A Study on Human Reflex-Based Biometric Authentication Using Eye-Head Coordination

A Study on Human Reflex-Based Biometric Authentication Using Eye-Head Coordination

Yosuke Takahashi[1], Masashi Endo[1], Hiroaki Matsuno[1], Hiroaki Muramatsu[1], Tetsushi Ohki[1], and Masakatsu Nishigaki[1]
[1] Shizuoka University, 3-5-1 Johoku, Naka, Hamamatsu, Shizuoka, 432-8011 Japan
E-mail: nisigaki@inf.shizuoka.ac.jp

Abstract. Biometric information can be easily leaked and/or copied. Therefore, the biometric information used for biometric authentication should be kept secure. To cope with this issue, we have proposed a user authentication system using a human reflex response. It is assumed that even if people know someone's reflex characteristics, it is difficult to impersonate that individual, as anyone cannot control his/her reflexes. In this study, we discuss a biometric authentication system using eye-head coordination as a particular instance of reflex-based authentication. The availability of the proposed authentication system is evaluated through fundamental experiments.

1 Introduction

The application of biometric authentication is popular because there is no possibility of forgetting or losing the information used for authentication. However, biometric information, such as fingerprint or iris pattern data, can be easily leaked, leading to the threat of impersonation by unauthorized people [1], [2], which is a large drawback considering security. To overcome such problems, biometric authentication techniques that use the following types of biometric information have been proposed:

(1) Biometric information that is difficult to leak
(2) Behavioral biometric information

Authentication systems that use biometric information (1) have already been put into practice, for example, authentication using the vein of a finger or the palm [3], [4]. In addition, information such as handwritten signatures are used in authentication systems that utilize biometric information (2) [5].

However, there are threats, such as hidden sensors or phishing websites, that can steal the biometric information of a user under unexpected circumstances. Thus, it cannot be denied that biometric information can be extracted illicitly, even if we use biometric information (1), which is considered to be difficult to leak under normal conditions. Similarly, if we use behavioral biometric information (2), which is based on volitional behaviors such as a handwritten signature, we can consider various threats, such as an impostor practicing the handwriting of an authorized user to master it.

There might be a demand for biometric authentication that is impossible to impersonate someone even when biometric information has been leaked, and cannot be mastered by practice or other means, particularly for situations such as authentication when accessing highly confidential information. A biometric reflex type of authentication has been proposed as an authentication method that could meet these requirements [6], [7]. Biometric reflexes provide involuntary biometric information that human beings find difficult to control; thus, if an impostor knows about them, we anticipate it would be difficult for him to impersonate a legitimate user.

However, there are not sufficient differences in certain reflexes (saccade response and pupillary light reflex) between different individuals, and hence it needs to be combined with physiological biometric information (blind spot) that differs for individual person, to enable its use in authentication [6], [7]. This study discusses a reflex with sufficient difference between individuals to realize a biometric authentication using human reflex itself. We use eye-head coordination as a particular instance for reflex-based authentication.

2 Authentication Based on Human Reflexes

Reflexes are involuntary responses that occur in muscles and other parts of the body when sensory organs are stimulated by external actions; they are always of a predetermined form and are expressed automatically, mechanically, and momentarily [8]. In this study, we define reflexes as bodily responses that human beings cannot control consciously.

The existence of various different reflexes has been confirmed in human beings and it is considered that such responses can be utilized for authentication, provided there is some degree of difference in response between different individuals. For example, it is known that when the intensity of light entering an eyeball increases suddenly, the pupil contracts as a reaction to the light [9].

Human reflexes are responses that are considered difficult for people to control consciously. Thus, it is predicted that even if authentication information, such as 'the pupil of user *P* contracts by *Q*% in response to a certain light stimulus *R*,' is leaked, a user other than user *P* would find it difficult to imitate (or master by practice) the human reflexes inherent to user *P* (the *Q*% contraction of the pupil with respect to that stimulus *R*). In other words, in user authentication based on differences in human reflexes between individuals, the biometric information should be something that is difficult for another person to imitate, even when the biometric information has been compromised.

