Projection Functions for Eye Detection

Zhi-Hua Zhou* and Xin Geng

State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China

Abstract

In this paper, the generalized projection function (GPF) is defined. Both the integral projection function (IPF) and the variance projection function (VPF) can be viewed as special cases of GPF. Another special case of GPF, i.e. the hybrid projection function (HPF), is developed through experimentally determining the optimal parameters of GPF. Experiments on three face databases show that IPF, VPF, and HPF are all effective in eye detection. Nevertheless, HPF is better than VPF, while VPF is better than IPF. Moreover, IPF is found to be more effective on occidental than on oriental faces, and VPF is more effective on oriental than on occidental faces. Analysis of the detections shows that this effect may be owed to the shadow of the noses and eyeholes of different races of people.

Keywords: Eye detection; Face detection; Face recognition; Projection function; Race effect

1. Introduction

Building automatic face recognition system has been a hot topic of computer vision and pattern recognition for decades. In general, an automatic face recognition task is accomplished in two steps, i.e. face detection and face recognition. The former determines whether or not there are any faces in the image and, if present, return the image location and extent of each face [1]. The latter identifies or verifies one or more persons in the scene using a stored database of faces [2, 3].

Roughly speaking, face recognition algorithms can be categorized into two classes [4]. In the first class, i.e. geometric, feature-based matching algorithms, salient facial landmarks such as eyes must be detected before any other processing can take place. In the second class, i.e. algorithms based on template matching, faces must be correctly aligned before recognition, which is usually performed based on the detection of eyes. In fact, as both the position of the eyes and the interocular distance are relatively constant for most people, detecting the eyes serves as an important role in face normalization, and thus facilitates further localization of facial landmarks [5]. Therefore, the detection of the eyes is a vital component in automatic face recognition systems.

According to Huang and Wechsler [5], there are two major approaches for eye detection. The first approach, i.e. the *holistic* approach, attempts to locate the eyes using global representations. Representatives are Pentland *et al.*'s modular eigenspaces [6] and Samaria's HMM based algorithm [7]. The second approach, i.e. the *abstractive* approach, extracts and measures discrete local features, and then employs standard pattern recognition techniques to locate the eyes using these features. Representatives are deformable template based algorithms presented by Yuille *et al.* [8] and extended by Lam and Yan [9].

In both holistic and abstractive approaches, after obtaining eye windows, i.e. image regions including eyes, image projection functions can be used to locate eye landmarks that are then used to guide the accurate detection

* Corresponding author. Tel.: +86-25-359-3163; fax: +86-25-330-0710.

E-mail address: zhouzh@nju.edu.cn (Z.-H. Zhou), gengxin@ai.nju.edu.cn (X. Geng).

of the eye position and shape [10, 11]. Among the image projection functions used in this purpose, the integral projection function (IPF) is the most popular one. However, in some cases it cannot well reflect the variation in the image, which is illustrated in Section 3. Although the variance projection function (VPF) is usually more sensitive to the variation in the image than IPF does [10], in some cases it cannot well reflect the variation in the image where IPF works well, which is illustrated in Section 3.

In this paper, the generalized projection function (GPF) is defined, which gracefully combines IPF and VPF. Both IPF and VPF can be viewed as special cases of GPF. Another special case of GPF, i.e. the hybrid projection function (HPF), which inherits both the robustness of IPF and the sensitiveness of VPF, is empirically developed. Experiments on three face databases show that all the special cases of GPF are effective in eye detection. Nevertheless, HPF is better than both IPF and VPF, while VPF is better than IPF. Moreover, it is found that IPF is more effective on the occidental face database than on the oriental face databases, and VPF is more effective on the oriental face databases than on the occidental face database. Analysis of the detections reveals that such a race effect may owe to the shadow caused by the noses and eyeholes of different races of people.

The rest of this paper is organized as follows. In Section 2, the process of eye detection with image projection functions is briefly illustrated. In Section 3, IPF, VPF, and GPF are presented. In Section 4, HPF is experimentally developed, and the performance of IPF, VPF, and HPF in eye detection are experimentally compared and analyzed. Finally in Section 5, the main contributions of this paper are summarized.

