

## CLUTCHING AND LAYER-SWITCHING: INTERACTION TECHNIQUES FOR PROJECTION-PHONE

S. SEO, B. SHIZUKI AND J. TANAKA

*Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba, Japan*

*Email address: {seo, shizuki, jiro}@iplab.cs.tsukuba.ac.jp*

**Abstract.** We made a prototype cellular phone that incorporates a projection function. This prototype of a projection-phone is a system based on a commercially available portable projector to simulate the *projection-phone* environment. This system uses a peek technique by using image-based motion estimation to display virtual information. We developed clutching and layer-switching as a way of enabling the user to interact with and manipulate the virtual information projected using the peek technique. We also present some scenarios in which the projection-phone can be used.

### 1. Introduction

A cellular phone is one of the most familiar portable devices. These devices now have many functions, such as connecting to the Internet and providing entertainment. The problem is that the size of a cellular phone is limited because it has to be handheld. This limitation also means the size of a cellular phone's built-in screen is small. This is a particular problem when the amount of information that has to be handled is large.

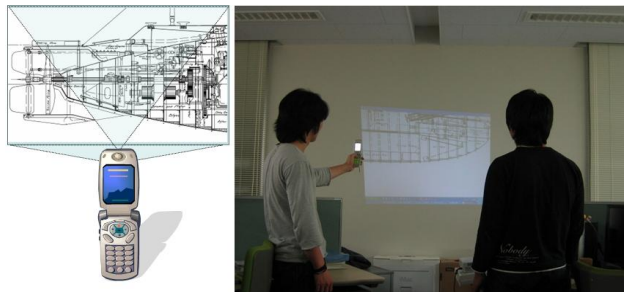


Figure 1. Projection-phone (Left) and system in use (Right).

We believe that recent developments in projection technology provide a solution to this problem. Recent projection technology enables projectors to be small enough to be incorporated in a portable device, like a cellular phone or a PDA. Therefore, a cellular phone could project information onto another surface and therefore the size of its built-in screen would not limit the amount of information it displays. In this paper, we call a cellular phone that has a projection function a *projection-phone* (Figure 1). This paper contains some interaction techniques for using this device.

## 2. Technique for Displaying Virtual Information.

We call this technique a *peek technique*. The peek technique creates the illusion for the user that he or she is *peeking* at only part of a complete projected image. To do this, a virtual image is created of a complete image by the projection-phone, but the projector actually only projects the part of that image at which the user points the projection-phone. In figures 2(a) and 2(b) the virtual image is comprised of the letters A to L. In figure 2(a) the projector is pointing to the part of the image containing the letters A to D, so these letters are all that is projected. In figure 2(b), the projector is pointing to the part of the image containing the letters H to L, so these letters are all that is projected.

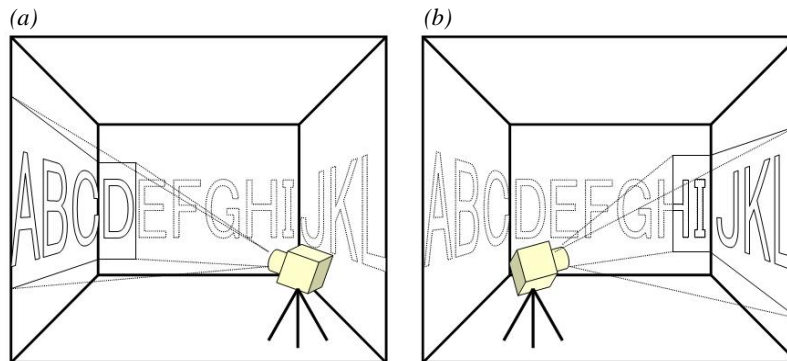


Figure 2. Peek Technique.

## 3. Interaction Techniques

We propose the following two interaction techniques to manipulate the virtual information displayed when the peek-technique is used.

- Clutching
- Layer-switching

### 3.1. CLUTCHING

Clutching is the operation used to look at any virtual information that falls outside the possible project area (Figure 3). The user presses the *clutch-button* on the projection-phone and can then scroll through the virtual information. The amount of scrolling is based on the phone's motion while the clutch button is being pressed. With this operation, the user can drag virtual information that is outside the possible projection area into the possible projection area (Figure 3) and view it.

