Ortholumen: Using Light for Direct Tabletop Input

Tommaso Piazza[‡,*]
[‡]Institut National Polytechnique de Grenoble
ENSIMAG
F-38402 St Martin d'Heres, France
E-mail: tommaso.piazza@gmail.com

Morten Fjeld[*]
[*]TableTop Interaction Lab (t2i Lab), www.t2i.se
CSE, Chalmers University of Technology
SE-41296 Gothenburg, Sweden
E-mail: morten@fjeld.ch

Abstract

Ortholumen is a light pen based tabletop interaction system that can employ all the pen's spatial degrees of freedom (DOF). The pen's light is projected from above onto a horizontal translucent screen and tracked by a webcam sitting underneath, facing upwards; system output is projected back onto the same screen. The elliptic light spot cast by the pen informs the system of pen position, orientation, and direction. While this adds up to six DOFs, we have used up to four at a time. In order to better separate input and output light we employ polarizing filters on the webcam and on the projector lens. Two applications, painting and map navigation, are presented. Ortholumen can be expanded to track multiple pens of the same or different colors. This would enable bi-manual input, collaboration, and placed pens as external memory. Visible light, as opposed to infrared or radio, may be perceived more directly by users. Ortholumen employs only low-cost parts, making the system affordable to home users.

1. Introduction

Most tabletop interaction relies exclusively on horizontal input. Recently, some efforts have been made to incorporate the third dimension above [1] and below [2] the table. While our work is an attempt to further explore interaction possibilities above the table surface, some of the solutions suggested here may also work for input underneath the table. Ortholumen[1], the system presented in this tech paper, is based on using visible light source(s) that are manipulated in the space above the table. Since light sources are defined by their position, orientation, and direction in space, they may give additional degrees of freedom (DOFs) to tabletop input. Our approach differs from laser pointer interaction [3]. As opposed to the light we use, laser is coherent collimated light.

To assure coinciding input and output spaces, any input device manipulated apart from the tabletop display surface should provide immediately perceivable meaningful mappings of its input cues, preferably in the output space. Such mappings can more easily be created employing visible light than alternative invisible media such as radio, ultrasound, or infrared. Using these alternatives may also be cumbersome due to their complex set-up. An added benefit of visible light is that it is perceived equally well by the human eye as by camera-based tracking systems.

By using visible light instead of invisible alternatives, Ortholumen complies with three of the five *direct manipulation* criteria [4] (two criteria, reversibility of actions and syntactic correctness, are less related to input technique):

- **visibility of the objects of interest**
  light does not in any way impede the visibility of the object of interest, except when concentrated
- **incremental action at the interface with rapid feedback for all actions**
  shape, size, and position of the elliptic light spot can be changed in a continuous way with just a small effort by the user and with immediate feedback
- **replacement of complex commands languages with actions to manipulate directly the visible object**
  all characteristics of the elliptic light spot can be used as parts of an interaction schemata

In the case of a handheld light source, there is a connection between perceiving projected light and sensing cues of the source. Capitalizing on this connection, Ortholumen allows for meaningful ways to combine intuitive handling of the tangible source with immediately perceived feedback to the user. Our work is focused on exploring potential

[1]This work was carried out at the t2i Lab, CSE, Chalmers. Project web site, including video: http://www.t2i.se/projects.php?project=ol

benefits of using distance and angle between the input device and table top. Since the most common handheld input device in a tabletop context is the pen, we chose a light source of similar size and form. Instead of tracking the yaw, pitch, roll, and position of the light emitting pen, the system relies on the elliptic light spot projected onto the table surface. Such a solution provides for marker-less tracking of untethered input devices with low power consumption.

2. Related Work

Horizontal interactive systems have garnered increasing attention from researchers and system designers. Device based, direct touch, and gesture [5][6] input possibilities are among those that have been widely explored as alternatives for interaction with objects at both close and far distances [7]. Local and remote pointing using laser or invisible media, such as radio, have also been explored [3][8]. Still, the physical space for interaction in most tabletop systems is restricted to the table surface; exceptionally, the spaces above [1] or below it [2] are employed. Only a few projects have employed the space between the user's body and the table surface [1][5]. In the first of these projects, two-handed gestures were combined with a vertical screen as an output device [5]. In the second project, offering multi-layer interaction on a horizontal screen, the space was divided into layers that the user can select and interact with by controlling distance between input device and table [1]. Inspired by these projects, we adopted a continuous approach to the space above the table. This is achieved by drawing on polarization techniques for image segmentation [9].

3. System Requirements and Set-Up

Ortholumen is designed to give the user immediate visual feedback. It integrates interaction and visualization spaces defined by the user's reach. Moreover, to avoid user obstruction of the system output, back projection was adopted. To enable better tracking of the light source we chose a webcam placed below the table top. Such camera placement combined with back-projection minimizes glare and bright disturbances caused by the projector. To further enhance tracking, we employed linear polarizing filter(s); one in front of the projector, in the case of a non-LCD projector, and one in front of the webcam lens.

