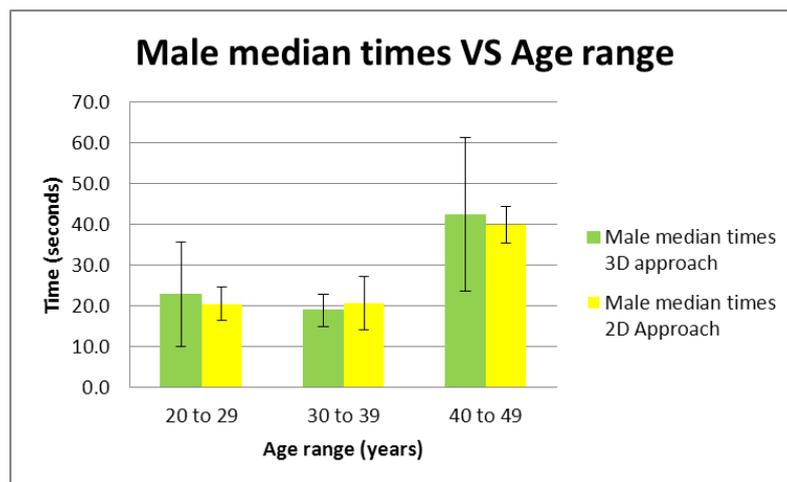
Users between 41 to 45 age range have the slowest time results, yet the highest median absolute deviation, which indicates high variations between the time performances between the users of this age group. In the opposite case, the users that have the best times (31 to 35 years old) have low median absolute deviation, indicating that the users with better results have a general good understanding and performance, which can be explained by a different level of knowledge on the use of interactive touch or gesture technologies.

Figures 9 and 10 present the results obtained by males and females respectively.



**Figure 9:** Performance time of female subjects based on their ages for both approaches (with median absolute deviation). The green bars show the performance times on the 3D approach; meanwhile the yellow bars show the performance times on the 2D approach by different clusters of users' ages.

In the case of females (Figure 9), the best performance for females is clearly in the age of 30 to 39 range for 3D and 2D approach, with a clear advantage for the 3D approach. That indicates a better understanding of the functionality and how to perform the different gestures, related with the clicking feature. It is also clear, that at this age range, the median absolute deviation of task's performance times is the lowest, which indicates the results in general are similar, especially for the 3D approach. For all age range, the 3D approach has better performance, which indicates the selection using clicking provides faster results than the swiping selection.



**Figure 10:** Performance time of male subjects based on their ages for both approaches (with median absolute deviation). The green bars show the performance times on the 3D approach; meanwhile the yellow bars show the performance times on the 2D approach by different clusters of users' ages.

For male subjects (Figure 10), the best result for the 3D and 2D approaches is in the range of 30 to 39 years of age, with an advantage for the 3D

approach (the same that happened with females) indicating that in general terms that age range have a better understanding of the use of 3D interfaces based on hand gestures. Also, the median absolute deviation for both approaches in this age range is low, indicating a low difference between the results obtained by each male user. In general, the 2D approach has better results, but the deviation of the data for the 3D approach in the range of 20 to 29 and 40 to 49 years of age is high, indicating some users in that age range may have more experience on this kind of interfaces than others.

According these results, it can be concluded that gender is not related to the performance in the group with best results (30 to 39 years old), since the results show the same tendency.

In figures 11 and 12 the time results for users with and without SQL knowledge are presented.

