CIBA: Continuous Interruption-free Brain Authentication

Vom Promotionsausschuss der
Technischen Universität Hamburg
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von
Florian Gondesen

aus
Preetz

2023

Gutachter:
Prof. Dr. Dieter Gollmann
Prof. Dr. Hoc Khiem Trieu

Datum der mündlichen Prüfung:
27.06.2022

Abstract

Authentication schemes which use biometric features of the electroencephalogram (EEG) or use an EEG-based brain-computer interface (BCI) to enter a secret suffer from a low signal to noise ratio due to the properties of the brain waves and the measurement on the scalp. This results in a low information transfer rate (ITR), which requires relatively long sessions to transfer a secret of sufficient entropy or sufficient discriminant information. Hence, there is a trade-off between security and usability. In cases where an electroencephalography (EEG) headset is not already worn, the usability is further reduced by a setup time of several minutes.

Aiming for a high ITR, we proposed a shoulder surfing resistant authentication scheme using a BCI based on the P300 component of the event-related potential (ERP), following the oddball paradigm. Rapid serial visual presentation (RSVP) is used to swiftly present a set of 100 images of which 5 are used as target or password stimuli. The authentication scheme was tested using a consumer grade EEG headset. Parameters were varied: simple drawings or photos as stimuli and allowing the user to select their own password. Varying these parameters did not exhibit significant differences. The experiments were repeated after some time to verify permanency, with 16 subjects participating in 3 sessions. The performance increased with repeated sessions, indicating a training effect. The P300 was classified by linear discriminant analysis (LDA) trained on several users and for comparision, individual classifiers were trained for each user. User-specific classifiers increased the performance, reaching equal error rates (EERs) below 0.09.

We furthermore investigated the permanency of a subject's brain responses to flickering lights (steady state visually evoked potential (SSVEP)). Experiments were conducted using consumer grade EEG headset. In five sessions we found the individual response spectra only to vary slightly and evaluate a biometric authentication system based on this feature. Among the four subjects an EER of 0 was reached.

As an alternative to increasing the usability of EEG-based authentication by increasing the ITR, we describe a scenario where the ITR is of lesser importance for the usability: the continuous authentication of a person wearing an EEG headset. In this scenario the user should be able to do regular work concurrently and should not be interrupted or distracted by the authentication process. Hence we conceptualize continuous interruption-free brain authentication (CIBA) and discuss how subliminal SSVEP or ERP paradigms could be used in a way that does not interrupt or distract the user or reduce the user's performance in the regular work.

Contents

1. Introduction

With an increasing number of diverse digital services used, the classic means of user authentication, an individual password, becomes less appropriate as users may be unable to cope with memorizing numerous secure passwords. Biometrics offer a convenient alternative and have therefore become popular for mobile devices. However, mainstream biometrics based on fingerprints or faces are very susceptible to presentation attacks ('spoofing attacks'). Templates to create a fake finger or mask can be captured with a smartphone camera [27]. Ubiquitous cameras operated by diverse entities may contribute to a vast availability of footage containing faces. Social networks may help to map the images to identities, even without the person being aware of this. Biometrics that cannot easily be captured remotely may be a way to mitigate presentation attacks. Emerging biometrics, using electrophysiological features, like electrocardiography (ECG) or electroencephalography (EEG), are difficult to capture remotely due to low frequencies and low voltages. To capture a sample, electrical sensors need to be placed on the subject's body. Unlike the electrocardiogram (ECG) that can hardly be voluntarily altered, the electroencephalogram (EEG) contains components that depend on voluntary activity, which may impede an adversary's efforts to covertly obtain a sample. These components of the EEG also allow creating a communication channel to the authentication system, rendering it a brain-computer interface (BCI). This BCI may be used to covertly enter a password to add a knowledge factor to the authentication system. Thus an EEG-based authentication system can be designed using biometrics and knowledge factors making it resistant to attacks aiming to visually obtain the biometric template or the password. Due to the noisy nature of the EEG, the information transfer rate (ITR) is quite low, requiring EEG recording sessions of several minutes to obtain sufficient discriminant information or to enter a secure password. Emerging consumer grade EEG headsets that would be key for mass adoption of EEG-based authentication, typically acquire signals with a quality inferior to a research grade electroencephalograph (EEG), further reducing the ITR. To tackle this issue, this thesis investigates how brain responses which are well-studied in the field of BCIs can be used to create secure authentication schemes which either cope with a short data sample or can be used in a continuous authentication scheme which does not distract the user from regular work.

1.1. Ethical Considerations

Biometric data is generally considered personal data or personally identifiable information (PII) [62]. Using EEG as biometric may be particularly sensitive due to relations between the electrical activity of the brain and the person's thought processes. Though currently reading a person's mind by EEG does not seem feasible, several studies showed that personal knowledge or traits may be derived from EEG. In specifically designed experiments, Martinovic et al. [60] obtained information from EEG helping to reduce the entropy in guessing attacks on a subject's bank, pin, known faces, month of birth, and geographic location. Vance et al. [112] used EEG to predict risk taking behavior. In a social pressure experiment, Trautmann et al. [108] found differences in the EEG between low and high autonomous subjects. Inzlicht et al. [42] found markers for religious conviction in the EEG. Age, certain illnesses,

addictions, or drug use may also influence the EEG [83, 99, 68]. While some of the above examples of personal data may require a specifically designed experiment, others may be extracted from EEG recordings from similar experiments or even any EEG recording. As a precaution EEG data needs to be treated as sensitive.

For the experiments conducted in the two studies presented in this thesis (see Sects. 4.1 and 5.1), the participants were first fully informed about the purpose and content of the respective study. The participants were able to agree to participate and did not belong to groups that are particularly vulnerable. No questions that are of an intimate nature were asked. It was not expected that the participants would suffer negative effects from the studies. The participants were informed about the processing of personal data. Wherever possible, the data were processed and stored in pseudonymized form. Participation was not remunerated.

1.2. Structure of the Dissertation and Contributions

This section provides short summaries of the chapters of this thesis, including, if applicable, the publications they are based on.

Chapter 1 - Introduction The motivation of exploring novel EEG-based authentication methods to cope with the relatively long data acquisition times has been given above.

Chapter 2 - Fundamentals The second chapter describes the fundamentals of EEG and the paradigms and methods used in this work. In addition, authentication and the respective security parameters are explained (Sect. 2.3).

The EEG related fundamentals are based on the fundamentals in “EEG-Based Biometrics”[33] by Gondesen, Marx, and Gollmann, in Biometric-Based Physical and Cybersecurity Systems (editors Obaidat, Traore, Woungang). The fundamentals for this publication were mainly contributed by Gondesen, the main author.

Chapter 3 - EEG-Based Authentication The third chapter surveys the research on EEG-based authentication, focusing on the findings relevant for this thesis.

The starting point for this survey was the literature review conducted for the section “Methods for EEG-Based Biometric Systems” in “EEG-Based Biometrics”[33] by Gondesen, Marx, and Gollmann, where Gondesen reviewed the literature on event-related potential (ERP)-based biometrics.

Chapter 4 - A P300 BCI for visual authentication We proposed a BCI for visual authentication that presents images as passwords in a shoulder surfing resistant way. Images are used to help the user remembering the password and to facilitate recognition of a huge variety of stimuli. Password images and non-password images are shown on the screen in fast succession. The user’s task is to silently count the occurrences of password images, which entails a P300 component in the ERP. We experimentally investigated several parameters:

- The influence of the number of repetitions of the stimuli set and hence the length of the authentication process on the error rates.
- Whether there is a difference in using simple drawings or photos as stimuli.

- Whether there is a difference when users are allowed to select their own password or not.
- The experiments were repeated after some time to investigate permanency and training effects.
- Usage of general classifiers and user-specific classifiers.

The main results of this study were published as "A Shoulder-Surfing Resistant Image-Based Authentication Scheme with a Brain-Computer Interface" [34], by Gondesen, Marx, and Kycler, presented on 4th October at the 2019 International Conference on Cyberworlds (CW) by Gondesen. This chapter additionally presents the survey, additional analyses and an extended discussion. Apparent differences in experiment results are mainly based on different sets of subjects considered. As the main author, Gondesen contributed to the study design, conducted the experiments, and analysed the results.

Chapter 5 - SSVEP Biometrics The fifth chapter reports on a small study that was conducted to evaluate the permanency of the brain responses to flickering lights (steady state visually evoked potential (SSVEP)) and proposes a biometric authentication scheme based on this.

Chapter 6 - CIBA: Continuous Interruption-free Brain Authentication In this chapter, we describe a scenario where a user is required to be continuously authenticated during security sensitive work. We conceptualize a continuous authentication scheme and discuss how SSVEP and ERP could be used in a way that does not interrupt or distract the user or reduces the user's performance in the regular work. The continuous interruption-free brain authentication (CIBA) concept was published as "CIBA: Continuous Interruption-free Brain Authentication" [32] by Gondesen and Gollmann, and presented as a poster at the BIOSIGNALS 2021 conference.

Chapter 7 - Conclusion In the concluding chapter we review the presented results and their implications.

2. Fundamentals

2.1. Electroencephalography

Electroencephalography (EEG) is the measurement of brain-originating electrical signals on the scalp. Due to attenuation by multiple layers of tissue and bone, synchronous activation of a large number of neurons is required to produce potentials measurable on the scalp. As these potentials are in the range of a few microvolts, precise measurement remains a challenge. Nevertheless, the first electroencephalogram (EEG), the graphical representation of EEG readings, of a human subject was recorded by Hans Berger in 1924 [9]. Recent EEG measuring devices record digital data, hence we use the terms EEG data and electroencephalogram synonymously. The EEG is typically sampled at relatively low rates between $128\,\mathrm{Hz}$ and $2048\,\mathrm{Hz}$. These rates are sufficient as the spectrum is similar to pink noise, having most of the power at low frequencies. Fig. 2.1 depicts the EEG of a subject resting with open eyes. The corresponding power spectral density is shown in Fig. 2.2. The frequency range below $21\,\mathrm{Hz}$ contains 90% of the total power. The time-frequency representation is shown in Fig. 2.3.

The EEG is often divided into certain frequency bands to capture rhythmic activity. Activity in those bands is associated with different mental states [45, 99]. The different bands are shown in Tab. 2.1.

2.1.1. Electrode Placement

To map the spatial distribution of the EEG, multiple electrodes (channels) can be used, allowing a spatial resolution in the range of few centimeters. The international 10-20 system provides rules for electrode placement and labelling (see Fig. 2.4). Capital letters 'F', 'P', 'O', 'T', and 'C' refer to the lobes of the brain: frontal, parietal, occipital, temporal, and central respectively (see Fig. 2.5). Electrode positions at the frontal pole region are identified by 'Fp', at the earlobe by a capital 'A'. Odd numbers refer to electrodes on the left, even numbers to electrodes on the right hemisphere. Electrode positions on the midline are identified by the letter 'z'.

2.1.2. Artifacts

Due to the weak signals measured, EEG is highly susceptive to noise. Components of the EEG not originating from the brain are called artifacts. To cope with the pervasive power grid, EEG systems

Table 2.1.: EEG frequency bands [33].

Band	Bandwidth (Hz)	Region	Associated with...
Delta (δ)	0.5 – 4	varies	deep (dreamless) sleep
Theta (θ)	4 – 7.5	varies	creative inspiration, deep meditation, drowsiness
Alpha (α)	8 – 13	occipital and parietal	physical and mental relaxation
Beta (β)	14 – 26	frontocentral	active thinking, active attention
Gamma (γ)	> 30	frontocentral	active information processing, processing of sensory stimuli

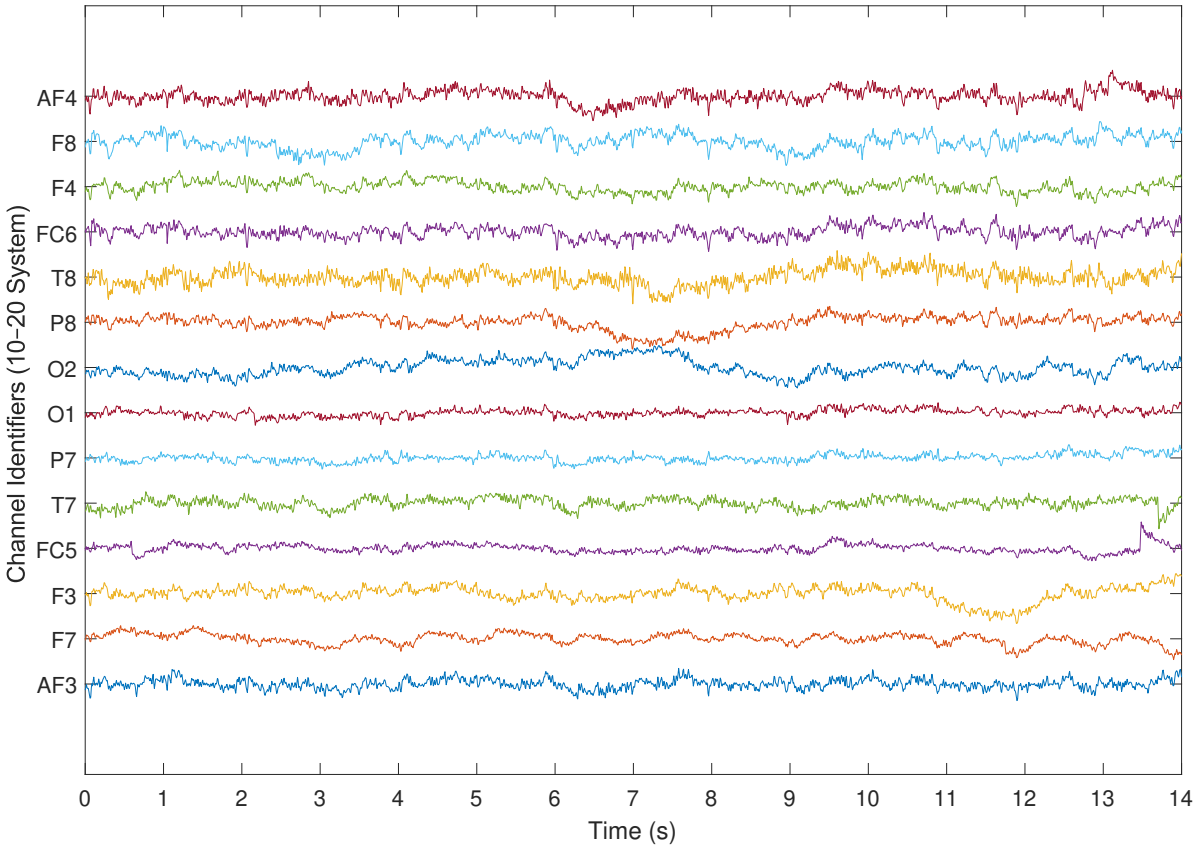


Figure 2.1.: EEG recorded with an EMOTIV Epoc EEG headset with 14 channels [33].

usually employ notch filters at $50\,\mathrm{Hz}$ and $60\,\mathrm{Hz}$. In addition, EEG experiments may be conducted in electrically shielded rooms. Physiological artifacts are caused by activities of the subject. Ocular artifacts, arising from eye blinks or eye movement, manifest themselves as spikes in the EEG at frontal electrodes.

