Brain Password: A Secure and Truly Cancelable Brain Biometrics for Smart Headwear

Feng Lin[1,2], Kun Woo Cho[1], Chen Song[1], Wenyao Xu[1], Zhanpeng Jin[1]
[1]University at Buffalo, the State University of New York, Buffalo, New York, USA
[2]University of Colorado Denver, Denver, Colorado, USA
{flin28,kunwooch,csong5,wenyaoxu,zjin}@buffalo.edu

ABSTRACT

In recent years, biometric techniques (e.g., fingerprint or iris) are increasingly integrated into mobile devices to offer security advantages over traditional practices (e.g., passwords and PINs) due to their ease of use in user authentication. However, existing biometric systems are with controversy: once divulged, they are compromised forever - no one can grow a new fingerprint or iris. This work explores a truly cancelable brain-based biometric system for mobile platforms (e.g., smart headwear). Specifically, we present a new psychophysiological protocol via non-volitional brain response for trustworthy mobile authentication, with an application example of smart headwear. Particularly, we address the following research challenges in mobile biometrics with a theoretical and empirical combined manner: (1) how to generate reliable brain responses with sophisticated visual stimuli; (2) how to acquire the distinct brain response and analyze unique features in the mobile platform; (3) how to reset and change brain biometrics when the current biometric credential is divulged. To evaluate the proposed solution, we conducted a pilot study and achieved an f-score accuracy of 95.46% and equal error rate (EER) of 2.503%, thereby demonstrating the potential feasibility of neurofeedback based biometrics for smart headwear. Furthermore, we perform the cancelability study and the longitudinal study, respectively, to show the effectiveness and usability of our new proposed mobile biometric system. To the best of our knowledge, it is the first in-depth research study on truly cancelable brain biometrics for secure mobile authentication.

CCS CONCEPTS

• Security and privacy → Authentication; **Biometrics**; **• Human-centered computing → Ubiquitous and mobile computing**;

KEYWORDS

Wearable computing, mobile authentication, cancelable biometrics, event-related potential.

ACM Reference Format:
Feng Lin[1,2], Kun Woo Cho[1], Chen Song[1], Wenyao Xu[1], Zhanpeng Jin[1]. 2018. Brain Password: A Secure and Truly Cancelable Brain Biometrics for Smart Headwear. In *MobiSys '18: The 16th Annual International Conference on Mobile Systems, Applications, and Services, June 10–15, 2018, Munich, Germany*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3210240.3210344

1 INTRODUCTION

In recent years, biometric authentication is taking over traditional passwords or PIN based authentication in mobile and wearable applications because of identification accuracy, convenience and seamless integration with personal devices. However, existing biometrics, such as fingerprint and face, are prone to prone to being hacked in everyday life or social media. For example, the Chaos Computer Club announced that one of its members had been able to replicate the fingerprint of German Minister of Defense Ursula von der Leyen, using only photographs taken of her finger [14]. Biometrics are unique to individual. Different from traditional passwords, once such biometric credentials are damaged or counterfeited, the user cannot cancel the pre-stored credentials or reset them with a different biometric input.

How to design a *truly cancelable* biometric system is an unsolved historical topic in the biometric research community. Cancelable biometrics are challenging because stability and cancelability in biometrics are at odds with each other. Stability requires that biometric traits are immutable and hard to change; cancelability requires that biometric traits are erasable and easy to change. According to our literature review, existing works on cancelable biometrics mainly focus on "soft cancellation", which means the biometric system only uses and saves transformed biometric credentials, such as images with random projection, in the database. Rather than generating a new biometric credential, once biometric credential in the database is divulged, soft cancalable biometric system will have users generate new biometric credentials with a different transformation formula. For example, Paul *et al*. [49] introduced a cancelable template generation algorithm, when previously transformed template is stolen, that produces a new transformed biometric template. The proposed algorithm can generate new templates unlinkable to the previous compromised template. Nevertheless, this soft-cancellation method is privacy-preserving in the biometric database and only works in case of database breaches. Once original biometric credentials are disclosed in either daily life or social media (e.g., stealing raw fingerprint patterns from a photograph), it still results in permanent biometric compromise in biometric systems. Therefore, to address this fundamental limitation of biometrics, we need to seek a new angle on cancelability study.

In recent years, physiological activities from human organs (e.g., brains [47]) receive increasing attention in biometric communities. The advantage of brain electric activity based biometrics is that they are biologically unique and less prone to forgery because of the

MobiSys '18, June 10–15, 2018, Munich, Germany

ACM ISBN 978-1-4503-5720-3/18/06.
https://doi.org/10.1145/3210240.3210344

dynamics of brain responses. For example, event-related potential (ERP) brainwave is one type of brain electrical signals and can be changed once different visual stimuli are presented [61]. This special feature of brain response offers the potential to design a truly cancelable biometrics, referring to “hard-cancellation". For example, if an ERP brainwave is produced in response to a series of images, that ERP brainwave can be canceled, and a new ERP brainwave can be generated in response to another series of image stimuli.

Here, we argue that the most secure cryptographic credential can be obtained by ERP brainwave signals. By definition, ERP is one of the brain biometric measures that is related to individual-specific characteristics. Besides its unique property of hard-cancellation, ERP also possesses another superior attribute compared with traditional biometrics. While conventional anatomical and behavioral biometrics, such as a fingerprint, voice, stroke, and gait, are not confidential to an individual or can easily be altered for imitation [13, 39], ERP biometrics are highly secure; one cannot reproduce or copy other person’s mental pass-phrase. Moreover, it is non-revealable and naturally less prone to spoofing and counterfeiting [52]. In summary, the ERP-based biometrics stand out with the following advantages:

- **Secure**: Traditional brainwaves biometrics require users to create thought patterns to generate the corresponding brainwave credentials [64]. In this case, brainwave credentials are consciously controlled by users, which can be revealed either purposely or unintentionally [42]. Instead, ERP is a non-volitional and involuntary brain response. This mechanism conceals conspicuous interactions and provides better security, i.e., even a user has no control of ERP generations.
- **Cancelable**: Part of what makes each brain unique is their knowledge and memories. The brain network that manages forming and accessing memories is large, and spans across many anatomical areas [5]. This provides a potential large capacity of various brain ERP responses. Therefore, if the ERP template database is breached, new user’s ERP credentials are possible to be generated by different stimuli sets. Note that ERP biometrics also require no memorization burden on users as other passwords (e.g., PIN, graphical pattern).

Based on the above arguments, we study a new psychophysiological approach for secure and trustworthy user authentication in a head-mounted display (HMD). An HMD is a computerized, information viewing device that is worn on the head. It consists of a small display optic in front of eyes, which covers the entire field of vision of the user and produces an imaginary screen that appears to be positioned away from eyes. Since both ERP acquisition sensor and HMD are mounted on the head (see Section 3 later), it is natural to employ ERP biometrics for the authentication of smart headwear.

In this work, we study ERP, a non-volitional and involuntary brainwave response to a specific sensory, cognitive, or visual stimulation, for HMD authentication. To generate distinct ERP patterns for biometric applications, we utilize a visual stimuli design consisting of the imagery patterns of animal, human face, and text as examples. Specially, a lightweight wearable brain-computer interface with three channels (i.e., P1, Pz, and P4) is developed for the brain activity data acquisition. Our main challenge is to figure out what is the effective strategy to reset and generate new and secure ERPs when the ERP credential is divulged. In this study, we present a novel stimuli update strategy that updates the in-use stimuli to evoke new stable ERPs. As an analogy to the case where the user is not allowed to use a password that is too close to a previous selection, we characterize the sequence of visual stimuli in a joint spatio-temporal domain and choose the ERP with the maximum proposed spatio-temporal warping distance as the new credential. As a result, the original and newly generated “brain passwords” are disparate enough that the original ERP cannot be cross-matched to access the system configured with new ERP credentials. Also, the system maintains stability as the new ERP retains immutability until it is divulged again. To validate the proposed approach, we further conduct a pilot study to evaluate the system security via f-score accuracy (f-1), receiver operating characteristic curve (ROC), equal error rate (EER), and time efficiency. With 179 adult participants, our system achieves a f-score accuracy of 95.46%, and EER of 2.503%. The cancelability evaluation proves that our stimuli update strategy is effective in revoking old ERP and reissuing new ERP derived from the same physical traits without degrading the authentication performance. Also, the unlinkability between old and new ERPs is discussed in this study.

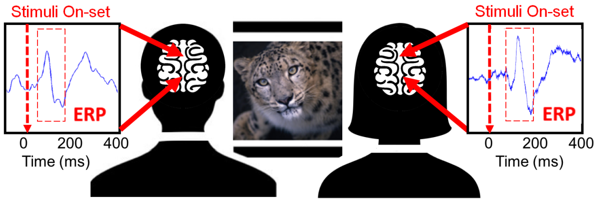


Figure 1: A single ERP signal is elicited by a specific sensory and cognitive event. ERP is unique for individuals that different people will have distinct response with the same stimulus.

To the best of our knowledge, this is the first in-depth study to explore secure and *truly cancelable* biometrics for mobile authentication. Our contribution is three-fold as follows:
1) We develop an end-to-end brain biometric system integrated with a head-worn device. We propose a secure and truly cancelable ERP-based authentication protocol with its application for smart headwears and study a sophisticated brain response model.
2) We study a joint spatio-temporal domain analysis-based stimuli update strategy to achieve the cancelability of our proposed biometric protocol. We empirically investigate the biometric capacity of brain response.
3) We validate the feasibility and effectiveness of our proposed system with multi-session pilot studies, including the performance study, cancelability study and longitudinal study in different user scenarios.