In existing authentication methods, pupillary light reflexes are used only for liveness detection [10], which simply detects if there is a response; thus, there have been alarming reports that a colored contact lens with printed false iris patterns could pass through an iris authentication device [11]. Moreover, pupillary light reflex cannot be stably obtained because it varies with psychological factors, such as stress [12]. Note that the proposed biometric authentication uses the differences in human

reflexes between individuals, and we make a clear distinction between that and the concept of liveness detection that simply detects presence of a response.

In this study, we propose a user authentication method that is based on the differences in human reflex itself. It is expected that spoofing will become further difficult, even when the biometric information has been leaked.

3 Authentication Method Using Eye-Head Coordination

3.1 Concept

As previously mentioned, in the existing human reflex-based authentication [6], [7], there are not sufficient differences in certain reflexes and therefore physiological biometric information is combinatorically used. In this study, we discuss a biometric authentication system using eye-head coordination as a particular instance of reflex-based authentication and evaluate the availability of the user authentication using human reflex response itself.

The eye-head coordination is a reflexive combination of eye and head movements when a human eye tracks a visual stimulus. It is known that "the ratio of sight line angle to head movement angle" is different for each subject. In the existing studies [13], [14], they use multiple sensors to calculate this feature. However, the use of such a rich sensor environment for authentication limits its application for the authentication method and is not realistic.

In this study, we use only the video of the eye taken by a corneal imaging camera shown in Fig. 1 and calculate the feature by the position of the eyeball when the user gazes at a stimulus. Consequently, the cost of authentication is reduced and the convenience of the user can be improved. It is expected that the application of this authentication method will also expand.

Fig. 1. Corneal imaging camera.

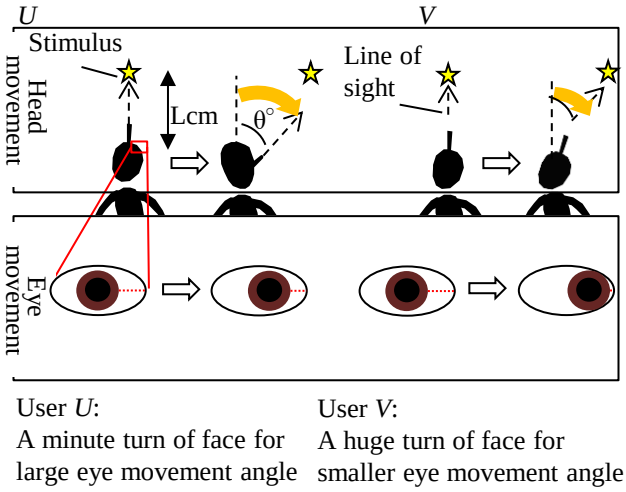

Fig. 2. Eye head coordination.

3.2 Feature Extraction

In this study, we apply "the ratio of sight line angle to head movement angle," which is called "the sharing rate," to the biometric authentication system. When a person gazes at a stimulus, the line-of-sight is represented by the vector sum of the eye and head movements. Therefore, when a user gazes at a stimulus, which is displayed to the right or left with the angle of θ degrees from the front, the angle of θ degree is the sum of the user's face and eye movement angles. Thus, when the user turns his/her face largely towards the stimulus, the viewing angle of his/her eyes is small and vice versa (Fig. 2). Therefore, by measuring the position of the cornea when gazing at the stimulus, it is possible to calculate the characteristic corresponding to "the sharing rate." Thus, we briefly calculate "the sharing rate" from only the video of the user's right eye taken by the corneal imaging camera.

