MITTEILUNGEN KLOSTERNEUBURG
65(2015) 4
www.mitt-klosterneuburg.com
Bimonthly ISSN: 0007-5922

An Automatic Image Capturing System Applied to an Identification Photo Booth

Yu-Xiang Zhao[a], Yi-Zeng Hsieh[b,*], Chien-Hsing Chou[c], Jia-Zai Zheng[a], Ting-Jou Ding[d,*], Yu-Sheng Su[e]

[a]Department of Computer Science & Information Engineering, National Quemoy University, Taiwan

[b]Department of Management and Information Technology, Southern Taiwan University of Science and Technology, Taiwan

[c]Department of Electrical Engineering, Tamkang University, New Taipei City, Taiwan

[d]Department of Materials and Energy Engineering, MingDao University, Taiwan

[e]Research Center for Advanced Science and Technology, National Central University, Taiwan

*Corresponding Author:
Yi-Zeng Hsieh, PhD
E-mail address: yizeng.hsieh@gmail.com
Tel: 886-6-253-3131 ext.8122
Ting-Jou Ding, PhD
E-mail address: tjding@mdu.edu.tw
Tel: 886- 4-887-6660 ext.8025

Abstract

With advances in technology, photo booths equipped with automatic capturing systems have gradually replaced the identification (ID) photo service provided by photography studios, thus assisting consumers in saving considerable amounts of time and money. Common automatic capturing systems employ text and voice instructions to guide users in capturing their ID photos; however, the capturing results may not conform to ID photo specifications. To address this concern, this study proposes an ID photo capturing algorithm that can automatically detect facial contours and adjust the size of capturing images. In the experiments, subjects were seated at various distances and heights for testing the performance of the proposed algorithm. The experimental results show that the proposed algorithm can effectively and accurately capture ID photos that satisfy required

specifications.

Keywords: Photo booth, Face detection, Identification photo, Chin Detection

1. Introduction

Traditionally, identification (ID) photos have been captured by photographers, and the films have been developed and printed in a photography studio. However, this process incurs additional costs and time because of the need for professional assistance. Consequently, an increasing number of consumers have begun using automatic machines to obtain ID photos. Automatic photo booths provide an enclosed space for photo capturing (http://www.dnpphoto.jp/). The hardware of photo booths generally includes a light-source module, camera, seat, printer, and an automatic capturing system to assist consumers with photo capturing. To capture ID photos, consumers must be seated on a fixed chair at the appropriate distance, and must follow the instructions provided by the system. Consumers must then adjust their faces in the specified range on the screen to generate the ID photos.

Although capturing an ID photo by using a photo booth can save time and money, the photo quality is difficult to maintain. The Ministry of the Interior of Taiwan makes provision for the specifications of an ID photo (http://www.boca.gov.tw). The face of the consumer must be positioned within the designated area of an appropriate size. Because the assistance of professional staff is not required when operating an ID photo booth, the capturing quality of photos depends entirely on the automatic capturing system of the machine. Most automatic capturing systems incorporate text and voice instructions to guide users to adjust their posture and position their face in the designated area. Although the users may follow all of the instructions, the captured pictures sometimes fail to meet the specifications of an ID photo because of an inappropriate distance or body height in relation to the camera. For example, babies and children generally have smaller faces and shorter heights than adults do. If babies and children are seated at the same distance as that appropriate for an adult, the size of the captured face in an ID photo may be too small. Such a photo would not conform to the specifications of ID photos according to the Ministry of the Interior of Taiwan.

To generate an ID photo, the most crucial step of an automatic capturing system is to detect the facial area in an image. Numerous previous studies have investigated face detection procedures (Yang, Kriegman, & Ahuja, 2002; Hjelmås, & Low, 2001; Zhang, & Zhang, 2010), yielding categories based on Haar-like features and their variations (Viola, & Jones, 2001; Lienhart, & Maydt, 2002, Li, Zhu, Zhang, Blake, Zhang, & Shum, 2002; Jones, & Viola, 2003; Viola, Jones, & Snow, 2003), pixel-based features (Baluja, Sahami, & Rowley, 2004; Abramson, Steux, & Ghorayeb, 2007), binarized features (Froba, & Ernst, 2004; Jin, Liu, Lu, & Tong, 2004; Zhang, Chu, Xiang, Liao, & Li, 2007), generic linear features (Meynet, Popovici, & Thiran, 2007; Chen, Gu, Li, & Zhang, 2001; Liu, & Shum, 2003; Wang, & Ji, 2005), and shape features (Opelt, Pinz, & Zisserman, 2006; Shotton, Blake, & Cipolla, 2005; Wu, & Nevatia, 2005; Sabzmeydani, & Mori, 2007). The proposed algorithms used in these studies have demonstrated satisfied performance in detecting faces in an image; however, two critical points should be considered in such research. First, the applied algorithm should be easily implemented and integrated into the relevant hardware of the photo booth, and the complex and computational load of the applied algorithm should be considered. In addition to detecting faces in an image, the more important concern is that the capturing of images should comply with formal ID photo specifications.

