ROBUST EYELID TRACKING FOR FATIGUE DETECTION

Fei Yang[1], *Xiang Yu*[1], *Junzhou Huang*[2], *Peng Yang*[1], *Dimitris Metaxas*[1]

[1]Rutgers University [2]University of Texas at Arlington

ABSTRACT

We develop a non-intrusive system for monitoring fatigue by tracking eyelids with a single web camera. Tracking slow eyelid closures is one of the most reliable ways to monitor fatigue during critical performance tasks. The challenges come from arbitrary head movement, occlusion, reflection of glasses, motion blurs, etc. We model the shape of eyes using a pair of parameterized parabolic curves, and fit the model in each frame to maximize the total likelihood of the eye regions. Our system is able to track face movement and fit eyelids reliably in real time. We test our system with videos captured from both alert and drowsy subjects. The experiment results prove the effectiveness of our system.

***Index Terms*—** fatigue detection, eyelid tracking

1. INTRODUCTION

Fatigue from chronic partial sleep deprivation, circadian misalignment, and work overload is a risk factor for people driving vehicles or performing critical tasks. People in fatigue exhibit certain visual behaviors observable from eyes and faces. Typical visual characteristics of fatigue include slow eyelid movement, smaller degree of eye openness, frequent nodding, sluggish in facial expression, sagging posture, etc [1]. Among them, tracking slow eyelid closures is one of the most reliable ways to detect lapse of attention [2].

Tracking eyelids to monitor fatigue is a challenging task. The challenge comes from the fact that the appearances of eyes are significantly different when the eyes are closed or half closed. A single template is insufficient to model eye appearances. Head rotations, shadows, reflection of glasses, motion blurs and image noises could also make this problem more difficult. Therefore, reliable eyelid tracking using optical cameras is still not resolved.

In this paper, we propose a new framework for robust eyelid tracking for fatigue detection. Instead of directly tracking the eyes, we first employ a robust face tracker, which could track the facial landmarks with various poses in real time. We use a deformable template to model the shapes of eyelids, and fit the model to maximize the total likelihood of the eye regions. We propose an efficient algorithm to fit the eye template and track the eyelids. Our system can successfully be used in fatigue detection.

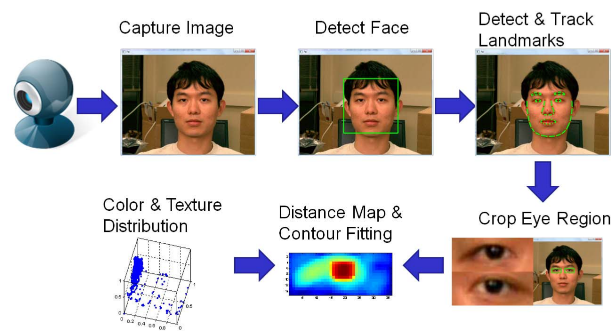

Fig. 1. System overview

The workflow of our system is shown in Fig. 1. We first detect face region from the image captured by a web camera. The facial features are detected and tracked by using our face tracker based on Active Shape Models. The eye regions are cropped, and a distance map is constructed by measuring the distance of each pixel to the distribution of the skin colors. We finally fit the deformable eye model to the distance map, which gives the positions of eyelids.

2. RELATED WORK

Eye tracking and blink detection have been studied extensively in recently years. The related work includes eye localization, blink detection, and infrared spectrum based systems.

Eye localization. Accurate eye localization is a key component of many computer vision systems. Previous methods localize eye centers by using the eye geometry [3], appearance features [4], and context information [5] [6].

Blink detection. Template based methods compare the eye appearance with a template trained from open eyes. Morris et al. [7] proposed a real-time detection system based on variance map and eye corners. Chau and Betke [8] developed a system by computing the correlation with an open eye template. These methods could only distinguish between two states (open or closed). The states in-between (e.g. half open) cannot be acquired. The performance of these systems also degrades with large head rotation.

The second group of methods uses statistical classifiers to detect eye closure. Pan et al. [9] developed a boosted classifier to detect the degree of eye closure. The changing of eye states is modeled by a Hidden Markov Model. Examples of typical eye motion are used for training the model, so that the classifier could handle partial eye closure.

Optical flow based methods have also been explored for blink detection. Sirohey et al. [10] proposed an approach for determining eye blinks by estimating the iris and eyelid motion using the normal flow. Their system is good in accuracy, but is not able to run in real-time. Divjak et al. [11] developed a system using optical flow for blink detection. Their system is able to work in realtime, since the optical flow calculation is offloaded to GPU.

Infrared spectrum based systems. Images captured in infrared spectrum can facilitate pupil localization and blink detection. Ji et al. [12] designed a real-time system to monitor driver's vigilance. They designed hardware to use active infrared illumination, and tracked eyes by combining the bright-pupil-based Kalman filter tracker with a mean shift eye tracker [13]. The degree of eye opening is characterized by the shape of the pupil. These systems requires special cameras or special illumination devices.

