World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:7, 2011

Fast Facial Feature Extraction and Matching with Artificial Face Models

Y. H. Tsai and Y. W. Chen

***Abstract*—**Facial features are frequently used to represent local properties of a human face image in computer vision applications. In this paper, we present a fast algorithm that can extract the facial features online such that they can give a satisfying representation of a face image. It includes one step for a coarse detection of each facial feature by AdaBoost and another one to increase the accuracy of the found points by Active Shape Models (ASM) in the regions of interest. The resulted facial features are evaluated by matching with artificial face models in the applications of physiognomy. The distance measure between the features and those in the fate models from the database is carried out by means of the Hausdorff distance. In the experiment, the proposed method shows the efficient performance in facial feature extractions and online system of physiognomy.

***Keywords*—**Facial feature extraction, AdaBoost, Active shape model, Hausdorff distance

I. Introduction

FACIAL feature extraction is becoming a key technology and challenge for many applications such as authentication, video surveillance, video conferencing, and intelligent robot. It is important that how to extract features from human faces in images or videos automatically. It requires fast algorithms for the needs of time limitations and the large amount of input data. As a result, the performance of facial feature extraction heavily influences the entire systems. Locating the precise facial features also plays an important role for other applications like eye tracking, 3D face reconstruction, facial expression analysis, and intelligent man-machine interfaces [1]. However, due to many kinds of variations such as colors, sizes, orientations, and lighting conditions, facial feature extraction is still a very challenging problem. For a real-time vision system, it has to deal with data streams in the range of several MBytes/s. How to Speed up the extraction and classification process is therefore of major concern when developing systems for real applications. They all require much more research effort to improve the current performance.Most of the algorithms existing in the literature can be divided into four groups: appearance-based, geometry-based knowledge-based, 3D vision-based. Matas et al. [2] proposed fiducial face points by PCA in the neighborhood of the Harris corner detection results. Hamouz et al. provided better results by replacing PCA with the Gabor filters [3]. Kim et al. [4] proposed a facial feature representation based on LDA although it is widely used in face recognition. Jee et al. applied SVM with RBF kernel to verify combinations of eye pairs [5]. Zhu and Ji [6] used SVM with Gaussian kernel to verify eye pairs under the difference of the IR light image and visible light image. Nguyen et al. [7] used SVM to analysis the pixels of eyes and reduced the number of processed pixels to 13% without loss of accuracy. Ryu and Oh performed eyes and mouth detection by NN and eigenvectors [8]. Duffner and Garcia [9] also proposed the architecture for facial feature detection based on 6-layer NN. Viola and Jones [10] presented a face detection algorithm based on AdaBoost classifiers. It combines many weak classifiers into one strong reliable classifier. Milborrow and Nicolls [11] detected 80 face fiducial points by Active Shape Models (ASM). They used one model for a coarse detection and another one to increase the accuracy of the found points. Zheng and Yang created separate model for each facial feature [12]. It is helpful to locate the precise position and shape of each facial feature. Skin color detection was also applied for the initial position of ASM [13].In this paper, we present a fast algorithm that can extract the facial features online such that they can efficiently give a satisfying representation of a face image and have great performance in applications. We apply AdaBoost algorithm to do the coarse detection step for each facial feature and ASM algorithm to increase the overall accuracy of the found points in the regions of interest. For evaluating the resulted facial features, an online matching system was developed to retrieve the most similar artificial face model in the generated database of physiognomy for the corresponding human face. The distance measure between the features and those in the fate models from the database is carried out by means of the Hausdorff distance. In fact, other distance measures can also be applied to the system. In the experiment, 50 persons were invited to test the proposed system and the results showed efficient performance in facial feature extractions and in the online system of physiognomy.The outline of the paper is listed as follows. Section 2 gives a brief overview of the facial feature detection by AdaBoost and ASM. In Section 3, we describe the proposed extraction and matching algorithm. Experimental results are presented in Section 4. Section 5 concludes this paper.

Y. H. Tsai is with the Department of Information Management, Hsuan Chuang University, Hsinchu 300, Taiwan (corresponding author to provide phone: 886-3-5302255 ext. 5340; fax: 886-3-5391252; e-mail: tyh@ hcu.edu.tw).