To address the above concerns, this study developed an automatic capturing algorithm for capturing ID photos. The remainder of this paper is organized as follows. Section 2 presents the proposed automatic capturing algorithm, which involves adopting techniques requiring minimal computational load for detecting the facial contours and positions of the lips and chin. Section 3 introduces the relevant hardware of the designed ID photo booth. In Section 4, the experimental results are described, indicating that the proposed algorithm can accurately capture ID photos to comply with formal specifications. Finally, Section 5 concludes the paper.

2. The Proposed Automatic Capturing Algorithm

This study proposes an automatic capturing algorithm designed to capture human faces accurately and comply with ID photo specifications. A flowchart of the algorithm is shown in Fig. 1. Because the algorithm must be incorporated into a photo booth, the computation time and accuracy are crucial for the algorithm design. The implementation of the proposed algorithm is detailed below.

MITTEILUNGEN KLOSTERNEUBURG
65(2015) 4
www.mitt-klosterneuburg.com
Bimonthly ISSN: 0007-5922

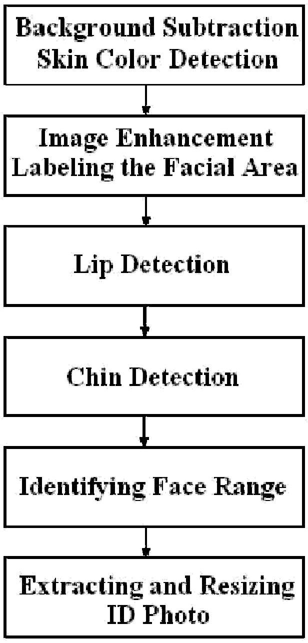
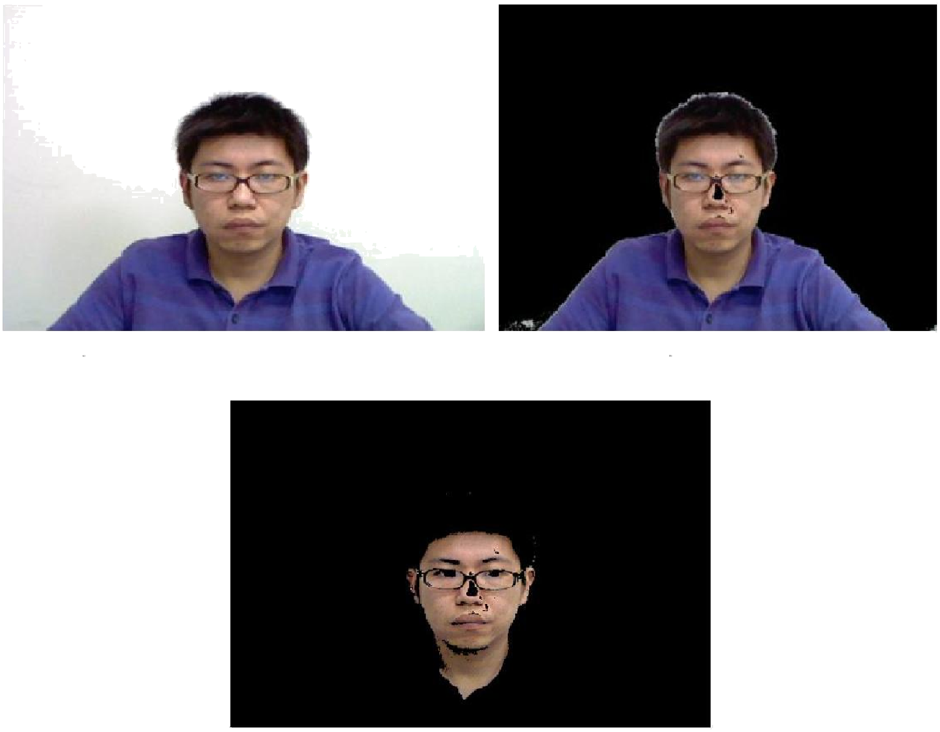


Fig. 1. Flowchart of the automatic capturing algorithm.