2. Detecting eyes with projection functions

In general, all image projection functions can be used to detect the boundary of different image regions. Suppose *PF* is a projection function, ξ is a small constant. If the value of *PF* rapidly changes from z_0 to $(z_0 + \xi)$, then z_0 may lie at the boundary between two homogeneous regions. In detail, given a threshold T, the vertical boundaries in the image can be identified according to:

$$\Theta_v = \max\left\{\left|\frac{\partial PF_v(x)}{\partial x}\right| > T\right\} \tag{1}$$

where Θ_v is the set of vertical critical points, such as $\{(x_1, PF_v(x_1)), (x_2, PF_v(x_2)), \ldots, (x_k, PF_v(x_k))\}$, which vertically divides the image into different regions. It is obvious that the horizontal critical points can be identified similarly. This property of *PF* can be well exploited in eye detection.

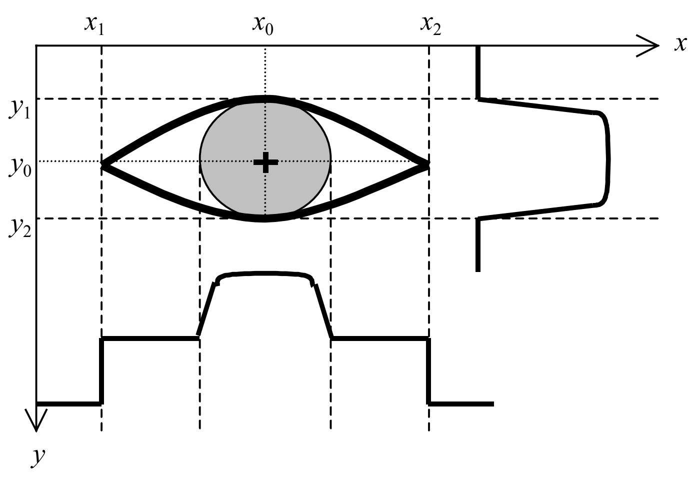


Fig.1 Eye model used in this paper

From the eye model shown in Fig.1, it is clear that in an eye window, the x-coordinate of the eye corners and the y-coordinate of the eyelids are necessary for accurately locating the central point of the eye. Fortunately, the value of PF_v can be expected to rapidly change at x_1 and x_2 because they are the vertical boundaries between different regions, and the value of PF_h can be expected to rapidly change at y_1 and y_2 because they are the horizontal boundaries between different regions. Therefore, after obtaining the values of x_1, x_2, y_1, and y_2, the position of the central point of the eye, i.e. (x_0, y_0), can be determined by:

$$x_0 = \frac{x_1 + x_2}{2} \tag{2}$$

$$y_0 = \frac{y_1 + y_2}{2} \tag{3}$$

An example for illustrating the process of locating the x-coordinate of the eye corners and the y-coordinate of the eyelids by an image projection function is shown in Fig.2, where the dark curve shows the projection function while the gray curve shows the first derivative of the projection function.

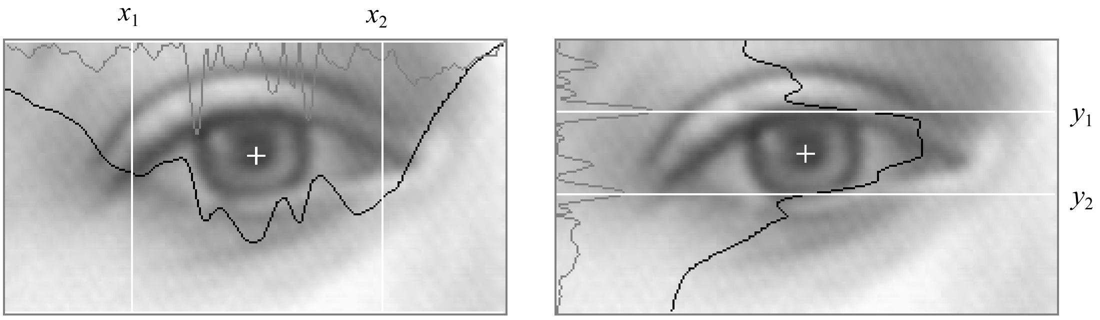


Fig.2 Use projection function to locate the x-coordinate of the eye corners and the y-coordinate of the eyelids