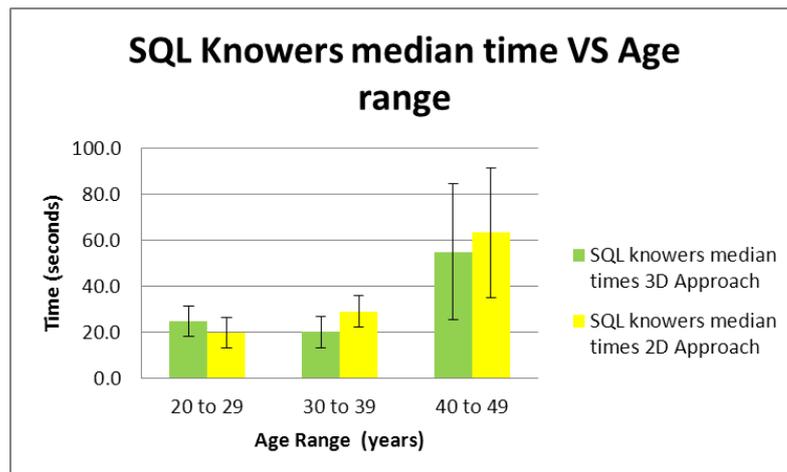
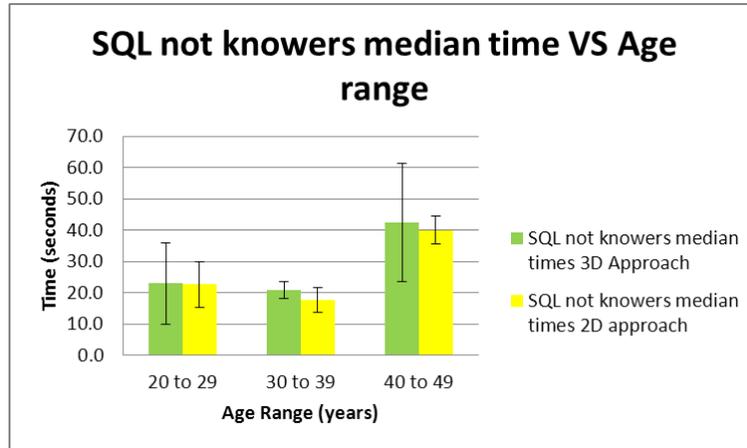


Figure 11: Performance times of subjects with knowledge of SQL for different age ranges (with median absolute deviation). The green bars show the performance times on the 3D approach; meanwhile the yellow bars show the performance times on the 2D approach by different clusters of users' ages.

For users with SQL knowledge (Figure 11), the best results in time performance using the 3D approach were obtained by those between 30 and 39 age range, while in the case of the 2D approach, the users between 20 and 29 years of age had the best results. In both ranges, the standard deviation was relatively low; indicating the performance of the users and that the level of understanding of the gesture interaction is similar for all the users in these age ranges. In the case of users in the range from 40 to 49 age, where the slowest results for both approaches were obtained, the 3D approach has an advantage, but

for both of them the standard deviation is high, which indicates the interaction was better understood for some users, which can be related to their previous use of similar interfaces.



**Figure 12: Performance times of subjects with no knowledge of SQL for different age ranges (with median absolute deviation). The green bars show the performance times of the 3D approach; meanwhile the yellow bars show the performance times on the 2D approach by different clusters of users' ages.**

In the case of users without SQL knowledge (Figure 12), the best results for both approaches are obtained by the users between 30 and 39 years old, with a clear advantage for the 2D approach. Also, this range of users presented lower median absolute deviation values in both cases. This advantage for the 2D gestures indicates users without SQL knowledge understand better swiping over clicking to perform a selection for the average of users in this age range, possibly related to previous use of similar technologies or better understanding of the interaction process.

Comparing the times obtained in both approaches, the statistical evaluation values obtained were *Wilcoxon Statistic* = 135,  $p < 0.05$  (one tailed), indicating that the median for the 3D results is significantly greater than the 2D ones.

### 5.2.2 User errors

In this section, the selection gestures are analysed. The overall median number of users' errors during the performance of the task for the 3D and 2D approaches are shown in Table 6.

Table 6: Median overall users' errors for 2D and 3D approaches.

		Median Errors
Median Error Selecting (amount)	3D (by clicking)	1.0
	2D (by swiping)	1.2

As it can be seen in the selection process, the 3D approach presents better results. This indicates that the process to select by clicking is easier to perform and the use of the depth information improves the overall interaction compared to the swipe movement. Also, this is due to the requirement of a more complex movement compared with the clicking which increases the possibility of errors. The median user errors for the 3D and 2D approach separated by gender and knowledge of SQL are shown in Table 7.

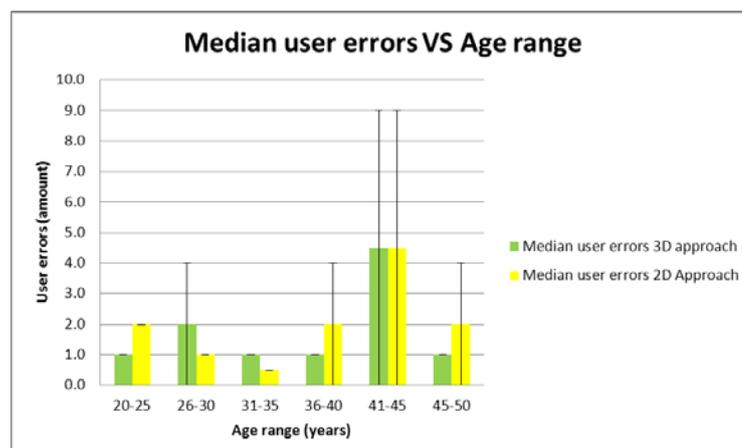
Table 7: Median amount of user errors for males, females, people who knew SQL and people who did not know it for 3D and 2D approaches.

