Pattern Recognition Letters 31 (2010) 1377–1387

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Eye/eyes tracking based on a unified deformable template and particle filtering

Yali Li, Shengjin Wang *, Xiaoqing Ding *

State Key Laboratory of Intelligent Technology and Systems, Department of Electronic Engineering, Tsinghua University, Beijing 100084, PR China

ARTICLE INFO

Article history:
Received 10 April 2009
Received in revised form 1 January 2010
Available online 25 February 2010
Communicated by A. Shokoufandeh

Keywords:
Eye/eyes tracking
Deformable template
Particle filter
Eye parameters capturing

ABSTRACT

This paper proposes a novel algorithm framework to solve the eye/eyes tracking problem. Two unified deformable templates are introduced for single eye tracking and double eyes tracking, respectively. Each deformable template can describe both open and closed eye states. Based on the templates the particle filtering framework is applied. A dynamic model considering both the geometrical transformations and state transitions of eye/eyes is given. Furthermore, a measurement model for contour tracking is also modified to adapt for eye/eyes tracking. The experimental results show that our approach can not only track the locations of eye/eyes accurately, but also obtain the eye contour parameters simultaneously.

1. Introduction

Eye/eyes tracking has long been recognized as an important task in computer vision. There are many applications that require robust and accurate eye/eyes tracking, such as human–computer interface (HCI), facial expression analysis, driver surveillance systems. Moreover, in some cases the application systems require not only to track the locations of eye/eyes but also to capture the detailed parameters of eye contour. For example, the changes of distance between upper and lower eyelids over time can provide a reliable measurement for driver fatigue state detection (Dinges and Grace, 1998). The accurate eye contour parameters are also needed to serve as the reference locations of eyeball for obtaining driver's gaze angle.

In this paper, a novel approach for eye/eyes tracking with a combination of unified deformable template and particle filtering is proposed. Firstly, deformable templates for single eye tracking and double eyes tracking are presented. Each template is designed to be capable of fitting the contour of both open and closed eye states. Thereby the eye/eyes contour can be described by using unified parameters without concerning on the eye state transitions. Then the particle filtering framework is applied and a dynamic model which involves geometrical transformations and state transitions of the eye/eyes is constructed. Furthermore, a measurement model for contour tracking (Isard and Blake, 1998a,b) is modified in order to satisfy the requirement of accurate eye contour tracking. The experimental validations show the capability and efficiency of our approach.

The remainder of this paper is organized as follows. A survey of typical existed eye tracking methods is presented in Section 2. The deformable template which is applied to describe eye/eyes states and its related parameters are introduced in Section 3. Then in Section 4, the detail algorithms of the proposed approach are presented, including the framework of particle filtering as well as the applied dynamic model and measurement model. Section 5 gives the experimental results and Section 6 gives the conclusions of the paper.

2. Related work

Numerous techniques have been proposed to locate and track the eye/eyes (Yuille et al., 1992; Deng and Lai, 1997; Tian et al., 2000; Jin et al., 2005; Zhu and Ji, 2005; Zheng et al., 2005; Tan and Zhang, 2006; Zhu et al., 2008; Bacivarov et al., 2008; Wu and Trivedi, 2008). These methods can be briefly classified into two categories: intrusive methods which require physical contacts with the users and non-intrusive methods that do not use physical contacts. In comparison with the intrusive methods, the non-intrusive methods free users from heavy and uncomfortable equipments and are much more convenient for many applications. Therefore, current research mainly concentrates on non-intrusive methods. However, the displacement of direct contacts also increases the degrees of freedom of eye positions and gets more challenges for eye detecting and tracking.

IR-based eye trackers are most widely reported since their encouraging experimental results (Zhu and Ji, 2005; Zhu et al.,

* Corresponding authors. Tel.: +86 10 62772044x208; fax: +86 10 62772044x211 (S. Wang).
E-mail addresses: liyali@ocrserv.ee.tsinghua.edu.cn (Y. Li), wsj@ocrserv.ee.tsinghua.edu.cn (S. Wang), dxq@ocrserv.ee.tsinghua.edu.cn (X. Ding).

0167-8655/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2010.02.013

2008). Under active IR (infra-red) illumination, the pixels constituting pupil are much brighter than neighboring pixels. Based on this so-called bright pupil effect, various methods are introduced to obtain the locations of eyes frame by frame, such as Kalman filtering and mean-shift method (Zhu and Ji, 2005; Zhu et al., 2008). Their examples show that even traditional problems like significant illumination changes are partly solved. However, this approach strictly requires the IR light source and user eye should be located on the optical axis of lens. Much more serious problems are that bright pupil effect does not appear in eye closed or face orientation changed images. Thus the IR-based eye trackers are prone to failure during eye blink or head movement procedure. These problems limit its applications.

In order to tackle with the dual-state (e.g. closed and open) eye tracking problem, new eye trackers are presented by considering the changes of appearance caused by eye blinks. Tian et al. (2000) develop an eye tracker combining feature point tracking with masked edge filtering techniques. In order to detect the eye state and recover the detailed eye parameters, deformable templates describing open and closed eyes respectively are introduced. Tan and Zhang (2006) modify Tian's approach by amending the energy function for eye state detection. Using autoregressive (abbreviated as AR) models the eyelid patterns are estimated by several former frames instead of only one to predict the eyelid patterns. These techniques do help to improve the accuracy of eye state detection. However, clear and high resolution eye region images are required in both referred approaches. Bacivarov et al. (2008) propose an eye tracking approach based on Active Appearance Models (abbreviated as AAM) (Cootes et al., 1998). They incorporate eye templates for open and closed eyes by using the same number of key points. Principal Component Analysis is applied to obtain both shape model and appearance model. The limitation of this approach is that its accuracy mainly relies on the training dataset of AAM.

In recent years particle filtering (Isard and Blake, 1998a,b) has been widely used to solve tracking problems. Therefore, several eye tracking methods based on this framework have been proposed. Jin et al. (2005) introduce a color-based multiple states particle filter for eye tracking. Both the locations of eyes and the transitions between open and closed eye states are considered in the dynamic model. Meanwhile the color histogram features are applied in the measurement model. Wu and Trivedi (2008) apply two interactive particle filters for open and closed eyes separately to track the locations and scales of eyes robustly. Both methods only detect eye state without capturing the detailed parameters. Therefore the application areas are restricted.

Our work is to develop an eye/eyes tracking method which can simultaneously obtain the eye locations and recover eye model parameters. For this purpose, unified deformable template which can fit the contour of eye or eyes in both open and closed states is introduced. Under the framework of dynamic systems modeling and particle filtering, a novel algorithm framework for eye/eyes tracking is proposed. The details and experimental evaluation results are all presented in the following sections.

3. Eye/eyes deformable template

Generally, eye blinks have always been a main obstacle on the progress of eye tracking research since they cause the significant appearance changes in the eye region of images. In this section, we present a unified deformable template to describe both open and closed eyes. The related parameters are also attached to the template. In addition, for the purpose of tracking two eyes at same time, another deformable template concerned with symmetry between two eyes is also given.

3.1. Eye deformable template

There are various deformable templates proposed to describe the eye features. Yuille et al. (1992) present an open eye template of which the iris boundary is described by a circle and the eyelid contour is fitted by parabolic curves. Tian et al. (2000) has modified Yuille's template by adding another model which uses the straight line to represent the closed eye for the purpose of dual-state parametric eye tracking. However, our experiments show that the contour of some closed eyes can be better fitted with parabolic curves rather than straight line. By considering this, we use two parabola curves with a same symmetrical axis to describe the upper and lower eyelids, respectively. Actually our model includes Tian's template since the straight line can be considered as one kind of special parabola with the quadratic factor set as zero. Besides, since the iris is often partially sheltered by eyelids and sometimes even invisible in low-contrast or blurred images, we eliminate the circle which Yuille et al. (1992) use to represent the iris boundary.

