Virtual Sculpture - Gesture-Controlled System for Artistic Expression

Mark Marshall*

*Interaction Design Centre
Computer Science and Information Systems Dept.
University of Limerick
Ireland
mark.t.marshall@ul.ie

Abstract

This paper describes the development of a gesture-based interface system for interactive multimedia applications. The system makes use of two-handed gestures, in the form of both free and direct gestures. This paper begins by describing the physical interface device that has been developed for the system, and then goes on to describe the first of the applications which have been developed to make use of the system in evaluating the use of different forms of gesture as a primary form of interaction for an interactive multimedia system.

1 Introduction

With the advent of ubiquitous computing the focus on mouse- and keyboard-based input for computing systems has begun to change towards a focus on less traditional forms of interaction which will allow user's to interact naturally and meaningfully with computer systems. As a result, there have been a number of systems to make use of various forms of gesture-based interaction. The natural, meaningful and rich form of control which gesture provides us with makes it especially useful for interactive multimedia systems.

As part of ongoing work in the area of gesture control, a system has been developed which allows for two-handed input in the form of both linguistic gesture, where the user's interaction is in the form of description of the action to be performed, and physically-based manipulation gestures, where the user directly grasps and manipulates the virtual objects.

Much of the work presented here builds upon earlier work performed as part of the Sounding Object (SOb) project. The Sounding Object Project[1] has pioneered recent attempts to couple physical simulations to efficient sound synthesis techniques. The European Commission funded this project to study new auditory interfaces for the Disappearing Computer initiative[2]. The work on this project included the investigation of gesture-based control of interactive sounding models, and the development of a virtual musical instrument, called the Virtual Bodhran (or Vodhran) (Marshall et al., 2002), which used gesture to interact with a sound model to simulate the playing of the traditional Irish instrument, the Bodhran.

The gesture control system which was created for the Vodhran, was limited in a number of ways. The system only tracked the position and orientation of the beater object held in the user's dominant hand, and the position of the user's secondary hand. From this data it extrapolated a secondary layer of data, such as direction of movement, speed of movement, and certain important events, such as sudden changes of direction. No provision was made for more detailed gesture capture, such as the detection of patterns of movement, or the detection of the positioning of the fingers of the hand. As a result of these deficiencies, it was decided to develop a more general and detailed gesture recognition system, the result of which is described here.

2 The Gesture-Interaction System

In order to allow for a broad range of gestures to be used as interaction to the system, it was decided to develop a system that could detect the movements of the hands, the orientation of the hands, and the position of the fingers. This would allow hand posture detection to be used to indicate commands, and hand rotation and movement to indicate the parameters of a command. The use of two-handed gesture was decided on for a number of reasons.

Firstly, the use of the second hand offers an intra-modal increase in interaction over a single-handed system (Bolt and Herranz, 1992). Also, the use of two hands for object manipulation tasks is more natural to the user. In his previous works on gesture, Hauptmann (1990) has found that for certain tasks, the use of two hands is more common than the use of one. More specifically he found that:

- For a translation task, users on average use 1.1 hands
- For a rotation task, users on average use 1.2 hands
- For a scaling task, users on average use 1.5 hands

Taking this in to account, we would expect a two-handed system to feel more natural for the user when ma-

[1] http://www.soundobject.org
[2] http://www.disappearing-computer.net

nipulating the objects, especially for the scaling manipulation.

2.1 The Gloves

In order to detect the posture of the hand, it is necessary to detect the position of the fingers. Many systems achieve this through the use of a number of bend-sensors placed along each finger. This allows for measuring of the bending at each joint in the finger and so can give a very accurate representation of the posture of the finger. However, for this system, this was deemed to be overly complex. By simply measuring whether or not a finger is bent, each finger can function as a simple on/off switch style of input. This gives us a total of 32 postures for each hand which, couple with information on the movement and rotation of the hand, should allow for a rich enough interaction with the system for many interactive applications.

This also allows for configuration of the system to match specific user's, including those with limited mobility of the fingers. Each bend sensor gives a range of output dependant on the amount of bending. While the system only uses the output from the sensors to detect whether a finger is bent or not, it is possible within the system to modify the threshold at which a finger is considered to be bent. Thus, while for a user with full mobility of the fingers, the sensor may have to indicate a bend of over 45° before the finger is considered bent, for a person of limited mobility of the fingers, this threshold could be set to a lower figure such as 10°.