Y. W. Chen is with the Department of Information Management, Hsuan Chuang University, Hsinchu 300, Taiwan.

II. The Overview of AdaBoost and Active Shape Models

The AdaBoost was first proposed by Freund and Schapire [14]. It can automatically select some weak classifiers from the weak classifier space to construct a strong classifier through the weighted integration of selected weak classifiers. After that, authors proposed new AdaBoost just like Schapire and Singer [15], and Friedman, Hastie and Tibshirani [16]. They was used extensively on face detection [10][17].In this paper, the proposed method is based on the Adaboost algorithm of Viola and Jones [10], which has excellent performance on robust real-time face detection. Toward this end they constructed a frontal face detection system which achieves detection and false positive rates. This system achieves high frame rates working only with the information present in a single grey scale image. These alternative sources of information can also be integrated with their system to achieve even higher frame rates. There are two main contributions in the detection framework. We will introduce the ideas briefly below.The first one is a new image representation called in integral image that allows for very fast feature evaluation. The summed area table (SAT) is shown in Fig. 1. The formula is listed in the following.

$$SAT(x,y) = \sum_{x' \le x, y' \le y} I(x', y'),$$

$$SAT(x,y) = SAT(x, y-1) + SAT(x-1, y)$$

$$+ I(x,y) - SAT(x-1, y-1),$$

$$SAT(-1,y) = SAT(x,-1) = 0,$$

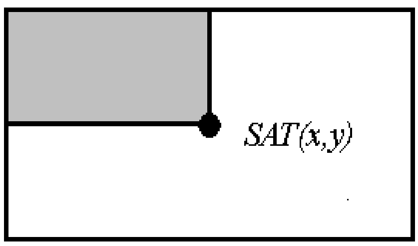


Fig. 1 The summed area table

The second one is a simple and efficient classifier that is built by selecting a small number of important features from a huge library of potential features using AdaBoost. Within any image sub-window the total number of Haar-like features is very large, far larger than the number of pixels. In order to ensure fast classification, the learning process must exclude a large majority of the available features, and focus on a small set of critical features. As a result each stage of the boosting process, which selects a new weak classifier, can be viewed as a feature selection process. AdaBoost provides an effective learning algorithm and strong bounds on generalization performance. The algorithm is listed in the following.

Algorithm

Given the image features $(x_1, y_1), (x_2, y_2)\ldots, (x_n, y_n)$.

Each y_i= 0,1 indicates the negative and positive pattern.

Initially the weighting is set by $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$.

The amounts of the negative and positive patterns are expressed by m and l.

For t = 1, 2, …, T

Normalize the weight

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j}^{n} = w_{t,j}},$$

Calculate the weighting to select weak classifier

$$\in_t = \min_{f,p,\theta} \sum_i w_i \left| h(x_i, f, p, \theta) - y_i \right|.$$

Define $h_t(x) = h(x, f_t, p_t, \theta_t)$, when the f_t, p_t and θ_t are minimum of the $\in_t$.

The way to update the weighting is $w_{t+1,i} = w_{t,i}\beta_t^{1-e_i}$, when the x_i is correct, $e_i = 0$, otherwise is $e_i = 1$, then $\beta_t = \frac{\in_t}{1-\in_t}$.

At last, the decision of the strong classifier is obtained by

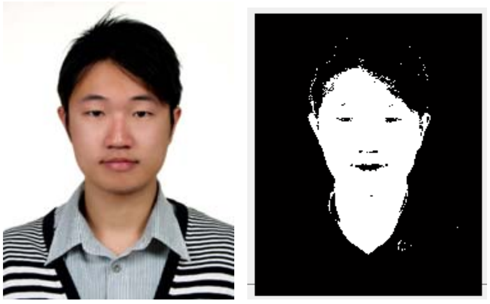
$$C(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \ge \frac{1}{2}\sum_{t=1}^{T} \alpha_t \\ 0 & otherwise \end{cases}$$

,where $\alpha_t = \log \frac{1}{\beta_t}$.

In ASM [11], each sample pattern contour is first marked to identify fiducial points. These points were subsequently normalized by rotation, scaling, and translation. Next, PCA was used to correlate points in order to get parameters for controlling object variations. An initial shape was placed on the image and adjusted by the above derived parameters to match strong edges and corners. The above two methods are adapted in the proposed face feature extraction.

III. The Proposed Method

Since skin color is an important apparent property for a face image, skin color detection is performed to get the initial placement of the extraction algorithm. But skin color may differ from race to race, and depend on the lighting conditions. A good modeling for skin color requires choosing an appropriate color space and identifying a cluster associated with skin color in this space. Some color models in the YCbCr space were proposed. Hsu et al. [18] developed a non-linear transformed YCbCr color space and an elliptical skin model in the transformed space. Hsu's Gaussian model can perform well in skin pixel classification. It is also observed that the distribution of Z in the transformed space changes with the small or large luminance. Skin color was also applied to other applications [19]. Fig. 2 shows the input image and the detected face region by skin color detection.



