

Research Article

Gabor Directional Binary Pattern: An Image Descriptor for Gaze Estimation

Hongzhi Ge

School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China

Correspondence should be addressed to Hongzhi Ge, hzge@jdl.ac.cn

Received 27 April 2010; Revised 26 July 2010; Accepted 24 August 2010

Academic Editor: William Sandham

Copyright © 2010 Hongzhi Ge. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes an image descriptor, Gabor Directional Binary Pattern (GDBP), for robust gaze estimation. In GDBP, Gabor magnitude information is extracted firstly from a cropped subimage. The local directional derivations are then utilized to encode the binary patterns in the given orientations. As an image descriptor, GDBP can suppress noises and robustness to illumination variations. Meanwhile, the encoding pattern can emphasize boundary. We use the GDBP features of eye regions and adopt the Support Vector Regression (SVR) to approximate the gaze mapping function, which is then used to predict the gaze direction with respect to the camera coordinate system. In the person-independent experiments, our dataset includes 4089 samples of 11 persons. Experimental results show that the gaze estimation can achieve an accuracy of less than 2° by using the proposed GDBP and SVR.

1. Introduction

In HCI (Human-Computer Interaction) scenario, eye gaze means the pointer from the viewer's two eyes to an object, and gaze is a very useful natural input modality. Combining with the sign language recognition or speech recognition, eye gaze tracking can highly improve usability for the disabled persons, and meanwhile it can be applied in some special fields such as, ophthalmology, neurology, and psychology. Many researchers in computer vision and pattern recognition community have been focusing on this topic, and some methods for gaze estimation can be found in related literature. These methods, by their approaches to represent the position of a pupil's center in the eye socket, are divided into two categories [1]: model-based methods and appearance-based methods. Model-based solutions, such as, Pirkinje image [2, 3] and limbus tracking [4], use an explicit geometric eye model and the geometric features to estimate the gaze direction. Appearance-based solutions treat an eye image as a high-dimensional feature, instead of using the explicit geometric characteristics [5]. These appearance-based approaches are usually more robust in experiments by better exploiting the statistical properties. Sugano et al. [1] take the cropped eye region as a point in a

local manifold model and make gaze estimation by clustering learning samples with similar head poses and constructing their local manifold model. In the approach proposed by Lu et al. [6], Local Binary Pattern (LBP) [7] represents the "pupil-glint" vector information related to gaze direction by obtaining the texture changes of eye images. In [8], an appearance-based method, Local Pattern Model (LPM), is presented. This model combines the improved Pixel-Pattern-Based Texture Feature (PPBTF) and LBP texture feature. Although the existing appearance-based methods have made significant progress in gaze estimation, their accuracy and robustness need to be further improved.

In this paper, we present an appearance-based gaze estimation method based on a novel image operator, Gabor Directional Binary Pattern (GDBP), and Support Vector Regression (SVR) [9]. In GDBP, multiscale and multiorientation Gabor wavelets are used to decompose an eye image, followed by the Directional Binary Pattern (DBP) operator. We use the GDBP operator to represent the texture changes of the eye images caused by the pupil centers which keep moving in the eye sockets, when people at a certain head pose gazes in different directions. With the advantages of Gabor filters [10] and the local directional differentiation information, GDBP is not only robust to illumination

variances, but also with much discriminating power. In applications, these patterns are useful in representing the horizontal and vertical pupil movements. As appearance-based features, GDBP is fed into SVR to approximate the gaze mapping function. The output gaze direction is represented in terms of Euler angles with respect to the camera coordinate system. Our experimental results show the validity of the proposed operator, and additionally we have achieved an accuracy of less than 2° .

The rest of the paper is organized as follows. In Section 2, we elaborate the computation of the proposed GDBP operator in detail, as well as some analysis on its robustness to the light variances and its different discriminating power in different orientations. Gaze estimation with fixed front head pose based on GDBP is presented in Section 3, followed by experimental results with comparisons with other approaches. In the last section, some brief conclusions are drawn with some discussions on the further work.

2. GDBP Operator

In this section, we first define the Directional Binary Pattern (DBP), and then extend it to GDBP, using multiscale and multiorientation Gabor filters. Finally, we analyze the robustness and discriminating power of the GDBP. The details are given as follows.

2.1. Directional Binary Pattern (DBP). We define the texture in a 3×3 pixels neighborhood of an image as a joint distribution of the gray levels in the 9 pixels $\{Z_c, Z_0, \dots, Z_7\}$ (see Figure 1). Z_c is the center pixel, around which are Z_0, \dots , and Z_7 .

