MVA2005 IAPR Conference on Machine VIsion Applications, May 16-18, 2005 Tsukuba Science City, Japan
8-14

Novel Pose-Variant Face Detection Method for Human-Robot Interaction Application

Taigun Lee, Sung-Kee Park
Intelligent Robotics Research Center, Korea Institute of Science and Technology (KIST)
Hawolgok-dong 39-1, Sungbuk-ku, Seoul 136-791, Korea
dungding@kist.re.kr , skee@kist.re.kr

Mignon Park
Department of Electrical and Electronic Engineering, Yonsei University
134 Sinchon-dong, Seodaemun-gu, Seoul 120-749, Korea
mignpark@yonsei.ac.kr

Abstract

We propose an effective facial features detection method for human–robot interaction (HRI) with indoor mobile robot. In case of human-robot interaction at mobile robot, its vision system has to cope with difficult problems such as pose variations, illumination changes, and complex backgrounds, which problems mainly arise from the movement of mobile robot. In this paper, in order to overcome such problems, we suggest a new facial feature detection approach based on local image region and direct pixel-intensity distributions. For this, we propose two novel concepts; the directional template for evaluating intensity distributions and the edge-like blob map image with multiple strength intensity. Using this blob map image, we show that the locations of major facial features – two eyes and a mouth – can be reliably estimated. Without boundary information of facial area, final candidate face region is determined by both obtained locations of facial features and weighted correlations with stored facial templates. Our approach is also flexible that can be applicable to both color and gray image. In case of color image, all detection tasks of both facial features and face are rapidly achieved by using the chromatic property of facial color. Experimental results from various color images and well-known gray level face database images show the usefulness of proposed method in HRI applications.

1 Introduction

The intelligent service robot has emerged as one of the recent trends in the research of robotics. One of the important functions of this service robot is the realization of the Human-Robot Interaction (HRI). The facial image interface in HRI (Fig.1), that uses a camera, requires only the minimal amount of user cooperation, and is capable of acquiring sufficient in-formation to interact with people. For this, human detection and their (user's) face detection play a vital role. In this case, first step of any face processing processes is detecting the locations in images where faces are present [8], and precise estimation method of facial features' locations are also needed. This paper proposes a new detection method of face location and the facial features estimation methods that are suitable for HRI.

Previous face detection researches [1, 3, 6, 9] mostly use a fixed camera; however, the face detection in HRI has some difficulties in its embodiment. First, the face in the acquired image has significant pose variation due to the robot platform's mobility. Second, the user's face image for HRI includes a complex background and a significant amount of illumination changes. Third, when the facial interface for HRI needs real time processing, the detection must be done in near real-time in order to locate the user instantaneously.

To overcome such constraints, we propose a new method of face and feature detection. First, we will propose a novel directional template for effective estimation of the locations of facial features such as the two eyes and mouth. This template will be applied to the input image to propose a new edge-like blob map with multiple intensity strengths. This blob map is robust in both facial pose and some illumination changes relatively, and it is also capable of estimating detailed locations of the facial features even without rough information on the facial area. Another merit of this approach is that it is independent to whether input image is color or gray image. The effectiveness of the proposed method is shown in the experiments using the well-known gray-level face databases and various color facial images obtained from camera of indoor mobile robot.

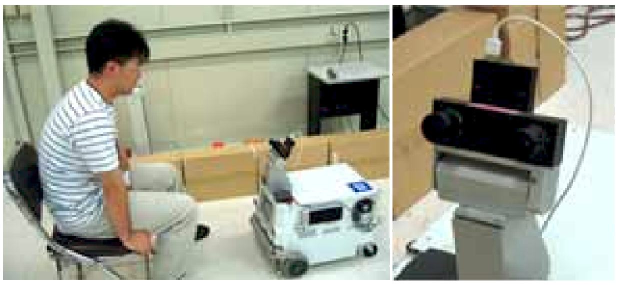


Fig. 1 Application of Human-robot interaction; face detection system with indoor mobile robot

2 Proposed Method

In this section, we present our method for detecting face and its feature locations. The overview of proposed algorithm is shown in Fig.2. One of the characteristics of our method is that this algorithm can be adapted to gray image as well as color image inputs. According to the image type, additional step for preprocessing facial image is included so that facial features can be detected more effectively.

