Extending Interactions into Hoverspace Using Reflected Light

Dmitry Pyryeskin, Mark Hancock, Jesse Hoey
University of Waterloo
Waterloo, Ontario, Canada
{dpyryesk, mark.hancock, jhoey}@uwaterloo.ca

ABSTRACT
Multi-touch tables are becoming increasingly popular and much research is dedicated to developing suitable interaction paradigms. There also exist multiple techniques aimed at extending interactions into the hoverspace—the space directly above a multi-touch table. We propose a novel hoverspace method that does not require any additional hardware or modification of existing vision-based multi-touch tables. Our prototype system was developed on a Diffused Surface Illumination (DSI) vision-based multi-touch set up, and uses light reflected from a person's palm to estimate its position in 3D space above the table.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

Keywords: Multimodal interaction, natural human computer interaction, surface computing, multi-touch, hoverspace.

INTRODUCTION
Most surface computing systems allow people to manipulate objects on the screen directly with their hands and fingers, which opens the possibility of interaction that is familiar from the physical world. Recent technology, such as the Kinect, has introduced the possibility of extending this physical interaction above or in front of the multi-touch display. The addition of hoverspace to touch input can provide another mode of interaction, while allowing seamless and effortless transition from one mode to another. Hoverspace gestures can be used as shortcuts to applications similarly to Hover Widgets [3]. Hover information makes it easy to assign each touch region from separate fingers to a hand, which can be used for user differentiation, table territoriality and security. Hoverspace images can also be used to create occlusion-aware interfaces [7].

RELATED WORK
The subject of hoverspace interactions has already been researched for some time and a number of techniques were proposed. Most of these methods require custom-built hardware or modification of existing multi-touch tables. Some systems rely on stereo cameras to calculate the 3D location of a person's hands and fingers [5]. Such systems require much more computational power than other vision-based systems and also require very precise camera calibration. Other systems use depth cameras [1]. There are also methods that use multiple layers of lasers to estimate the position of a person's hands and fingers [6], and methods that position infrared emitters above the table to create and track shadows of people's hands together with touch regions [2].

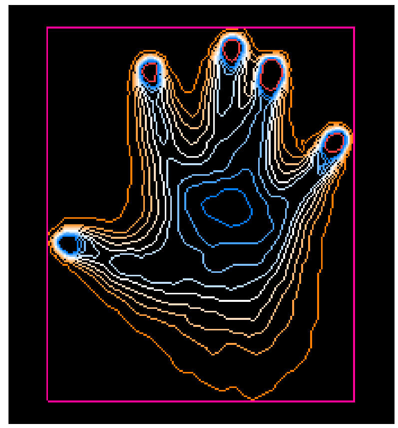

Figure 1: An image is split into 10 layers based on the brightness and blobs are found in each layer.

PROPOSED METHOD
Hardware
The proposed method uses a standard unmodified DSI vision-based multi-touch table setup. The setup utilizes EndLighten acrylic that scatters infrared (IR) light evenly throughout the table's surface. A piece of acrylic of size 81 cm $\times$ 61 cm forms the surface of the table. It is edge-illuminated by 850 nm IR diodes. A rear-projection film, applied to the acrylic surface, allows it to act as a screen for rear-mounted short-throw projector. A Unibrain Fire-i^{TM}camera equipped with an 850 nm band-pass filter is mounted behind the screen. The table is powered by a Windows XP computer. When an object, such as person's hand, approaches the table's surface, it reflects scattered IR light back through it to be detected by the camera. The amount of reflected light is inversely proportional to the distance of the object from the surface.

Software
The software that was developed for the proposed system is based on similar principles as some of the other vision-based multi-touch trackers, i.e. Community Core Vision [1] and reac-

[1]http://ccv.nuigroup.com/

Conference ITS'11, November 13-16, Kobe, Japan.

TIVision [2]. Unlike other trackers, it splits each frame into a number of layers, based on their brightness (as shown in Figure 1). The layers are extracted by applying threshold filter to the original image. Ten thresholds were used to create 10 binary images, each of which is analyzed using a blob-finding algorithm. Brighter blobs are caused by fingers touching the surface, while dimmer blobs represent hands hovering above the surface. The software then extracts the background, eliminates the noise using smoothing filter and applies a high-pass filter to detect touch regions. Finally, it analyzes sizes and positions of blobs in various layers to assign each finger to a hand. Each blob is tracked across multiple frames and its position, speed and acceleration in 3D space are calculated. These parameters may be used to develop 3D gestures. The tracker then sends touch and hoverspace data to multi-touch client applications. The ability to switch between 2D and 3D profiles defined in TUIO [4] protocol allows the application to be compatible with existing client applications.

