A New Method of Detecting Human Eyelids by Deformable Templates

Yuwen WU,   Hong LIU,   Hongbin ZHA

National Lab. on Machine Perception, Peking University, 100871, Beijing, China

E-mail: wuyw@cis.pku.edu.cn   liuhong@cis.pku.edu.cn   zha@cis.pku.edu.cn

ABSTRACT

In this paper, a new method of automatically detecting human eyelids using deformable templates is proposed. Deformable templates are often used in eye's feature detection although there are several drawbacks such as high computational complexity, unexpected shrinking of the templates and rotation of the template. Based on the analysis on these problems, an improved template for eyelid edge detection with several new ideas is presented. The problems of unexpected shrinking of templates and the rotations of template are overcome by using a new kind of potential field and accurate estimations of the template parameters. Dynamic threshold selection and eye corner field are adopted for better performance of eyelid edge detection and two kinds of new energy terms are used to control the eyelid template optimization. Experiments show that the eyelid template and the optimization method are effective even for complex changes in facial images, such as eye opening and closing, obvious irises movements and rotations of heads in small angles.

1. INTRODUCTION

Human robot interaction in a friendly way is very important in many intelligent systems especially for service robots working in human life. Service robots should understand the masters' intentions from their behaviors. Human facial expressions are one of the main interactive ways between the robot and masters. Eye feature provides significant and reliable information of human facial expressions when a person has some intentions or emotions to express. Many researchers have been actively pursuing the task of building automatic systems for detecting eye features recently. Deformable template based methods are rather popular for detecting these features which include irises, eye corners and eyelids.

Yuille, Hallinan and Cohen used deformable templates to detect eye features [1]. Their system works well after the appropriate values of all the weights were found. However, even with an elaborate adjustment, it cannot guarantee the eye template get into a good position and size. For example, when the valley energy term or the intensity energy term is overweighed, the template will easily shrink to a point at the darkest part of the iris. When the initial size of the template is much smaller than the size of the eye in the images, there is no chance to recover it to the proper size. The key problem here is how to determine the weights of each energy term. The solution is spending much time to find these appropriate weights. The weights of these energy terms cannot be generalized and they must be adjusted by experiments. The method by Yuille et al has another problem. It used the two center points of the white areas of sclera for adjusting the rotation of the eye template. It will be easily disturbed by noises. When the iris moves one side extremely, only one peak region exists, and there is no solution available to the problem in the paper [1].

In this paper, a new method is proposed to overcome these problems in Yuille's method. The shrinking of template and the rotation of template are controlled by potential fields and accurate estimations of template parameters. Firstly, this method uses two parabolas as eyelid template instead of the eye template proposed by Yuille et al. Secondly, the intensity transformation and the histogram concavity analysis are used to get dynamic thresholds for getting perfect energy field. Thirdly, a new energy field is introduced to localize two corners of eyelid template. Finally, two new energy terms are used to control the eyelid template shape. These two energy terms are constructed by estimations of eyelid template parameters. The shape of eyelid template can be controlled accurately by these two energy terms. Because many other factors will influence the expected result, some constraints must be introduced. The face image sequences are almost completely viewed frontally in our experiments, because when a head is turned to one side,

eyelids cannot be described by only planar curves.

This paper is organized as follows: In Section 2, the approach of finding the position of eye is presented. In Section 3, the new method of improved deformable templates is proposed. Then some experimental results are shown in Section 4. Finally, the conclusions are given in Section 5.

2. LOCALIZING EYE

For detecting the eyelid in a nearly frontal-view color face image, the position of the eye should be localized firstly. There are a lot of approaches that can be used to localize the eye. Here color information is used to localize the eye.