2.1 Preprocessing for detecting facial features

We present a novel approach that uses gray intensity of facial features, irrespective of their color characteristics. Two eyes have such an intensity property that the eye region is darker than near facial area. Ordinary near-eye region has distinctive shape of intensity. That is, the width

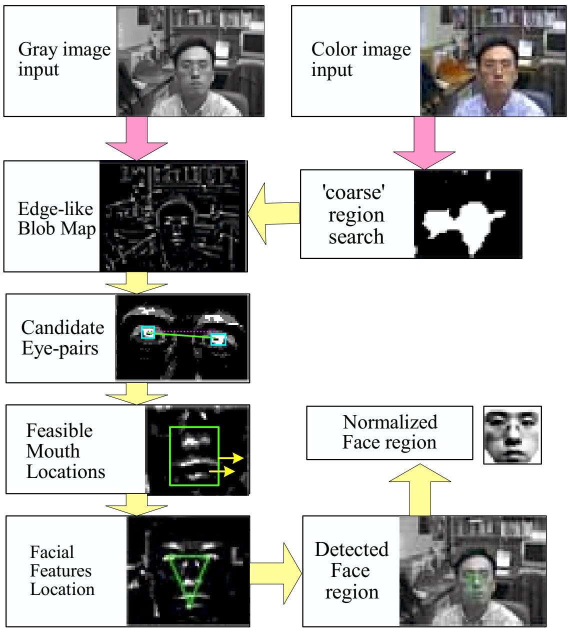


Fig. 2 Summary of face detection and facial feature localization algorithm

of small darker intensity area of eye is longer than the height of area; such shape is just similar to horizontal 'blob' edge. Near-mouth region also has some darker intensity area like near-eye region. Despite the variety of mouth's region size, darker regions of mouth are not generally different from those of eyes. Besides, horizontal blob edge is more distinct around the mouth region. Therefore, considering these intensity characteristics of facial features, we pro-pose the new 'directional blob template' and estimate candidate locations of features. Template size is determined according to facial area size, which is intended as appropriate area to detect in this method, as big as the eye blob region. Width of template is larger than height like as Fig.3.

At the pixel $P(x,y)$ in obtained intensity image (size: $W \times H$), the center pixel $P_{cent}(x_c, y_c)$ is defined as the pixel at which the template is placed. From this pixel P_{cent}, average intensity $\overline{I}_i$ (1) of eight-directions of feature template (size: $h_{FF} \times w_{FF}$) are obtained, and intensity differences ΔI_j , between $\overline{I}_i$ and I_{cent} (intensity value of P_{cent}), are also found as (2). Example of directional template of facial features is shown in Fig.3. The intensity value that has most large magnitude of intensity gradient is defined as principal intensity I_{Pr} (3). Now, using principal intensity value I_{Pr}, edge-like blob map with multiple strength intensity is created as follows. For all locations of pixel $P(x,y)$ in the entire obtained intensity image, the masking operation with above directional template is applied to intensity image. Using a threshold value that is weighted on principal intensity I_{Pr}, multiple intensity strength of each pixel in entire image is determined. For intensity difference ΔI_j of both sides of horizontal direction at a pixel $P(x,y)$; if certain pixel intensity is larger than each weighting thresholds, +1 level intensity strength is assigned, respectively. From this process, the entire gray intensity image is converted into edge-like blob map image that has different 5-level intensity strengths (0, +1, …, +4). Most bright edge-like blob pixels have its intensity level, +4. Each intensity value of strength level is defined as 0, 40, 80, 120, and 200. Fig.4(c) is a negative image of edge-like blob map for clearly showing differences of edge strengths rather than original blob map image, Fig.4 (b).