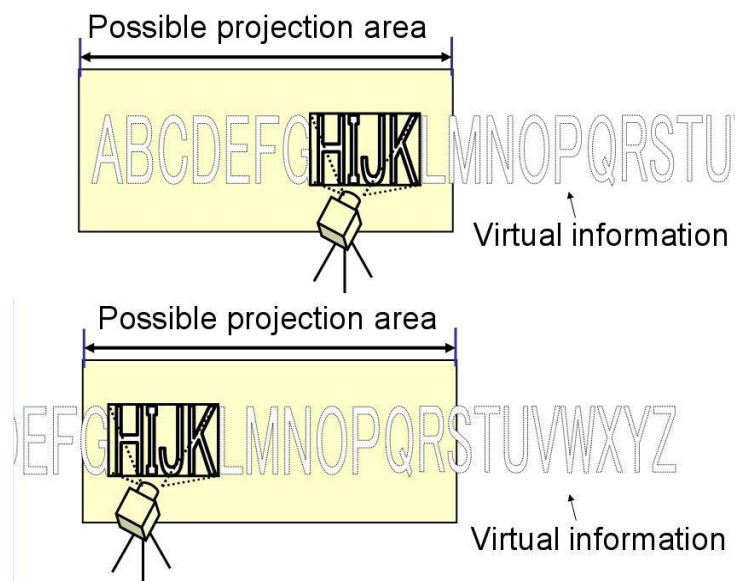


Figure 3. Clutching enables scrolling through virtual information.

### 3.2. LAYER-SWITCHING

We define the information-layer as a virtual plane that includes the virtual information projected using the peek-technique. The prototype system has two kinds of information-layer: one is an inspection layer and the other is a selection layer (Figure 4). The inspection layer is used for browsing the virtual information. The selection layer is used for selecting the virtual information indicated in the inspection layer. Layer-switching is an operation that switches between these two information-layers. This user presses the *layer-button* on the projection-phone to switch the indicated layer from the inspection layer to the selection layer and vice versa. In the selection layer, the virtual information is selected by moving the cursor over the center of the projected image.

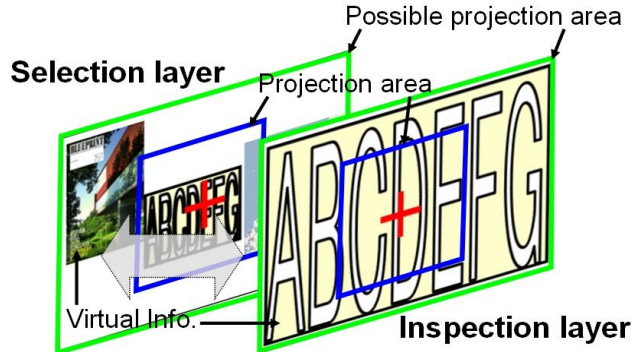


Figure 4. Prototype system has two kinds of information-layer: inspection layer (right) and selection layer(left).

#### 4. Usage Scenario

Virtual information displayed using the peek technique can be information as complex as a blueprint or a map. In addition, the virtual images can be high-resolution images (such as drawings, pictures and text) that contain more than 4000X3000 pixels. For instance, an aerial photograph or an A0 size poster is the kind of virtual information that can be displayed using the peek technique.

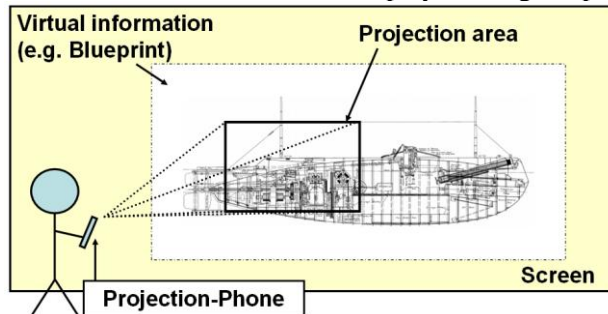


Figure 5. Indoor scenario: browsing large blueprints.

In indoor scenario, the projection-phone is used to project a blueprint, a map, or a similar high-resolution image onto a screen or wall. For instance, after putting the wiring diagram into the projection-phone for each room of an apartment under construction, an electrician can check the actual wiring of the apartment during construction by projecting the wiring diagram onto the wall of each room as he or she makes an inspection. In another example, a ship designer in conference with a client can project the blueprint of ship

similar to one in which the client has interest. This will help the client and designer discuss the client's needs (Figure 5).