2 BACKGROUND AND RATIONALE

2.1 HMD Authentication

Significance: In recent years, HMDs have been widely developed and improved for a variety of purposes. Main applications include

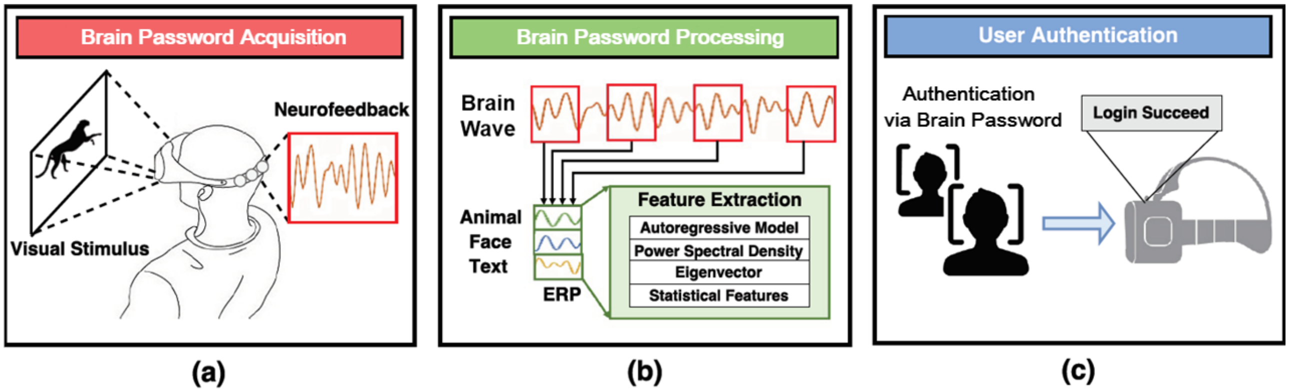


Figure 2: A Brain Password-based HMD authentication system framework is illustrated. (a) When a user attempts to access a head-mounted device (such as a VR headset or Google Glass), a set of visual stimuli will be displayed on the display optic, and the dry sensors implemented in the device will measure brain-responses. (b) Obtained brain signals will be processed and analyzed. (c) The ownership will be identified by comparing with pre-stored templates of the device owner.

virtual reality (VR) for simulation of user's presence in artificial environments (Samsung VR [57]) and realistic experience of 3D games (PlayStation VR [50]). Also, some HMDs provide an augmented reality (AR) to integrate digital information with user's real-world environment (Google Glass [27]), medical visualization for surgeon's natural view of the operation [37], and military simulation and training for either dangerous or costly situation [63]. According to the analysts [41], the HMD market is expected to reach up to 15.25 billion dollars by 2020.

Challenges: To date, existing authentication approached for HMD are limited in multiple aspects. Since HMDs are lacking in either physical keyboards or touchscreen, current authentication systems in HDM often rely on additional mobile devices, which must be carried along, registered, and paired via a wireless connection (e.g. Bluetooth). For hands-off devices, this authentication mechanism is not only inconvenient but also vulnerable to hacking if the paired device is lost or stolen. In fact, technological advancements provide better security mechanisms using biometrics, such as eye blinking [55], head movement [36], and hand gesture [16], for authentication in HMD. Yet, addressed methods are not perfectly trustworthy because a majority of biometrics can be surreptitiously duplicated or revealed by attackers [8].

2.2 Brain Response to Visual Stimuli

2.2.1 ERP Rationale. ERP is a stereotyped brainwave response to a specific sensory, cognitive, or motor event. Part of what makes each human unique is their memory. No two people have had exactly the same experiences. Importantly, no two people interpret similar events exactly the same way. Each person's interpretation of an event is based on their semantic memory, a part of memory that includes a person's knowledge about what images depict and how they relate to their own experiences [48]. Thus, semantic memory is individually unique in this way, and the activity of semantic memory is visible in the scalp-recorded ERP, as shown in Fig. 1.

2.2.2 Characteristics of ERP. In this part, we will discuss the key properties of ERP in biometric applications, including three aspects:

Cancelable: In traditional authentication systems, users can easily replace the password when their credential is divulged. As an analogy to this, we argue that hard-cancellation can be achieved with ERP biometrics by changing visual stimuli. No person has exactly the same experience and memory on different events. Since the ERP is a stereotyped response to a particular event, we claim that changing the event can alter the characteristics (e.g., shape, occurrence duration) of individual's ERP signal and provide new ERP signatures for the password reset.

Stable: Electroencephalogram (EEG) is a type of brainwaves that is often collected without stimulation. Therefore, the performance of EEG biometrics can be highly unstable as it depends on individual's emotional and physical states at the moment of authentication. Moreover, Ruiz-Blondet *et al*. [56] demonstrated that typical EEG signals cannot reflect narrow, specific and cognitive processes as they are not captured time-locked to any stimulus. In our study, we present a much more stable authentication method by utilizing the ERP signal, a stimulus-averaged signal that is time-locked to a specific event.

Non-volitional: In the absence of stimulation, EEG can be volitionally modulated. For instance, a volitional control of neural activities can be achieved by real and imagined movements and cognitive imagery [23]. Thus, without stimulation, EEG can be controlled by conscious thinking of the user, which denotes that EEG is less secure to be used for authentication in case that users intentionally disclose their EEG credentials. In contrast, ERP biometrics are evoked by the stimulus, and therefore it is not under control of the user. This characteristic prevents the user from manipulating the brainwave contents purposely [56].

3 ERP AUTHENTICATION FRAMEWORK

3.1 Framework Overview

Our proposed system comprises of three modules: visual stimuli selection, ERP signal acquisition, and signal pattern analysis. Primarily, a series of stimuli are selected according to our designated stimuli selection strategy. Brainwave signals are then acquired and

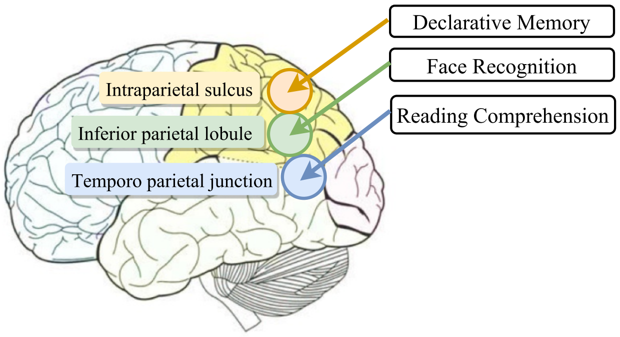


Figure 3: Certain areas of human brain largely influence certain cognitive functions.

averaged into the stimulus-averaged ERP signal. Then, the ERP signals are filtered, and the features are extracted via autoregressive model (AR), power spectral density (PSD), and eigenvector. Lastly, the classification of feature vectors is performed via support vector machine. The illustration of the ERP-based authentication system is shown in Fig. 2.

3.2 Visual Stimuli Design

Design Fundamentals: To generate effective ERP biosignals, we use a distinct stimulation protocol that consists of a large set of various stimuli. As an analogy to a strong personal identification number (PIN) that requires a mix of numbers, letters, and special characters, (e.g., 1E@2R!3P), our brain password design also includes a mixture of various visual stimuli to enhance the "brain password" strength.

The criterion of stimuli selection is that the chosen stimuli must stimulate certain brain areas and reflect certain functional capabilities of the human brain. In this way, our brain password can satisfy the design diversity, thus forming a secure and robust credential. As shown in Fig. 3, three special areas exist at the back of the human brain, including intraparietal sulcus, inferior parietal lobule, and temporo parietal junction, each of which corresponds to a dedicated function of human brain. Specifically, intraparietal sulcus controls the declarative memory [66], inferior parietal lobule processes the face recognition [29], and temporo parietal junction manages the reading comprehension [35]. When a certain function is evoked, a distinct characteristic of the brain waveform is exhibited. In our design, pictures of *animal, celebrity human face, and the segment of texts* are selected as the effective stimuli for aforementioned brain areas to process declarative memory, face recognition, and reading comprehension, respectively. The examples of the three visual stimuli are shown in Fig. 4.

Declarative memory is the memory of facts and events, and refers to those that can be consciously declared [22]. It can be further sub-divided into episodic memory and semantic memory, in which semantic memory is a structured record of knowledge about the external world that we have acquired, including general factual knowledge, shared and independent of personal experience [65]. The rationale for choosing pictures of animal for the declarative memory is that one's semantic memory on the appearance of a certain animal is highly individualized [68]. For example, a person who has suffered a spider bite will react differently to a spider picture than a person who has never been bitten by a spider. Moreover, the brain activation of people with particular emotion to a certain category of animal is different from the brain activation of people who don't possess such emotional state when the visual representation of that category of animal is exposed [68]. As for the human face, neurophysiology studies [33, 70] prove that the unique subject-specific brain signals can be obtained during the human face recognition process. For instance, face stimuli elicit a larger peak of the negative brain potential at 170 ms (N170) compared to the ERP evoked by non-face stimuli [62]. Furthermore, texts are used to elicit semantic memory as it is extremely unlikely for any two people to have same ability to comprehend text. Also, texts are known to elicit a distinctive negative brain potential for each individual [6].

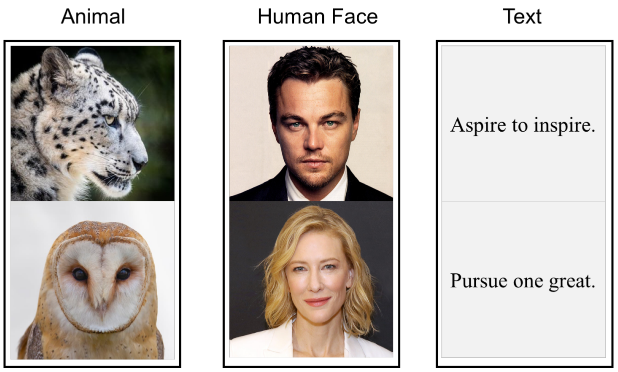


Figure 4: Examples of visual stimuli, including animals, celebrity human faces and texts.

