

Eye-Movements as a Biometric

Roman Bednarik, Tomi Kinnunen, Andrei Mihaila, and Pasi Fränti

Department of Computer Science,
University of Joensuu,
P.O.Box 111, FI-80101 Joensuu, Finland
{bednarik,tkinnu,amihaila,franti}@cs.joensuu.fi

Abstract. We propose the use of eye-movements as a biometric. A case study investigating potentials of eye-movement data for biometric purposes was conducted. Twelve participants' eye-movements were measured during still and moving objects viewing. The measured data includes pupil sizes and their dynamics, gaze velocities and distances of infrared reflections of the eyes. For still object viewing of 1 second duration, identification rate of 60 % can be obtained by using dynamics of pupil diameters. We suggest an integration of the eye-movement-based identification into general video-based biometric systems.

1 Introduction

Biometric person authentication [1] refers to identifying or verifying persons' identity based on their physical and/or behavioral (learned) characteristics. Examples of physical biometrics are *fingerprints* and *facial images*, and examples of behavioral biometrics include *voice* and *signature*. There are numerous applications of biometrics, including forensics, access control to physical facilities, border control, identity verification in e-commerce, and personalizing user profiles in a mobile device.

A perfect biometric should be unique, universal, and permanent over time, easy to measure, cheap in costs, and have high user acceptance. No single biometric fulfills all these requirements simultaneously. For instance, fingerprints and retina are known to be highly unique, but they require dedicated sensors and are not user friendly. On the other hand, voice and facial geometry are not as unique, but they require only a cheap microphone or a camera as a sensor, and they are unobtrusive. Numerous studies have demonstrated that the combination of several complementary biometrics can provide higher recognition accuracy than any individual biometric alone [2,8].

Biometric authentication tasks can be subdivided into *identification* and *verification* tasks. In the identification task, an unknown biometric sample is compared to whole database of known individuals, and the best matching template is selected (*1:N matching*). On the other hand, the verification task, or *1:1 matching*, consists of verifying whether the provider of the biometric sample (*claimant*) is the one who (s)he claims to be. In both cases, a "no decision" option is also possible. Verification systems are well-suited for applications having high security requirements, but poorly suited for cases when user friendliness has higher priority. In verification, the user is required to give the identity claim as well as a biometric sample. In identification, the authentication process can be ubiquitous and the user does not even know that he is being authenticated.

Person's eyes provide several useful biometric features both for the high security and user convenient applications. First, iris and retinal patterns are known to be among the most accurate biometrics. Second, eyes are often used as anchor points in geometrical approaches for face detection [10], and the geometric differences (e.g. distance between eyes) can be utilized directly as features in face recognition [3].

In this paper, we propose to use *eye movements* as an additional biometric that can be integrated with other biometrics. To our knowledge, eye-tracking systems have not been considered as a possible solution for a biometric system. The main goal of the present paper is therefore to investigate the potential of eye-tracking as a biometric, as we had no preconceived hypothesis about whether and how the features of eye-movement signal are discriminative or not. This paper reports on a case study conducted by using Tobii ET-1750 eye-tracker in laboratory conditions.

The rest of this paper is organized as follows. In Section 2, we briefly review the current technology for the eye-movement tracking, and discuss its usefulness for biometric application. In Section 3, we consider different alternatives for potential features extracted by the eye-tracker, and describe the feature extraction and classification methods chosen. Experiments are carried out in Section 4, and results reported in Section 5. Discussion is in Section 6, and conclusions are drawn in Section 7.

2 Eye-Movement Tracking

Humans move their eyes in order to bring an image of inspected object onto fovea, a small and high-resolution area of the retina. Once the image of the object is stabilized on the retina, the information can be extracted. *Eye-tracker* is a device that records these movements. Most of the current eye-trackers use infrared light emitters and video image analysis of the pupil center and reflections from cornea to estimate the direction of gaze, see Figure 1. The infrared reflections are seen as the points with high intensity inside subject's pupil.