2D and 3D Gesture approach Errors					
		<u>Male</u>	<u>Female</u>	<u>SQL</u>	<u>No SQL</u>
Median Error Selecting (amount)	3D (by clicking)	1.0	1.5	1.0	1.0
	2D (by swiping)	1.0	2.0	1.5	1.5

Table 7 clearly demonstrates female users made more mistakes using the swiping gesture (2D) than the clicking one (3D). Comparing with the results in Table 5 (median times), it can be said there is a direct relation between user errors and the time taken to perform the tasks.

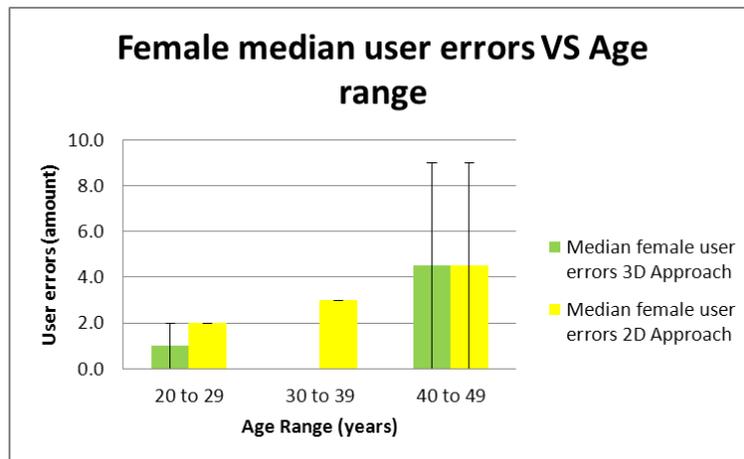
For users with SQL knowledge, regarding the selection, both groups of users had better performance in the 3D approach. These results (compared with the ones in Table 5) reinforce the theory of a direct relation between users' errors and time performance.

In order to understand the users' behaviour and the distribution of their errors, the total errors according to their age ranges are displayed in Figure 13.



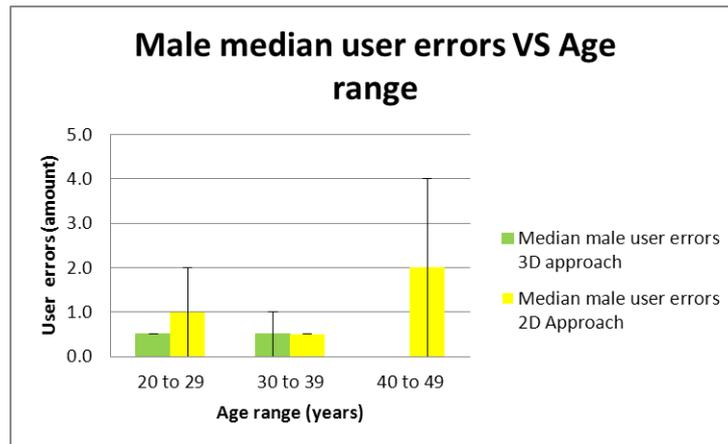
**Figure 13: Median user errors according the age range (with median absolute deviation). The green bars show the amount of user errors on the 3D approach; meanwhile the yellow bars show the amount of user errors on the 2D approach by different clusters of users' ages.**

Analysing the median user errors in each approach, presented in Figure 13, it can be seen that the worst performance was achieved by the group between 40-45 years old, but with a notably high median absolute deviation, showing that the median is not a clear indicator of the performance of the users and that some of them were capable of using the interface fairly easily. The best results for the 3D set of gestures can be seen in the ranges between 20 to 25, 31 to 40 and 45 to 50 years old, while the 2D interaction has fewer errors for users between 31 and 35 years old, but with more uneven results in general. These good performances on the 3D approach can be related to a better understanding of the clicking gesture.



