Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Facial Landmark Detection System using Interest-region Model and Edge Energy Function

MiYoung Nam
Dept.of Computer Engineering and Information
Inha University
Incheon, South Korea
nammiyoung@gmail.com

Zhan Yu
Dept.of Computer Engineering and Information
Inha University
Incheon, South Korea
yshmzhan@hotmail.com

Gi Han Kim
Dept.of Computer Engineering and Information
Inha University
Incheon, South Korea
morninggloomy@gmail.com

Phill Kyu Rhee
Dept.of Computer Engineering and Information
Inha University
Incheon, South Korea
pkrhee@inha.ac.kr

***Abstract*—In this paper, we proposed a new facial landmark-detection system using as edge energy function. The facial landmark-detection system is divided into a learning stage and a detection stage. The learning stage creates an interest-region model, to set up a search region of each landmark, as pre-information necessary for a detection stage and creates a detector for each landmark to detect a landmark in a search region. The detection stage sets up a search region of each landmark in an input image with an interest-region model created in the learning stage. Because a landmark to detect from a system has the characteristics of an edge as both edge of an eye, both edge of a mouth and both edges of eyebrows, we have detected a landmark by applying an edge energy function to the Bayesian discrimination method. We have implemented aforementioned technique by abstracting 800 impassive images from the FERET database and have measured data in which the normalized average error distance is less than 0.1 occupying 98 % of the total data.**

***Keywords*—Energy function, local area, onject recognition**

I. INTRODUCTION

Recently, many techniques for biometric recognition exist on the world. Above all, a face recognition technology is getting a lot of interests due to non-coercive convenience. For accurate face recognition, we extract facial feature exactly. This is one of the factors which are important to improve face recognition rate. The feature extraction method is divided into global approach and local feature extraction based methods. The global approach represents face using statistic method like PCA [1, 2] and local feature extraction based method represents face by extracting feature from face image. For accurate face representation, detecting accurate locations of salient points like edges of eyes, edges of mouths and edges of eyebrows are important issues for a system. The fiducial point detection method is divided into the texture based method and the shape based method. The texture based method models local texture around fiducial points. Shape based methods consider face fiducial points as face shapes and search shapes for new faces. Typical texture based methods include neural-network-based eye-feature detection[3], facial feature points detection using Gabor feature based boosted classifiers[4], two-stage facial point detection using a hierarchy of Gabor wavelet networks[5]. Typical shape-based methods include ASM using statistical methods [6]. A number of approaches combining texture- and shape-based methods have been proposed as well. These are AAM method combining texture with ASM [7] and EBGM method using graph structure [8].

In this paper, we propose a texture and shape based facial landmark detection system using energy function and interest-region model. Facial landmark detection system operates basically on two stages, learning stage and detection stage. Learning stage generate interest-region model and detector. Detector for landmark is generated by Bayesian learning using extracted feature vectors. The Detection stage setup searches region using interest region model and detect landmark using Bayesian discriminant

This research work is organized as follows: Section 2 describes landmark and non landmark modeling using Bayesian method. Section 3 describes facial landmark detection system. Section 4 describes the experimental result of the proposed system. Conclusion is given in Section 5.

II. LANDMARK AND NON LANDMARK MODELING USING BAYESIAN METHOD

System can detect accurate landmark using both landmark and non landmark patterns rather than implementing only landmark pattern because landmark and around landmark patterns are similar to each other [9, 10].

If conditional density function of landmark and non landmark is multi-dimension normal distribution, then the equation can be written as,

$$p(x \mid w_p) = \frac{1}{(2\pi)^{N/2} \mid \Sigma_p \mid^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu_p)^t \Sigma_p^{-1}(x-\mu_p)\right\}$$
$$p(x \mid w_n) = \frac{1}{(2\pi)^{N/2} \mid \Sigma_n \mid^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu_n)^t \Sigma_n^{-1}(x-\mu_n)\right\} \quad (1)$$