In outdoor scenario, the projection-phone can be used to project a map onto the floor or wall of a nearby building of a place where the user is actually standing.

## 5. Implementation

We made a prototype system based on a pocket projector (Figure 6). To simulate the cellular phone's projection function, we used a Toshiba TDP-FF1 Pocket Projector. A USB camera (320X240 pixels, 24bits/pixel, 30fps) was used in place of a cellular phone's built-in camera. We used a Bluetooth-enabled Nokia Cookie sensor as a 3-axis accelerometer. These three components were connected to a laptop PC, which produced the projection image and handled the interaction. The clutch-button and the layer-button were keys on the keyboard of the PC.

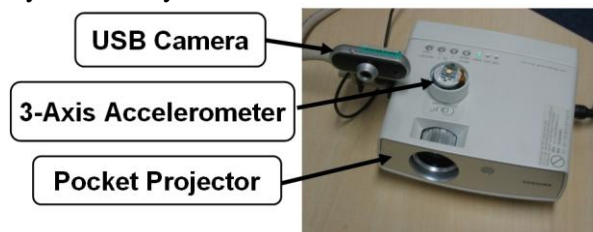


Figure 6. Projection-phone prototype.

### 5.1. IMAGE-BASED MOTION ESTIMATION

Information from the USB camera was used to determine any changes in orientation and position of the pocket projector.

Any discrete features that were detected in the images were sifted to find the corresponding matches between each sequential image pair from the USB camera. These features were used to determine the motion estimate of the pocket projector between the image pair. Although there are a number of feature-detectors available that can be used to extract these features, we used the renowned Kanade-Lucas-Tomasi (KLT) algorithm (Shi and Tomasi, 1994) because it can be used on any background image from a USB camera.

The KLT features were used as the starting locations for window-based intensity correlation matching (Wang et. al., 2006). Over this correlation matching, we reduced the search space by restricting the range of each search to those features in the current image that lay within  $R$  pixels of the given feature in the previous image. In our implementation,  $R$  was equal to 3.

The next stage of correlation matching was to take a random number of those matched pairs, calculate the Fundamental Matrix between them, and see whether that matrix was appropriate in consideration of all the pairs. We solved this stage by following the well-known RANSAC procedure, introduced by Fischler and Bolles (1981). Seven feature correspondences were randomly selected from the set of all the feature tracks. Then, we computed the fundamental matrix  $F$  for the image pair, using Hartley and Zisserman's 7-point Algorithm (2000). When the procedure was complete, the best estimate of  $F$  was used to compute the robust  $F$  by removing all features that were considered outliers.

This fundamental matrix  $F$  and the camera's pre-computed calibration matrix  $K$  were used to calculate the essential matrix  $E$ , which is given by

$$E = K^T F K. \quad (1)$$

The next stage of the motion calculation was to extract the translation and rotation parameters from  $E$ . The essential matrix  $E$  can be decomposed to give the rotation matrix  $R$  and the translation vector  $T$ , which is given by (2)

$$E = [t]_x R. \quad (2)$$

This correlation equality can be expressed in matrix form as

$$x'^T R S x = x'^T E x = 0, \quad (3)$$

where  $S$  denotes the  $(3 \times 3)$  skew symmetric matrix for which  $Sx = t \times x$ . The rotation matrix  $R$  and  $S$  matrix can be determined from the singular value decomposition  $E = U \text{diag}(1, 1, 0) V^T$  as follows from Hartley and Zisserman's work (2000).

$$R = U W V^T \text{ or } R = U W^T V^T, \quad S = U Z U^T, \quad (4)$$

where

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (5)$$

## 5.2. KEYSTONE CORRECTION

We used a 3-axis accelerometer as a tilt sensor to measure the first orientation of the projector and, by also using the previous motion estimation step, to obtain the absolute rotation of the pocket projector. The first orientation of

the projector is used to initialize the starting point on the virtual information. We used the absolute rotation in order to correct any errors, which are often present in relative movement information.

