Face Detection from Four Captured Images Related to Intelligent Room for the Deaf

Young-joon OH, Jong-in KIM, Taeh-yun YOON, Kee-chul JUNG
(*Dept. of Digital Media, Soongsil University, Seoul* 156-743, *South Korea*)

Abstract – The intelligent environment needs Human-Computer Interactive technology (HCI) and a projector projects screen on wall in the intelligent environments. We propose the front-face detection from four captured images related to the intelligent room for the deaf. Our proposal purpose is that a deaf user faces wall displaying everywhere. The system gets the images from four cameras, and detects the user region from a silhouette image using a different method, detects and cuts a motion body region from a different image, and cuts the vertex-chest region from the cut body region image. The system attempts to find front-face using Haar-like feature, and selects a detected front-face image from the vertex-chest region. We estimate the front-face detection of recognition rate, which shows somewhat successfully.

Key words – *face detection*; *hand gesture*; *intellegence room*; *everwhere dispalys*; *image processing*

Manuscript Number: 1674-8042(2010)04-0338-05

dio: 10.3969/j.issn.1674-8042.2010.04.09

1 Introduction

1.1 Background of the study

The three dimension-based the intelligent room environment needs human-computer interactive technology, such as 3D graphic technology, artificial intelligence and image processing. The intelligent room has whole system with information and communication technologies and presents creating interactive environments to physical space[1], and that can control video camera, projector, lighting apparatus and home appliances using ubiquitous computing. Not only a video camera perceives user access automatically, analysis user movement and grasps user's aim, but can detect user's body and face using Hear-like feature. A projetor projects to wall or the floor of a room by deflected mirror, and the system detects hand motion from user on wall screen for screen object touch using video camera. A user can have interaction with the everywhere displays.

Recently, many researchers developed several user interface device for the disabled in the intelligent room or virtual environment. For example, Do et al developed the soft remote control system using hand pointing gestures, and the user commands homes appliance using hand pointing and the system recognized hand pointing from 3 color cameras[2]. However, some researchers did not consider the specialized technology for deaf people's convenience and did not research visual communication method through everywhere display. Deaf people cannot hear or speak over because already adopted sign language, writing or lips-reading into mainly communication[3]. In the present, the existing deaf-related technology is sign interpreting system, text messaging system, video cellular phone, TV close-caption decoder and facsimile[4].

1.2 Everywhere displays

In 2001, Claudio Pentanes developed the everywhere projector system that projects display and controls move and rotate of every surfaces on wall screen as "touch screen"[5]. He asserted that the system is composed of a projector, a mirror and a camera. Choi, et al proposed that the real-time geometric calibration system as shown in Fig. 1[6]. A user touches a processed image on any sides of the display wall and controls image moving to other side. Their system generated the corrected images from projector by the degree angle of a mirror. Their proposed system can be applied to the ubiquitous space like the interactive environment or the entertainment playing.

1.3 The purpose of our study

In this paper, we propose the face detection from four captured images related to the intelligent room for the Deaf. The purpose of our study is that a deaf user

※ **Received**: 2010-09-12
Project supported: This research was supported by the Ministry of Knowledge Economy, Korea, the ITRC (Information Technology Research Center) support program (NIA-2009-(C1090-0902-0007)), the Contents Technology Research Center support program
Corresponding author: Young-joon OH(yjooh@ssu.ac.kr)

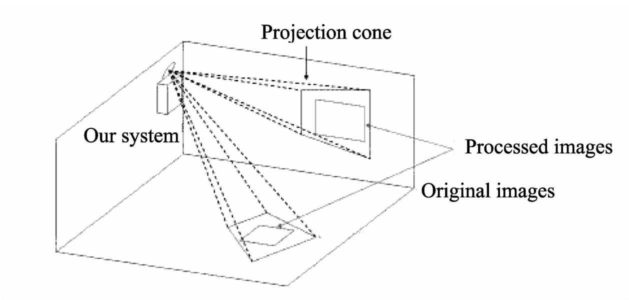

Fig.1 Everywhere system for automatic calibration[6]