(a) (b)

Fig. 2 (a) The original image (b) the detected face region

For an image, features of eyes, nose and mouth are usually significant in a face image. Among the above features, we should generate a more stable one under different illumination conditions. In this section, we apply some image processing techniques in [20] to find the features. Sharpening and average are used to enhance the difference between face features and the other area in an input image. Morphological operations are also used to locate the area of each face feature. They are performed by combining the gray values of pixels in a small neighborhood. A mask is defined with the size of the neighborhood. The elementary combination of those pixels in the neighborhood is given by an operation which multiplies each pixel, $g(m, n)$, in the range of the mask with the corresponding weighting factor of the mask, adds up the products, and writes the sum to the position of the center pixel, $g'(m, n)$.

$$g'(m,n) = \sum_{m'=-r}^{r} \sum_{n'=-s}^{s} h(m',n') g(m-m', n-n')$$

The size of the mask for image sharpening is 3×3, i.e., $r=s=1$. $h(0,0)$ is set to be 6. $h(0,-1)$, $h(0,1)$, $h(-1,0)$, and $h(1,0)$ are set to be -1 and the others are set to be 0. If the value of $g'(m, n)$ is larger than 255, it is replaced by 255. After the image sharpening, a smoothing operation is used to get a more uniform distribution of pixel values by averaging the pixels in the 3×3 neighboring area. The pair of sharpening and smoothing operations is applied two times in the preprocessing step to enhance the area of the input image except the area of the number digits. After the preprocessing, we get a new image I_T. An example of I_T is shown in Fig. 3.

Closing operation by a structuring element with size of 5×5 is performed to fill the intensity valley in I_T. Applying different operation on I_T and the closing result, the facial features remain on the resulting image. Opening operation with size of 3×3 is used to remove noise on the resulting image. Note that the size of the structuring elements determines the number and size of the face feature segments. An analysis was done over 1000 images found randomly on the WEB.

Then, a connected component labeling process is applied to locate the facial feature blob to be the candidates. For an input image, each candidate is denoted by E_i, i=1, 2, …, Ne, where Ne is the number of total candidates. Each E_i records the center coordinates (x, y) of the corresponding facial feature blob. The facial feature blob is represented with black pixels in Fig. 3. In most cases, the facial feature blob is not of the same size with its corresponding facial feature, but we get an approximation of the location of the features.

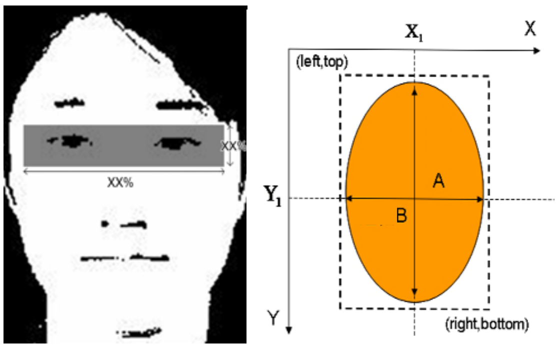
Fig. 3 The detected face feature segments

For getting the accurate Haar-like features in the face image, each feature is cropped without the background and resized. Besides, the negative example of the non-face patterns is captured from the natural scene. Their size is the same with the face image. As usual, the number of the negative examples is more than that of the positive examples.

The final detector is scanned across the testing image at multiple scales and locations. The scaling process is achieved by scaling the detector itself, rather than scaling the image. This process makes sense because the features can be evaluated at any scale with the same cost. Good detection results were obtained using scales which are a factor of 1.25 apart. The detector is also scanned across different locations. Different locations are obtained by shifting the window some number of pixels Δ along the rows sequentially. This shifting process is also affected by the scale of the detector. If the current scale is s, the window is shifted by $[s\Delta]$, where [ ] represents the rounding operation.

The choice of the number of pixels Δ affects both the speed of the detector as well as the accuracy. Since the training images may have some translational variability in the cropping process, the learned AdaBoost detector can also achieve good detection performance in spite of small shifting in the image. As a result, the detection window can be shifted more than one pixel tends to decrease the detection rate slightly while also decreasing the number of false positives for the overall performance in the online system. For getting the accurate Haar-like features in the human face, we have to do the same preprocessing just like the previous one for extracting training features. At last, all the candidates are measured by the AdaBoost classification. Fig. 4 shows an example of the detected eyes pattern from AdaBoost classification and the total face region applied for the proposed algorithm.

