Video pupil tracking for iris based identification

W. Ketchantang[1,2], S. Derrode[1], S. Bourennane[1] and L. Martin[2]

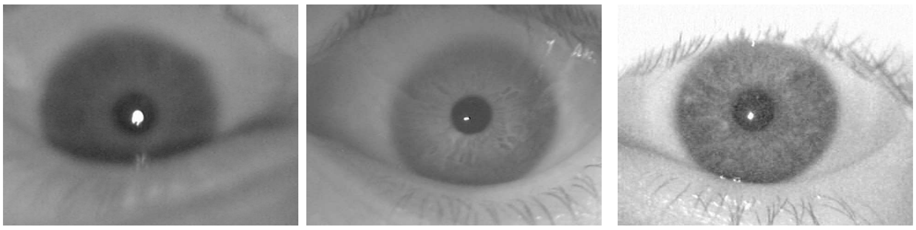
[1] Univ. Paul Cézanne, Institut Fresnel (CNRS UMR 6133),
Dom. Univ. de Saint Jérôme, F-13013 Marseille cedex 20, France
william.ketchantang@fresnel.fr
[2] ST Microelectronics, ZI Rousset BP 2, F-13106 Rousset, France
lionel.martin@st.com

Abstract. Currently, iris identification systems are not easy to use since they need a strict cooperation of the user during the snapshot acquisition process. Several acquisitions are generally needed to obtain a workable image of the iris for recognition purpose. To make the system more flexible and open to large public applications, we propose to work on the entire sequence acquired by a camera during the enrolment. Hence the recognition step can be applied on a selected number of the "best workable images" of the iris within the sequence. In this context, the aim of the paper is to present a method for pupil tracking based on a dynamic Gaussian Mixture Model (GMM) together with Kalman prediction of the pupil position along the sequence. The method has been experimented on a real video sequence captured by a near Infra-Red (IR) sensitive camera and has shown its effectiveness in nearly real time computing.

1 Introduction

Person identification from its iris is known to be one of the most reliable biometric technique, amongst face, fingerprint, hand shape, etc, based methods [1]. Iris texture is a relatively stable physical characteristic over years (and quite hard to falsify), that can be used even to guaranty high-level security access, by using a number of different iris signature [2,3,4].

Iris coding and comparison technics are relatively mature and show nice performances, in terms of both False Acceptance Rate (FAR) and False Rejection Rate (FRR). However, one major drawback in such iris identification systems comes from the eye acquisition procedure which needs a strict cooperation of the user in order to get a good quality image. Several snapshots of the iris are generally necessary to obtain a ready-to-process image. Examples in Fig. 1 show classical near IR iris acquisition problems obtained during enrolment, with blurred, noisy or defocused images. These problems are emphasized when using low-cost IR camera for large public applications. To increase the flexibility and to make the system more friendly, we propose to work on the entire sequence of images acquired by a camera during the enrolment, and to automatically select the "best workable image(s)" of the iris within the sequence before applying the recognition procedure. The "best workable images" are images showing a



(a) Blurred (b) Out-of-focus (c) Noisy

Fig. 1. Three kinds of degradations encountered in iris IR image acquisition.

clear iris (without degradations shown in 1) where partial occlusions provoked by eyelids and eyelash are small.

For this purpose, it is necessary to track the pupil along the sequence, in order to locate it efficiently and exactly, and to measure iris quality in the vicinity of the pupil. This problem is different from the -more studied- eye following problem in face sequence [5,6], typically encountered in driver attention level surveillance [7]. Such pupil sequences generally present images whose quality varies during time of acquisition, depending on the user motion:

- jerky motion of the user head resulting in quick and large scale translations of the pupil,
- translation of the user head along the optical axis resulting in both (i) scale variations of the pupil size and (ii) alternative focus and out-of-focus series of images.