3. Projection functions

3.1. Integral projection function

Suppose $I(x, y)$ is the intensity of a pixel at location (x, y), the vertical integral projection $IPF_v(x)$ and horizontal integral projection $IPF_h(y)$ of $I(x, y)$ in intervals $[y_1, y_2]$ and $[x_1, x_2]$ can be defined respectively as:

$$IPF_v(x) = \int_{y_1}^{y_2} I(x, y)\, dy \tag{4}$$

$$IPF_h(y) = \int_{x_1}^{x_2} I(x, y)\, dx \tag{5}$$

Usually the mean vertical and horizontal projections are used, which can be defined respectively as:

$$IPF_v^{'}(x) = \frac{1}{y_2 - y_1} \int_{y_1}^{y_2} I(x, y)\, dy \tag{6}$$

$$IPF_h^{'}(y) = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} I(x, y)\, dx \tag{7}$$

For convenience of discussion, if without notification, in the rest of this paper we do not distinguish IPF_v from IPF_v*'* and IPF_h from IPF_h*'*.

Although IPF is the most commonly used projection function, there are cases where it cannot well reflect the variation in the image. An example is shown in Fig.3, where IPF_v cannot capture the vertical variation of the image.

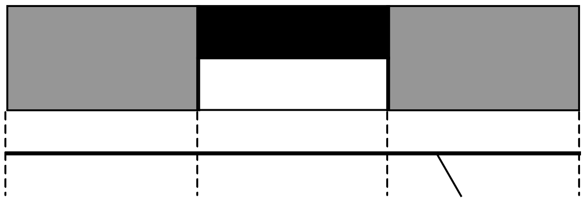


Fig.3 Scenario where IPF_v cannot capture the vertical variation

3.2. Variance projection function

VPF was proposed by Feng and Yuen [10]. Suppose $I(x, y)$ is the intensity of a pixel at location (x, y), the vertical variance projection $VPF_v(x)$ and horizontal variance projection $VPF_h(y)$ of $I(x, y)$ in intervals $[y_1, y_2]$ and $[x_1, x_2]$ can be defined respectively as:

$$VPF_v(x) = \frac{1}{y_2 - y_1} \sum_{y_i = y_1}^{y_2} \left[I(x, y_i) - IPF_v^{'}(x) \right] \tag{8}$$

$$VPF_h(y) = \frac{1}{x_2 - x_1} \sum_{x_i = x_1}^{x_2} \left[I(x_i, y) - IPF_h^{'}(y) \right] \tag{9}$$

Although VPF usually is more sensitive to the variation in the image than IPF does [10], it does not mean that VPF always works well. Fig.4 shows an example where VPF_v cannot well reflect the vertical variation of the image.