Fig. 1 illustrates the related parameters of our proposed eye deformable template. Let P_l and P_r denote the left and right eye corners, P_u and P_d represent the apexes of upper and lower eyelids, respectively. Then, we choose the following parameters to describe the model. Let (x_c, y_c) refer to the center coordinate of the eye, which is also the cross point of line $\overline{P_lP_r}$ and $\overline{P_uP_d}$. w is the distance between eye corners and θ is the orientation of the template. h_u and h_d refer to the point-to-line distance which are specified as the point P_u and P_d to line $\overline{P_lP_r}$, respectively. To make sure of the one-to-one correspondence between the values of h_u, h_d and points P_u, P_d , we define h_u, h_d as directed distances. That is, if the line is expressed as an equation $ax + by + c = 0$, then h_u and h_d can be expressed by Eq. (1):

$$h_u = \frac{ax_u + by_u + c}{\sqrt{a^2 + b^2}}, \quad h_d = \frac{ax_d + by_d + c}{\sqrt{a^2 + b^2}}, \tag{1}$$

where (x_u, y_u) and (x_d, y_d) denote the locations of P_u and P_d, respectively. To sum up, the mentioned six parameters make up a vector $\boldsymbol{X}$ as Eq. (2):

$$\boldsymbol{X} = (x_c, y_c, w, \theta, h_u, h_d). \tag{2}$$

Fig. 2 shows the fitting results for eyes in different states based on the deformable template.

3.2. Eyes deformable template

For some applications it is inadequate to track only one eye. Besides, tracking of double eyes on approximately frontal face can be

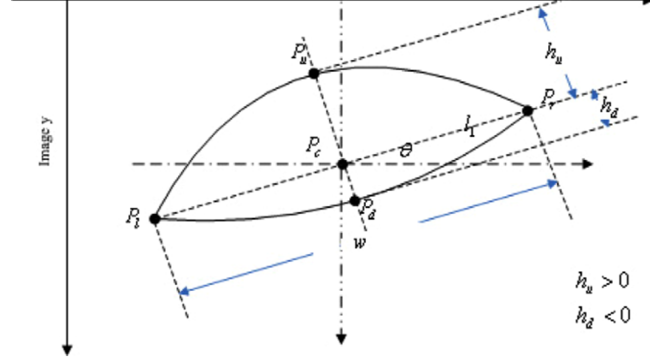


Fig. 1. Related parameters of eye deformable template. Note that the parameters h_u and h_d are both directed distances.

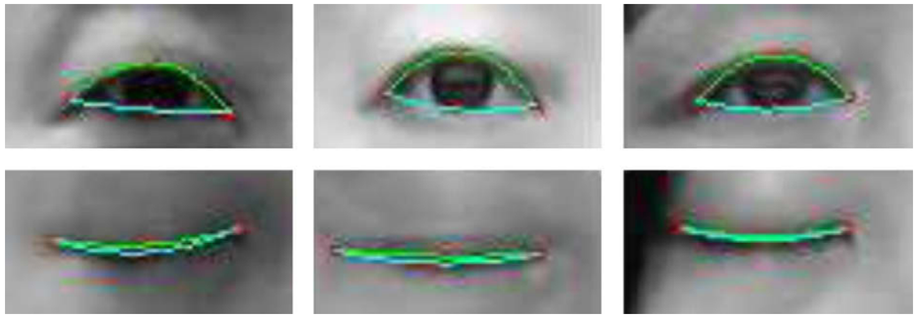


Fig. 2. Fitting results of eyes. The eye samples come from CAS-PEAL Face Database (Gao et al., 2005). The green curves indicate the eyelids while the red-dots mark the eye corners and apexes of eyelids.

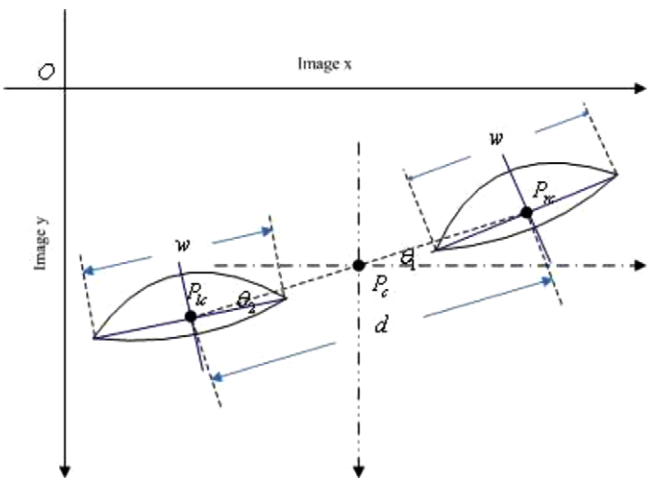


Fig. 3. Related parameters of two eyes deformable template. It is similar to single eye template in Fig. 2.

more robust since much more observation features are added. A straightforward way for eyes tracking is to track two eyes individually. However, the extra computational costs are imposed and the relations between eyes are also ignored. By the consideration for overcoming these shortages, we propose a new deformable template to represent two eyes at one time, which is shown in Fig. 3.

Our proposed deformable template for eyes tracking is based on two assumptions. One is that people's two eyes are approximately symmetrical. The other is that the left and right eyes blink in a closely synchronized fashion. These assumptions are also reasonable from the physiological acknowledge. As it is shown in Fig. 3, the following related parameters are used to describe the eyes deformable template. (x_c, y_c) refers to the location of center point of the two eyes and d refers to the distance between left and right eyes. θ_1 is the orientation of the line connecting two eyes and θ_2 is the angle between the line connecting two eyes and the line connecting the corners of left eye. In addition, the denotations of w, h_u and h_l are identical with single eye template described in Section 3.1. The corresponding vector is composed by the seven parameters in Eq. (3):

$$\boldsymbol{X} = (x_c, y_c, w, \theta_1, \theta_2, h_u, h_d). \tag{3}$$

4. Eye/eyes tracking

In this section, the details of the proposed eye/eyes tracking approach are presented. Brief review of particle filter is described. The applied dynamic model and measurement model are also given. It should be noted that since our target is to track the eye/eyes, we do not concern too much with the approaches used for detection. The existed methods for eye detection on gray images (Deng and Lai, 1997; Yuille et al., 1992) or color images (Zheng et al., 2005) can be adapted according to the applications. Moreover, the precise parameters in the initial frame are not indispensable since our method can automatically adjust them in subsequent frames.

4.1. Particle filter

Generally, a dynamic system is composed by a dynamic model which indicates the state transitions over time and a measurement model which indicates the relations between the states and observations. The aim of object tracking is to estimate the states sequentially. Particle filtering presents a good solution framework for the tracking of stochastic movement (Isard and Blake, 1998a,b). Let $\boldsymbol{x}_t$ denote the state vector and $\boldsymbol{z}_t$ denote the observation vector at a given time t. Then the dynamic model of the system has the form $\boldsymbol{x}_t = f_t(\boldsymbol{x}_{t-1}, \boldsymbol{w}_t)$ and the measurement model has the form $\boldsymbol{z_t} = h_t(\boldsymbol{x}_t, \boldsymbol{v}_t)$, where $\boldsymbol{w}_t$ and $\boldsymbol{v}_t$ denote noises. The particle filtering approach approximates the probability distribution of state $\boldsymbol{x}_t$ with a weighted sample-set $\{(\boldsymbol{s}_t^{(n)}, \pi_t^{(n)})\}$ and applies the following iterative process to the sample-set to obtain object positions sequentially. That is, from the "old" sample-set $\left\{\left(\boldsymbol{s}_{t-1}^{(n)}, \pi_{t-1}^{(n)}\right)\right\}$ at time $t-1$, a "new" sample-set $\left\{\left(\boldsymbol{s}_t^{(n)}, \pi_t^{(n)}\right)\right\}$ for time t is constructed through three steps (Isard and Blake, 1998a,b):

1. **Select** a sample $\boldsymbol{s}_t'^{(n)}$ as follows: Generate a random number j with the probability proportional to $\pi_{t-1}^{(j)}$ by binary subdivision using cumulative probabilities. Then, we set $\boldsymbol{s}_t'^{(n)} = \boldsymbol{s}_{t-1}^{(j)}$.
2. **Predict** by sampling from $p(\boldsymbol{x}_t | \boldsymbol{x}_{t-1} = \boldsymbol{s}_t'^{(n)})$ to choose each $\boldsymbol{s}_t^{(n)}$.
3. **Measure** and weight the new position with image features $\boldsymbol{z}_t$ as $\pi_t^{(n)} = p(\boldsymbol{z}_t | \boldsymbol{x}_t = \boldsymbol{s}_t^{(n)})$, then normalize so that $\sum_n \pi_t^{(n)} = 1$.