Fig. 2.6 shows the EEG of a subject blinking after three visual cues. Approximately half a second after each cue, the EEG contains spikes prominent in the frontmost electrodes AF3 and AF4. The corresponding time-frequency representation is depicted in Fig. 2.7 [33]. If the experiment permits, a fixation cross (see Fig. 2.8) may help to reduce eye movements. A subject may also be instructed only to blink in designated breaks and to refrain from moving and swallowing, which may also produce artifacts otherwise. But such instructions need to be carefully considered as they may divert attention from the actual task of the subject. If artifacts cannot be prevented with these methods, trials containing artifacts may be excluded from analysis. Artifacts may be detected based on their typical features by an automated process according to certain thresholds or by visual inspection by the experimenter. The latter poses a risk of introducing a bias.

2.1.3. Event-Related Potentials

An ERP is the electrical response of the brain to a sensory, motor, or cognitive event. Compared to the background activity of the brain, ERP amplitudes are relatively small, ranging from $1\,\mu\mathrm{V}$ to $30\,\mu\mathrm{V}$ [99, p.125]. Therefore it is common practice to capture the EEG of multiple identical events and average the data time-locked to the onset of the event, averaging out EEG components independent of the event and yielding the ERP waveform.

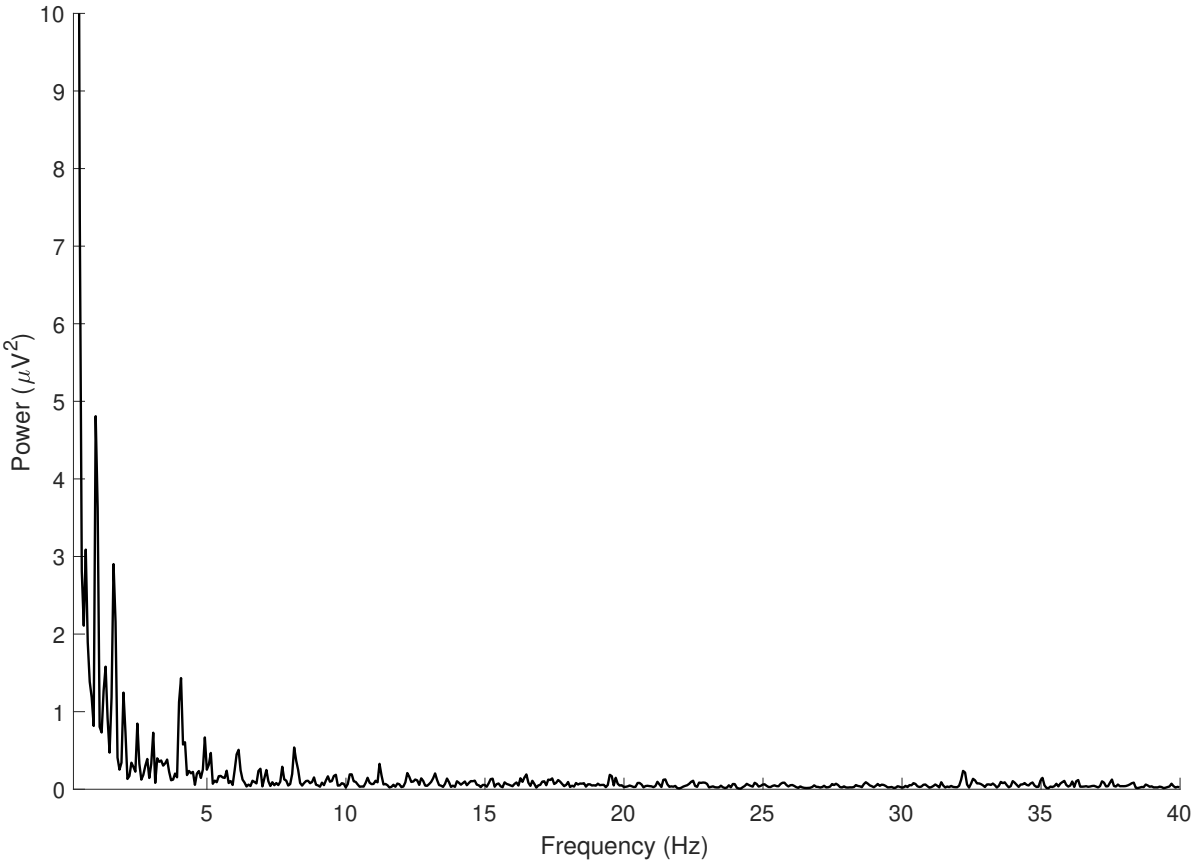


Figure 2.2.: Spectrum, averaged over all channels of the EEG shown in Fig. 2.1 [33].

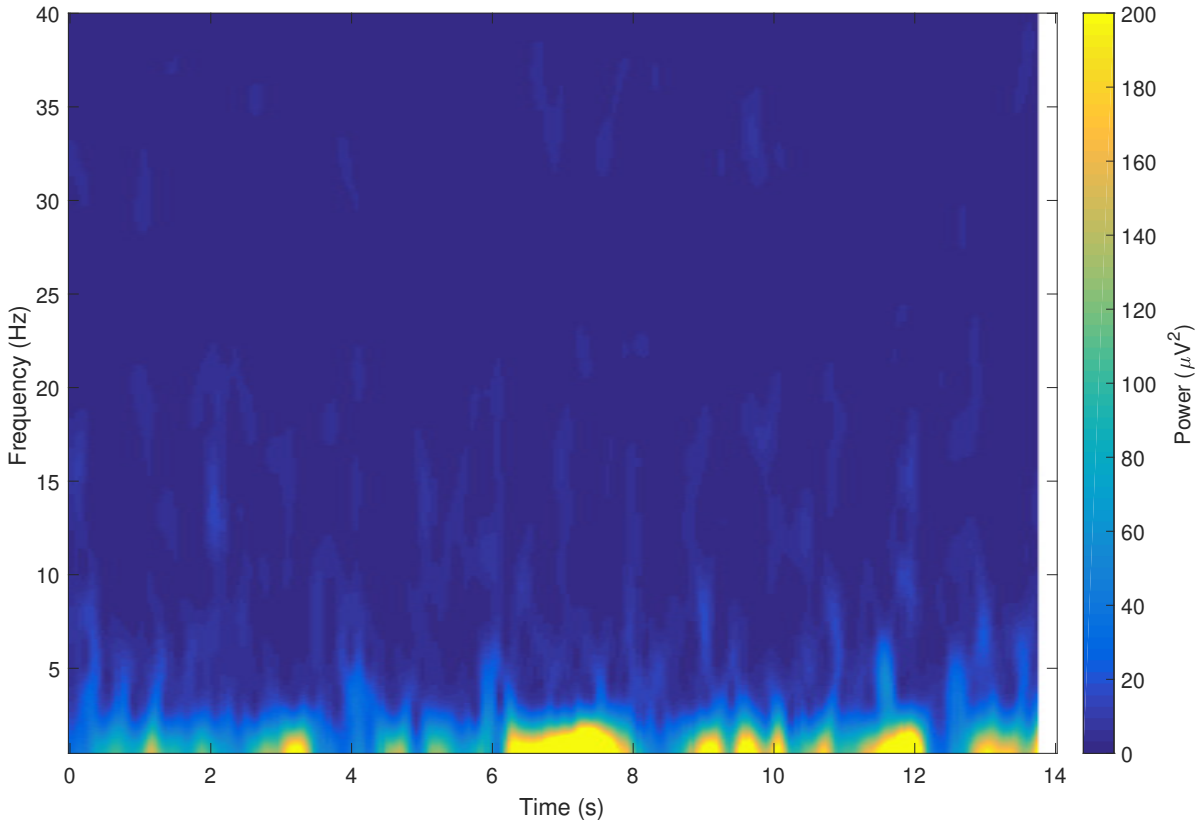


Figure 2.3.: Time-frequency representation of the EEG readings shown in Fig. 2.1 averaged over all channels. A good portion of the light yellow areas represent power beyond scale [33].

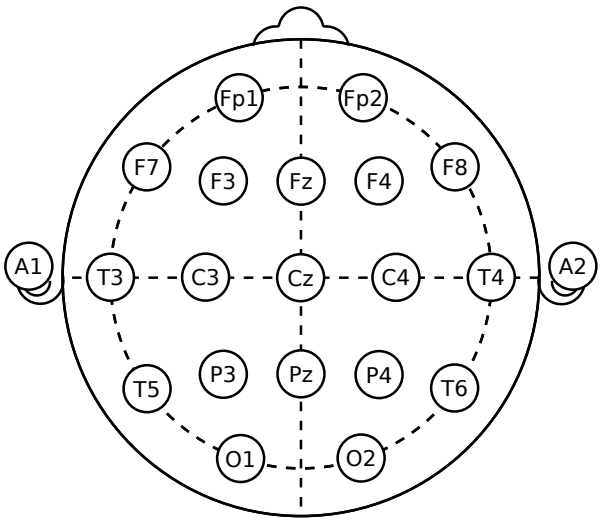


Figure 2.4.: Alignment of 21 electrodes by the 10-20 system [6]

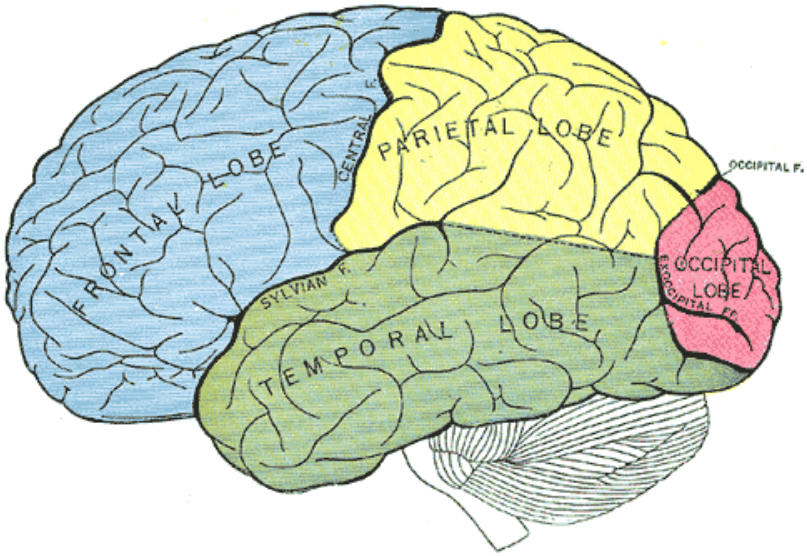


Figure 2.5.: Lobes of the brain [35, Fig. 728]

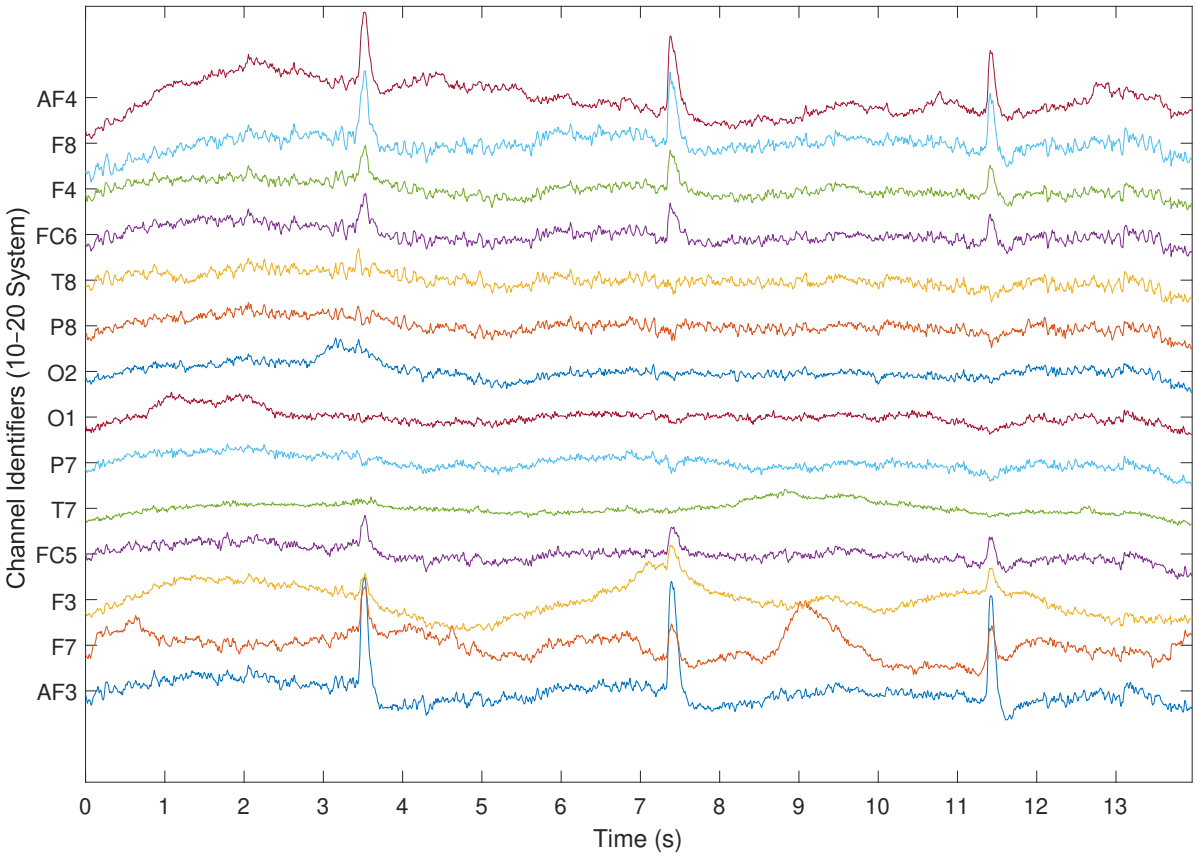


Figure 2.6.: EEG recorded with an EMOTIV Epoc EEG headset with 14 channels in the 10-20 system. The subject was resting and cued to blink at seconds three, seven and eleven [33].