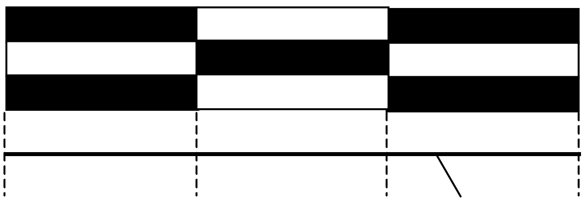


Fig.4 Scenario where VPF_v cannot capture the vertical variation

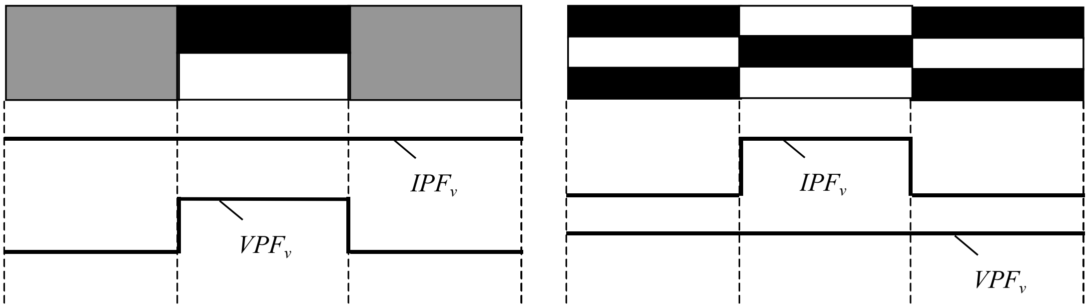
3.3. Generalized projection function

It is easy to find that IPF and VPF can be complementary because IPF considers the mean of intensity while VPF considers the variance of intensity. Such a complementary effect is shown in Fig.5, where VPF works better than IPF on Fig.5 a) but worse on Fig.5 b).

Combining IPF and VPF results in a new projection function, i.e. GPF. Suppose $I(x, y)$ is the intensity of a pixel at location (x, y), the vertical generalized projection $GPF_v(x)$ and horizontal generalized projection $GPF_h(y)$ of $I(x, y)$ in intervals $[y_1, y_2]$ and $[x_1, x_2]$ can be defined respectively as:

$$GPF_v(x) = (1 - \alpha) \cdot IPF_v^{'}(x) + \alpha \cdot VPF_v(x) \tag{10}$$

$$GPF_h(y) = (1 - \alpha) \cdot IPF_h^{'}(y) + \alpha \cdot VPF_h(y) \tag{11}$$



a) VPF_v is better than IPF_v b) IPF_v is better than VPF_v

Fig.5 Complementary effect of IPF_v and VPF_v in capturing the vertical variation

where $0 \leq \alpha \leq 1$ is used to control the relative contribution of IPF and VPF.

It is obvious that both IPF and VPF are special cases of GPF where $\alpha = 0$ or 1, respectively. Other special case of GPF, such as the hybrid projection function (HPF) where $\alpha = 0.6$[1], may work better than IPF and VPF in some cases. For example, HPF can gracefully tackle the problem shown in Fig.5, as shown in Fig.6.

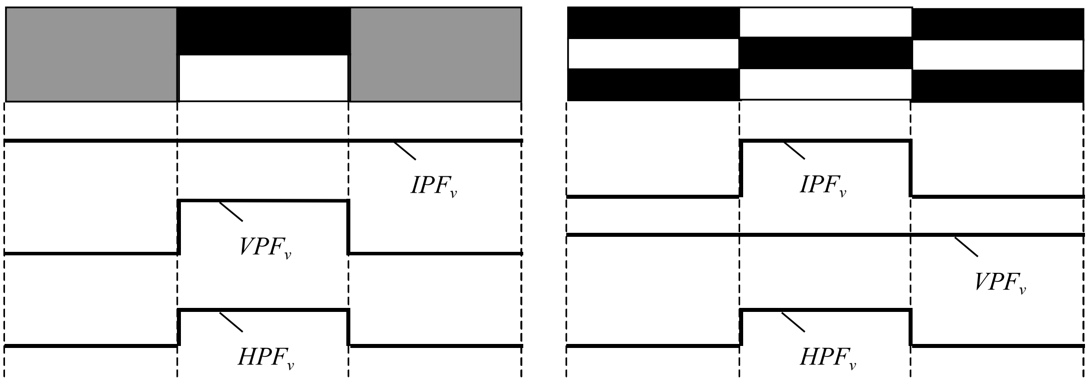


Fig.6 HPF_v works better than both IPF_v and VPF_v in capturing the vertical variation

4. Experiments

4.1. Databases

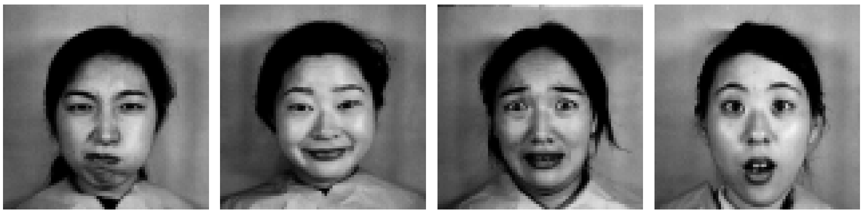
Three databases, i.e. BioID, JAFFE, and NJUFace, are used in our experiments. All images in these databases are with head-and-shoulder faces.