Visual Stimuli Design and Selection: To choose effective images from three stimuli types for each person, we require the ERP signal from each type of stimuli to be distinct from the ones from other types of stimuli, such that each ERP signal can significantly reflect the attributes of their corresponding brain areas. Therefore, we aim at selecting stimuli whose ERP signals can achieve maximization of the dissimilarity among them. Specifically, let p(t) be the continuous-time 2D ERP signal and T_s be the sampling period. The discrete ERP sample for each stimulus can be written as $p_i = p\left(iT_s\right)$. For the jth ERP signal from animals stimuli, it can be written as:

$$\mathbf{a}_j = \left\{p_1^{a,j}, p_2^{a,j}, \cdots, p_{N_s}^{a,j}\right\}^T, j = 1, 2, \cdots, N, \qquad (1)$$

where N_s denotes the number of the sample size in the ERP signal, and N denotes the total number of the ERP for each type in the pool of collected data. The superscript a indicates that the signal belongs to the animals stimuli category. Likewise, ERP signals from texts and human faces can be written as $\mathbf{t}_j$ and $\mathbf{f}_j$.

ERP signals corresponding to the same stimulus can be expressed and mapped as a dot in a high-dimensional space, where each point has the dimensionality of N_s. For ease of representation, we depict the geometric relationship of ERP signals in a 3D space, as shown in Fig. 5. The ERP signals from the same type of stimuli are aggregated as a set, namely, **A** for animals, **T** for texts, and **F** for celebrity faces.

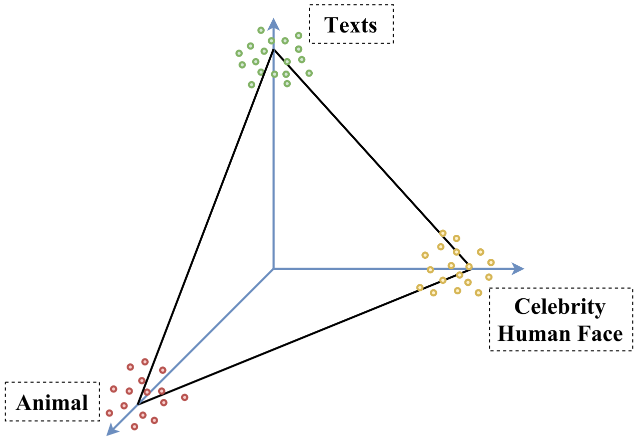


Figure 5: A geometric illustration on visual stimuli selection. Images of animals, celebrity human faces, and texts are distributed in the 3D space as three clusters. We aim to find three dots from clusters with the maximum perimeter.

To maximize the diversity among the ERP signals from different types, we aim to find a triangle, as shown in Fig. 5, which has the largest perimeter. Thus, the visual stimuli selection can be formulated as follows:

$$\underset{i,j,k}{\text{maximize}} \ \left\|\mathbf{a}_i - \mathbf{t}_j\right\|_2 + \left\|\mathbf{a}_i - \mathbf{f}_k\right\|_2 + \left\|\mathbf{t}_j - \mathbf{f}_k\right\|_2, \qquad (2)$$

$$\text{s.t.} \ \ \mathbf{a}_i \in \mathbf{A}, \mathbf{t}_j \in \mathbf{T}, \mathbf{f}_k \in \mathbf{F}, i, j, k = 1, 2, \cdots, N. \qquad (3)$$

By solving the above formulation, we can use the solution set $\{i, j, k\}$ as the ERP stimuli set.

Password Set Expansion: We can define the size of the ERP stimuli set by finding the sub-optimal solution with a certain dimension in Eq. (2) and (3). This is similar to expanding the PIN password length from "1@a" to "1@a2!b". In this study, we define the size of the ERP password set as N_p, where we consider one combination of three stimuli types (one triangle) as one password set ($N_p = 1$). The performance of various N_p values are evaluated and discussed in Section 7.3.4.

3.3 ERP Acquisition Protocol

In our ERP acquisition protocol, three types of images are presented in a certain order. The order of the stimulus presentation is from Animal, human Face to Text (short for A-F-T). When this stimuli sequence with certain images repeats for four times, the acquired EEG signal undergoes the ERP processing method (see Section 5) and produces a single stimulus-averaged ERP, which we simply refer to as an ERP signal, for each stimulus type. During the data collection task, participants were instructed to pay attention to the image. Each image is flashed for only 200 ms to avoid the use of exploratory eye movements, and 200 ms interval is applied in between two images to make each stimulus independent of the previous stimulus (see Fig. 6). In our experimental protocol, the acquisition of ERPs for the animal, human face, and text took approximately 4.8 seconds. The appropriate duration of stimulus presentation is further investigated in Section 7.3.4.

4 SYSTEM IMPLEMENTATION

4.1 System Overview

Fig. 7 shows the flowchart of our proposed system. A set of visual stimuli is selected from the database and displayed to the user through the VR headset. The generated ERP signal is extracted and analyzed for later matching with the owner record. If it matches, the user is considered as the owner. Otherwise, the user is rejected as the intruder.

4.2 ERP Acquisition Device

To capture the ERP data, our team has developed an ERP brain sensor headset, which is equipped with dry electrodes. Such electrodes utilize a set of angled legs and permit the legs to flex outward under pressure which help push aside hair for better contact. The sensors are coated with metallized paint for conductivity, providing low impedance contact (100-500 kΩ) to suppress noise in the ERP acquisition. The headset employs the channel P3, Pz, and P4 (International 10 − 20 System) with two grounds (Fp1 and Fp2) and reference on A1 (See Fig. 8). The brain sensor headset can conveniently collect brainwave signals at the sampling rate of 1000Hz. The collected data can be saved locally or streamed to a computer via Bluetooth.

4.3 Electrode Placement

In standard practice, 32 to 64 electrodes are used for ERP measurement, and the number of electrodes can increase up to 256 to obtain detailed information [59]. However, the implementation of multiple electrodes in the HMDs is problematic due to the heavy weight, low cost-efficiency, and highly complex data acquisition process [8]. Therefore, we customized a sensory headset that is suitable for HMD applications. Our brain sensor device contains three channels (i.e., P3, Pz, and P4) on the parietal lobe.

According to previous studies [19, 30, 40], brain-computer interface (BCI) classification accuracy can be significantly increased by utilizing the parietal electrodes P7, P3, P4, Pz, and P8 because the negative peak of ERPs in the parietal region is unique compared to other regions. Also, since the parietal lobe has an important role in the recollection of episodic memory [10], the parietal electrodes are highly recommended as an alternative to using the complete

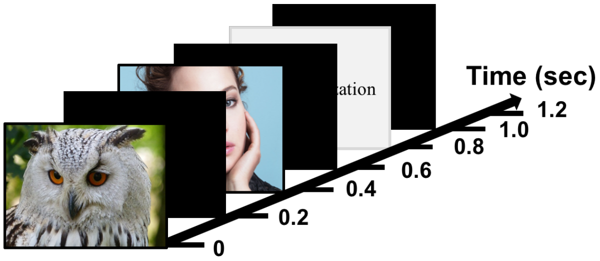


Figure 6: The time interval between images. Each image flashes for 200 milliseconds, and it takes another 200 milliseconds to switch to the next image. This image sequence is shown for four times, and four brain responses from each image are combined into an aggregated ERP response representation.

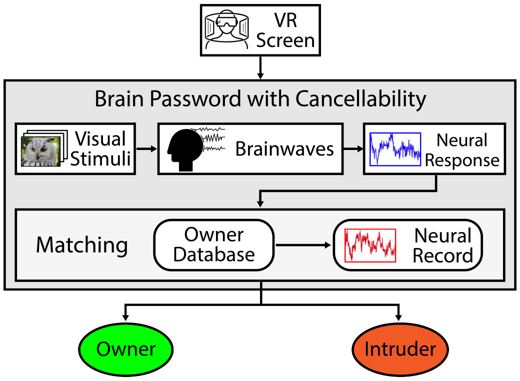


Figure 7: The flowchart of the proposed ERP-based brain password system.

EEG channel set. More importantly, as shown in Fig. 8(a), P3, Pz and P4 are placed on the brain areas addressed in Section 3.2. Also, since the headband of HMD is typically placed on the back, these electrodes can be easily implemented in the headband, providing more convenient and non-invasive data acquisition process.

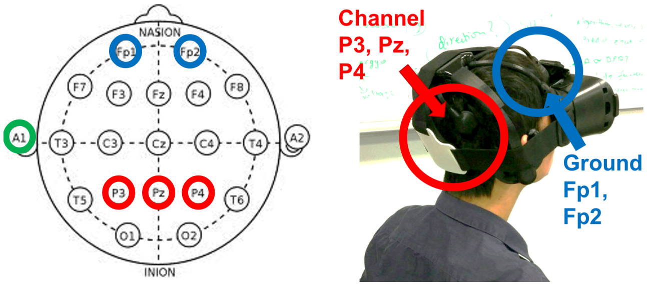
4.4 Motion Artifacts Suppression

Motion artifacts generated by the head movement may compromise the ERP recordings. However, this is inevitable while wearing smart headwears. In our proposed method, the automatic epoch rejection removes the data epoch with extreme artifact noises using visual inspection and measurement statistics including mean, standard deviation, skewness, kurtosis, and median. Then, an infinite impulse response filter reduces high-frequency noises. To further compensate for artifact noises, we applied a channel-based artifact template regression procedure and subsequent spatial filtering approach [28], which removes the ambulation-related movement artifacts. After that, an adaptive independent component analysis (ICA) mixture model parses the EEG signals into maximally independent components (IC), which undergoes the component-based template regression procedure. The feasibility of this approach is proved by the data collected while walking and running on a treadmill.

5 ERP PROCESSING

5.1 Pre-processing

Pre-processing is applied to improve the resolution of brain signals. After obtaining a full EEG waveform, the signal is segmented from the start to the end of the stimulus hit. Thus, each ERP segment has a length of 200 milliseconds. Automatic epoch rejection [21] is applied at the probability threshold of 2.5 to remove the segments with abnormal electrode activity. Then, four ERP segments of the same type are averaged into a single stimulus-averaged ERP signal. These ERP signals of animal, face, and text type are combined into one vector. Therefore, there is 600 milliseconds stimulus-averaged ERP template per channel per subject. Then, an infinite impulse response (IIR) Butterworth filter is employed to produce a zero phase-shift. The diagram that shows the whole ERP processing is illustrated in Fig. 2 (b).