PRELIMINARY EVALUATION

A preliminary study was conducted in which we collected data from three participants. Each participant was asked to keep their palm steady at certain distances above the surface for 2–4 seconds (or 60–120 frames at 30 fps). A laser level was set up above the surface to assist participants in staying within a given height. A video was recorded using the table's camera for each participant/height combination. These videos were analyzed with the experimental software that found blobs of varied brightness in each frame. Brightness level data extracted from the videos was then matched to the actual height of the participants' palms and a correlation was calculated. Linear regression applied to the resulting data showed that there is a clear correlation between the actual height of a palm and perceived brightness levels. Moreover, the analysis shows that a simple linear model fits well ($R^2 > 0.95$ - see Figure 2). The experiments also show that the amount of reflected light varies by individual, which means a calibration step may be required. Another experiment focused on the system's ability to predict the actual height of participant's palm, based on the collected data and the perceived brightness. A leave-one-out cross-validation procedure was applied to each participant's data. A linear fit to all but one data point was used to predict the surface distance of the left-out point from its brightness data. Accuracy within different errors (5, 10, 15 and 20 mm) were averaged by applying the same procedure to each data point. The system was able to predict 78.6%–81.8% of data points with the precision of 10 mm (see Figure 3).

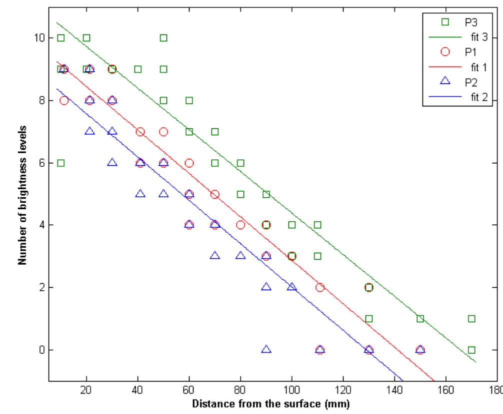

Figure 2: Linear regressions for 3 participants, with $N = 1508, 1178, 1654$ for participants $1, 2, 3$, resp.

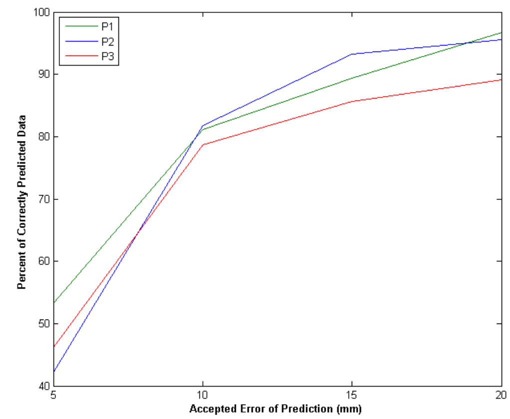

Figure 3: Height prediction accuracy for the 3 subjects

CONCLUSION

The proposed method has several benefits that set it apart from related work. The most important one is the compatibility with existing hardware with no additional modifications required. A hoverspace tracker can be installed on any DSI table (and theoretically on any Rear Diffused Illumination table). Another benefit of the proposed system is its ability to define a large number of hoverspace layers (limited by the table's processing power) to track position of a hand in 3D space as opposed to other techniques which have only one [2] or three [6] hoverspace layers.

In future, our system prototype will be improved to increase accuracy in calculating the height of a person's palm. We will then conduct a study to evaluate the effectiveness and precision of the proposed method. Several client applications utilizing the hoverspace tracking will be developed.

[2] http://reactivision.sourceforge.net/

REFERENCES

1. H. Benko and A.D. Wilson. Depthtouch : Using depth-sensing camera to enable freehand interactions on and above the interactive surface. Technical report, Technical Report MSR-TR-2009-23, Microsoft Research, 2009.
2. F. Echtler, M. Huber, and G. Klinker. Shadow tracking on multi-touch tables. In *Pro. of AVI '08*, pages 388–391, New York, NY, USA, 2008.
3. T. Grossman, K. Hinckley, P. Baudisch, M. Agrawala, and R. Balakrishnan. Hover widgets: using the tracking state to extend the capabilities of pen-operated devices. In *Proc. of CHI '06*, pages 861–870, New York, NY, USA, 2006.
4. M. Kaltenbrunner, T. Bovermann, R. Bencina, and E. Costanza. Tuio: A protocol for table-top tangible user interfaces. In *Proc. of the The 6th Intl Workshop on Gesture in Human-Computer Interaction and Simulation*, Vannes, France, 2005.
5. B. Leibe, T. Starner, W. Ribarsky, Z. Wartell, D. Krum, B. Singletary, and L. Hodges. The perceptive workbench: Toward spontaneous and natural interaction in semi-immersive virtual environments. In *Proc. of IEEE VR '00*, pages 13–, Washington, DC, USA, 2000.
6. Y. Takeoka, T. Miyaki, and J. Rekimoto. Z-touch: an infrastructure for 3d gesture interaction in the proximity of tabletop surfaces. In *Proc. of ITS '10*, pages 91–94, New York, NY, USA, 2010.
7. D. Vogel and R. Balakrishnan. Occlusion-aware interfaces. In *Proc. of CHI '10*, pages 263–272, New York, NY, USA, 2010.