The BioID face database [12] consists of 1521 frontal view gray level images with a resolution of 384×286 pixel. This database features a large variety of illumination and face size, and background of images is very complex. The BioID face database is believed to be more difficult than some commonly used head-and-shoulder face database without complex background, e.g. the extended M2VTS database (XM2VTS) [13]. In [14], when the same detection method and evaluation criteria were applied to both XM2VTS and BioID face databases, the successful detection rates are 98.4% and 91.8% respectively. Some images from BioID are shown in Fig.7 a).

The JAFFE face database [15] consists of 213 frontal view gray level images with a resolution of 256×256 pixel. The images are with a large variety of facial expressions posed by Japanese females. This database has

[1] HPF will be experimentally developed in Section 4.3.

a) Sample images from BioID



b) Sample images from JAFFE

c) Sample images from NJUFace

Fig.7 Some images in the experimental databases

been used in the processing of facial expressions [16]. Some images from JAFFE are shown in Fig.7 b).

The NJUFace database consists of 359 color images, which has been transformed to gray level images with a resolution of 380×285 pixel. The images are with a large variety of illumination, expression, pose, and face size. All the subjects in the images are Chinese. Some images from NJUFace are shown in Fig.7 c).

4.2. Methodology

In our experiments, special cases of GPF are used to accurately detect the central point of the eyes in eye windows. Here the eye windows are obtained through roughly locating the eye positions and then expanding a rectangle area near each rough position. The algorithm used to locate the rough eye positions was proposed by Wu and Zhou [17].

Suppose the rough eye positions are l and r, and the distance between them is d. Then the eye windows are rectangles of the size of $0.8d \times 0.4d$, centered at l and r, respectively, as shown in Fig.8 where the circulars denote the rough eye positions while the crosses denote the accurate central points of the eyes.

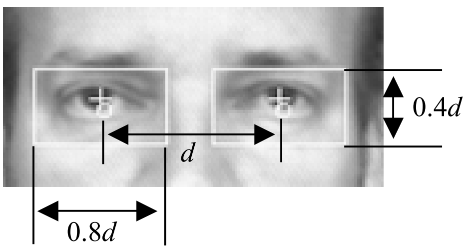


Fig.8 Eye windows used in experiments

The criterion of [14] is used to judge the quality of eye detection, which is a relative error measure based on the distances between the detected and the accurate central points of the eyes. Let C_l and C_r be the manually extracted left and right eye positions, C_l' and C_r' be the detected positions, d_l be the Euclidean distance between C_l and C_l', d_r be the Euclidean distance between C_l' and C_r'. Then the *relative error* of this detection is defined as:

$$err = \frac{\max\left(d_l, d_r\right)}{d_{lr}} \tag{12}$$

If $err < 0.25$, the detection is considered to be correct. Note that $err = 0.25$ means the bigger one of d_l and d_r roughly equals half an eye width. Thus, for a face database comprising N images, the detection rate is defined as:

$$rate = \sum_{\substack{i=1 \\ err_i < 0.25}}^{N} 1/N \times 100\% \tag{13}$$

where err_i is err on the i-th image.

It is worth mentioning that the projection functions compared in our experiments will always return some central points of the eyes because they work in eye windows, and the problem concerned is whether the returned points are correct or not. In other words, it is assumed that when the compared functions are utilized, the eye windows have already been obtained and the goal of the compared functions is to refine the eye location. Therefore, the quality of the detection can be measured as how accurate the central points of the eyes are located with the projection functions. According to Eq.(13), a detection is regarded as an erroneous one if the distance between the detected central point and the manually extracted central point is bigger than half an eye width.

4.3. Hybrid projection function

The eye detection results of GPF on three experimental face databases are tabulated as Table 1, where the value of the parameter α in Eq.(10) and Eq.(11) is increased from 0.0 to 1.0 with 0.1 as the interval.

