Image Browsing in Mobile Device Using User Motion Tracking

Sunghoon Yim, Jane Hwang, Seungmoon Choi, and Gerard Jounghyun Kim

***Abstract*. Recent mobile devices often store massive images. However, typical interfaces such as 2D display with a limited size and discrete button inputs make the browsing and manipulation of such massive images difficult and time-consuming. Expecting that a continuous input interface and a suitable 3D visualization may mitigate this difficulty, we designed and implemented an image browser that features with user motion tracking and 3D layouts, and carried out a usability study. The results showed that although intuitive and fun, the motion-tracking platform still has much room for further improvements.**

***Index Terms*—Image browsing, Mobile device, Motion interface**

I. Introduction

Mobile devices such as digital cameras, cellular phones, PDAs (personal digital assistants) and UMPCs (ultra mobile PCs) often store an enormous number of images. This makes browsing and manipulating images, such as seeking for a particular image, very difficult and time-consuming [1]. This difficulty is mostly due to the facts that the current mobile devices provide discrete user interfaces only (e.g., buttons) and that their limited screen size restricts the amount of image information that can be displayed simultaneously. The image handing difficulty is usually commensurate with the amount of the image data. Considering the rapidly increasing image data in the mobile device, there exists an immediate need for improved interfaces.

Toward efficient image browsing in the mobile device, we have applied motion-based interaction that can provide a continuous and three dimensional (3D) interface. It was shown that the motion-based interaction with mobile devices can be intuitive and natural in our previous work [2]. To maximize the effectiveness, we designed two image viewing methods and three motion mapping modes. A viewing method determines how images are aligned in the 2D display, and a motion mapping mode relates a user's motion to a predefined image handing command. The performance of the motion-based interaction was also evaluated through a usability experiment.

Sunghoon Yim is with the Haptics and Virtual Reality Laboratory in the Department of Computer Science and Engineering at POSTECH, South Korea (e-mail: algorab@postech.ac.kr).

Jane Hwang is with the Department of Computer Science and Engineering, Korea University, South Korea (e-mail: jane@postech.ac.kr).

Seungmoon Choi is with the Haptics and Virtual Reality Laboratory in the Department of Computer Science and Engineering, POSTECH, South Korea (e-mail: choism@postech.ac.kr)

Gerard Jounghyun Kim is with the Department of Computer Science and Engineering, Korea University, South Korea (e-mail: gjkim@korea.ac.kr).

II. Design and Implementation

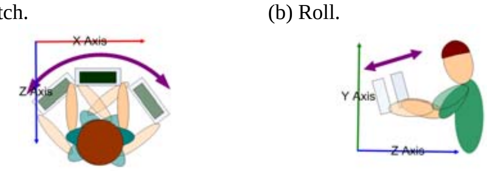
We designed two types of image views (tiled and fish-eye views) using a cylindrical layout. In the cylindrical layout (see Fig. 1(a)), images are placed in the 3D motion space beneath a virtual cylinder with the user located at its center. A viewing method determines which images should be shown in the mobile device screen from the layout and how they are displayed. In the tiled view, the images were spaced uniformly in the grids (Fig. 1(b)). The fish-eye view gives more emphasis on images in the center of the display (Fig. 1(c)). The fish-eye view was shown to be effective at browsing in the mobile device [3].

(a) Cylindrical layout. (b) Tiled view. (c) Fisheye view.

Fig. 1. Image layout and views used in the present study.

In our previous work [2, 4] , we implemented a hybrid tracking method that uses a 3D accelerometer and a camera to track the four degrees of freedom (DoF) movement of the mobile device (Fig. 2). The tracking method provides three tilt angles and forward/backward movement position. With this motion information, we can map the mobile device motions to manipulation commands for the image in the 3D layout space.

(a) Pitch. (b) Roll.



(c) Yaw. (d) Forward / backward.

Fig. 2. Tracking the four DoF movement of the mobile device.

The user motion was mapped to browsing actions in the image layout space with three different modes. The first one is direct mapping, where the position and orientation of the virtual camera that sees images were proportional to those of the mobile device. This mode allows accurate cursor control and can be quite effective for browsing images in a small region. However, it can be inefficient for long distance navigation. For example, the user might be required to turn several times to see all of the images distributed on the virtual cylinder depending on the scaling gain. The second mode is tilt-acceleration, where the roll and pitch angles are respectively mapped to the acceleration of the left/right and up/down movements of the virtual camera. The tilt-acceleration mode makes it easy to browse a wide area with a small tilting movement of the user, but often exhibits difficulty for the precise position control of the virtual camera. The last mode is a combination of the two modes. In this hybrid mapping mode, the mapping mode is set to direct mapping for tilting angles measured to be smaller than certain thresholds. Otherwise, the tilt-acceleration mode is used for easy long-distance browsing.