Step 1: Background Subtraction and Skin Color Detection

Typically, white is the standard background color used in an ID photo booth. The background subtraction is applied to obtain the foreground image (Fig 2(b)). Skin color is then detected from the foreground image. Numerous methods have been proposed for skin color detection in various color spaces (Kakumanu, Makrogiannis, & Bourbakis, 2007; Chaves-González, Vega-Rodríguez, Gómez-Pulido, & Sánchez-Pérez, 2010; Soriano, Martinkauppi, Huovinen, & Laaksonen, 2003; Vezhnevets, Sazonov, & Andreeva, 2003; Garcia, & Tziritas, 1999; Wang, & Yuan, 2001; Hsu, Abdel-Mottaleb, & Jain, 2002), including the RGB (Soriano et al., 2003; Vezhnevets, et al., 2003), HSV (Garcia, & Tziritas, 1999; Wang, & Yuan, 2001), and YCbCr (Hsu et al., 2002), which have all presented satisfactory results. However, when the HSV or YCbCr color space is employed, the RGB values of each pixel must be converted to the HSV or YCbCr color space. This generates additional computational load, which burdens real-time computing systems. Therefore, this study adopted the RGB color space (Vezhnevets, et al., 2003) for skin color detection, as expressed in the equation below. Fig. 2(c) shows the result of skin color detection.

$$R > 95,\ G > 40,\ B > 20,\ \textit{and}\\ \max\{R,G,\ B\}-\min\{R,G,\ B\} > 15,\\ \textit{and}\ |R-G|>15,\ R>G\ \textit{and}\ R>B \tag{1}$$

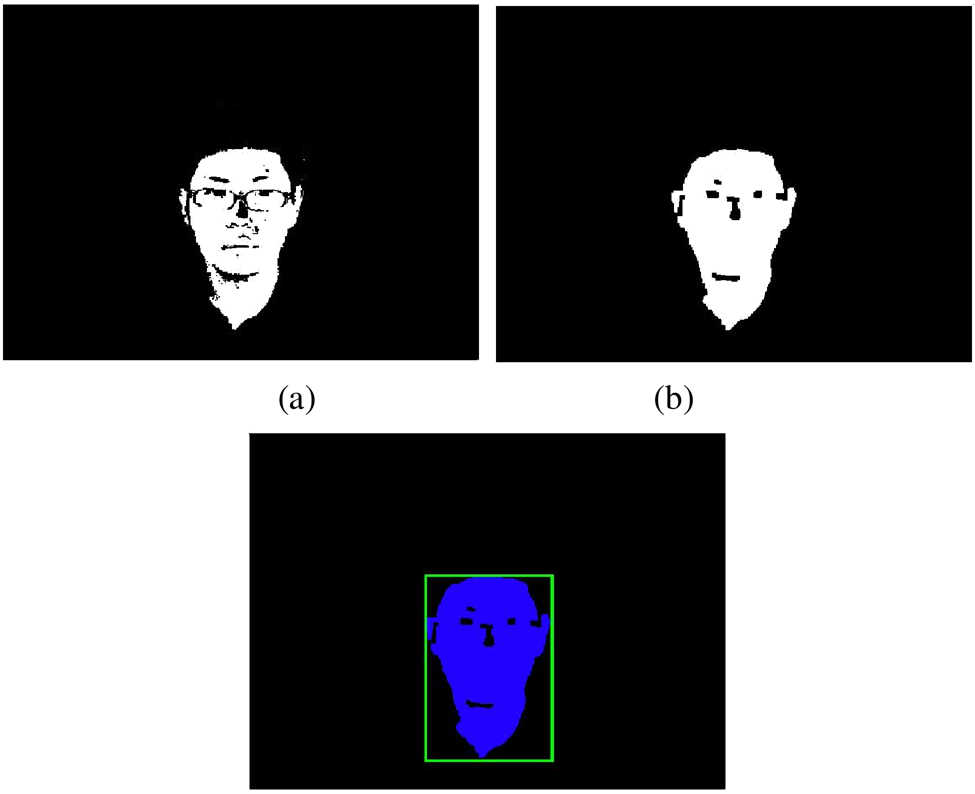


(c)

Fig. 2. (a) Original Image; (b) Extracting the foreground image by applying background subtraction; (c) Detecting the skin color area from the foreground image.