(a) Standard EEG location

(b) Actual positions

Figure 8: The hardware setup in Brain Password. Fig. 8(a) the standard electrode location in International 10-20 System. The electrode in green represents the reference, the electrodes in blue are grounds, and the electrodes in red reflect the channels used. Fig. 8(b) the placement of the Dry EEG Headset.

5.2 Feature Extraction

Each channel has 280 feature elements, and the feature vector of the channel is attached to the feature vector of other channels. Therefore, the final length of one feature vector is 840. The following features are extracted for each feature vector:

Autoregressive Model: We utilize three 6th order autoregressive (AR) models [34] to extract ERP features. AR model is advantageous with short data segments because the frequency resolution of AR spectrum is infinite and does not depend on the length of analyzed data [2]. Since our ERP signals are short data segments, AR model is suitable for our system. By definition, the AR model is a linear difference equation in the time domain:

$$X_t = \sum_{i=1}^{p} a_i x_{t-i} + \varepsilon_t, \tag{4}$$

where X_t is the signal at the sampled point t, p is the order of the model, a_i is the AR coefficient, and ε_t is an independent and identically distributed white noise input [32]. To obtain normalized autoregressive (AR) parameters, we employ the Yule-Walker method [24], which exploits the approximate of the autocorrelation data function. Then, the Burg method [12] is utilized to reduce linear prediction errors. Lastly, the covariance and modified covariance methods are used to minimize the forward and backward prediction errors. Since each model consists of six parameters, 24 AR coefficients are obtained for each channel. With all three channels, there are 72 features attached to the vector.

Power Spectral Density: To accurately detect the spread of power with respect to frequency, the power spectral density (PSD) estimate is obtained by the Welch's overlapped segment averaging estimator [3]. First, ERP signals are divided into frames of 128 to utilize periodogram method for ERP application. Then, the Welch power spectrum estimates the PSD by averaging modified periodograms. We extract 128 features from the estimates for each channel and consequently attach 384 features to the feature vector.

Eigenvector: Since the skin electrode interfaces in dry EEG may induce signal noises, the eigenvector spectral estimation method is used to compensate the effect of the noises. The eigenvector method is known to provide a suitable resolution for artifact corrupted signals by calculating a pseudo-spectrum estimation, which is defined as [2, 58]:

$$P(f) = \frac{1}{\sum_{j=i+1}^{N} \left| V_j^H e(f) \right|^2 / \lambda_j}, \quad (5)$$

where $V_j^H e(f)$ represents a Fourier transform, N is the dimension of the eigenvectors, i indicates the integer value of the dimension of the signal subspace, and λ_j represents the eigenvalue of the matrix. ERP signals are divided into frames of 128, and the pseudo-spectrum is measured by estimates of the eigenvectors. We extract 128 features for each channel, and a total of 384 features are obtained for the feature vector.

5.3 User Authentication

The user authentication process is described as below. Initially, the owner's template is stored in the system. Then, the anonymous user attempts to access the system by wearing the smart headwear device. After detecting the user presence, the system provides a series of stimulus and elicits brain signals of the unknown user. The stimulus-averaged ERP signal from the corresponding user is then verified against the pre-stored templates. During the authentication process, we employ support vector machine (SVM) with a radial basis function (RBF) kernel [31] for the classifier. SVM with RBF kernel enables classification operation in a high-dimensional, implicit feature space without ever computing the coordinates of the data in the input space, where two parameters γ and C dominates the kernel function. γ can be seen as the inverse of the radius of influence of samples selected by the model as support vectors and C trades off misclassification of training examples against simplicity of the decision surface. In our study, γ and C of RBF function are chosen as 0.001 and 10000, respectively. The LIBSVM library for SVM [15] is used for the calculation and decision making. The details on cross validation and evaluation is described in Section 7.2.

6 CANCELABILITY AGAINST ATTACK

Traditionally, once a human biometric, such as iris or fingerprint, is divulged, the authentication system is compromised and no longer safe to use. Comparing with these biometrics, ERP-based brain password is superior because the originally stored credential of brainwave can be canceled if divulged or suffered attack. In other words, our system updates the in-use stimuli to avoid any potential risk. In practice, when a user need to change their password, the system will present a large number of images from the pool to the user and record the brainwave signal, then there is an offline phase where a new password is chosen corresponding to a subset of the images where the selection of that subset follows a stimuli update strategy. In the following, we will deliberate the stimuli update strategy to cancel ERP credentials.

6.1 Stimuli Update Strategy

The candidates for new visual stimuli must satisfy two conditions:

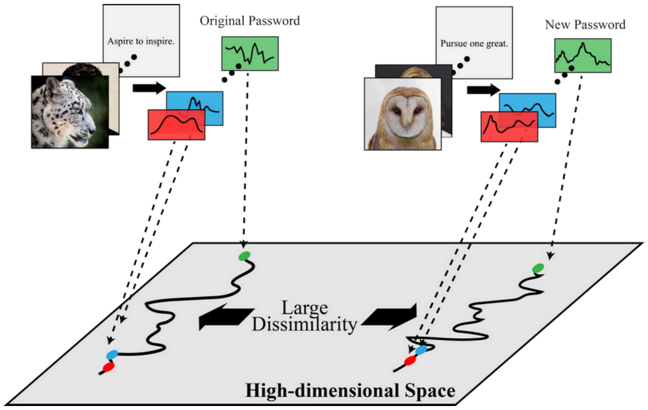


Figure 9: Illustration of stimuli update strategy in Brain Password. The original password design is depicted on the left, and the new one is depicted on the right. The update strategy intends to maximize the difference (i.e., the designated distance) between the original ERP-based biometric credential and the newly generated one.

(1) the new brain password should achieve comparable authentication performance comparing with the original one. Therefore, the new stimuli should also comprise of images from the three diverse categories separately.

(2) the ERP signals evoked by these images should be distinct from the ones evoked by the original images, which is analogical to the case where we are not allowed to use the previously used passwords when resetting passwords. In this way, we guarantee the two passwords are disparate enough that the original brain password is not accessible to the system configured with the new brain password. In other words, we aim to maintain an extremely low false acceptance rate by preventing unauthorized access.

Base on the above discussion, the update strategy is illustrated in Fig. 9, where the original password design is depicted on the left, and the new one is depicted on the right. As previously depicted in Fig. 6, visual stimuli and the corresponding ERP signals can be considered as time series signals. In the meantime, for a specific image, its ERP signal can be expressed as a dot in high-dimensional space (see Fig. 9). Therefore, the time series of ERP signals exhibit spatio-temporal attribute. To quantify the dissimilarity of ERP signals that are generated by the original and new selection of images, we propose a dissimilarity metric, i.e., spatio-temporal warping distance, and compare the two password designs (i.e., ERP stimuli sets) in the joint spatio-temporal domain. Our goal is to find the maximum dissimilarity between two password design in terms of the spatio-temporal warping distance.

6.2 Dissimilarity Metrics

In the following, we will elaborate the design of *spatio-temporal warping distance* as the dissimilarity measurement metric.

Spatial Domain Analysis: Suppose the jth images are considered for both original and new ERP signals, and the ERP signals can be represented in the form of vectors:

$$\gamma_j = \left\{ \mathrm{a}_j^T, \mathrm{t}_j^T, \mathrm{f}_j^T \right\}^T = \left\{ p_1^{\gamma,j}, p_2^{\gamma,j}, \cdots, p_{3N_s}^{\gamma,j} \right\}^T, j = 1, \cdots, N, \quad (6)$$

and likewise $\widetilde{\gamma}_j$, where each element is as defined in Eq. (1). The superscript γ indicates the element belongs to γ. Both γ_j and $\widetilde{\gamma}_j$ have the dimension of $3N_s$.

For the pair of γ_j and $\widetilde{\gamma}_j$, each element in the vector is normalized by dividing the sum of all elements in the vector, written as: $q_i^j = \frac{p_i^{\gamma,j}}{\sum_{i=1}^{3N_s} p_i^{\gamma,j}}$, $\widetilde{q}_k^j = \frac{\widetilde{p}_k^{\gamma,j}}{\sum_{k=1}^{3N_s} \widetilde{p}_k^{\gamma,j}}$. Here, we use q_i^j and $\widetilde{q}_k^j$ to denote the normalized value, and the superscript γ is removed since there is no ambiguity for the symbol q and $\widetilde{q}$. Then we define the cost c_{ik} of transporting between ith data from γ_j, which is q_i^j, and kth data from $\widetilde{\gamma}_j$, which is $\widetilde{q}_k^j$. Specifically, we use the Euclidean norm for the cost definition.

The next task is to find a flow, $\mathbf{F}(i,k) = f_{ik}$, such that the matching work between two datasets γ_j and $\widetilde{\gamma}_j$ will have the least cost:

$$\text{minimize} \sum_{i=1}^{3N_s} \sum_{k=1}^{3N_s} c_{ik} f_{ik}, \tag{7}$$

s.t. $\sum_{i=1}^{3N_s} q_i^j = \sum_{k=1}^{3N_s} \widetilde{q}_k^j, f_{ik} \geq 0, 1 \leq i \leq 3N_s, 1 \leq k \leq 3N_s, \sum_{k=1}^{3N_s} f_{ik} \leq q_i^j, 1 \leq i \leq 3N_s, \sum_{i=1}^{3N_s} f_{ik} \leq \widetilde{q}_k^j, 1 \leq k \leq 3N_s, \sum_{i=1}^{3N_s} \sum_{k=1}^{3N_s} f_{ik} = \min\left(\sum_{i=1}^{3N_s} q_i^j, \sum_{k=1}^{3N_s} \widetilde{q}_k^j\right)$.