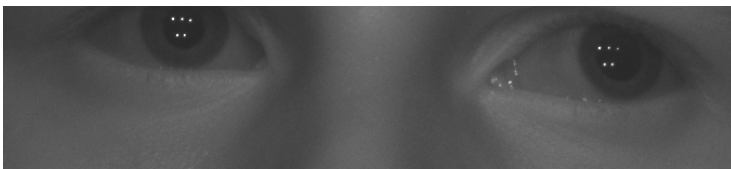


Fig. 1. Example image from eye-tracker's camera. The bright spots in the middle of the eyes are infrared illumination used by the eye-tracker device

The accuracy of current commercially available eye trackers ranges around 1 degree, while the data is sampled at rates of 50–500Hz. Modern eye-trackers are relatively cheap and able to reliably and unobtrusively collect the gaze data. The usability of eye-tracking is high in controlled laboratory conditions. Their application in real-life situations, however, is still limited due to the need for calibration of each of the users before the recording. An example of such system is shown in Figure 2.

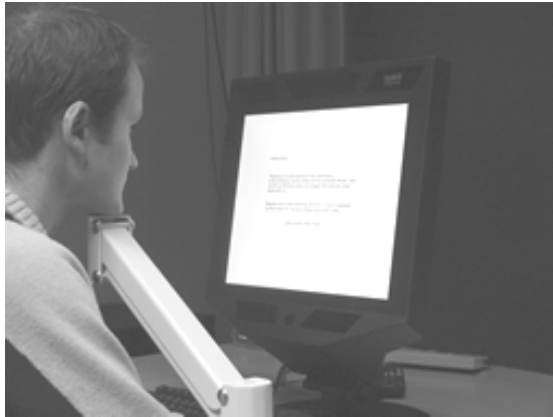


Fig. 2. Eye-tracker used in the experiments

Previous research has established a solid evidence about the relation between eye movements, visual attention and underlying cognitive processes [6,7]. Knowing which objects have been visually inspected and in which order and context, one can attempt to infer what cognitive processes were involved to perform a task related to these objects. However, to our knowledge, no attempts were made in order to distinguish individuals based on the properties of eye-movements seen as time-signals.

Eye is never perfectly still. Similarly, pupil diameter is never constant, but oscillates around certain value. This fact is used by the iris biometric systems to enforce the measurement of true authentic eyes. It could be hypothesized, that the mass of the eye-ball, the muscles responsible for movements of the eyes, and the muscles controlling the dilatation of the pupil are anatomically individual. In other words, such a complex system can exhibit high degree of uniqueness. Although the eye-tracker does not provide the direct measures of the muscular systems, the overt movements of eyes and especially the involuntary movements could be thought to reflect the underlying anatomical organization.

Current eye-trackers satisfy most of the desired features for an unobtrusive biometric. In the present eye-tracking systems, the detection of the position of eyes is based on video-imaging, the same technology used in face recognition. Therefore, there is a possibility to join such biometric systems into one, eye-movement based and face-feature based. With a high-resolution sensor the systems could be further combined with iris recognition systems.

The current eye-movement trackers are still too costly, but as the price of the technology becomes lower, wider inclusion into current computer and video-systems is expected. It will, therefore, be possible to simultaneously use all eye-movement measures in combination with other biometrics.

3 Person Identification Using Eye-Movements

The eye-tracker employed in the present study provides several measurements for both eyes. First, the eye-tracker outputs normalized coordinates within the camera's

view, as well as estimated gaze-direction measurements. The latter ones are computed through an interpolation between calibration points, whereas the former ones are the coordinates of the infrared light reflections on the cornea of an eye. The device provides also pupil diameters and the distance of the subject from the monitor.