Therefor, two gloves were built by attaching a single bend-sensor to the inside of each of the fingers of two gloves. This allowed for the measurement of the bending of a single joint in each finger. The output from each of the bend sensors was measured using a 10 channel, 10 bit analog-to-digital converter at a sampling rate of 10kHz, and sent to the PC. Tracking of the position and orientation of the hands was performed using a Polhemus Fastrak[3] position tracking device, with each glove having a single Fastrak sensor attached to the back of the hand. The Fastrak sensors allow us to track each hand with 6 degrees-of-freedom, and an update rate of 120Hz.

2.2 Gesture Recognition Software

There are two main forms of gesture interaction which the system is required to recognise. These forms are physically-based manipulation, where gesture is used to provide a multi-dimensional direct control over objects, and linguistic gestures, where gesture is used as a symbolic language. Weimer and Ganapathy (1987) viewed these two forms as being at opposite ends of the direct manipulation spectrum, however for the system described here it was decided that both forms of input must be recognisable by the system, although actual applications need not make use of both forms.

[3] www.polhemus.com

Each of these forms of gesture is based around the use of postures. A posture is a specific configuration of the fingers of the hand. For physically-based manipulation gestures the user places their hand (or hands) into a specific posture to indicate the operation being performed, and then moves the hand(s) to indicate the parameters of the operation. For linguistic gestures, the user makes a series of postures, to indicate a command to be issued to the system. For instance, an open hand posture, followed by a closed hand posture, followed once more by an open hand posture might indicate selection of an object under the hand.

In order to detect postures, the software system compares the input data from each of the bend sensors to the threshold for that sensor, to determine which fingers are bent, and which are not. Once this is determined, the system compares the data to the list of stored postures to determine which posture is currently being maintained.

To determine which gesture is then being made involves the use of a *feature tree*, such as the partial one shown in Figure 1. This method is based upon that of Jones et al. (1993).

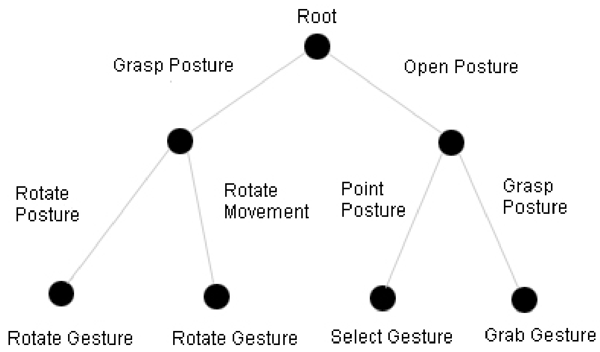
For instance, if the user makes and maintains a grasping posture with one hand, the traversal takes the leftmost branch. Then, if while maintaining this posture, the user rotates their hand (i.e. a rotation movement), then the traversal takes the rightmost branch which indicates a rotate gesture. This shows interaction with the system through a physically-based manipulation. However if the user were to perform the same action in a linguistic gesture based fashion, the user would first make a grasping posture, again causing traversal to the left, then a rotate posture, which would again cause traversal to the left, finally leading to a rotate gesture. Therefor the same application can activate a gesture in a number of ways, using either form of interaction. While the list of postures available is part of the gesture recognition system itself, the feature trees must be created by the specific applications, to indicate which gestures they will respond to.

3 The Virtual Sculpture Application

In order to demonstrate and test the gesture control system an application had to be developed which made use of gesture as its primary method of interaction, and which allowed for both of forms of gesture which the system recognises to be used either separately or together. As gesture is generally used as a form of interaction for the manipulation of physical objects in the real world it was decided to create an application which allowed the user to create and manipulate virtual objects using gesture.

A number of systems have previously been created which allow the user to create and interact with virtual objects using gesture, and these systems have used a number of different methods for the creation of the 3-dimensional

Figure 1: A sample Feature Tree, linking postures and movements to form gestures



objects by the user. The modelling of 3-dimensional objects using two-handed interaction was initially pioneered by Kruger (1993), whose VIDEODESK application allowed the user to control the shape of the objects being created using their own hands. The system made use of computer vision technologies to detect certain features, such as the positioning of the user's fingers and thumb.

