Go'Then'Tag : A 3D point cloud annotation technique

Manuel Veit*
Holo3
Saint Louis, France

Antonio Capobianco†
ICube
University of Strasbourg,
France

ABSTRACT

Performing complex tasks within a 3-D virtual environment, such as labeling 3-D point clouds, is still an ongoing research challenge. We propose Go'Then'Tag (GTT), a tool that helps users to edit complex 3-D data sets, and label them at different levels. It allows to quickly select a subset of points in the 3-D point cloud and edit a tag associated to the selection. It also comes with a tool to edit and manipulate a tree of tags, that allows overlapping hierarchies of annotations. The interaction is handled through a tracked multi-touch device that combines 2-D and 3-D interaction techniques in a single device.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [User Interfaces]: Input devices and strategies—Interaction styles

1 INTRODUCTION

We propose Go'Then'Tag (GTT), a system that allows users to annotate 3-D point clouds. GTT relies on two core ideas: the use of progressive refinement to facilitate the selection and the use of a tracked multi-touch interaction device.

- *Pre-selection:* To reduce the complexity of the selection task, we want to implement the idea of progressive refinement [2] . For this purpose, we propose a pre-selecting algorithm that generates a tree of sets of points. The user should be able to easily select one or several nodes of the tree for annotation. Our tool should also allow to easily edit the current selection in order to add or remove single points. The more efficient the pre-selection, the easier the labeling task.

- *Tracked smart-phone :* The intrinsic 3-D nature of the tasks of manipulation and selection of objects should be performed with a 3-D pointer. However, labeling data sets, manipulating labels and editing text are more easily achieved using a 2-D interface. For this reason, we want to explore the design possibilities given by the combination of multi-touch input and 6-DOF tracking. While a previous study led to the conclusion that tracked multi-touch mobile devices do not bring performance advantages when compared to best-practice techniques for 3D interaction [3], we think that this approach can be interesting for annotating 3-D data sets. It will allow us to use the tracked device as a 3-D pointer for interaction tasks with the data set. The tactile screen of the device will be used for editing texts and editing the hierarchy of tags.

2 TECHNIQUE DESIGN

The standard manipulation scenario with our tool separates the edition and manipulation of the labels from the selection and manipulation of sets of points. The latter is performed in the 3-D environment using a tracked multi-touch device and a ray-casting technique [1]. The former is performed through the tactile screen of the device. It allows to easily edit text and manage the hierarchy of labels. Multi-touch techniques are also used to orientate the 3-D point cloud.

2.1 Creating a set of points

We defined several possible actions to select and edit sets of points.

- *Navigating in the pre-selected hierarchy:* at the beginning of the task, all the points are selected and are divided in subparts. In our proof of concept tool, this division is performed using an octree generation algorithm. However, the division algorithm could be any other algorithm that might infer from the data-set plausible groups of points based on the geometry of the object, the proximity of the points, etc. The user can navigate in the subdivision by selecting a node of the subdivision with the 3-D pointer. When a node is selected, the subdivision of the current selection is displayed. With this approach, the user successively refines its selection, to select the node that contains the group of points that is the closest to the actual data set he wants to select.

- *Select stand alone points or define stand alone groups of points:* The user might want to select stand alone points or small groups of points that do not fit the pre-selecting algorithm. He can do so in order to either remove them form the current selection, or add them to the group of points he is creating to adjust finely the current selection. To this purpose we created a pencil tool that allow easy selection of stand alone points or groups of points. The user can select simple geometric shapes (sphere, cube, etc.) to shape this tool. The shape is attached to a ray that is controlled by moving the device in the environment. The size of the shape can be modified at will by the user using multi-touch interaction. When in selection mode, any point that is included in the radius of influence of the tool is included in the selection. The user can then choose to remove the points selected with the pencil from the current set of points. He can also choose to add this group of points to a set that has already been selected.

- *Orientation:* In order to select a node or any stand alone point, the user must be able to orientate the data set. When a node is selected, the position and scale of the current data-set are automatically calculated to provide a point of view adapted to the current selection. The user can then finely adjust the orientation by rotating the selection around the x and y axis through sliding gestures on the tactile screen.

- *Merging:* If the group of points are in several nodes, the user can make multiple selections of nodes and then merge them in a custom group of points.

At any time during the interaction with the data-set, multimodal feedback, including visual, audio and vibrotactile information are given to the user.

