1st International Workshop on Artificial Life and Robotics **89 - 95**

Human Faces: A Review from Detection to Recognition

Mohamed Rizon[+], M. Karthigayan[+], Sazali Yaacob[+], R. Nagarajan[+], Puteh Saad[*] and P.L. EhKan[*]

[+]*School of Mechatronics Engineering*
[*]*School of Computer and Communication Engineering*
Northern Malaysia University College of Engineering
01000 Kangar, Perlis, Malaysia

Abstract - ***In this paper, we review researches on human faces from aspect of face recognition and face detection. Face emotion recognition are also reviewed. The study of human faces has much assistance on computer vision and machine learning research. We can identify the age, gender, and ethnic origin from faces. Face recognition systems are divided into the holistic approach and the analytic approach. The holistic approach treats a human face as 2D pattern of intensity variation. The analytic approach recognizes a human face using the geometrical measurements taken among facial features, such as eye and mouth. The face emotions are applied in robot to helf patients, senior citizens and disable people. Some time the robots are known as welfare robot. The face emotions are recognized using features such as eye, eye brow and mouth. These features play a vital role in emotion recognition. Face detection is a crucial step for face recognition and previously reported is classified into two groups: algorithms for detecting human faces in intensity images and algorithms for detecting human faces in color images.***

Keywords: face recognition, face detection, and computer vision.

1. INTRODUCTION

The face recognition has many applications such as personal identification, criminal investigation, security work, and login authentication. Automatic recognition of human faces by computer has been approached in two ways: holistic and analytic. The holistic approach [1-6] treats a human face as 2D pattern of intensity variation. The analytic approach [1,7-9] recognizes a human face using the geometrical measurements taken among facial features, such as eye and mouth. Face gesture can be used for the recognition of the face emotion. Incase, if there is any abnormal emotion from the acquired image, it can be indicated by any thing such as robot, alarm etc. The Face Emotion Recognition includes the capturing of the real-time image using digital camera. The inputs are considered for the recognition of emotion. The preprocessing, feature extraction and classification are done simultaneously and a decision on the emotion recognition can be achieved.

Face detection is a crucial step for face recognition. The face regions have to be detected from the images before the face recognition process. The face recognition algorithms previously reported are classified into two groups: algorithms for detecting faces in intensity images [12-15] and algorithms for detecting faces in color images [16-20].

2. FACE RECOGNITION

We first show face recognition systems using the first approach. Chellappa et al. [10] and Samal et al. [11] gave excellent surveys of face recognition research prior to 1995. But, the face recognition area has become very active since 1995. Brunelli et al. [1] used template matching for face recognition. The algorithm prepares a set of four masks representing eyes, nose, mouth and face for each registered person (see Fig.1). To identify the unknown person in the image, the algorithm first detects eyes using template matching and then normalizes position, scale and rotation of the face in the image plane using the detected eye position. Next, for each person in the database, the algorithm places his four masks on their positions relative to eye position and computes the cross-correlation values between the four masks and the image. The unknown person in the image is classified as the person giving the greatest sum of the cross-correlation values of the four masks.

Beymer [2] proposed a template based system for recognizing faces under varying poses. The algorithm produces several 2D views for each registered person by varying the face rotation out of the image plane. In the face recognition phase, the algorithm finds a 2D view that has the best match to the image and classifies the unknown person in image as the person corresponding to this view. Doi et al. [3] proposed a face identification system for automatic lock control. The basis idea of this system is the same as that of [1]. But, Doi et al. proposed a new template matching method which was robust to lighting fluctuation. Sato et al. [4] used a neural network instead of template matching. In the neural network, output units correspond to registered persons and input units correspond to pixels of the input image. Sato et al. trained the neural network using three face templates for each person. In the face recognition phase, the neural network computes an output vector from each test image. And, the unknown person in the image is classified as the person corresponding to the output unit that has the maximum value of the output vector if the maximum value is greater than a threshold value.