Objects can be recognized using their color by color indexing [2][3] , a technique developed by Swain and Ballard. Human skin colors do not fall randomly in color space, but a cluster in a small region of color space. Human skin information is more affected in intensity than in color. Under a certain lighting condition, a skin-color distribution can be characterized by a multivariate normal distribution in the normalized color space [4] . Because human skin colors have these features, facial color information can be used to localize eyes. After detecting the face region using skin color information, two coarse regions of interest in the upper left and upper right half of face image can be defined to detect eyes. Here, holes in the skin-segmented image and the facial symmetry are used to localize eyes [5]. After position of eyes is localized, center and region of two eyes can be found. Using the eye region, eye image can be got from the facial image. An eye image is shown in Figure 1.
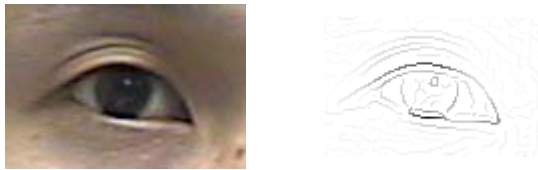

Fig. 1. Eye image and its edge field for computing the potential energy.

3. IMPROVED DEFORMABLE TEMPLATES

3.1 Eyelid template

The eyelid template proposed in this paper differs from the one proposed by Yuille et al. Because detecting eye features is for detecting eye states and eye states can be detected by eyelid states and the iris isn't necessary for

detecting eye states, the circle template is deleted in our template. Without circle template of iris, processing cost less time and accurate result of eyelid detection can be got. After eyelid is detected, the iris can be detected more accurately. Because eye states are got from detecting eyelid, the iris can be detected easily with eye states. Known the eye states, in close eye images and almost close eye images the iris is not detected. In this way, precision of detecting iris can be increased. The eyelid template consists of two parabolas, one for upper eyelid, and the other for lower eyelid. The template is illustrated in Figure 2.
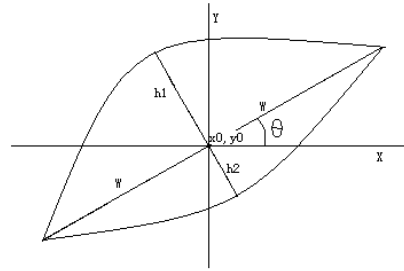

Fig.2. The eyelid template is composed of two parabolas. W is half width of template, h1 is upper height of the template, h2 is lower height of the template, (x0, y0) is center point of the template, θ is orientation of the template.

3.2 Potential fields

The deformable templates are parameterized active model: it interacts with the image in both the geometric aspect and the intensity aspect [6] . The measure of fitness determines the actions of the template. The approach of the measurement is modeled as energy minimization. This need construct appropriate energy function using appropriate potential fields. In our method, first three potential fields are the edge field ϕ_e, the valley-peak field ϕ_{vp}, and the eye corner field ϕ_c. So these three fields are constructed firstly.

The edge field is the edge map of eye. There are many methods to get the edge map. Here, the Canny edge operator [7] is used to get the eye edge map. An eye edge map is shown in Figure 1 before.

For accurately estimating the template's parameters, some works are done on the valley-peak field. First, the eye's binary image is got from its gray image. Here, the threshold is a key problem. For the influence of illuminative condition, color, different persons and the quality of input image, we cannot find a constant as threshold to get binary image of eye gray image. For

overcoming this problem, a method getting dynamic threshold is presented. Some eye images and their histograms are shown in Figure 3. Figure 4. Shows some binary images of these images with different thresholds. These three histograms haven't united form and fixed threshold is not fit for getting appropriate binary image of eye gray image. For automatic processing and getting dynamic threshold, the histograms must have united form or similar form.
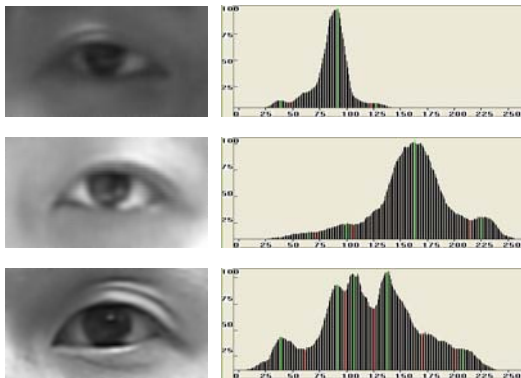

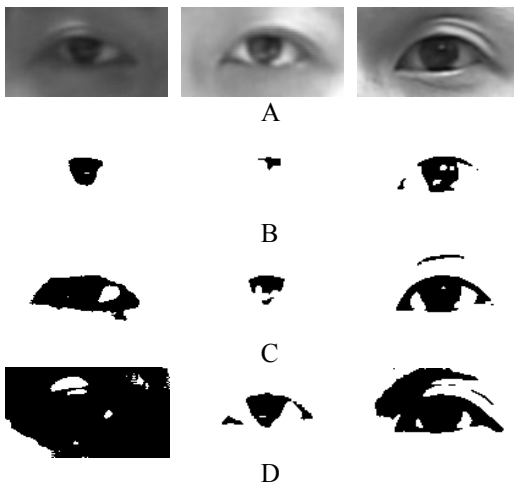**Fig. 3. Some eye images and their histograms.**

Fig. 4.   A: Original images, B: Threshold is 50, C: Threshold is 75, D: Threshold is 100.