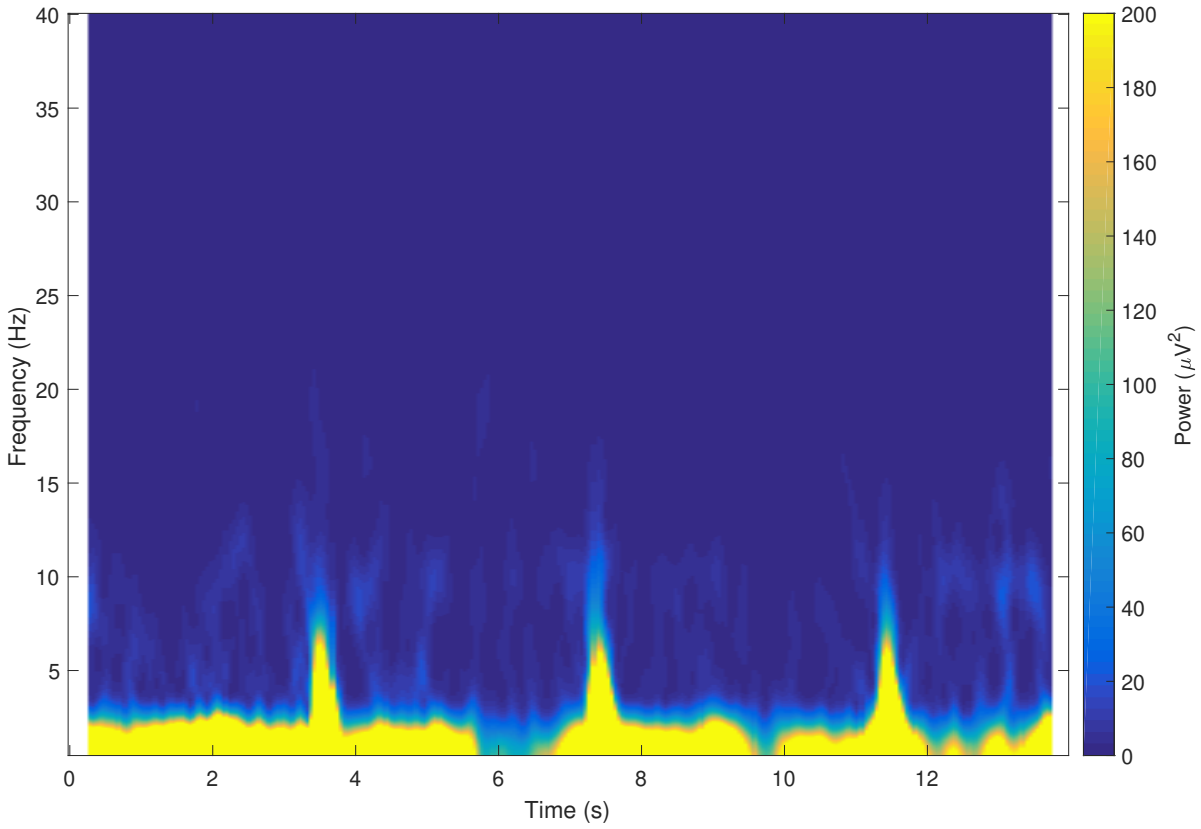


Figure 2.7.: Time-frequency representation of the EEG readings shown in Fig. 2.6, averaged over all channels. A good portion of the light yellow areas represent power beyond scale [33].

Figure 2.8.: Fixation cross on gray background [33].

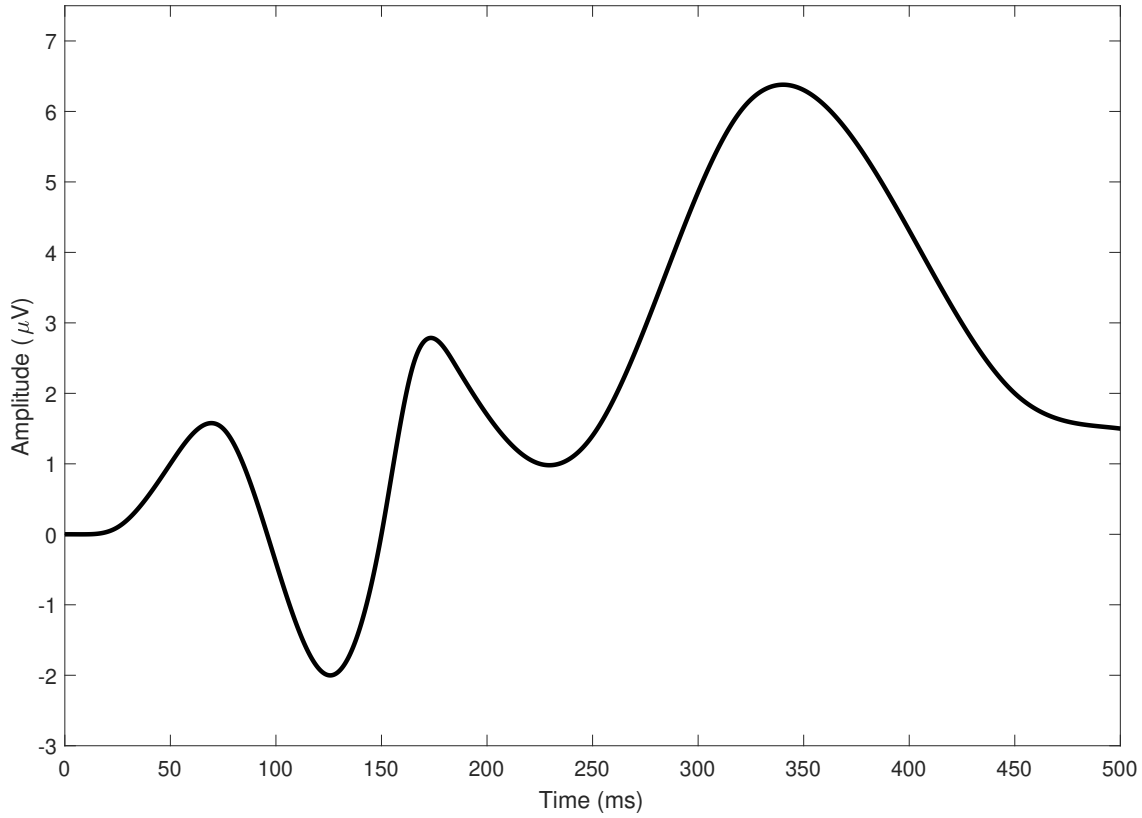


Figure 2.9.: ERP with P300 component [33].

Components of the ERP are named P or N, for positive and negative deflections respectively, followed by the number of the respective deflection or alternatively by peak latency in milliseconds after the event. Components are characterized by a certain time window around the peak latency and the position of maximum amplitude on the scalp. They also depend on certain conditions of the event. Latencies, amplitudes, and topographies vary with the event and may also be influenced by the subject's reactions or attitudes towards a stimulus and by the subject's physiology.

P300

The P300 or P3 is a cognitive ERP component in a time window between $220\,\mathrm{ms}$ and $500\,\mathrm{ms}$ after stimulus onset [83]. Fig. 2.9 shows an exemplary P300. The P300 can be divided into two overlapping subcomponents, P3a and P3b, with average peak latencies of $240\,\mathrm{ms}$ and $350\,\mathrm{ms}$, and peak amplitude locations frontocentral and parietocentral respectively [103]. The P3b, the 'classic' P300, can be found in an experiment following the oddball paradigm. It consists of frequent irrelevant standard stimuli and infrequent target stimuli. The subject is instructed to react upon the target stimuli, for example counting the occurrences. The P3b is elicited in the rare and relevant target condition. To distinguish the P3a, the experiment may be extended by an additional stimulus class that is irrelevant to the user's task but appears infrequently and deviates from the standard non-target stimuli. The P3a is elicited by the unexpected deviant stimuli, but repeated presentation entails habituation which reduces the P3a amplitude. Therefore the P3a is sometimes called 'novelty P3'. The P3b is not affected by habituation, but by the subject's attention and the frequency of target occurrences.

Figure 2.10.: Person wearing an EMOTIV Epoc EEG headset [33]

2.1.4. SSVEP

SSVEPs are the response to repetitive visual stimuli like a flickering light. The EEG of a subject typically exhibits activity at the frequency of the stimulus and some of its (sub)harmonics. SSVEP is mainly recorded at the occipital electrodes, where the visual cortex resides. Simulation frequencies used for SSVEP are typically in the range of about $3.5\,\mathrm{Hz}$ to $75\,\mathrm{Hz}$ [41, 3]. Similar to the pink noise characteristics of the EEG, the response amplitudes decrease with increasing frequency. Therefore, when comparing the SSVEP to the non-stimulus baseline, the optimal signal-to-noise ratio (SNR) mainly depends on the individual. Other stimulus parameters to optimize the SNR include duty cycle, color, and contrast between the alternating stimuli [100].

2.1.5. EMOTIV EPOC

The EPOC, a wireless consumer grade EEG headset by EMOTIV, is shown in Fig. 2.10. It has 16 electrodes which use felt pads soaked in saline solution at the positions AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, P3, and P4. The electrodes at P3 and P4 are used for referencing while the other 14 channels are sequentially sampled by a single analog-to-digital converter (ADC) with $2048\,\mathrm{Hz}$. This is internally downsampled by the device, only allowing to record data at a sample rate of $128\,\mathrm{Hz}$. The EPOC+ variant allows to alternatively record data with $256\,\mathrm{Hz}$, though both variants have a bandwidth of $0.2\,\mathrm{Hz}$ to $45\,\mathrm{Hz}$ [24]. To record EEG data on a PC, Emotiv provides a dedicated USB receiver and the software 'Testbench'. Unlike research grade EEG systems, the EPOC is not equipped with a hardware trigger channel. To synchronize experiments with the data, Testbench allows to add an eight bit trigger channel, called marker, which can be fed from a serial port.

EMOTIV claims "fast set up" for the EPOC of $3\,\mathrm{min}$ to $5\,\mathrm{min}$ [25]. This matches our finding of average $197\,\mathrm{s}$, a standard deviation of $138\,\mathrm{s}$, and additionally approximately $3\,\mathrm{min}$ for the preparation of the headset [34].

The EPOC headband is designed to fit all sizes. In the study presented in Ch. 4, subjects wore the headset for sessions of approximately $50\,\mathrm{min}$ during which 9% of the subjects reported pain or pressure inflicted by the headset (see Sect. 4.2.7).

2.2. Brain-Computer Interfaces

A BCI is a system that enables a user to transmit data to a machine without using peripheral nerves or muscles [116]. Instead, a data acquisition device captures data from brain activity. The user is required to do a mental task which enables the user to intentionally change features of the brain activity. These features are extracted, classified, and translated to a suitable output, which can for example be a command for an electric wheelchair to move forward. In this example, the user would experience the forward movement as feedback. By definition, feedback is not a necessity for BCIs but typically at least the output, for example the character typed, is presented for usability. Continuous feedback allows the user to train the mental task. Some BCI paradigms rely on presenting stimuli to the user. These paradigms typically depend on the stimulus timing, requiring synchronization with data acquisition. For this purpose, data acquisition systems offer a trigger channel.

Fig. 2.11 depicts the components of a BCI. For each component there exist multiple approaches or methods. The choice of these building blocks is to some degree interdependent. Every paradigm requires an acquisition system and signal processing which are actually able to measure and extract the task dependent feature.

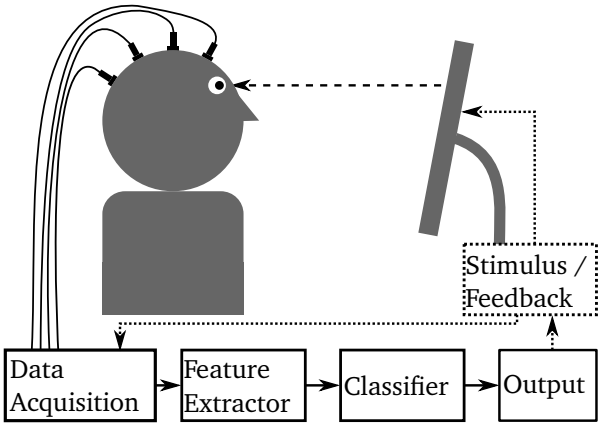


Figure 2.11.: Components of a BCI. The user performs a mental task. Brain activity is acquired to extract features for a classifier to determine the output. The methods for the individual components can usually be selected independently of each other. Presentation of stimuli or feedback are optional.