Our operator is not only a gray scale texture operator, but also an encoding of directional differential patterns, which are given by the directional differential equations:

$$\Delta I(Z_c)_\theta = I(Z_c) - I(Z_i) \quad \left(i = 0, \dots, 7, \theta = \frac{i\pi}{4} \right), \quad (1)$$

where θ is the direction of the directional differentiation and i is the corresponding pixel. In a given direction θ , the directional differentiation information around Z_c is summarized in a Directional Binary Pattern (DBP):

$$\text{DBP}_\theta(Z_c) = \{\text{sign}(\Delta I(Z_0)_\theta), \dots, \text{sign}(\Delta I(Z_7)_\theta)\}, \quad (2)$$

where

$$\text{sign}(\Delta I(Z_i)_\theta) = \begin{cases} 1 & \text{if } \Delta I(Z_i)_\theta > 0, \\ 0 & \text{if } \Delta I(Z_i)_\theta \leq 0. \end{cases} \quad (i = 0, \dots, 7). \quad (3)$$

As stated above, eight neighbor pixels are used in the differential results of the same direction, which can be formed into a byte to represent 256 different modes. These modes can be easily encoded as follows:

$$\text{DBP}_\theta(Z_c) = \sum_{i=0}^7 \text{sign}(\Delta I(Z_i)_\theta) 2^i, \quad (4)$$

Z_3	Z_2	Z_1
Z_4	Z_c	Z_0
Z_5	Z_6	Z_7

FIGURE 1: A 3×3 pixels neighborhood.

where (2) is used to compute a decimal value for an 8-bit binary string as (4). By using this encoding method, we can get a decimal number for each pixel corresponding to its directional pattern. It ranges from $[0, 255]$, and it is easy to visualize it as a gray-level image. The directional differential pattern is calculated in 0° (the x -coordinate direction) and the formula is the following:

$$\Delta I(Z_c)_0 = I(Z_c) - I(Z_0). \quad (5)$$

Similar to the eight-pixel neighborhood of Z_c as shown in Figure 1, the other directional derivatives are given by the following:

$$\begin{aligned} \Delta I(Z_c)_{\pi/4} &= I(Z_c) - I(Z_1), \\ \Delta I(Z_c)_{\pi/2} &= I(Z_c) - I(Z_2), \\ \Delta I(Z_c)_{3\pi/4} &= I(Z_c) - I(Z_3). \end{aligned} \quad (6)$$

As shown in Figure 1, eight neighboring points around Z_c correspond to eight orientations. However, four directional derivatives of $\theta = \{\pi, 5\pi/4, 3\pi/2, 2\pi\}$ are calculated when the centers are Z_4, Z_5, Z_6 , and Z_7 . The whole DBP consists of DBP_θ s at 4 directions:

$$\text{DBP}(Z_c) = \left\{ \text{DBP}_\theta(Z_c) \mid \theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4} \right\}. \quad (7)$$

Similar to the LBP operator, DBP encodes the local binary pattern. In addition, DBP also represents directional differential information (see Figure 2). With these advantages, DBP can be applied broadly.

2.2. Extending DBP with Gabor Filters. Gabor wavelets with multiscale and multiorientation are widely applied in image processing and pattern recognition. We extend DBP to GDBP to enhance the capability of object representation. Gabor wavelets (kernels, filters) are defined as $\psi_{u,v}(z)$ in [11]. Gabor map is defined as

$$G_{u,v}(Z) = I(Z) * \psi_{u,v}(Z), \quad (8)$$

where $Z = (x, y)$ denotes the image position and “ $*$ ” is the convolution operator. For frequency v in $\{0, 1, 2, 3, 4\}$ and orientation u in $\{0, \dots, 7\}$, the directional differentiation of $G_{u,v}(z)$ in direction θ at location Z_c is computed as

$$\Delta G_{u,v,\theta}(Z_c) = G_{u,v}(Z_c) - G_{u,v}(Z_i) \quad \left(i = 0, 1, 2, 3, \theta = \frac{i\pi}{4} \right). \quad (9)$$

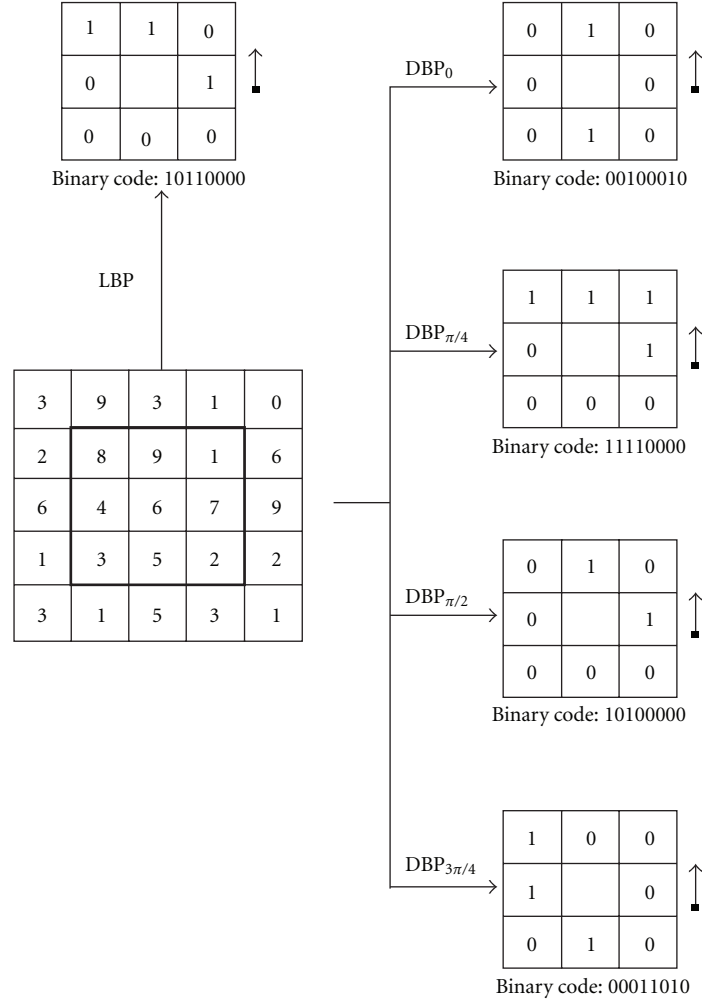


FIGURE 2: A comparison of LBP and DBP. Bottom-left: the input image with the neighborhood of interest highlighted. Top-left: the binary code by LBP. Right-4-rows: the binary codes by DBP.

