Development of CAD Interface Using Leap Motion

Shrey Pareek  
Mechanical Engineering  
University at Buffalo, State University of New York  
Buffalo, USA  
shreypar@buffalo.edu

Vaibhav Sharma  
Mechanical Engineering  
University at Buffalo, State University of New York  
Buffalo, USA  
vsharma6@buffalo.edu

Abstract – The following study proposes a Leap Motion based CAD interface that uses number, position, velocity and direction of fingers as input data in order to draw, extrude, scale, translate and rotate an object in the 3D space. The system allows the user to generate basic CSG primitives and advanced geometries (geometries that cannot be realized using the CSG primitives) and to perform basic the CAD operations described above. As opposed to traditional systems wherein gestures are used to carry out operations, the proposed system uses switches that operate on simple binary principles thus reducing the computational cost of the system by eliminating the use of a classifier scheme to a high extent. A user study involving 2 subjects is also presented in order to determine the qualitative and quantitative efficacy of the system.

Keywords - Leap Motion; CAD; interface; 3D

I. INTRODUCTION

The use of a pen and paper to lay down the conceptual design is a method adopted by architects, engineers and designers around the world. Design issues (if any), are first resolved on paper before recreating them using computer design software. This technique though simple and effective, leads to a certain degree of redundancy during the design process and delays the use of sophisticated modern engineering until the later stages of the process. This problem has been attributed to the lack of computer user interfaces that combine the intuitiveness of pen-paper system along with the accuracy of CAD interface.

Development of such computer based interfaces has been carried out by various researchers. Sketch based systems have been developed by Schmidt et al. [1], Sezgin et al.[2] and Herold et al. [3]. These systems usually employ a touch based system and allow the designer to ‘pen in’ the design onto the interface. The interface then processes the data or ink. Processing may include classification [4]-[6] as well as segmentation [3], [7], [8]. These sketch based systems are limited to two dimensions only and realization of 3D geometries is a complicated task. Since most CAD geometries involve the use of the 3D design space these systems need to be modified accordingly.

These modifications generally employ the use of optical 3D sensors. These sensors allow acquisition of 3D data and are used for industrial tasks, people and object tracking, motion analysis, character animation, 3D scene reconstruction and gesture based user interfaces [9], [10]. Optical sensors can be classified into the following categories on the basis of their operating principles - Structured Light, Stereo Vision and Time of Flight. Structured light sensors such as Microsoft Kinect® (http://www.microsoft.com/en-us/kinectforwindows/) analyze the deformation of known pattern on an unknown surface (object) in order to determine the three dimension space. Stereo Vision Cameras use two optical 2D cameras with known extrinsic parameters. Depth of the scene is determined on the basis of correspondence points between the two 2D images[10].

The following study focuses on the use of Stereo Vision based Leap Motion® (https://www.leapmotion.com/) for the development of a 3D CAD interface. The Leap Motion Controller along with the current Application Programmer Interface (API) identifies and returns the position of objects such as finger tips, pen tips etc. It uses three IR emitters and two IR cameras (Figure 1) for position identification. The controller is typically used for hand gesture recognition using the above finger position data. The origin of the Cartesian System is located at the center of the Leap Motion.

![Schematic representation of Leap Motion Controller](http://www.leapmotion.com/)

Figure 1-Schematic representation of Leap Motion Controller [10].

The following study proposes a Leap Motion based CAD interface that uses number, position, velocity and direction of fingers as input data in order to draw, extrude, scale, translate and rotate an object in the 3D space. Using basic Constructive Solid Geometry (CSG) primitives (cone, sphere, triangular prism, cylinder and torus) the interface allows the user to perform basic CAD operations described above. An advanced drawing mode is implemented, enabling the user to draw and extrude irregular geometries or geometries that cannot be realized using the basic CSG primitives. A user study involving 2 subjects is also presented in order to determine the effectiveness of the system. It should be noted that the study presents a basic proof of concept of the CAD interface and gesture recognition extension into other CAD operation can be realized using future studies.
II. The Interface