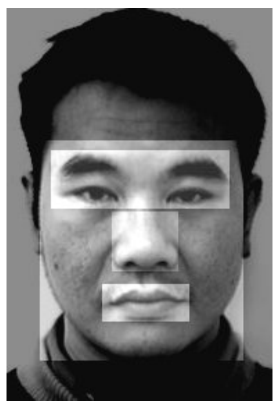

Fig.1. Four masks used in the template matching.

Turk et al. [5] used eigenspace method instead of template matching. This method constructs an eigenspace for each registered person using sample face images. In the face recognition phase, the test image is projected onto the eigenspaces of all registered persons to compute the matching errors. And, the unknown person in the image is classified as the person corresponding to the eigenspace giving the smallest matching error. Kirby et al. [6] proposed a similar method. Turk et al. [5] reported that the eigenspace method was relatively robust to variations in position, scale and pose of the face if the eigenspace of each person was constructed by using face images with different positions, scales and poses. But, if variations in position, scale and pose of the face are not so small, the method requires the normalization of the face. That is, the face has to be normalized to standard position, scale and pose.

Not only eigenspace method but also template matching and neural network-based methods require the face normalization. The face normalization is a crucial step for holistic face recognition systems if the face in the image has unknown position, scale and pose. Next, we show face recognition systems using analytic approach. Although there are many face recognition systems using analytic approach [1,7-9], they are similar if we disregard the methods for detecting facial features. That is, the point of these systems is facial feature detection. So, we show the systems of [1] as an example of analytic face recognition systems because this system has been referred to by many researchers.

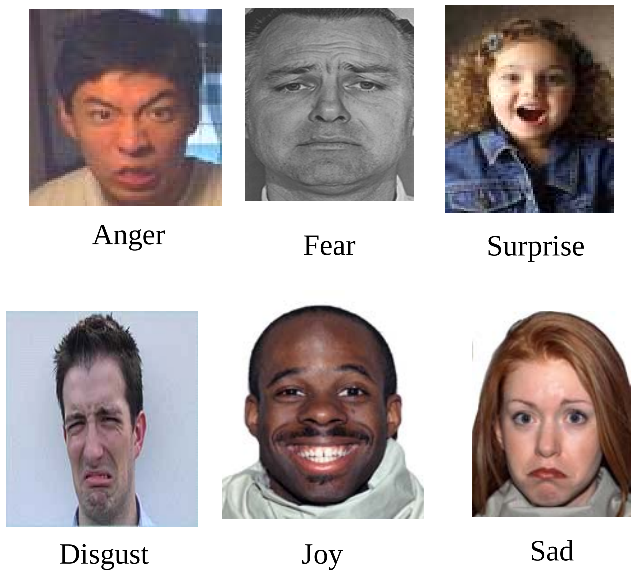

Fig 3. Various Emotions

The results have been found to be satisfactory [27]. Robot-Assisted Activity (RAA) to provide physical and mental support and the emotion driven model for RAA has been proposed [28]. Japan community feels that robots can give more support for senior citizens. More researchers are concentrating on welfare robot for aged people in many countries. The robot has been proposed for the aged people where the robot can feel the emotion of these people and to support them. The construction of robot model uses hierarchical knowledge. The robot can perform complex actions based on the robot's internal and external environmental information. Many robots are involved to support physically and mentally for the senior citizens. These robots work on the basis of face emotion control.

The Face emotion recognition includes the capturing of the real-time image using digital camera. The preprocessing, feature extraction and classification are done simultaneously and a decision on the emotion recognition is achieved. A robot, named as IFBOL, understands a partner's feeling [29]. IFBOL is a Sensibility Technology Communication Robot, can detect emotion, recognize conversation, go by affection of conversation partner and speak with sentiment. Mental health is very important to lead a safe and secured life. This IFBOL robot is able to communicate with people upon their feeling and stress. lFBOL expresses one's feelings with eyelid and eye, besides learning the habit and personality of the conversing partner, so as to respond to the partner.