3.3 Authentication Procedure

During the experiment, the user wears a corneal imaging camera (Fig. 1) for the detection of the position of the cornea. The display was set so that the visual distance from the right eye of the user to the stimulus was about L cm, and the visual angle for the stimulus from the front of the user was about θ° (Fig. 2). The stimuli are displayed at the height of the eyes of each user. Our prototype system works as follows:

1. A system displays the initial gaze target (stimulus S) at the front of a user.
2. The user is instructed by the system to follow the stimulus with his eyes during the authentication.

3. Further, the system displays the stimulus at right (stimulus S'_R) or left (stimulus S'_L) randomly. The dynamic movement of the right eye of the user is captured in the video with the corneal imaging camera.
4. For each frame of the video obtained in step 3, the time series data calculated by measuring the position of the cornea is regarded as the feature.
5. If the features extracted in the authentication phase are close enough to the features extracted in the enrollment phase, the user is authenticated as the legitimate. The "closeness" used in this study is described in the following section.

3.4 Verification

We define the similarity between two data by distance, which is determined by the dynamic time warping (DTW) algorithm [15], and calculate the similarity between legitimate users as well as that between a legitimate user and others. The DTW algorithm measures similarity between two temporal sequences. The DTW algorithm produces an intuitive similarity measure, allowing similar shapes to match even if they are out of phase in the time axis.

In this system, the DTW score is obtained by comparing two sequences $X = (x_1,x_2,\ldots,x_\alpha)$ and $Y = (y_1,y_2,\ldots,y_\beta)$. A warping path is a sequence of grid points $F = f_1,f_2,\ldots,f_K$ on the $\alpha \times \beta$ plane. Let the distance between the two values, x_{ik} $(1 \le i \le \alpha)$ and y_{jk} $(1 \le j \le \beta)$, be $d(f_{(k)}) = | x_{ik} - y_{jk} |$. Further, we calculated the matching score by using DTW(X,Y), where X is the time series data of the enrollment phase and Y is the time series data of the authentication phase, which is calculated by the following equation:

$$\mathrm{DTW}(X,Y) = \frac{min(\sum_{k=1}^{K} d(f(k)))}{k} \quad (1)$$

.

The closer the score is to zero, the higher the similarity is. DTW(X,Y) < t, where t is a threshold for authenticating legitimate users.

Considering that the eye-head coordination has large variance even within individuals, each subject is asked to conduct several trials at the enrollment phase. The DTW algorithm is also used to select the optimal template from among the template data of N trials for each subject. Specifically, the optimal template data is compared with DTW(T_{ij}, $T_{ij'}$) ($j \neq j'$), where the trial number j ($1 \le j \le N$) of subject i is denoted as $T_{i,j}$, and the one with the best score is selected.

3.5 Tolerance Against a Masquerade Attack

An active attacker would attempt to practice and imitate the eye and head movements of a legitimate user. However, the eye-head coordination is a remarkable involuntary movement for a human; head movements are generally dealt with successfully by counter-rotation of the eyes induced by the combined actions of the vestibulo-ocular reflex (VOR) and the optokinetic reflex [16]. Such complicated coordinated movements are biometric information that is considered to be very difficult for people

themselves to control, and it makes a clear distinction between that and the behavioral biometrics which simply use unconscious response. Thus, even if an impostor matches his/her eye and head position to "the position of the eye and head of the legitimate user when a legitimate user gazes at a stimulus," it is considered impossible for the attacker to imitate the feature at all times when gazing at a moving stimulus.

4 Basic Experiment

4.1 Experiment Purpose

To explore the feasibility of the proposed biometric authentication system based on human reflexes, we conducted a fundamental experiment on the differences between individual people in eye-head coordination. The test subjects were ten students from a university (subjects *A*–*J*). The experimental equipment has the camera only to the right side; thus, the experiment is conducted using the right-eye video acquired by the corneal imaging camera.

It is necessary to verify that it is impossible to imitate the eye-head coordination movements illicitly, even when the authentication information of a legitimate user is leaked. However, we have not conducted experiments on spoofing because our study is the verification stage that the availability of the proposed authentication system is evaluated by eye-head coordination.