2.2.1. P300 Speller

A common variant of a P300-based BCI is the P300 speller. A matrix of characters is shown on a screen. The user focuses on the letter to be typed. In random order, single columns or rows are lit up for a short time. If the letter to be typed up is lit up, this creates the rare and relevant event entailing a P300 component in the EEG. By detecting a P300 component, the system can conclude from the timing which column or row caused the P300. Thus, to identify a single character, it must be lit up at least once columnwise and once rowwise. Additional trials may be included for averaging to increase P300 detection accuracy at the cost of an extended response time. The system is typically tuned to achieve the maximum ITR, which depends on accuracy and response time. Fig. 2.12 shows screenshots of the operation of a P300 speller. The rate of people who are not able to accurately use a BCI based on P300 is very low, i.e., the BCI illiteracy is low [23].

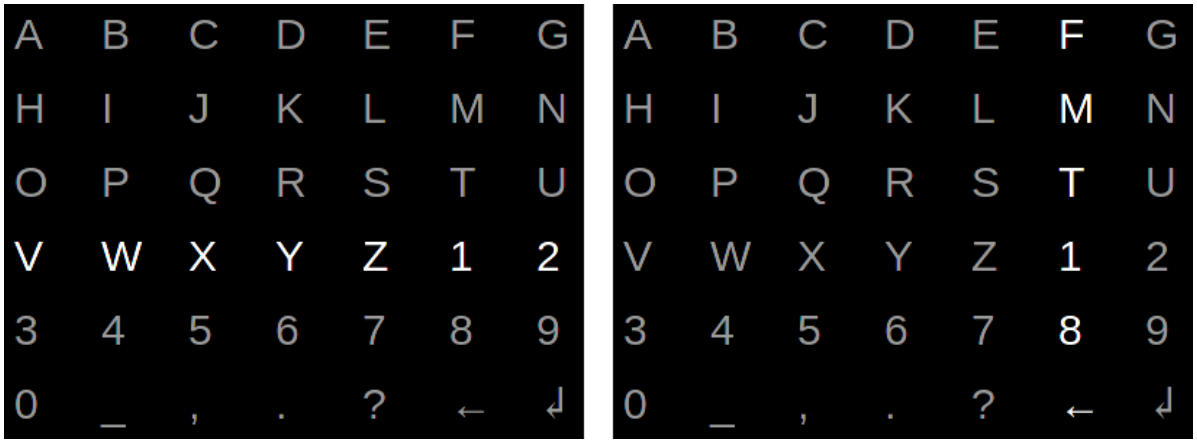


Figure 2.12.: Screenshots of a P300 speller. The symbols are randomly lit up column- and rowwise.

2.2.2. SSVEP BCI

BCIs based on SSVEP use multiple different flicker stimuli, simultaneously delivered by a screen or precisely controllable lights. The user observes the flicker stimuli, each contributing an SSVEP response to the EEG. The stimuli differ in frequency, phase, or sequence pattern, so that the elicited SSVEP responses can be distinguished. To input a symbol, the user focuses on a single stimulus, which relatively decreases the power of the SSVEP responses caused by the unattended stimuli, making the SSVEP response of the attended flicker stimulus dominant. Thus, by analysing shifts in the power of the SSVEP responses, the user's intention can be identified. As all stimuli are presented simultaneously, a SSVEP BCI can in principle be operated at the user's pace. In practice, the response time depends on the time the system requires to accurately identify the attended stimulus, which depends on how well the SSVEP responses can be differentiated. Therefore, to optimize the ITR, SSVEP BCIs typically operate at frequencies previously found to have the strongest SSVEP response effect for the respective user. The system design may limit the usable frequencies, not only due to technical limitations of the screen or lights used, but also due to practical considerations of the analysis as very similar stimuli may be hard to discriminate. Also, if the system relies on the analysis of harmonics, it is more difficult to discriminate stimulus frequencies, which are harmonics of another stimulus frequency used. The rate of people who are not able to accurately use a BCI based on SSVEP is very low, i.e., the BCI illiteracy is low [23, 36].

2.3. Identification and Authentication

Authentication is a process where a user provides proof for a claimed identity. Stating the identity is referred to as identification in this context. The most common means of user authentication is typing a password, typically preceded by a user name. This is an example for authentication based on a knowledge factor. Other possible factors are based on location, possession or are inherent to the users. Anyone in control of each factor of a user may authenticate as this user. This is straightforward for location or possession factors: The impostor must stay at the location or posses the authentication token. Distinct properties of knowledge and inherence factors are described in the following.

2.3.1. Password Authentication

A very commonly used knowledge factor is a password, which is a shared secret between a user and the system. Ideally, this is a 1:1 relation where the password is only stored in a secure place of the system

and in the user's memory. For authentication, the system asks the user for the password. If the reply matches the previously established password, the user is authenticated.

While machine to machine authentication allows to use shared secrets of very high entropy, in user authentication the entropy is limited by the user's ability and willingness to memorize a high entropy password, as there is a trade-off between the memorability and the entropy of a password. Due to the often low entropy of passwords used, guessing or brute force attacks are common. Passwords may also be obtained by intercepting a password during an authentication process. While in remote authentication the password can be kept secret by using a challenge response protocol with cryptographic hash functions, the transmission of the password from the user to the machine requires different methods that depend on the means of communicating the password.

Shoulder Surfing Learning a secret by observing a user entering it into a system, for example typing a password on a keyboard, is referred to as shoulder surfing. A standard countermeasure is visual blocking, for example when entering a PIN to a keypad, covering it by the other hand. One may extend the definition of shoulder surfing from visual to other means of observing the entering process.

2.3.2. Biometrics

An inherence factor may be based on biometrics, which means measurements of physical characteristics of individuals aiming to extract information that allows to tell individuals apart. We refer to this as **discriminant information**. The amount of discriminant information is sufficient if it allows to discriminate between all individuals in a group of a certain size with an acceptable accuracy. This is often referred to as **uniqueness**. For example, in a group of five people, the height may contain sufficient discriminant information. With an increasing group size it will become more likely that the group includes people of the same height, thus more discriminant information is required. This can to some extent be achieved by more precise measurements. Alternatively more features can be added, for example the weight. A feature that is fully unique for every individual contains maximal discriminant information.

In order to reproducibly discriminate individuals, biometric features must not change beyond a certain limit within a certain time frame. We refer to this as **permanency** (also **permanence**). While the height of a person may be permanent for many years, it varies when considering the whole life span.

For practical applications, features should be efficiently measurable. This is referred to as **measurability** or **collectability**. Measurement errors should be low but always need to be expected.

Practical biometric applications also require **universality** which means that each individual should exhibit the biometric feature. For example, a person without hands does not have fingerprints.

A non-technical requirement for a practical biometric application may be **acceptability**. Depending on the biometric feature, people may not accept the system that uses it, for example due to privacy concerns.

For biometric authentication, the sample acquired by the measurement of the feature is compared to a previously recorded template of the claimed identity (1:1). The decision whether a sample is accepted or rejected is performed by a binary classifier and may simply use a threshold on the biometric feature vector. The components of a biometric authentication scheme are shown in Fig. 2.13. Due to various factors, a legitimate sample may be too distant from the template which leads to a false rejection. Also, a sample taken from a different individual may be within the threshold so that a login attempt would lead to a false acceptance. False rejection rate (FRR) and false acceptance rate (FAR) are important

characteristics to evaluate a biometric authentication scheme. Both depend on the threshold used, which can be varied. The intersection of FRR and FAR over the threshold is referred to as equal error rate (EER).

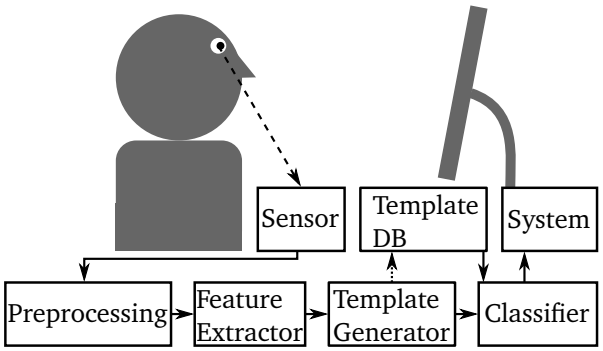


Figure 2.13.: Components of a biometric authentication system. The user wants to login to a system. A sensor captures a biometric measurement. After preprocessing, features are extracted to create a template. This is compared by a classifier or matcher to a template of the user that was stored in the template database in a previous enrollment phase. The user is authenticated to the system if the fresh sample matches the stored template within a threshold.

Due to the measurement errors, remote authentication cannot be realised by simply using a biometric sample as secret in a challenge response protocol as applying the cryptographic hash functions prevents to determine the distance of the sample to template, unless it exactly matches. One solution is correcting the measurement by public helper data algorithms, which have the disadvantage of entropy leakage [22].

In the context of biometrics, the term identification is not used for stating an identity. Instead it describes the process of comparing a biometric measurement to multiple stored templates (1:n) to retrieve the closest match or all matches within a certain threshold. A classical use case of biometric identification is crime investigation.

Presentation Attacks An adversary can be authenticated by a biometric system by presenting features to the sensor that are similar enough to legitimate ones. For example, a rubber finger can be created with a copy of a legitimate user's fingerprint. To counter such attacks, biometric systems may include **liveness detection** mechanisms, which also may be circumvented by simulating the liveness features accordingly. In principle, every biometric system can be defeated by creating a sufficiently accurate facsimile. In practice, the security of a biometric system depends on how difficult it is to forge a facsimile. This difficulty also depends on how easy it is to obtain the biometric sample from the victim. For example, obtaining a someone's fingerprint, which is often left on surfaces touched, is easier than obtaining a palm veins pattern.

Once an adversary has obtained biometric samples and is able to create a facsimile that is accepted by the system, this biometric feature cannot be used securely any more. Unlike password authentication schemes biometrics do not offer **changeability**. To mitigate leaks from biometric template databases, templates can be transformed to make it hard to recover the original template. This is either done by a non-invertible transformation or by biometric salting, where the user needs to supply the transformation parameters together with the biometric sample, which can be considered a secret component. In both cases, for each distinct database, individual transformation parameters are used which can be

reset in case of a leak. These approaches are referred to as **cancelable biometrics** [90].

2.3.3. User Non-Compliance

In order to be secure, authentication mechanisms typically require the user to comply to policies, such as, not writing down a secret, not keeping a token in an easily accessible place, or simply not authenticating to give access to a third person. There is always a risk that users do not comply. Reasons may be negligence or lack of understanding but also attacks. In a **coercion attack**, a user is forced to either login to system for the attacker or to reveal a secret or token to allow the attacker to login or use the secret for cryptographic operations. Defeating cryptographic protection by coercion is sometimes referred to as rubber-hose cryptography or a rubber-hose attack.

The goal of getting the user to reveal a secret or logging in can also be reached by **social engineering**. In a social engineering attack, the user is deceived by the adversary and the user is not aware of the attacker's true intention [111]. As coercion and social engineering attacks aim at weaknesses of the user it is hard to achieve resistance by technical means.

2.3.4. Continuous Authentication

In common user authentication schemes, the user is only authenticated once at the beginning of a session, subsequently allowing the session any operation in accordance with the user's privileges. This poses the risk of session hijacking. Continuous authentication schemes try to continuously verify the authenticity of the user, which can be done by monitoring biometric features.

3. EEG-Based Authentication

Electroencephalography can be used for an authentication scheme to obtain relevant information from a user's brain. This information may be knowledge that a user reveals in a BCI scheme, discriminant information in a biometric scheme, or a combination of both. This chapter discusses previous work in the fields of EEG-based password spellers and biometric authentication systems. Also, relevant contributions on EEG-based biometric identification are discussed, as these systems can in principle also be used for biometric authentication (see Sect. 2.3.2). We focus on research relevant to the approaches presented in this thesis, i.e., ERP and SSVEP. Also, as permanency is an important property for biometrics, we focus on contributions that included multiple subsequent sessions. We structure our overview by the paradigms used and discuss common performance and security properties.

Several review articles on EEG-based biometrics have been published in recent years [89, 120, 37, 33, 11, 44]. As the proposed schemes are very different and have been evaluated in different ways, a comparison is challenging. To ease this, the review of Yang et al. [120] introduced a usability score U that is computed from the number of subjects included, N, the number of channels used, K, the time used for recording the training data set, Tr, and for recording the test set, Te, and the correct recognition rate *CRR*.

$$U = \frac{N \times CRR}{Tr + K \times Te} \tag{3.1}$$

This usability score may still not reflect the value of the different approaches. In some ERP studies for example, a relatively high interstimulus interval (ISI) is used, probably to eliminate effects of overlapping trials, while the approach may also work with rapid serial visual presentation (RSVP). Although the score incorporates the number of channels used, which is typically limited when using consumer grade EEG systems, it does not incorporate the signal quality of the system used which can be expected to be lower for consumer grade devices. Note that the CRR and hence the usability score of a certain biometric system may be different for an identification or an authentication scheme due to the different classification tasks. Therefore the performance of identification and authentication studies cannot easily be compared.

