A stereo camera and mini-projector for video communication in cell-phone calls

Gulomjon SANGIROV*, Yongqing FU*, Ikechi Augustine UKAEGBU**, Jamshid SANGIROV**

*Information and Communication Engineering Dept, Harbin Engineering University (HEU), Harbin, China

**Electrical Engineering Dept., KAIST, South Korea

sgulom@rambler.ru

Abstract - **A novel cost-effective video communication technology for cell phones using mini-projector and stereo-camera has been introduced. In our proposed technique a virtual screen is created by displaying an image or video. And using the displayed images it is possible to interact with the virtual screen. To demonstrate the capability of our proposed technique, we have displayed dialling numbers to make a phone call. Where dialling numbers are displayed by the mini - project and using these dial-in numbers it is possible to dial the destination number and accomplish a call. To press a given number the finger point should stay over the displayed number till the finger point is detected. For detecting which number has been pressed first the captured images are converted to binary image and then finger detection is done using edge detection algorithm. The stereo camera and mini-projector can be mounted on top of a vision glass. The proposed technique uses a flexible display for projecting the dial-pad for cell-phone calls. Hence, the virtual screen is created by the mini - projector, where dial-up is accomplished by pressing the numbers on the virtual screen.**

***Keywords*— stereo camera and projector, cell phone dial, distance estimation, virtual screen**

I. INTRODUCTION

As the technology develops low-cost mini-projectors are being produced, hence is allowed the building of cheaper and more flexible virtual 3D displays and virtual screens [1-2]. And further improvements of distance measuring using mono- and stereo-camera algorithms have been proposed [3-5], and using object tracking algorithms [6]. The mono- and stereo-camera algorithms have been significant developments by Pranav Mistry related to virtual screens, interactive screens and object tracking [7-10]. The development of sixth sense proposal for reading and interacting with interactive and smart visual contents [9-10] was an important contribution which showed the feasibility of applying interactive virtual screen to many applications. And demonstration of the hands free blinkBot control for moving objects by blinking an eye showed the importance of possibility for using a visual content for disabled people [8].

However, using more complex algorithms for creating an interactive virtual screen may require more memory and stronger computational resources. Therefore, we have proposed a simple way of implementing an interactive virtual screen based on edge detection technique. In our proposed interactive virtual screen we have used a standard low resolution cell-phone camera. Thus, the proposed technique can be implemented using a low-cost cameras and a mini-projector.

Our environment friendly virtual display for video communication technology in cell phones uses mini-projector and camera, where mini-projector is used to display the dial pad numbers and camera is used to capture and detect a finger point over dialling numbers. For dialling a cell phone number pressing a physical dial pad is not required, hence it is done by the gesture of the finger pointing instantaneously. It gives relief to fingers by not pressing the dialling pad continuously and using the gesture over the displayed numbers. This technique uses two stereo cameras to measure distance of dialling numbers projected by miniature projector. For detecting which number has been pressed binary edge detection algorithm has been deployed [11].

II. THE DESIGN STRUCTURE OF THE PROPOSED SYSTEM

The stereo cameras mounted on top of glasses are used to capture images shown in Fig. 1. Then the distance of projected numbers is computed. According to a distance the pressed number by finger point is determined. The object distance measure can be generalized into two groups: one is active and the other one passive. Inactive distance measure separate transmitter and receiver are required. Where the transmitter sends a signal and the receiver receives the reflected signal from the target object and finds the distance by comparing the time received with transmitted signals. The passive distance measuring system to estimate the distance only received or captured data (by the camera). The passive distance measure is less complicated and cheaper to implement since the process can involve only captured images rather than sending and receiving the information.

The dial-pad numbers are projected inside rectangle to ease finding the finger point. The size of the rectangle may differ according to a distance of projection from camera as shown in Fig. 2. Thus the size of the rectangle is saved for various distances. When the finger point detection starts, the distance of projected numbers from the camera is calculated by stereo camera system. Then the camera starts recording to find the indicated number to fix it and display the result. This process continues until all numbers is pressed and call button is pressed.