In preprocessing, some processes such as gray scaling, edge detection and filtering [30] can be performed. These processes are necessary for a good feature extraction and also for removing the unwanted noise in the images. The Gaussian filtering and a two level filtering can be applied to the image [31]. After preprocessing, feature extraction has to be carried out. The features of eye, lips, cheek, eye brows, mouth and skin contraction are extracted. There are several feature extraction methods available in the literature. Among the feature extraction methods, wavelet transform, eigenface, projection profile can be considered [32, 33]. These extracted features play a vital role in identifying the emotion of the face. These features are classified for the emotion classification. The emotions are classified out by using neural network, fuzzy logic and neurofuzzy [34, 35, 36 and 37].

3. FACE DETECTION

Face detection algorithms previously reported are classified into two groups: algorithms for detecting faces in intensity images [12-15] and algorithms for detecting faces in color images [16-20]. [14] used a region-growing method to extract the face region from an intensity image. But, when the intensity difference between background and skin-region of the face is small, it is difficult to correctly extract the face region by the region-growing method.

A neural network-based algorithm [12] requires many sample images for training the neural network. Specially, the number of nonface images needed to train the neural network is huge. In addition, the algorithm can detect only the right front faces. To detect faces with varying poses, we have to train the neural network using sample faces with many different poses. But, as the number of sample face images increases, the discrimination of face images from nonface images becomes difficult.

The eigenspace method also requires many sample images for the construction of the eigenspace. If the number of sample images is not sufficiently large, the method can detect only faces whose patterns are similar to sample faces. In addition, to detect faces with varying poses, we have to construct many eigenspaces, each of which corresponds to one pose. If we construct only one eigenspace for faces with varying poses, the discrimination power of the eigenspace decreases. Thus, the eigenspace method requires so much time to detect faces with varying poses.

Moghaddam et al. [13] used eigenspace method for face detection. This method constructs the eigenspace of the faces using sample face images with the same size. In face recognition phase, sub images with the same sizes as sample images are cut off from the test image and they are projected onto the eigenspace to compute the matching errors. The sub image with the smallest matching error gives the position of the face.

Lin and Wu [14] used a region growing method to detect faces. This method can be used only when the background is considerably darker than skin of the face. Starting from a seed bright region detected inside the face skin, the method grows the seed region into a large connected bright region corresponding to the face skin.

Lam et al. [15] used integral projections of the intensity image and its edge image for face detection. The top boundary of the head was found by horizontal integral projections of the two images. The left and right boundaries of the head were detected by vertical integral projections of the two images.

Faces are significantly characterized by specific skin-color. Thus, the use of color information is useful for face detection. But, the skin has different colors if cameras or lighting conditions vary. Thus, extraction of skin-color pixels is not easy task. Many color spaces such as YCrCb [16,17], rgb [19], HSV (Hue-Saturation-Value) [20] and STV (Saturation-Tint-Value) [21] were used to extract skin-color pixels.

Chai et al. [16] first extracts skin-color pixels using the conditions of 133 Cr 173 and 77 Cb 127 and then applies morphological operations to the regions of skin-color pixels to detect the face region. Nagata et al. [17,18] approximates human heads by upright ellipses and defines a skin region and a hair region inside each ellipse. Generating ellipses, [17] computes a score for each ellipse. The score is given by the sum of three terms. The first term is the density of edge pixels on the boundary of the hair region. The second term is the density of skin-color pixels in the skin region. The third term is the density of dark pixels in the hair region. Heads are chosen from the generated ellipses according to nonincreasing order to their scores.