Step 2: Image Enhancement and Labeling the Facial Area

After the skin color area is detected, morphological opening and closing operations are used to remove small objects and holes. Fig. 3(b) shows the enhancement result by applying the opening and closing operations. Because the capturing environment of a photo booth is simple, the connected-component labeling algorithm is employed to identify the largest area of skin color, which consists of the facial area. Fig. 3(c) shows that the largest skin color area is identified using the connected-component labeling algorithm.



(a) (b) (c)

Fig. 3. (a) Area of skin color; (b) Image enhancement by applying the opening and closing operations; (c) Labeling the largest skin color area.

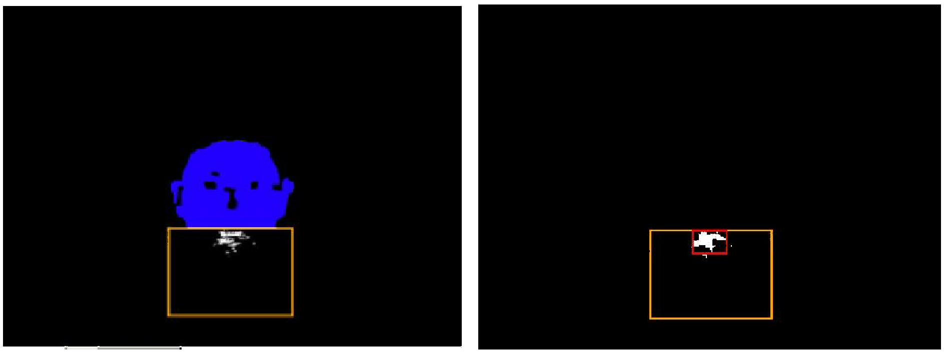
Step 3: Lip Detection

To identify the facial area accurately in Step 3, the position of the lips is located as reference information for determining the facial area. Because the lips are located in the lower half of the face, the search region is limited to the lower half of the skin color area to minimize computational load. In contrast to detecting the positions of the eyes or nose, which requires using complex algorithms, detecting the position of the lips can be accomplished by examining the lip color in the image. A new lip color detection formula is proposed as follows and applied to detect lip positions:

$$(2)\quad 85 \le R \le 170,\ 40 \le G \le 100,\ \text{and}\ R - G \ge 45$$

The orange rectangle in Fig. 4(a) represents the search region. By using the new formula (Eq. (2)), the lip area is detected and indicated as the white region in

Fig. 4(a). Because the color at the junction between the upper and lower lips is duller, the morphological closing operation was performed to connect the upper and lower lip areas (Fig 4(b)). Finally, the connected-component labeling algorithm was applied to identify the largest area of lip colors, and the resulting area (red rectangle) was considered as the lip position.



(a) (b)

Fig. 4. (a) Detecting lip color area with Eq. (2); (b) Locating the lip position (red rectangle) after applying the closing operation and connected-component labeling algorithm.

Step 4: Chin Detection

In this step, the position of the chin is located. First, a downward search is performed within the area from the lower boundary of the lip area to the bottom of the skin color area, which is the white rectangle shown in Fig. 5(a). The position of the chin is searched in this rectangular area (Fig. 5(b)), which is then rotated 90° counterclockwise, as shown in Fig. 5(c), and divided into several vertical blocks with one-pixel width. For each vertical block, the total number of skin color pixels is accumulated. In Fig. 5(c), the histogram represents the numbers of skin color pixels for different vertical blocks. Because the color at the junction between the chin and the neck is generally duller, several pixels of non-skin color appear around the junction. As shown in Fig. 5(c), it is a reasonable assumption that the local minimum number of the histogram is the position of chin.