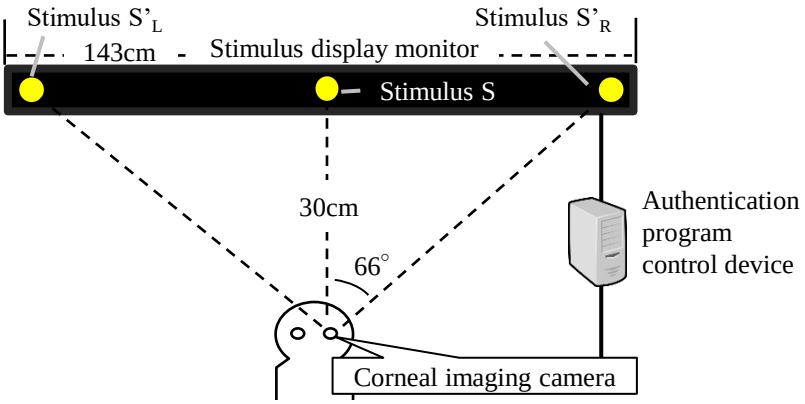

Fig. 3. System arrangement.

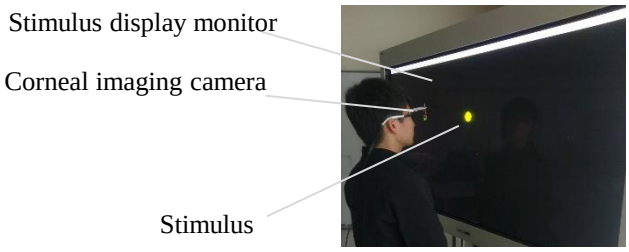

Fig. 4. Experiment environment.

4.2 Experiment Overview

Overviews of the system arrangement and the experiment environment are shown in Fig. 3 and Fig. 4. The details about each device are given below:

- Stimulus display monitor: This device displays the visual target. Using a 65 inch display TY-TP65P10S (developed by Panasonic), which has a resolution of 32,000 × 18,000 pixels.
- Corneal imaging camera: This device takes the video of the right eye of the subject. The subjects wear the corneal imaging camera (Fig. 1) during the experiment. We use NCM13-J (developed by Nippon Chemicon, 640 × 480 pixels, 15 fps).
- Authentication program control device: This device measures the corneal position. It contains a program written in JAVA, to implement each of the enrollment and authentication phases. This program runs on a PC (CPU: Intel Core i5 2.6 GHz, Memory: 8 GB, OS: Windows 10 HOME).

4.3 Experiment Method

This experiment was conducted with L = 30 and $\theta = 66$ using the method described in section 3.3, as shown in Fig. 3.

As described in section 3.4, the eye-head coordination has a large variance even within individuals. To mitigate it, each subject was asked to undergo a total of five trials in the enrollment phase. The optimal template from among the five trials for each subject was selected and used in the authentication phase. This means we obtained 100 sample videos (ten subjects × five trials × two directions (S'_R or S'_L)) in the enrollment phase.

The experiment was conducted for two months. For the authentication phase, five trials were conducted at about two minutes after the enrollment phase, one month after, and two months after. Thus, we obtained 300 sample videos (ten subjects × five trials × two directions (S'_R or S'_L) × three days) in the authentication phase.

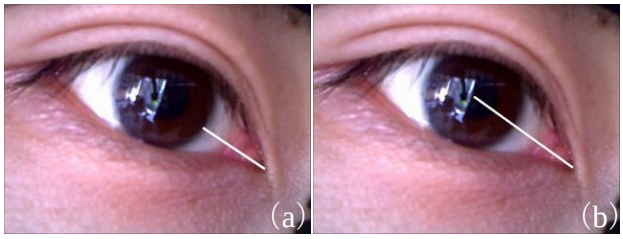

Fig. 5. Features. (a) Distance between the inner canthus and cornea boundary, and (b) Distance between the inner canthus and center of the cornea.