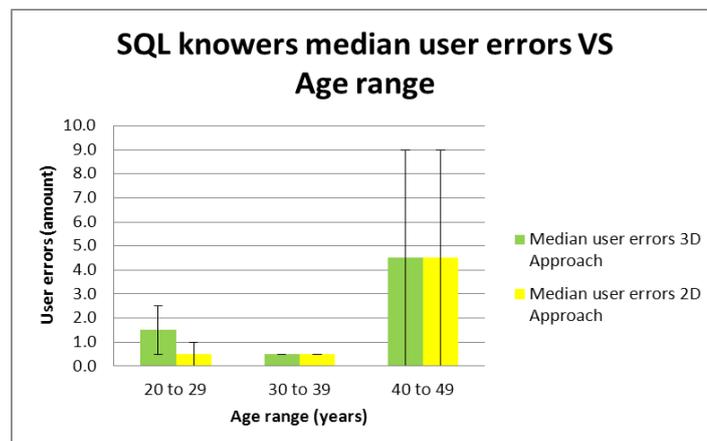
**Figure 14: Median female user errors according the age range (with median absolute deviation). The green bars show the amount of female user errors on the 3D approach; meanwhile the yellow bars show the amount of female user errors on the 2D approach by different clusters of users' ages.**

Female users made fewer errors on 3D gestures in the age range between 30 and 39 years of age, while female users between 20 and 29 years old had fewer mistakes in the 2D interface (see Figure 14). Also, the users between 40 and 49 age range present the highest median number of errors, but with high median absolute deviation, indicating that some users had a better understanding of the interface and manipulation than others, possibly related to previous use of gesture based technology.



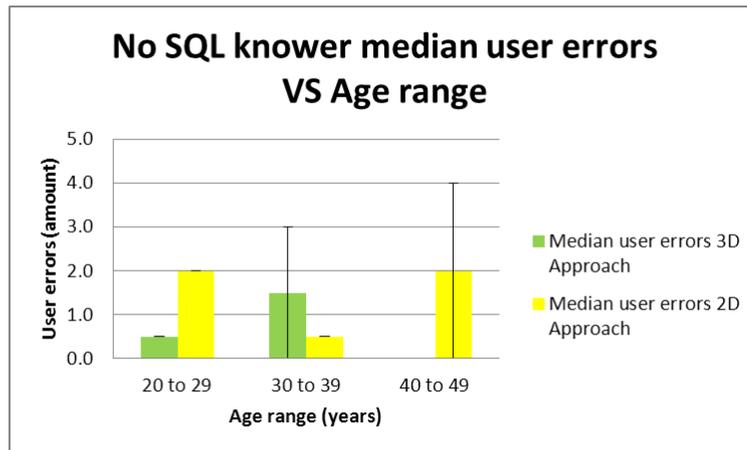
**Figure 15: Median male user errors according the age range (with median absolute deviation). The green bars show the amount of male user errors on the 3D approach; meanwhile the yellow bars show the amount of male user errors on the 2D approach by different clusters of users' ages.**

In the case of male users (see Figure 15), the best results were obtained by users between 40 to 49 for the 3D approach (with no errors, indicating accurate performance for this age range), and in the case of the 2D approach, the best results were obtained by the users between 30 and 39 years old. Also, the deviations were high in almost all the cases, which can be related to several factors, such as previous experience with gestural interfaces or better understanding in the use of the gestures for some users.



**Figure 16: Median errors for user with SQL knowledge according the age range (with median absolute deviation). The green bars show the amount of user errors on the 3D approach; meanwhile the yellow bars show the amount of user errors on the 2D approach by different clusters of users' ages.**

In the case of users with SQL knowledge (Figure 16), the lowest number of errors for the 3D and 2D approaches belonged to the group of users between 30 and 39 years old with low deviation. As it happened previously, the oldest users-40 to 49 years old have the worst results with the highest deviation.



**Figure 17: Median errors for user without SQL knowledge according the age range (with median absolute deviation). The green bars show the amount of user errors on the 3D approach; meanwhile the yellow bars show the amount of user errors on the 2D approach by different clusters of users' ages.**

In the case of users without SQL knowledge (Figure 17), the best result for the 3D approach belongs to the group between 40 to 49 years old with no errors, and in the case of the 2D approach, the group with best performance was the group of users between 30 and 39 years old. The slowest performances were for users between 30 and 39 years old for the 3D case; and in the group of users with ages from 40 to 49 in the 2D approach. Also, the deviations were high in some cases, which can be related to similar causes as the ones presented in the other cases (such as understanding of the gestures or previous experience with gestural interfaces).