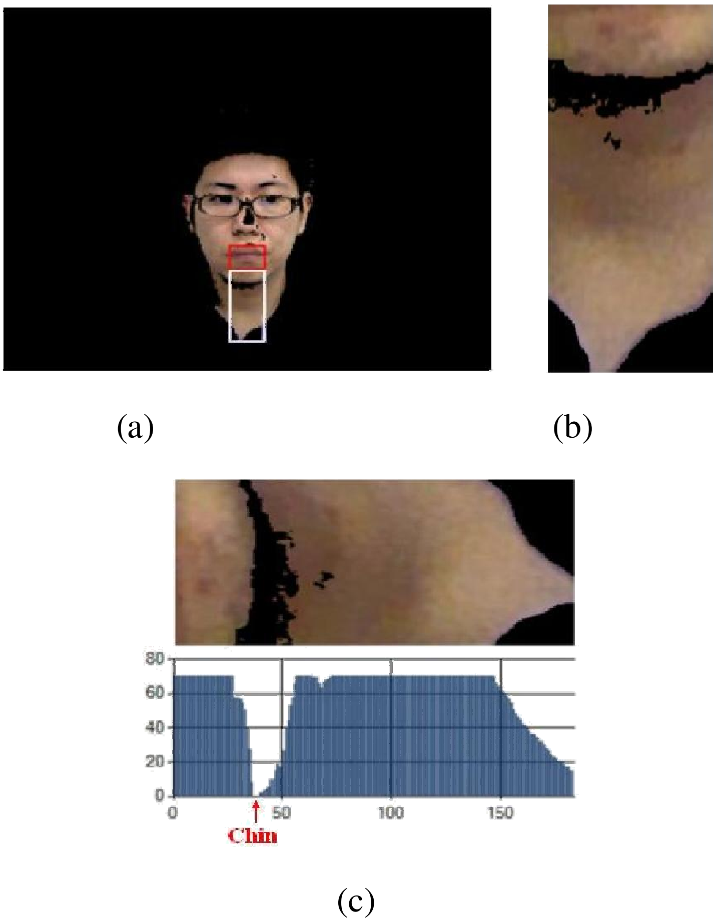
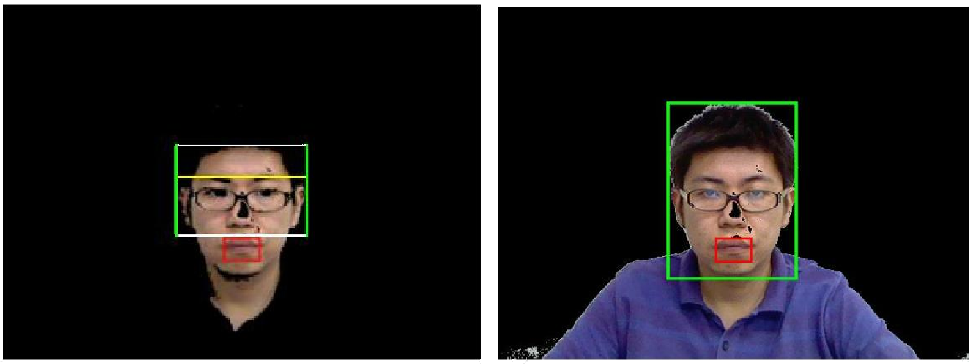


Fig. 5. (a) Searching the chin position from the lower boundary of the lip area to the bottom of skin color area. (b) Area for searching the chin; (c) Histogram represents the numbers of skin color pixels for different vertical blocks.

Step 5: Identifying the Facial Range

In this step, the range of the face is identified. First, as shown in Fig. 6(a), the vertical position (yellow line) at two-thirds of the height between the upper boundary of the skin color area and the upper boundary of the lip area (white lines) is marked. This vertical height is located near the eyebrows. From this vertical height, the left and right boundaries of the facial area are determined by extending rightward and leftward until reaching the boundaries of the foreground image. The green lines in Fig. 6(a) represent the left and right boundaries of the facial area. Subsequently, the height stretching upward from the center of lip area to the upper boundary of foreground image is set as the upper boundary of the face. Finally, the chin position identified in Step 4 is set as the lower boundary of the face. Fig. 6(b) shows the final range identified as the face (green rectangle).



(a) (b)

Fig. 6. (a) Identifying the left and right boundaries of face; (b) Green rectangle is identified as the range of face.