Pupil tracking is a difficult problem also because of image degradations due to light variations and specular reflections in iris area (see white spots in the three images in Fig. 1). However, we can use some prior knowledge about the pupil which is approximatively circular and dark. In this context, the aim of the paper is to present a method for pupil tracking based on a GMM, which is dynamically and automatically updated along the time of acquisition. This method is combined with a Kalman filter to predict the pupil position in the next frame. The method has been experimented on a real sequence captured by an IR sensitive camera and has shown its effectiveness in nearly real time computing.

This paper is organized as follows. In section 2, we describe the principles of the algorithm based on GMM and Kalman prediction of the pupil modeled by a circle. Section 3 presents experimental results on an iris IR sequence and section 4 draws conclusion and perspectives.

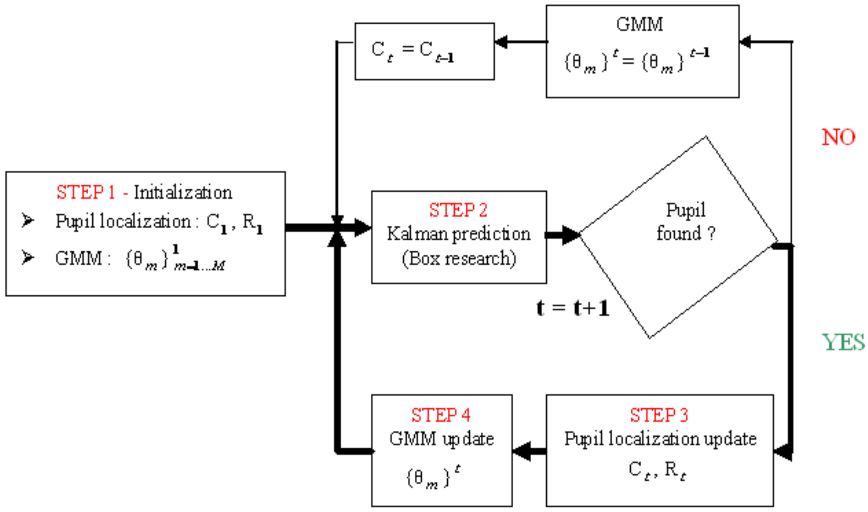


Fig. 2. Overview of the pupil tracking module.

2 Pupil tracking algorithm

This section is devoted to the description of the tracking method, whose principle is presented in Fig. 2. The pupil is modeled by a circle, with center $C(x, y)$ and radius R which vary along the sequence.

First, the pupil is detected and localized in the first frame, thanks to the Gradient Hough Transform (GHT) and the integro-differential operator detailed in [3]. The gray-level mixture of pixels inside the pupil is then modeled by a GMM whose parameters are learned by an EM algorithm (section 2.1). Pixels from white spots inside the pupil are thresholded and did not take into account in the mixture. In step 2, a Kalman filter is used to predict the box where the pupil has to be searched in the current image (section 2.2). If the pupil is not found in the predicted box, we simply copy both mixture and pupil parameters of previous image to the next image. Conversely, if the pupil is found in the box, we compute precisely the new pupil position C_t and size R_t (section 2.3), and update GMM parameters according to the new localized pupil at time t. This way the gaussian mixture of the pupil is updated dynamically along the sequence.

2.1 Parameters estimation of the GMM

The gray-level distribution of pixels inside the pupil is modeled by a mixture of M Gaussian distributions:

$$P(x/pupil) = \sum_{m=1}^{M} \alpha_m \, P_m(x \,|\mu_m, \sigma_m) \tag{1}$$

Hence, mixture parameters are defined by the M means μ_m, the M variances σ_m of the Gaussians $\mathcal{N}_m$ $(P_m(x/\mu_m, \sigma_m))$ and the M relative weights α_m of each Gaussian in the mixture; $\theta_m = \{\alpha_m, \mu_m, \sigma_m\}$ and $\sum_{m=1}^{M} \alpha_m = 1$. These parameters are estimated thanks to the well-known iterative Expectation - Maximisation (EM) method [8,9], whose update equations are given by:

$$
\begin{aligned}
\theta_m^{old} &= \theta_m^{new} \\
P\left(m \left| x_n, \theta_m \right.\right) &= \frac{\alpha_m^{old}\, P_m\left(x_n \left|\mu_m, \sigma_m\right.\right)}{\displaystyle\sum_{m=1}^{M} \alpha_m^{old}\, P_m(x_n \left|\mu_m, \sigma_m\right.)} \\
\alpha_m^{new} &= \frac{1}{N} \sum_{n=1}^{N} P\left(m \left| x_n, \theta_m \right.\right) \\
\mu_m^{new} &= \frac{\displaystyle\sum_{n=1}^{N} x_n\, P\left(m \left| x_n, \theta_m \right.\right)}{\displaystyle\sum_{n=1}^{N} P\left(m \left| x_n, \theta_m \right.\right)} \\
\sigma_m^{new} &= \frac{\sum_{n=1}^{N} \left(x_n - \mu_m^{new}\right)\, \left(x_n - \mu_m^{new}\right)\, P\left(m \left| x_n, \theta_m \right.\right)}{\sum_{n=1}^{N} P\left(m \left| x_n, \theta_m \right.\right)},
\end{aligned} \tag{2}
$$

where N is the number of pixels inside the pupil. This kind of model has been used successfully to track face or hand human skin in video sequences [10,11,7].

Thus, we get parameter values that maximise the pupil data likelihood. Due to local convergence of the EM procedure, good initial parameter estimates are important. For the first frame of the sequence, we implemented a K-means algorithm whereas, for other frames, we used parameter values estimated at previous frame, guarantying a convergence of EM in a few number of iterations since the pupil distribution is supposed to vary slowly between two consecutive images. Since the pupil area is dark and almost homogeneous, we take only two Gaussians. The number of EM and K-means iterations were respectively set to 10 and 5.

2.2 Kalman prediction of pupil position

Kalman filter [12] is used to predict the position of the pupil center C_{t+1} in next frame from its positions in previous frames. In general, the Kalman filter describes a system with a system state X_t and a measurement model C_t as follows

$$
\begin{aligned}
X_{t+1} &= A\, X_t + W_t \\
C_t &= H\, X_t + V_t
\end{aligned} \tag{3}
$$

where W_t and V_t denote respectively the model and measurement noises. They are supposed to be independent and their variance-covariance matrices are respectively Q and R. A and H are respectively the transition and measurement

matrices.
Kalman filter is one of the most popular estimation techniques in motion prediction because it provides an optimal estimation method for linear dynamic systems with Gaussian noise.

Filtering equations are given by

$$\begin{aligned} K_t &= P_{t/t-1}\, H_t^T\, (R + H_t\, P_{t/t-1}\, H_t^T)^{-1}, \\ P_{t/t} &= (I - K_t\, H_t)\, P_{t/t-1}, \\ X_{t/t} &= X_{t/t-1} + K_t\, \left(C_t - H_t\, X_{t/t-1}\right), \end{aligned} \tag{4}$$

where K_t and $P_{t/t}$ are respectively the Kalman gain and error covariance matrices at time t. Finally, prediction equations are given by

$$\begin{aligned} P_{t+1/t} &= A\, P_{t/t}\, A^T + Q, \\ X_{t+1/t} &= A\, X_{t/t}. \end{aligned} \tag{5}$$

As pupil moves slowly, we can assume that the first derivatives of position (x_t, y_t) is constant. Then, A and H are defined through

$$A = \begin{pmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \tag{6}$$

where ΔT is the time of frame acquisition ($\Delta T = 1/25s$) with as state vector $X_t(x_t, y_t, \dot{x}_t, \dot{y}_t)$ and measurement vector $C_t(\hat{x}_t, \hat{y}_t)$.

These equations allow us to predict the position of the pupil in next frame.