In general terms, the previous figures show advantages for the 3D approach, especially in the case of the users with the best performance (30 to 39 years old). Also, the previous mentioned group presents in general the lowest deviation, indicating the consistency with the results presented in the section 5.2.

Finally, to evaluate the influence of the user errors on performance, in Figure 18 a comparison considering the number of errors is shown. In general, the number of errors was less than 3 (see Table 8). Also, it can be seen a correlation between the number of errors and the time required to accomplish the task.

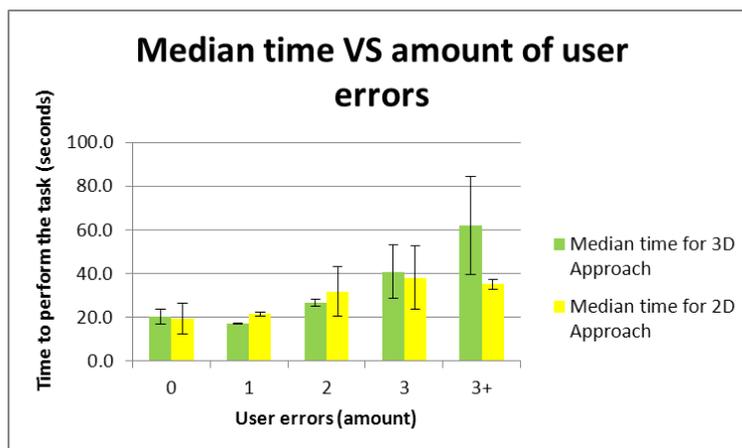


Figure 18: Median time according the amount of user errors (with median absolute deviation). The green bars show the time on the 3D approach; meanwhile the yellow bars show the time on the 2D approach by different clusters of amount of user errors.

Table 8 Amount users according amount of user errors.

		Amount of user errors				
		0	1	2	3	3+
Amount of users	3D approach	17	3	3	3	3
	2D approach	15	3	5	2	4

The general conclusion in this section presenting a new 3D interaction approach is that there is a clear correlation between the number of user errors and the time required to perform the task, but this is not always clear in the case of the 2D approach. Also, it has been shown that a 2D approach is more prone to errors and the use of a third dimension can help to improve functionality and reduce the number of users' errors during the execution of tasks in a 3D visual environment.

Comparing the amount of users' errors obtained in both approaches, the Wilcoxon signed-rank test values obtained were  $Wilcoxon\ Statistic = 122, p < 0.05$  (one tailed), indicating that the median for the 3D results is significantly greater than the 2D ones for the amount of users' errors during the execution of the evaluation task.

### 5.2.3. System Errors

The final metric to be analysed corresponds to the errors generated with the system in the selection process. These errors are associated to failures in identifying a gesture correctly performed. The results for this metric are shown in Figure 19.

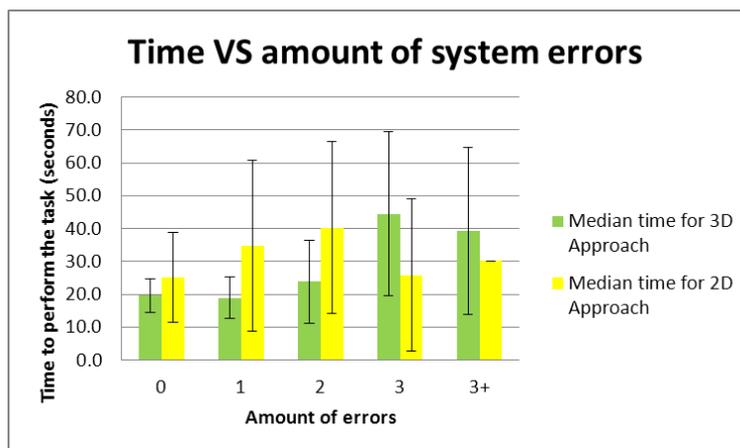


Figure 19: Median time according the amount of system’s errors (with median absolute deviation). The green bars show the time on the 3D approach; meanwhile the yellow bars show the time on the 2D approach by different clusters of amount of system errors.

As seen in Figure 19, there is a correlation between the number of system errors and the time to complete the tasks in the 3D approach. However, that is not clear in the case of the 2D interactions. Also, the number of errors in identification 3D and 2D gestures is concentrated between 0 and 2 (as it can be seen on Table 9). Figure 19 also shows that the 3D approach presents in general lower deviation than the 2D one, especially in the cases of fewer errors.

