Multi-pointing Method Using a Desk Lamp and Single Camera for Effective Human-Computer Interaction

Taehoun Song[1], Thien Cong Pham[1], Soonmook Jung[1], Jihwan Park[1], Keyho Kwon[2], and Jaewook Jeon[2]

[1] School of Information and Communication Engineering, Sungkyunkwan University, Suwon, Korea
{thsong,pham,kuni80,fellens}@ece.skku.ac.kr
[2] School of Information and Communication Engineering, Sungkyunkwan University, Suwon, Korea
{khkwon,jwjeon}@yurim.skku.ac.kr

Abstract. Multi-pointing has become an important research interest, and is used in many computer applications to allow users to interact effectively with a program. Multi-pointing is used as an input method, and can also be fun and very user-friendly. However, in order to use the method, a complex and expensive hardware configuration is required. This paper presents a new and low cost method of multi-pointing based on a simple hardware configuration. Our method uses dual hand recognition, a table lamp, and a single CMOS camera. The table lamp provides a steady illumination environment for image processing, and the CMOS camera is mounted to maintain good stability. A single camera is used for dual hand recognition to achieve multi-pointing. Therefore, image processing does not require intensive computing which allows us to use a stand-alone system (including a 32 bit RISK processor). The results of the proposed method show that effective control navigation of applications such as Google Earth or Google Maps can be achieved.

Keywords: Multi-Pointing, Hand Recognition, Human-Computer Interaction.

1 Introduction

Multi-pointing on a computer screen is an effective inputting method that is generating considerable research interest. Multi-pointing studies have increased recently due to the development of touch-pads and sophisticated camera applications. The latest display screens, used in personal computers, are typically liquid crystal displays(LCDs), which are also used as the displays in embedded systems. Embedded systems usually use seven inch LCD screens with touch-pads. Unfortunately, the normal display systems of personal computers cannot use such touch-pads because, in

J.A. Jacko (Ed.): Human-Computer Interaction, Part III, HCII 2009, LNCS 5612, pp. 520–525, 2009.
© Springer-Verlag Berlin Heidelberg 2009

general, PC' display resolution is much greater then the resolution in embedded systems.

Research on camera-based multi-pointing is limited since camera-based methods can increase the frame rate of the system, thereby increasing hardware configuration requirements for real-time performance. The input devices used in PC systems normally include keyboards and mice, and possibly joysticks and tracker-ball devices [1]-[4]. These input methods do not support multi-pointing and commercial touch-pads have limited multi-pointing ability with which to support desktop PC display systems.

2 Motivation

The proposed multi-pointing method is performed using dual hand recognition, using a table lamp and a single CMOS camera. The aim is to support multi-pointing using a normal LCD monitor through a simple hardware configuration without intensive image processing. Section III describes multi-pointing environments and the multi-pointing strategy, including skin color information extraction, dual hand recognition, and hand position estimation. Section IV discusses the implementation scenario and the required interface commands. Section V presents results from our proposed multi-pointing system, and some conclusions from our research.

3 Simple Strategy for Multi-pointing

3.1 Overview of Multi-pointing Environments

The experimental environments considered in this study are shown in Fig. 1. A CMOS camera mounted with a table lamp views the LCD monitor of the PC'. The images captured by the camera are transmitted to a stand-alone system consisting of a Marvell's PXA320 processor through DMA, via the PXA320's Quick Capture Interface port. Dual hand recognition can be obtained through the received CMOS camera's image using the multi-pointing strategy described in this section. We show that the Google Earth application program can be controlled by the commands generated from the dual hand recognition system. This allows screen control commands to move forward and backward, turn left and right, and zoom in and out on the axis of the PC screen. These screen-based control commands can be control Google Earth program, and can also control the navigational aspects within the Google Maps application.

Figure 2 shows a block diagram for the experimental environments. The blocks of Fig. 2 represent all development and applications. The main components of proposed multi-pointing method are shown in five bold rectangle blocks. Our proposed multi-pointing method system is not heavy computing performance, and can be embedded to small size.