Where x is a d-component column vector, μ is the d-component mean vector, Σ is the d-by-d covariance matrix. Equation (1) can be rewritten as equation (2) as,

$$g_p(x) = -\frac{1}{2}(x-\mu_p)^t \Sigma_p^{-1}(x-\mu_p) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_p| + \ln p(\omega_p)$$
$$g_n(x) = -\frac{1}{2}(x-\mu_n)^t \Sigma_n^{-1}(x-\mu_n) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_n| + \ln p(\omega_n) \quad (2)$$

$$g_p(x) = -\frac{1}{2} r_p^2 + \alpha_p$$
$$g_n(x) = -\frac{1}{2} r_n^2 + \alpha_n \quad (3)$$

terms except for first term of discriminant is unrelated with x. So squares of Mahalanobis distance r_p, r_n and α_p, α_n in constant about landmark and non landmark samples represent as:

Because α_p, α_n is unrelated with pattern x, discriminant is related with Mahalanobis distance (r_p, r_n). So discernment of input pattern perform method like (4).

III. FACIAL LANDMARK DETECTION SYSTEM

A. Facial landmark detection system structure

We propose a facial landmark-detection system for accurate face recognition. The facial landmark-detection system is divided into a learning stage and a detection stage. A learning stage creates an interest-region model to set up a search region of each landmark as pre-information necessary for a detection stage and creates a detector for each landmark to detect a landmark in a search region. A detection stage set up a search region of each landmark in an input image with an interest-region model created in the learning stage. Since a landmark to detect from a system has the character of an edge as both edges of an eye, both edges of a mouth and both edges of eyebrows, we detected a landmark by applying an energy function composed of intensity and gradient to the Bayesian discriminant.

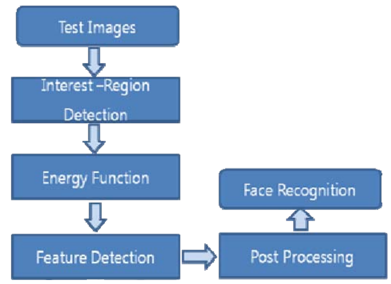

Figure 1. Facial landmark detection system structure

3.2 Learning stage for generating interest-region model and detector

3.2.1 Landmark definition

When a face image is compared with another face image, landmark is a selected location which is distinct from another image. We define locations like edges of eye, edges of nose and edges of eyebrow which are landmarks of a given face image. This feature has more discriminant than others due to high frequency.

Figure2. Landmark location

B. Interest-region model

For facial landmark detection, system search region of eye, nose and mouth. In this paper, for detecting interest-region of landmarks system, make Interest-region and use more accurate landmark detection. We propose interest-region model using fact that eye, nose and mouth exist at a certain distance. Interest-region model is generate by 5 step

Step 1. Make coordinate of landmarks vector.

First Landmark Coordinate x	First Landmark Coordinate y	Second Landmark coordinate x	…….	sixteenth Landmark coordinate x	sixteenth Landmark coordinate y

Figure3. Coordination vector

Step 2. Normalize coordinate vector of landmarks. Center of coordinate set up center of fifth landmark and tenth landmark. Normalize that distance of center and fifth landmark is 1.

Step 3. Interest-region model is average coordinate vector and region vector. Region vector is generated by changeable distance from average coordinate vector of face landmarks and average location of landmark.

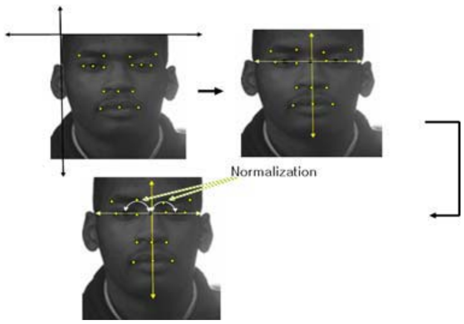

Figure 4. Normalization method of landmark coordination vector

Step 4. Collect landmark and non landmark pattern.

System needs two class models for classifying landmark using Bayesian method. These are landmark model and non landmark model.

In this paper, 40 element vectors for landmark model are extracted by Gabor wavelets having 8 frequencies and 5 orientations about 3×3 pixel. The vector for non landmark model is extracted by separating location 6 pixels apart from landmark to ease detecting landmark in landmark interest-region.

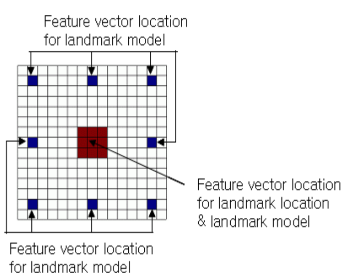

Figure 5. Collect landmark and non landmark pattern

Step 5. Detector generation

In this paper, the landmark detector is generated by Bayesian method. Discriminant feature use Mahalanobis distance formulation because landmark and non landmark patterns show multivariate normal density.