We consider the following features as the candidates for a biometric cue:

- Pupil diameters (ϕ_L, ϕ_R)
- Distance between the reflections $d(L,R)$
- Velocities of gaze (v_L, v_R)
- Delta pupil diameters ($\Delta\phi_L, \Delta\phi_R$)

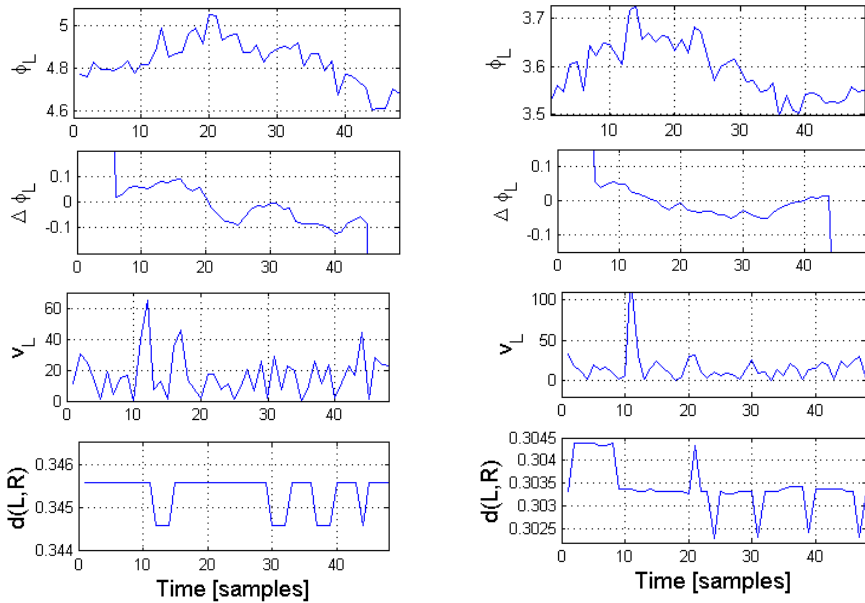


Fig. 3. Examples of (top to bottom) left pupil diameters, delta pupil diameters, velocity of the left eye and distance between the reflections, for subjects 3 (left) and 9 (right) in task C1

Pupil diameters are directly provided by the eye-tracker, while the distance, the velocities and the delta pupil diameters are computed from the raw measurements. Distance between the infrared reflections $d(L,R)$ (hereafter referred to as distance) is computed as the Euclidean distance between the two coordinates (x_L, y_L) and (x_R, y_R) within eye-tracker’s camera view. The velocities are calculated by estimating the first time derivative of the Euclidean distance between two consecutive points. The delta pupil diameters are computed using linear regression. Examples of eye-movement data collected in the present study are shown in Figure 3.

3.1 Feature Extraction

We consider the time track of a single measurement as a vector $\mathbf{x}=(x_1, x_2, \dots, x_T)$ of dimensionality T . This vector is reduced to a smaller dimensional space using the following approaches:

- *Fourier spectrum* (FFT) [5]
- *Principal components analysis* (PCA)[4]
- FFT followed by PCA (FFT+PCA)

We utilize a fast Fourier transform for a Hamming-windowed sequence, where the purpose of windowing is to suppress the spectral artifacts arising from the finite length effects. We retain the magnitude spectrum for further processing. As an example, the magnitude spectrum applied to pupil diameter data captures slow and fast variations of the pupil size that are encoded in the lower and higher frequencies, respectively.

The Principal component analysis (PCA) is a widely used dimensionality reduction method. It maps the data onto the directions that maximize the total scatter across all classes. In this study, we apply PCA directly to the raw data as well as to its Fourier spectrum. For PCA, we retain the directions of the largest variance as measured by the leading eigenvalues.

3.2 Classification and Fusion

In this study, we limit our experiments to the identification task for simplicity. We consider the different features both individually and in combination, and apply the k -nearest neighbor classification using Euclidean distance. For fusion we combine the individual distances by their weighted sum. We use *leave-one-out cross-validation* to estimate the error rates [4].

4 Experimental Setup

For the experiments, we collected a database of 12 volunteering participants (1 female, 11 male) recruited from research and teaching staff from the authors' department. All participants had normal or corrected-to-normal vision. Three of the participants were wearing eyeglasses.