The interface consists of two windows or phases, viz. the initialization window and the CAD window. The interface has been designed to completely eliminate the use of keyboard and mouse inputs, and is solely controlled using hand gestures. The initialization window (Figure 2), comprises of a graphical space and a set of three hand positions. The user needs to perform these gestures in order to advance to the CAD phase of the interface. These three actions are used to record the working position of the hand, which is self-selected by the user. This position is then used as a reference for drawing operations to be carried out during the CAD phase. Detailed explanation of the initialization phase is provided in subsequent sections.

The second window (Figure 3) or the CAD window, is the actual interactive interface where the user performs the required CAD operations. The window consists of three graphical areas and one text box. The Sketch Window (Top Left) is where the user can choose between the basic CSG primitives by drawing Arabic numerals between 1 and 6, or draw a different geometry using the advanced drawing mode. Once a number is drawn, it is classified using the classifier developed by Kara et al. [5]. The primitive related to each number (1-6) and the drawing gesture have been explained in later sections. If a ‘drawing’ is not classified as a number, the advanced drawing mode is activated.

The advanced mode (Figure 4) has been designed to incorporate geometries other than the basic primitives described in the previous section. Once the user has finished drawing the 2D geometry, the required geometry is displayed in the Advanced Window (bottom left). The user then uses the extrude gesture (in the Sketch Window) to extrude the given geometry in 3D space. The extrusion can be carried out along a straight line or a Bezier Curve. In order to facilitate easier visualization of the finger position w.r.t. the Leap Motion, a ‘virtual’ Leap Motion is centered at the origin of the graphical spaces.

Once the initial 3D geometry (basic or advanced) has been realized, the user activates the Main Window (right graphical space) using right swipe gesture. It is here that the user can perform CAD operations such as scaling (uniform and non-uniform), rotation and translation using their respective gestures. The interface allows the user to switch back to the drawing mode (left swipe gesture), wherein the user can draw a new shape and carry out operations on it. Further, an assembly comprising of two separate parts can also be realized using the system. Finally the user can close the interface by using a

---

1 The x-y-z axes shown in the figure have been inserted to promote a better understanding of the system for the reader. It should be noted that these axes are not displayed in the original interface.
downward swipe gesture in the Sketch Window. As mentioned previously, these gestures are explained in detailed in the subsequent sections. Also, the primitives associated with each number and the number of fingers required for each gestures are described using a text box in the far right corner of the CAD Window.

III. CLASSIFICATION SCHEME

Following the work of Kara et al. [5], an image based classifier is used for character recognition. They present a rotation invariant ensemble based classifier. Classification is based on the sum of the normalized scores of four classifiers viz. Hausdorff Distance, Modified Hausdorff Distance, Tanimoto Coefficient and Yule Coefficient. The classifier uses only one sample per class for training, is rotation invariant, is unaffected by over-stroking and is independent of the starting and end point of the stroke. These advantages make it a viable choice for classification and also give a higher degree of freedom to the user for drawing the required character.

Feature based classifiers such as Rubine’s Classifier [4] and Adaptive Boosting [6], were also considered. However due to higher inaccuracies on account of missing time stamps for some data points (preventing the calculation of features requiring time data), these classifiers were rejected for use. Further these classifiers require that the user follow the same stroke path as described in the training data, thus making the system less intuitive.

IV. GESTURES

The following section explains the gestures used by the interface and their initialization using the initialization window. The number of fingers required for each gesture along with the graphical area in which the gesture needs to be performed is explained. The interface uses a set of eight gestures or switches (refer Appendix 1) in order to carry out the required operations. The use of switches enables the system to operate in a binary condition eliminating the use of a classifier for identification and invariably decreasing the computation cost. It should be noted that the aforementioned eight gestures/switches are different from the six (1-6) numbers that are drawn to identify the primitives.

