The Point Walker Multi-label Approach

Conference Paper · March 2014
DOI: 10.1109/3DUI.2014.6798884

6 authors, including:

Theodoro Mota
Universidade Federal do Rio Grande do Sul
2 PUBLICATIONS  2 CITATIONS

Matheus Toazza Tura
Universidade Federal do Rio Grande do Sul
2 PUBLICATIONS  1 CITATION

Anderson Maciel
Universidade Federal do Rio Grande do Sul
92 PUBLICATIONS  366 CITATIONS

Luciana Nedel
Universidade Federal do Rio Grande do Sul
159 PUBLICATIONS  880 CITATIONS

Some of the authors of this publication are also working on these related projects:

iMIS^2 (individualized Minimally Invasive Surgery Simulation) View project

HCHCI – Haptic Communication for Human-computer Interaction View project

All content following this page was uploaded by Luciana Nedel on 24 July 2014.

The user has requested enhancement of the downloaded file. All in-text references underlined in blue are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.
The Point Walker Multi-label Approach

Hernandi Krammes∗ Marcio M. Silva† Theodoro Mota‡ Matheus T. Tura§ Anderson Maciel¶
Luciana Nedel∥

Instituto de Informática (INF), Universidade Federal do Rio Grande do Sul (UFRGS)

Abstract
This paper presents a 3D user interface to select and label point sets in a point cloud. A walk-in-place strategy based on a weight platform is used for navigation. Selection is made in two levels of precision. First, a pointing technique is used relying on a smartphone and built-in sensors. Then, an ellipsoidal selection volume is deformed by pinching on the smartphone touchscreen in different orientations. Labels are finally selected by pointing icons and a hierarchy of labels is automatically defined by multiple labelling. Voice is used to create new icons/labels. The paper describes the concepts in our approach and the system implementation.

Index Terms: H.5.2 User Interfaces [Input devices and strategies]: 3D Interaction—

1 Introduction
In times of big data, sensor devices such as laser scanners can provide huge amounts of geometrical information in the form of point clouds. To be useful for a number of applications in many areas, these data must be visualized, manipulated and classified by effective usable interaction techniques.

In this work we introduce the use of a 3D interface combining navigation, manipulation and multiple selection to hierarchically label point clouds of different sizes and topographies. The interface relies on a weight platform to track user steps for navigation within the environment and around objects. It also explores the sensor capability of smartphones to provide a pointing device for coarse selection and to manipulate an ellipsoidal selection volume. Upon selection, the point sets can be labelled by pointing and selecting labels on a list. New labels can be created using voice. Navigation through hierarchical label levels is possible by leaning forward or backward on the weight platform. The interface has been built to be used with an immersive display system. We tested with a conventional HMD, as depicted in Fig. 1a.

In the remaining of the paper we describe our interaction approach, the system implementation and a few results.

2 Approach Overview
Three-dimensional point clouds can vary from the surface of a small object that we explore by moving around, to large city environments that we explore by flying over or walking inside. Thus, we wanted to conceive a system which allowed great mobility for easily turning and visually inspecting all portions of the data. Our system consists of a navigation model similar to the real world walking and rotating. To avoid heavy installed infrastructure, we use a portable weight platform to track velocity and direction of a user walking on top of it.

For about a decade, smartphones evolved and are now ubiquitous. They contain enough sensors to work as very precise pointing devices. A combination of inertial sensors and magnetometer coupled with multi-finger touch gestures allows to use them for 3D selection and manipulation. In our approach, selection of multiple points is achieved by pointing with the device to control a 3D cursor. The pointing direction defines a ray and the intersection of the ray with the data defines the cursor position, similar to the Disambiguation Canvas [1]. The cursor is an ellipsoidal volume that selects all the points inside it. It can be deformed in any direction by pinching on the smartphone screen while the device is oriented in the direction wanted for the deformation. As the selection volume is deformed, additional cloud points are covered by the volume and are then flagged for selection. When the user is satisfied with the selection, they can assign a label to the set of points.

A voice recognition framework in the smartphone can be activated when a new label is to be created. As our design expects an
immersively display using voice for typing helps keeping the
immersion. Notice that nothing is displayed on the smartphone’s
screen. It is used merely as an input device. Points can receive
many labels. For example, a point in an eye is also part of a head
that is part of a body. Assignment of multiple labels gives rise to a
hierarchy of labels. With our approach, the user navigates through
label levels by zooming in and out, similarly to a map where coun-
try, state, major city and small town labels appear as we zoom in.
The user action to zoom into the hierarchy is to lean forward or
backward, a movement quickly understood by the weight platform.

Next section details the implementation of these principles.

3 IMPLEMENTATION

The system has been implemented in C++ with OpenGL and using
the following hardware: a PC, an Android based smartphone; a
Sensics HMD; 4 Wiimote balance boards.