From the sequentially constructed sample-set, the tracked position at each time t can be estimated as Eq. (4):

$$\varepsilon[g(\boldsymbol{x}_t)] = \sum_{n=1}^{N} \pi_t^{(n)} g(\boldsymbol{s}_t^{(n)}). \tag{4}$$

4.2. Dynamic model

In general, the geometrical transformations of eye are mainly caused by head movements. Thus the influenced parameters are the location (x_c, y_c), the scale w and the orientation θ. By contrast, the eye state transitions change the shape of template. Thus the influenced parameter is the proportionality factor between h_u and w. In view of these, we build up the dynamic model by considering both geometrical transformations and state transitions of eye.

A second-order autoregressive (AR) model which has been widely used in particle filter tracking approaches (Isard and Blake, 1998a,b) is used to estimate eye location changes. It can be written as

$$\boldsymbol{x}_t^c = A_x(\boldsymbol{x}_{t-1}^c - \bar{\boldsymbol{x}}^c) + \bar{\boldsymbol{x}}^c + B_x\boldsymbol{\mu}_t^x, \tag{5}$$

$$\boldsymbol{y}_t^c = A_y(\boldsymbol{y}_{t-1}^c - \bar{\boldsymbol{y}}^c) + \bar{\boldsymbol{y}}^c + B_y\boldsymbol{\mu}_t^y, \tag{6}$$

where

$$\boldsymbol{x}_t^c = \begin{pmatrix} x_t^c \\ x_{t-1}^c \end{pmatrix}, \quad \boldsymbol{y}_t^c = \begin{pmatrix} y_t^c \\ y_{t-1}^c \end{pmatrix}. \tag{7}$$

$\bar{\boldsymbol{x}}^c$ and $\bar{\boldsymbol{y}}^c$ are the corresponding mean values for $\boldsymbol{x}^c$ and $\boldsymbol{y}^c$. Furthermore, (A_x, A_y) and (B_x, B_y) are matrices representing the deterministic and stochastic components, respectively. They can be obtained by maximum likelihood estimation method or manually set by prior knowledge. $\boldsymbol{\mu}^x$ and $\boldsymbol{\mu}^y$ are the independent and identically distributed (i.i.d) Gaussian noises.

When considering the transition of the orientation θ, we note that it is confined by neck locations. Thereby a first-order AR model is used to model the orientation change, that is

$$\theta_t = \bar{\theta} + C_\theta(\theta_{t-1} - \bar{\theta}) + D_\theta v_t. \tag{8}$$

Similarly, C_θ and D_θ are the parameters describing deterministic and stochastic components, respectively. $\bar{\theta}$ is the mean value of orientation and v is the i.i.d Gaussian noise. Since the scale transition is caused by the distance changes between the eye and its monitoring camera, we introduce the model as Eq. (9) from a perspective point of view:

$$w_t = w_{t-1}(1 + \eta_t). \tag{9}$$

where η is the i.i.d Gaussian noise. In fact, a random multiplicative disturbance term is added to model the scale variation.

Besides, the shape changes caused by eye state transitions will also need to be modeled. Intuitively, eye state transitions mainly influence the proportionality factor between h_u and w. The former refers to the directed distance between upper eyelid apex and eye center while the latter denotes the distance between two eye corners. Since the lower eyelid does not change too much, h_d/w is assumed to be constant even when eye state transition happens. To simplify the following expression, we represent h_u/w with γ. Considering that the eye blink is a continuous changing procedure, γ can be viewed as a continuous random variable. However, we divide its value into three intervals and label them as state o (open), state h (half open) and state c (closed), respectively. For each state, a Gaussian random variable is attached. Therefore, γ is a random variable with the probability distribution shown in Eq. (10):

$$\gamma \sim \alpha_o N(\mu_o, \sigma) + \alpha_h N(\mu_h, \sigma) + \alpha_c N(\mu_c, \sigma). \tag{10}$$

$N(\mu_i, \sigma)$ is Gaussian distribution with mean μ_i and the standard deviation σ. α_o, α_h and α_c are the mixing coefficients and they satisfy the equation $\alpha_o + \alpha_h + \alpha_c = 1$. μ_o, μ_h and μ_c refer to the mean value of γ in state *o*, *h* and *c*, respectively.

A discrete random variable noted as Q_t is introduced to denote the label of eye state at time t. Thus the probability for state *o*, *h* and *c* can be expressed as $P(Q_t = o)$, $P(Q_t = h)$ and $P(Q_t = c)$, respectively. We then use a transitional probability matrix (TPM) $\Pi = \{p_{ij}\}$ to predict the eye state at the next time $t+1$. The component p_{ij} indicates the probability of the transition from state i to state j, that is

$$p_{ij} = P\{Q_{t+1} = j | Q_t = i\}, \quad i,j \in \{o,h,c\}. \tag{11}$$

Hence the TPM Π is a 3×3 matrix in Eq. (12):

$$\Pi = \begin{bmatrix} p_{oo} & p_{oh} & p_{oc} \\ p_{ho} & p_{hh} & p_{hc} \\ p_{co} & p_{ch} & p_{cc} \end{bmatrix}. \tag{12}$$

We make a reasonable assumption that Q_t is merely dependent to γ_t while independent to γ_{t+1}. Therefore the transition of γ_t can be expressed as

$$p(\gamma_{t+1}|\gamma_t) = \sum_{Q_t} \sum_{Q_{t+1}} p(Q_t|\gamma_t) p(Q_{t+1}|Q_t) p(\gamma_{t+1}|Q_{t+1}). \tag{13}$$

The posteriori probability $p(Q_t|\gamma_t)$ can be estimated by Bayesian formula. However, when it is implemented under the framework of particle filter, the delta function is chosen to approximate $p(Q_t|\gamma_t)$ in order to reduce computational cost. The maximum-likelihood state Q_t corresponding to γ_t is chosen and the state Q_{t+1} is sampled from the distribution $p(Q_{t+1}|Q_t)$, which has been indicated in Π. Finally, the new state value γ_{t+1} is sampled from the Gaussian distribution $N(\mu_{Q_{t+1}}, \sigma)$.

Now the dynamic model for single eye tracking has been built up. Moreover, when considering double eyes tracking, the transition models for x_c, y_c, w, h_u, h_d are similar to those of single eye tracking mentioned above. In addition, the transition model for d is similar to w, and θ_1 is similar to θ. Since θ_2 is determined by the property of two eyes and does not change too much over time, no extra transition model is required. Therefore, though the dimension of state vector increases by two, the number of actual variable parameters is only added by one.

4.3. Measurement model

Theoretically, the entire set of visible features can all be the observations and be taken into account to construct the measurement model. However, since the deformable eye/eyes template coincides with the contour of eye/eyes, edge features are chosen to build up the measurement model.

Isard and Blake (1998a) have already proposed a systemic measuring approach for contour tracking with edge features and experiment results have also been presented to validate the capability. Our measurement model for eye/eyes tracking is based on their model yet we replace the B-spline curves with parabolic curves described by the template parameters. For each parabolic curve, several points are uniformly sampled and at each sample point there exists a normal line. Edges are detected along the normal lines and the coordinate values of corresponding edge points are used as the features to calculate the posterior measurement probability. Besides, since the intersection points of two parabolas are singular and no associated normal lines exist, we use a line to connect the two eye corners instead (see Fig. 4).