System generate covariance and average vectors using landmark and non landmark patterns from the learning stage Because Mahalanobis distance is measured by ~~using~~ covariance and average vector of learning pattern.

C. *Detection stage using energy function*

1) *Region setting of landmark using interest-region model*

Our system sets up interest-region of landmark using interest-region model for detecting landmark from learning stage. For setting interest-region, it is adjusted by adapting face of input image and moving image accordance axis from center of eyes.

$$M = (x_1, y_1, x_2, y_2, ..., x_{16}, y_{16})$$
$$R = (dx_1, dy_1, dx_2, dy_2, ..., dx_{16}, dy_{16}) \quad (5)$$

M is average coordinate vector. R is region vector.

$$M = M \times dist + (x_{center}, y_{center})$$
$$D = D \times dist \quad (6)$$

Transform of image axis represent like equation (6)

'dist' is distance of center between eyes and pupil of eye, and (x_{center}, y_{center}) is coordinate of center between eyes.

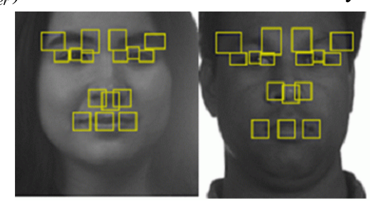

Figure 6. Interest region of facial landmark

2) *Bayesian method using energy function*

a) *Energy function*

Energy function which is used in the general vision systems is used when system detects edge of object[11]. It has internal energy and external energy. Energy function can be written as

$$E = \sum_{i=1}^{n} E_{internal}(i) + E_{external}(i) \quad (7)$$

. The external energy has intensity energy and gradient energy. The internal energy function is intended to enforce a shape on the deformable contour and to maintain a constant distance between the points in the contour. The external energy function attracts the deformable contour to interesting features, such as object boundaries, in an image. In this paper, we use the external energy at Bayesian discriminant. The external energy can be written

$$E_{external} = \alpha E_{intensity} + \beta E_{gradient} \,|$$
$$E_{intensity} = I(x, y) \quad (8)$$
$$E_{gradient} = -|\nabla(I(x, y))|$$

b) *Landmark detection method*

In this paper, defined landmark is located at edges of eye, edges of eyebrow, edges of nose and edges of mouth. Because the less energy magnitude is, the nearer location of edge is. So we use the external energy function at discrimination.

$$g(x,y) = g_p(x,y) - g_n(x,y) - (\alpha I(x,y) - \beta\,|\nabla(I(x,y))|) \qquad (9)$$

$g_p(x,y)$ and $g_n(x,y)$ of Equation(9) is discrimination of landmark and non landmark. After we detect landmark candidates using equation (9) until lank 10, detects landmark using post-processing.

3) Post-processing

We detect landmark candidates using Bayesian method. If detected candidates are near the landmarks, we don't need post-processing. In this paper, to detect accurate landmark, we use a window of a proper size. We search interest-region using this window. If there exists most detected candidates in the window, we consider landmark as average location of candidates.

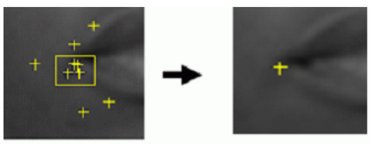

Figure 7. Post-processing

IV. Experiment Results

In this paper, we measure data proportion by normalized mean of error distance. The normalized mean of error distance is measured by dividing the mean of distance between the n detected location and true location into distance between eyes.

$$m = \frac{1}{n d_{eye}} \sum_{i=1} \sqrt{(x_{di} - x_{ti})^2 + (y_{di} - y_{ti})^2} \qquad (10)$$

We use 800 impassive images from FERET Database fa and fb. Data is used by reconstructing face size 128x 128 in experiment.

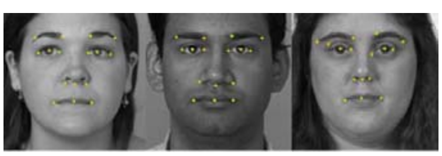

Figure 8. Landmark detection results

A. Edge energy function weight experiment

We select edge energy function weight at Bayesian method through experiment. Landmark detection rate is varied by Intensity and gradient weight of edge energy function. So we need optimized weight of edge energy function. Figure 9 present gradient weight experiment of edge energy function and comparison of the proportion of detected data less than each normalized mean error distance in total data. When weight is 0.3, graph shows least performance. Although difference of performance of weight 0.1 and 0.2 is low, performance is best when weight is 0.2.