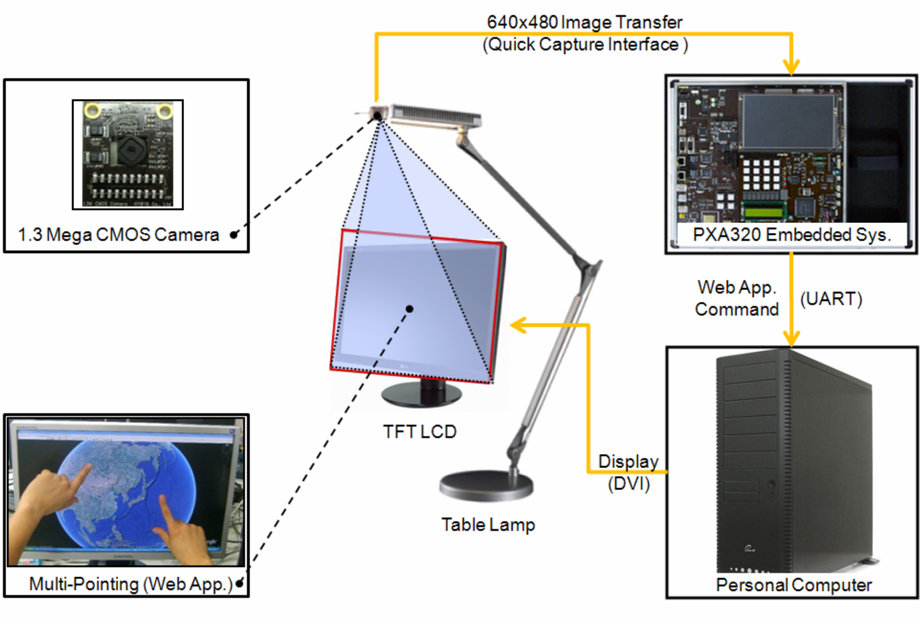

Fig. 1. Multi-pointing experimental environments

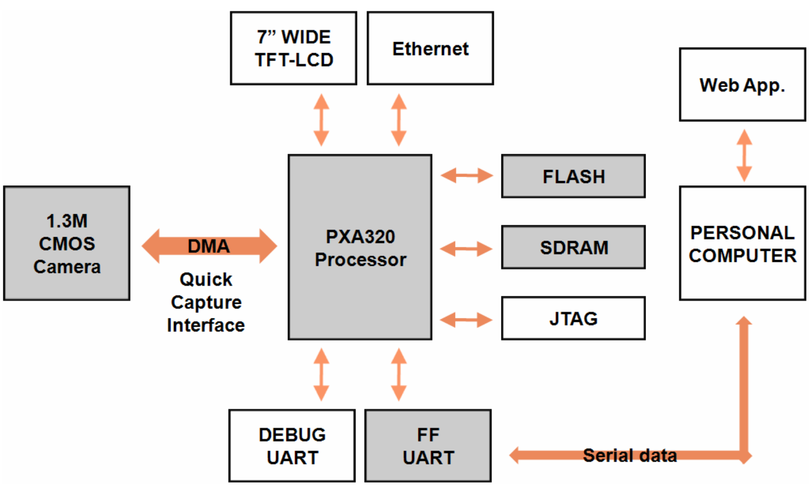

Fig. 2. Multi-pointing system block diagram

3.2 Multi-pointing Strategy

A. Skin color extraction

Our skin color extraction approach uses a red-green-blue (RGB) color-based threshold method, whereas previous studies used the HSV color-space for better illumination tolerance [5], [6]. The working conditions in [5], [6] changed very quickly, the authors detected the skin in real videos, but the algorithm needed an extensive skin color database requiring many skin sample images. To avoid this dependence on an extensive database, our approach uses a much narrower working environment and requires much higher accuracy. The skin color extraction can be used on an input image by manually extracting a trial color so that the RGB characteristics of the hand's color can be obtained, and the threshold values can be used (Fig. 3).

The proposed method is performed in two steps:

1. Hand skin color extraction.
2. Skin color threshold value created by the proposed method: sample images are captured in the same working environment as the target hand color extraction program.

After the skin color extraction step, background and noise effects remain in the resulting image, and future work will address these problems by adopting appropriate filtering of the affected pixels.

B. Background and small noise removal

This method is referred to in [7] and follows these steps to remove any small noise effects:

1. Every connected component is extracted.
2. Small connected components containing less than a pre-defined number of pixels are removed.

End of this step, there are still some big components remaining; these will be filtered next.

C. Hand recognition

The relative information between any two components is considered in the recognition of the hands (Fig. 5). The hand extraction process described in [7] states that many conditions exist for effective hand recognition, and thus many evaluation functions must be applied to assess the result. However, our proposed method removes many of these hand recognition conditions by using a virtual grid, as shown in Fig. 4. In this way, an evaluation function to recognize the hands can be used to detect some key features that form the important connected components.