We used the remote, binocular Tobii ET-1750 eye-tracker (Fig. 2), sampling at the rate of 50Hz. The eye-tracker is built into a 17'' TFT panel so no moving or otherwise disturbing parts can be seen or heard. In a pilot experiment, we noticed that the accuracy was degraded when participants moved their head. These caused some inaccuracies and additional reflections, for instance from eye-glasses. Therefore, we constructed a stand where the subjects rested their chin. The chin-rest was mounted approximately 80 centimeters from the screen, centered in the middle. The accuracy of the eye-tracker was greatly improved by fixing the head position.

4.1 Procedure

Experiments were conducted in a quiet usability laboratory with constant light conditions. The subjects were not informed that the data will be used for a biometric study. Instead, they were told that the data is needed for calibrating a new eye-tracker. Before a recording, a required automatic calibration procedure had to be conducted. If needed, the calibration was repeated to achieve the highest possible accuracy.

After a successful calibration, the participants were instructed to follow the instructions appearing on the screen. The test consisted of several different tasks (see Fig. 4), including text reading (T1, T2), tracking of a moving red cross (M) for 15.9 seconds, and watching a static gray-scaled image (I). After each of these stimuli, a cross was displayed on the middle of the screen (C1, C2, C3, C4) for 1 second. Competitions of the reading tasks were signaled by participants pressing a key, the time to view the static image was not restricted. Completion of the whole task took less than 5 minutes per participant.

4.2 Data Preprocessing

To preprocess the data, erroneous measurements were first removed. These consisted mostly from blinking, and in few cases from unintentional head-movements. After a preliminary experimentation, we realized that the gaze measurements were much noisier than the raw camera data, so we decided to keep only the camera-based measurements. An explanation of this might lie in inaccuracies created during the calibration procedures.

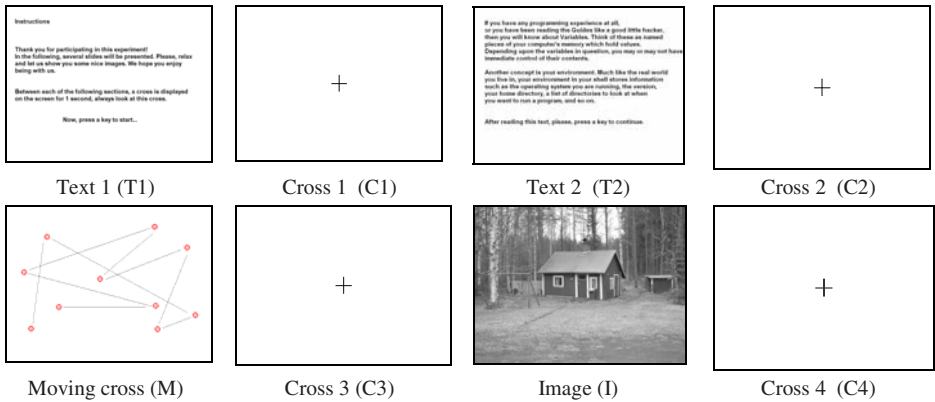


Fig. 4. Tasks and their order

5 Results

We restricted our experiments to the static crosses tasks (C1-C4) for two reasons. First, as we were interested in studying the inter-person variations, we should factor out the possible influence of the task. Second, this task contained approximately the same amount of data per person, which made the data compatible. For the leave-one-out cross-validation procedure, one of the four tasks was classified at a time using a 3-nearest neighbor classifier; the rest three samples were acting as the representatives for the class. Table 1 reports the number of data vectors recorded for each of the participants during all of the tasks in this experiment.

Table 1. The number of data vectors for each of the tasks

Subject: Task ID:	Text		Static cross				Moving cross	Image	Mean
	T1	T2	C1	C2	C3	C4	M	I	
1	643	1201	41	41	28	43	752	1057	423
2	384	936	44	44	50	42	744	549	311
3	546	1146	50	50	49	50	775	455	347
4	534	794	50	48	49	36	748	783	338
5	561	1253	48	49	49	51	701	436	350
6	626	1414	50	50	50	50	787	624	406
7	623	1292	45	41	35	38	779	149	334
8	843	1480	50	49	50	50	771	726	447
9	997	1777	49	50	50	49	790	501	475
10	532	1219	47	44	44	44	741	258	327
11	577	1090	44	44	49	50	782	451	344
12	537	878	50	50	49	50	754	375	306
Mean	617	1207	47	47	46	46	760	530	