By solving the above formulation, we can find the optimal flow $\mathbf{F}$, the spatial matching (SM) metric is found as the matching work normalized by the total flow:

$$\text{SM}\left(\widetilde{\gamma}_j, \gamma_j\right) = \frac{\sum_{i=1}^{3N_s} \sum_{k=1}^{3N_s} c_{ik} f_{ik}}{\sum_{i=1}^{3N_s} \sum_{k=1}^{3N_s} f_{ik}}. \tag{8}$$

Temporal Domain Analysis: Suppose the password set size $N_p > 1$, which means there are more than one image set from three clusters, we can incorporate the temporal domain analysis in addition to the spatial domain analysis. To measure the similarity between these two sequences of images illustrated in Fig. 6, an $N_p \times N_p$ matrix $\mathbf{D}$ is created, called *distance matrix*. The value of the (m^{th}, n^{th}) element in $\mathbf{D}$ represents the distance $d\left(\widetilde{\gamma}_\mathbf{n}, \gamma_\mathbf{m}\right)$ between two sets of ERP signals $\widetilde{\gamma}_\mathbf{n}$ and $\gamma_\mathbf{m}$. Then the SM defined above is adopted as the distance metric, and we can obtain: $\mathbf{D}(n,m) = d\left(\widetilde{\gamma}_n, \gamma_m\right) = \text{SM}\left(\widetilde{\gamma}_n, \gamma_m\right)$. With the guidance of the distance matrix, the shortest warped path through the matrix can be derived [54]:

$$cd(n,m) = d\left(\widetilde{\gamma}_n, \gamma_m\right) + \min \begin{cases} cd(n-1, m-1) \\ cd(n, m-1) \\ cd(n-1, m) \end{cases} \tag{9}$$

and $1 \leq n \leq N_p, 1 \leq m \leq N_p$, where $cd(n,m)$ is the current minimum cumulative distance for $\mathbf{D}(n,m)$, and the initial setting is $cd(0,0) = 0, cd(0,m) = cd(n,0) = \infty$.

After that, the overall minimized cumulative distance $cd\left(N_p, N_p\right)$ can be found. Finally, the spatio-temporal warping distance is calculated as:

$$Dist = cd(N_p, N_p). \tag{10}$$

Overall, our aim is to find a new design that has the maximum *Dist* to the original design.

7 PERFORMANCE EVALUATION

7.1 Participants

In the pilot study, brainwaves are obtained from 179 adult participants with a mean age of 29.85 and standard deviation of 7.72. Among 179 participants, 93 of them are male participants, and 86 of them are female participants. Consent forms for participation in the research study were obtained at the time of the study, and all participants have received a comprehensive description of the experimental procedures. As mentioned above, electroencephalography is a safe monitoring method with no side effects [11]. Moreover, our headset is in dry form that does not require gel or other fluids. To alleviate possible eye irritation that may occur due to the various stimuli used in the procedure, we avoided the use of extremely bright colors and flashing lights.

As described above, the system evaluation relies on a strategically developed experiment that will involve a cohort of participants. We hold an existing active IRB protocol from both the University at Buffalo and University of Colorado Denver, which allows for recording brainwave from human participants for user authentication.

7.2 Description of Experiment

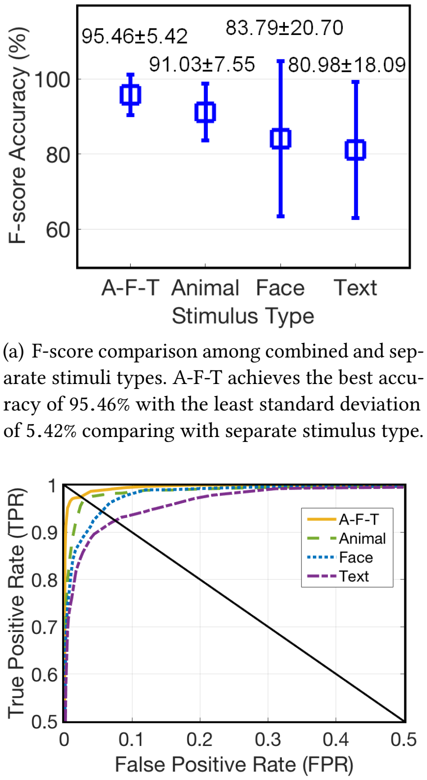
The data are collected in three sessions. The data from the first session is used to evaluate the system performance and cancelability, and the data from the second and third sessions are used for a longitudinal study. Among 179 participants, 80 have participated in the second session (short-term study), and the third session (long-term study). Because some data from 2 participants are damaged, the valid participants for the longitudinal study is 78 with the average age of 27.36.

As there are a total of 179 participants, one of the subjects acts as an owner once while the remaining subjects act as attackers. This process repeats for all subjects. Here, 10-fold cross validation is used to prevent overfitting. The data set is randomly separated into 10 equal-sized subsets. For each trial, one of the 10 subsets is used as a test set, and the remaining subsets are used as a training set. Cross-validation is repeated with each of the subsets.

For each session, the data collection task is organized in a series of 300 images with 100 images for each stimulus type. As mentioned in Section 3.3, a series of same images repeats for four times. Thus, there are 25 different images among 100 images for each type. In other words, the number of stimulus-averaged ERP (N) in the pool of each animal, human face, and text set is 25, which corresponds to the total number of dots in each cluster. For the authentication, one dot for every cluster (one triangle) is used for one-set password (N_p = 1), two dots for every cluster (two triangles) are used for two-set password (N_p = 2), and three dots (three triangles) for every cluster are used for three-set password (N_p = 3). The maximum number of set is N, which is equivalent to 25. We used the one-set password for all evaluations except for Section 7.3.4. To produce multiple ERP templates, we repeat the data collection task 20 times for each participant.

7.3 System Performance

7.3.1 F-score accuracy. The accuracy (ACC) [44] is predominantly used for the statistical classification. However, ACC is an



(a) F-score comparison among combined and separate stimuli types. A-F-T achieves the best accuracy of 95.46% with the least standard deviation of 5.42% comparing with separate stimulus type.

(b) The average ROC curve comparison among combined and separate stimuli types, in which A-F-T appears in the most upper-left corner that indicating our system is robust.

Figure 10: System performance evaluations of F-score and ROC curve.

inappropriate accuracy metric when negative and positive classes are not balanced. Thus, to avoid an unbalanced accuracy measurement, we evaluate our system performance based on f-score accuracy (F_1), which is preferred for the sake of non-sensitivity to class imbalance. Fig. 10(a) depicts the f-score comparison among various stimulus types. As shown, A-F-T indicates the combination of animal, face, and text stimuli that is designed based on our visual stimuli model (see Section 3.2). The stimuli for animal, face, and text types are identical to the pictures used in A-F-T. Among four types, A-F-T achieves the best accuracy of 95.46% with the least standard deviation (STD) of 5.42%. The accuracy of A-F-T is higher than that of animal, face, and text stimuli by 4.43%, 11.67%, and 14.48%, respectively. Moreover, the STD of A-F-T is lower than other three types by 2.13%, 15.28%, and 12.67%, respectively. The results prove that our visual stimuli model improves security and robustness of the brain password by satisfying the design diversity.

7.3.2 Receiver operating characteristic curve. For a comprehensive evaluation of the system performance, a receiver operating characteristic curve (ROC) is investigated. By definition, it visualizes the sensitivity or TPR (true positive rate) against FPR (false positive rate) as the threshold is varied. As the curve follows the top-left portion of the graph, the system has a high sensitivity and specificity and is more accurate. In Fig. 10(b), the average ROC curve of A-F-T, animal, face, and text stimulus type are plotted. Among four curves, A-F-T follows the most upper-left portion of the graph, indicating that our system is robust and feasible.

7.3.3 Equal error rate. The equal error rate (EER), a rate that corresponds to an equal probability of an acceptance error and rejection error, can be derived from the average ROC curve. Specifically, the x-axis value of intersection point between the curve and the diagonal of the unit square is known as EER. More specifically, the EER value of A-F-T is 2.503% and the EER of animal, face, and text are 3.114%, 5.559%, and 7.517%, respectively (derived from Fig. 10(b)). Again, A-F-T achieves lowest EER, which indicates that our visual stimuli model increases the system performance.

7.3.4 Optimization of authentication time efficiency. Since our authentication system targets smart headwear application, the optimization of the authentication time is essential. Thus, we examine several methods to optimize the authentication time efficiency.

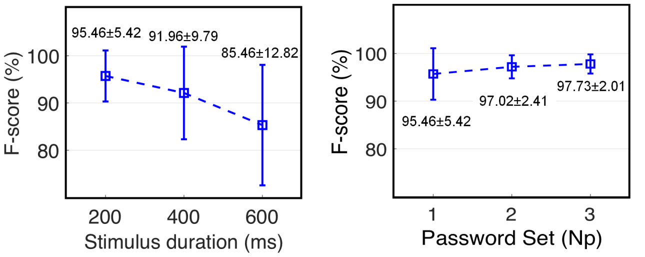
Stimulus Duration: In the experiment, each stimulus is presented for 200 ms, and the black screen is displayed for 200 ms to separate each stimulus. By discovering the optimal stimulus duration, the authentication time can be reduced. As shown in Fig. 11(a), the accuracy declines by 3.5% and the STD increases by 4.37% as the stimulus duration increases from 200 ms to 400 ms. Similarly, when the duration of stimulus exceeds 600 ms, the accuracy reaches 85.46%, which is 10% lower than the accuracy of 200 ms. Also, the STD increases by 7.4% at 600 ms. The reason for this phenomenon is because stimulus presented for more than 200 ms will induce exploratory eye movements, which in some extent will compromise the collected EEG signal dedicated as a response to the visual stimuli. Although the accuracy increases with decreasing duration, the stimulus duration less than 200 ms is too instantaneous for the average human reaction time to the onset of a visual stimulus [4]. Thus, the optimal stimulus duration is 200 ms.

Password Set: As described in Section 3.2, we can optimize the authentication time efficiency by adjusting the size of password set (see Fig. 11(b)). For one-set password ($N_p = 1$), the system accuracy reaches 95.46%. When two-set password ($N_p = 2$) is used, the accuracy increases by 1.56% and the STD decreases by 3.01%. When three-set password ($N_p = 3$) is employed, the accuracy is increased by 0.71% and the STD is reduced by 0.4%. This result indicates that the accuracy and stability of the system increase as the size of password increases.