and $GDBP_{u,v,\theta}$ can be written as

$$GDBP_{u,v,\theta}(Z_c) = \sum_{i=0}^7 \text{sign}(\Delta G_{u,v,\theta}(Z_c)) 2^i. \quad (10)$$

Finally, the whole directional differential patterns of GDBP are formulated as

$$GDBP(Z_c) = \left\{ GDBP_{u,v,\theta}(Z_c) \mid \theta = 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4} \right\}. \quad (11)$$

As stated above, a cropped eye image is encoded into GDBP by the following procedures. (1) The image is normalized and transformed to obtain multiple Gabor magnitude maps in frequency domain by applying multiscale and multiorientation Gabor filters. (2) The Directional Binary Pattern (DBP) is extracted from these maps.

2.3. Robustness Analysis of the GDBP. We analyze the robustness of the proposed GDBP because a good image representation should be robust to lighting variations. To evaluate GDBP's robustness to lighting variations, we compared the

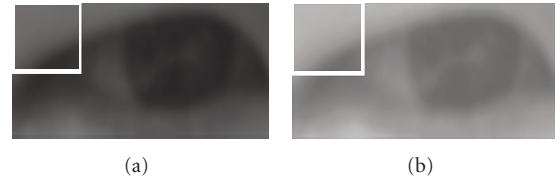
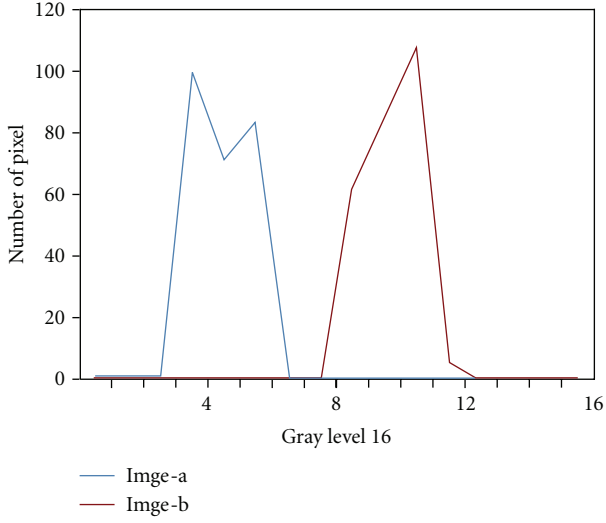
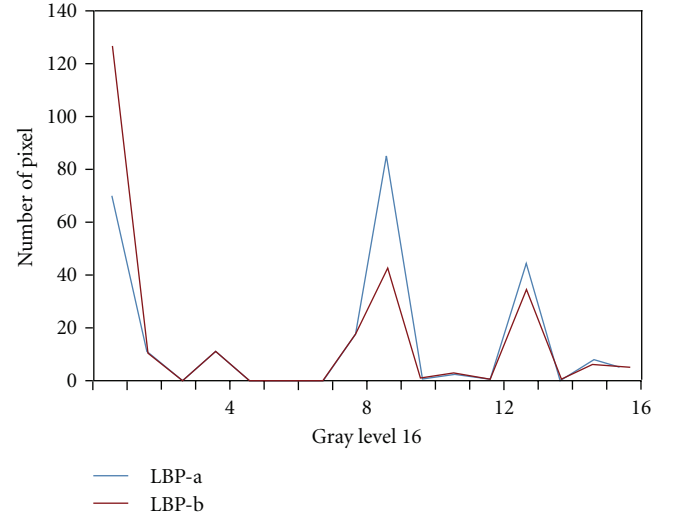


FIGURE 3: Two eye images from the same subject with different illumination.

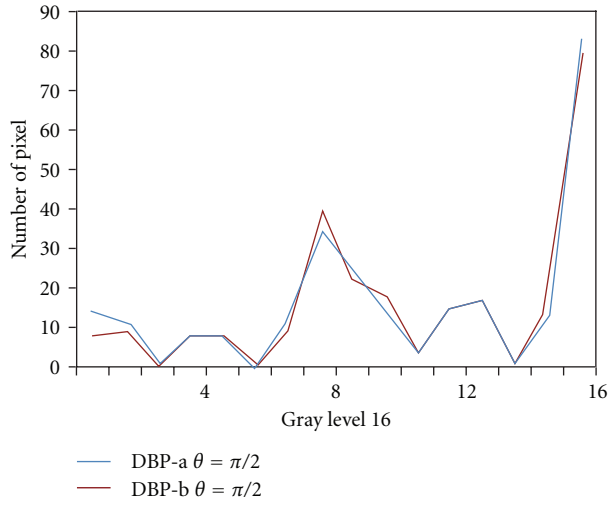
histograms of six representations extracted from the two images of the same eye with different lighting (see Figure 3). In this paper, the parameters of GDBP are set as $v = 2$, $u = 7$, and $\theta = 0, \pi/2$ for a trade-off between speed and accuracy. As shown in Figure 4, the six representations are, respectively, the original image intensity, LBP of the image, $DBP_{\theta=0,\pi/2}$ of the image, and $GDBP_{\theta=0,\pi/2}$ of the image. Two white regions of the same window size 16×16 pixel array as shown in Figures 3(a) and 3(b) are, respectively, selected to extract different histograms, and the results are shown in Figure 4.