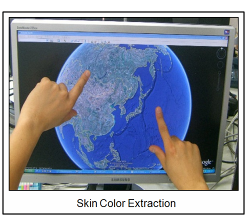

Fig. 3. Color Extraction

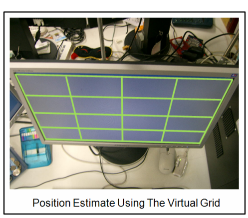

Fig. 4. Virtual Grid

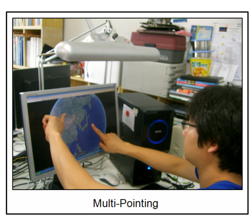

Fig. 5. Multi-Pointing

The proposed hand recognition is performed in three steps:

1. Size of a component A: The number of pixels in component A.
2. Size index of a component A: The index of A, using its size information, compared to the other remaining components.
3. Estimate the start and end positions of a component A: Estimate the component's start and end positions using the virtual grid on the screen.

4 Implementation

The intended implementation of the proposed multi-pointing method is to control the Google Earth application. Google Earth is a global map searching program, based on the World Wide Web. Our implementation can control Google Earth in the following ways: move up image (by the forward command); move down image (by the backward command); turn left image (by using the left command); and turn right image (by using the right command). It can also control the image zoom in and out by using the zoom-in/out commands. Each command generated by the results of the hand recognition procedure is shown in Fig. 6.

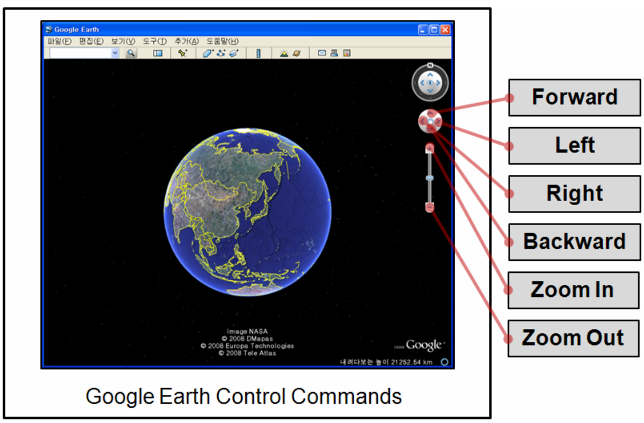

Fig. 6. Implementation Google Earth Control

5 Conclusion

The proposed multi-pointing method, using a desk lamp and single camera, can generate four directional commands: forward, backward, left, and right. In addition, it can generate zoom-in and zoom-out commands, based on human hand position and figure extraction. The Google Earth application controlled using the new approach working in real-time. This study focused on developing an effective input method using a simple multi-pointing approach which can promote the user experience and be fun to use when communicating with computer applications. It is hoped that the proposed multi-pointing method can be a useful and effective way to replace traditional input devices such as keyboards, mice, and joysticks.

Acknowledgments. This research was supported by MKE, Korea under ITRC IITA-2009-(C1090-0902-0046).

References

1. Zhai, S.: User performance in Relation to 3D Input Device Design. Computer Graphics 32(4), 50–54 (1998)
2. Zhai, S., Kandogan, E., Barton, A.S., Selker, T.: Design and Experimentation of a Bimanual 3D Navigation Interface. Journal of Visual Languages and Computing, 3–17 (October 1999)
3. Ku, J.Y., Hong, J.P.: A Study on the Controlling Method for Remote Rehabilitation Assisting Mobile Robot Using Force Reflection Joystick. Journal of the Institute of Electronics Engineers of Korea 40, 26–34 (2003)
4. Lapointe, J.F., Vinson, N.G.: Effects of joystick mapping and field-of-view on human performance in virtual walkthroughs. In: Proceeding of 1st International Symposium on 3D Data Processing Visualization and Transmission Padova, Italia, June 18-21 (2002)
5. Sigal, L., Sclaroff, S., Athitsos, V.: Skin Color-Based Video Segmentation under Time-Varying Illumination. IEEE Trans. Pattern Analysis and Machine Intelligence, 862–877 (2004)
6. Dadgostar, F., Barczak, A.L.C., Sarrafzadeh, A.: A Color Hand Gesture Database for Evaluating and Improving Algorithms on Hand Gesture and Posture Recognition. Research Letters in the Information and Mathematical Sciences 7, 127–134 (2005)
7. Pham, T.C., Pham, X.D., Nguyen, D.D., Jin, S.H., Jeon, J.W.: Dual Hand Extraction Using Skin Color and Stereo Information. In: IEEE International Conference on Robotics and Biomimetics, Bangkok, Thailand, February 21-26 (accepted, 2009)