Furthermore, ASM is adapted to increase the precision of the position and shape of each facial feature. Some initial fiducial points are set for each facial feature. These points were subsequently normalized by rotation, scaling, and translation. Next, PCA was used to correlate points in order to get parameters for controlling object variations. At last, an initial shape was placed on the image and adjusted by the above derived parameters to match strong edges and corners. The generated facial features include eyes, nose and mouth.



(a) (b)

Fig. 4 (a) the detected eyes pattern (b) the face region for the AdaBoost detection algorithm

The resulted facial features are evaluated by matching with artificial face models in the applications of physiognomy. Hence, a matching system is developed by Hausdorff distance [21]. The Hausdorff distance was originally proposed for binary image comparison. Unlike most shape comparison methods, the Hausdorff distance can be calculated without the explicit pairing of points in their respective data sets, and can be extended to allow partial matching. It is because of this desirable property that we use it to measure the proposed algorithm. Given two finite points sets $A=\{a_1, a_2, \dots, a_m\}$ and $B=\{b_1, b_2, \dots, b_m\}$, the Hausdorff Distance is defined as

$$H(A,B) = \max\{h(A,B), h(B,A)\},$$

where

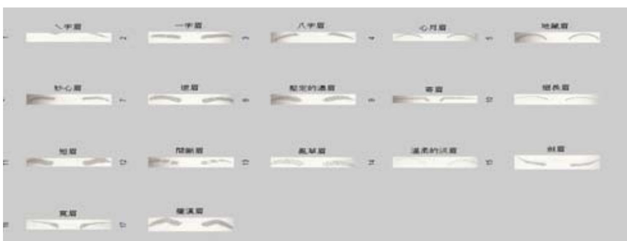
$$h(A,B) = \max_{a\in A} \min_{b\in B} \|a-b\|$$

is called the directed Hausdorff distance from A to B.

The feature vector of a human face consists of 21 feature values. Length, width and angle to the x-axis are calculated from eyebrow, eyes, nose and mouth. Furthermore, we divide the features of eyebrow, eyes and mouth into left and right part to describe detailed information of the human face.

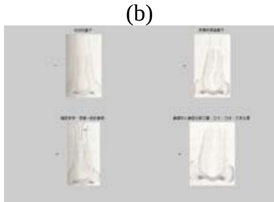
IV. Experimentations

A testing image database containing images from 50 persons was used to test the performance. They were captured under three different illumination conditions. The experimental platform is a PC with an Intel P4 2.8GHz CPU. The artificial face models in the applications of physiognomy are illustrated in Fig. 5. For more detailed information about the face models in physiognomy, please visit the website [22].

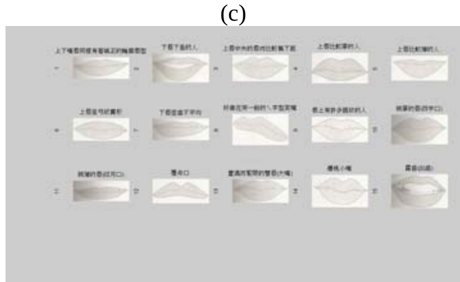


(a)

(b)



(c)



(d)

Fig. 5 Face features of the face models in physiognomy

The online fortune telling system based on face feature matching is shown in Fig. 6. Our demo system summarized the patterns in physiognomy according to some books and websites. The system first captures the face image of the user and extracts the corresponding facial features. After the matching process, it will retrieve the most similar artificial face model in the database and report the result of physiognomy for the corresponding human face. For the measuring the performance of the system, each matching result was examine with people's vision. The precision rates for each single face feature are listed in Table I. Since there is much difference between the real face and an animated face, it is hard to improve the precision rate of the system. If the artificial face model can be modified to approach real human faces, the performance of the entire system will be more efficient than the current version.

Fig. 6 Face features of the face models in physiognomy

TABLE I
UNITS FOR MAGNETIC PROPERTIES

	Precision rate
Eyebrow	85%
Eyes	81%
Nose	75%
Mouth	72%
Average	78.25%

V. CONCLUSION

In this paper, we present a fast algorithm that can extract the facial features online such that they can give a satisfying representation of a face image. The resulted facial features are evaluated by matching with artificial face models in the applications of physiognomy. It is believed that the pattern of face features has strong relationship with human characteristics in physiognomy. Using the proposed method, the system can successfully provide online fortune-telling services through efficient face feature extraction and matching algorithm.