A method to create freeform polygonal surfaces was shown by Shaw and Green (1997) which used two-handed interaction to create and manipulate control points on the shape through direct manipulation. A number of approaches to creating objects using *superquadrics* have also been proposed, such as Yoshida et al. (1996) which used a statistical method to calculate deformations of the shape from hand gestures, and Nishino et al. (1998) which allowed the user to create complex objects using two-handed gestures to perform blending and axial deformation of superquadrics.

Our Virtual Sculpture application makes use of a simpler form of object creation and manipulation. Rather than the use of superquadrics, or of meshes and control points, the system makes use of only a small number of geometric primitives, such as the sphere, cylinder, cone, torus and rectangular box. These objects are manipulated by simple axial transformations, which include scaling, rotation, and shearing. Objects can also be joined together to create a composite objects.

Objects are created by moving the hand over an icon representing the object at the top left of the display, creating a grasping posture, and "dragging" onto the display in the appropriate position.

Once an object is on the screen, it can be manipulated through a combination of physically-based manipulation and linguistic gestures. The system can be configured to use only linguistic gesture based controls, only physically-based manipulation controls or a mixture of both. This allows us to examine the use of the different forms of gesture to perform the same tasks.

Figure 2: Two methods of forming the rotation gesture

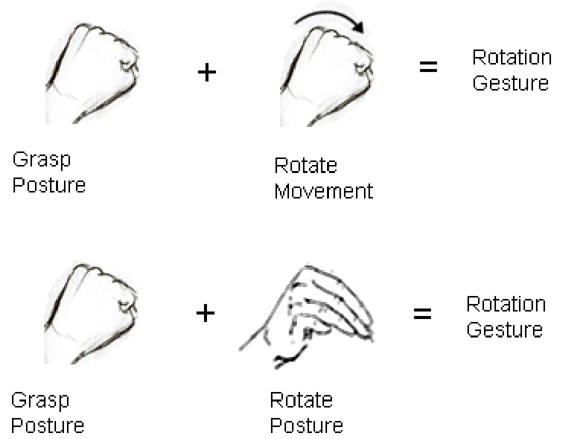
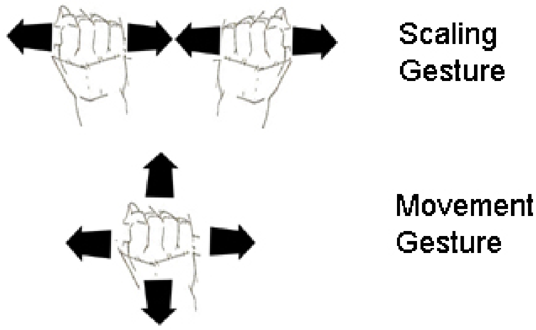


Figure 3: Movement and scaling gestures



When making use of the physically-based manipulation controls for the system, the user appears to act directly on the object. For instance, to rotate the object the user places a hand over the object and makes a grasping posture to indicate "grabbing" the object. The user then rotates their hand which causes the object to rotate through the same angles. When operating in a linguistic gesture based mode, the user indicates selection of an object by "grabbing" it as previously described, but then creates a number of different rotation postures, to rotate the object to the desired position. Figure 2 illustrates the two different approaches to rotating an object in the system. Some examples of other control gestures can also be seen in Figure 3.

3.1 Virtual Sculpture Installation

The installation for the Virtual Sculpture is a large, public space based installation. It makes use of a 100cm by 75cm display, which is built into a wall. The size of the display was chosen to allow for most people to comfortable reach the majority of the display area. The Virtual Sculpture application is then back projected onto this display. The gloves are placed on a small table in front of the display and connected to the wall of the installation. The user interacts with the system from a position standing in front of the display.

The decision to use this form of installation was made for a number of reasons. The size of the display allows most users to be able to reach the majority of the display, and also allows for the objects created to be fairly large, making it easier for the user to interact with the objects. It also makes it easier to observe the user during the testing process. Finally, it allows for others to observe the objects being created by the user, perhaps allowing for collaboration between the user and those watching. While the user interacts with the system from a standing position in this installation, it would also be possible to lower the screen and allow the user to work from a sitting position. This was not deemed necessary for this particular installation, as it was felt that interaction with the system would not be for so long a time that the user would begin to feel tired.