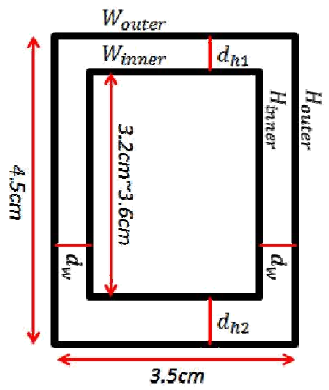
Step 6: Extracting and Resizing the ID Photo

After identifying the facial range, this information can be used to calculate the reasonable range for capturing an ID photo. To comply with the ID photo requirements specified by the Ministry of the Interior of Taiwan, two rectangular areas are designed for capturing an ID photo from an image (Fig. 7(a)). The ID photo specification suggests that the height of the facial region be from 3.2 cm to 3.6 cm; therefore, in this study, the height of the green rectangle was set at 3.5 cm after resizing the facial region. In Fig. 7(a), H_{inner} represents the height of the facial region, H_{outer} and W_{outer} respectively represent the height and width of the ID photo, d_{h1} indicates the distance between the facial region and the upper boundary of ID photo, and d_{h2} indicates the distance between the facial region and the lower boundary of the ID photo. The ratio of d_{h1} to d_{h2} is set at 1:3. Based on the proportion provided in Fig. 7(a), the calculating equation is expressed as follows:

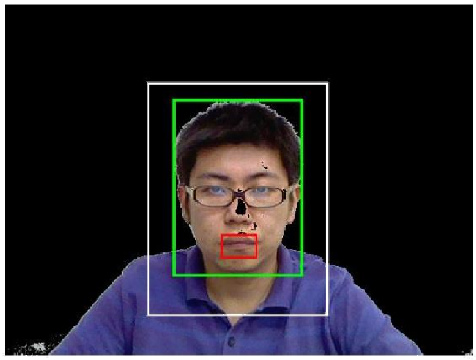
$$= H_{outer} \,,\, W \left(\frac{4.5}{3.5}\right) \left(\times_{inner} = H_{outer} \; H \right)$$

$$3\times(H_{inner}) - H_{outer} \quad \frac{(H}{4} = {}_{h2}\,,\, d \quad = {}_{h1}\, d \tag{3}$$

As shown in Fig. 7(b), the green rectangle corresponds to the facial region previously identified in Step 5, and the white outer rectangle is the image of the ID photo by using Eq. (3). Finally, the image of the white rectangle is extracted and resized at 3.5 cm×4.5 cm. Fig. 7(c) shows the ID photo extracted using the proposed automatic capturing algorithm.



(a)



(b)

(c)

Fig. 7. (a) Two rectangular areas are designed for capturing an ID photo; (b) Green inner rectangle indicates the facial range and white outer rectangle represents the image of ID photo; (c) ID photo extracted by the proposed algorithm.

3. Hardware of the Designed Photo Booth

In Fig. 8, the hardware of the designed photo booth includes a light-source module, seat, printer, personal computer, and two cameras. The proposed automatic capturing algorithm is implemented on the personal computer with Intel Core i5 CPU, memory of 4G bytes, and Windows 7 OS. In contrast to other photo booths, the designed photo booth can automatically detect facial contours and generate ID photo. In addition, the photo booth provides two cameras for capturing (Fig. 8). For standing users, the upper camera is used to capture photos. However, children, wheelchair users, and seated users can use the lower camera to capture photos without adjusting their height. In addition to the seat of the photo booth being collapsible, this design is convenient for a wheelchair user to enter the photo booth.

Fig. 8. Hardware of our photo booth.

4. Experimental Results

In the experiments, the performance of the proposed algorithm was tested by examining six subjects. Each subject was asked to sit at two distances, which were normal and far distances, between the seat and camera. For a specified distance, the subject placed his or her face in five positions during the capturing of a test image: the middle, upper half, lower half, right side, and left side of the image. Finally, 60 images were generated for testing the proposed algorithm. Fig. 9 shows some of the experimental results. Two of the subjects dressed in turtleneck tops, and one of them had long hair. The experimental results illustrate that the proposed algorithm can accurately capture ID photos in various situations. The average computational time of processing a test image is 0.61 second for generating an ID photo. This is an acceptable time for integrating this algorithm into the relevant hardware of a photo booth.

