Slice-n-Swipe: A Free-Hand Gesture User Interface for 3D Point Cloud Annotation

Felipe Bacim[1], Mahdi Nabiyouni[2], and Doug A. Bowman[3]
Center for Human-Computer Interaction and Department of Computer Science, Virginia Tech

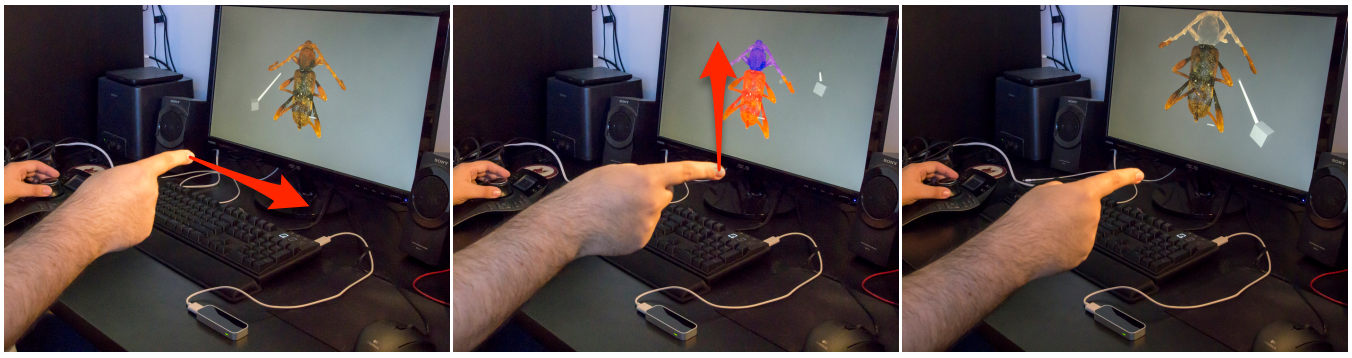

Figure 1: Three stages of the Slice-n-Swipe technique: the user prepares to slice the dataset (left); the user slices the dataset, resulting in two subsets (center); the user swipes away the unwanted subset (right).

ABSTRACT

Three-dimensional point clouds are generated by devices such as laser scanners and depth cameras, but their output is a set of unstructured, unlabeled points. Many scenarios require users to identify parts of the point cloud through manual annotation. Inspired by the current generation of "natural user interface" technologies, we present Slice-n-Swipe, a technique for 3D point cloud annotation based on free-hand gesture input. The technique is based on a chef's knife metaphor, and uses progressive refinement to allow the user to specify the points of interest. We demonstrate the Slice-n-Swipe concept with a prototype using the Leap Motion Controller for free-hand gesture input and a 3D mouse for virtual camera control.

Keywords: Free-hand gesture, 3D interaction techniques, natural user interfaces, progressive refinement.

1 INTRODUCTION

This year's 3DUI contest required entrants to design a system for the annotation of 3D point clouds. This is an important task due to the preponderance of point cloud data coming from laser scanners, depth cameras, and other common technologies. The data produced by such devices may be highly accurate, but it also has significant limitations due to lack of structure and semantics. In other words, 3D point clouds are simply collections of 3D points without any information about what structure(s) each point belongs to. Thus, it is important to have a user interface allowing knowledgeable users to annotate these point clouds.

Such an interface requires attention to three key interaction tasks: *viewpoint manipulation* (i.e., camera control) to obtain a reasonable view of the region of interest, *selection* of the set of points to be annotated, and *text entry* for the annotation. Of these tasks, selection is the most difficult and complex.

{[1] fbacim, [2] nabiyoun, [3] bowman}@vt.edu

IEEE Symposium on 3D User Interfaces 2014
29 - 30 March, Minneapolis, Minnesota, USA
U.S. Government work not protected by U.S. copyright

Although Yu et al. [5] provide a solution for selecting multiple points in point clouds and Coffey et al. [2] introduce a navigation and annotation method for 3D datasets, both use 2D input devices such as touchscreens. A 3D user interface might provide a more direct mapping for the selection of multiple 3D points.

In designing our solution, we were inspired by the current generation of so-called "natural user interface" (NUI) technologies [4], which track the user's hands, fingers, or entire body in 3D, without the encumbrances of wires or devices that must be held or worn. The Leap Motion Controller is an example of a NUI device designed for high-precision hand and finger tracking in *free-hand* gesture interfaces.

Free-hand gestures can provide effective and natural methods for 3D interaction with 3D datasets, which can result in fluidity and immediacy of the interaction [3]. However, simply because the interface is based on "natural" gestures does not reduce the need for careful interface design. In fact, because of the inherent ambiguity and lack of precision of everyday gestures, designers must think deeply about how to make free-hand gesture interfaces usable and effective, while still feeling natural and fluid.

Using the Leap, we designed a point cloud annotation system we call Slice-n-Swipe. Based on a chef's knife metaphor, Slice-n-Swipe allows users to iteratively cut the dataset and remove unwanted points until the desired selection is achieved.

2 TECHNIQUE DESIGN

Our design process included brainstorming about devices and metaphors, paper prototyping, and iteration based on discussion of sketches and storyboards. Our prototype system enables users to perform all three subtasks involved in 3D point cloud annotation: selection, viewpoint manipulation, and text entry. For each subtask, we describe our design and its rationale.