1. Draw
   Fingers: 2 (Recommended: Thumb and Index)
   Window: Sketch
   Gesture: Make Thumb and Index Parallel.
   Operation: If the angle between the two fingers is less than a threshold, the user is able to start the data collection process for classification.

2. Extrusion
   Fingers: 3 (Recommended: Thumb, Index and Middle)
   Window: Sketch
   Gesture: Make the three fingers parallel to each other.
   Operation: If the dot product between the 3 fingers is less than a threshold, the user is able to define the path for extrusion operation.
   If the path has 2 vertices the extrusion is performed along a straight path.
   If the path has more than 2 vertices the extrusion is performed along a Bezier curve approximating the path.

3. Translation
   Fingers: 1 (Recommended: Index)
   Window: Main
   Gesture: Point the finger directly at the screen (point inclusion not required).
   Operation: The user is able to translate the 3D object.

4. Rotation
   Fingers: 2 (Recommended: Thumb and Index)
   Window: Main
   Gesture: Make the two fingers parallel to each other.
   Operation: The user is able to rotate the 3D object about its geometric center -
   a. Along X – If the mean direction of the two fingers is along the X axis.
   b. Along Y - If the mean direction of the two fingers is along the Y axis.

5. Uniform Scaling
   Fingers: 3 (Recommended: Thumb, Index and Middle)
   Window: Main
   Gesture: Spread or converge the fingers.
   Operation: The user is able to uniformly scale the 3D object. If the distance between the three fingers is less than a threshold, the object is down scaled as vice versa.
6. **Non-Uniform Scaling**  
   Fingers: 4 (Recommended: Thumb and Index of Left and Right hand)  
   Window: Main  
   Gesture: Hold out the fingers and move the hands closer or further away.  
   Operation: The user is able to scale the 3D object about the Y-axis.  
   If the distance between the hands is less than a threshold, the object is down scaled as vice versa.

7. **Close**  
   Fingers: 1 (Recommended: Index)  
   Window: Sketch  
   Gesture: Swipe swiftly downward.  
   Operation: The user is able to close the window and terminate the program if the downward velocity is greater than the threshold.

8. **Switch windows**  
   a. **Right-swipe**  
      Fingers: 1 (Recommended: Index)  
      Window: Sketch  
      Gesture: Swipe swiftly rightward.  
      Operation: The user is able to switch to the main window to perform 3D operations if the rightward velocity is greater than the threshold.  
   b. **Left-swipe**  
      Fingers: 1 (Recommended: Index)  
      Window: Main  
      Gesture: Swipe swiftly leftward.  
      Operation: The user is able to switch to the sketch window to perform drawing operation if the leftward velocity is greater than the threshold.

---

**Initialization Phase**  
In order to make the program adaptive to each user. The first window is used to decide the threshold for draw gesture. In this mode, the user has to extend the Index and Thumb in the L-shape. The user is then instructed to perform the gestures in the same order as shown in Figure 5. Upon Completion the user performs the close (down) gesture. The angle threshold for the draw gesture is determined using the mean of the maximum angle and the minimum angle between the two fingers.

In order to promote a better understanding of the system, the steps required to operate the interface have been listed as follows.

Step 1.  
Use draw gesture to draw-  
1 for Cylinder  
2 for Cone  
3 for Prism  
4 for Cube  
5 for Sphere  
6 for Torus  
Any other shape for advanced mode

Step 2.  
   a. If sketch is classified as a number between 1 and 6, the corresponding 3D geometry is created in the Main Window.  
   b. If sketch is not classified as a number, the advanced drawing mode is activated and extrusion operation is carried out in the Sketch Window.

Step 3.  
Perform right-swipe gesture to switch to Main Window. The user can perform CAD operations viz. translation, scaling or rotation.

Step 4.  
Perform Left-swipe gesture to return to the Main Window to create the next shape.

Repeat steps 1 to 4 until the user performs the close gesture.