Table 9: Amount users according amount of system’s errors.

		Amount of system errors				
		0	1	2	3	3+
Amount of users	3D approach	8	7	4	2	6
	2D approach	13	6	6	3	1

Table 9 shows that the occurrence of system errors tends to be less than 2 for the 3D and 2D approach, but, in the case of the 3D approach, there are more users with more than three system errors than in the case of 2D. Also, the total median amount of errors (Table 10) shows a similar result, where the 2D approach has less system errors than the 3D one.

Table 10: Average overall system errors for the 2D and the 3D approach.

		Average Errors
Median. Error Selecting (amount)	3D (by clicking)	2.0
	2D (by swiping)	1.0

However, since the selection gesture (clicking) in the 3D approach can be performed faster than the same gesture in the 2D approach (swipe), the influence

of the system errors in the 3D approach is less in the time execution of the task. Also, new devices with more accurate movement detection and depth resolution (e.g. Kinect 2) could reduce these types of system errors, which can improve significantly the results in a 3D gesture interface.

Comparing the amount of system's errors obtained in both approaches, the values obtained for the statistical evaluation were *Wilcoxon Statistic* = 77,  $p < 0.05$  (one tailed), indicating that the 3D approach's results are statistically significant better than the 2D approach for the amount of system's errors during the execution of the evaluation task.

Table 11 shows the result obtained for the Wilcoxon statistical evaluation test for each of our experiments.

Table 11: Statistical significance values for the experimental data obtained.

Experimental Data	Amount of subjects involved (N)	Wilcoxon Statistical	P-Value
Qualitative questionnaire Section 1	29	133	0.04182
Qualitative questionnaire Section 2	29	173	0.35197
Qualitative questionnaire Section 3	29	156	0.14231
Average performance time	29	135	0.03515
Average amount of users' errors	29	122	0.03144
Average amount of system	29	77	0.00114

#### **5.2.4. Qualitative and quantitative results' comparison**

The comparison between the qualitative evaluation given by the users and quantitative results obtained by them (times and errors) are presented in this section.

The case of the qualitative results presents some interesting results. Figures 5 and 6, demonstrate users between 35 and 40 years of age give the best evaluation to both approaches to do with aspects of interaction and interface (sections 1 and 2 of the questionnaire respectively), with an advantage to the 3D approach. In the case of the comparison with a traditional SQL interface, the best evaluation to both approaches was given by the users between 41 and 45 age range. In the three graphs presented, there is not a clear correlation between age and preference over one specific approach, but it is clear that all the evaluations in the three sections are positive.

Figure 8, points out that the best time results for users' average time performance were obtained by those users in the age range between 30 and 40 years old. In the case of the users' errors (Figure 13), there is a similar result for the group with the highest number of errors, but that tendency is not that clear in the case of the best performances. This fact can be related to the users' speed to perform the gestures and the general task defined. In general terms it can be said the users with best performances are those with ages between 20 and 40 years old.

Comparing qualitative and quantitative results, there is not a clear relation between preferences by the users (qualitative evaluation) and their performances. This can be related to the novelty of the interaction approaches (especially in the case of the 3D approach) that makes the users feel comfortable with the interface, regarding the time required to perform the tasks or the errors committed during the interaction process.

## **6. Conclusions**

In this paper a 3D gesture based interface for 3D database interaction was presented. This interface is based on hand gesture interaction in 3D environments providing a more natural interaction to the end user. Furthermore, this approach incorporates different interaction methodologies, which were analysed providing all the details of their mechanics, allowing the definition of more complex interaction systems for data manipulation in future work.

To evaluate and validate this framework, two sets of interaction experiments were performed, using Microsoft Kinect to capture the hand gestures. The experiments indicate differences over the users' experience for the two models that were introduced (2D and 3D hand gesture interaction) operating both under the same interface. The experiments show a clear preference for the hand gesture interaction over the traditional keyboard and mouse based interfaces, and especially the 3D based approach. In general these interfaces operated by the users' hands, are regarded as a more suitable and effective input methods mainly for 3D than 2D tasks. Finally, it can be concluded that the proposed 3D hand gesture interface is more intuitive and less amount of time and training effort are required to understand and apply it on different tasks and mainly on 3D databases.

The range of ages that presented better results for all cases is between 20 and 40 years old, who required lower amount of time to perform the tasks resulting in a lower number of errors in general.