The main reason of histogram's difference is that the intensity of image is difference. So the intensity of image is first changed by intensity transformation, which is defined as:

$$s = c \log(1 + |r|)$$

Where c is a scaling constant, and the logarithm function performs the desired compression. Figure 5 shows some eye images transformed by the above formula and their histograms. These eye images are those shown in Figure 4. Their histograms have similar form. Then the histogram concavity analysis[8] is used to select the dynamic threshold. Results of histogram analysis are shown in Figure 6.

Up to now, primary valley-peak field is got. To get final valley-peak field, the region of sclera must be filled with black pixel in final eye valley-peak field. Color information is used again. The YES color space is used, which is defined as:

$$\begin{cases} Y = 0.253R + 0.684G + 0.063B \\ E = 0.5R - 0.5G \\ S = 0.250R + 0.250G - 0.500B \end{cases}$$
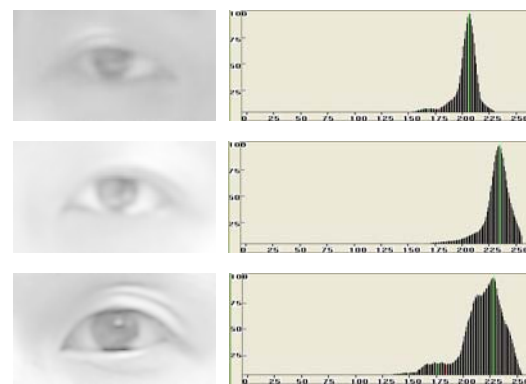

Fig. 5. Eye images by intensity transformation and their histograms.
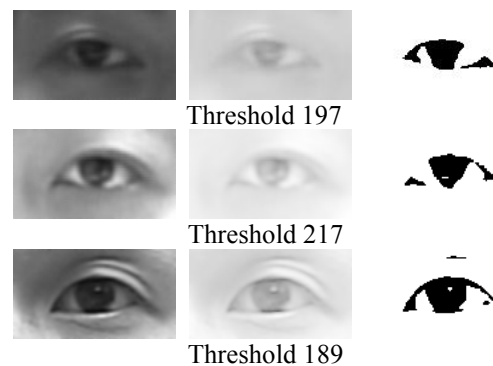

Threshold 197

Threshold 217

Threshold 189

Fig. 6. Column 1: original eye images; Column 2: result images of intensity transformation of column 1 eye images; Column 3: Binary images with threshold got by histogram concavity analysis.

With YES color model, the binary image of E component image of eye color image is used to fill the primary valley-peak field. The binary image of E

component image is got by the histogram concavity analysis, too. Figure 7 shows E component images of some eye color images and their binary images. After filled, the valley-peak field still has some faults. In regions of the iris and the sclera, some holes are not filled; in the other region, some black clips should be erased. These faults are corrected by morphological transformation. Filled image is dilated by binary dilation. And then connectivity is used to erase some small black clip. The entire connective regions are selected. If the area of connect region is less than a constant, this connect region is deleted. In this way, the small black clip can be erased. Finally, binary erosion is used to erase burrs around the remained connective regions. Now the final valley-peak field is got. Some middle results of eye valley-peak fields and final eye valley-peak fields are shown in Figure 8.
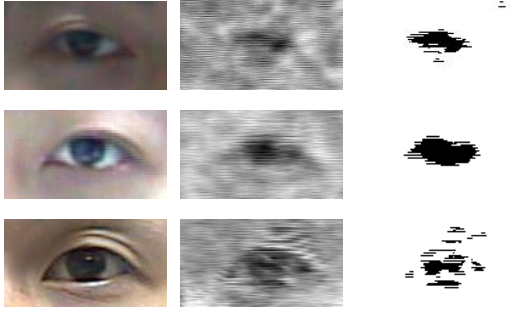