Time Efficiency: As the stimulus duration and size of password set increase, the authentication time increases as well. In brief, the optimal time can be calculated as:

$$Time\ (s) = N_p \cdot N_{avg} \cdot 3\ (stimu\_duration + 0.20), \tag{11}$$

where N_p indicates the size of password set, and N_{avg} represents the number of the segments that are averaged into a stimulus-averaged ERP, which is 4. In the formula, the interval duration (0.20 second) and the number of stimulus type (3 for animal, face, and text) are included. With the optimal stimulus duration (200 ms),



(a) Stimulus duration impact. The 200 ms duration setting achieves the best F-score of 95.46% with standard deviation of 5.42. Larger duration will have lower F-score and larger standard deviation.

(b) Password length impact. Longer password set will have higher F-score. With the simplest one-set of password, we achieves 95.46% F-score.

Figure 11: The authentication time optimization via stimulus duration and password length adjustment.

one-set password takes approximately 4.80 seconds, two-set password takes 9.60 seconds, and three-set password takes 14.4 seconds. Also, more computation is necessary for higher N_p value. Since the authentication for smart headwear devices must be reasonably fast, we select the one-set password ($N_p = 1$) and the optimized time is 4.80 seconds.

8 CANCELABILITY ANALYSIS

To properly revoke and reissue the credential, the cancelability must satisfy two properties: *revocability* and *unlinkability* [45].

8.1 Revocability

Objectives: In this section, we verify the revocability of ERP in two ways. First, we prove that new ERP is distinguished from the original ERP, thereby corroborating its robustness against the attack using the original password. Second, we demonstrate that new ERP generated according to our stimuli update strategy has a high accuracy to serve as a new brain password.

Description of Experiment: The updated stimuli set is given to the participants, and 20 new ERP templates are obtained per subject. Again, each subject acts as an owner and the rest act as an attacker. The SVM classifiers are used with 10 fold cross-validation. For the second objective, we assume the user generates new ERPs according to the stimuli update strategy and updates the user profile. The attacker uses the replication of user's original ERP to access the system configured with the new ERP. For evaluation, we randomly select a portion of new ERPs to create the updated profile and test the performance by authenticating with the remaining new ERP templates and original ERP templates from Section 7.3. We employ SVM with a 10 fold cross-validation. This repeats for each subject and the FRR and FAR is averaged of all subjects.

Results and Discussion: The evaluation results are shown in Table 1, where it reveals that the original visual stimuli will result in true negatives when adopting them to a system configured with new stimuli. The new ERP credential provides the recall, precision, and f-score of 94.64%, 95.62% and 94.87%, correspondingly. The STD are 6.03%, 5.11%, and 3.69%. Although the recall, precision, and f-score of the original ERPs are slightly higher by 1.04%, 0.29%, and 0.59%, these discrepancies are not significant. Therefore, the updated strategy does not degrade our system performance. As shown in Table 2, our second revocability task achieves a high recall and precision value of 99.20% and 99.05% with low FRR and FAR of 0.775% and 0.789%. These results indicate that our updated ERPs are highly distinct from the originals such that the replicated original credential is unlikely to be used to access the system configured with new credential. This result validates our two hypotheses. First, the ERP biometrics are truly cancelable as the change of the visual stimulus alters the characteristics of ERP. As mentioned previously, the reason is that no one has exactly the same memory on different images. For instance, a person's memory of the spider is highly likely to be different from the memory of the dog. Hence, changing the stimulus from the spider picture to the dog image elicits new characteristics in ERP. Second, our stimuli update strategy amplifies such alteration by finding the maximum dissimilarity among ERPs in response to a larger pool of images.

8.2 Unlinkability

Description of Experiment: We employ the original and new ERP data from Section 8.1 and specifically use the Pearson's correlation coefficient, R, which is defined by the following [38]:

$$R_{i,j} = \frac{1}{N-1}\sum_{n=1}^{N}\frac{(a_{i_n}-\mu_{a_i})}{\sigma_{a_i}}\frac{(b_{j_n}-\mu_{b_j})}{\sigma_{b_j}}, \tag{12}$$

where a_{i_n}, b_{j_n} are the feature element of the original and new ERP template. μ_{a_i} and σ_{a_i} represent the mean and STD of all feature elements of the corresponding original ERP template while μ_{b_j} and σ_{b_j} signify the mean and STD of the elements of the new ERP template. Every template is composed of 840 feature elements as mentioned in Section 5.2, and thus N equals to 840. To avoid increasing the correlation coefficient in all ERP templates, we suppress the feature trend by normalizing each template with the mean of all templates as below:

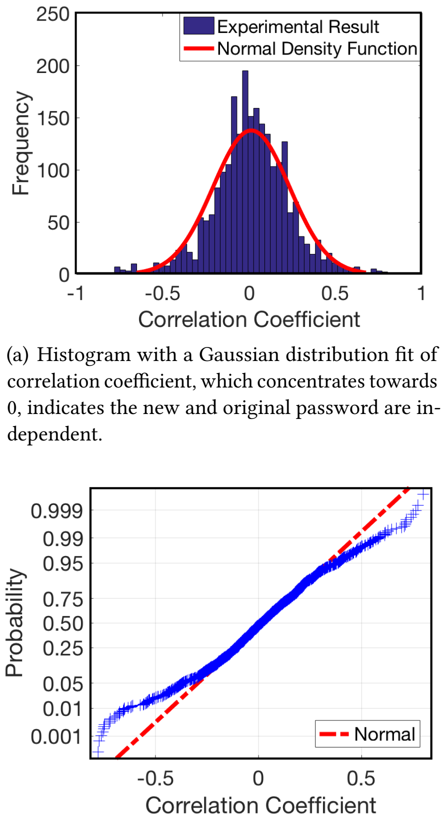
$$Normalized(A_i) = \frac{A_i}{\frac{1}{k}\sum_{p=1}^{k} A_p};\ 1 \le i \le k, \tag{13}$$

$$Normalized(B_j) = \frac{B_j}{\frac{1}{k}\sum_{p=1}^{k} B_p};\ 1 \le j \le k, \tag{14}$$

where A_i and B_j represent the original and new ERP template, respectively. k is the total number of templates for each subject experimented on the same stimuli set, which is equivalent to 20. The correlation coefficient, R, is computed by comparing each normalized template of the old stimuli set with every normalized template of new stimuli set ($k \times k$ comparison).

Table 1: Performance table for each stimuli set.

Trial	Recall (%)	Precision (%)	F-score (%)
Original ERP	95.68±6.89	95.91±4.91	95.46±5.42
New ERP	94.64±6.03	95.62±5.11	94.87±3.69



(a) Histogram with a Gaussian distribution fit of correlation coefficient, which concentrates towards 0, indicates the new and original password are independent.

(b) Normal probability plot. The probability from 0.05 to 0.95 has the correlation ranging from −0.3 to 0.3, which is considered as a weak strength of association.

Figure 12: The correlation test between original and new ERP-based brain password.

Results and Discussion: As shown in Fig. 12(a), the Gaussian curve of the results centers at zero, which indicates that the original ERP and updated ERP are highly independent. At 95% confidence interval ($\alpha = 0.05$), an estimate of the mean is 0.0130 and an estimate of the STD is 0.2212. Moreover, the lower bound of the confidence intervals for the mean is 0.0040, and the upper bound is 0.0219. In addition to the frequency distribution histogram, Fig. 12(b) shows the normal probability plot to identify any substantive departure from normality. The dotted line in red provides the reference for a perfect normality. The upper end of the plot bends below the diagonal line while the lower end bends above that line, forming an S shaped-curve, which indicates a light-tailedness. In other words, our correlation results have less variance than expected. In this graph, we can also observe that approximately 90% of the data has a weak association because the probability from 0.05 to 0.95 has the correlation ranging from −0.3 to 0.3, which is considered as a weak strength of association. Thereby, we prove the independence between two ERPs and ensure that attackers are unlikely to link the old ERP to the new ERP.

Table 2: Authentication of the system configured with the new ERP.

Recall (%)	Precision (%)	FRR (%)	FAR (%)
99.20±1.829	99.05±2.034	0.775±1.805	0.789±1.775

9 LONGITUDINAL STUDY

Description of Experiment: We follow the same experimental settings as Section 7.3. In the enrollment phase, we randomly select a part of owner data and use them to create a profile of the user. Then, we test the performance of the classifier by authenticating the user with the owner templates and all attacker templates. Here, we refer the authentication phase as a pre-trial and re-test for either short-term or long-term study as a post-trial. For the short-term study, the interval between the pre-trial and post-trial is five days. For the long-term study, participants are experimented five months after the pre-trial. The average time interval is 142.8 days. During the short-term study, the participants are familiarized with the stimuli set by observing the set before the experiment. For each subject, the profile of user remains the same and newly collected data are used for login attempts. Each subject acts as the owner once, and the rest acts as the attacker. This test repeats for every subject. Thus, there are total 78 tests for short-term study and 78 tests for long-term study with each test consisting 77 user attempts and 77 attacks from each attacker.

Results and Discussion: The overall performance change is summarized in Table 3. The f-score is increased by only 0.02% during the short-term study. The possible reason is that our stimulus presentation is too fast to properly trigger a short-term memory, and therefore an intrinsic reaction from the semantic memory, a portion of long-term memory, overrides the response from the short-term memory. Conversely, the performance is declined by 1.01% during the long-term study. This change is slightly higher than the change observed in the short-term study. However, it should be noted that this change is still insignificant.

10 DISCUSSION

Liveness Detection: To prevent spoofing attacks, the authentication system must differentiate real biometrics from counterfeits. Most promising way to distinguish them is to detect physiological signs of liveness. Existing methods [9, 25] either request the user to provide signs of liveness or force user to interact with the system continuously, which decrease the user comfort. In contrast, our proposed ERP-based approach is a dynamic and continuous biometric credential, which itself provides the physiological sign of liveness as the active EEG must always come from living individuals. Therefore, the dynamic nature of the brain response [69] provides us a potential method to distinguish the recorded replay attacks injected through the electrodes.