Yang et al. [19] produces a Gaussian distribution model G(r,g) of skin colors (r,g) using the skin regions selected from sample images. And, color values (r,g) are determined to be skin colors if G(r,g) exceeds a threshold value. After applying morphological operation, [19] selects the largest connected components of skin-color pixels as the face region. Sobottka et al. [20] first extracts skin-color pixels using a Gaussian distribution model of skin colors (H,S) and then selects the elliptical connected components as the candidates for the faces. After that, the face candidates are verified by searching for facial features inside of the connected components.

4. CONCLUSION

Face recognition systems are divided into the holistic approach and the analytic approach. The holistic approach treats a human face as 2D pattern of intensity variation. The analytic approach recognizes a human face using the geometrical measurements taken among facial features, such as eye and mouth. The some review of research in face emotion is done. The face emotion system plays vital role in robot to handle patient, disable people and senior citizens. The main beneficiaries of this work are hospital patients and also the hospital nurses and doctors. Face detection is a crucial step for face recognition and previously reported is classified into two groups: algorithms for detecting human faces in intensity images and algorithms for detecting human faces in color images. Color information is useful to detect the face region. But, we cannot use color information for intensity images. For example, most of surveillance cameras nowadays installed in shops and airports are still intensity cameras. And, the electronic processing of mug-shot or identity card databases requires to detect faces from intensity pictures printed on paper.

REFERENCES

[1] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. 15 (10) (1993) 1042-1052.

[2] D. J. Beymer, Face recognition under varying pose, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp.756-761.

[3] M. Doi, K. Sato, K. Chihara, A robust face identification against lighting fluctuation for lock control, Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, 1998, pp.42-47.

[4] K. Sato, S. Shah, J.K. Aggarwal, Partial face recognition using radial basis function networks, Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, 1998, pp.288-293.

[5] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neuroscience 3 (1) (1991) 71-86.

[6] M. Kirby, L. Sirovich, Application of the Karhunen-Loeve procedure for the characterization of human faces, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 103-108.

[7] Y. Kaya, K. Kobayashi, A basic study on human face recognition, in Frontiers of Pattern Recognition, Academic Press, New York, 1972.

[8] T. Kanade, Computer recognition of human faces, Birkhauser, Basel and Stuttgart, 1977.

[9] S.Y. Lee, Y.K. Ham, R.H. Park, Recognition of human front faces using knowledge-based feature extraction and neuro-fuzzy algorithm, Pattern Recognition 29 (11) (1996) 1863-1876

[10] R. Chellappa, C.L. Wilson, S. Sirohey, Human and machine recognition of faces: a survey, Proceedings of IEEE, 83 (5) (1995) 704-740.

[11] A. Samal, P. Iyengar, Automatic recognition and analysis of human faces and facial expressions: a survey, Pattern Recognition, 25 (1992) 65-67.

[12] H.A. Rowly, S. Baluja, T. Kanade, Neural network-based face detection, IEEE Trans. Pattern Anal. Mach. Intell. 20 (1) (1998) 23-28.

[13] B. Moghaddam, A. Pentland, Probabilistic visual learning for object representation, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 696-710.

[14] C.H. Lin, J.L. Wu, Automatic facial feature extraction by genetic algorithms, IEEE Trans. Image Process. 8 (6) (1999) 834-845.