4 System Testing

The main aim of creating the Virtual Sculpture application and installation was to test the gesture interface system which has been developed, and to evaluate the use of different types of gesture in an interactive application.

In order to achieve this, testing is being performed with a number of users, in order to evaluate the ease of use of the gesture control system, and to examine which forms of gesture are suitable for which tasks. The testing involves the performance of certain object manipulation tasks, first using only linguistic gesture based controls, and then using only physically-based manipulation controls. Finally, the user's are asked to configure the controls with the gestures which the find most suit each manipulation, and to perform the tasks a final time.

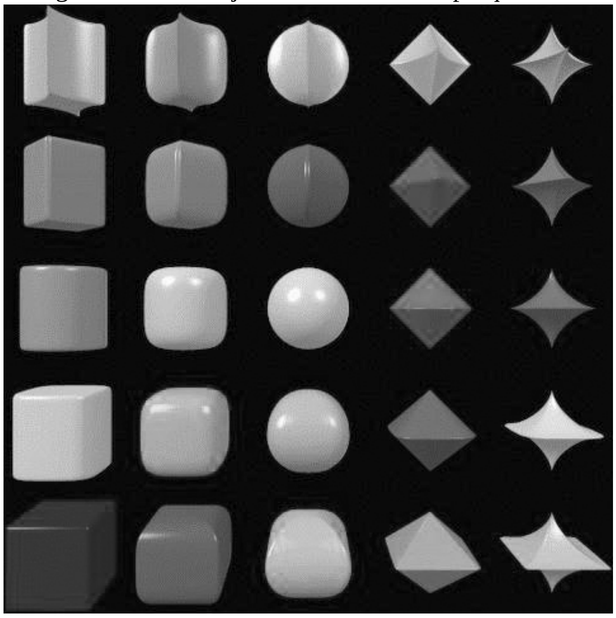
The task provided range from simple manipulations, such as creating an object and moving an object, to more complex tasks such as rotation only in one dimension by a set distance, to the most complex tasks which involve the creation of certain compound objects from the simple primitives provided.

By evaluating the results of this testing it will be possible to examine the suitability of a gesture interface to this form of interactive application, and also to determine whether user's find it most intuitive to use a physically-based manipulation gesture system, a linguistic gesture based system or a combination of each.

One strength of the system which has become particularly clear throughout its development and testing is that of the two-handed input. Performing manipulation of virtual objects using two-handed gesture is a very natural method of interaction for the user. This has been shown clearly from the comments of users as they interact with the application.

One manipulation which shows this in particular is that of scaling. Scaling an object using a two-handed gesture, where the hands are held in the grasping posture, and moved along the direction in which we would like the object scaled has proven to be extremely natural and no single-handed gesture has been found which feels as natural to the user. This bears out our initial reasoning behind the use of two-handed input, and also confirms the findings of Hauptmann (1990).

Figure 4: Some objects created from superquadrics



5 Further Development

The use of simple objects and transformations in the Virtual Sculpture system means that it is not particularly suited to the creation of intricate or detailed 3-dimensional objects, but rather to the creation of more simplistic objects. This decision was initially made to allow for concentration on the interaction with the system and on the use of gesture-based controls in particular.

In order to make the application more useful for the creation of 3-dimensional objects, and especially more useful as a form of artistic expression, a more complex set of objects will be required, along with a larger set of transformations and manipulations.

The use of superquadrics rather than the current simple geometric primitives would be one way of allowing the user to create more complex objects. Objects created in this way would allow for a greater control over the shape of the object, and would also respond to a larger set of transformations, including bending and twisting transformations which are not available in our system at present. See Figure 4 for examples of some shapes created from superquadrics.

However, to allow for the creation of really complex objects, an even greater degree of control might be required. In this case it may be necessary to develop the system to create objects from 3-dimensional meshes, so that particular points on an object can be manipulated, allowing for the distortion of the surface of an object as well as transformation of the object as a whole.

For further developments in the area of the gesture control interface itself, it may become necessary to treat the input from the bend sensors on the gloves as an analogue value rather than as a binary switch value. While the decision to use the inputs as switches was made for the reasons given earlier, namely that of ease of use for those with limited mobility, and as it gave a rich enough vocabulary of gestures for many applications, it still provides us with a limit to the number of different postures which the system can detect.