Aging Effect: Most biometrics (e.g., fingerprint, iris, and face) spontaneously morph over the lifetime in a extremely slow pace. Similarly, age-related alterations of brainwave includes the overall EEG power decrease, slower alpha frequency, and slight diminution in P3 amplitude. Yet, they are orthogonal to ERPs obtained from visual stimuli [51]. As shown in the longitudinal study, we do not

Table 3: Overall performance variation (f-score)

Duration	Pre-trial (%)	Post-trial (%)	Change(%)
Short-term	96.43±3.99	96.45±4.32	+0.02
Long-term	96.00±5.81	94.99±6.30	-1.01

observe any significant mutation of brain signal in a long period. Nevertheless, regular ERP profile update can be a potential solution.

Privacy Preservation: In the context of privacy concerns, one natural question is “will this brain biometrics leak privacy information?" The answer is “No". Previous works indicate that brain leakage requires a satisfactory data, such as high-fidelity brainwaves with a professional device (e.g., BCI2000-64 channels [42]) or invasive measures by embedding chips into brain [53]. On the contrary, our system only requires three channels with a small information disclosure. Moreover, our system only collects ERP P/N200 (i.e., within 200ms post-stimulus onset response), while most of the semantic memory attacks require the relatively long-term brainwaves (e.g., non-ERP sections from seconds to minutes [43]).

Further motion artifacts cancellation: While our method described in subsection 4.4 is effective for non-continuous motion artifact noises, it could be vulnerable when extreme physical activities continue throughout the authentication. Thus, further processing to counterbalance artifacts from the continuous gait events are needed. With a three-dimensional accelerometer, the system can detect artifacts induced by head movements and remove the brainwave synchronous with the recorded acceleration above certain threshold [20]. To describe in a more detailed way, head accelerations are measured relative to the initial position, and ICA identifies EEG components that are statistically independent. Then, components that correlate with the recorded acceleration above certain threshold are removed [20].

Future Work: Though we have utilized the Pearson's correlation analysis for the unlinkability property assessment, we plan to provide a more comprehensive evaluation to prove that the reissued brainwave biometrics is indeed unlinkable. Specifically, Spearman's rank order correlation [26], Kendall rank correlation [1], and Hausdorff distance [67] will be employed for the analysis. At the current stage, we validated the feasibility of our brain password with 177 adult participants, a further study with a much larger sets of participants to verify the uniqueness and stability of the brain biometrics is in our plan. Another promising research direction to pursue is to investigate the impact of visual stimuli protocols, such as full color versus black and white, designated visual stimuli under other different categories.

11 RELATED WORK

Headwear User Authentication: In recent year, how to authenticate users in untraditional personal device, such as head-mounted displays, has been increasingly explored in both mobile and security research communities. Chauhan *et al.* [16] developed a touch gesture-based continuous authentication for wearable devices like Google Glass. Similarly, Li *et al.* [36] proposed an authentication system for head-worn devices using user's unique head movement patterns in response to music. Also, Rogers *et al.* [55] presented the method to identify an HMD user based on the user's unconscious blinking and head movement. Other existing techniques, such as eye movement biometrics [60], can be conveniently integrated into HMD devices. However, such physiological and behavioral characteristics are prone to compromise in daily life and thus can be surreptitiously duplicated and counterfeited.

Authentication via Brainwaves: Most brainwave authentications have used EEG as the biometric. Chuang *et al.* [17] presented an subject authentication scheme based on single-channel EEG signals. Similarly, Ashby *et al.* [7] employed EEG signals for person authentication with AR model and power spectral density. However, their results are limited to the controlled condition as the regular EEG signal is sensitive to factors such as human emotion. In contrast, the proposed ERP signal is stimulated based on the inherent human experience. One nascent work [6][56] brings up the concept of ERP-based user authentication, but there is no in-depth exploration regarding the biometric cancelability. While this work focuses on the biometrics cancelability including update strategy design and cancelability analysis.

Cancelable Biometric Systems: Cancelability is one of the most desired features in biometrics. Connie *et al.* [18] proposed a method which uses existing biometric palmprint features with a set of pseudo-random data to generate a unique discretized code for every individual. Similarly, Paul *et al.* [49] developed a cancelable biometric template generation algorithm using random projection and transformation-based feature extraction for multi-modal face and ear biometrics. Further, Ouda *et al.* [46] exploited the feature domain transformation for protecting IrisCode. The feature transformation is accomplished by IrisCode generation, consistent bits extraction, and cancelable BioCode generation. However, these methods are based on a soft-cancellation, which generates a cancelable biometric through the alteration and transformation of existing templates. For the first time, we introduced the notion of hard-cancellation as a generation of totally new bio-features.

12 CONCLUSION

In this paper, we presented the first study to explore secure and usable authentication to headwear devices using cancelable ERP biometrics. The evaluation results show that our approach achieves the f-score accuracy of 95.72%, and equal error rate (EER) of 2.503%. Thus, for the first time, we have validated the feasibility of using unique, non-volitional components of brainwave response for authentication of smart headwear users. Also, we introduced cancelability to the brainwave biometrics through a novel stimuli update strategy. A further cancelability analysis in terms of revocability and unlinkability is conducted to prove the effectiveness of the reissued biometrics credential.

ACKNOWLEDGEMENT

We thank our shepherd, Dr. Landon Cox, and all anonymous reviewers for their insightful comments on this paper. This work was in part supported by the National Science Foundation under grant No. 1266183, 1423061/1422417, 1564104/1564046.

REFERENCES

[1] Hervé Abdi. 2007. The Kendall rank correlation coefficient. *Encyclopedia of Measurement and Statistics. Sage, Thousand Oaks, CA* (2007), 508–510.

[2] Amjed S Al-Fahoum and Ausilah A Al-Fraihat. 2014. Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. *ISRN neuroscience* (2014).

[3] Ahmet Alkan and M Kemal Kiymik. 2006. Comparison of AR and Welch methods in epileptic seizure detection. *Journal of Medical Systems* 30, 6 (2006), 413–419.

[4] Kaoru Amano, Naokazu Goda, Shin'ya Nishida, Yoshimichi Ejima, Tsunehiro Takeda, and Yoshio Ohtani. 2006. Estimation of the timing of human visual perception from magnetoencephalography. *Journal of Neuroscience* 26, 15 (2006), 3981–3991.

[5] David G Amaral. 1987. Memory: Anatomical organization of candidate brain regions. *Comprehensive Physiology* (1987).

[6] Blair C Armstrong, Maria V Ruiz-Blondet, Negin Khalifian, Kenneth J Kurtz, Zhanpeng Jin, and Sarah Laszlo. 2015. Brainprint: Assessing the uniqueness, collectability, and permanence of a novel method for ERP biometrics. *Neurocomputing* 166 (2015), 59–67.

[7] Corey Ashby, Amit Bhatia, Francesco Tenore, and Jacob Vogelstein. 2011. Low-cost electroencephalogram (EEG) based authentication. In *5th IEEE International Coference on Neural Engineering*. IEEE, 442–445.

[8] Daniel V Bailey, Markus Dürmuth, and Christof Paar. 2014. Typing passwords with voice recognition: How to authenticate to Google Glass. In *Proc. of the Symposium on Usable Privacy and Security*.

[9] Denis Baldisserra, Annalisa Franco, Dario Maio, and Davide Maltoni. 2006. Fake fingerprint detection by odor analysis. In *International Conference on Biometrics*. Springer, 265–272.

[10] Marian E Berryhill, Lisa Phuong, Lauren Picasso, Roberto Cabeza, and Ingrid R Olson. 2007. Parietal lobe and episodic memory: bilateral damage causes impaired free recall of autobiographical memory. *Journal of Neuroscience* 27, 52 (2007), 14415–14423.

[11] BetterHealth. [n. d.]. EEG test. https://www.betterhealth.vic.gov.au/health/-conditionsandtreatments/eeg-test Accessed by September 17, 2017.

[12] Robert Bos, Stijn De Waele, and Piet MT Broersen. 2002. Autoregressive spectral estimation by application of the Burg algorithm to irregularly sampled data. *IEEE Transactions on Instrumentation and Measurement* 51, 6 (2002), 1289–1294.

[13] Kai Cao and Anil K Jain. 2016. Hacking Mobile Phones Using 2D Printed Fingerprints. *PasswordResearch* (2016).

[14] Chaos Computer Club (CCC). 2014. Fingerprint biometrics hacked again. http://www.ccc.de/en/updates/2014/ursel. Accessed by May 13, 2017.

[15] Chih-Chung Chang and Chih-Jen Lin. [n. d.]. LIBSVM – A Library for Support Vector Machines. https://www.csie.ntu.edu.tw/~cjlin/libsvm/

[16] Jagmohan Chauhan, Hassan Jameel Asghar, Anirban Mahanti, and Mohamed Ali Kaafar. 2016. Gesture-Based Continuous Authentication for Wearable Devices: The Smart Glasses Use Case. In *International Conference on Applied Cryptography and Network Security*. Springer, 648–665.

[17] John Chuang, Hamilton Nguyen, Charles Wang, and Benjamin Johnson. 2013. I think, therefore i am: Usability and security of authentication using brainwaves. In *International Conference on Financial Cryptography and Data Security*. Springer, 1–16.

[18] Tee Connie, Andrew Teoh, Michael Goh, and David Ngo. 2005. Palmhashing: a novel approach for cancelable biometrics. *Information processing letters* 93, 1 (2005), 1–5.

[19] Bernardo Dal Seno, Matteo Matteucci, and Luca Mainardi. 2008. A genetic algorithm for automatic feature extraction in P300 detection. In *IEEE International Joint Conference on Neural Networks*. IEEE, 3145–3152.

[20] Ian Daly, Martin Billinger, Reinhold Scherer, and Gernot Müller-Putz. 2013. On the automated removal of artifacts related to head movement from the EEG. *IEEE Transactions on neural systems and rehabilitation engineering* 21, 3 (2013), 427–434.

[21] A Delorme, S Makeig, TZ Jung, and TJ Sejnowski. 2001. Automatic rejection of event-related potential trials and components using independent component analysis. In *Society for Neuroscience Abstracts*, Vol. 27.

[22] Howard Eichenbaum. 2000. A cortical–hippocampal system for declarative memory. *Nature Reviews Neuroscience* 1, 1 (2000), 41.