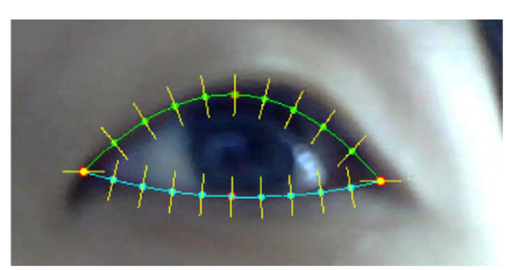
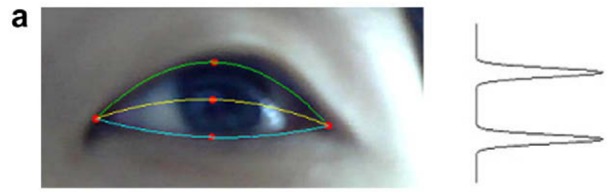
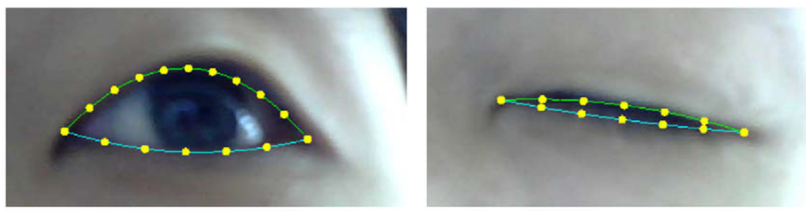


Fig. 4. Measurement procedure: the edges are detected along the yellow lines and the coordinate values of corresponding points are used as the features to calculate the posterior measurement probability.

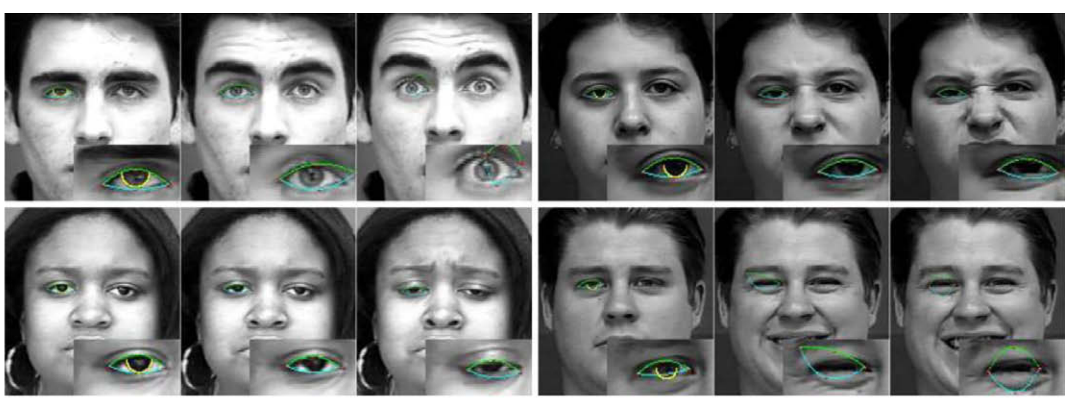


(a)Problem: The features corresponding to upper and lower eyelids are indistinguishable. Therefore, when processing an open eye image, the measuring probability distribution is bimodal. Consequently, the estimated location for upper eyelid is as indicated by the middle yellow curve, far from the actual location as indicated by upper green curve.

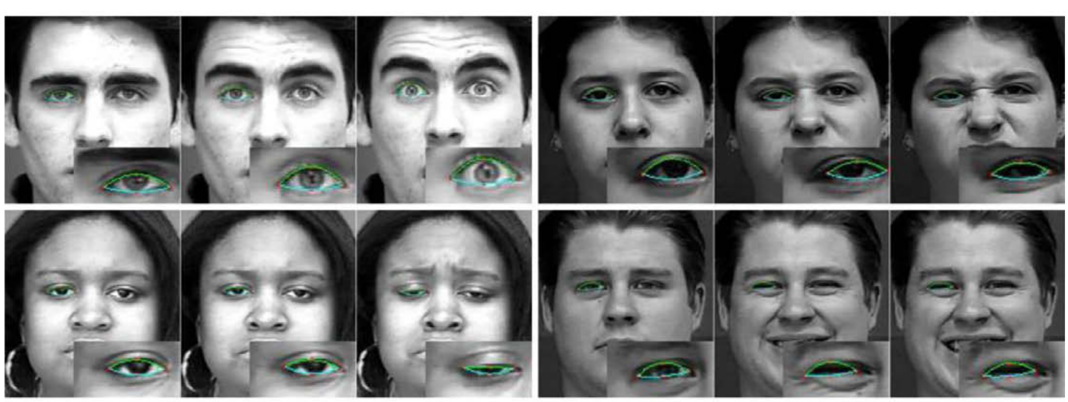


(b) Our solution: The number of sampled points is not constant. It changes along with the length of eye contour. That is, in an open eye image there is a long contour and more points are sampled to calculate the measured probability, while fewer points are sampled from a short contour in a closed eye image.

Fig. 5. Encountered problem and applied adaptive number of sample-points technique.



(a) The tracking results of the method proposed by Tan and Zhang (2006)



(b) The tracking result of our method

Fig. 6. Examples of left eye tracking on image sequences from Cohn-Kanade database (Kanade et al., 2000).

If M specified points s_m $(m = 1, \dots, M)$ are sampled from the parabolic curve and the fixed cast length along the normal line for edge detection is μ, the measurement probability is then expressed as

$$p(\boldsymbol{z}|\boldsymbol{x}) \propto \exp\left\{-\sum_{m=1}^{M} \frac{1}{2M\sigma^2} f(\boldsymbol{z}_1(\boldsymbol{s}_m) - \boldsymbol{r}(\boldsymbol{s}_m); \mu)\right\}, \quad s_m = m/M, \tag{14}$$

Table 1
The statistic results of eye state detection.

	Accuracy	True positive rate	True negative rate
Our method	0.97	0.93	0.98
Method proposed by Tan and Zhang (2006)	0.88	0.65	0.96
Method proposed by Tian et al. (2000)	0.61	0.74	0.57

where σ is the standard deviation, $\mathbf{r}(s_m)$ is the location of the mth sampled point and $\mathbf{z}_1(s_m)$ is the closest associated feature to $\mathbf{r}(s_m)$. $f(v;\mu) = \min(v^2, \mu^2)$ is the function to take account of the search scale μ . The details for derivation can be found in (Isard and Blake, 1998a). When considering double eyes tracking, we repeat the same measuring procedure for another eye.

However, according to our experiments, there is a problem in the measurement model above. That is, the associated features for upper and lower eyelids are indistinguishable. As a result, when processing an open eye image, the measured probability

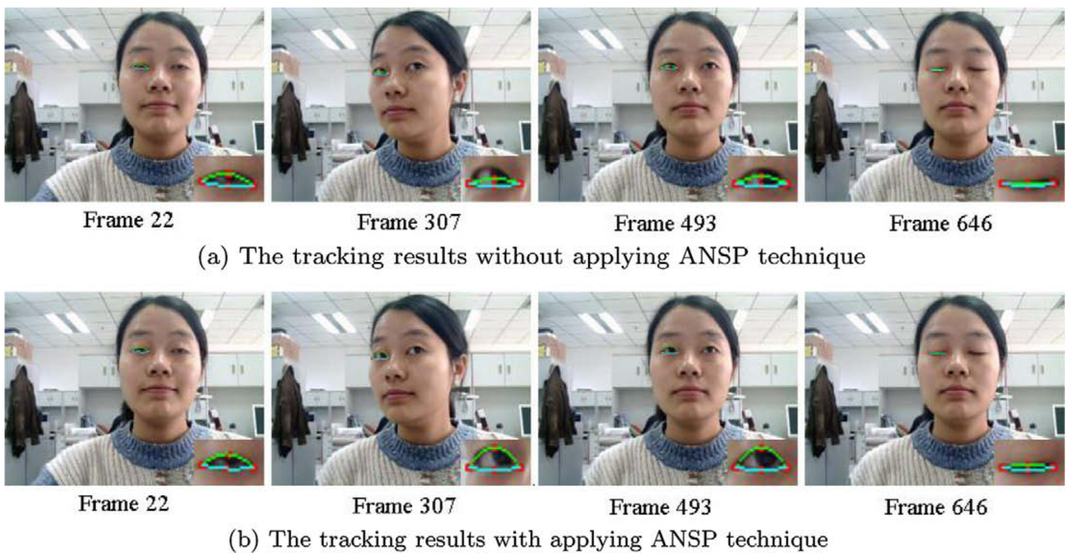


Fig. 7. An example of left eye tracking to demonstrate the importance of ANSP technique.