communicates with a hearing person via visual messages through operating of the projected everywhere display when gaze projected everywhere display. A deaf user faces wall with everywhere display directly and the system recognize hand gesture commands from face-detected user in priority. A deaf user commands the program execution, controls movement of objects using hand/finger pointing and touches button on projected screen using finger when he/she gaze everywhere display. In further, we consider to develop system is that face-gazed user can controls home appliances using hand pointing Four cameras project interior space of room, and the system gets four capture images from four cameras, and detects the user region from a silhouette image using the difference method, and detects front-face using Hear-features. In finally, the system selects a detected front-face image from four images. We describe this paper as follows. We discuss specification of the everywhere display projector and the intelligent room in section 2. We emonstrate body detection and Hear-like feature in section 3, and demonstrate the face detection and the face classification for front-face detection in section 4. We show experimental results in section 5. Finally, we present conclusions in section 6.

2 Intelligent room

2.1 Concept of the intelligent room

Rodney A. proposed a concept of the intelligent room at the MIT in 1991[7]. The intelligent room have computational environment and a user did not use keyboard, mouse, monitor but can be control by body gesture, hand gesture, face feature that are able to be recognized by perceptual interfaces in the intelligent room and some interfaces assist user's control method of home appliances and lighting goods[8]. The intelligent room with ubiquitous computing can be adopts to smart home or smart office, and acknowledges the intelligent room environment and modulates illumination and heating conditions and grasps states of users and home appliances continuously[9].

2.2 Intelligent room with four cameras

We illustrate the environment of the intelligent room as follows. Four cameras are fixed in four corner ceilings of a room. As shown in Fig.2, it is captured image of camera and a camera looks down room scenery in oblique angle but the camera image view is similar with convex mirror and have the part region of image corner is curved because camera lens shape is circle. The system traces user's body walking and stopping, and detects bodies and faces from perfected images at same time during image calibration.

Fig.2 Oblique view of the intelligent room

3 Body motion detection

3.1 Difference image

We describe the procedure of the difference image as shown in Fig.3 and the system subtracts a current frame image as shown in Fig.3 (b) from a previous frame image as shown in Fig.3(b) but remains dynamic marks on the difference image as shown in Fig.3(c). As shown in Fig.3 (d), the system converts from the difference image to the silhouette image using color-grey converting, and dilates white points on silhouette image through the threshold operation and the dilate operation, and the dilated image detects several noises and white region, and the system selects big white region as body motion region from the dilated image.

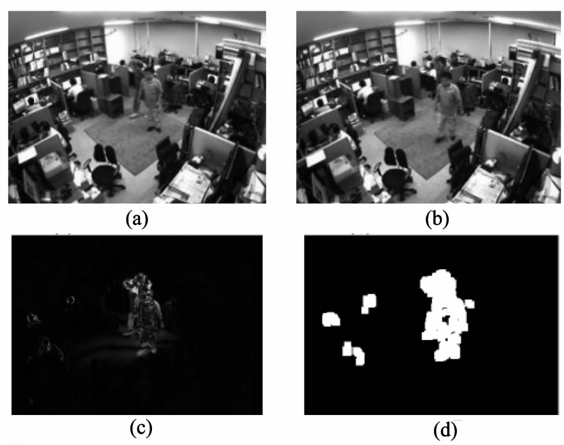

Fig.3 Differece image: (a) Previous frame, (b) Current frame, (d) Difference image, (d) Dilate image

3.2 Body detection

As shown in Fig. 4, the system separates foreground and background from a captured image in order to detect the user body region, and we describe the procedure of body detection through the difference image as follows.

1) The system captures image from camera and a captured image is shown in Fig. 4 (a).

2) The system gets the image differencing using mask operation is shown in Fig. 4 (b).

3) The system converts from the difference image to the silhouette image using threshold operation but detects several noises as shown in Fig. 4(c) and selects a big noise as the body region as shown in Fig. 4(c) on the silhouette image.

4) The system contours body feature with a rectangle line as the selected big-noise region what is shown in Fig. 4(d).