REFERENCES

[1] G. P. Campadelli and R. Lanzarotti, "Fiducial point localization in color images of face foregrounds," *Image and Vision Computing*, vol. 22, pp. 863–872, 2004.
[2] J. Matas, P. B´ılek, M. Hamouz, and J. Kittler, "Discriminative regions for human face detection," in *in Proceedings of Asian Conference on Computer Vision*, 2002.
[3] M. Hamouz, J. Kittler, J.-K. Kamarainen, P. Paalanen, H. Kalviainen, and J. Matas, "Feature-based affine-invariant localization of faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 9, pp. 1490–1495, 2005.
[4] T. Kim, H. Kim, W. Hwang, S. Kee, and J. Kittler, "Component-based LDA face descriptor for image retrieval," in *The 13th British Machine Vision Conference*, 2002.
[5] H. Jee, K. Lee, and S. Pan, "Eye and face detection using svm," in *Intelligent Sensors, Sensor Networks and Information Processing Conference*, pp. 577–580, 2004.
[6] Z. Zhu and Q. Ji, "Robust real-time eye detection and tracking under variable lighting conditions and various face orientations," *Comput. Vis. Image Underst.*, vol. 98, no. 1, pp. 124–154, 2005.
[7] M. H. Nguyen, J. Perez, and F. D. la Torre Frade, "Facial feature detection with optimal pixel reduction svms," in *8th IEEE International Conference on Automatic Face and Gesture Recognition*, September 2008.
[8] Y.-S. Ryu and S.-Y. Oh, "Automatic extraction of eye and mouth fields from a face image using eigenfeatures and ensemble networks," *Applied Intelligence*, vol. 17, no. 2, pp. 171–185, 2002.
[9] S. Duffner and C. Garcia, "A connexionist approach for robust and precise facial feature detection in complex scenes," in *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, pp. 316–321, Sept. 2005.
[10] P. Viola, M. J. Jonse, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
[11] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," in *ECCV 08*, pp. IV: 504–513, 2008.
[12] Z.-L. Zheng and F. Yang, "Enhanced active shape model for facial feature localization," in *2008 International Conference on Machine Learning and Cybernetics*, vol. 5, pp. 2841–2845, July 2008.
[13] M. Haj, J. Orozco, J. Gonzalez, and J. Villanueva, "Automatic face and facial features initialization for robust and accurate tracking," in *International Conference on Pattern Recognition*, pp. 1–4, 2008.
[14] Y. Freund and R. E. Schapire, "Experiments with a New Boosting Aalgorithm," in 13*th International Conference on Machine Learning*, pp. 148-156, 1996.
[15] R. E. Schapire and Y. Singer, "Improved Boosting Algorithms Using Confidence-rated Predictions." *Machine Learning*, vol. 37, no. 3, pp. 297-336, 1999.
[16] J. Friedman, T. Hastie, and R. Tibshirani, "Additive Logistic Regression: a Statistical View of Boosting," *The Annals of Statistics*, vol.28, no.2, pp.337-407, 2000.
[17] L. L. Huang, A. Shimizu, "A multi-expert approach for robust face detection," *Pattern Recognition*, vol. 39, pp. 1695-1703, 2006.
[18] R. L. Hsu, M. A. Mottaleb, A.K. Jain, Face detection in color images, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 695– 705, 2002.
[19] A. Cheddad, J. Condell, K. Curran, P. Kevitt, "A skin tone detection algorithm for an adaptive approach to steganography," *Singal Processing*, vol.89, pp. 2465-2478, 2009.
[20] R.C. Gonzalez and R.E. Woods, Digital image processing, Addison-Wesley Publishing Company, Inc., 1992.
[21] D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge, Comparing images using the Hausdorff distance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993.
[22] http://www.y28predictions.com/program/index.php

Yao-Hong Tsai received the M.S. and Ph.D. degrees in information management from the National Taiwan University of Science and Technology, Taipei, Taiwan in 1994 and 1999, respectively. He is a Assistant Professor with the Department of Information Management, Institute of Information Management, Hsuan Chuang University, Hsinchu, Taiwan. His current research interests include image processing, computer vision, and pattern recognition.

Yung-Wei Chen is a student for the M.S. degree in information management from the Hsuan Chuang University, Hsinchu, Taiwan. His current research interests include image processing, computer vision, and pattern recognition.