[15] K.M. Lam, H. Yan, Fast algorithm for locating head boundaries, Journal of Electronic Imaging 3 (4) (1994) 351-359.
[16] D. Chai, K.N. Ngan, Locating facial region of a head-and-shoulders color images, Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition, 1998, pp.124-129.
[17] R. Nagata, T. Kawaguchi, Face detection using skin color model and elliptical head model, Proceedings of the Meeting on Image Recognition and Understanding, Vol.II, 2000, pp.121-126 (in Japanese).
[18] R. Nagata, T. Kawaguchi, Face detection using color, intensity and edge information, Proceedings of International Technical Conference on Circuits/ Systems, Computer and Communications, 1999, pp.249-252.
19] J. Yang, A. Waibel, Real-time face tracker, Proceedings of the 3rd IEEE Workshop on Applications of Computer Vision, 1996, pp.142-147.
[20] K. Sobottka, I. Pitas, Extraction of facial regions and features using color and shape information, Proceedings of International Conference on Pattern Recognition, Vol.III, 1996, pp.420-425.
[21] Nicu Sebe, Michael S. Lew, Ira Cohen, Ashutosh Garg and Thomas S. Huang, **Emotion recognition using a Cauchy Naive Bayes classifier.** Proceedings of 16th International Conference on Pattern Recognition. vol.1, 2002. pp. 17 – 20.
[22] Thomas Mc Kenna. Video Surveillance and Human Activity Recognition for Anti Terrorism and Force Protection. Proceedings of the IEEE conference on Advance Video and Signal based Surveillance. 2003.
[23] Ka Leung Lee and Yangsheng Yu, Real-Time Estimation of facial Expression Intensity. Proceedings of the IEEE International conference on Robotics and Automation. Taiwan. 14-19 Sept 2003: pp-2567-2572.
[24] Katoh. A and Fukui. Y, Classification of facial Expressions using Self-organizing Maps. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 29 Oct.-1 Nov. 1998, vol.2, pp. 986 - 989
[25] Raouzaiou A, Karpouzis K and Kollias S, Emotion Representation for online Gaming. Proceedings of International Conference on Multimedia and Expo. ICME ’03, vol.1, 6-9 July 2003, pp. 417-20.
[26] Cohen I, Sebe N, Garg, A.; Lew, M.S.; Huang, T.S, Facial Expression Recognition from Video Sequences. Proceedings of IEEE International Conference on Multimedia and Expo. ICME ’02. 26-29 Aug 2002, vol.2. pp.121 – 124.
[27] Fuji Ren, Matsumoto k, Mitsuyoshi S, Kuroiwa S and Gai Lin, **Researches on the Emotion Measurement System.** IEEE International Conference on Systems, Man and Cybernetics, Vol 2 , 5-8 Oct 2003, pp. 1666 – 1672.
[28] Tomomi Hashimoto, Toshimitsu Hamada and Toshiko Alolzawa. 2003. Proposal of Emotion-driven Control Model for Robot-Assisted Activity. Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation. July 16-20. Kobe, Japan: 125-129.
[29] Kenji Tamural, Shohei kat, Takanori Yamakita, Kenji Kimurad and Hidenori Itoh, A System of Dialogical Mental Health Care with Sensibility Technology Communication Robot. International Symposium on Micromechatronics and Human Science, 2003, pp. 67-71.
[30] Rafael C. Gonzalez and Richard E. Woods. Digital Image Processing: Prentice-Hall Inc, USA, 2002, pp. 672-674.
[31] De Silva L.C. and Suen Chun Hui. Real-Time Facial Feature Extraction and Emotion. Proceedings of the 2003 Joint Conference of the Fourth International Recognition. Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia, vol.3, 15-18 Dec 2003, pp. 1310 – 1314.
[32] Liu. Y, Schmidt. K. L, Cohn. J. F and We aver R. L. Facial Asymmetry Qualification for Expression Invariant Human Identification. Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR’02). 20-21 May 2002, pp. 198-204

[33] Rizk M. R. M and Taha A, Analysis of Neural Network for Face Recognition Systems with Feature Extraction to Develop an Eye Localization Based Method. Proceedings of Ninth International Conference on Electronics, Circuits and System, 15-18 September 2002., Vol 3, pp. 847-850
[34] George F. Luger and William A Stubblefield, Artificial Intelligence: Addison-Wesley Longman Inc. 1998
[35] Laurene Fausett. Fundamentals of Neural Network: Prentice-Hall Inc, 1994.
[36] Lotfi A. Zadeh. Fuzzy Logic. Computer. Vol. 21(4), 1998, pp. 83-93.
[37] Stamatios V. Kartalopoulos, Understanding Neural Networks and Fuzzy Logic: IEEE press, 1996.