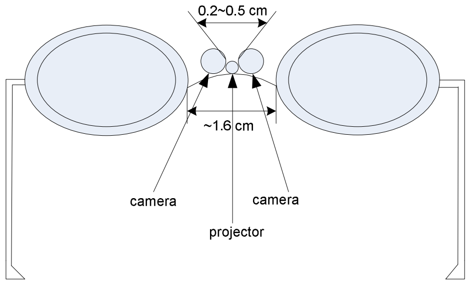
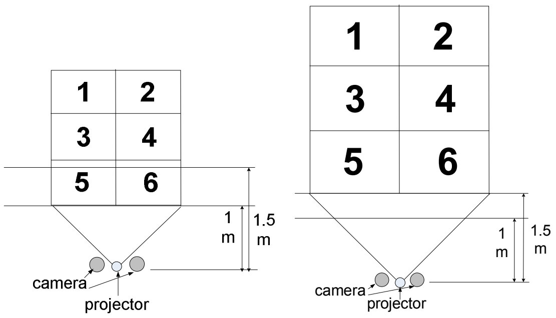


Figure 1. Glasses with camera, mini-projector and radar system



(a) (b)

Figure 2. Number projection from distance of a) 1m and b) 1,5 m

III. A STEREO CAMERA SYSTEM FOR DISTANCE ESTIMATION

In order to find the distance of displaying image a small object can be used in the middle or on top of the dial-in numbers. There are some works on distance measuring based on captured images using mono camera [3], [4] and stereo camera [5]. The red Laser spot and green LED are used for mono camera distance measurement [3]. In stereoscopic image based distance measure projecting the colour spot (green LED or red Laser) can be used to improve the accuracy of distance measurement. The distance can be estimated by founding the pixel offset using the stereoscopic measurement method [5]. The distance estimation equation based on an object is derived in Eq. (1):

$$D = (B * x_0)/(2\tan(\varphi/2) * offset), \qquad (1)$$

Where B is the distance between cameras, x_0 is the number of the horizontal pixels, φ is viewing angle, and *offset* is the pixel offset of an object from the left and right side cameras.

The pixel offset is a varied distance of the same object taken by right camera and left camera, the pixel offset is illustrated in Fig. 3. The equation of pixel offset is estimated by adding two distances of the same object from a central point:

$$offset = x1 + x2, \qquad (2)$$

where *x1* is a distance of an object from center taken by right camera and *x2* is the distance of an object from center taken by left camera. The distances from camera to object and between left and right camera are shown in Fig. 3.

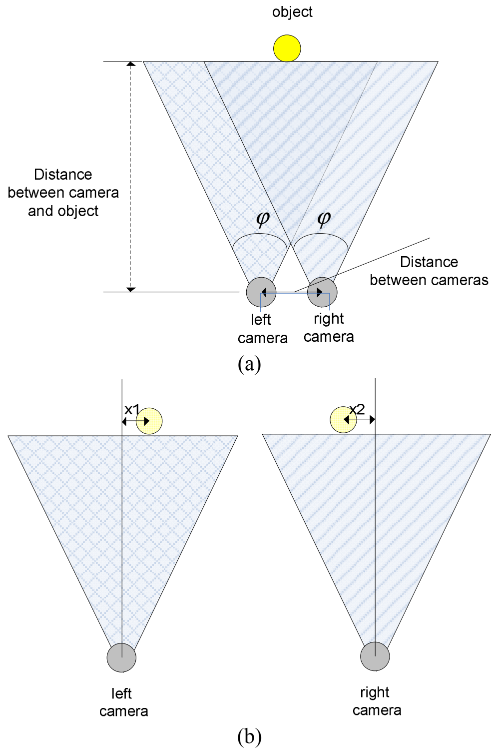


Figure 3. a) Stereo camera and view angle with object and b) distance varaiance of the same object between two pictures.