Fig. 7. Column 1：Eye color images; Column 2: E component images of eye color images; Column 3: Binary images of E component images.
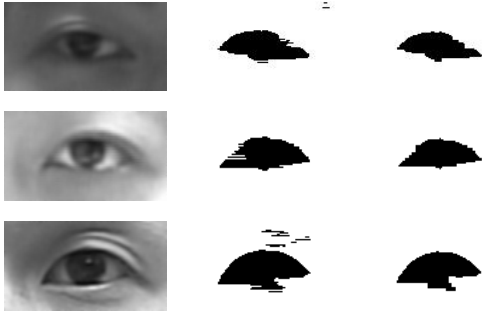

Fig. 8. Column 1: original eye images; Column 2: middle results of eye valley-peak fields; Column 3: final eye valley-peak fields.

Because there is not energy term to control eye corner in former methods of detecting eye features using deformable templates, the eye corners is not accurately detected, especially for eye closed. The eye corners are important eye's features. Accurately localizing the eye inner corner is very important for some other work, such

as facial normalization of recognizing facial expressions. Then the eye corner field is introduced to accurately localize eye corners. Because the region of the eye corner has two intersectant edges, the KLT algorithm[9] can be used to construct the eye corner field. Smaller eigenvalue of the 2×2 coefficient matrix of every pixel is constructed the eye corner field. Introduced this kind of eye corner field overcome the problem of accurately locating eye corner, which is not mention in former detecting eye features by using deformable templates. Some eye corner fields are shown in Figure 9. With the eye corner energy term constructed by this kind of eye corner field, the eye corner can be localized more accurately.
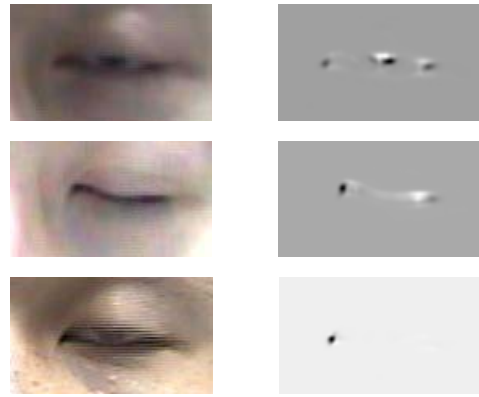

Fig. 9. Eye images and their corner fields.

3.3 Energy function

Our energy function consists of five energy terms. First energy term is the edge energy term. It is defined as:

$$E_{edge} = \frac{1}{\left\| L_{para} \right\|} \int_{L_{para}} \phi_e ds$$

L_{para} is two parabolas of eyelid. This energy term is used to localize the edge of eyelid. The second energy term is valley-peak energy term, which defined as:

$$E_{vp} = \frac{1}{\left\| A_{para} \right\|} \int_{A_{para}} \phi_{vp} dA$$

A_{para} is the region between two parabolas. This energy term is used to localize the position of eyelid template and control the size of eyelid template. The third energy term is the eye corner energy term, defined as:

$$E_{cor} = \frac{1}{\left\| A_{cor} \right\|} \int_{A_{cor}} \phi_c dA$$

A_{cor} is the region of template corner. This energy term is used to localize the position of the eye corner.

In addition, there are other two energy terms, template length energy term and template orientation energy term. For overcoming the problem of the shrinking of template and the rotation of template, two estimations of eyelid template's parameters are got from eye valley-peak field, which shows in Figure 8. One is the half-length of template, W. The other is orientation of the eyelid template, θ. In the valley-peak field, the pixel with minimal abscissa and the pixel with maximal abscissa are found. The distance between two pixels is two times of estimation of W. The estimation of θ comes from the slope of the line passing these two pixels. The template length energy term and the template orientation energy term are defined as:

$$E_w = \frac{W}{W_0} \qquad E_\theta = \left| \theta - \theta_0 \right|$$

Here W is the half-length of template; W_0 is the estimation of W. Similarly, θ is the orientation of the template and θ_0 is the estimation of θ. With these two energy terms, the shape of template can be controlled. The template length energy term can control the shrinking of template, and the template orientation energy term can overcome the problem of the rotation of eyelid template. So the final energy function of eyelid template is defined as:

$$E_T = k_1 E_{edge} + k_2 E_{vp} + k_3 E_{cor} + k_4 E_w + k_5 E_\theta$$

Constants $k_1, ..., k_5$ are the weights of the energy terms. The weights of the template length energy term and the template orientation energy term are larger than the other weights, because the estimations of template parameters are got accurately by using our method.