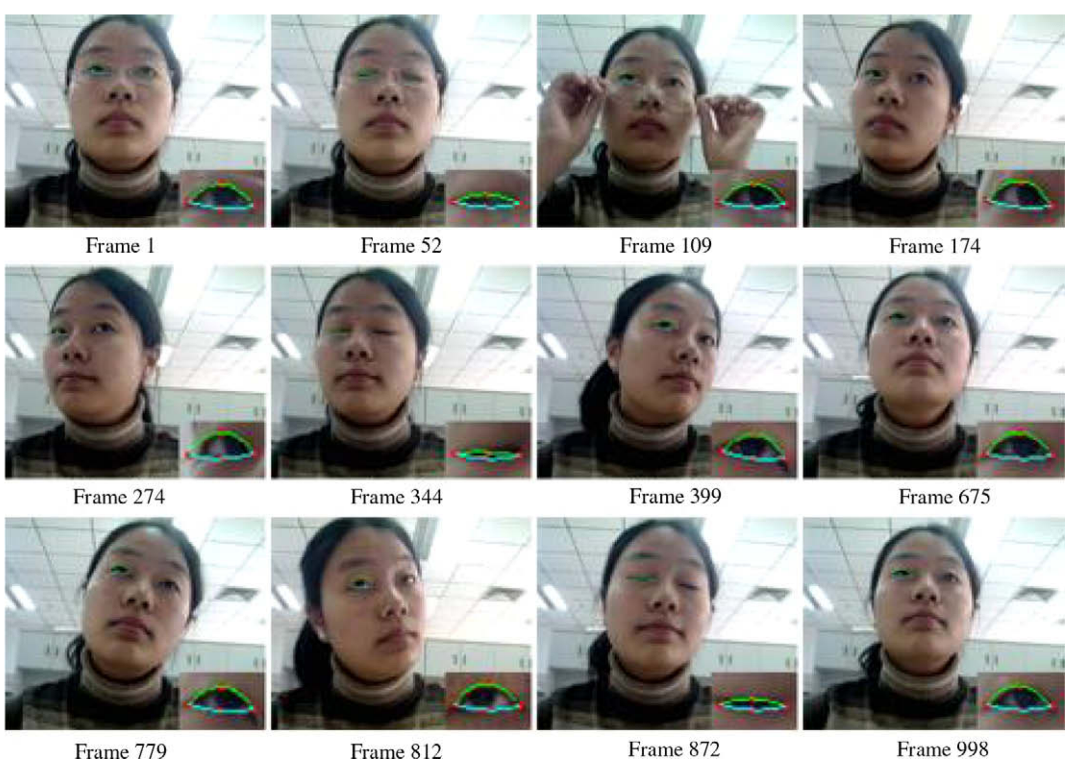


Fig. 8. An example of left eye tracking on indoor video with complex head movements and appearance changes.

distribution is bimodal. One peak corresponds to actual location of upper eyelid and the other corresponds to the location of lower eyelid. Consequently, the tracking location of upper eyelid estimated by Eq. (14) is much less than actual value. We might use prior knowledge to distinguish these two kinds of features. But it is hard for realization. Here, we introduce an effective technique

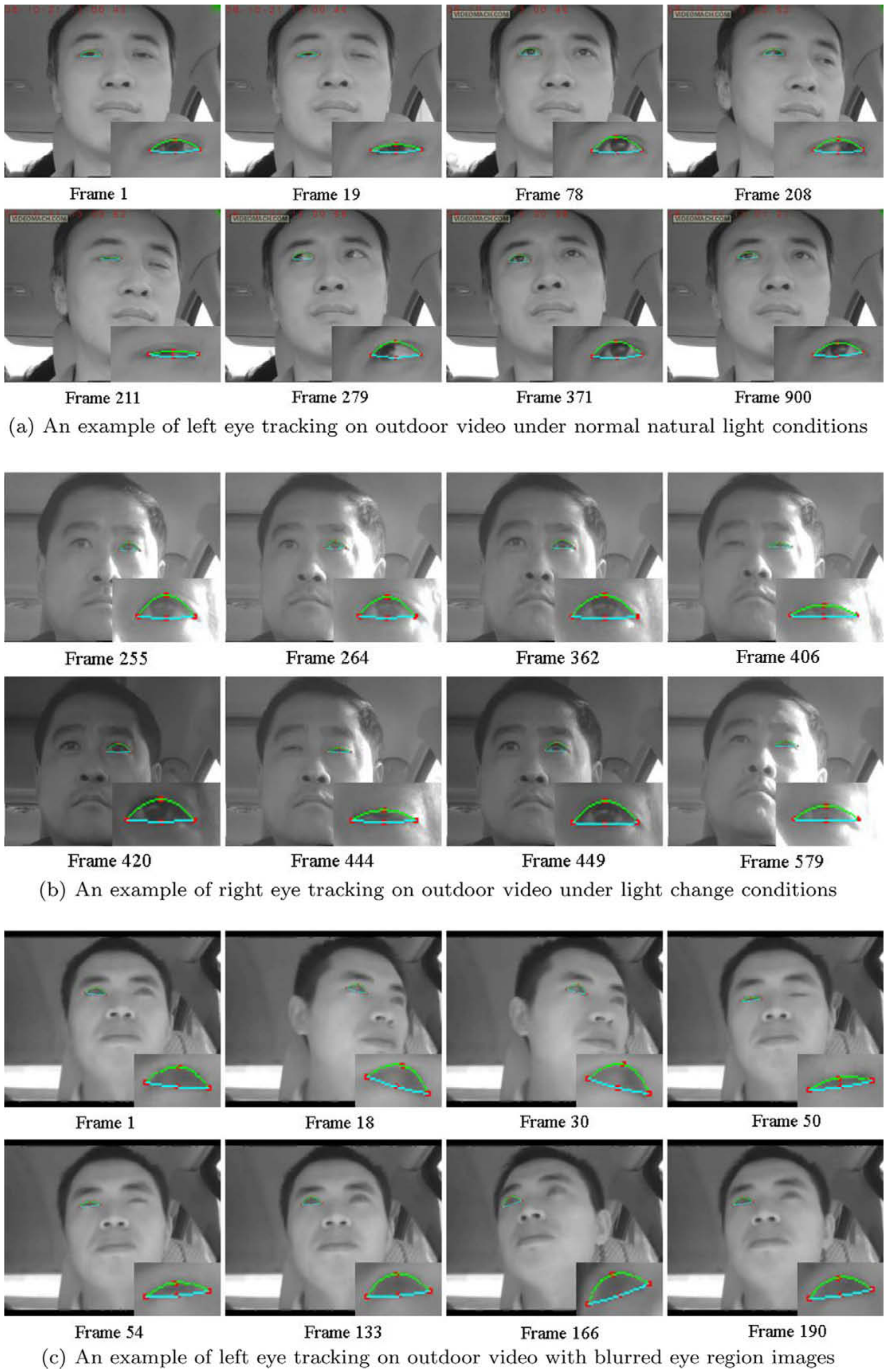


Fig. 9. Examples of eye tracking on outdoor videos.

to tackle this problem and name it as adaptive number of sample-points (abbreviated as ANSP) technique. Based on this technique, the number of sampled points is no longer a constant. It is variable and determined by the length of eye contour. When the parameters of the particle indicate the open eye state, there is a long eye contour and thus more points are sampled to calculate the measured probability. By contrast, for the closed eye particle, fewer points are sampled from a short contour. The details of the encountered problem and the associated solution are illustrated in Fig. 5. In actual implementation, we do not calculate the length of eye contour but just use the normalized eyelids distance $|h_u - h_d|/w$ instead.