3.1. Paradigms for EEG-based Biometric Systems

3.1.1. Event-Related Paradigms

As ERPs are time-locked to an event and it is difficult to determine the exact moment of an internal event such as a certain thought, typically, external stimulus events are used in ERP biometrics. Although in principle stimuli could be applied to all senses, most studies use visual or rarely auditory stimuli, which can be easily presented by a PC or smartphone. To obtain a high number of trials for averaging in a short time, visual stimuli can be delivered quickly via RSVP. Any image used as visual stimulus entails a visually evoked potential (VEP). Early components of this type of ERP are associated with sensory processing. More complex images or more complex associated tasks may elicit later ERP components which are associated with higher levels of processing. An important example is the cognitive P300

component (see Sect. 2.1.3). The positions and amplitudes of the ERP components, sometimes referred to as landmarks, may vary with the user due to different knowledge, memories, experiences, or attitudes and hence contain discriminant information.

Visually Evoked Potentials

Chen et al. [16] used a picture set of single easily recognizable objects [93] that resemble the set of Snodgrass and Vanderwart [102] in an oddball protocol. After selecting three pictures as a password, the subjects had to count the occurrences of password pictures in a randomized stream of pictures to elicit P300s for the password pictures. This is in principle a P300 BCI for password spelling. But as individual classifiers per user were trained, the system becomes partially biometric. Thus, impostors knowing the victim's password, only got accepted in about 15% of the cases.

Das et al. [18, 19] employed an oddball paradigm on two different types of stimuli: geometric shapes and characters. In each type, one stimulus acted as target stimulus. Average ERPs of targets and non-targets where analysed separately. Users could be classified even when only regarding the first $300\,\mathrm{ms}$ which barely includes the P300. User recognition rate was generally higher on the non-target data, supposedly due to the higher number of trials added to the average. This shows that early VEPs contain discriminant information.

In several studies Palaniappan et al. [75, 76, 79, 80] used line drawings of single easily recognizable objects from the original Snodgrass and Vanderwart set [102] to elicit VEPs. Different classifiers and feature extraction methods were tested. Usually the features were based on the power spectrum. In an identification experiment with 102 participant an accuracy of 98% could be reached, but there were no permanence tests [80].

Armstrong et al. [5] used acronyms as stimuli in a study on 45 subjects. It was assumed that people know different sets of acronyms and that the processing of acronyms in semantic networks results in an N400 component. Out of four different classifiers tested, cross correlation was found to work best. Permanence was shown by repeating experiments after about two weeks and about half a year. Based on this work, Ruiz-Blondet et al. [94] evaluated polarizing foods and celebrities and compared them to other visual stimuli as well as different tasks: oddball, pass-thought, and resting state. The topography of discriminant information was found to depend on the type of stimulus. While it is strongest over the occipital cortex for foods, it is more central for celebrities. The permanence of the approach for up to 74 days was analysed in a later work [95].

Studies by Yeom et al. [123] and Wu et al. [119, 118, 125] used pictures of the user's face and other faces as stimuli. It is shown that the ERP of self-faces are different from non-self faces and also slightly different from the ERP of impostors who know the user's face [125]. In a small study with five subjects, Harshit et al. [39] also used self-faces and non-self-faces. The latter included both familiar and unfamiliar faces, which were analysed separately. In addition, auditory stimuli were used, self-voice, familiar voice, and unfamiliar voice. It was shown that the response to a familiar voice has a much higher amplitude than to an unfamiliar voice.

3.1.2. SSVEP

Although SSVEP has been one of the most important fields of BCI research, the paradigm was only applied for biometrics in recent years. Studies are rare, often with few subjects and not assessing permanency. The frequencies tested are typically low in the range of $5\,\mathrm{Hz}$ to $24\,\mathrm{Hz}$. There are protocols using a single stimulus frequency [30, 1], a sequence of different frequencies [82, 107], and simultane-

Table 3.1.: Authentication schemes based on ERP

Ref.	N	K	Tr(s)	Te(s)	Stimulus	Features	Class.	Sess.	CRR	U
[123]	10	18	6856	776	face vs. self-face	averaged signals	SVM	2 ($\geq$2 d)	0.86	0.0004
[16]	29	28	180	7.2	SVLO	averaged signals	LDA	1	0.99	7.5236
Ch. 4	16	4	1800	600	SVLO, photos	averaged signals	LDA	3 (4–9 m)	0.88	0.0034
[52]	78	3	120	120	faces, text, animals	averaged signals, AR, PSD	SVM	3 (5 m)	0.97	0.1576
[125]	15	16	600	600	face vs. self-face	spatial LDA, temporal LR	GA-HDCA	2 (2 w)	0.94	0.0013
[18]	50	6	504	504	characters	averaged signals	cosine dist.	3 (5 w)	0.88	0.0125

Table 3.2.: Identification schemes based on ERP

Ref.	N	K	Tr(s)	Te(s)	Stimulus	Features	Class.	Sess.	CRR	U
[80]	102	61	50	50	SVO	PSD	ENN	1	0.98	0.0323
[5]	9	6	188	188	acronyms	N400	XCOR		0.83	0.0057
[94]	50	3	1194	81	words, faces, foods		XCOR	1	1.00	3.4795
[95]	20	4	597	1194	words, faces, foods		XCOR	7–74 w	1.00	0.3722
[119]	8	16	1080	6	face vs. self-face		CNN	2 (7 d)	0.91	0.0062
[19]	40	17	336	336	geometric shapes	bandpass 0.5–8 Hz	CNN	2 (7 d)	0.99	0.0066

ously presented frequencies [115, 124, 26]. Although it may seem obvious to use features based on the SSVEP response at stimulus frequency and harmonics, only few studies extract these features [26, 30]. Others use autoregressive (AR) [82, 1] or convolutional neural network (CNN) on the raw EEG data [124].

Two studies investigated the applicability of SSVEP without proposing a biometric scheme. Tokovarov et al. [107] found individual differences in the SNR at the stimulus frequencies but did not assess the permanency. Wei et al. [115] suggests that more discriminant information can be found in transient components of the SSVEP response at low frequencies. Therefore, in a follow-up work [124] only the low-pass filtered data are used to train a CNN for an authentication scheme. An overview of proposed authentication schemes is shown in Tab. 3.3.

Table 3.3.: Authentication schemes based on SSVEP

Ref.	N	K	Tr(s)	Te(s)	Stimulus	Features	Sess.	CRR	U
[26]	8	32	480	480	6 Hz to 15 Hz (6 simul.)	variance at stimulus frequency and harmonics	3 (3 w)	0.96	0.0005
[30]	6	4	190	10	5 Hz	5 harmonics in power spectrum	4 (3 m)	0.89	0.0232
[124]	8	9	240	10	8 Hz to 15.8 Hz (40 simul.)	EEG samples (low-pass 8Hz stopband edge)	2 ($\geq$2 d)	0.97	0.0235
[1]	16	4	91	9	5 Hz	Multivariate AR	1	0.90	0.1134
Ch. 5	4	2	910	910	15 Hz to 41 Hz	5 harmonics in power spectrum	5 (53 d)	1.00	0.0015
Ch. 5	4	2	228	228	15 Hz to 41 Hz	5 harmonics in power spectrum	5 (53 d)	0.90	0.0053

In an identification study with 15 subjects, Piciucco et al. [82] presented trials with SSVEP stimuli at 6 Hz, 12 Hz, 18 Hz and 24 Hz for one minute each. The first session was used for enrollment (training) and the second session, which took place about 15 days later, as test. Features extracted by Mel Frequency Cepstral Coefficients (MFCCs) and AR reflection coefficients were compared. The classification was based on the minimum Manhattan distance. A maximum correct recognition rate (CRR) of 100% could be achieved when using all 19 recorded channels. The set of channels Fz, Cz, Pz, O1, O2 achieved up to 96%. This indicates that the SSVEP topographies contain discriminant information. This is also substantiated by Akiyama et al. [1] who found an increased accuracy for four instead of two channels used.

3.1.3. Non-Event-Related Paradigms

Non-event-related paradigms work without a time-locking event like stimulus presentation that allows averaging EEG data from multiple precisely defined trials. Instead, the user is performing a longer mental task, while EEG data are acquired. Various tasks have been tested, e.g., **motor imagery** [59, 54], **speech imagery** [14, 54], **mathematical calculation** [54], **resting state** [88, 87, 81, 77, 92, 78, 48, 53, 58, 114, 56, 55]. Resting state can be used with eyes open or eyes closed. Maiorana et al. found resting state to perform better with eyes closed [56, 55]. A few notable works on authentication schemes using resting state with eyes closed are shown in Tab. 3.4.

Maiorana et al. recently published a study on a **task-independent** EEG authentication [54]. In multiple-task enrollment, six different tasks were included in training CNNs. Multiple sessions were used to assess the permanency. With multiple session enrollment, an EER of 12% could be achieved in

Table 3.4.: Authentication schemes based on resting state with eyes closed

Ref.	N	K	Tr(s)	Te(s)	Features	Class.	Sess.	CRR	U
[49]	9	3	45	15	AR	linear (MSE)	2 (1–3 w)	1.00	0.1000
[56]	50	7	240	45	AR	fusion of L1, L2, cosine distance	3 (34 d)	0.83	0.0748
[71]	20	3	300	300	PSD	linear regression	2 (15 m)	0.88	0.0147

the long term for cross-task authentication.

3.2. Discussion

3.2.1. Performance and Permanence

A basic requirement for biometrics is that templates contain discriminant information and it is possible to acquire a matching sample even after a longer time, i.e., collectability and permanency are required. EEG signals are very low in amplitude which makes them prone to artifacts. The EEG itself is influenced by mental states or medication. As the brain adapts due to experiences, creating new memories, related features of the EEG may also change. Research grade EEG systems typically have a high number of channels and relatively high sampling rates, creating large data sets. Studies often include only fewer than 20 subjects. Finding a feature set that separates 20 data sets in the high dimensional EEG data is almost certain. Thus, studies with a small number of subjects and a single recording session are of little validity. Therefore, the performance overview tables (Tabs. 3.1 to 3.4) focus on studies with many subjects and multiple sessions. Studies with multiple sessions typically yield lower usability scores. All studies evaluated that analysed the development of accuracies over subsequent sessions, found a decrease. The usability score is often also decreased as in later sessions fewer subjects participate. Note that the entries in Tabs. 3.1 to 3.4 are typically computed from the best performing stimuli, channel sets, features, and classifiers and reflect the accuracy and number of subjects in the last session. To mitigate lower accuracies which may result both from short term and long term changes in the EEG, some studies use multi session enrollment, where data of multiple sessions are used to create templates or train classifiers. With this method, Maiorana et al. [54] could decrease the EER from 10% to 4%.

Channel Optimization As pointed out by Yang et al. [120], the number of EEG channels used affects the usability of a scheme and is therefore a factor of the usability score (see Eq. 3.1). Usability is mainly seen to be affected by more channels requiring a longer set up time. But also the cost of the acquisition device and processing of more data is a negative effect. In several works a high number of channels is used for recording but only a small optimized subset is later processed [125, 94, 17, 49, 56]. Yeom et al. [123] found different optimal channels for the different subjects. Although more channels may offer more discriminant information, higher dimensional feature vectors may complicate classification. Also, basing discriminant information on spatial features requires to be able to deal with slightly varying positions of the sensors in different sessions.

3.2.2. Security

Studies on EEG biometrics often claim that this kind of biometrics offers all important security related features, often without further investigation.

Shoulder Surfing Resistance As a BCI is by definition a means to enter data without any muscle movement (see Sect. 2.2), the input cannot be observed from the user's behavior unless the input is presented in form of a feedback, for example the letters typed on a screen. This in principle also applies to an EEG-based biometric system. Lin et al. [52] propose an authentication system with a head-mounted display (HMD) to deliver stimuli.

Presentation Attack Resistance EEG signals are claimed to be hard to covertly acquire and replay to a biometric sensor and also have an inherent **liveness detection** feature [56, 92, 52]. Recording an EEG template requires the user to wear an EEG headset and, if it is based on a certain task, also getting the user to perform this task. It may be possible, for example, to hide EEG sensors in a headphone, but this would normally be removed when mounting the real EEG headset for authentication. This makes acquisition of a template more complicated than in common biometrics such as fingerprints or face. Still, templates may be obtained by breaches of the template database.

Replaying a template could be realized by a wig that shields the impostor's EEG and is able to replay signals with the correct amplitude and spatial distribution. This can also be regarded as more complicated than creating a rubber finger, but a replay attack is a possible threat. To counter the replay attack threat, Gui et al. [38] suggested that exactly matching templates should generally not be accepted, as it is unlikely that exactly matching samples occur, and they proposed a method to detect samples where an attacker added noise to a template.

Changeability It is often argued that task dependent EEG biometrics schemes are inherently cancelable, as the task or stimuli can be changed when a template is leaked [7, 38, 94]. Chen et al. [16] showed for their proposed P300 scheme that uses a knowledge factor that the password image set can be replaced and that the presentation the old password does not elicit a strong P300 component any more. Lin et al. [52] showed cancelability in their proposed ERP-based scheme and presented a method to find replacement stimuli eliciting ERPs with a maximum distance to the old ones.

Coercion Resistance Several works argue that EEG-based biometrics may be coercion resistant, as a coercion attack induces stress on the user and stress alters the EEG [123, 118, 59]. To our knowledge there are no works that systematically investigate this claim. Although measuring stress by EEG has been shown [51], the works on EEG-based authentication systems typically do not assess the impact of stress on the features used. The claim that coercion induces stress may also not hold under all circumstances. Su et al. [106] propose to include a covert warning feature to the EEG-based authentication system where the user can covertly trigger an alarm by clenching teeth three times. The activation of the muscles induces strong artifacts into the EEG which can easily detected. In their EEG-based authentication system that includes a knowledge factor, Chen et al. [16] show that the user can deliberately enter a different secret to prevent a forced login.