2.3 Pupil parameters

We now search for the real position of the pupil center C_t in a square box centered on the predicted position (predicted box : PB), with a side equal to $2\, R_{t-1}$. We should use the method employed to determine the position of the pupil in the first frame.
We prefer to use the gray level distribution of pupil pixels modeled by a GMM to compute the center of mass of pixels x_n $(C_t(\hat{x}_t, \hat{y}_t))$ with $P\,(x_n\,|\theta\,)$ lower than a given threshold.

$$P\left(j\left|\theta\right.\right)=\sum_{m=1}^{M}\alpha_m P_m(j/\theta_m) \tag{7}$$

$$\hat{x}_t=\frac{\sum\limits_{j\in PB} x_j\, P\left(j\left|\theta\right.\right)}{\sum\limits_{j\in PB} P\left(j\left|\theta\right.\right)}$$

$$\hat{y}_t=\frac{\sum\limits_{j\in PB} y_j\, P\left(j\left|\theta\right.\right)}{\sum\limits_{j\in PB} P\left(j\left|\theta\right.\right)}$$

If all pixels verify $P(x_n/\theta) < threshold$ then the pupil is considered lost in this frame, and tracking is delayed to the next frame.

As the pupil size also varies with time (due to translation of the head along the optical axis and to natural size variations), we also compute the pupil radius R_t to have a tracking robust to scale changes. The radius R_t is computed as follows

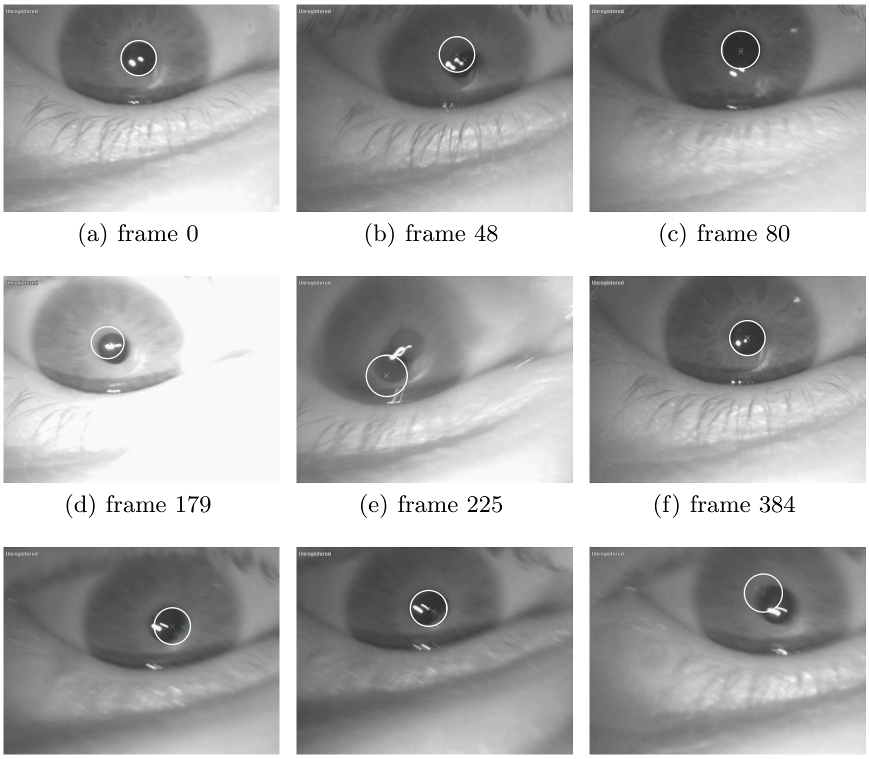
$$R_t=\underset{R}{Argmax}\left|\frac{\partial}{\partial c}\oint_c I(x,y)ds\right| \tag{8}$$

where c is the edge of the circle which gets as radius R, and $I(x,y)$ is the gray level of pixels belonging to c. R varies between R_{min} and R_{max}.