5.1 Individual Features

As we were interested to investigate whether there were any individual differences in eye-movement dynamics, we had to remove the static properties from the signal. Thus, for a comparison, we created a *static* user template by taking the time averages for each subject. As long-term statistics, these were expected to carry the information about the physiological properties of the subject's eyes. Figure 5 shows the time-averaged features with their 95 % confidence intervals.

The *dynamic* user templates were formed by considering the time signal as a feature vector as explained in Section 3.1. The identification rates for both the static and dynamic cases are reported in Table 2. For the dynamic features, we studied both the original and mean-removed signals. By removing the long-term mean, we expected the features to discriminate mainly based on their dynamics.

Table 2. Identification rates (%) for single category of features

Method	Eye	Static	Dynamic					
		Mean	FFT		PCA		FFT + PCA	
			Original	Mean removed	Original	Mean removed	Original	Mean removed
Pupil	L	31	38	8	32	15	38	8
	R	33	33	13	38	8	33	13
	L+R	38	38	19	46	15	38	17
Delta Pupil	L	4	42	56	44	48	42	56
	R	4	46	56	50	50	46	56
	L+R	8	50	54	50	60	50	54
Velocity	L	17	19	10	15	13	19	10
	R	25	10	6	8	10	10	6
	L+R	21	13	6	13	15	19	8
Distance	-	83	90	6	83	4	90	8

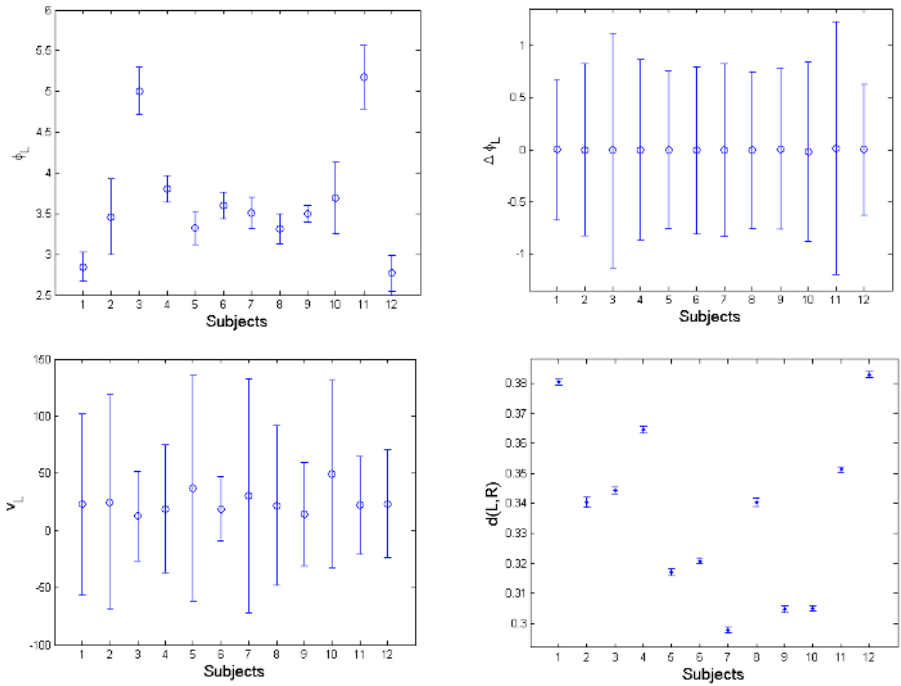


Fig. 5. Time-averaged features with their 95 % confidence intervals

5.2 Fusion of the Features

By combining complementary features, we expected to improve the accuracy further. After preliminary experiments, we decided to select the combinations as shown in Table 3. The fusion weights were set with an error-and-trial procedure by taking into account the individual performances.