The fixed parts of the set-up consist of a table, holding a translucent screen, an LCD projector, a mirror, and a low-cost web camera (Fig. 1). In this installation, one or multiple LED pens are held by hand in the space above the table or in direct contact with the surface. The light

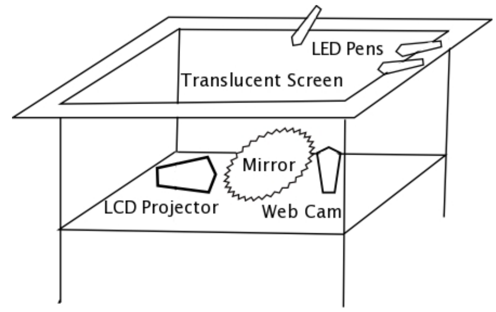

Figure 1. The set-up including table with translucent screen, LCD projector, mirror, web cam, and LED pens.

of the LCD projector is reflected by the mirror onto the translucent screen. The webcam, mounted next to the mirror, captures the screen image combining pen input and projector output. In front of the camera lens there is a polarizing filter. The filter is aligned orthogonally to the polarization direction of the projector's light. All components are of consumer quality, making the system affordable to most users:

- **Projector**
  The video projector used is of the LCD type, a *Sony VPL-CS5* 1800 lumen.
- **Camera**
  The web camera used is of the CCD type, an *Apple iSight* with 640x480 pixel resolution.
- **Polarizing Filters**
  A neutral gray linear polarizing film with 99% polarizing efficiency, 38% single, 30.1% parallel, and 0.0045% transmittance efficiency was chosen because of its superior transmissive characteristics between 400 and 700 nm.
- **Light Emitting Pen**
  The LED pen used has tip-mounted color LEDs which can be activated one at a time or in combination by pressing a side-mounted button (Fig. 2). The LED light employed is incoherent, not collimated, and mostly monochromatic which benefits tracking. Solid state light sources require no external reflector to collect light because the solid package can be designed to concentrate light. This, in addition to its solid-state nature, provides for resistance to shock and vibration.

4. Tracking

One of the first problems approached was how to track a pen's light as it would mix with the projector and

Figure 2. Side and front view of the LED pen.

environmental light. Since the light coming from the pen is colored and blends with projected images from the system, plain color histogram tracking is generally not precise enough to allow our system to work well.

The goal is to track the position, orientation, and size of the elliptic light spot cast by the pen. In a scenario where the pen's light is mostly diffuse and often of the same color as the background, tracking according only to color is unsatisfactory. For this reason we adopted the HSV color space, which enables the system to distinguish between color and brightness. A two-dimensional hue-value histogram of the region of the image known to be the pen is computed and then back-projected to the image's two planes of interest (H and V). The CamShift algorithm [10] as implemented in the OpenCV[2] library is then applied.

The primary assumption made is that the pen's light spot would appear as one of the brightest sources in the image. This assumption should also hold while lifting or lowering the pen and for small pen surface distances. Still, once the light spot becomes too weak, due to pen distance, tracking terminates. Interference by other sources of light present in the environment is prevented by the hue-value tracking combination since normal light is of neutral color and the tracked light has specific hue values. As the system has no learning capacities start-up involves manual selection of the light spot to be tracked.

5. Polarizing Filters

The role of the filter(s) is to enhance tracking sensitivity and precision. Following the principles of polarization the idea is to prevent as much light as possible from interfering with the tracking system. Interference is extra bright zone(s) in the image or glare(s), making the camera unable to discriminate the pen's light spot. For these reasons, one polarizing filter is laid in front of the projector lens, in the case of a non LCD projector, and a second one orthogonal to the first one in front of the camera lens. Since the projector lens used to focalize the light could alter the polarization, the filter sits after the lens. For the camera, there is no need to preserve polarization after the lens.

[2] http://www.intel.com/technology/computing/opencv/

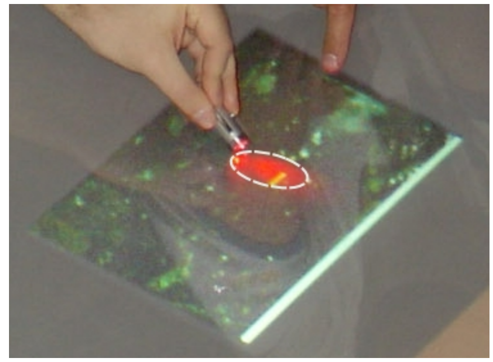

Figure 3. A view of the map application.