3. FACE TRACKING

Our face tracker is based on the Active Shape Models (ASMs) [14] together with a novel nonlinear shape subspace method to handle large head rotations. The need for this comes from the fact that feature shapes differ significantly across poses of varying tilt, pitch and yaw angles. The learned model allows the complex, non-linear region of the facial shape manifold to be approximated in a piecewise fashion, as a combination of smaller linear sub-regions. Each sub-region defines a hyper-ellipsoid on this manifold. Facial shapes of similar pose are constrained to lie in the same linear subspaces.

The mechanism for the facial shape search iteratively modifies the current shape by searching along the landmark points and simultaneously constraining the overall shape to lie on the shape manifold, so that the tracked shape is a valid face shape. To perform tracking in realtime, we employ KLT tracker [15] to track facial landmarks, then apply the shape model to constrain the positions of individual landmarks with the shape manifold. The algorithm automatically initializes and continues tracking automatically. The face shape is dynamically adjusted over time to fit the shape model to current target's appearance.

4. EYELID FITTING

The eye region given by the face tracker is refined by an deformable eye contour template. The template is defined as two parabolic sections intersecting at two points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$, as shown in Fig. 2(a). The center points bisecting the parabolic sections are denoted as $P_3(x_3, y_3)$ and $P_4(x_4, y_4)$. The objective is to deform the template to best fit the eye image. Since each parabola is uniquely defined by three points. The objective becomes to place four landmarks $P_1, ..., P_4$ to best fit the eye image.

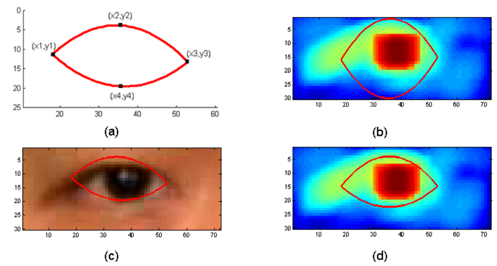

Fig. 2. Fitting eye contours. (a) Eye template. (b) Distance map and initial positions of the curves. (c) Curves with overlapping images. (d) Curves after a few iterations.

The likelihood of the eye region is defined as

$$L(\theta|I) = \int_{x=x_1}^{x_2} \int_{y=y_1(x)}^{y_2(x)} s(x, y)dydx \tag{1}$$

in which $s(x, y)$ is pixel-wise likelihood function. We compute the pixel-wise likelihood by clustering the color and texture descriptors for the first few frames, and then compute the Mahalanobis distances to the distribution of the skin pixels. The parabolic sections are defined as quadratic functions $y_i = a_i x^2 + b_i x + c_i, (i = 1, 2)$ with coefficients $\theta_i = [a_i, b_i, c_i]^T$. Let θ_1 be the upper section, and θ_2 be the lower section. To maximize the likelihood $L(\theta|I)$, we take its derivative to θ_i,

$$\frac{\partial L}{\partial \theta_i} = (-1)^i \begin{bmatrix} \int_{x_1}^{x_2} y_i(x)x^2 dx \\ \int_{x_1}^{x_2} y_i(x)x dx \\ \int_{x_1}^{x_2} y_i(x) dx \end{bmatrix}, (i = 1, 2) \tag{2}$$

A parabolic curve can be uniquely defined by three points located on it, in the form of the linear system:

$$\begin{bmatrix} y_{i_1} \\ y_{i_2} \\ y_{i_3} \end{bmatrix} = \begin{bmatrix} x_{i_1}^2 & x_{i_1} & 1 \\ x_{i_2}^2 & x_{i_2} & 1 \\ x_{i_3}^2 & x_{i_3} & 1 \end{bmatrix} \cdot \theta_i \tag{3}$$

Therefore, we have

$$\frac{d\theta_i}{dy} = A^{-1}, \text{ and } \frac{d\theta_i}{dx} = -A^{-1}B \tag{4}$$

where

$$A = \begin{bmatrix} x_{i_1}^2 & x_{i_1} & 1 \\ x_{i_2}^2 & x_{i_2} & 1 \\ x_{i_3}^2 & x_{i_3} & 1 \end{bmatrix}, \text{ and } B = \begin{bmatrix} 2x_{i_1} & 1 & 0 \\ 2x_{i_2} & 1 & 0 \\ 2x_{i_3} & 1 & 0 \end{bmatrix} \tag{5}$$

Therefore, the derivatives of L to x_j and y_j are

$$\frac{dL}{dy_j} = \sum_{i=1}^{2} \frac{\partial L}{\partial \theta_i} \cdot \frac{d\theta_i}{dy_j}, \quad \text{and} \quad \frac{dL}{dx_j} = \sum_{i=1}^{2} \frac{\partial L}{\partial \theta_i} \cdot \frac{d\theta_i}{dx_j} \tag{6}$$

In practice, since the two parabolic curves have 6 degrees of freedom (DOF), and the 4 landmarks have 8 DOF, we fix x_2 and x_4 to be the mean value of x_1 and x_3. The contour fitting algorithm is summarized in Algorithm 1.