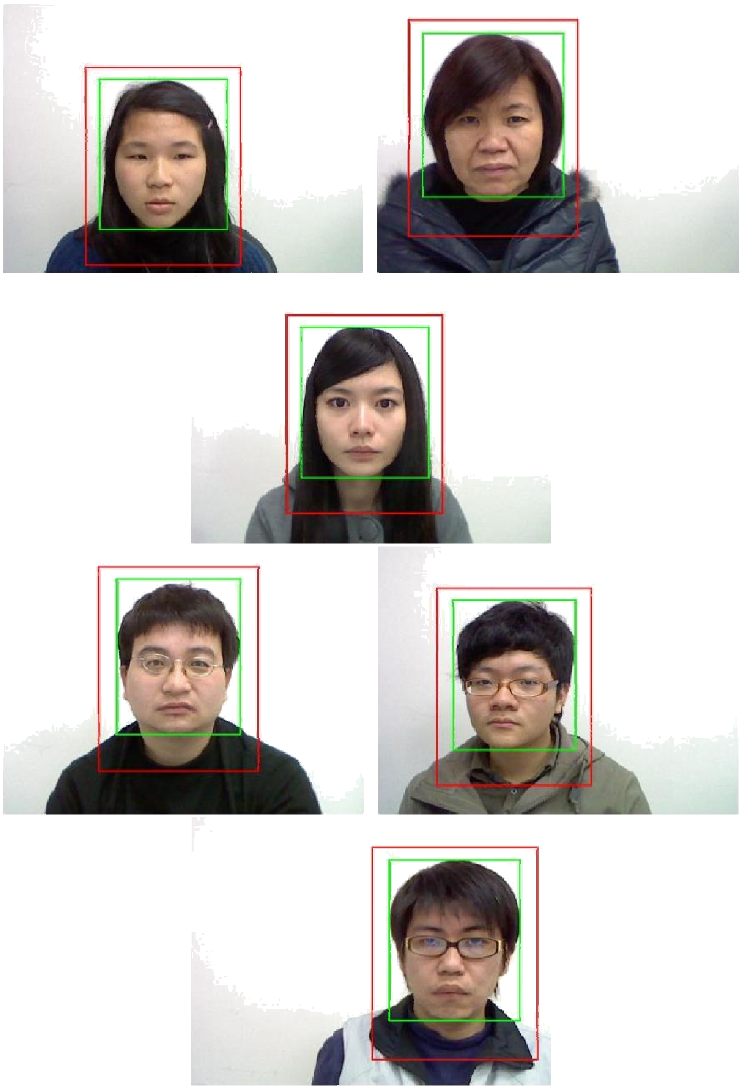


Fig. 9. Some experimental results of proposed algorithm.

According to the specifications of the Ministry of the Interior of Taiwan, the area of the face must be encompassed approximately 70% to 80% of the area of an ID photo. To examine this specification, for each ID photo, the height and width of the facial area are artificially measured, as shown in Fig. 10. The area of the face is then calculated to examine whether this specification is met. Table 1 lists the experimental results, indicating that the performance of the proposed algorithm is satisfactory, and that the ID photos extracted from all of the test images meet the specifications of the Ministry of the Interior of Taiwan.

MITTEILUNGEN KLOSTERNEUBURG
65(2015) 4
www.mitt-klosterneuburg.com
Bimonthly ISSN: 0007-5922

Fig. 10. Artifically measure the height and width of facial area. Table.

1. Ratio of the face area holds in the ID Photo for 60 test images.

Average	Subject ID						Position	Distance
	6	5	4	3	2	1		
74.0	72.8	74.5	72.8	75.0	73.9	75.0	Middle	Normal
73.9	72.8	75.0	73.4	74.5	73.9	73.9	UP	
74.2	72.8	73.9	73.9	75.0	75.0	74.5	Down	
73.7	73.4	75.0	73.9	74.5	71.7	73.9	Right	
74.7	74.5	72.8	75.0	76.1	75.0	75.0	Left	
73.6	73.4	73.9	72.3	73.9	73.9	73.9	Middle	Far
74.4	75.0	75.0	74.5	73.4	74.5	73.9	UP	
75.1	73.4	76.6	76.1	74.5	74.5	75.5	Down	
74.2	72.8	74.5	73.9	74.5	74.5	75.0	Right	
74.6	74.5	75.0	75.0	74.5	73.9	75.0	Left	

5. Conclusion

This study proposes an automatic capturing algorithm to detect facial contours in an image accurately and extract ID photos. The users can sit at various distances and positions during photo capturing. The experimental results show that the proposed algorithm can accurately capture ID photos that satisfy required specifications and be successfully integrated into relevant hardware for constructing an enhanced photo booth a photo booth.

References

Abramson, Y., Steux, B., & Ghorayeb, H. (2007). Yet even faster (yef) real-time object detection. International Journal of Intelligent Systems Technologies and Applications, 2(2), 102-112.

Baluja, S., Sahami, M., & Rowley, H. A. (2004, October). Efficient face orientation discrimination. In Image Processing, 2004. ICIP'04. 2004 International Conference on (Vol. 1, pp. 589-592). IEEE.