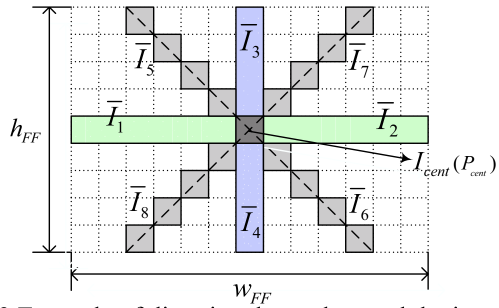


Fig. 3 Example of directional template and the intensity differences of 8-directions from center pixel of template

$\overline{I}_i \; (i = 1, 2, ..., 8)$ ………………………………………… (1)

; *avg. brightness of neighborhood pixels in template for each direction*
(directions : $i = 1$ (*left*), 2 (*right*), 3 (*top*), 4 (*bottom*), 5 (*NW*), 6 (*SE*), 7 (*NE*), 8 (*SW*))

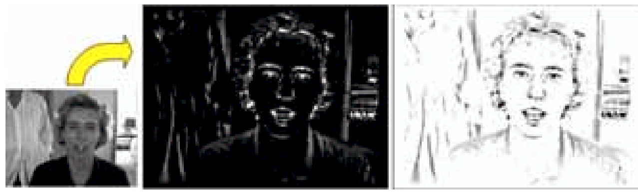
$$Ex) \; \overline{I}_1 = \left\{ \sum_{i = X_C - w_{FF}/2}^{X_C} \sum_{j = Y_C}^{Y_C} P(x - i, y) \Big/ (w_{FF}/2) \right\}$$

$$\Delta I_1 = \mathrm{avg}\,[\, I_{cent} | \; \overline{I}_1, \; \overline{I}_2 \,] \quad (|\, I_{cent} - \overline{I}_1 | + | I_{cent} - \overline{I}_2 |\,)/2 \; \cdots\cdots \; (2)$$

$$\Delta I_2 = \mathrm{avg}\,[\, I_{cent} | \; \overline{I}_3, \; \overline{I}_4 \,], \quad I_3 = \mathrm{avg}\,[\, I_{cent} | \; \overline{I}_5, \; \overline{I}_6 \,],$$

$$\Delta I_4 = \mathrm{avg}\,[\, I_{cent} | \; \overline{I}_7, \; \overline{I}_8 \,]$$

$$I_{\mathrm{Pr}} = \max_{j=1,\ldots,4} \{\Delta I_j\} \quad \cdots\cdots\cdots\cdots \quad (3)$$



(a) (b) (c)

Fig. 4 Edge-like blob map from original gray image and its negative image

2.2 Finding all candidate pairs of eyes and estimation of probable locations of mouth

All From above edge-like blob map, eye analogue blob regions are marked and all probable eye-pair region sets are selected. The eye-like region has more dark intensity property than other feature regions i.e. mouth. So, we choose level +4 edge strength pixels only for candidate eye pixels. All probable eye-pairs are composed from above selected eye blobs, and only candidate eye pairs are selected according to whether facial geometric conditions could be satisfied. Length of eye pair distance, direction of eye-pairs, and ratio of two eye regions' area size are considered as geometric conditions. (Refer to Fig.5)

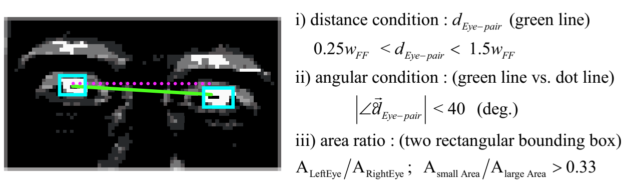


Fig. 5 Facial geometric conditions for Eye-pair blob

For the case that varying pose of face, if facial bounding area is precisely cut off, this area shape is quite a variable form of rectangular area. Due to this reason, estimating the location of mouth as well as eye-pair locations is also needed for detecting precise facial region. The normalization of this precisely selected facial region means a fixed standard face image patch, so this can be effective in increasing the rate of correct face recognition, for future work.

Due to the variety of mouth feature's shape, edge strengths of mouth in edge-like blob map are not sufficiently prominent rather than those of eyes. On the other hand, the horizontal narrow edge is more distinctive around mouth region. Therefore, estimation of mouth location is carried out by obtaining multiple probable locations that have strong and suitable horizontal edges. And several candidate facial regions are decided by both pre-obtained eye pairs and above probable locations of mouth. Summary of the procedures that estimating probable locations of mouth is as follows. (Also refer to Fig. 6)

(Step-1) Find the area range of probable mouth locations, considering of location and direction of eye pair vector.
(Step-2) Rotate this area and make intensity histogram by projection to vertical direction.
(Step-3) Select candidate locations of mouth, if the histogram intensity is higher than predefined threshold.
(Step-4) Choose up to three locations from the bottom (two positions are obtained in Fig.6).