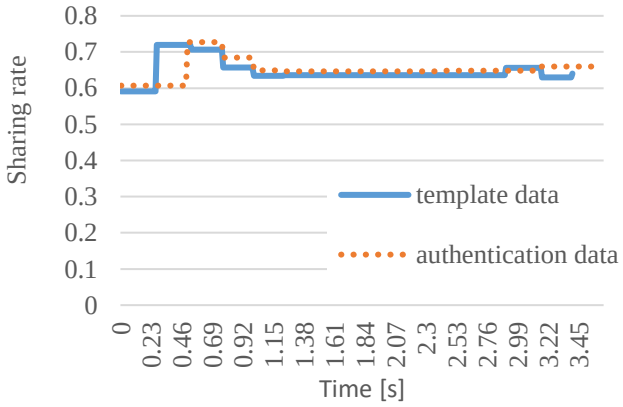

Fig. 6. Change of sharing rate for subject.

As described in section 3.2, in this study, we calculate the position of the eyeball as a feature corresponding to "the sharing rate." Specifically, we measured the distance between the inner canthus and the cornea boundary (Fig. 5 (a)) in each frame of the video and considered it as "the sharing rate." To eliminate the difference in eye size between subjects, this distance is normalized by dividing it by the distance between the inner canthus and center of the cornea (Fig. 5 (b)).

As a specific example of data, the first template data of the subject *A*'s enrollment phase and the first authentication data of the authentication phase are shown in Fig. 6. The time count starts when the stimulus is onset on the display.

4.4 Evaluation

Using the optimum template data, the DTW score described in section 3.4 along with each authentication data are calculated. Considering the experiment setting that the corneal imaging camera is put on the right eye of the subject, the authentication data of each subject is classified into two types: the right side stimulus S'_R and the left side stimulus S'_L. The authentication accuracy was evaluated in the following four

decision modes, where the authentication data when the stimulus is presented on the right side of the subject is denoted as "authentication data S'_R" and the data when the stimulus is presented on the left side as "authentication data S'_L":

1. Use only authentication data S'_R
2. Use only authentication data S'_L
3. Use authentication data S'_R or S'_L
4. Use authentication data S'_R and S'_L

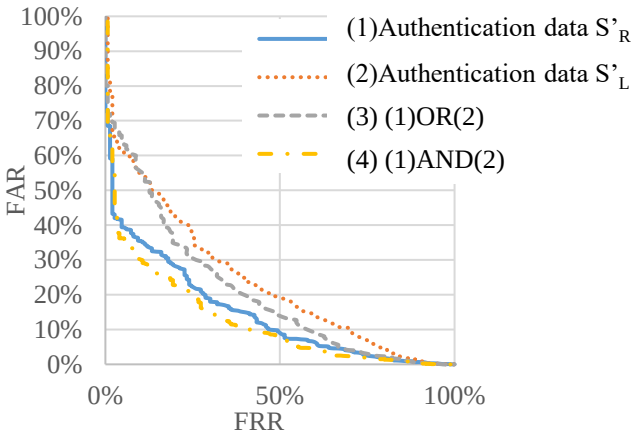

Fig. 7. FAR and FRR.

Table 1. Authentication accuracy.

	(1)	(2)	(3)	(4)
FRR	24.0%	30.0%	28.7%	22.0%
FAR	23.9%	30.7%	28.3%	19.4%
EER	23.9%	30.3%	28.5%	21.8%

Authentication accuracy. To investigate the authentication accuracy of the proposed method, the matching scores between each subject's template data and all subjects' authentication data acquired within two months (at two minutes after the template enrollment, and one month after, and two months after) were calculated. We evaluate our attempt in terms of false acceptance rate (FAR), false rejection rate (FRR), and equal error rate (ERR). The results with the four decision modes are illustrated in Table 1. Fig. 7 shows the trade-off between FRR and FAR at different cut-off points; this is called as receiver operating characteristics (ROC).