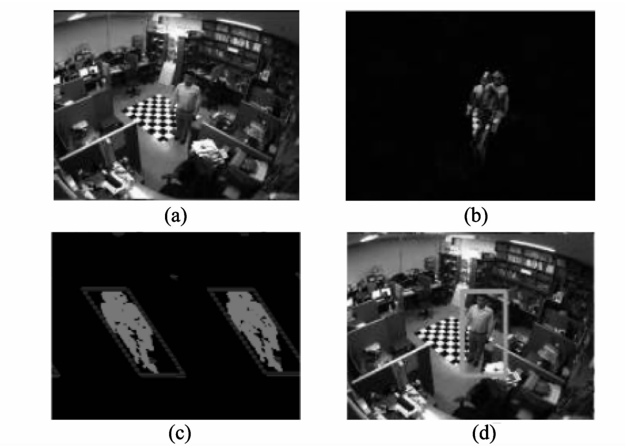

Fig. 4 Procedure of the body detection: (a) Captured image, (b) Difference image, (c) Silhouette image, (d) Body detection

4 Front-face detection

4.1 Haar-like feature

Viola et al. proposed a multi-stage classification procedure. It reduces the processing time during realizing details and complex single stage classifier[10]. Hear-like features are recognized image features using the Hear wavelets algorithm. Hear-like feature is the difference of the total of pixels of the rectangle area and can be used to face detection, body detection and other etc. Rainer Lien hart demonstrated a set of rotated Haar-like features and a basic set of simple Haar-like features for the efficient calculation[11]. As shown in Fig. 5, they indicate that Haar-like features prototypes composes of four edge features(a), eight line features (b) and two center-surround features (c). Their features have one dark region and over one light region[11]. For example, Tomoya et al. developed the multi-viewpoint person tracking system with face detection using Haar-like features in order to can detect the variety of poses such as front-dace and side-head etc[12].

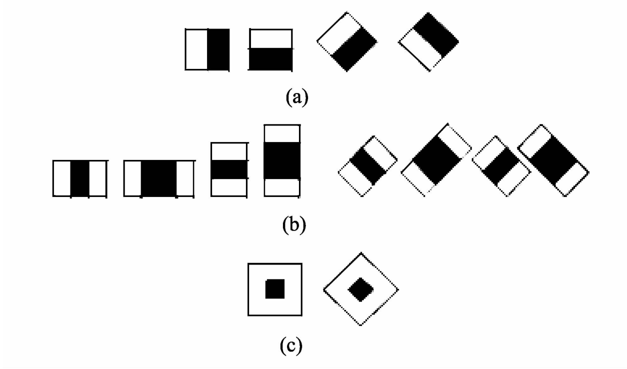

Fig. 5 Haar-like features prototypes[11]: (a) Edge feature, (b) Line features, (c) Center-surround features

4.2 Procedure of the front-face detection

As shown in Fig. 6, we describe the procedure of the front-face detection through the Haar-like features as follows:

1) The system separates a contoured the user body region from a captured image to a cut body region image as shown in Fig. 6(a).

2) As shown in Fig. 6(b), the system contours the vertex-chest region with a rectangle line from a cut body region image. We explain the reason to contours the vertex-chest region, the system can detect front-face easily and can reduce the misunderstanding of front-face based on Haar-like feature. For instance, the system misunderstands any object of captured image as face feature.

3) The system separates a contoured the vertex-chest region and detects front-face as full face obviously using Haar-features as shown in Fig. 6(c). Be the way, we explain classification for the detected front-face in the following and display results of the front-face detection at full length in clause B, Section 5.

4) The system gets to pixel values of rectangle of a contoured front-face as shown in Fig. 6 (c) and duplicates a contoured rectangle to a captured image as shown in Fig. 6(d).