Continuous Authentication Wang et al. [113] proposed a continuous authentication system based on feature fusion of face recognition and ERP biometrics based on an oddball paradigm. The ERP is only tested every two minutes and requires the user's attention. Nakanishi et al. proposed a continuous authentication scheme monitoring the EEG during a driving simulator task [70] and a scheme using subliminal visual stimuli [69]. The latter showed an effect of the visibility on the ERP, but for authentication band ratios of the power spectral density (PSD) were used as feature.

4. A P300 BCI for visual authentication

The P300 paradigm is often used in BCIs, as it allows to infer the user's will from the EEG (see Sect. 2.2.1). Thus, a BCI allows the user to enter information into the system without any movement. Hence it can be used to enter passwords without being prone to shoulder surfing. We proposed such an authentication scheme using images instead of characters as a password. In the proposed scheme, the size of the image set limits the password entropy, which determines the security against brute force attacks. To be able to deliver a sufficient number of stimuli, a short ISI of $200\,\mathrm{ms}$ is used.

4.1. Methods

We propose an EEG-based authentication system that uses RSVP to quickly present a set of images in random order. A small subset of these images is regarded as the password. These password images act as targets. They are rare and also relevant to the user, as the user is instructed to silently count their occurrence. These are the basic conditions to create a P300 response (see Sect. 2.1.3). P300 components in the EEG are identified by the system which allows to conclude which images the user had in mind as password set. If it matches the stored password set the user is authenticated. When P300 classifiers are trained on data from a sufficient number of subjects, the authentication is solely based on the knowledge of the password, not using any inherence factor.

4.1.1. Basic Study Design Considerations

To achieve a reasonable password entropy, we decided to use sets of one hundred images out of which five are used as password set. Although the work of Standing et al. [104] suggests that many more targets would be possible for such a task, as a precaution, we chose a number in the range of the 7 ± 2 items Miller et al. [66] found an average person can hold in working memory.

The 95% non-target images are more than required for the rarity requirement of the P300, but a higher number decreases the chance of adjacent images being targets and increases the entropy of the password.

Using 100 different images for an ERP study which requires averaging over multiple presentations of each image would become a very lengthy experiment if a proper distance between the stimuli was maintained. Therefore, we accepted the disadvantages of overlapping ERPs and used RSVP with five images per second, assuming that the increased rate overcompensates the quality loss due to ERP overlap and increases the ITR.

As the properties of the P300 vary with many factors (see Sect. 2.1.3), experiments were conducted with two different types of stimuli:

photos showing animals, landscapes, people, plants, fruits etc.

simple objects grayscale drawings of items that can be described with a single word such as bed, hand, elefant, sandwich

The photos were randomly selected from the public domain imagery Pixabay [85]. The simple objects were grayscale versions of a set by Rossion and Pourtois [93], resembling the set of Snodgrass and Vanderwart [102]. Hence we call them Snodgrass and Vanderwart like objects (SVLOs). The full sets of photo and SVLO stimuli are shown in Appendices A.1 and A.2 respectively.

4.1.2. Experimental Setup

Data were recorded at a sampling rate of $128\,\mathrm{Hz}$ using the EMOTIV EPOC+ headset and the corresponding software 'Testbench' (see Sect. 2.1.5).

The experiment was controlled by a Matlab Script using PsychToolBox [12, 46] to ensure precise timing of stimulus presentation. To sync the stimuli to the EEG, markers were sent to the recording software over a virtual serial port. The markers used values allowing to identify the different stimuli. By connecting a phototransistor circuit to the marker channel we found a constant delay of $180\,\mathrm{ms}$ between flipping the screen and the appearance of the marker in the data. In data analysis, the timestamps were corrected for this delay.

Subjects were comfortably seated approximately $1\,\mathrm{m}$ in front of a 24" screen with a refresh rate of $60\,\mathrm{Hz}$ and a resolution of 1920x1200 pixels. The size of the stimuli was limited to 400x300 pixels, corresponding to a vision angle of $6.2°$, keeping the stimuli approximately in the foveal area of the eye allowing the subject to easily view the whole stimulus without eye movement. In addition, to reduce eye movement, a black cross marked a fixation point. Stimuli were presented on a gray background to keep the contrast to the stimuli low.

4.1.3. Participants

In total 26 subjects (9 female, 17 male), mainly students or staff of Hamburg University of Technology, participated in the study. Their average age was 28 with standard deviation of 5.59, ranging from 20 to 34. 22 of the 26 subjects were naïve to the paradigm. 18 subjects attended a second session, 16 also a third. The average delay between the first two sessions was 68 days with standard deviation of 33 days. The third session was on average conducted 168 days after the first with standard deviation of 58 days.

4.1.4. Structure of Experiments

Experiment sessions consisted of two runs, one for each stimulus type, SVLOs and photos, in randomized order to avoid introducing a bias by fatigue or training. One run consisted of 50 bursts. In each burst, each of the 100 images was presented for $200\,\mathrm{ms}$, with no ISI. The bursts were separated by displaying a fixation cross for $4\,\mathrm{s}$. The sequence is depicted in Fig. 4.2. The appearance of one image on the screen constituted a single trial. Thus there were 50 trials for each image in total.

To examine possible benefits of self-selected passwords, the subjects were divided into four groups. One group was allowed to select their own passwords for both stimuli types. Two groups were either allowed to choose their own password for the photo set or the SVLO set. The fourth group had to use the default passwords (see Fig. 4.1) for both stimuli types.

After the adjusting the EEG headset, the subjects were instructed to produce muscle, movement, and eye artifacts and to observe the effect in the EEG timeline to explain why producing artifacts during the experiment run should be limited. The main task was set as silently counting the occurrences of target images, restarting with each burst. For motivation the subjects were briefed on the importance

Figure 4.1.: The left column shows the default set of password images (targets) of type photo (from top to bottom p1–p5), the right column shows the set of default targets of type SVLO (from top to bottom s1–s5) [34].

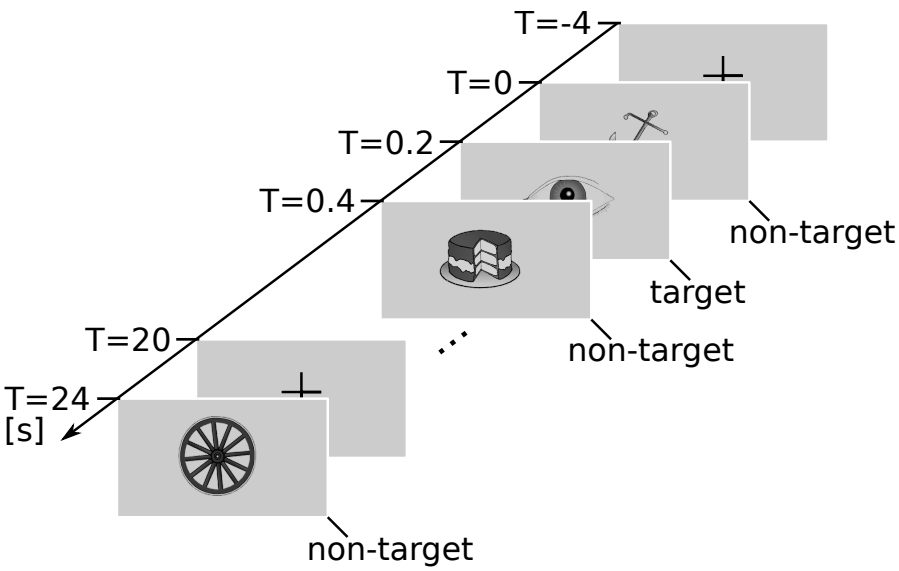


Figure 4.2.: The sequence of the stimuli using RSVP of target and non-target images.

of attention and they were told that they would be informed about their own relative performance after the study.

Prior to the actual experiment, each target image was shown individually for a period at the discretion of the subject. Test bursts followed, requiring the subject to enter the number of targets counted afterwards. If the correct number was entered, the subject was regarded as being able to identify all targets, allowing to start the actual experiment consisting of 50 bursts.

After each experiment run, the experimenter used a questionnaire to assess the attention and performance and possible other issues during the experiment run. The subjects were asked to rate their attentional control (scaled one to ten) and for each target, how often they were able to recognize it on average. To be able to detect issues like malfunctioning stimulus presentation, distraction or difficulties complying to the experiment protocol, the subjects were asked two general questions:

- Did you notice anything unusual?
- Has anything disturbed you?

From the second session on, participants were also asked how they performed compared to previous sessions and how well they could remember their password.

4.1.5. Data Analysis

Data were analysed using MATLAB [61] and the FieldTrip toolbox [74]. With the help of FieldTrip, raw EEG data were cut into trials of $1.5\,\mathrm{s}$, beginning $0.5\,\mathrm{s}$ before each stimulus onset to use these $0.5\,\mathrm{s}$ of data for baseline correction. Trials classified as containing eye artifacts by the z-value-based automatic artifact rejection of FieldTrip (channels AF3 and AF4, threshold 10) were excluded from further processing. Trials adjacent to a target stimulus or the end of a burst were excluded as well.

The preprocessing stage included linear trend removal, a bandpass of $0.1\,\mathrm{Hz}$ to $9.5\,\mathrm{Hz}$, and a bandstop $4.5\,\mathrm{Hz}$ to $5.5\,\mathrm{Hz}$. These filters remove steady state responses at the stimulus frequency of $5\,\mathrm{Hz}$ and the harmonics from $10\,\mathrm{Hz}$. Up to 50 trials were averaged for each stimulus image yielding 100 averages. To be able to evaluate the influence of the number of bursts on the classification, averages over the first n bursts only were computed.

4.1.6. Classification

To identify trial averages containing a P300, only the interval $0.22\,\mathrm{s}$ to $0.5\,\mathrm{s}$ was considered. Data were downsampled to $32\,\mathrm{Hz}$ and averaged over the channels F7, F8, FC5, and FC6, yielding nine amplitude samples which were directly used as feature vector.

As there is a skew between the distribution of targets and non-targets, having 95% of non-targets, a trivial classifier which always predicts a non-target would reach an accuracy of 95%. Hence, to assess the classifiers we use the F_1 score.

The F_1 score is the harmonic mean of precision and sensitivity. As both precision and sensitivity are 0 when there are no true positives, the F_1 score is not defined in this case because of the division by zero, we use the continuous extension $F_1(0,0) = 0$ (see equation 4.2).

$$F_1 = 2 * \frac{precision * sensitivity}{precision + sensitivity} \tag{4.1}$$

$$F_1(precision, sensitivity) = \begin{cases} 2 \cdot \frac{precision \cdot sensitivity}{precision + sensitivity} & , \ sensitivity > 0 \\ 0 & , \ sensitivity = 0 \end{cases} \tag{4.2}$$

The 100 feature vectors of each experiment were classified by linear discriminant analysis (LDA), using three different types of training data:

1. To estimate the class separability within one experiment run, leave-one-out **cross-validation (CV)** was used. We refer to it as CV-score.

2. For each experiment run a **general classifier (GC)** was trained with all other experiment runs with a CV-score of 0.7 or above, including experiment runs of other subjects. For the evaluation only including the first 25 bursts, all experiment runs with a CV-score of 0.5 or above were used.

3. For each experiment run an **individual classifier (IC)** was trained with all other experiment runs of the respective user with a CV-score of 0.7 or above. If less than three runs reached the threshold, the experiment runs with the three highest CV-score were used.

4.1.7. Statistical Tests

As the scores of our classifiers do not follow a normal distribution, non-parametric statistical tests were used.

For pairwise comparisons, like SVLO vs. photo, the scores of all experiment runs were divided into respective groups to conduct a Wilcoxon signed-rank test. The comparison between given and self-selected password images was performed with a Wilcoxon rank-sum test. The three groups of scores corresponding to the three experiment sessions and the three types of classifiers were compared by a Friedman test.

As several subjects reported that the task was easier with either SVLO or photo, three groups of experiment runs were created: the ones with the stimulus type perceived easier, harder, and where subjects did not report a difference. On these groups a Kruskal-Wallis test was conducted. A twotailed Spearman correlation was used to test for a correlation between the scores and the reports of the subjects on how well they could focus on the task in each experiment run.

4.1.8. Authentication

To estimate the error rates of the proposed authentication scheme, login attempts were simulated from the experiment data. All experiment runs were used as a challenge for each user. This was done by classifying the 100 images of the experiment run into targets and non-targets. These predictions were evaluated against all class label sets (password image sets) used in the study by using the F_1 score. A user was granted access if the F_1 score of the experiment run is above a threshold. If the actual password image set did not match, this was counted as false acceptance. If the actual password image set matched, but the F_1 score was below the threshold, this was counted as false rejection. The error rates were computed for a varying F_1 score threshold. The authentication error rates were analysed separately for the two stimulus types.

4.2. Results

Unless otherwise specified, the results presented originate from the 16 subjects who participated in all three sessions. The scores of different groups of experiment runs are generally shown in boxplots. The edges of the box show the 25th and 75th percentiles. The 50th percentile (median) is indicated by the central red line. The whiskers extend to data points not considered outliers. Outliers are represented by the '+' symbol.

4.2.1. Number of Bursts

Fig. 4.3 shows the grand average F_1 scores of the cross-validation (CV) by the number of bursts regarded. The red graph is based on all experiment runs by all 26 subjects. These experiment runs are split up into the 17.5% runs having a score over 0.7 (blue) and the remaining experiment runs (orange). The F_1 scores increase with the number of bursts from below 0.1 to 0.34 (0.83 and 0.23 respectively). Around thirteen bursts, all curves have an inflection point from concave upward to downward around.