Chen, X., Gu, L., Li, S. Z., & Zhang, H. J. (2001). Learning representative local features for face detection. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-1126). IEEE.

Chaves-González, J. M., Vega-Rodríguez, M. A., Gómez-Pulido, J. A., & Sánchez-Pérez, J. M. (2010). Detecting skin in face recognition systems: A colour spaces study. Digital Signal Processing, 20(3), 806-823.

Froba, B., & Ernst, A. (2004, May). Face detection with the modified census transform. In Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on (pp. 91-96). IEEE.

Garcia, C., & Tziritas, G. (1999). Face detection using quantized skin color regions merging and wavelet packet analysis. Multimedia, IEEE Transactions on, 1(3), 264-277.

Hjelmås, E., & Low, B. K. (2001). Face detection: A survey. Computer vision and image understanding, 83(3), 236-274.

Hsu, R. L., Abdel-Mottaleb, M., & Jain, A. K. (2002). Face detection in color images. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(5), 696-706.

Jones, M., & Viola, P. (2003). Fast multi-view face detection. Mitsubishi Electric Research Lab TR-20003-96, 3, 14.

Jin, H., Liu, Q., Lu, H., & Tong, X. (2004, December). Face detection using improved LBP under bayesian framework. In Multi-Agent Security and

Survivability, 2004 IEEE First Symposium on (pp. 306-309). IEEE.

Kakumanu, P., Makrogiannis, S., & Bourbakis, N. (2007). A survey of skin-color modeling and detection methods. Pattern recognition, 40(3), 1106-1122. Lienhart, R., & Maydt, J. (2002). An extended set of haar-like features for rapid object detection. In Image Processing. 2002. Proceedings. 2002 International Conference on (Vol. 1, pp. I-900). IEEE.

Li, S. Z., Zhu, L., Zhang, Z., Blake, A., Zhang, H., & Shum, H. (2002). Statistical learning of multi-view face detection. In Computer Vision—ECCV 2002 (pp. 67-81). Springer Berlin Heidelberg.

Liu, C., & Shum, H. Y. (2003, June). Kullback-leibler boosting. In Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on (Vol. 1, pp. I-587). IEEE.

Meynet, J., Popovici, V., & Thiran, J. P. (2007). Face detection with boosted Gaussian features. Pattern Recognition, 40(8), 2283-2291.

Opelt, A., Pinz, A., & Zisserman, A. (2006). A boundary-fragment-model for object detection. In Computer Vision–ECCV 2006 (pp. 575-588). Springer Berlin Heidelberg.

Soriano, M., Martinkauppi, B., Huovinen, S., & Laaksonen, M. (2003). Adaptive skin color modeling using the skin locus for selecting training pixels. Pattern Recognition, 36(3), 681-690.

Shotton, J., Blake, A., & Cipolla, R. (2005, October). Contour-based learning for object detection. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on (Vol. 1, pp. 503-510). IEEE.

Sabzmeydani, P., & Mori, G. (2007, June). Detecting pedestrians by learning shapelet features. In Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on (pp. 1-8). IEEE.

Vezhnevets, V., Sazonov, V., & Andreeva, A. (2003, September). A survey on pixel-based skin color detection techniques. In Proc. Graphicon (Vol. 3, pp. 85-92).

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-511). IEEE.

Viola, P., Jones, M. J., & Snow, D. (2003, October). Detecting pedestrians using patterns of motion and appearance. In Computer Vision, 2003.

Proceedings. Ninth IEEE International Conference on (pp. 734-741). IEEE.

Wang, Y., & Yuan, B. (2001). A novel approach for human face detection from color images under complex background. Pattern Recognition, 34(10), 1983-1992.

Wang, P., & Ji, Q. (2005, June). Learning discriminant features for multi-view face and eye detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 373-379). IEEE.

Wu, B., & Nevatia, R. (2005, October). Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on (Vol. 1, pp. 90-97). IEEE.

Yang, M. H., Kriegman, D., & Ahuja, N. (2002). Detecting faces in images: A survey. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(1), 34-58.

Zhang, L., Chu, R., Xiang, S., Liao, S., & Li, S. Z. (2007). Face detection based on multi-block lbp representation. In Advances in biometrics (pp. 11-18). Springer Berlin Heidelberg.

Zhang, C., & Zhang, Z. (2010). A survey of recent advances in face detection. Tech. rep., Microsoft Research.