Table 1. Eye detection rates of GPF on three experimental face databases. Different rows present the detection rates of GPF with different value of α in Eq.(10) and Eq.(11). The 2nd to 4th columns present the detection rates of GPF on different databases. The detection rates are computed according to Eq.(13).

α	BioID	JAFFE	NJUFace
0.0	93.69%	96.71%	92.48%
0.1	93.82%	96.71%	92.76%
0.2	93.82%	97.18%	94.99%
0.3	94.21%	96.71%	95.26%
0.4	94.41%	97.18%	95.26%
0.5	94.41%	97.18%	95.54%
0.6	<u>94.81%</u>	<u>97.18%</u>	<u>95.82%</u>
0.7	91.85%	97.18%	95.54%
0.8	94.54%	97.18%	95.54%
0.9	94.48%	97.18%	95.26%
1.0	94.41%	97.18%	95.54%

It is amazing that when α is set to 0.6, the best detection rates of GPF are obtained on all three databases. This is the reason why HPF is defined as:

$$HPF_v(x) = 0.4 \cdot IPF_v'(x) + 0.6 \cdot VPF_v(x) \tag{14}$$

$$HPF_h(y) = 0.4 \cdot IPF_h'(y) + 0.6 \cdot VPF_h(y) \tag{15}$$

In general, human eye area has two distinct characteristics. The first is that eye area is darker than its neighboring areas, which has been exploited by IPF. The second is that the intensity of eye area rapidly changes, which has been exploited by VPF. It is evident that HPF has exploited both of these characteristics so that it can obtain better performance. It is also worth mentioning that the value of α in HPF, i.e. 0.6, reveals that VPF is slightly more useful in eye detection than IPF does, which supports the claim made by Feng and Yuen [10].

Note that in obtaining HPF, the value of α of GPF is empirically determined to be 0.6. Such a choice is with no theoretical justification. Therefore, although HPF obtains the best performance on BioID, JAFFE, and NJUFace our experimental face databases, it is possible that on other face databases the best performance of GPF may be obtained by setting α to other values. Nevertheless, the success of HPF indicates that the combination of IPF and VPF can be more powerful than sole IPF or sole VPF in eye detection.

4.4. Further exploration

Since IPF, VPF, and HPF are all special cases of GPF, it is easy to get the comparison of their detection rates on the three experimental face databases from Table 1, as shown in Table 2.

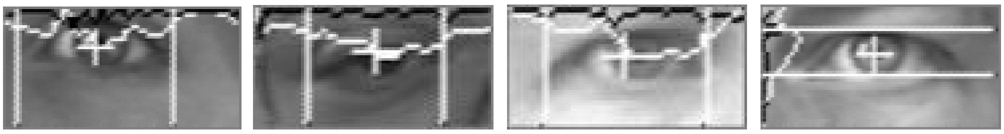
Table 2. Eye detection rates of IPF, VPF, and HPF. Different rows present the detection rates of different projection functions. The 2nd to 4th columns present the detection rates of the projection functions on different databases. The detection rates are computed according to Eq.(13).

func	BioID	JAFFE	NJUFace
IPF	93.69%	96.71%	92.48%
VPF	94.41%	97.18%	95.54%
HPF	94.81%	97.18%	95.82%

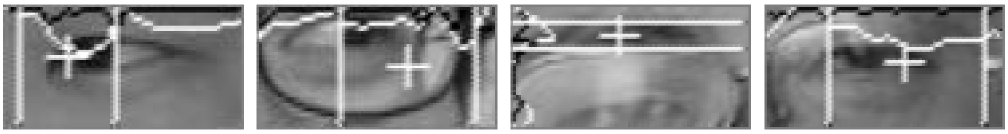
Table 2 shows that IPF, VPF, and HPF are all effective in eye detection because their detection rates on the experimental face databases are all beyond 90%. Nevertheless, the performance of VPF is significantly better than that of IPF on all three databases, while the performance of HPF is significantly better than that of IPF on all three databases and is significantly better than that of VPF on both BioID and NJUFace. This is not strange because in fact HPF is obtained through experimentally determining the optimal parameters of GPF.