Table 3. Identification rates (%) for fusion

Method Features	Fusion Weights	FFT	PCA	FFT + PCA
Pupil + Velocity	0.90 / 0.10	42	42	42
Delta Pupil + Velocity	0.90 / 0.10	46	44	46
Pupil + Distance	0.04 / 0.96	90	90	90
Velocity + Distance	0.10 / 0.90	83	83	83
Pupil + Velocity + Distance	0.10 / 0.05 / 0.85	92	88	92

6 Discussion

From the static features, the distance between the eyes shows high discrimination (90%) as expected. The distance can be accurately measured by an eye-tracker and used as a feature for biometric. However, similar measurement could be also obtained from a picture taken by a regular camera as well, as it does not include any dynamics of the eye-movement.

Static features and the original data (without mean removal) perform better compared to the eye dynamics. Nevertheless, the mean-removed features are above the chance level (8.35 %) in most of the cases. From the dynamic features, the delta pupil shows the best performance of 50-60 % and the velocity of the eyes 6-15 %. A possible explanation can be that the stimuli were static and displayed only for 1 second. Considering the feature extraction, FFT and PCA performed equally well and their combination did not improve accuracy further.

Considering the fusion of the features, no further improvement was achieved. The distance provided already an accuracy of 90 %, and dominated the results. The identification rate of the fusion for the dynamic features was around 40-50%. Given that the number of subjects was low, no statistically significant weight selection could be done.

7 Conclusions

This paper presents a first step towards using eye-movements as a biometric. We have conducted a case study for investigating the potential of the eye-tracking signal. The distance between eyes turned out to be the most discriminative and stable measurement, yielding identification rate of 90 % for twelve subjects. However, this feature could be measured without an eye-tracking device, and it does not truly reflect the behavioral properties of the eyes. The best dynamic feature was the delta pupil size (60 %), which corresponds to the variation of the pupil size in time. Interestingly, the pupil size itself provides rather weaker discrimination (40 %). Combination of different features gave only marginal improve in accuracy of identification.

In summary, the results indicate that there is discriminatory information in the eye-movements. Considering that both the training and test signals had the duration of 1 second only, the recognition accuracy of 40-90 % can be considered high, especially taking into account the low sampling rate (50 Hz). We expect improvements of the proposed system by having longer training and/or testing data, as well as a higher sampling rate.

As the eye-tracking systems are expected to become more widely available, we expect that they can be integrated into general video-based systems. In our future studies, we intend to create a larger database of recordings and tasks, and to study the inter-session variability of eye-movements as biometric features, i.e. the stability of the features over time. Other important future considerations are the dependence of the recognition accuracy on the task and the proper weight selection for fusion. Ultimately, we should have a feature that is consistent over different tasks.

References

1. R. Bolle, J. Connell, S. Pankanti, N. Ratha, and A. Senior, *Guide to Biometrics*, Springer, 2004.
2. R. Brunelli, D. Falavigna, Person identification using multiple cues, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17 (10): 955-966, October 1995.
3. R. Brunelli, T. Poggio, Face recognition: feature versus templates, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15 (10): 1042-1052, October 1993.
4. R. Duda, P. Hart and D. Stork, *Pattern Classification*, 2nd edition, Wiley Interscience, New York, 2000.
5. E.C. Ifeachor and B.W. Lewis, *Digital Signal Processing - a Practical Approach* (2nd edition), Pearson Education Limited, Edinburgh Gate, 2002.
6. M.A. Just and P.A. Carpenter. Eye fixations and cognitive processes, *Cognitive Psychology*, 8: 441-480, 1976.
7. K. Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124: 372-422, 1998.
8. A. Ross, A. Jain, Information fusion in biometrics, *Pattern Recognition Letters* 24, 2115-2125, 2003.
9. P. Verlinde, G. Chollet, M. Acheroy, Multi-modal identity verification using expert fusion, *Information Fusion* 1 (1): 17-33, Ed. Elsevier, 2000.
10. M. Yang, D. J. Kriegman, N. Ahuja, Detecting faces in images: a survey, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24 (1): 34-58, January 2002.