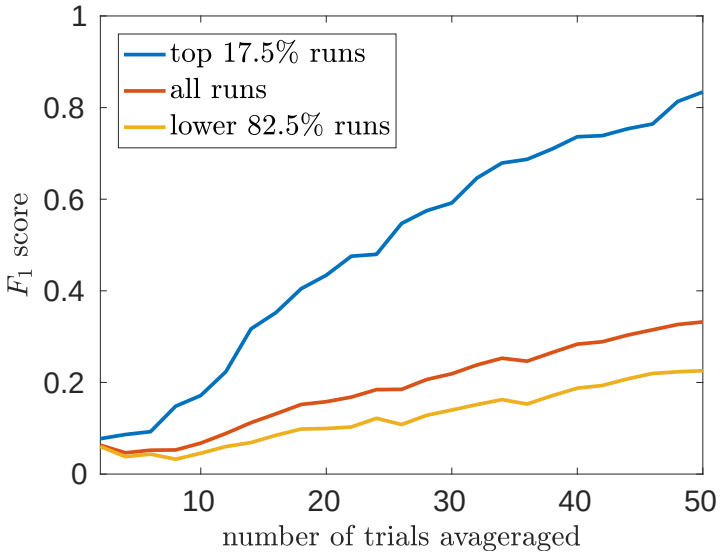


Figure 4.3.: Grand average cross-validation F_1 scores over the number of trials (bursts) [34].

4.2.2. Sessions

Mean and standard deviation of the CV F_1 scores by session are shown in Fig. 4.4. From session one to session three the scores increase significantly ($p = 0.0218$). The mean $\bar{x}$ increases from 0.33 to 0.48.

The respective attentional control (see Fig. 4.5) shows no significant difference ($p = 0.46$). After the second session twelve out of 18 subjects reported that performing the task was easier or required less concentration. Eight of 16 subjects reported this for the third session.

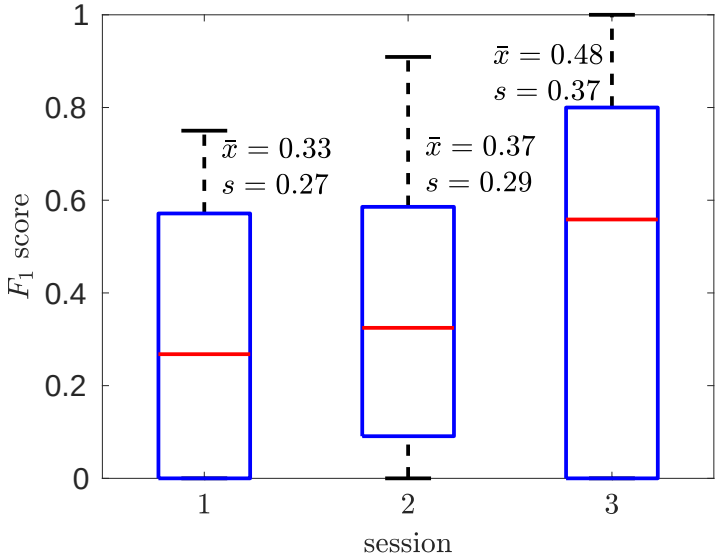


Figure 4.4.: Cross-validation F_1 scores of all subjects with 3 sessions for each session.

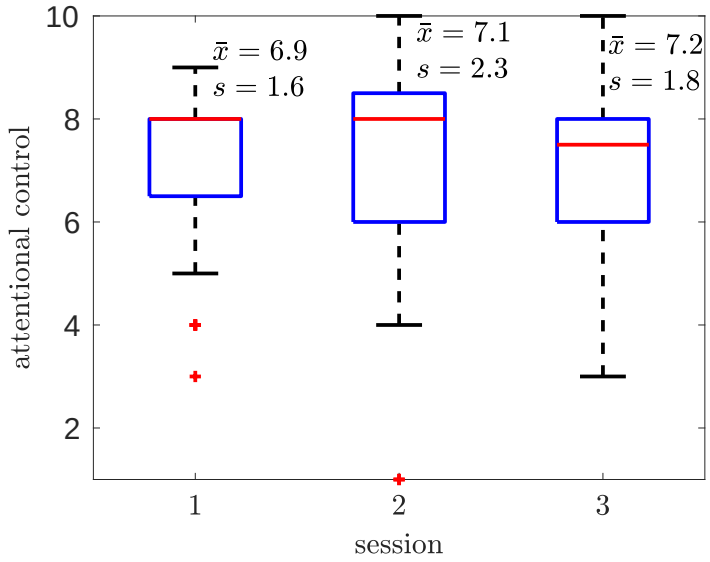


Figure 4.5.: Reported attentional control of all subjects with 3 sessions for each session.

4.2.3. SVLO vs. Photos

The bipartite selection of the experiments with stimulus type SVLO and photo fail to show significant differences in the scores ($p = 0.732$, see Fig. 4.6 and Tabs. 4.1 and 4.3). The reported attentional control is slightly higher for photo stimuli (see Fig. 4.7), but not showing statistical significance ($p = 0.16$). Several subjects reported that the task is harder with a certain stimulus type ($SVLO = 10$, $photos = 6$, $indifferent = 10$). There was no significant difference in the scores of the task perceived easier or more difficult ($p = 0.29$).

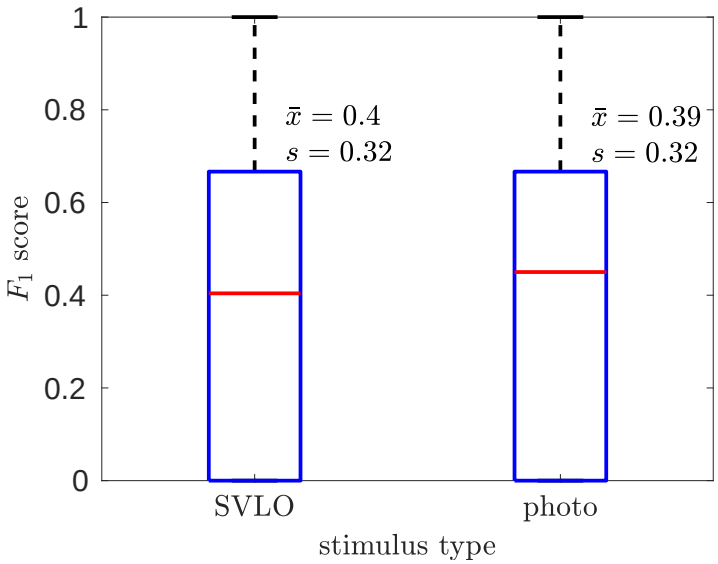


Figure 4.6.: The cross-validation F_1 scores of SVLO and photo stimuli.

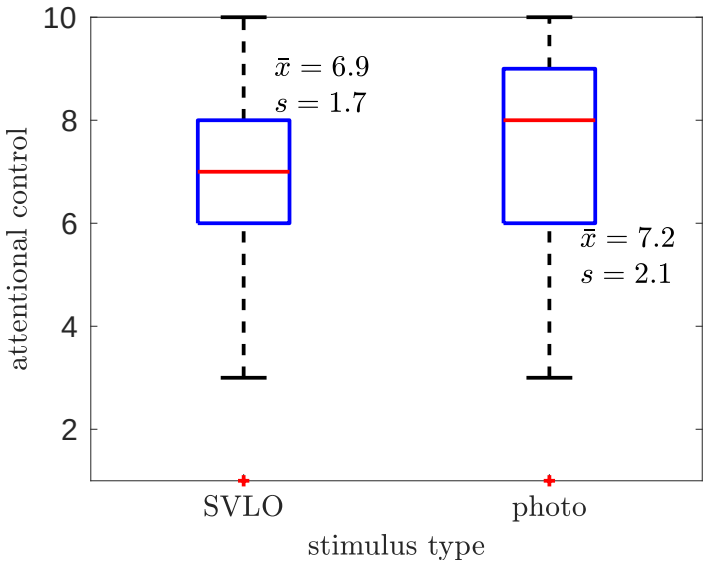


Figure 4.7.: Reported attentional control of experiment runs with SVLO and photo stimuli.

4.2.4. User Chosen vs. Default Password

Scores of self selected (own) passwords are not significantly ($p = 0.4339$) different from default (given) passwords (see Fig. 4.8 and Tabs. 4.1 and 4.3). The reported attentional control does not significantly differ ($p = 0.53$) between using self-selected and default passwords (see Fig. 4.9).

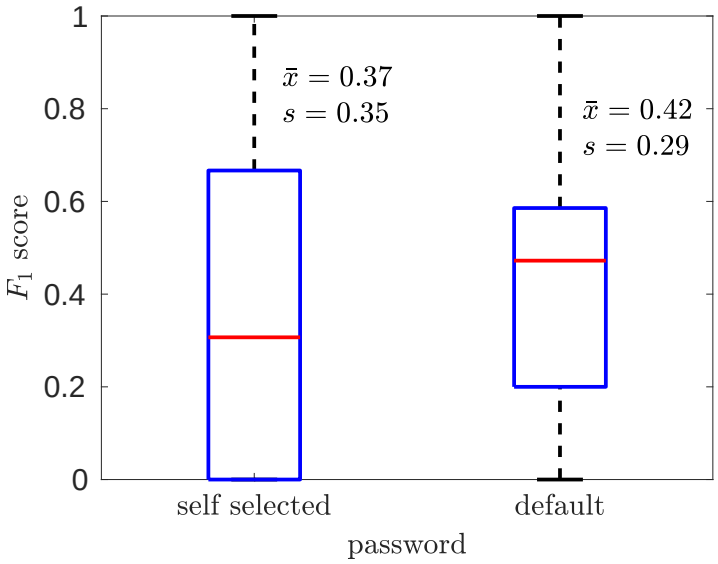


Figure 4.8.: The cross-validation F_1 scores of self-selected vs. default passwords.

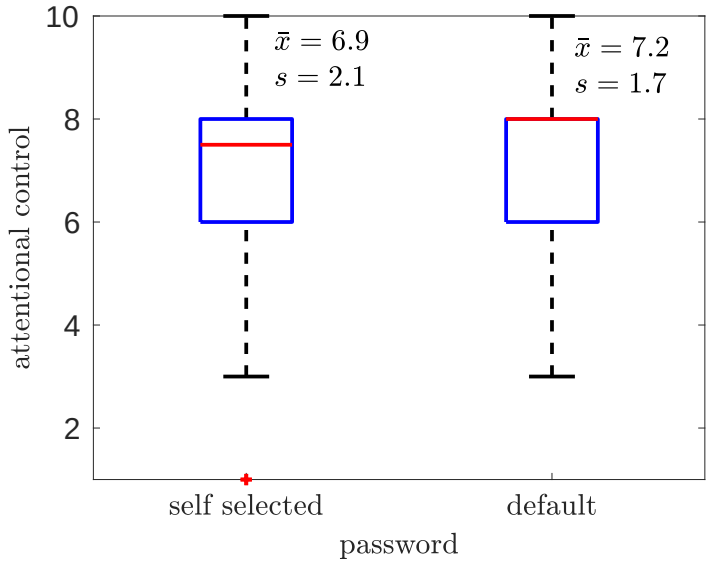


Figure 4.9.: Reported attentional control of self-selected vs. default passwords.

4.2.5. First vs. Second Part of a Session

The scores of experiment runs within the same session did not significantly differ ($p = 0.4262$). The respective statistics are shown in Fig. 4.10 and Tabs. 4.1 and 4.3. The subjects' reports on how well they could focus on the task (attentional control) are significantly ($p = 0.006$) higher for the first run (see Fig. 4.11).

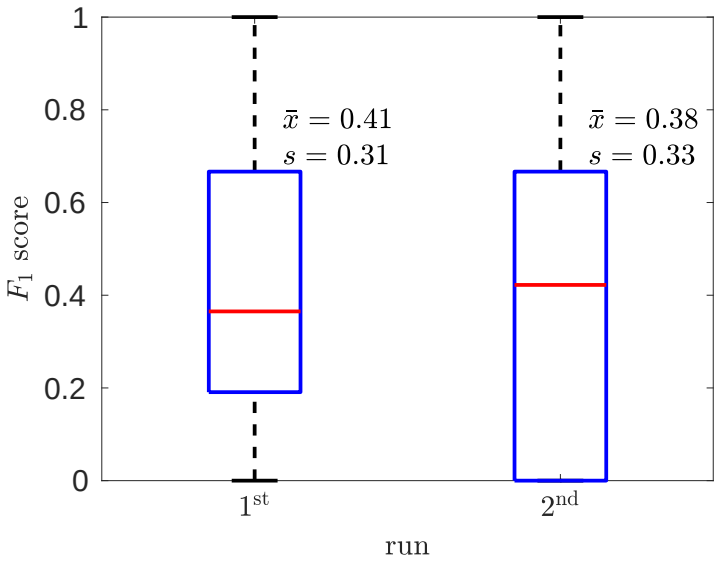


Figure 4.10.: The CV F_1 scores of the first and the second run during one session.

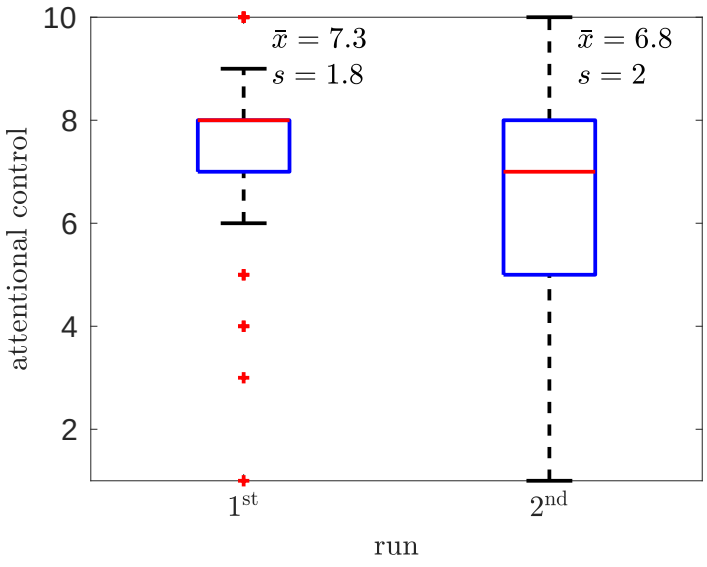


Figure 4.11.: The reported attentional control in the first and the second run during one session.