Algorithm 1 *Fitting Eyelids*

Build likelihood map for the eye region.
Initialize template using $\mathbf{x}^0$ and $\mathbf{y}^0$.
repeat
    Update $y_j (j = 1, ..., 4)$: $y_j^{(k+1)} = y_j^{(k)} - c\frac{dL}{dy_j^{(k)}}$
    Update $x_j (j = 1, 3)$: $x_j^{(k+1)} = x_j^{(k)} - c\frac{dL}{dx_j^{(k)}}$
    and $x_2 = x_4 = 0.5(x_1 + x_3)$
until $\mathbf{x}$ and $\mathbf{y}$ converge.

5. EXPERIMENTS

We perform experiments on a fatigue database. In this database, 28 healthy adult subjects (14 males and 14 females) completed a 3-night controlled laboratory experiment and were randomized to either acute total sleep deprivation (0 hour in bed) or no sleep deprivation (9 hours in bed) on the second night. Subjects completed a 20-minute Psychomotor Vigilance Task (PVT) every 2 hours while awake. Images of the face were recorded during the PVT test by a digital camera mounted above compute monitor.

The screenshots of our system are shown in Fig. 3. The eye closure scores for the same subject is shown in Fig. 4. The top two figures show the scores when the subject is alert, and the bottom two figures show the scores when the subject is drowsy. The scores generated by our system are shown in blue curves, which closely match the ground truth (human annotations) shown in green lines. In addition, the scores clearly reveal two typical characteristics of fatigue: frequent blinking and long eye closure.

We perform quantitative evaluation of our method, using two videos with human annotated ground truth. In the first video, the subject is alert and performs a total of 167 blinks. In the second video, the subject is drowsy and performs 411 blinks in total. The hit rate and false detection rate are summarized in Table 1.

We also perform experiments to show the PERCLOS (percentage of slow eyelid closures), which is recognized to be the most valid ocular parameter for monitoring fatigue. As shown in Fig. 5. The top figure shows the scores of two videos captured of subject 1. In one video (shown in red) the subject

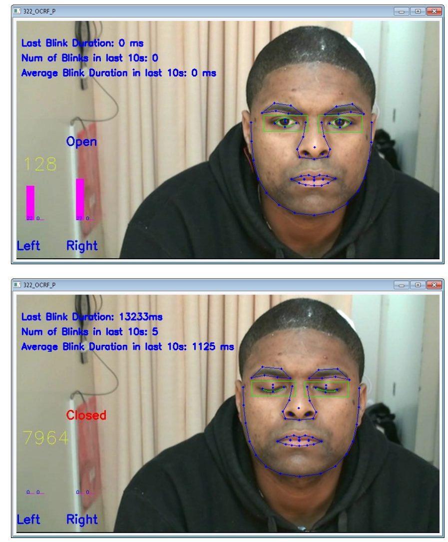

Fig. 3. Examples of the eyelid fitting. **Top:** when the eyes are open. **Bottom:** when the eyes are closed.

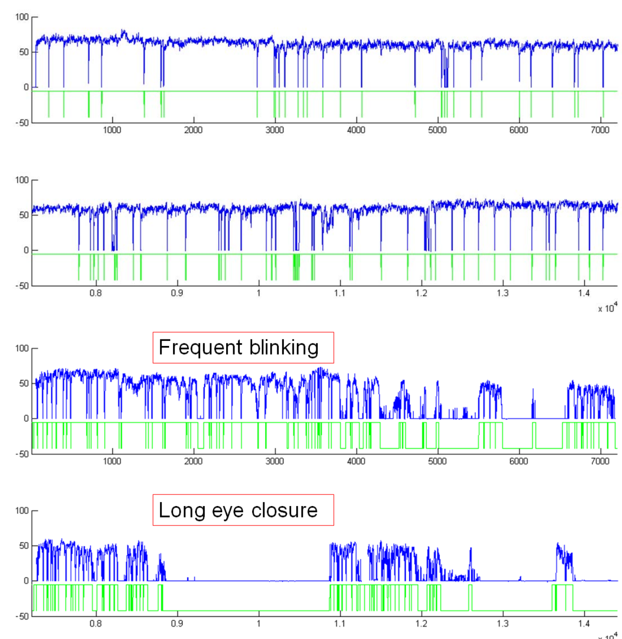

Fig. 4. Eye closure scores evaluated by our system (blue) and eye blinks labeled manually (green), for one subject in the database. **Top two:** when the subject is alert. **Bottom two:** when the subject is drowsy. Two typical characteristics of fatigue, frequent blinking and long eye closure are illustrated.