It is interesting that Table 2 exhibits that the detection rate of VPF on NJUFace is better than its rate on BioID, but the detection rate of IPF on NJUFace is even worse than its rate on BioID. Since BioID consists of occidental faces while NJUFace consists of oriental faces, such a difference may be caused by some anthropological reasons. Analysis of the detections shows that it may owe to the facial shadow. In detail, occidental faces are often with higher noses and deeper eyeholes so that there is much shadow on the face; oriental faces are often with lower noses and shallower eyeholes so that there is little shadow on the face. Since shadow around eye area may disturb the rapid changes of intensity of this area, occidental eyes may be more difficult than oriental eyes to be detected by VPF. Since shadow around eye area may reduce the mean intensity of this area, occidental eyes may be easier than oriental eyes to be detected by IPF. Some detection results are shown in Fig.9.

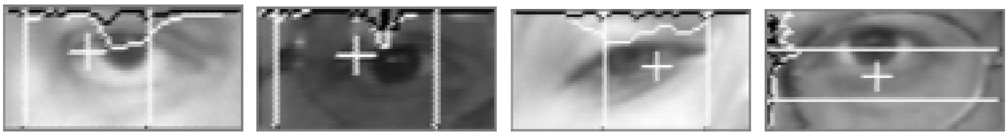
Note that such a race effect has not been exposed on JAFFE. This may because that the faces in JAFFE are with little variety of illumination, pose, and face size so that the eyes are relatively easy to be detected, which is supported by the fact that the detection rates of all the projection functions on JAFFE are better than 96.5%.



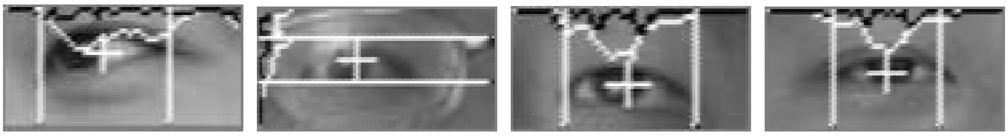
a) Sample eyes in BioID that are correctly detected by IPF



b) Sample eyes in NJUFace that are wrongly detected by IPF



c) Sample eyes in BioID that are wrongly detected by VPF



d) Sample eyes in NJUFace that are correctly detected by VPF

Fig.9 Some detection results on BioID and NJUFace

Moreover, it is worth mentioning that although IPF prefers occidental eyes to oriental eyes while VPF prefers oriental eyes to occidental eyes, the performance of VPF is still better than that of IPF on either occidental or oriental eyes, because in general VPF is more powerful than IPF.

5. Conclusion

In this paper, the generalized projection function, i.e. GPF, is defined. Both the IPF and VPF are special cases of GPF with the parameter α being set to 0 and 1, respectively. Another special case of GPF, i.e. HPF, is empirically developed through setting α to 0.6. HPF exploits two characteristics of eye areas that have been individually exploited by IPF and VPF. That is, eye area is darker than its neighboring areas, and the intensity of eye area rapidly changes. So, its eye detection performance is better than that of IPF and VPF.

Note that although GPF with α being set to 0.6 obtains the best performance on BioID, JAFFE, and NJUFace, it is possible that on other face databases the best performance of GPF may be obtained by setting α to other values. Developing some mechanism to determine the appropriate value of α for concrete tasks is an important issue to be explored in future work. Nevertheless, this paper shows that the combination of IPF and VPF can be more powerful than sole IPF or sole VPF in eye detection.

It is interesting to find that IPF is more effective on occidental faces than on oriental faces, and VPF is more effective on oriental faces than on occidental faces. Analysis of the detections shows that this effect may owe to the shadow caused by the noses and eyeholes of different races of people. This reminds us not to overlook anthropological factors in developing face processing or other kinds of biometrics techniques.

Acknowledgements

The comments and suggestions from the anonymous reviewers greatly improved this paper. The authors wish to thank Michael J. Lyons for providing the JAFFE face database. This work was supported by the National Outstanding Youth Foundation of China under the Grant No. 60325207.