Another improvement which has been suggested for the system would be the ability of the user to train the system to certain gestures, rather than have the gestures set only by the application. This enhancement to the system would allow a user to set their own control gestures, which would be useful in the case of gestures which were not felt to be natural enough for the user, and might also be useful for users of limited mobility, which was one of the initial ideas behind the development of the system. By allowing users with limited mobility to set their own gestures, we eliminate the possibility of the system making use of a gesture which the user is physically unable to perform, or which might in any way be uncomfortable for the user.

As well as reducing any discomfort for the user, the tailoring of the system to the users mobility may allow for easier interaction with the system by the users. Tailoring the motion input of a system to the user has proven in the past to produce an increase in the success rate of the users at manipulating the system, to the point of producing success rates for disable users which clearly overlapped those of non-disabled users for a large portion of their range (Pausch et al., 1992).

Finally, the addition of some form of haptic feedback to the system might result in some advantage. Some users commented on the lack of any "real feeling" when using the system, which might be accomplished by the addition of some haptic feedback, to indicate when the hand is over an object, and also to give some impression of grasping objects. This would increase the realism and therefor would be hoped to increase the naturalness of the interface for the user (Burdea, 1996). Haptic feedback has been shown to be a major integral part of many tasks, and Dai (1998) has shown that for the process of object prototyping, which is a very similar process to that used in our system, it is a necessity.

6 Conclusion

This paper presented a recently developed gesture interaction system which can be used as a main form of input to an interactive multimedia application. It also presented an application, called Virtual Sculpture, which demonstrates the use of the system as the main input into an interactive application, and which can be used to evaluate the suitability of the system as an input device, and also to examine the forms of gesture most suitable for use in these interactions. It discussed the use of this system for such an evaluation, and detailed an installation of the system in a public environment for use as a tool for artistic expression. Finally, some possible enhancements to the system were presented, which may be incorporated into a future version of the system.

References

R. Bolt and E. Herranz. Two-handed gesture in multimodal natural dialog. In *Proceedings of the 5th Annual ACM Symposium on User Interface Software and Technology*, pages 7–14, Monteray, California, USA, November 1992.

G.C. Burdea. *Force and Touch Feedback for Virtual Reality*. John Wiley and Sons, Inc., New York, USA, 1996.

F. Dai. *Virtual Reality for Industrial Applications*. Springer, New York, USA, 1998.

A. Hauptmann. Speech and gestures for graphic image manipulation. In *Proc. Of CHI89 Human Factors in Computing Systems*, pages 241–245, New York, USA, 1990. ACM Press.

M. Jones, R. Doyle, and P. O'Neill. The gesture interface module, GLAD-IN-ART deliverable 3.3.2. Technical report, Trinity College, Dublin, Ireland, January 1993.

M.W. Kruger. Environmental technology: Making the real world virtual. *Communications of the ACM*, 36 (7):36–37, 1993.

M. Marshall, B. Moynihan, and M. Rath. The vodhran - design and development of a virtual instrument. In *Proceedings of ICMC 2002*, Goteborg, Sweden, 2002.

H. Nishino, K. Utsumiya, and K. Korida. 3d object modelling using spatial and pictographic gestures. In *Proc. Of ACM Symposium on Virtual Reality Software and Technology*, pages 51–58, New York, NY, USA, 1998. ACM Press.

R. Pausch, L. Vogtle, and M. Conway. One dimensional motion tailoring for the disabled: A user study. In *Proc. Of SIGCHI Conference on Human Factors in Computing Systems*, pages 405–411, Monterey, California, USA, 1992. ACM Press.

C. Shaw and M. Green. THREAD: A two-handed design system. *ACM Multimedia Systems*, 5:126–139, March 1997.

D. Weimer and S.K. Ganapathy. Interaction techniques using hand tracking and speech recognition. In M. Blattner and R. Dannenberg, editors, *Multimedia Interface Design*, Frontier Series, pages 109–126. Addison-Wesley Publishing Company, 1987.

M. Yoshida, Y. Tijerino, T. Miyasato, and F. Kishino. An interface system based on hand gestures and verbal expressions that generate shapes for 3d virtual objects. *Tech. Report of IEICE, MVE95-64*, pages 33–40, 1996.