Additionally, there is a correlation between the amount of user errors and system errors with the time required to perform the tasks in the case of the 3D set of gestures. However, the number of errors in general is very low and does not affect significantly the users' performance and the users' satisfaction.

It can be argued that after the experiments presented in this paper, 3D interfaces based on hand gestures improve the users' experience, reducing the required learning time and the overall task procedure. Also, the application of these interfaces in data manipulation systems indicate a better understanding of the tasks by the users, especially when the hand gesture interactions support 3D gestures, providing a more natural and less complex gestures to interact with a system, simplifying the overall interaction process.

This paper presented a 3D approach to interact with databases in a 3D environment, aiming to improve the developer's interaction. But there is still a need to extend this kind of technology for other software development areas. Since the application development process for 3D interactive interfaces is better achieved in a 3D environment, experiments on that area are highly desirable.

## References

1. J.P. Wachs, M. Kölsch, H. Stern and Y. Edan, Vision-based hand-gesture applications, *Communications of the ACM* **54-2**(2011), 60-71.
2. N.T. Dang, M. Tavanti, I. Rankin and M. Cooper, A comparison of different input devices for a 3D environment, *International Journal of Industrial Ergonomics* **39-3**(2009), 554-563.
3. S. Clerici, C. Zoltan and G. Prestigiacomio, NiMoToons: a Totally Graphic Workbench for Program Tuning and Experimentation, *Electronic Notes in Theoretical Computer Science* **258-1**(2009), 93-107.
4. M. Rotard, D. Weiskopf and T. Ertl, A combined introductory course on human-computer interaction and computer graphics, *Computers & Graphics* **29**(2005), 267-272.
5. M. Harris, B. Buxton, W.T. Freeman, H. Ishii, M. Lucente and M.J. Sinclair, Interfaces for humans (panel): natural interaction, tangible data, and beyond, *SIGGRAPH 98 Conference abstracts and applications*(1998), 200-202.

6. P. Clifton, A. Mazalek, G.T. LCC, M. Nitsche, J. Sanford, G.T. COA and J. Murray, WorkTop: A Multi-touch Tabletop Collaborative Sketching Application for Interdisciplinary Design, *Master Thesis Georgia Institute of Technology* (2010).
7. A.F. Blackwell, Dealing with new cognitive dimensions, *Workshop on Cognitive Dimensions: Strengthening the Cognitive Dimensions Research Community., University of Hertfordshire* (2000).
8. L. Chittaro and R. Ranon, Web3D technologies in learning, education and training: Motivations, issues, opportunities, *Computers & Education* **49**(2009), 3-18.
9. P. Barr, J. Noble and R. Biddle, Video game values: Human-computer interaction and games, *Interactive Computers* **19-2**(2007), 180-195.
10. M. Wu and R. Balakrishnan, Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays, *ACM Special Interest Group on Computer Graphics and Interactive Techniques* (2003), 193-202.
11. K.P. Fishkin, A taxonomy for and analysis of tangible interfaces, *Personal and Ubiquitous Computing*, **8-5**(2004), 347-358.
12. M. Conway, S. Audia, T. Burnette, D. Cosgrove and K. Christiansen, Alice: lessons learned from building a 3D system for novices, *Proceedings of the SIGCHI conference on Human factors in computing systems* (2000), 486-493.
13. Y. Boussemart, F. Rioux, F. Rudzicz, M. Wozniowski and J.R. Cooperstock, A framework for 3D visualization and manipulation in an immersive space using an untethered bimanual gestural interface, *Proceedings of the ACM symposium on Virtual reality software and technology*(2004), 162-165.
14. P.L. Davidson and J.Y. Han, Synthesis and control on large scale multi-touch sensing displays, *Proceedings of the 2006 conference on new interfaces for musical expression* (2006). pp. 216-219.
15. T. Döring, A. Sahami Shirazi and A. Schmidt, Exploring gesture-based interaction techniques in multi-display environments with mobile phones and multi-touch table, *Proceedings of the International Conference on Advanced Visual Interfaces* (2010), 419.
16. J. Hernández-López, A. Quintanilla-Olvera, J. López-Ramírez, F. Rangel-Butanda, M. Ibarra-Manzano and D. Almanza-Ojeda, Detecting objects using color and depth segmentation with Kinect sensor, *Procedia Technology* **3-0**(2012), 196-204.
17. Guettard, E., et al. "Virtual reality from the keyboard/mouse couple to Kinect." *Annals of Physical and Rehabilitation Medicine* 54.10 suppl 1 (2011): 239.
18. Zhou Ren; Junsong Yuan; Jingjing Meng; Zhengyou Zhang, "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor," *Multimedia, IEEE Transactions on* , vol.15, no.5, pp.1110,1120, Aug. 2013.
19. Alexandros Pino, Evangelos Tzemis, Nikolaos Ioannou, Georgios Kouroupetroglou, Using Kinect for 2D and 3D Pointing Tasks: Performance Evaluation, Human-Computer Interaction. Interaction Modalities and Techniques, *Lecture Notes in Computer Science Volume 8007*, pp 358-367, 2013.
20. Stratos Idreos, and Erietta Liarou. "dbTouch: Analytics at your Fingertips." *CIDR*. 2013.