The result is not a sufficient level of accuracy, however, Fig. 7 and Table 1 show the possibility that the eye-head coordination can be a feature for user identification. Moreover, Table 1 shows that the authentication data S'_R when presenting the stimulus on the right side has a higher authentication accuracy than that of the authentication data S'_L when presenting the stimulus on the left side. Generally, humans use the right eye for the stimulus on the right side and left eye for the stimulus on the left side. It

seems probably because the corneal imaging camera is placed on the right-eye side, it can be confirmed that the identity is significantly found in the movement of the right eye that is mainly associated with gazing at a right side stimulus S'_R. Furthermore, it can be confirmed from Table 1 that the authentication accuracy is improved by using both the authentication data S'_R and S'_L. From the fact, it is considered that the identity is also included to a certain degree in the movement of the right eye when gazing at a left side stimulus S'_L.

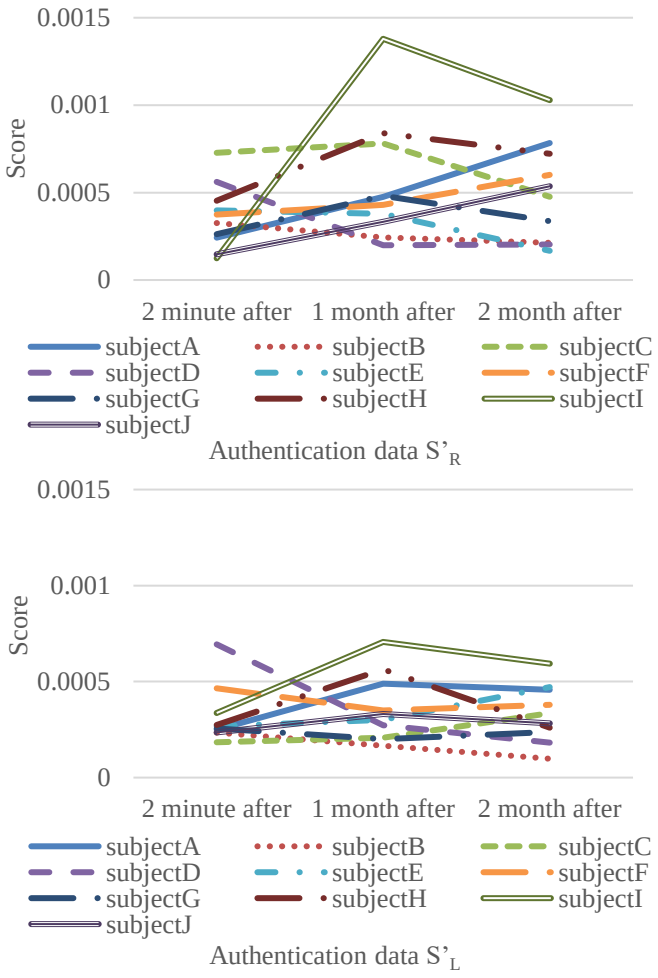

Fig. 8. Long term experimental result

Influence Against Change Over Time. To investigate the long-term stability of the proposed method, the matching scores between each subject's template data and all subjects' authentication data acquired at two minutes/one month/two months after the

template enrollment were respectively calculated. The change of the average score over time for each subject is depicted in Fig. 8.

Thus, it was confirmed that the matching score of some subjects change over time. Therefore, at the current stage of this research, the application of our method will be limited to short-term authentication systems. A possible strategy is to update the template each time authentication is performed. It is necessary to continue discussing methods that can be verified over a longer term.

5 Conclusion

In this study, we calculate the feature from the dynamic behavior of a user, when he/she follows a moving stimulus with his eyes, for the biometric reflex authentication using the individual difference of the eye-head coordination and verified it through a basic experiment. As described in section 3.5, the eye-head coordination is a remarkable and complicated involuntary movement for humans; thus, it is considered impossible for an attacker to imitate it consistently with the movements of a legitimate user at all times while following a moving stimulus with his/her eyes.

