Using Motion Capture to Manipulate and Edit Meshes

Joe Istead[*] Roy Eagleson[†] and Sandrine de Ribaupierre[‡]

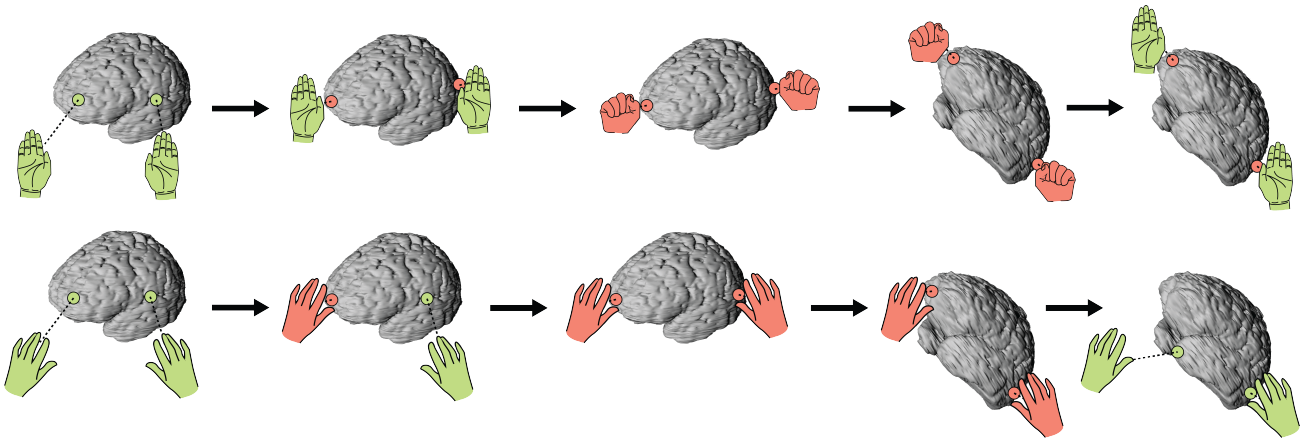

Figure 1: *Two sequences of gestures for mesh rotation: Grab and Release (top); Pin and Manipulate (bottom).*

CR Categories: I.3.4 [Computer Graphics]: Graphics Utilities—Graphics editors; I.3.4 [Computer Graphics]: Graphics Utilities—Virtual device interfaces; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality;

Keywords: hand gestures, motion capture, mesh manipulation and editing

Links: DL PDF

1 Introduction

We are in the initial stages of the design of a system that uses Microsoft's Kinect interface to support gesture-based interaction with segmented 3D medical imagery. By harnessing the skeletal information provided by the Kinect motion capture device, our application enables spatial interaction between the user avatar and meshes using hand motions and gestures. In particular, the user selects vertices nearest their hands and then scales, translates and rotates the mesh with natural, intuitive gestures. This system of interaction enables the user to view and manipulate complex graphical objects in real-time, a tool that we envision will allow medical students to interact with 3D biomedical scans, and share 3D collaborative spaces in order to plan and reason about surgical procedures.

More specifically, we demonstrate two different gesture systems for selecting vertices and transforming the mesh that only rely on the position of the user's hands. The first system employs an intuitive "Grab and Release" metaphor that distinguishes between open and closed hands–a task limited by the lack of hand configuration data. The second system employs a less intuitive "Pin and Manipulate" metaphor that is better suited to hand position-only motion capture data. Our initial analysis focuses on speed and accuracy tradeoffs; however, later analysis demonstrates a Fitts-like Index of Performance for selecting vertex positions in 3D.

[*]e-mail: jistead@uwo.ca
[†]e-mail:eagleson@uwo.ca
[‡]e-mail:sderibau@uwo.ca

Copyright is held by the author / owner(s).
SIGGRAPH 2012, Los Angeles, California, August 5 – 9, 2012.
ISBN 978-1-4503-1435-0/12/0008

2 Grab and Release

The user touches their hands against the surface of the mesh, closing their hands to grasp it. Then, while still grasping the mesh with closed hands, the user repositions their hands to scale, rotate and translate the mesh. The operation completes when the user opens their hands, releasing the transformed mesh.

While this gesture system represents a very natural and intuitive approach to manipulating meshes, it must distinguish between open and closed hands, an impossible task given only the position of each hand. Instead of registering the hand configuration from the Kinect camera or using voice recognition, our grab and release system used a sequence of hold times: after moving into vertex proximity, the user holds their hands still for a short duration to close the hands; after performing the transformation, the user holds the hands still for another duration to open the hands, releasing the mesh and ending the operation. Even after fine-tuning the radius of proximity and hold times, we found this gesture system unsuitable for relatively small transformations due to spatial jitter during the hold times. The approach also requires more physical and mental effort compared to the pin and manipulate gesture system.

3 Pin and Manipulate

The Pin and Manipulate gesture system employs a "pin the tail on the donkey" metaphor where the user fixes an initial vertex in place and then manipulates a second vertex about that fixed point. The system requires the user to hold the pinned hand in place until they wish to terminate the transformation by moving the hand away from the pinned vertex.

This system is suitable for smaller transformations since there are no hold times; furthermore, such an approach requires less physical and mental effort. Unfortunately, fixing one vertex reduces the transformation space, so this requires the user to perform extra transformations. As well, combining translations and rotations is taxing since the rotation must maintain a fixed distance; similarly, translating and scaling in one transformation is impossible.