21. Vicknair, Chad, et al. "A comparison of a graph database and a relational database: a data provenance perspective." *Proceedings of the 48th annual Southeast regional conference*. ACM, 2010.
22. Gómez, Leticia I., Silvia A. Gómez, and Alejandro A. Vaisman. "A generic data model and query language for spatiotemporal olap cube analysis." *Proceedings of the 15th International Conference on Extending Database Technology*. ACM, 2012.
23. Mehdi, Muntazir, et al. "On-the-fly generation of multidimensional data cubes for web of things." *Proceedings of the 17th International Database Engineering & Applications Symposium*. ACM, 2013.
24. Bohak, Ciril, and Matija Marolt. "Kinect Web Kiosk Framework." *Human Factors in Computing and Informatics*. Springer Berlin Heidelberg, 2013. 785-790.
25. Hirte, Steffen, et al. "Data3--a kinect interface for olap using complex event processing." *Data Engineering (ICDE), 2012 IEEE 28th International Conference on. IEEE*, 2012.
26. Echtler, Florian, and Gudrun Klinker. "A multitouch software architecture." *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. ACM, 2008.
27. Kühn, Thomas. "The Kinect sensor platform." *Advances in Media Technology* (2011): 1.
28. Kaltenbrunner, Martin, et al. "TUIO: A protocol for table-top tangible user interfaces." *Proceedings of the The 6th International Workshop on Gesture in Human-Computer Interaction and Simulation*. 2005.
29. Smisek, Jan, Michal Jancosek, and Tomas Pajdla. "3D with Kinect." *Consumer Depth Cameras for Computer Vision*. Springer London, 2013. 3-25.
30. Xia, Lu, Chia-Chih Chen, and J. K. Aggarwal. "Human detection using depth information by Kinect." *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on. IEEE*, 2011.
31. Frati, Valentino, and Domenico Prattichizzo. "Using Kinect for hand tracking and rendering in wearable haptics." *World Haptics Conference (WHC), 2011 IEEE. IEEE*, 2011.
32. Tang, Matthew. "Recognizing hand gestures with microsoft's kinect." *Palo Alto: Department of Electrical Engineering of Stanford University:[sn]* (2011).
33. Vogel, Daniel, and Ravin Balakrishnan. "Distant freehand pointing and clicking on very large, high resolution displays." *Proceedings of the 18th annual ACM symposium on User interface software and technology*. ACM, 2005.
34. Ren, Zhou, et al. "Robust hand gesture recognition with Kinect sensor." *Proceedings of the 19th ACM international conference on Multimedia*. ACM, 2011.
35. Ren, Zhou, et al. "Robust part-based hand gesture recognition using kinect sensor." *Multimedia, IEEE Transactions on* 15.5 (2013): 1110-1120.
36. Li, Yi. "Hand gesture recognition using Kinect." *Software Engineering and Service Science (ICSESS), 2012 IEEE 3rd International Conference on. IEEE*, 2012.
37. Jadav, Jigna J., and Mahesh Panchal. "Association rule mining method on OLAP cube." *International Journal of Engineering Research and Applications (IJERA)* 2.2 (2012): 1147-1151.

38. Ludwick, Dave A., and John Doucette. "Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries." *International journal of medical informatics* 78.1 (2009): 22-31.
39. Lee, Unseok, and Jiro Tanaka. "Finger identification and hand gesture recognition techniques for natural user interface." *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction*. ACM, 2013.
40. Lewis, James R. "IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use." *International Journal of Human-Computer Interaction* 7.1 (1995): 57-78.
41. Rosner, Bernard, Robert J. Glynn, and Mei-Ling T. Lee. "The Wilcoxon signed rank test for paired comparisons of clustered data." *Biometrics* 62.1 (2006): 185-192.