6. Applications

As a proof-of-concept, we built two single-user applications employing one up to four of the pen's six DOFs. Firstly, we designed a *painting application* where the pen plays the role of a brush. The light spot's x and y coordinates determine the brush coordinates. The pen's distance from the table top, inferred from the characteristics of the elliptic light spot, determines the size of the brush. Hence, it is possible to interpret the pen's distance from the table top either in a visually-coherent (VC) mode or in a position-coherent (PC) mode. In VC mode, lifting the pen increases the size of the brush just as the user-perceived size of the light gets bigger; lowering the pen reduces the size of the brush. In PC mode distance affects size in the opposite way. Secondly, we built a *map navigation application* by connecting Ortholumen and Google Earth[3] (Fig. 3). Here the pen acts much like a joystick: distance and angle to the center of the display jointly activate and control horizontal map movement; elevation determines the zoom factor. Tilt around the display's x axis controls tilt in the map. Much like in the painting application, leaving or approaching the table surface affects zoom according to one of two alternative modes. In PC mode leaving the surface results in zooming out the map while approaching results in zooming in. In VC mode the pen works like a magnifying glass. When lifting the pen the system reacts as if a magnifying glass was lifted from the surface, zooming in the map.

7. Discussion and Future work

Ortholumen derives much of its value from a three-dimensional interaction space and the use of light as interactive media. While this is less complex than other tabletop interaction technologies, our approach also has a number of limitations.

[3] http://earth.google.com/

Since the approach relies on visual tracking and interpretation of the light's color and brightness in order to determine changes in position, size, and shape, the detection of the correct light source representing the pen depends on the quality of the image's segmentation. While segmentation as such is difficult, it becomes even more a challenge in uncontrolled viewing circumstances such as mobile scenarios (e.g. map navigation). While our prototype facilitates segmentation by using polarization, back projection and the use of LCD projector technology limit the power of the technique. Very bright zones are likely to fool the system.

The overall quality of the estimation of the pen's distance from the table greatly depends on the diffusion of the light reaching the table top. At a certain height, typically above 10 cm, light beamed to the surface becomes too diffuse to be tracked by a webcam-based system. Small angles, however, are less of an issue as long as the pen projects onto the surface. The use of a more concentrated light source could partially help solving these issues in future prototypes.

Tilting the pen to change the shape of the projection radically influences the size in most cases. This becomes an issue if there is a need to use both size and other characteristics of the shape simultaneously, because it relies on the user's ability to keep one parameter constant while varying the other. To avoid such problems in future prototypes decoupling of size and tilt may be reached by introducing color coding. Finally, tracking the pen's rotation around its main axis remains a challenge in realizing full six DOFs input.

In future work, multiple pen realization of Ortholumen may enable bi-manual input, collaboration, and pens placed upright as external memory. Connecting multiple DOFs with a set of active and placed pens will require a systematic investigation. Infrared light could be added to the system by integrating two infrared LEDs into the pen. This would make computation of the pen's rotation around its main axis easier.

8. Conclusion

We have realized a tabletop interaction technique based on visible light which extends interaction to the space above the table surface. This gives meaning to user input actions such as lifting or lowering a hand held device, as well as changing its yaw, pitch, and roll. Unlike other techniques, our technique does not involve specialized hardware and gives the user immediate visual feedback. Two applications using up to four DOFs have been realized. We would like to further explore the Ortholumen technique in multiuser environments and other contexts where multiple tangible input devices can be of use.

9. Acknowledgments

Andreas Kunz and Martin Kuechler gave good advice on the use of polarization for image segmentation. Erik Tobin, Fredrik Gustafsson, Sven Berg Ryen, and Wayne Brailsford helped in proofreading.

References

[1] Subramanian, S., Aliakseyeu, D., and Lucero, A. “Multi-layer interaction for digital tables”. In Proc. UIST'06, 2006, pp. 269-272.

[2] Wigdor, D., Leigh, D., Forlines, C., Shipman, S., Barnwell, J., Balakrishnan, R., and Shen, C. “Under the table interaction”. In Proc. UIST'06, 2006, pp. 259-268.

[3] Olsen, D. R. and Nielsen, T. “Laser pointer interaction”. In Proc. CHI'01, 2001, pp. 17-22.

[4] Shneiderman, Ben. “Direct manipulation: a step beyond programming languages”. In IEEE Computer 16(8), 1983, pp. 57-69.

[5] Wilson, A. D. “Robust computer vision-based detection of pinching for one and two-handed gesture input”. In Proc. UIST'06, 2006, pp. 255-258.

[6] Wu, M. and Balakrishnan, R. “Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays”. In Proc. UIST'03, 2003, pp. 193-202.

[7] Vogel, D. and Balakrishnan, R. “Distant freehand pointing and clicking on very large, high resolution displays”. In Proc. UIST'05, 2005, pp. 33-42.

[8] Parker, J. K., Mandryk, R. L., and Inkpen, K. M. “TractorBeam: seamless integration of local and remote pointing for tabletop displays”. In Proc. Graphics interface 2005, 2005, pp. 33-40.

[9] Kuechler, M., Kunz, A. “Imperceptible projection blanking for reliable segmentation within mixed reality applications”. In Proc. 9th IPT/11th EGVE, 2005, pp. 23-30.

[10] Bradski, G. R.“Computer vision face tracking tor use in a perceptual user interface”. In Intel Technology Journal Q2'98, 1998.