3 Pupil tracking results

The algorithm has been coded in C++, on a Pentium IV (2.6 GHz) PC platform and performs tracking at approximately $f \simeq 16$Hz. Experimental results in Fig. 3 show the pupil tracking algorithm in action along an IR video sequence. Globally, the method seems to be efficient and relatively robust to the image degradations mentioned in introduction. Indeed, our algorithm localizes correctly the pupil for not too fast eye motion, despite specular reflections (white spots). Even with out-of-focus images, the tracking algorithm manages to localize the pupil, provided that eye motion is slow (frame 475).

On the other hand, when pupil motion is quick and jerky, the algorithm does not localize it correctly (frames 48, 225 and 510) since the pupil position is not well predicted by Kalman filtering. This can be explained by the constant velocity hypothesis which is too strick and can not be assumed for all images in the sequence. When we take into account the velocity variations by integrating acceleration $(\ddot{x}, \ddot{y})$ in state vector, we verified experimentally that the performances and stability of the tracker decrease, because slow eye motion are not well predicted. In addition, our algorithm can loose the pupil when illumination variations are strong (frame 179). Indeed, an intensive illumination



(a) frame 0 (b) frame 48 (c) frame 80

(d) frame 179 (e) frame 225 (f) frame 384

(g) frame 466 (h) frame 475 (i) frame 510

Fig. 3. Pupil tracking results.

variation modifies considerably the pupil gray-level distribution, and the GMM is not able to update the model instantaneously.

An important point to note is that the algorithm is able to recover (frame 384) the pupil even if it was lost before (frame 225), as soon as the assumption on slow eye motion is verified.

4 Conclusion and perspectives

In this paper, we have presented a method for tracking a pupil in an IR iris sequence of images. The pupil geometry is modeled by a circle and the pupil distribution by a mixture of two Gaussians. The pupil position is predicted according to a Kalman filter. The GMM allows a computation of the pupil center which is robust to illumination change, and the Kalman filter is used to predict the area where to search the pupil in the next frame. The results we obtained are really encouraging. Nevertheless, we can think of making the tracking more robust by taking into account brief jump of velocity in the state evolution equation. Thus, the constant velocity assumption of the Kalman filter will be applied in slow pupil motion, and modified in fast eye motion parts of the sequence. By selecting the best workable image(s) of the pupil along the sequence, we can expect to improve the recognition rates and identification.

References

1. Ross, A., Jain, A.K.: Information fusion in biometrics. Pattern Recognition Letters **24** (2003) 2115–2125
2. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE trans. on PAMI **15** (1993) 1148–1161
3. Tisse, C.L., Martin, L., Torres, L., Robert, M.: Person identification technique using human iris recognition. In: 15th Int. Conf. on Vision Interface, Calgary, CA (2002) 294–299
4. Ma, L., Tan, T., Wang, Y., Zhang, D.: Efficient iris recognition by characterizing key local variations. IEEE trans. on Image Processing **13** (2004) 739–750
5. Sirohey, S., Rosenfeld, A., Duric, Z.: A method of detecting and traking irises and eyelids in video. Pattern Recognition **35** (2002) 389–401
6. Zhu, Z., Ji, Q.: Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. Computer Vision and Image Understanding **38** (2005) 124–154
7. Zhu, Y., Fujimura, K.: Driver face tracking using Gaussian mixture model. In: Proc. of the IEEE Intelligent Vehicles Symp. (2003) 587–592
8. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. of the Royal Statistical Society **39** (1977) 1–38
9. Bilmes, J.A.: A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. report tr-97-021, International Computer Science Institute, Berkeley, CA, USA (1997)
10. McKenna, S., Gong, S., Raja, Y.: Modelling facial colour and identity with Gaussian mixtures. Pattern Recognition **31** (1998) 1883–1892
11. Yang, J., Lu, W., Waibel, A.: Skin color modeling and adaptation. In: Proc. of the Third Asian Conf. on Computer Vision. Volume 2. (1998) 687–694 Lecture Notes In Computer Science - Vol. 1352.
12. Minkler, G., Minkler, J.: Theory and Application of Kalman filtering. Magellan Book Company, Palm Bay, FL, USA (1993)