IV. IMPLEMENTATION

The implementation of the algorithm is shown in Fig. 4. The pseudo-code is described as follows:

- *Begin: System starts*
- *Step1: Capture image and convert to binary*
- *Step2: Find the distance of the virtual screen*
- *Step3: Run binary edge detection*
    - ***If** (find finger point over number pads)*
        - *go to step3*
    - ***else***
        - *go to step1*
- *Step4:*
    - ***If** (finger point over number pad is determined)*
        - *go to step4*
    - ***else***
        - *go to step1*
- *Step5:*
    - ***If** (still pressed for **M** frames over same number)*
        - *go to step5*

- *else*
  - *go to step1*
- *Step6: Displays the number*
- *Step7: Press call when all numbers are chosen*
- *End: System stops*

First the distance of the projected dial-in numbers is determined. Then edge detection algorithm starts by converting the captured image to a binary image. To maintain the accuracy of the binary edge detection algorithm, the minimum distance from mini-projector to displayed image should be less than 2 meters and no closer than 30 cm. To finding finger point more precisely the size of number pads is saved in the table for different distances with 10 cm gap from 30 cm to 2m. Thus, the searching area can be reduced to the smaller area to save a time for searching whole image for finger point. If the images are taken in colour they should be converted to binary image. Finally, the number with higher ones is chosen from binary image, if the finger point stays over a certain number, within an ***M*** number of image frames, over a certain number. The ***M*** can be around 8~10 frames according to calculate precisely the finger point over a certain number. For mobile phones the 20 image frames per second is captured. Then, that number will be selected and displayed, and there'll be a gap for pressing another number. If any wrong number is pressed then the delete can be pressed to delete the unnecessary number. After all necessary numbers are pressed, and then the call can be pressed. When mounting two cameras vertically, it is important to align them accurately to avoid vertical error [5].

The equation for edge detection for an M number of image frames is expressed as follows:

$$D_{number} = \sum_{i=1}^{M} \left(N_{i,ones} / M \right) \tag{3}$$

where D_{number} is an average number of ones for each number display in a square, $N_{i,ones}$ estimated number of ones for each i frame.

To select the number from dialling pad the maximum D_{number} is chosen from estimated numbers and compared to Constanta B:

$$K_{number} = \sum_{i=0}^{9} \max\left(D_{j,number} \right)$$
$$B \le K_{number} \le D \tag{4}$$

where K_{number} is the selected number from dialling pad, $D_{j,number}$ an average number of binary ones estimated for each j, numbers from 0 to 9, and B is a lower limit and D is an upper limit, they can be estimated empirically to improve the accuracy of the edge detection algorithm.

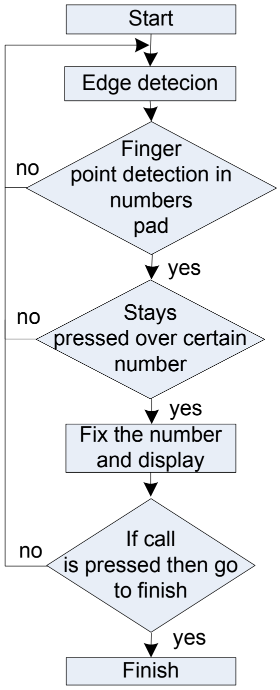


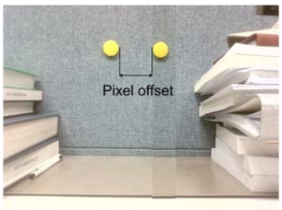
Figure 4. Algorithm.

V. Experimental Results

For capturing the images for distance estimation and finding the offset pixels between stereo cameras, the standard Samsung cell-phone camera (SCH-W330) has been used. The reason for using this camera is to show that this technique could be implemented by small and ordinary cell-phone camera. The images in Fig. 5(a) and (b) have been captured by two cameras, one is called left side and the other called right side cameras to find the amount of offset pixels. The pixel offset in Fig. 5(c) is used for finding the distance between camera and projected image, this offset is from images taken by left and right side cameras. In order to find the pixel offset, the view angle is needed. If the detailed specification for cell-phone camera is not given and angle view is not known. Then, it is also possible to calculate the view angle of the camera by comparing the pixel offset based on Eq. (1). Hence, the reverse calculation has been done for finding the view angle of the camera in our experiment. And using Eq. (1) we have calculated view angle of the camera which is equal to 31^{o}.

Fig. 6 shows the error rate for distance measure between camera and displayed object. As the distance from camera is increasing the error rate is also increasing. The distance between cameras is 10 cm while displayed object varies from 30 cm to 2m.

(a) (b)



(c)

Figure 5. (a) Image by left side camera, (b) image with right side camera and (c) the distance between objects.