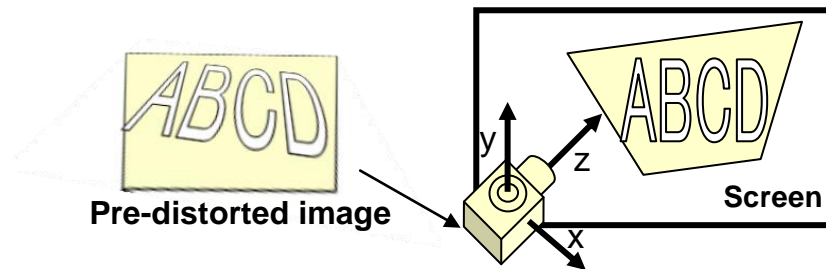


Figure 7. Keystone correction.

Using this sensor's information and the image-based motion information, the system updates the position of the portion of the display on the virtual information. Because the plane of projection and the projector are not always maintained in a horizontal position for the motion of the projector, the system also estimates the form of the plane of projection using the computed motion information. When the estimation of the form of the plane of projection is complete, the system pre-distorts the projected image to compensate for the distortion caused by the inclination of the projector (Figure 7).

## 6. Related Works

Forlines et al. (2005), Beardsley et al. (2005), and Cao and Balakrishnan (2006) have all developed systems that use handheld projectors. In some of these systems, it is necessary to set up an expensive tracking system beforehand (Cao and Balakrishnan, 2006) and to paste the marker on the projection plane side (Forlines et al., 2005; Beardsley et al., 2005) in order to manipulate the virtual information. However, because our system does not need an expensive tracking system to be set up beforehand, it is possible to use our system in a wide-ranging environment. The Hotaru system (Sugimoto et al., 2005) simulates projection from a portable device and the interaction is facilitated using the tip of a finger.

## 7. Discussion

Movement and relative posture in three dimensions of the portable projector can be measured if there are 2 sequence images taken by the USB camera. Furthermore, similar information can be obtained using the 3-axis

acceleration sensor and a magnetic sensor. The problem is that a magnetic sensor is easily affected by the environment and the 3-axis acceleration sensor does not easily handle measurements other than those in the direction of gravity. Therefore our approach uses a USB camera and 3-axis acceleration sensor, although using the image from a USB camera requires a complex calculation in order to estimate the movement and the posture of the projector.

## 8. Conclusion

We made a prototype projection-phone system based on a commercially available portable projector. We implemented the peek technique by using image-based motion estimation. We proposed using clutching and the layer-switching as a way of enabling the user to interact with and manipulate the virtual information that was projected using the peek technique. We presented some scenarios in which the projection-phone can be used.

## References

- Beardsley, P., Baar, J. V., Raskar, R. and Forlines, C.: Interaction using a handheld projector, *IEEE Computer Graphics and Applications*, **25**(1), pp. 39-43.
- Cao, X. and Balakrishnan, R.: 2006, Interacting with dynamically defined information spaces using a handheld projector and a pen, In *UIST2006*, ACM Symposium on User Interface Software and Technology, pp. 225-234.
- Cao, X., Forlines, C. and Balakrishnan, R.: 2007, Multi-User Interaction using Handheld Projectors, In *UIST2007*, ACM Symposium on User Interface Software and Technology, pp 43-52.
- Fischler, M. A. and Bolles, R. C.: 1981, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Comm. of the ACM*, **24**, pp. 381-395.
- Forlines, C., Balakrishna, R., Beardsley, P., Baar, J. V. and Raskar, R.: 2005, Zoom-and-pick: facilitating visual zooming and precision pointing with interactive handheld projectors, In *UIST2005*, ACM Symposium on User Interface Software and Technology, pp. 73-82.
- Hartley, R. I. and Zisserman, A.: 2000, *Multiple View Geometry in Computer Vision*, Cambridge University Press, ISBN: 0521623049.
- Shi, J. and Tomasi, C.: 1994, Good Features to Track, In *CVPR94*, IEEE Conference on Computer Vision and Pattern Recognition, pp. 593-600.
- Sugimoto, M., Miyahara, K., Inoue, H. and Tsunesada, Y.: 2005, Hotaru: Intuitive Manipulation Techniques for Projected Displays of Mobile Devices, *INTERACT 2005*, pp 57-68.
- Wang, J., Zhai, S. and Canny, J.: 2006, Camera Phone Based Motion Sensing: Interaction Techniques, Applications and Performance Study, In *UIST2006*, ACM Symposium on User Interface Software and Technology, pp. 101-110.