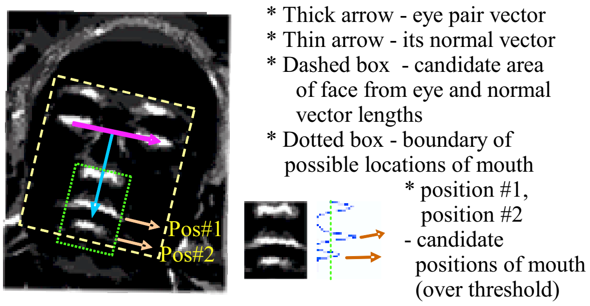


Fig. 6 Estimation of probable mouth locations

2.3 Detection of face region and localization of facial features

From above multiple candidate facial areas that are acquired by eye-pair and mouth location from previous sections, the final location of face and its features can be determined. The similarity measure between candidate area and predefined template of standard face patch is calculated.

At first, normalized form of candidate facial area are created for measuring similarity with predefined template. Variable shape of rectangle of obtained facial area is scaled to fixed square patch of size 60*60 pixels. Similarity measure between candidate facial areas and standard face templates is based on basic form of correlation equations (4).

$$\rho_{I_{FD},I_{tmpl}} = \frac{E[I_{FD} I_{tmpl}] - E[I_{FD}]\ E[I_{tmpl}]}{\sigma_{I_{FD}} \sigma_{I_{tmpl}}} \quad (4)$$

(*where* I_{FD} : obtained facial area, I_{tmpl} : face templates)

3 Experimental results

To present more practical results at various facial images, we have constructed three face image test sets and evaluated our method on these various face image databases. First set is composed at our laboratory, for 10 persons, of pose variant images that view directions of face are naturally random. For second test set, BioID face database [14] is adopted. In recent re-searches [11, 13] of face detection, BioID face database is advantageous for describing more real environments of world. And this is also a head and shoulder images in complex backgrounds, so suits our target task, human-robot interaction. Third set is constructed for our experiment from the World Wide Web- internet news photos [16] and the Carnegie Mellon University (CMU) face images [15]. Some examples of face and features detection on second test set of BioID face database are shown at below Fig. 7.

In order to present comparison results with previous methods, this paper uses relative error measurements in [13]. The error measurement, d_{eye}, is the estimation error between the two eyes' position and their actual positions. This is defined by (5).

$$d_{eye} = \frac{\max(d_l, d_r)}{\|C_l - C_r\|} \quad (5)$$

(C_l, C_r : true eye centers in facial image)

(d_l, d_r : distance between C_l, C_r and estimated eye positions)

Table 1 shows face and its feature detection results on two types of evaluations. One of the evaluation result is 'face area detection ratio' (DR) whose error measurement (d_{eye}=0.25) is adapted as the previous face detection study [13]. This result includes some errors in locating for facial features' positions. The others of result is 'facial feature detection ratio' (FDR), which finds facial features more exactly than above result. This results adopt d_{eye}=0.1 as error measurement. Also, the obtained locations of facial features are shown as a shape of triangle in binary image. An average required time is also suggested at P4-2.8G PC. Table 1 shows the results from the three sets of tests performed using such measurements.

Table 1. Detection results for three test sets

	±	±	±	±

The results show that errors in facial features detection appear in the images that do not clearly show the eyes or mouths. Such cases mostly occur when the eye is not seen to be darker than its surroundings, due to glare from glasses, or when location of mouth area becomes ambiguous due to a beard surrounding it. Additional results of detection performance that excludes such image cases are presented for the Second Test Set. In this case, results show approximately a 97% of successful face detection and 94% rate in facial features detection (Table I). In evaluating the proposed method's detection performance, the previous studies [12,13] from face detection executed on the BioID face database [14] are compared with our method. In Fig.8, distribution function of correct detection rates against relative error measure values for test set #2 (BioID database)