References

[1] M.-H. Yang, D. J. Kriegman, N. Ahuja. Detecting faces in images: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(1) (2002) 34-58.

[2] R. Chellappa, C. L. Wilson, S. Sirohey. Human and machine recognition of faces: a survey. *Proceedings of the IEEE* 83(5) (1995) 705-740.

[3] W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips. Face recognition: a literature survey. Technical Report: UMD CfAR-TR-948, University of Maryland, College Park, MD, 2000.

[4] R. Brunelli, T. Poggio. Face recognition: features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(10) (1993) 1042-1052.

[5] J. Huang, H. Wechsler. Visual routines for eye location using learing and evolution. *IEEE Transactions on Evolutionary Computation* 4(1) (2000) 73-82.

[6] A. Pentland, B. Moghaddam, T. Starner. View-based and modular eigenspaces for face recognition. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, Seattle, WA, 1994, pp.84-91.

[7] F. S. Samaria, A. C. Harter. Parameterization of a stochastic model for human face identification. In: *Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, FL, 1994, pp.138-142.

[8] A. L. Yuille, P. W. Hallinan, D. S. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision* 8(2) (1992) 99-111.

[9] K. M. Lam, H. Yan. Locating and extracting the eye in human face images. *Pattern Recognition* 29(5) (1996) 771-779.

[10] G. C. Feng, P. C. Yuen. Variance projection function and its application to eye detection for human face recognition. *Pattern Recognition Letters* 19(9) (1998) 899-906.

[11] G. C. Feng, P. C. Yuen. Multi-cues eye detection on gray intensity image. *Pattern Recognition* 34(5) (2001) 1033-1046.

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

[13] K. Messer, J. Matas, J. Kittler, J. Luettin, G. Maitre. XM2VTSDB: the extended M2VTS database. In: *Proceedings of the 2nd International Conference on Audio- and Video-based Biometric Person Authentication*, Washington, DC, 1999, pp.72-77.

[14] O. Jesorsky, K. Kirchberg, R. Frischholz. Robust face detection using the Hausdorff distance. In: J. Bigun, F. Smeraldi Eds. *Lecture Notes in Computer Science 2091*, Berlin: Springer, 2001, pp.90-95.

[15] The JAFFE database [http://www.mis.atr.co.jp/~mlyons/jaffe.html].

[16] M. J. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba. Coding facial expressions with Gabor wavelets. In: *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, 1999, pp.200-205.

[17] J. Wu, Z.-H. Zhou. Efficient face candidates selector for face detection. *Pattern Recognition* 36(5) (2003) 1175-1186.

About the Author – Zhi-Hua Zhou received his B.Sc., M.Sc. and Ph.D. degrees in computer science from Nanjing University, China, in 1996, 1998 and 2000, respectively, all with the highest honor. At present he is an associate professor and director of the AI Lab of the Computer Science & Technology Department, Nanjing University, China, and an honorary fellow of the Faculty of Science & Technology, Deakin University, Australia. His current interests are in machine learning, data mining, neural computing, pattern recognition, information

retrieval, and evolutionary computing. In these areas he has published over 40 technical papers in refereed journals or conferences. He has won the Microsoft Fellowship Award (1999) and the National Distinguished Ph.D. Dissertation Award of China (2003). He is an associate editor of the journal *Knowledge and Information Systems* (Springer), editorial board member of the journal *Artificial Intelligence in Medicine* (Elsevier), and reviewer for over ten international journals including several IEEE Transactions. He serves as grant reviewer for the National Natural Science Foundation of China, the Research Grants Council of Hong Kong, and NWO-The Netherlands Organisation for Scientific Research. He chaired the organizing committee of the 7th Chinese Workshop on Machine Learning (2000), and served as program committee member for many international conferences. He is an executive committee member of Chinese Association of Artificial Intelligence (CAAI), the chief secretary of CAAI Machine Learning Council, and a member of IEEE and IEEE Computer Society.

About the Author – Xin Geng received his B.Sc. degree in Computer Science from Nanjing University, China, in 2001. Currently he is a graduate student in the Computer Science & Technology Department, Nanjing University. His research interests include pattern recognition and machine learning.