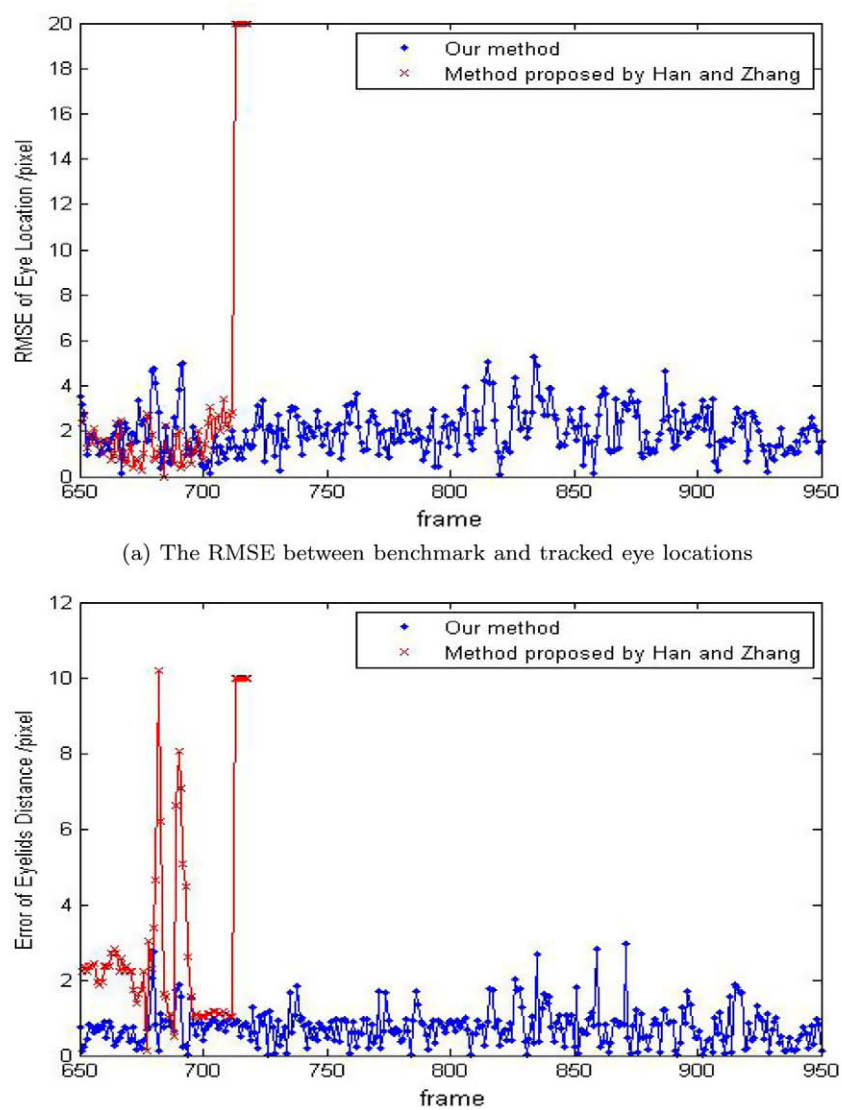
5. Experimental results

The proposed eye/eyes tracking method is implemented by using Matlab on a PC with Pentium (R) 4 2.80 GHz CPU and a 1.00 GB RAM. To evaluate the performance, both image sequences from public available Cohn-Kanade database (Kanade et al., 2000) and those collected under indoor and outdoor environments are tested. Several examples are also presented in the remainder part of this section. Besides, quantitative evaluations such as eye tracking accuracy and eye state detection rate are also shown. Generally, the number of particles used for single eye tracking is generally set to 150–200, while about 300 particles are used for double eyes tracking.

5.1. Performance of eye tracking

5.1.1. Comparative results for eye tracking

Since the aim of our method is much alike with that proposed by Tan and Zhang (2006), we have re-implemented their method for comparison. The tested image sequences are chosen from Cohn-Kanade database (Kanade et al., 2000) in which subjects of different races with various facial expression changes have been videotaped. The average width of single eye region is about 50 pixels. For initial frames, the locations of the eye corners and the apexes of eyelids are manually marked. Several tracking results are illustrated in Fig. 6. The green curves are used to mark the eyelid according to the tracking results while the red-dots point out



(a) The RMSE between benchmark and tracked eye locations

(b) The error of eyelids distance recovering

Fig. 10. An evaluation of tracking accuracy. The horizontal axis for both figures is the frame index and the vertical axis are the tracking error of eye locations and eyelids distance, respectively.

the key points. The results demonstrate that eyes under different facial expressions can well be tracked with our method and more accurate eyelids tracking results are obtained.

Besides, we also conduct the quantitative measurements for eye state detection and compare those with the statistical data listed in Tan and Zhang's paper. Here, we define closed eye as that whose eyelids distance is less than 20% of normal open eye. Three measurements defined in Tan and Zhang's paper named accuracy, true positive rate and true negative rate are listed in Table 1. We can see that the performance of eye state detection is improved dramatically.

Two reasons are regarded to cause the distinctive experimental results. First, in both the methods proposed by Tian et al. (2000) and Tan and Zhang (2006), one half-circle mask consisting of edge and intensity features is applied to detect iris boundary and discriminate eye state. However, the edge map for eye region is so complex that it is difficult to accurately locate the position of iris and furthermore is difficult to detect eye state. Second, in their methods, pyramid Lucas-Kanade trackers (Lucas and Kanade, 1981) are applied to track the eye corners and the apexes of eyelids. The trackers assume that the intensity values do not change but merely shift from one position to another for any given region. However, the patterns of eye region might change a lot under different facial expressions. In contrast, our approach uses eyelids distance to distinguish eye state. Furthermore, particles expressing various deformable template parameters are measured by edge features to obtain the eye contour frame by frame. Therefore, our approach can extract the eye feature more accurately and acquire higher detection rate of eye state.

5.1.2. Influence of ANSP Technique

In Section 4.3 we have described a technique called ANSP to solve the measuring problem caused by indistinguishable features for upper and lower eyelids. By this technique the number of

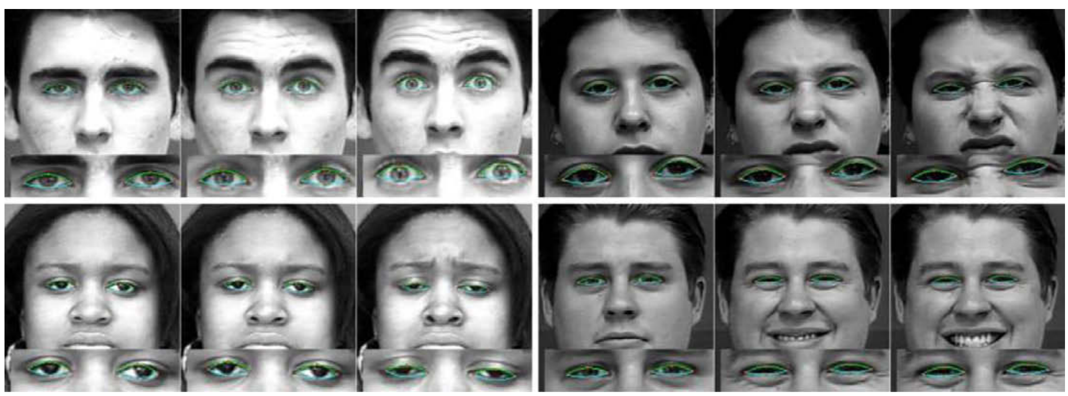


Fig. 11. Examples of eyes tracking on Cohn-Kanade database (Kanade et al., 2000).