*m.veit@holo3.com
†a.capobianco@unistra.fr

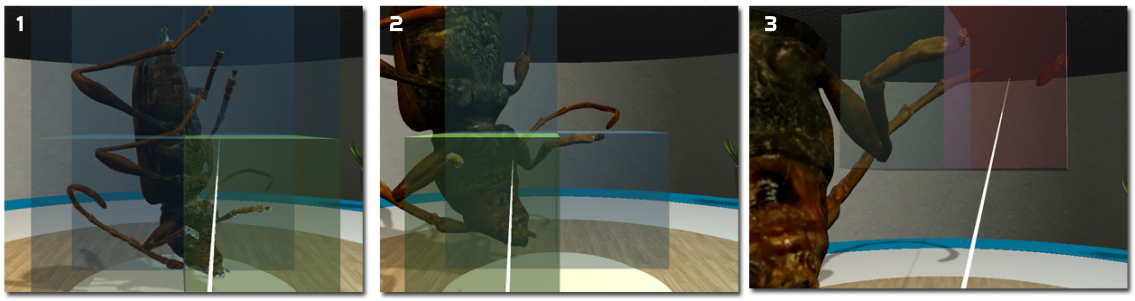

Figure 1: The user select successive nodes in the pre-selection to navigate through the data. Green boxes represent nodes that contains other pre-selections and red boxes represent leaves.

2.2 Creating and managing labels

Once the user has defined a group of points that he wants to annotate, he can choose to enter in annotation mode. The set of actions that are available are the following:

- *Define a label:* in this mode, the user can edit a label using the keyboard of the device. Through the multi-touch device this action can be performed efficiently without creating a break-out in the interaction.
- *Edit the hierarchy:* a visual interface is given to the user to handle the hierarchy of labels. Using this interface, the user can suppress, modify or move any item in the hierarchy of labels.

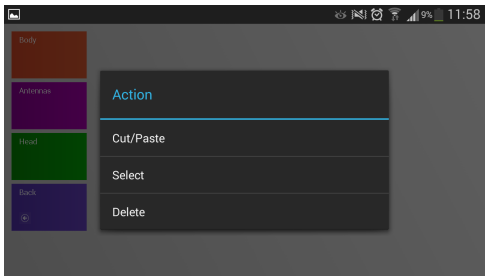

Figure 2: The hierarchy can be accessed through the smart-phone. In this case, the longhorn contains three parts (Body, Antennas and Head) that can be manipulated using the menu invoked by a long press on one item.

3 Technique development and novelty

3.1 Development environment

The environment is composed of one 3-D stereoscopic display with an optical tracking system capturing the head and device positions. The multi-touch device used in this environment is a Samsung Galaxy Note 2 running Android 4.3. The bi-directional communication between the smart-phone and the application has been developed specifically for this application. The software has been developed on Unity3D and MiddleVR for Windows 7.

3.2 Novelty

We think that the novelty of our technique relies on the combination of its two core ideas. The use of a tracked multi-touch device allows to combine advantages of both 2-D and 3-D interaction in a single interaction paradigm.

The 3-D pointer is used to perform the tasks that are intrinsically 3-D, namely the selection of items in the data-set. Since 3-D point clouds can be overwhelming data, due to the density of the information displayed, we implemented the idea of progressive refinement.

The tactile interface is used to perform the tasks that are more easily performed using a 2-D interface. These tasks are: editing the text of the labels and handling the hierarchy of tags.

We try to give greater comfort to the user by letting him chose the visualization context whenever possible. For example, the hierarchy of tags can be manipulated and visualized both in the environment, and on the screen of the smart-phone.

Using these interaction principles, we are able to create a tool that is versatile, easy to use, efficient for the task at hand, and that can adapt to the users preferences.

4 Future work

4.1 Evaluation

The tool was informally tested by several non expert users but no quantitative data were collected during these evaluations. We are now planning to more formally evaluate this tool on several data-sets (those proposed in the context of the 3DUI contest and others). The methodology's definition is still ongoing.

4.2 Design

We plan to work on an improved use of multi-touch interaction during the coarse and fine selection (for richer view positioning possibilities on the current data-set). We are also working on finding a way to better integrate the smart-phone in the environment. This can be done, for instance, by visualizing (using animations) the exchanges between the virtual environment and the device (e.g. when an annotation is created on the smart-phone and sent to the environment, or when the annotation hierarchy is edited).

We also want to try several other algorithm, based for example on the proximity of the points or on the geometry of the original data set. This would allow our tool to offer pre-selections that could be the most relevant according to the data properties.

References

[1] T. Grossman and R. Balakrishnan. The design and evaluation of selection techniques for 3d volumetric displays. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 3–12. ACM, 2006.

[2] R. Kopper, F. Bacim, and D. A. Bowman. Rapid and accurate 3d selection by progressive refinement. In *3D User Interfaces (3DUI), 2011 IEEE Symposium on*, pages 67–74. IEEE, 2011.

[3] C. B. Wilkes, D. Tilden, and D. A. Bowman. 3d user interfaces using tracked multi-touch mobile devices. In *Joint Virtual Reality Conference of ICAT-EGVE-EuroVR*, pages 65–72. The Eurographics Association, 2012.