Table 1. Accuracy of blink detection

	Hit rate	False detection rate
Alert	94.0%	7.3%
Drowsy	87.3%	10.6%

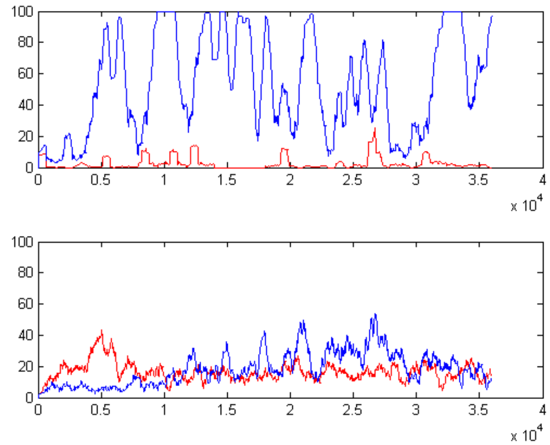

Fig. 5. PERCLOS scores of two subjects. **Top:** the scores of two videos of subject 1. In one video (shown in red) the subject is alert. And in the other video (shown in blue) the subject is very drowsy. **Bottom:** the scores of two videos of subject 2. The subject is drowsy in both videos.

is alert. And in the other video (shown in blue) the subject is very drowsy. The bottom figure shows the scores of two videos of subject 2. The subject is drowsy in both videos. The PERCLOS scores computed by our system also effectively detect fatigue from normal cases.

6. CONCLUSION

In this paper, we propose a new framework for robust eyelid tracking for fatigue detection. Instead of directly tracking the eyes, we employ a robust face tracker, which could track facial landmarks with large face rotations in real time. We proposed a deformable eye template, and propose an efficient algorithm to fit the template to the image. The experiments in this paper prove the effectiveness of our method.

7. REFERENCES

[1] Q. Ji and X.J. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real Time Imaging*, vol. 8, no. 5, pp. 357–377, 2002.

[2] D.F. Dinges, G. Maislin, J.W. Powell, and M.M. Mallis, "Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management," *Final report for the U.S. Department of Transportation, National Highway Traffic Safety Administration*, 1998.

[3] Alan L. Yuille, Peter W. Hallinan, and David S. Cohen, "Feature extraction from faces using deformable templates," *International Journal of Computer Vision*, vol. 8, no. 2, pp. 99–111, 1992.

[4] M. Hamouz, J. Kittler, J.K. Kamarainen, P. Paalanen, H. Kälviäinen, and J. Matas, "Feature-based affine-invariant localization of faces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 9, pp. 1490–1495, 2005.

[5] X.Y. Tan, F. Song, Z.H. Zhou, and S.C. Chen, "Enhanced pictorial structures for precise eye localization under uncontrolled conditions," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 1621–1628.

[6] F. Yang, J.Z. Huang, P. Yang, and D. N. Metaxas, "Eye localization through multiscale sparse dictionaries," in *the 9th Conference on Automatic Face and Gesture Recognition (FG)*, 2011, pp. 514–518.

[7] T. Morris, P. Blenkhorn, and F. Zaidi, "Blink detection for real-time eye tracking," *Journal of Network and Computer Applications*, 2002.

[8] M. Chau and M. Betke, "Real time eye tracking and blink detection with usb cameras," *Boston University Computer Science Technical Report No.2005-12*, May 2005.

[9] G. Pan, L. Sun, Z.H. Wu, and S.H. Lao, "Eyeblink-based anti-spoofing in face recognition from a generic webcamera," in *11th IEEE International Conference on Computer Vision*, 2007, pp. 1–8.

[10] S.A. Sirohey, A. Rosenfeld, and Z. Duric, "A method of detecting and tracking irises and eyelids in video," *Pattern Recognition*, vol. 35, no. 6, pp. 1389–1401, 2002.

[11] M. Divjak and H. Bischof, "Eye blink based fatigue detection for prevention of computer vision syndrome," in *the IAPR Conference on Machine Vision Applications*, 2009, pp. 350–353.

[12] Q. Ji, Z.W. Zhu, and P. Lan, "Real time non-intrusive monitoring and prediction of driver fatigue," *IEEE Trans. Vehicular Technology*, vol. 53, pp. 1052–1068, 2004.

[13] Dorin Comaniciu, Visvanathan Ramesh, and Peter Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proc. CVPR*, 2000, vol. 2, pp. 142–149.

[14] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models—their training and application," *Compututer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[15] Jianbo Shi and Carlo Tomasi, "Good features to track," in *Proc. CVPR*, 1994, pp. 593 – 600.