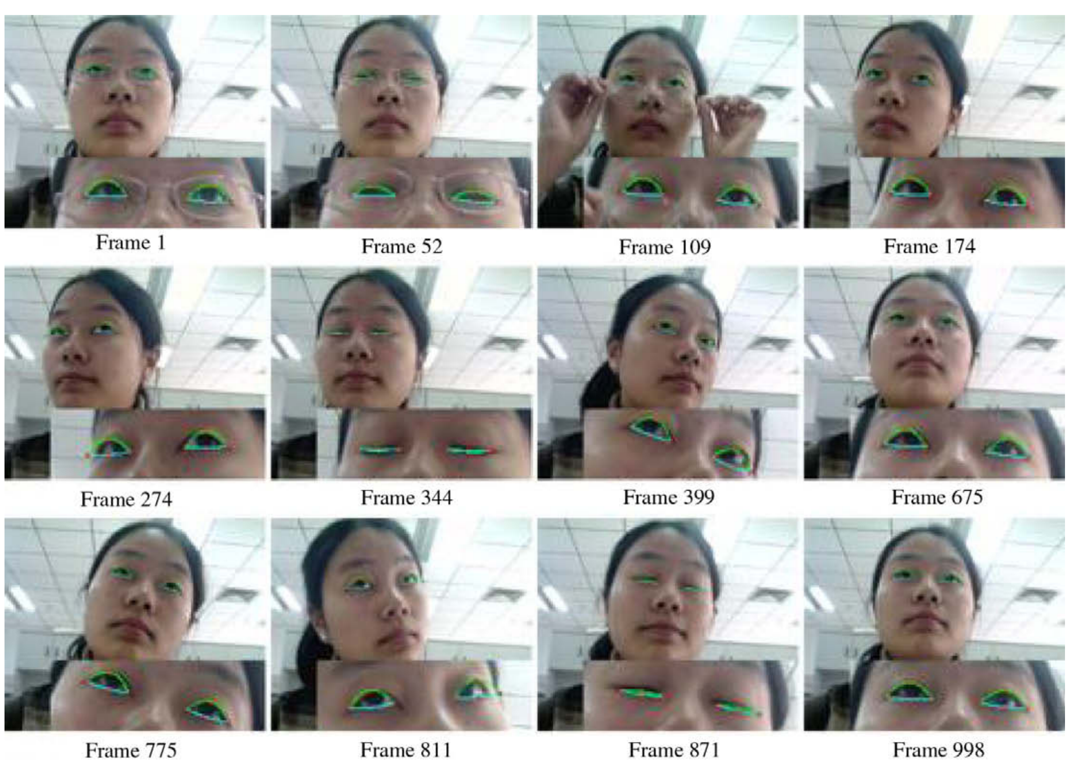
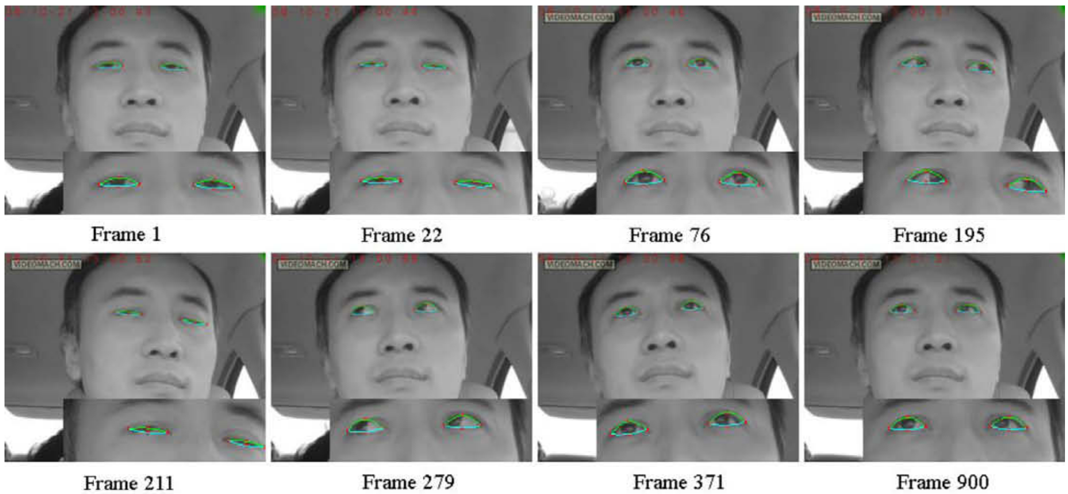
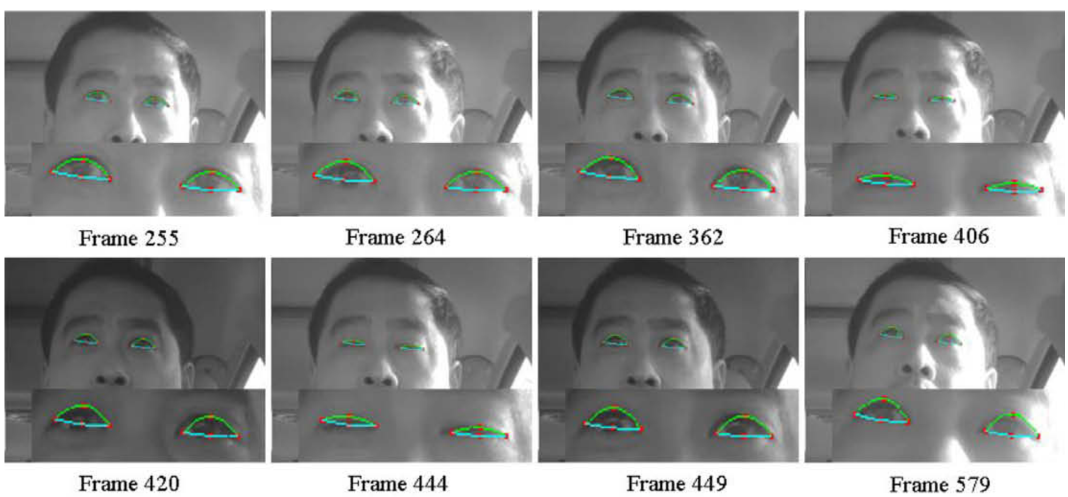


Fig. 12. An example of eyes tracking on indoor video under artificial light conditions.

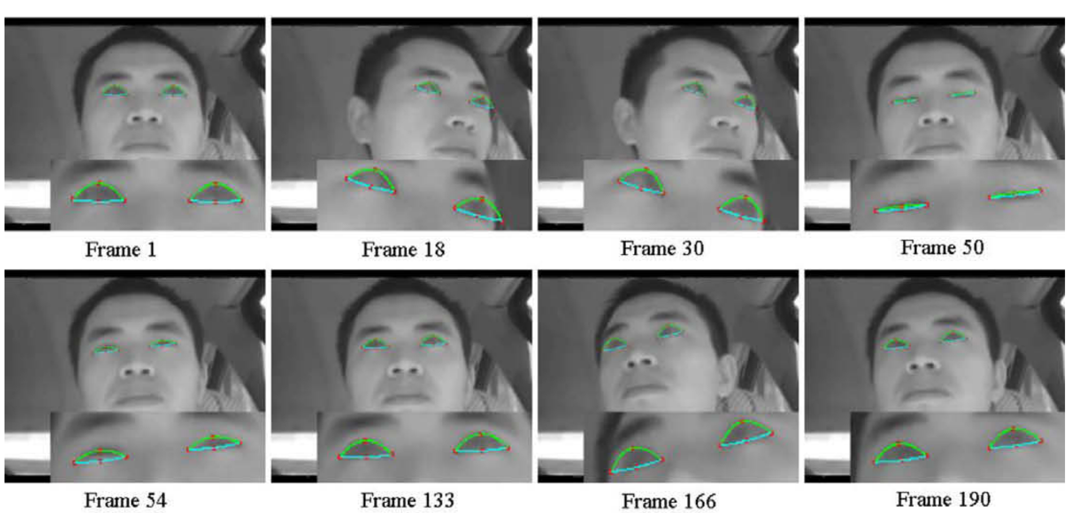
sampled points used in the measurement model is adaptively set in term of the particle parameters. Comparative results are shown in Fig. 7 to demonstrate the importance of ANSP technique. In Fig. 7a, the tracking results without applying ANSP technique are given. We find that the recovered eyelid distance values for open eye states are far from the actual values. In comparison, the results with applying ANSP technique are given in Fig. 7b. It is shown that the contour for both open and closed eyes is tracked precisely.