3.4 Minimization

In Yuille et al's minimization scheme, the iterative template fitting process implements through a series of carefully designed epochs. Our template matching is achieved by applying the downhill simplex method[10] for minimizing the global energy cost. In the minimization scheme, all the weights are of fixed values and each energy term is normalized. It allows each energy term to take action simultaneously during the minimization process and helps the algorithm to converge rapidly. Our minimization scheme is not divided into several epochs and all the parameters of the eyelid template are updated in iteration.

4. EXPERIMENTAL RESULTS

To examine this method, experiments with different color images of different face image sequences were collected. In Figure 10, some experimental results are shown. Testing images are 480×540 facial color images. These images in the same row are the continuous frames of facial image sequence. This method operates continuously at 15 images per second on a 1700 megahertz PC. When the iris moves one side of eye, the Yuille et al's method cannot control the rotation of template. But our method can control the orientation of eyelid template and localize the edge of eyelid accurately by using the estimation of eyelid template's orientation. The row one of Figure 10 shows the results of iris moving to one extreme side. With the template length energy term, the shrinking of template is overcome. Figure 10 shows final templates have not shrink. The illuminance has serious effect on results of deformable templates. In our method, because the intensity transformation and the histogram concavity analysis are used to construct the valley-peak field, the effect of illuminance is decreased. The image sequences showing in Figure 10 are captured in different illuminance. In different illuminate condition, Figure 10 show that this method can get satisfactory results. In Figure 10, we can see that the eye corners are localized accurately, especially the eye closed. The last two rows of Figure 10 shows that the edge of eyelid can be detected accurately, when the rotation of head is in small angles.

5. CONCLUSION

In this paper, an improved deformable template has been proposed for detecting eyelid. The particular feature of this method is constructing appropriate potential fields and finding accurate estimations of eyelid template parameters. Some effective methods are used to construct eye potential fields. Estimations of eye template parameters are got from potential field to construct two new energy terms. Using these two energy terms, some problem of deformable templates can be overcome, such as the shrinking of template and the rotation of template. In addition, a new potential field is introduced, eye corner field. Using eye corner field, the position of eye corners is localized more accurately. Because this method improves the minimization, time cost by this method is reduced. Experiments show that this method is effective for detecting eyelid with complex change in facial image sequences.

REFERENCES

[1]. A. Yuille, P. Hallinan, and D. Cohen, Feature Extraction from Faces Using Deformable Templates, International Journal of Computer Vision, 8(2): 99-111, 1992.

[2]. M. J. Swain and D. H. Ballard, Color indexing, International Journal of Computer Vision, 7(1): 11-32, 1991.

[3]. B. V. Funt and G. D. Finlayson. Color constant color indexing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(5): 522-528, 1995.

[4]. S. A. Shafer, Optical Phenomena in Computer Vision, Proc. Canadian Soc. Computational Studies of Intelligence, 572-577, 1984.

[5]. T. C. Chang, T. S. Huang, C. Novak, Facial feature Extraction from Color Images, Pattern Recognition, 2:39-43 1994.

[6]. Andrew Blake and Michael Isard, Active Contours, Springer Press, 1998

[7]. J. Canny, A Computational Approach to Edge Detection, IEEE Trans. Pattern Analysis Mach. Intell. 8(6): 679-698, 1986.

[8]. A.Rosenfeld, P.Torre. Histogram Concavity Analysis as an Aid in Threshold Selection. IEEE Trans. Systems, Man, and Cybernetics SMC-13: 231--235, 1983.

[9]. S. Birchfield, Derivation of Kanade Lucas Tomasi Tracking Equation, Unpublished, 1997.

[10]. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, Cambridge, 1998.

Fig. 10. Some experimental results.