It can be confirmed that the combining feature is improved by using an AND of authentication data when presenting the stimulus on the right side and that on the right side, and the best authentication accuracy is obtained with an EER of 21.8.

However, the difference between the user and others did not appear greatly in the biometric information of reflex; thus, further improvement is necessary. In addition, it is confirmed that the matching score of a user changes with time. These issues need to be addressed in future work.

Acknowledgments. We, the authors of this study, are grateful to Prof. Atsushi Nakazawa of Kyoto University, Japan for his advice on the authentication method and provision of experimental equipment in progressing this research.

References

1. T. Matsumoto, “Gummy and conductive silicone rubber fingers importance of vulnerability analysis”, Advances in Cryptology–ASIACRYPT 2002, pp.574–575, Springer, 2002.
2. J. Daugman, “Recognizing Persons by their Iris Patterns”, Advances in Biometric Person Authentication, Springer, pp.5-25, 2004.
3. M. Watanabe, “Palm vein authentication”, Advances in biometrics, pp.75-88, 2008.
4. D. Mulyono and H. S. Jinn, “A study of finger vein biometric for personal identification”, Proc. Int. Symp. Biometrics & Security Technologies, ISBAST 2008. IsIamabad, pp.1-8, 2008.
5. A.K. Jain *et al.*, “On-line signature verification”, Pattern Recognition, Vol.35, No.12, pp.2963-2972, 2002.
6. M. Nishigaki and D. Arai, “A user authentication based on human reflexes using blind spot and saccade response”, International Journal of Biometrics, Vol.1, No.2, pp.173-190, 2008.

7. M. Nishigaki and Y. Ozawa, “A user authentication using blind and papillary light reflex”, Information Processing Society of Japan Journal, Vol.48, No.9, pp.3039-3050, 2007. (in Japanese)
8. A.A Leis *et al.*, “Behavior of the H-reflex in humans following mechanical perturbation or injury to rostral spinal cord”, Journal of Muscle & Nerve, Vol. 19, No.11, pp.1373–1382, 1996.
9. R.S. Young *et al.*, “Transient and sustained components of the pupillary responses evoked by luminance and color”, Journal of Vision Research, pp.437–446, 1993.
10. J. Daugman, “Demodulation by complex-valued wavelets for stochastic pattern recognition”, Int’l J. Wavelets, Multiresolution and Information Processing, Vol. 1, No. 1, pp.1–17, 2003.
11. B. Sabarigiri and D. Suganyadevi, "Counter Measures Against Iris Direct Attacks Using Fake Images and Liveness Detection Based on Electroencephalogram (EEG)", World Applied Sciences Journal, Vol.29, pp.93-98, 2014.
12. J. Jomier *et al.*, “Automatic quantification of pupil dilation under stress”, in Proceedings of the IEEE International Symposium on Biomedical Imaging:Macro to Nano, pp.249–252, 2004.
13. M.A. Gresty, “Coordination of head and eye movements to fixate continuous and intermittent targets”, Vision Research, Vol.14, No.6, pp.395-403, 1974.
14. G.R. Barns, “Vestibulo-ocular function during co-ordinated head and eye movements to acquire visual targets”, The Journal of Physiology, Vol.287, No.1, pp.121-147, 1979.
15. H. Sakoe and S. Chiba, “Dynamic programing algorithm optimization for spoken word recognition”, IEEE Transactions on acoustics, Speech and Signal Processing, Vol.26, No.1, pp.43-49, 1978.
16. J.L. Vercher *et al.*, “Eye-head movement coordination: vestibulo-ocular reflex suppression with head-fixed target fixation”, Journal of Vestibular Research, Vol.1, pp.161-170, 1991.