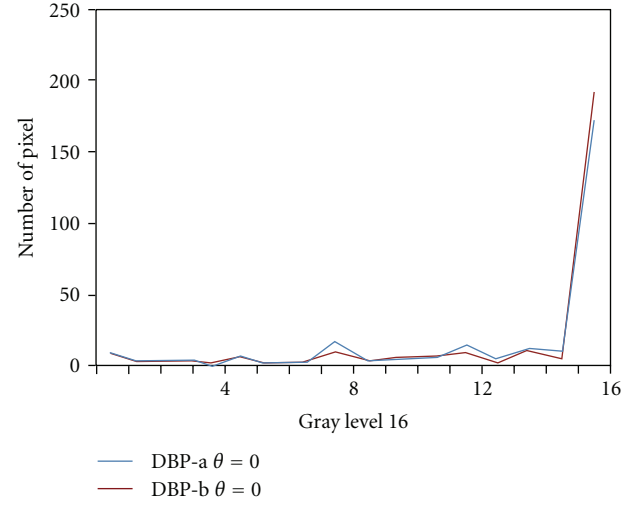
(a)



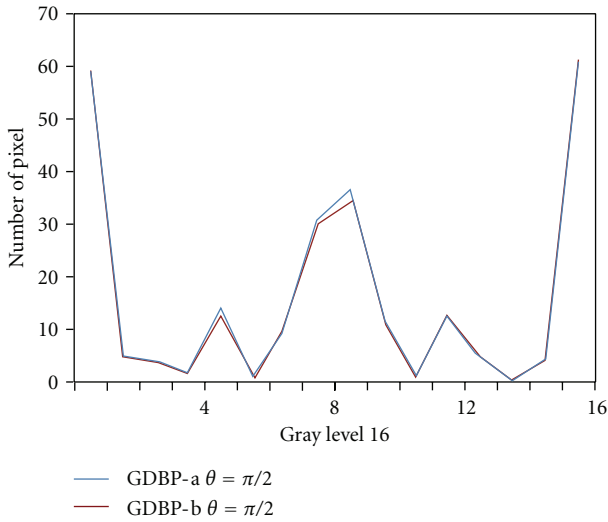
(b)



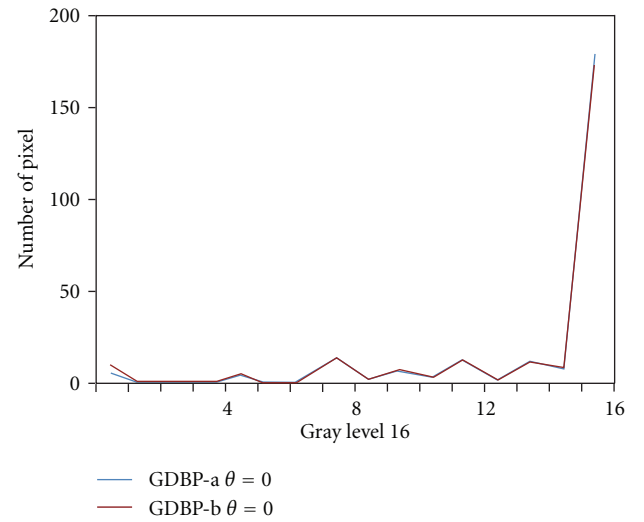
(c)



(d)



(e)



(f)

FIGURE 4: Robustness of the different histogram to images with lighting variation. GDBP (a, b), LBP (a, b), and DBP (a, b) mean that three operators are, respectively, applied on the white regions of image a and image b (see Figure 3), and θ is the direction of GDBP and DBP.

TABLE 1: Errors of three gaze directions under three lighting variations (degree).

Gaze	Feature		
	LBP	DBP	GDBP
$E_\alpha(\mu \pm \sigma)$	2.1 ± 1.5	1.9 ± 1.3	1.7 ± 1.3
$E_\beta(\mu \pm \sigma)$	2.0 ± 1.7	2.0 ± 1.4	1.7 ± 1.4
$E_\gamma(\mu \pm \sigma)$	2.1 ± 1.6	1.8 ± 1.4	1.6 ± 1.2

We can see clearly that the histograms of GDBP ($\theta = 0, \pi/2$) are the most similar. We also can see that GDBP has better robustness properties than either of DBP or LBP. This implies that eyelid representation of GDBP is robust to the lighting variations, which can lessen the effect of reflections on eyelids and benefit for gaze estimation.