All of the image views and motion mapping modes are implemented in a motion-based image browser using a SONY UMPC. A 3D accelerometer and a LOGITECH QuickCam were attached to the UMPC and communicated via USB 2.0. The image browser combined with custom-defined buttons can provide usual operations such as browse, select, copy, paste, cut, and delete, in an intuitive manner.

III. Usability Study

We carried out an experiment to compare the relative benefits of different image views and interface methods in the image searching task. Twenty four subjects participated in the experiment. The participants were all engineering students (18 males and 6 females) and aged between 17 and 25 with an average of 21. We used a two-factor within-subject experimental design. The independent variables were image viewing method and interface method. The tiled and fish-eye views were used for viewing method. For interface method, the tilt acceleration and hybrid modes were used in addition to the conventional button-based interface. The dependent variables were task completion time, ease of use, preference, intuitiveness, naturalness, and enjoyment. Each dependent variable was obtained by questionnaire except for the completion time.

For each combination of the image views and the interfaces, the subject performed two kinds of image searching tasks. One task was to search an image the overall location of which was provided to the subject in terms of a mini-map. The other task was to find an image with its location totally unknown to the subject.

The experimental results showed that the task completion time of the tiled view is lower than the fish-eye view and the task completion time of the button interface is lower than those of other interface modes in all experimental conditions (see Fig. 3). These results indicated that the task performance of the conventional interaction method was higher than those of newly tested methods. The results of other qualitative measures (ease of use, preference, intuitiveness, naturalness, and enjoyment) are presented in Fig. 4 for each combination of experimental factors. The subjects tended to prefer the conventional interaction method, except for enjoyment.

Subject debriefing revealed why the motion-based interfaces were shown to be inferior to the traditional button-based interface. It was reported that since our motion-recognizing UMPC was still heavy, the subjects could not quickly manipulate the device. Another reason was the inaccuracy of the current motion sensing capability, which often made the precise control of a screen cursor very difficult.

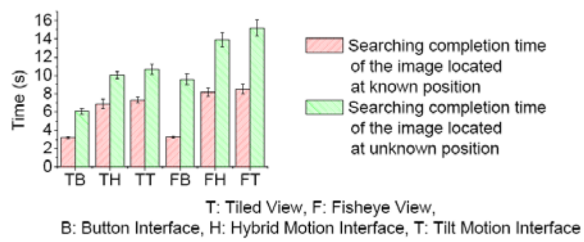


Fig. 3. Task completion time of image searching.

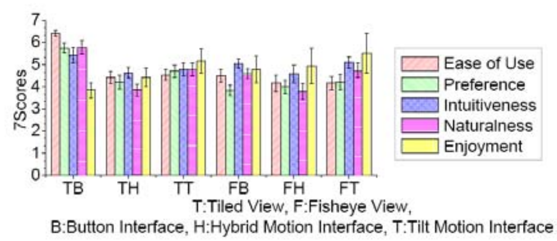


Fig. 4. Qualitative measures estimated in the experiment.

IV. Conclusion

In this paper, we presented a new paradigm of browsing and manipulating images in the mobile device using the user's motion. Two image views and three motion mapping modes were designed and implemented in a motion-based image browsing program in a hand-held PC. The user study showed that although our motion-based image browser is intuitive and fun to use, there are still technical challenges to be resolved for better performance.

References

[1] D. Patel, G. Marsden, M. Jones, and S. Jones, "Improving Photo Searching Interfaces for Small-screen Mobile Computers," in Proceedings of the 8th conference on Human-computer interaction with mobile devices and services: ACM Press, 2006, pp. 149 - 156.

[2] J. Hwang, J. Jung, and G. J. Kim, "Hand-held Virtual Reality: A Feasibility Study," in Proceedings of the ACM symposium on Virtual reality software and technology: ACM Press, 2006, pp. 356 - 363

[3] C. Gutwin and C. Fedak, "Interacting with Big Interfaces on Small Screens: a Comparison of Fisheye, Zoom, and Panning Techniques," in Proceedings of Graphics Interface: Canadian Human-Computer Communications Society, 2004, pp. 145-152.

[4] J. Hwang, G. J. Kim, and N. Kim, "Camera based Relative Motion Tracking for Hand-held Virtual Reality," presented at NICOGRAPH International 2006.