(a) An example of eyes tracking on outdoor video under normal natural light conditions



(b) An example of eyes tracking on outdoor video under light change conditions



(c) An example of eyes tracking on outdoor video with blurred eye region

Fig. 13. Examples of eyes tracking for outdoor videos.

5.1.3. Performance on collected videos

We also test the eye tracking method on collected indoor and outdoor videos. The indoor videos are recorded with USB camera under artificial light conditions while the outdoor videos are recorded for drivers with IR camera under natural light conditions. Generally, the width of single eye region is about 25–30 pixels. The left eye tracking results on an indoor video is presented in Fig. 8 and several tracking results on outdoor videos are presented in Fig. 9. As can be seen from Fig. 8, our method performs well even on the facial video with complex head movements (e.g. head turns, tilts and rotations) and appearance changes (e.g. with and without glasses). From Fig. 9, we can find the encouraging results on outdoor videos under distinct scenarios as normal light conditions (e.g. Fig. 9a), light change conditions (e.g. Fig. 9b) and blurred eye region (e.g. Fig. 9c).

5.1.4. Eye tracking accuracy

In order to evaluate the accuracy of our eye tracking approach, an image sequence with several head movements is used. The width of single eye region is about 32 pixels and the distance between eyelids for open eye is about 10 pixels. Besides, we define the midpoint between eye corners as eye location and the variation range is about 35 pixels in *Y*-direction and 100 pixels in *X*-direction. To obtain the required benchmark data, we manually calibrate the eye corners and apexes of eyelids and then calculate the eye locations as well as eyelids distance frame by frame. Also the parameterized tracking results are acquired. Fig. 10a illustrates the Root Mean Squared Error (RMSE) between the benchmark and tracked eye locations, and Fig. 10b shows the error between the benchmark and recovered eyelids distance. It can be seen from Fig. 10a that our method (blue curves with dots) can track the eye for more than 300 frames while Tan and Zhang's method (red curves with cross) fails to capture the eye within less than 100 frames. Meanwhile Fig. 10b shows that our method provides a more accurate detection rate of eye state. Both figures indicate that our method can track eye locations and obtain eye closure rate accurately.

5.2. Performance of eyes tracking

The method for double eyes tracking is presented above. Since it is similar to single eye tracking except for tiny difference in describing template, we also present several tracking examples here. The tracking results of eyes on Cohn-Kanade database (Kanade et al., 2000) are shown in Fig. 11. The results show that eyes under different facial expressions are well tracked. Besides, Figs. 12 and 13 illustrate tracking results of eyes on collected indoor and outdoor videos. Both figures indicate the efficiency of our eyes tracking approach under various light conditions. However, since our eyes deformable template is constructed based on the assumption that the two eyes are approximately symmetrical, the accuracy of eyes tracker is prone to decrease when the user swings his/her head in a drastic mode. Moreover, the eyes tracker will fail when nearly profile faces are captured. The experimental results show that our eyes tracking method works well within the yaw angle range from -30° to 30°.

6. Conclusions

This paper describes an algorithm framework for eye/eyes tracking. Unified deformable templates for both single eye tracking and double eyes tracking are presented. The particle filtering framework is introduced to track the eye/eyes contour. Dynamic model based on the template parameters is given while the existed measurement model for contour tracking is also modified. The experimental results on public database and collected videos also show that our proposed method can not only track the locations of eye/eyes accurately, but also recover the detailed eye contour parameters without extra post-processing.

However, some problems still need to be solved. Firstly, the proposed eye/eyes tracking method cannot deal with low resolution image sequences. The width of single eye is at least 20 pixels so that the eye contour can be extracted. Besides, the eyes tracker is prone to fail when the range of head motion from left to right becomes wider. Second, although our method can process the slightly blurred image sequences, the bad illuminated conditions still cause tracking failures. Lastly, more robust measurement model needs to be settled to improve the eye corner location. Future work will be done to tackle with those problems and obtain a more robust algorithm.

Acknowledgements

This work is supported by the National Basic Research Program of China (973 program) under Grant No. 2007CB311004 and the National High Technology Research and Development Program of China (863 program) under Grant No. 2009AA11Z214.

References

Bacivarov, I., Ionita, M., Corcoran, P., 2008. Statistical models of appearance for eye tracking and eye-blink detection and measurement. IEEE Trans. Consumer Electron. 54 (3), 1312–1328.

Cootes, T.F., Edwards, G.J., Taylor, C.J., 1998. Active appearance models. Lect. Notes Comput. Sci. 1407, 484.

Deng, J.Y., Lai, F.P., 1997. Region-based template deformation and masking for eye-feature extraction and description. Pattern Recognit. 30 (3), 403–419.

Dinges, D., Grace, R., 1998. PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. In: US Department of Transportation, Federal Highway Administration, TechBrief, FHWA-MCRT-98-006.

Gao, W., Cao, B., Shan, S., Zhou, D., Zhang, X., Zhao, D., 2005. The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations. <http://www.jdl.ac.cn/peal/index.html>.

Isard, M., Blake, A., 1998a. Condensation-conditional density propagation for visual tracking. Internat. J. Comput. Vision 29 (1), 5–28.

Isard, M., Blake, A., 1998. A mixed-state condensation tracker with automatic model switching. In: Internat. Conf. on Computer Vision, pp. 107–112.

Jin, X., Jiang, Z., Liu, J., Feng, H., 2005. Multiple states and joint objects particle filter for eye tracking. SPIE, Wuhan, China, Bellingham WA, WA 98227-0010, United States.

Kanade, T., Cohn, J.F., Tian, Y., 2000. Comprehensive database for facial expression analysis. In: Proc. 4th IEEE Internat. Conf. Automatic Face and Gesture Recognition (FG'00), Grenoble, France, pp. 46–53.

Lucas, B., Kanade, T., 1981. An iterative image registration technique with an application in stereo vision. In: The 7th Internat. Joint Conf. on Artificial Intelligence, pp. 674–679.

Tan, H., Zhang, Y.J., 2006. Detecting eye blink states by tracking iris and eyelids. Pattern Recognition Lett. 27 (6), 667–675.

Tian, Y.L., Kanade, T., Cohn, J., 2000. Dual-state parametric eye tracking. In: Proc. Internat. Conf. on Face and Gesture Recognition, vol. 2000, pp. 26–30.

Wu, J., Trivedi, M.M., 2008. Simultaneous eye tracking and blink detection with interactive particle filters. Eurasip J. Adv. Signal Process. 2008, 823695.

Yuille, A.L., Haallinan, P., Cohen, D.S., 1992. Feature extraction from faces using deformable templates. Internat. J. Computer Vision 8 (2), 99–111.

Zheng, Z., Yang, J., Yang, L., 2005. A robust method for eye features extraction on color image. Pattern Recognition Lett. 26 (14), 2252–2261.

Zhu, Z., Ji, Q., 2005. Eye gaze tracking under natural head movements. In: Proc. IEEE Comput. Soc. Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 918–923.

Zhu, Z., Ji, Q., Fujimura, K., 2008. Combining Kalman filtering and mean shift for real time eye tracking under active IR illumination. In: Proc. Internat. Conf. on Pattern Recognition, vol. 16, pp. 318–321.