In the off-line person-dependent gaze estimation, we use three sets under different lighting conditions to evaluate the performance of LBPs, DBPs, and GDBPs. Three data sets are collected from three persons by the method described in Section 3.1. Each set has the 742 samples, Set 1 with the left lighting, Set 2 with the frontal lighting, and Set 3 with the right lighting, as shown in Figure 5. Three sets are cross-validated with each other in order to evaluate the robustness of various features by estimation accuracy. When one is used as a training set, the others are used as the testing sets. The final estimation accuracy is calculated on three cross-validated experiments. From Table 1, there is a similar accuracy between LBP and DBP, while with the advantages of Gabor filters [10], GDBP has the best robustness to lighting variations.

GDBP is not only robust to the variations of imaging condition but also with much discriminating power. GDBP has the different discriminating power in different orientations. A white region of the window size 16×16 pixel array, as shown in Figure 3(a), is selected to extract the gray histograms by $\text{GDBP}_{\theta=0,\pi/2}$. As shown in Figure 6, we can see clearly that the histograms of horizontal and vertical GDBP contain different discriminative information. This implies that the representation of GDBP can meet the special requirements, namely, GDBP is useful in representing the horizontal and vertical movement of the pupils for gaze estimation.

3. Experiments

3.1. Experimental Data Collection. Gaze estimation aims to calculate the direction of subject's attention from an image, which is represented in terms of Euler angles between the gaze vector \vec{v} and three axes of camera's coordinate system. In this paper, we apply the proposed descriptor to estimate the gaze direction with a basic assumption: head poses are fixed front view. In order to gain the samples labeled by the accurate Euler angles, a designed studio is setup (see Figure 7), which contains a camera C for collecting the face images, a monitor S as a gazed object and a localizer (a commercial FASTRAK [12]) to calibrate the objects (user's eyes, observed cursor on the monitor, camera) of our studio.

The FASTRAK has a transmitter (the world coordinate system $O_W X_W Y_W Z_W$. The origin is located in the centre of the transmitter) and four receivers, and only three receivers are used in our data collection procedure (Each receiver owns a local coordinate system). Three receivers are mounted on the head of the viewer, the camera, and top-left corner of the screen, respectively. The data from a receiver is its position and orientation related to the transmitter, which are six values: X, Y, Z in cm, and Azimuth, Elevation, Roll in degree. And then our system can calculate the receiver's translation and rotation matrices related to the transmitter's coordinate system. In Figure 7, one receiver is mounted on the top of the camera, and the output of its translation and calculated rotation matrices are \mathbf{T}_{Cam} and $\mathbf{R}_{\text{Cam}} = (r_1^{\text{cam}} \ r_2^{\text{cam}} \ r_3^{\text{cam}})$ related to the transmitter's coordinate system. $\mathbf{r}_i^{\text{Cam}}$ ($i = 1, 2, 3$) are the axes of camera's coordinate system. Suppose that the translation and rotation matrices from the receiver's coordinate system to camera's coordinate system are $\mathbf{T}_C = (0 \ y_c \ 0)^T$ and $\mathbf{R}_C = \mathbf{I}$ (\mathbf{I} is a 3×3 unit matrix), respectively. The second receiver is mounted on the top-left corner of the screen, and the output of its translation and calculated rotation matrices are \mathbf{T}_{Scr} and \mathbf{R}_{Scr} related to the transmitter's coordinate system. For each generated cursor as a gazed point, we assume the translation and rotation matrices from the receiver's coordinate system to screen's coordinate system are $\mathbf{T}_S = (x_s \ y_s \ 0)^T$ and $\mathbf{R}_S = \mathbf{I}$, respectively. The third receiver is mounted on top of the viewer's head, and the output of its translation and calculated rotation matrices are \mathbf{T}_{Head} and \mathbf{R}_{Head} related to his transmitter's coordinate system. Assume that the centre of the two eyes has a translation of $\mathbf{T}_V \approx (0, -15, 8)$ in cm related to the third receiver ($(0, -15, 8)$ is a statistical average and the error of different centre can be ignored and tested by the experiments). We keep the X direction of the receiver paralleling to the line between the two eyes and the Y direction upright, and the Z direction parallels the direction of head pose.

As stated above, we have the translation and rotation matrices from the three receivers. Then, the position of the observed target is $\mathbf{T}_{\text{Scr}} + \mathbf{R}_{\text{Scr}} \mathbf{T}_S$, and the position of the centre of two eyes is $\mathbf{T}_{\text{Head}} + \mathbf{R}_{\text{Head}} \mathbf{T}_V$ with respect to the world coordinate system $O_W X_W Y_W Z_W$. Therefore, the gaze vector \vec{v} is calculated by

$$\mathbf{v} = \frac{\mathbf{T}_{\text{Scr}} - \mathbf{T}_{\text{Head}} + \mathbf{R}_{\text{Scr}} \mathbf{T}_S - \mathbf{R}_{\text{Head}} \mathbf{T}_V}{\|\mathbf{T}_{\text{Scr}} - \mathbf{T}_{\text{Head}} + \mathbf{R}_{\text{Scr}} \mathbf{T}_S - \mathbf{R}_{\text{Head}} \mathbf{T}_V\|}. \quad (12)$$