3.1 Navigation

An array of balance boards compose the walking platform. The sys-
tem also works with a single board, but due to the user immersion,
4 boards provide a safer and more comfortable surface. The boards
communicate with the PC through a bluetooth protocol. We imple-
mented an algorithm that merges the readings of 4 balance boards
to compute the user’s body center of mass projected on the walking
surface (C). This is a generalization of Wang [3]. The center of
mass is calculated as \( \bar{C}(t) = \sum_{i=1}^{M} \bar{u}_i a_i(t) \), where \( \bar{u}_i \)
is the position of a sensor in the array of boards, \( a_i(t) \) is raw sensor i data over
time and \( M \) the total number of sensors. Fig. 2 shows the matrix
configuration tested with two boards (eight sensors).

![Figure 2: Balance board array and parametrization.](image)

The pattern of change in the center of mass \( \bar{C} \) while walking al-
 lows us to estimate pace and body direction. The center of mass is
derived over time to generate a directional impulse response when
the user steps in some direction (\( G(t) = d\bar{C}/dt \)). Such operations
occur on the plane. Direction is obtained as follows: a perpendicular
vector is first calculated from \( \bar{N}(t) = \text{perpendicular}(G(t)) \) and
then the walking direction angle over time, \( \alpha(t) \), is computed.

Moreover, to avoid drifting between head orientation obtained by
the HMD and feet orientation obtained by the boards, we applied
the comfort pose algorithm [2]. This solution does not require any
additional tracking sensor or infrastructure. The algorithm relies on
the assumption that people tend to align body and head comfortably
a few seconds after looking to either side.

3.2 Selection

A Samsung Galaxy SIII Android phone is the mobile device used
for selection. An app has been implemented that works as a server
reading sensor data and sending them through the network. The
data is grouped in a string sent through TCP over wifi. The string
is read by the PC client that parse it and use the sensor data to
construct a pointing vector.

The vector is then used to cast a ray. Each point of the point cloud
is set with a radius defining a small sphere. Intersection between the
ray and the spheres define a selection region. The closest sphere to
the user is computed and an ellipsoidal cursor is displayed at that
sphere’s position (Fig. 1b). Cursor position is updated in real time
as the user points towards the cloud. When the user finds some
points they want to select, they tap on the handheld device, causing
the cursor to stop follow. From this moment, a pinch movement
on the device’s touchscreen scales the ellipsoid one the direction of
the plane corresponding to the device’s orientation (Fig. 1c). The
user can then reorient the device and keep deforming the ellipsoid
until he/she is satisfied with the points being covered. Another tap
brings back the ray cast allowing to select one among a set of icons
corresponding to labels. This action will assign the label to the
selected set of points.

3.3 Labels and Hierarchy

As soon as a set of points is selected, the same pointing device is
used to select a label from a list of icons. The respective label is
then assigned to those points. To create a new label, the user selects
‘new’, activating the voice recognition system based on the Google
API in the smartphone. The user then speaks a name for the label,
which will be promptly converted to a string and sent together with
the sensor information in the TCP package. The name is used to
create a new label ready to be used.

As the user adds more labels to points, a hierarchical data struc-
ture is automatically created, placing smaller point sets in lower
levels of the hierarchy. For example, if points of an eye have been
previously labeled ‘EYE’, when the whole head is labeled ‘HEAD’,
the head will assume a higher level in the hierarchy.

For hierarchy visualization, the user action is to lean forward/backward. This causes a weight displacement on the weight
platform that triggers changes in the label hierarchy level being dis-
played. This is a natural interface as in the real world people lean
forward to see more details and backward to grab more context.

4 RESULTS AND FINAL COMMENTS

We performed the labeling task on a couple of models provided by
the contest organization for demonstration. Fig. 1d brings an
example.

We presented an involving and usable approach for multiple 3D
point selection an labeling. Our solution proposes a full body user
experience where hand-pointing and pinching is used to specify an
ellipsoidal selection volume. Navigation, in turn, uses an improved
walk in place technique that includes walking direction in a natu-
ral way. Hierarchical visualization of labels uses the well known
zoom-into-map approach. As future work, we are planning user
studies to confirm the usability and effectiveness of the approach.
We also plan tests with an annotation application in medicine.

We believe that the technique can be implemented with even sim-
pler hardware, e.g., the Oculus Rift, being mobile and accessible for
a huge number of applications that combine navigation selection
and manipulation.

ACKNOWLEDGEMENTS

We acknowledge the support of CAPES, CNPq-Brazil through
PIBIC, grants 311547/2011-7 and 305071/2012-2, and AES Sul.

REFERENCES

[1] H. Debarba, J. Grandi, A. Maciel, L. Nedel, and R. Boulie. Dis-
ambiguation canvas: A precise selection technique for virtual envi-
ronments. In P. Kotze, G. Marsden, G. Lindgaard, J. Wesson, and
M. Winckler, editors, Human-Computer Interaction - INTERACT 2013,
volume 8119 of Lecture Notes in Computer Science, pages 388–405.

and body orientation using a comfort pose function. In Virtual and Aug-

and elastic board interfaces for locomotion in vr. In 3D User Interfaces