---

**V. USER STUDY**

In order to get a qualitative and quantitative estimate of the efficacy of the system a user study was conducted on two subjects (one left handed and one right handed). The subjects were initially trained on the system, following which they were instructed to perform predefined operations on the interface. Their responses were recorded after completion of the study.
Following this the users were instructed to draw numbers from 1-6 and classification accuracy of the system was recorded. Finally the subjects were instructed to insert a cylinder into a torus (Figure 6). The above assembly requires the use of all the aforementioned eight gestures to be realized completely. The number of trials required to obtain the required geometry was noted.

VI. RESULTS AND DISCUSSIONS

The system shows an average drawing gesture recognition accuracy of 92.85% (13/14). The average classification accuracy of the system is observed to be 81.25% (13/16). The first subject required 3 trials to correctly realize the required assembly. The second subject however required only one trial to arrive at the same results. It should be noted here that the second subject was present during the training and testing phases of the first subject as well. Thus it can be safely assumed that the second subject had higher expertise in terms of understanding of the system and was thus able to perform the operations better (both qualitatively and quantitatively) as compared to the first subject. Use of a right-handed and a left-handed subject asserts the ambidextrous abilities of the system.

Also higher accuracy of drawing gesture/switch recognition as compared to the classification accuracy of the image based classifier[5] points towards the advantage of using switches as compared to traditional gestures. It should however be noted that these switches can be used only in applications where the number of operations are relatively lower. Use of a large number of switches makes the system less intuitive and requires more training for the subjects.

VII. CONCLUSION AND FUTURE WORK

The study proposes a novel switch based CAD Interface that uses finger count, position, velocity, and direction as inputs in order to draw, extrude, scale, translate and rotate an object in the 3D space. The interface uses Leap Motion® for data collection. As opposed to traditional systems wherein gestures are used to carry out operations, the proposed system uses switches that operate on simple binary principles thus reducing the computational cost of the system by eliminating the use of a classifier scheme to a high extent.

The system allows the user to generate basic CSG primitives and advanced geometries (geometries that cannot be realized using the CSG primitives). Employing an Image Based Classifier[5] during the drawing mode enables the system to differentiate between the basic and the advanced drawing modes. The system also enables the user to generate an assembly comprising of two parts. Further the interface is self-adapting for different users, which makes it more intuitive and accurate.

Finally a user study involving two subjects is presented, that gives a qualitative and quantitative assessment of the system. The system is ambidextrous, as both left-handed and right-handed users are able to use it with relative ease. A higher accuracy in case of switches as compared to classifications of numbers points to the aforementioned advantage of using switches over traditional gestures.

It should be noted that the current system is presented just as a proof of concept. CSG operations such as subtraction and intersection have not been covered in this study and can be explored in future works. Also the system is limited to assemblies comprising of at most two parts. In order to make the system applicable as a CAD package, the assemblies need to be more practical. The recently launched Leap Motion 2®, enables differentiation of the right hand from the left hand and can also identify the fingers being used. These features along with the new pinch gesture can be used to make the system more intuitive for the user. Use of a multi modal system that may use voice commands and brain signals in addition to the hand gestures as inputs is another vista being explored. This multi modal system will present higher accuracy and intuitiveness to the system.

REFERENCES

# APPENDIX 1

Table of Gestures

<table>
<thead>
<tr>
<th>Draw</th>
<th>Extrude</th>
<th>Translate</th>
<th>Rotate</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Draw" /></td>
<td><img src="image2" alt="Extrude" /></td>
<td><img src="image3" alt="Translate" /></td>
<td><img src="image4" alt="Rotate" /></td>
</tr>
<tr>
<td>Uniform Scale</td>
<td>Close</td>
<td>Right Swipe</td>
<td>Left Swipe</td>
</tr>
<tr>
<td><img src="image5" alt="Uniform Scale" /></td>
<td><img src="image6" alt="Close" /></td>
<td><img src="image7" alt="Right Swipe" /></td>
<td><img src="image8" alt="Left Swipe" /></td>
</tr>
<tr>
<td>Non-Uniform Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>