Meanwhile the camera's axes are $\mathbf{r}_1^{\text{Cam}} = (r_{11}^{\text{cam}} \ r_{21}^{\text{cam}} \ r_{31}^{\text{cam}})^T$, $\mathbf{r}_2^{\text{Cam}} = (r_{12}^{\text{cam}} \ r_{22}^{\text{cam}} \ r_{32}^{\text{cam}})^T$, and $\mathbf{r}_3^{\text{Cam}} = (r_{13}^{\text{cam}} \ r_{23}^{\text{cam}} \ r_{33}^{\text{cam}})^T$. Therefore, gaze direction represented in terms of three Euler angles (α, β, γ) of gaze vector \vec{v} with respect to the camera coordinate system is calculated by

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \arccos(\mathbf{v}, \mathbf{r}_1^{\text{Cam}}) \\ \arccos(\mathbf{v}, \mathbf{r}_2^{\text{Cam}}) \\ \arccos(\mathbf{v}, \mathbf{r}_3^{\text{Cam}}) \end{pmatrix}. \quad (13)$$

Our system synchronizes the image capturing and the computation of gaze direction. In the data collection, the

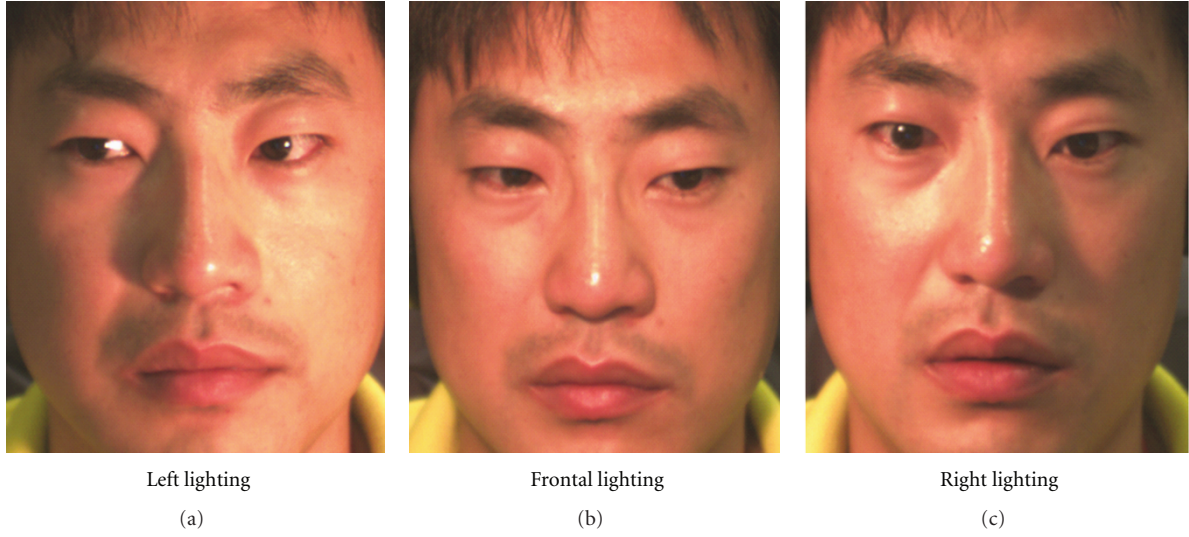


FIGURE 5: Samples under three lighting variations.

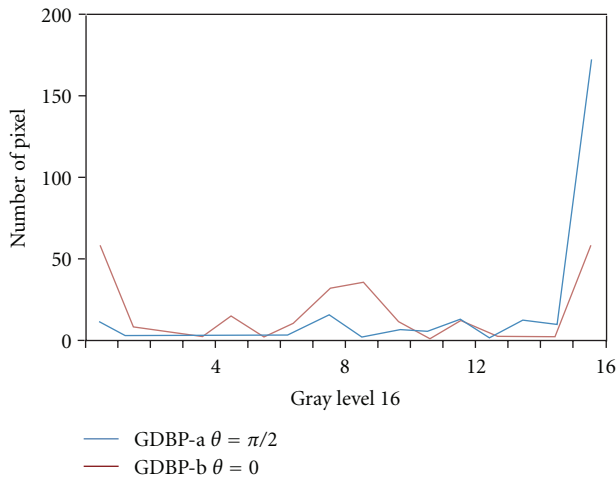


FIGURE 6: Histograms of horizontal and vertical GDBP extract from the white region image as shown in Figure 3(a).

distance between subject's heads and the screen is around 600 mm. In this paper, in order to simplify the experiments, we keep the head pose in front view by holding the Z direction of receiver fixed on the top of head. In the data collection, at each time only one of predefined 16 points appears on the monitor and its world coordinate is calculated from the screen coordinate by the translation and rotation matrices. We provide a database for further research.