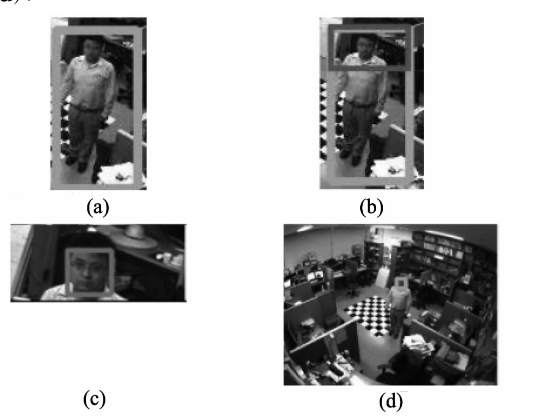

Fig. 6 Procedure of the front-face detection: (a) Body separation, (b) Vertex-chest separation, (c) Front-face detection 1, (d) Front-face detection 2

4.3 Classification for front-face detection

We present classification diagram for the detected front-face as shown in Fig. 7. The system executes the vertex-chest region function, and it attempts to check head on the vertex-chest region through Haar-like feature. In the face detection diverge, the head pose is classified into side-head, back-head, hang-head, nodding-head, tilt-head and front-head, and corresponds to user's movement, and presents head diversity in the camera viewpoint. If the system succeeds in face detection then detects Front face and accepts hand pointing from user and he can express the hand pointing control on everywhere displays. If the system fails in face detection, return to the face-finding in order to checks front face in repetitive.

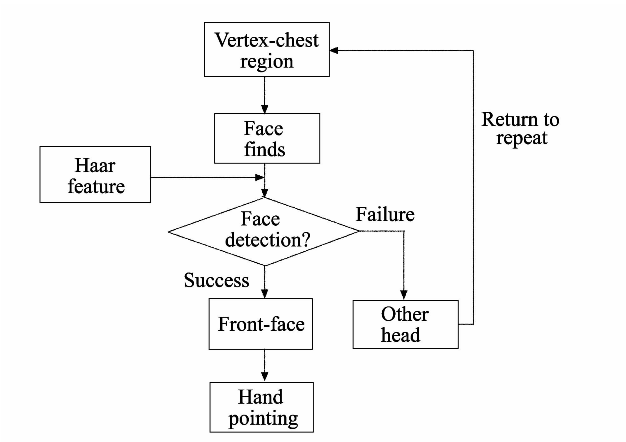

Fig. 7 Classification diagram for the detected front-face

5 Experimental results

5.1 Experimental environment

We use the system that fixed 2.8 GHz CPU and 2 Gbyte memory and connected to four Sony 1/2″ CCD ICX267 cameras of each corner ceilings of a room. The experimental room was width 5.0 m × length 4.0 m. We demonstrate the results of the front-face feature extraction from capture images as shown in Fig. 8. The system traces and detects user moving and stopping and presents each contoured user's body with rectangle line in camera 1～4 images as shown in Fig. 8 but the body region size is gradually varying in corresponding to user's room movement. The system presents a detected front-face and a contoured outline of camera 3 image among four captured images as shown in Fig. 8. We describe user's length location related to camera 1 of Fig. 8 what is shown in Fig. 9, and verify face-detection performance of the system about distance difference. Accordingly, the system can detect little face of user who is located far from a user-tracing. The system measures the distance between user's vertex and camera 1 what was fixed room corner ceiling as shown in Fig. 9. The system measured value of the distance of Fig. 9(a) is 3.5 m and Fig. 8(b) is 2.5 m.

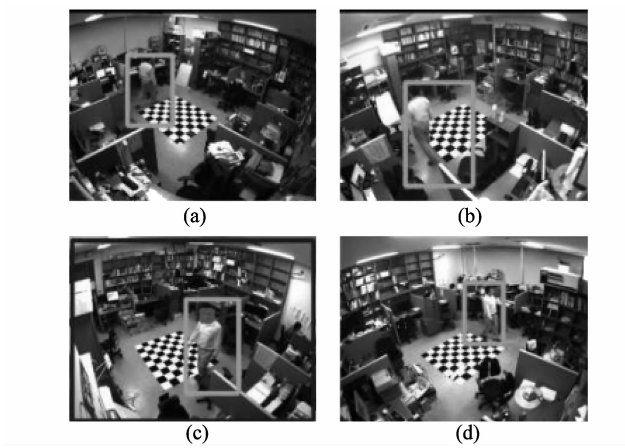

Fig. 8 A front-face detection in four cameras: (a) Camer 1, (b) Camer 2, (c) Camer 3, (d) Camer 4