The number of binary ones was 22% and 28% for dial number "1" from captured and converted to binary image as shown in Fig. 7 (b) and (d). From these empirical results we can conclude that the upper and lower limits, *B* and *D,* can be set to 15% to 50% to get more robust results in real-time system implementation. The captured images were taken at approximately maximum and minimum required distances. The image shown in Fig. 7(a) is captured from a 30cm distance and the image of Fig. 7(c) is captured from 2m distance.

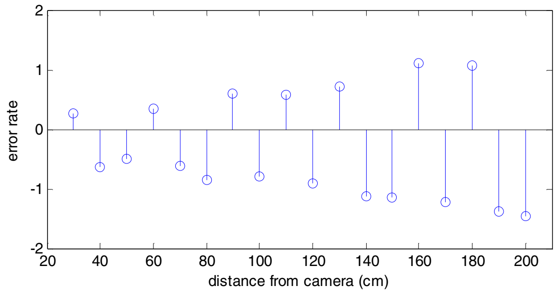


Figure 6. The error rate for distance measure between camera and displayed object.

Since, the colour pixel range is 0 to 255 bits for each colour (RGB colour). The threshold for converting to binary image from colour image for binary "one" are pixels from 100 to 150, the rest is set to "zero". The estimated elapsed time for converting coloured image with size of 480x640 pixels shown in Fig. 7(a) to binary image in Fig. 7(b) is equal to 27.642 ms and the second image shown in Fig. 7(c) has been converted to Fig. 7(d) with estimated time of 27,267 ms. The conversion of colour image to binary image has been accomplished using MATLAB software tool. To get more precise estimation of the distance between the camera and displayed numbers the background needs to be monotonic.

(a) (b)

(c) (d)

Figure 7. Edge detection.

VI. Conclusion

A cost-effective new technique for creating interactive virtual screen for cell-phone dialling based on mini-projector and stereo camera has been introduced. The experimental results are done to prove the applicability of the proposed method using a standard cell-phone camera and mini-projector. Therefore, the simplicity of this technique can be advantageous for implementing of a low-cost virtual screen based cell-phone video communication system.

Acknowledgment

This paper is funded by the International Exchange Program of Harbin Engineering University for Innovation-oriented Talents Cultivation.

References

[1] J Joseph., Jr. LaViola, Michael Katzourin, "An Exploration of Non-Isomorphic 3D Rotation in Surround Screen Virtual Environments" IEEE Symposium on 3D User Interfaces, 2007, pp. 49-54.

[2] M. Birla, and etl, "TouchX- Mouse Simulation Using Gestures on a Virtual Screen" , *IEEE Students' Conference on Electrical, Electronics and Computer Science*, p1-4, 2012.

[3] Masahiro Takatsuka, Geoff A. W. West, Svetha Venkatesh, Terry M. Caelli, "Low-cost Interactive Active Monocular Range Finder", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 444-449, 1999.

[4] Ashfaqur Rahman, "An Image Based Approach to Compute Object Distance", *International Journal of Computational Intelligence Systems*, Vol. 1, no. 4, pp. 304-312, Dec. 2008.

[5] J. Mrovlje, D. Vrancic, "Distance measuring based on stereoscopic pictures", *International PhD Workshop on Systems and Control: Young Generation Viewpoint*, pp. 1-6, 2008.

[6] Jong-Sung Kim, Myung-Gyu Kim, "Golf Ball Detection from Multi-Exposure Images under Heavy Occlusion", *IEEE ICACT*, pp. 1276-1279, 2012.

[7] P. Mistry, P. Maes, "Mouseless", *In the Proceedings of UIST,* pp. 441-442, 2010.

[8] P. Mistry, K. Ishii, M. Inami, T. Igarashi, "BlinkBot - Look at, Blink and Move", *In the Proceedings of UIST*, pp. 397-398, 2010.

[9] P. Mistry, P. Maes, "SixthSense – A Wearable Gestural Interface", *In the Proceedings of SIGGRAPH Asia*, 2009.

[10] P. Mistry, "ThirdEye - Multiple viewers can see different things on a single screen", *In the Proceedings of SIGGRAPH Asia*, 2009.

[11] Hiroyuki Tsuji, Suguru Saito, Hiroki Takahashi, Masayuki Nakajima, "Estimating object contours from binary edge images", *IEEE International Conference on Image Processing (ICIP)*, vol. 3, pp. III-453-6, 2005.