[23] Eberhard E Fetz. 2007. Volitional control of neural activity: implications for brain–computer interfaces. *The Journal of physiology* 579, 3 (2007), 571–579.

[24] Benjamin Friedlander and Boaz Porat. 1984. The modified Yule-Walker method of ARMA spectral estimation. *IEEE Trans. Aerospace Electron. Systems* 2 (1984), 158–173.

[25] Javier Galbally, Fernando Alonso-Fernandez, Julian Fierrez, and Javier Ortega-Garcia. 2012. A high performance fingerprint liveness detection method based on quality related features. *Future Generation Computer Systems* 28, 1 (2012), 311–321.

[26] Thomas D Gautheir. 2001. Detecting trends using Spearman's rank correlation coefficient. *Environmental forensics* 2, 4 (2001), 359–362.

[27] Google. 2016. Google Glass. https://www.google.com/glass/start/.

[28] Joseph T Gwin, Klaus Gramann, Scott Makeig, and Daniel P Ferris. 2010. Removal of movement artifact from high-density EEG recorded during walking and running. *Journal of neurophysiology* 103, 6 (2010), 3526–3534.

[29] James V Haxby, Leslie G Ungerleider, Barry Horwitz, Jose Ma Maisog, Stanley I Rapoport, and Cheryl L Grady. 1996. Face encoding and recognition in the human brain. *Proceedings of the National Academy of Sciences* 93, 2 (1996), 922–927.

[30] Ulrich Hoffmann, Jean-Marc Vesin, Touradj Ebrahimi, and Karin Diserens. 2008. An efficient P300-based brain–computer interface for disabled subjects. *Journal of Neuroscience methods* 167, 1 (2008), 115–125.

[31] Shujie Hou and Robert Caiming Qiu. 2014. Kernel feature template matching for spectrum sensing. *IEEE Transactions on Vehicular Technology* 63, 5 (2014), 2258–2271.

[32] Shikha Jain and Gopikrishna Deshpande. 2004. Parametric modeling of brain signals. In *Biotechnology and Bioinformatics, 2004. Proceedings. Technology for Life: North Carolina Symposium on*. IEEE, 85–91.

[33] Boutheina Jemel, Michèle Pisani, Marco Calabria, Marc Crommelinck, and Raymond Bruyer. 2003. Is the N170 for faces cognitively penetrable? Evidence from repetition priming of Mooney faces of familiar and unfamiliar persons. *Cognitive Brain Research* 17, 2 (2003), 431–446.

[34] Søren Johansen. 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society* (1991), 1551–1580.

[35] Torkel Klingberg, Maj Hedehus, Elise Temple, Talya Salz, John DE Gabrieli, Michael E Moseley, and Russell A Poldrack. 2000. Microstructure of temporo-parietal white matter as a basis for reading ability: evidence from diffusion tensor magnetic resonance imaging. *Neuron* 25, 2 (2000), 493–500.

[36] Sugang Li, Ashwin Ashok, Yanyong Zhang, Chenren Xu, Janne Lindqvist, and Macro Gruteser. 2016. Whose move is it anyway? Authenticating smart wearable devices using unique head movement patterns. In *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 1–9.

[37] David Liu, Simon A Jenkins, Penelope M Sanderson, Perry Fabian, and W John Russell. 2010. Monitoring with head-mounted displays in general anesthesia: a clinical evaluation in the operating room. *Anesthesia & Analgesia* 110, 4 (2010), 1032–1038.

[38] Joseph K Liu and Ron Steinfeld. 2016. *Information Security and Privacy: 21st Australasian Conference, ACISP 2016, Melbourne, VIC, Australia, July 4-6, 2016, Proceedings*. Vol. 9723. Springer.

[39] Jukka Määttä, Abdenour Hadid, and Matti Pietikäinen. 2011. Face spoofing detection from single images using micro-texture analysis. In *International Joint Conference on Biometrics*. IEEE, 1–7.

[40] Joseph N Mak, Dennis J McFarland, Theresa M Vaughan, Lynn M McCane, Phillippa Z Tsui, Debra J Zeitlin, Eric W Sellers, and Jonathan R Wolpaw. 2012. EEG correlates of P300-based brain–computer interface (BCI) performance in people with amyotrophic lateral sclerosis. *Journal of neural engineering* 9, 2 (2012), 026014.

[41] Markets and Markets. 2015. Global Head-Mounted Display Market 2016-2020. http://www.marketsandmarkets.com/Market-Reports/head-mounted-display-hmd-market-729.html.

[42] Ivan Martinovic, Doug Davies, Mario Frank, Daniele Perito, Tomas Ros, and Dawn Song. 2012. On the Feasibility of Side- channel Attacks with Brain- computer Interfaces. In *Proceedings of the 21st USENIX conference on Security symposium*. USENIX Association.

[43] Richard Matovu and Abdul Serwadda. 2016. Your substance abuse disorder is an open secret! Gleaning sensitive personal information from templates in an EEG-based authentication system. In *IEEE 8th International Conference on Biometrics Theory, Applications and Systems*. IEEE, 1–7.

[44] Tom M Mitchell et al. 1997. Machine learning. WCB.

[45] Karthik Nandakumar and Anil K Jain. 2015. Biometric template protection: Bridging the performance gap between theory and practice. *IEEE Signal Processing Magazine* 32, 5 (2015), 88–100.

[46] Osama Ouda, Norimichi Tsumura, and Toshiya Nakaguchi. 2010. Tokenless cancelable biometrics scheme for protecting iris codes. In *20th International Conference on Pattern Recognition*. IEEE, 882–885.

[47] Ramaswamy Palaniappan and Danilo P Mandic. 2007. Biometrics from brain electrical activity: A machine learning approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 4 (2007), 738–742.

[48] Ken A Paller and Anthony D Wagner. 2002. Observing the transformation of experience into memory. *Trends in cognitive sciences* 6, 2 (2002), 93–102.

[49] Padma Polash Paul and Marina Gavrilova. 2012. Multimodal cancelable biometrics. In *IEEE 11th International Conference on Cognitive Informatics & Cognitive Computing*. IEEE, 43–49.

[50] PlayStation. 2016. PlayStation VR. https://www.playstation.com/en-us/explore/playstation-vr/.

[51] John Polich. 1997. EEG and ERP assessment of normal aging. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 104, 3 (1997), 244–256.

[52] M Poulos, M Rangoussi, N Alexandris, A Evangelou, et al. 2002. Person identification from the EEG using nonlinear signal classification. *Methods of information in Medicine* 41, 1 (2002), 64–75.

[53] Rodrigo Quian Quiroga and Stefano Panzeri. 2009. Extracting information from neuronal populations: information theory and decoding approaches. *Nature reviews. Neuroscience* 10, 3 (2009), 173.

[54] Toni M Rath and Raghavan Manmatha. 2003. Word image matching using dynamic time warping. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2. IEEE.

[55] Cynthia E Rogers, Alexander W Witt, Alexander D Solomon, and Krishna K Venkatasubramanian. 2015. An approach for user identification for head-mounted displays. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 143–146.

[56] Maria V Ruiz-Blondet, Zhanpeng Jin, and Sarah Laszlo. 2016. CEREBRE: A Novel Method for Very High Accuracy Event-Related Potential Biometric Identification. *IEEE Transactions on Information Forensics and Security* 11, 7 (2016), 1618–1629.

[57] Samsung. 2015. Samsung VR. http://www.samsung.com/us/explore/gear-vr/?cid=van-mb-cph-0716-10000089.

[58] Ralph Schmidt. 1986. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation* 34, 3 (1986), 276–280.

[59] Yogendra Narain Singh, Sanjay Kumar Singh, and Amit Kumar Ray. 2012. Bioelectrical signals as emerging biometrics: Issues and challenges. *ISRN Signal Processing* (2012).

[60] Chen Song, Aosen Wang, Kui Ren, and Wenyao Xu. 2016. "EyeVeri: A Secure and Usable Approach for Smartphone User Authentication". In *IEEE International Conference on Computer Communication (INFOCOM'16)*. San Francisco, California, 1 – 9.

[61] Samuel Sutton, Margery Braren, Joseph Zubin, and ER John. 1965. Evoked-potential correlates of stimulus uncertainty. *Science* 150, 3700 (1965), 1187–1188.

[62] James W Tanaka, Tim Curran, Albert L Porterfield, and Daniel Collins. 2006. Activation of preexisting and acquired face representations: the N250 event-related potential as an index of face familiarity. *Journal of Cognitive Neuroscience* 18, 9 (2006), 1488–1497.

[63] Jason I Thompson. 2005. *A three dimensional helmet mounted primary flight reference for paratroopers*. Technical Report. DTIC Document.

[64] Julie Thorpe, Paul C van Oorschot, and Anil Somayaji. 2005. Pass-thoughts: authenticating with our minds. In *Proceedings of the 2005 workshop on New security paradigms*. ACM, 45–56.

[65] Endel Tulving et al. 1972. Episodic and semantic memory. *Organization of memory* 1 (1972), 381–403.

[66] Anthony D Wagner, Benjamin J Shannon, Itamar Kahn, and Randy L Buckner. 2005. Parietal lobe contributions to episodic memory retrieval. *Trends in cognitive sciences* 9, 9 (2005), 445–453.

[67] Liang Wang and David Suter. 2006. Analyzing human movements from silhouettes using manifold learning. In *IEEE International Conference on Video and Signal Based Surveillance*. IEEE, 7 – 7.

[68] Julia Wendt, Martin Lotze, Almut I Weike, Norbert Hosten, and Alfons O Hamm. 2008. Brain activation and defensive response mobilization during sustained exposure to phobia-related and other affective pictures in spider phobia. *Psychophysiology* 45, 2 (2008), 205–215.

[69] JJ Wright. 1999. Simulation of EEG: dynamic changes in synaptic efficacy, cerebral rhythms, and dissipative and generative activity in cortex. *Biological cybernetics* 81, 2 (1999), 131–147.

[70] Seul-Ki Yeom, Heung-Il Suk, and Seong-Whan Lee. 2013. Person authentication from neural activity of face-specific visual self-representation. *Pattern Recognition* 46, 4 (2013), 1159–1169.