Experiment result for selecting optimized weight showed that, when gradient weight is 0.2 and intensity is 0.4, performance is the best.

B. facial landmark detection system performance comparison .

Figure 9 compare Bayesian method using edge energy function, general Bayesian method and Gabor bunch method by landmarks when normalized error distance is 0.8. Figure 12 presents that Bayesian method using edge energy function is the best performance at landmark except for second landmark. When we compare our method with other methods, difference of performances is low at eyebrow region (landmark 1, 2, 6, 7) and eye region (landmark 3, 4, 5, 8, 9, 10), but difference of performances is high at nose region (landmark 11, 12, 13) and mouth region (landmark 14, 15, 16). We consider our method better for nose region and mouth regions.

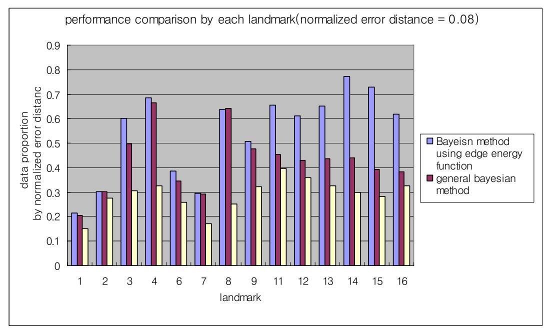

Figure 9. Performance comparison by each landmark

We compare performance among Bayesian method using edge energy function, general Bayesian method, Gabor bunch method. Our method shows superior performance than other methods. We confirm that high performance especially at mouth region.

.

V. Conclusion

In this paper, we propose a facial landmark detection system using interest-region model and energy function for accurate landmark detection. Facial landmark detection system is divided into learning stage and detection stage. In the learning stage, system generates interest-region model and detector. In the detection stage, system sets up interest-region and detected landmark by using energy function at Bayesian method. In the experiment result, we confirm more performance then other systems.

Facial landmarks have different characteristic. We will concentrate our next research for adaptive vector extraction methods and adaptive detection methods at each landmark.

REFERENCES

[1] K.I. Diamantaras and S.Y. Kung. "Principle component neural networks: theory and application," Published by johj Wiley and Sons, 1996.

[2] M. Turk and A. Pentland, "Eigen faces for recognition," Journal of Cognitive Neuroscience, vol. 3, no.1, 1991.

[3] M.J.T. Reinders, et al., "Locating Facial Features in Image Sequences using Neural Networks," Proc. IEEE Int'l Conf. Face and Gesture Recognition, pp230-235, 1996.

[4] Danjjela Vukadinovic and Maja Pantic "Fully Automatic Facial Feature Point Detection Using Gabor Feature Based Boosted Classifiers," IEEE Conference on System, Man and Cybermetics 2005.

[5] R.s. Feris, et al., "Hierarchical Wavelet Networks for Facial Feature Localization," Proc. IEEE Int'l Conf. Face and Gesture Recognition, pp.118-123, 2002.

[6] T.F. Cootes, C.J.Taylor, D.H. Cooper, and J. Graham "Active Shape Models-Their Training and Application," Computer vision and Image understanding Vol. 61, No.1, January, pp.38-59, 1995.

[7] T.F. Cootes, G.J. Edwards and C.J. Taylor "Active Appearance Models," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 23, No. June 2001.

[8] L. Wiskott, et al., "Face recognition by elastic bunch graph matching," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 775-779, 1997.

[9] K. K. Sung and Tomaso Poggio, "Example-Based Learning for View-Based Human Face Detection," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 20, No.1, January 1998.

[10] C. Liu, "A Bayesian Discriminating Features Methods for Face Detection," IEEE Trans. Pattern Anal. mach. Intell., Vol. 25, pp.725-740, 2003.

[11] [11] M. Kass, A. Witkin, and D. Terzopoulos“Snakes: Active contour models,” Int. J. Comput. Vis., vol. 1, pp. 321–331, 1988.