2.1 Selection

The foundation of a successful point cloud annotation system must be a technique that allows the user to select the set of points that she wishes to annotate. In designing a selection technique based on free-hand gesture, we wanted a metaphor that was easy to understand and that lent itself to gestures that could be easily detected by the Leap device. We considered various volume

selection approaches in which the user specified the size, shape, and position of a selection volume, but found that these required precise input by the user and could involve difficult-to-detect gestures. To mitigate these problems, we proposed an approach based on progressive refinement [1], in which the user roughly indicates the region of interest and then refines that selection via additional steps. The metaphor we settled on was that of a chef's knife, which is used to quickly cut off and swipe away unwanted parts of the food until the desired portion is all that remains.

This metaphor has several advantages: 1) it does not require precision from either of the gestures (slice or swipe); 2) it is easy for users to understand; 3) it avoids the specification of parameters like shape and size of a selection volume; 4) it does not require tracking of the orientation of the hand or the positions of multiple fingers. This last point is critical when working with the Leap device. It excels at tracking the position of a single finger, but has difficulty continuously tracking multiple fingers or the orientation of the hand, due to occlusion and range issues.

Figure 1 illustrates the basic flow of the Slice-n-Swipe technique. The user inserts a single finger into the Leap workspace, and the 3D position of the fingertip and the hand are tracked by the Leap and displayed in the virtual workspace. A rapid movement of the finger is detected as a slice, and the system then creates a plane containing three points: the positions of the fingertip at the beginning and end of the movement, and the position of the hand at the end of the movement. The plane is infinite, so that precise depth or positioning of the finger is not required. The plane divides the 3D points into two subsets, which are highlighted with different colors.

A second rapid movement of the finger, roughly perpendicular to the plane, is detected as a swipe, and the system removes the subset of points in the direction of the swipe from the selection set. The viewpoint then automatically zooms in to the set of selected points. Alternatively, the user can undo the slice by waiting a couple of seconds without performing any gestures.

The user can continue to slice and swipe until the desired set of points is selected, or she can perform an "open hand" gesture to reset the selection to the entire dataset.

2.2 Viewpoint Manipulation

To enable easy use of the Slice-n-Swipe technique, the user must be able to view the region of interest from a comfortable camera location. In our system, we use a 3DConnexion SpacePilot Pro 3D mouse for viewpoint manipulation with the non-dominant hand, which enables the user to perform the more precise task of selection with the dominant hand.

The 3D mouse provides six-degree-of-freedom isometric input, which enables camera translation and rotation in the 3D environment. Translational force causes the virtual camera to translate in the scene, while rotational force causes the camera to rotate about the center point of the 3D dataset. This makes it easy for users to view the dataset from any angle and zoom toward or away from the dataset.

2.3 Text Entry

Once the desired set of points has been selected via Slice-n-Swipe, the user provides an annotation by typing on the keyboard (Figure 2). The annotation appears in the 3D environment near the set of points to which it refers, and stays near those points as the user manipulates the virtual camera. The system allows as many annotations as desired, and points can have multiple annotations. For example, in the point cloud of a beetle shown in Figure 2, the user can select the head with both antennae and label those points as head, and then select each antenna to label them individually,

Figure 2: The user performing an annotation on the selected points

creating a hierarchy of annotations. The order in which annotations happen does not affect the final annotation hierarchy.

3 Discussion and Future Work

The goal of our work was to provide a solution for selection and annotation of multiple points in a point-cloud dataset and have a better understanding of how to design free-hand gestures for the selection task. We believe that our techniques provide the user with an effective and natural method to manipulate the viewpoint and select the desired subset of points in the dataset. The primary innovations and contributions of our work include the following:

- The design of a free-hand gesture interface based on a novel chef's knife metaphor
- The use of progressive refinement to avoid requiring precision in free-hand gestures
- Avoiding the limitations of the Leap device while still enabling interaction that feels natural and fluid
- Providing a two-handed interface to allow rapid and controllable viewpoint manipulation and selection

We will continue to improve and expand on this work by exploring different free-hand gesture metaphors and tools. Improved visual feedback mechanisms are needed to help the user understand how the dataset will be sliced and which points will be removed. We plan to refine the viewpoint manipulation technique to work effectively in a variety of scenarios. We also hope to investigate alternate viewpoint manipulation possibilities through the use of a tracked head-mounted display. Of course, we will iterate our design through usability studies to measure performance, find usability problems, and gather feedback about the effectiveness of our system.

References

[1] F. Bacim, R. Kopper, and D. A. Bowman. "Design and Evaluation of 3D Selection Techniques Based on Progressive Refinement." *International Journal of Human-Computer Studies*, 71, no. 7-8 (2013): 785-802.

[2] D. Coffey, N. Malbraaten, T. Bao Le, I. Borazjani, F. Sotiropoulos, A. G. Erdman, and D. F. Keefe. "Interactive Slice WIM: Navigating and interrogating volume data sets using a multisurface, multitouch VR interface." *Visualization and Computer Graphics, IEEE Transactions on* 18, no. 10 (2012): 1614-1626.

[3] G. Ren and E. O'Neill. "3D selection with freehand gesture." *Computers & Graphics* (2013).

[4] D. Wigdor and D. Wixon. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Morgan-Kaufmann (2011).

[5] L. Yu, K. Efstathiou, P. Isenberg, and T. Isenberg. "Efficient Structure-Aware Selection Techniques for 3D Point Cloud Visualizations with 2DOF Input." *Visualization and Computer Graphics, IEEE Transactions on* 18, no. 12 (2012): 2245-2254.