3.2. Gaze Estimation Based on GDBP and Experimental Results. In our experiments, we use the GDBP operator to encode the eye images as the appearance-based feature. Our algorithm gets a captured image as input, and outputs the gaze direction. In our method, the eye centers and the contour points of the eyelids are located by the method proposed in [13]. The eye image is then cropped among

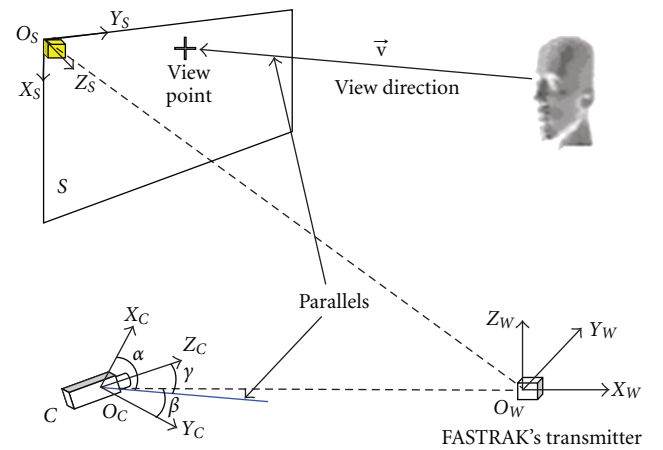


FIGURE 7: Studio setup.

the outline points of the eyes (see Figure 8(a)). The cropped two eyes image is divided into 16 nonoverlap rectangular regions. Our experiments show that when a single eye is divided into 8 regions and double eyes 16 regions, we achieve the best accuracy; see Table 2. GDBP features are computed (see Figures 8(b) and 8(c)), where the parameters are set as $v = 2$ and $u = 7$. The parameters are so set for a trade-off between speed and accuracy. Gray histograms are computed from each region. All histograms (256 bins) are used to replace the classical "pupil-glint" vector. GDBP features, as well as the histograms, are fed to SVR to predict the gaze direction. In our experiments, we calculate GDBP features only in horizontal and vertical directions ($\theta = 0, \pi/2$) to improve the computational efficiency.

In this paper, we use front view samples for training and testing. The cropped eye sample distribution in the gaze space looks like a net, which is a double unicom space as shown in Figure 9. This implies that the eye gaze is a successive movement and we use SVR to map the

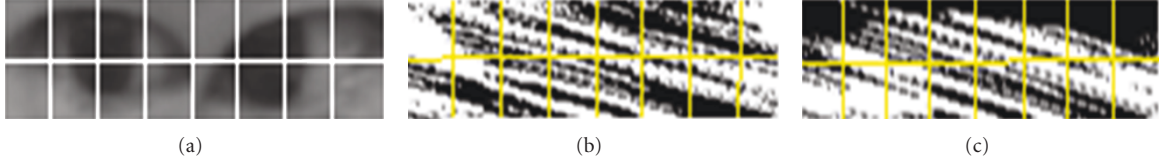


FIGURE 8: (a) Eye image (16 regions of double eyes). (b) Visualization of $GDBP_0$. (c) Visualization of $GDBP_{\pi/2}$.

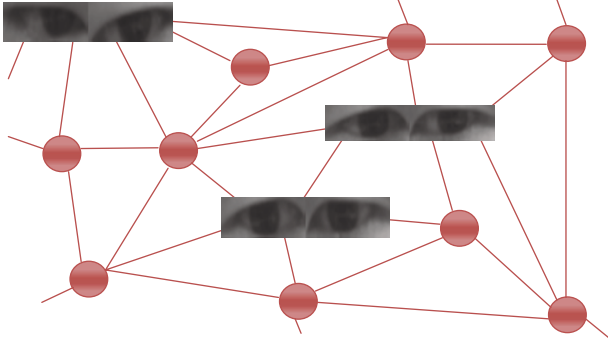


FIGURE 9: The distribution of eye samples corresponding to the different gaze directions.

TABLE 2: Errors of different regions (degree).

Gaze	Regions			
	3×2	4×2	3×3	4×3
$E_a(\mu)$	2.4	2.1	2.3	2.7

TABLE 3: Errors of three gaze directions (degree).

Gaze	GDBP	Feature DBP	LBP
$E_a(\mu \pm \sigma)$	2.1 ± 1.1	2.4 ± 1.5	2.5 ± 1.7
$E_\beta(\mu \pm \sigma)$	2.2 ± 1.5	2.7 ± 1.7	2.8 ± 1.8
$E_\gamma(\mu \pm \sigma)$	2.0 ± 1.6	2.5 ± 1.9	2.6 ± 1.8

model of nonlinear successive motion. Our sample dataset consists of 4089 samples from 11 persons. It is divided into three sets: Set 1 with 1666 samples from 4 persons, Set 2 with 1140 samples from 3 persons, and Set 3 with 1283 samples from 3 persons. Our off-line person-independent experiment is cross-validated on the three sets. We compared our GDBP and DBP operators to LBP operator by applying each operator on the same datasets. From Table 3, we see clearly that GDBP performs the best and achieve an accuracy of less than 2° .