Fig. 7 Some examples of detection results from BioID database

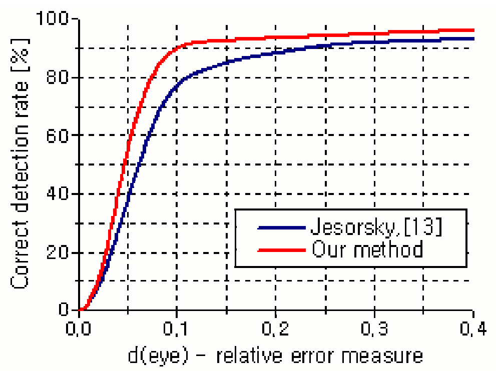


Fig. 8 Distribution function of correct detection rates for BioID database

4 Conclusion

We have presented a facial feature detection method for pose-varying face in complex backgrounds image. Using the intensity characteristics of neighborhood area of main facial features, a novel directional template and an edge-like blob map are proposed. With the various well-known face databases, we also showed that the pro-posed method can achieve better detection perform-ances compared to previous researches. For its characteristics of robustly detecting facial features' loca-tions, our method may be suitable to face recognition and facial expression recognition applications.

Acknowledgments This research was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy of Korea.

References

[1] A.Nikolaidis, I.Pitas, "Facial feature extraction and pose determination", Pattern Recognition, vol.33, pp.1783~1791, 2000.

[2] H.A.Rowley, S.Baluja, and T.Kanade, "Neural Net-work-based Face Detection", IEEE Trans. Pattern Analysis and. Machine Intell., vol.20, no.1, pp.23 ~38, Jan.1998

[3] M.H. Yang, N. Ahuja, and D. Kriegman, "Mixtures of Linear Subspaces for Face Detection," Proc. Fourth Int'l Conf. Automatic Face and Gesture Recognition, pp.70-76, 2000

[4] Anil K. Jain, "Fundamentals of Digital Image Processing", Prentice Hall, 1989

[5] T. Lee, S.-K. Park and M. Kim, "Development of pose-invariant face recognition system for mobile robot ap-plications," Proceedings of Int'l Conf. of Control, Automation and Systems Engineers (ICCAS), pp.783~788, 2003

[6] C. Garcia and G. Tziritas, "Face Detection using Quantized Skin Color Regions Merging and Wavelet Packet Analysis", IEEE Trans. Multimedia, vol.1, no.3, Sep. pp.264-277, 1999

[7] M.H.Yang, D.Kriegman, N.Ahuja, "Detecting face in Images: Survey", IEEE Trans. Pattern Analysis and Machine Intell. vol.24, pp.33-58, Jan.2002

[8] R.L.Hsu, Mohamed Abdel-Mottaleb, Anil K. Jain, "Face detection in Color images", IEEE Trans. Pattern Analysis and. Machine Intell., vol.24, pp.696-706, May 2002

[9] Y. Wang, C.-S. Chua, Y.-K. Ho, "Facial feature detection and face recognition from 2D and 3D images", Pattern Recogni-tion Letters, vol.23, pp.1191~1202, 2002

[10] Y. Gao, M. K.H. Leung, "Face Recognition using Line Edge Map", IEEE Trans. Pattern Analysis and. Machine Intell., Vol.24, No.6, June, pp.764-779, 2002

[11] Z.H.Zhou, X.Geng, "Projection functions for eye detection", Pattern Recognition, vol.37, Issue 5, pp. 1049-1056, May 2004,

[12] B. Fröba and C. Küblbeck, "Robust Face Detection at Video Frame Rate Based on Edge Orientation Features", Proc. 5th Int'l Conf. Automatic Face and Gesture Recognition, pp.327-332, 2002

[13]O.Jesorsky, K.J.Kirchberg, and R.W.Frischholz, "Robust Face detection using the Hausdorff Distance", Proc. 3rd Int'l Conf. on Audio- and Video-based biometric person Authenti-cation, pp.90-95, 2001

[14]The BioID face database;
http://www.bioid.com/downloads/facedb/facedatabase.html

[15]Carnegie Mellon University, Image database, face;
http://vasc.ri.cmu.edu/idb/html/face/index.html

[16] Yahoo! News, News photos;
http://news.yahoo.com/news?tmpl=index2&cid=702.