Fig. 9 User's length location of camer 1: (a) Long length, (b) Short length

5.2 Results of the front-face detection test

As shown in Fig. 10, a user expresses a change of head pose based on head movement to camera with corner ceiling when change working according to his direction. We practice head tracing from cut vertex-chest image based on each camera using Haar-like feature. Of course, we execute front-face detection verification and show results of side-head image (a), back-head image (b), hang-head image (c) and front-face image (d) as shown in Fig. 10. In finally, the system selects front-face image (d) using Haar-like feature because, user gazes at camera directly and front-face has face elements such as eyes, eyebrows, nose and lip. But, the system fails to detect the face as shown in Fig. 10 (a), (b) and (c).

Fig. 10 Head tracing for front-face detection: (a) Side-head, (b) Back-head, (c) Hang-head, (d) Front-head

5.3 Recognition rate

We evaluate capability and performance of the proposed system five rounds. We estimate that front-face detection rate is 86.8% as shown in Tab.1.

Tab.1 Front-face detection of reognition rate(%)

Test round	Four cameras				Average per camera
	1	2	3	4	
1st	95	83	75	73	81.5
2nd	91	75	67	86	79.75
3rd	92	93	94	80	89.75
4th	94	91	93	71	87.25
5th	90	92	96	97	93.75
Average per round	93.6	86.8	85	81.4	86.7

6 Concucions

This work shows that front-face detection from four capture images related to the intelligent room in real-time and we consider utility and convenience of minorities such as the aged and the disabled, and develop a specialized interactive technology of the intelligent room. We address the intelligent room, body detection, face detection from captured images, evaluate capability of the proposed system, and show results of the front-face feature. We consider it to be used as development of the everywhere display system for the deaf. In further, we will study that the next system controls program cursor using arm/finger/laser pointing on the everywhere displays.

References

[1] Steventon, Wright, 2004. Intelligent Spaces: The Application of Pervasive ICT. Springer.

[2] Jun-hyung Do, Jung-bae Kim, Kwang-hyun. Park, Wonchul Bang, Z. Zenn Bien, 2002. Soft remote control system using hand pointing gesture. *Int. Journal of Human-friendly Welfare Robotic Systems*, 3(1): 27-30.

[3] Wikipedia, Deaf _ community. http://en. wikipedia. org/wiki/ Deaf _ community.

[4] Elizabeth keating, Gene mirus, 2003. American sign language in virtual space: interactions between deaf users of computer-mediated video communication and the impact of technology on language practices. Facebook Language in Society, Cambridge University Press, p. 693-714.

[5] C. Pinhanez, 2001. The everywhere displays projector: a device to create ubiquitous graphical interfaces. Proc. UbiComp: Ubiquitous omputing, LLCS 2201, Springer-Verlag, Berlin. p. 315-331.

[6] Hyunchul Choi, Dongwuk Kyoung, Keechul Jung, 2007. Real-time image correction for interactive environment. *HCI International*, 4554: 345-354.

[7] A. Rodney. Intelligence without reason. MIT AI Lab Memo 1293.

[8] Kota Irie, Masaki Wada, Kazunori Umeda, 2008. Construction of an Intelligent Room using distributed camera system. Stereo Vision edited by Asim Bhatti, InTech Education and Publishing, p. 121-130.

[9] A. Rodney, Brooks, Darren Dang, Jeremy De Bonet, Joshua Kramer, John Mellor, Polly Pook, Chris Stauffer, Lynn Stein, Mark Torrance, Michael Wessler, 1997. Intelligent Room Project. Proceedings of the Second International Cognitive Technology, p. 271-279.

[10] P. Viola, M. Jones, 2001. Rapid object detection using a boosted cascade of simple features. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Hawaii, p. 511-518.

[11] Rainer Lienhart, Jochen Maydt, 2007. An extended set of haar-like features for rapid object detection. IEEE ICIP p. 900-903.

[12] Yu Ito, Atsushi Yamshita, Toru Kaneko, 2008. A multi-viewpoint person tracking system based on face detection using Haar-like features and a mean-shift tracking. Asia International Symposium on Mechatronics, Hokkaido University. Japan.