In real-time application, we can also map the gaze point to the coordinate of the screen similar to the classic “glint-pupil” methods. Our gaze estimation system operates in a desktop environment, and a user sits in front of a PC monitor away from 600 mm to 650 mm. Our system consists of a camera (uEye-1540x) and a Windows PC with a monitor, a 3.0 GHz CPU and 1 GB memory (the localizer, FASTRAK,

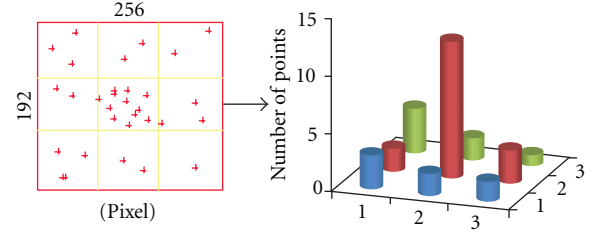


FIGURE 10: Left red rectangle is 256×192 pixels whose center is (640, 480) of the monitor coordinate. Right histogram is a statistical result on 9 subregions (small yellow regions).

is not used here). The camera is mounted on the top of monitor. We only need an initial calibration that calculates the range of sight related to the four corners of the monitor. The distance of two eyes' center is used as a parameter according to the depth of field. The monitor is divided into 16 subregions and subjects gaze at the center of each sub-region. The geometric errors are $(\mu, \sigma) 2.7^\circ \pm 1.8^\circ$, between the centers of the subregions and the estimated gaze points, for all of the subregions. As shown in Figure 10, one of 16 subregions is in size of 256×192 pixels whose center is (640, 480) of the monitor coordinate. Subjects gaze at the center of this sub-region while his head pose is front view for simplifying real-time experiment. The experimental results of three subjects show that the accuracy is 85% of the calculated points located in this rectangle, and the points represent a Gaussian distribution. By utilizing the off-the-shelf devices, gaze estimation is executed in 0.85 second from capturing an image to the output of the gaze direction without little code optimizations. Although most commercial gaze estimation solution provides the accuracy less than 1 degree, they usually depend on the assistance of LEDs, as shown in the head mounted EyeLink II [14]. Our method is nonintrusive without any help of LED. The remote monocular nonintrusive gaze estimation is a tendency of gaze research.

In our experiments, the kernel function of SVR is the Gaussian kernel function. The sixteen regions ($4 \times 2 \times 2 = 16$) of double eye images are used and the average error is around 2° . It is important to note that our eye gaze method is noninvasive, fast, and stable. It is stable due to the robustness of our novel features to the light variances.

4. Conclusions

In this paper, a robust image descriptor, GDBP, is proposed for gaze estimation. GDBP captures not only the local binary

pattern, but also the texture change information related to the given directions. Other advantages of GDBP include noise restrain and robustness to lighting variations. GDBP features are finally fed into SVR to estimate the gaze direction with respect to the camera coordinate system. In the future, we will investigate how to match two GDBPs and how to apply the discriminative capacity of the GDBP operator for other tasks.

References

- [1] Y. Sugano, Y. Matsushita, Y. Sato, and H. Koike, "An incremental learning method for unconstrained gaze estimation," in *Proceedings of the European Conference on Computer Vision*, vol. 5304 of *Lecture Notes in Computer Science*, pp. 656–667, 2008.
- [2] T. N. Cornsweet and H. D. Crane, "Accurate two-dimensional eye tracker using first and fourth Purkinje images," *Journal of the Optical Society of America*, vol. 63, no. 8, pp. 921–928, 1973.
- [3] T. Ohno, N. Mukawa, and A. Yoshikawa, "FreeGaze: a gaze tracking system for everyday gaze interaction," in *Proceedings of the Eye Tracking Research and Applications Symposium (ETRA '02)*, pp. 125–132, March 2002.
- [4] Y. Matsumoto and A. Zelinsky, "An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement," in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (AFGR '00)*, pp. 499–504, 2000.
- [5] D. Hansen and Q. Ji, "In the eye of the beholder: a survey of models for eyes and gaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 478–500, 2010.
- [6] H.-C. Lu, C. Wang, and Y.-W. Chen, "Gaze tracking By binocular vision and LBP features," in *Proceedings of IEEE 19th International Conference on the Pattern Recognition (ICPR '08)*, pp. 1–4, August 2008.
- [7] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [8] H. Lu, G. Fang, C. Wang, and Y. Chen, "A novel method for gaze tracking by local pattern model and support vector regressor," *EURASIP Signal Processing*, vol. 90, no. 4, pp. 1290–1299, 2010.
- [9] A. Smola, B. Scholkopf et al., "A tutorial on support vector regression," Tech. Rep. NeuroCOLT NC-TR-98-030, Royal Holloway College, University of London, London, UK, 1998.
- [10] S. Shan, W. Gao, Y. Chang, B. Cao, and P. Yang, "Review the strength of gabor features for face recognition from the angle of its robustness to mis-alignment," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, pp. 338–341, August 2004.
- [11] B. Zhang, Z. Wang, and B. Zhong, "Kernel learning of histogram of local Gabor phase patterns for face recognition," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 469109, 8 pages, 2008.
- [12] Polhemus FASTRAK, http://www.polhemus.com/?page=Motion_Fastrak.
- [13] Z. Niu, S. Shan, X. Chen, B. Ma, and W. Gao, "Enhance ASMs based on AdaBoost-based salient landmarks localization and confidence-constraint shape modeling," in *Proceedings of the International Workshop on Biometric Recognition Systems (IWBR '05)*, vol. 3781 of *Lecture Notes in Computer Science*, pp. 9–14, 2005.
- [14] EyeLink II, <http://www.sr-research.com/EL-II.html>.