Table 4.1.: Mean F_1 scores of subjects with three sessions [34].

(a) using 50 bursts

Set	**CV**	**IC**	**GC**
All	0.39	0.37	0.28
SVLO	0.4	0.36	0.27
Photo	0.39	0.37	0.29
1st run	0.41	0.36	0.27
2nd run	0.38	0.38	0.29
Own	0.37	0.34	0.29
Given	0.42	0.39	0.28
Session 1	0.33	0.19	0.15
Session 2	0.37	0.46	0.37
Session 3	0.48	0.46	0.33

(b) using first 25 bursts

Set	**CV**	**IC**	**GC**
All	0.22	0.2	0.17
SVLO	0.21	0.21	0.14
Photo	0.24	0.19	0.2
1st run	0.21	0.22	0.17
2nd run	0.23	0.18	0.16
Own	0.25	0.19	0.17
Given	0.2	0.21	0.16
Session 1	0.14	0.12	0.06
Session 2	0.19	0.22	0.22
Session 3	0.34	0.26	0.22

Table 4.3.: p-values for cross-validation (CV), individual classifier (IC) and general classifier (GC) [34].

Set	**CV**	**IC**	**GC**
SVLO vs. Photo	0.7323	0.757	0.688
1st vs. 2nd run	0.4262	0.393	0.287
Own vs. Given	0.4339	0.276	0.76
Sessions	0.0218	0.06096	0.05217

4.2.6. Individual vs. General Classifier

The mean F_1 scores over all experiment runs when using CV, individual classifier (IC), and general classifier (GC) are 0.39,0.37, and 0.28 respectively, with the standard deviations 0.32, 0.31, and 0.3. There is a significant difference ($p = 0.00137$) between CV, IC, and GC (see Fig. 4.12). When only including the first 25 bursts, the F_1 scores of CV, IC, and GC are 0.22,0.2, and 0.17 respectively, with the standard deviations 0.27, 0.25, and 0.23. There is no significant difference ($p = 0.291$) between the scores of the different classifiers (see Fig. 4.12).

The FARs and FRRs of the simulated login attempts are depicted in Figs. 4.14 to 4.17. The graphs are similar for the tested parameters stimulus types, classifiers, and numbers of bursts used: The FAR falls quickly down to 0 at a threshold of about 0.3. The FRR reaches a plateau below 0.4 at a threshold of 0.5. Their intersections, the EERs are between 0.0825 and 0.142. When only including the first 25 bursts, the error rates are generally higher (see Tab. 4.4 and Figs. 4.16 and 4.17).

Table 4.4.: Equal error rates

(a) including 50 bursts

Set	IC	GC
SVLO	0.0825	0.109
Photo	0.0825	0.123

(b) including first 25 bursts

Set	IC	GC
SVLO	0.118	0.124
Photo	0.142	0.113

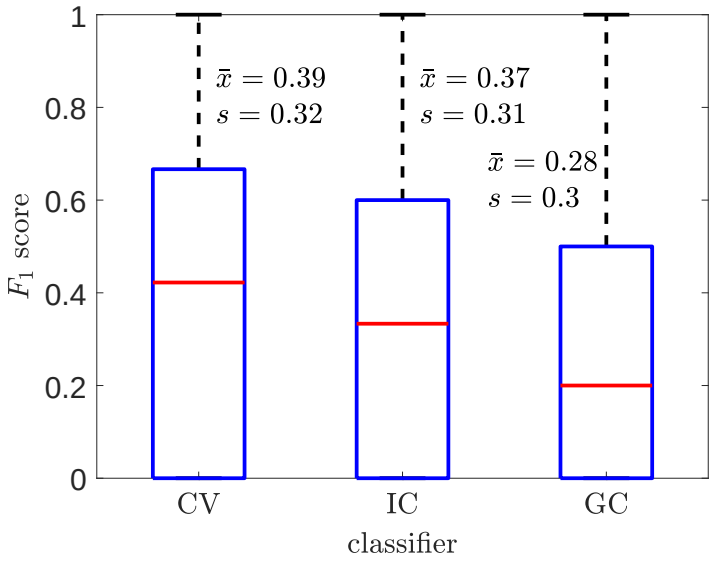


Figure 4.12.: F_1 scores of cross-validation (CV), individual classifier (IC), and general classifier (GC) [34].

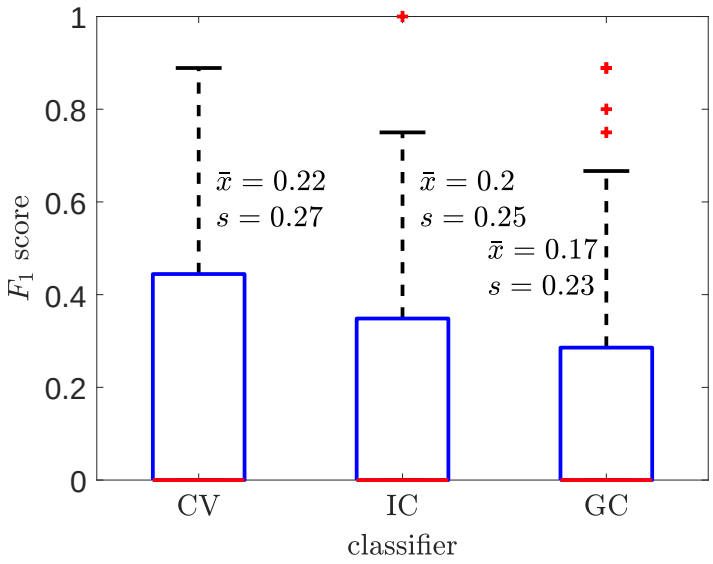


Figure 4.13.: F_1 scores of cross-validation (CV), individual classifier (IC), and general classifier (GC) when only including the first 25 bursts.

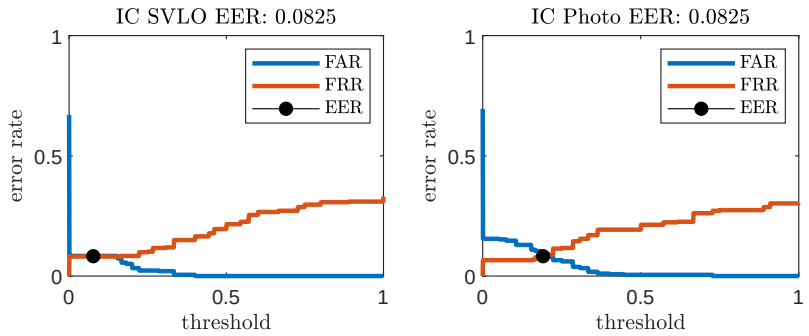


Figure 4.14.: Error rates of the authentication system using the individual classifier (IC) [34].

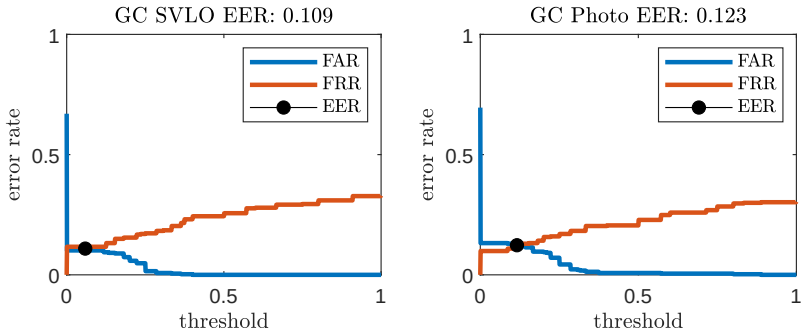


Figure 4.15.: Error rates of the authentication system using the general classifier (GC) [34].

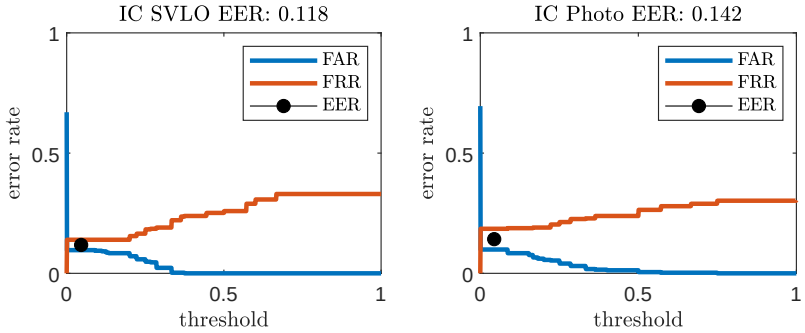


Figure 4.16.: Error rates of the authentication system using the individual classifier (IC) and only including the first 25 bursts [34].

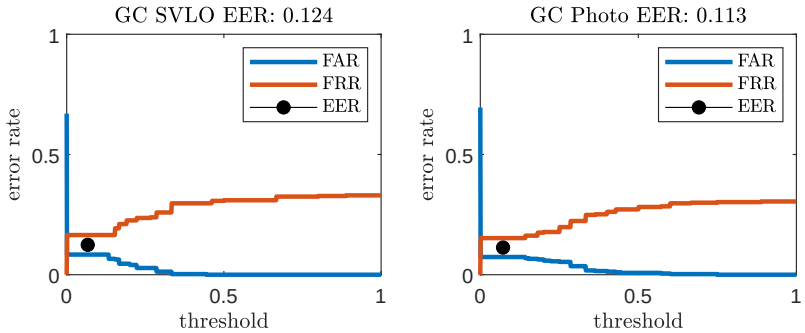


Figure 4.17.: Error rates of the authentication system using the general classifier (GC) and only including the first 25 bursts.

4.2.7. Survey

The correlation between the scores and the subjects' reports on how well they could focus on the task did not reach a significant level (twotailed Spearman $p = 0.4947$).

For all experiment runs with default password (see Fig. 4.1) we added up how often each target was reported as 'difficult', meaning as being recognized less often than approximately all appearances (see Tab. 4.4).

Table 4.6.: The number of targets reported as 'difficult' for both stimulus types in each of the three sessions. p1–p5 and s1–s5 were the default passwords. See Fig. 4.1. The last row shows the total number of participants using the default password per session.

target	1	2	3	target	1	2	3
p1	1	0	0	**s1**	4	1	0
p2	7	1	1	**s2**	2	0	0
p3	1	0	0	**s3**	0	0	0
p4	8	4	1	**s4**	0	1	0
p5	0	0	0	**s5**	0	0	0
total	14	8	8		13	9	8

On the questions on the disturbances the subjects reported for the 120 experiment runs eleven cases of perceiving pressure or pain inflicted by the headset and twelve cases of noise from outside the experimentation room.

4.3. Discussion

4.3.1. Images as Password

Images are widely believed being easy to remember. People are often able to recognize images they had seen a long time ago, Nickerson et al. [72] showed this for periods of up to one year.

Unlike in a normal password scheme, our system is presenting the secret components to the user. This may allow a user to recognize the password and to log in even if the password could not be recalled before. The study was not designed to test the effects of a not fully known password, instead the password set was presented at the beginning of each experiment run. However, after the second session twelve of 18 and after the third session five of sixteen subjects reported that they did not fully remember their password before the experiment, but quickly recalled it when the target images were shown.

It is likely that users will recognize images in an authentication sequence which have previously been used as part of the password set. Chen et al. [16] found in a similar study that there is no significant difference in the ERP of non-target images and images previously used as password. Still, as a precaution, it may be advisable to change the whole image set at password changes. In case of a password leak the password set should fully be changed to avoid the risk of a false acceptance just with the known images. If this policy is enforced, including leaked password images as non-targets reduces entropy.

4.3.2. Performance and Usability

The presented experiment uses 50 bursts of 100 images, presented for $200\,\mathrm{ms}$ each, followed by an ISI of $4\,\mathrm{s}$. This results in an experiment run duration of $20\,\mathrm{min}$. Fig. 4.3 shows an increasing classification

score over the number of trials averaged. The number of trials where the classification scores converge seems to be beyond the 50 trials we tested. As the increase of scores is steepest at the beginning, reducing the number of trials to decrease the experiment time might be an option, but in our experiment, using only the first 25 bursts approximately halves the scores (see Tabs. 4.1 and 4.4). The resulting experiment time of $10\,\mathrm{min}$ is still not suitable for a practical authentication scheme. The respective usability score (see Ch. 3 and Eq. 3.1) increases from 0.0018 to 0.0034 for IC and SVLO, which is still comparatively low (see Tab. 3.1). The experiment time could also be decreased by presenting each image for a shorter duration. The lower bound for the trial duration in RSVP is assumed to be at $50\,\mathrm{ms}$ [86]. However, decreasing the trial duration will increase the overlap of the ERPs complicating the classification. BCIs using RSVP can address this by ensuring that there is a certain minimum distance between the presentation of targets. This is not suitable for our authentication system as it would allow shoulder surfing: when observing a number of authentication runs, targets are identifiable by the property that they are never presented closer to each other than the minimum distance.

There is research on single trial P300 detection [21, 43, 97]. Manor et al. [57] used a CNN in a $100\,\mathrm{ms}$ RSVP paradigm. If this is applicable to the proposed authentication scheme, the experiment duration could be reduced to $10\,\mathrm{s}$.

The experiment duration could further be reduced by using a smaller image set. A selection of 5 out 50 would halve the duration and even 5 out of 30 would not exceed the maximum target rate [91]. Certainly, the reduction of the image sets reduces the entropy of the password, but together with single trial detection the experiment duration could be reduced to a few seconds which may be short enough for a real world application.

4.3.3. Advantages of Individual Classifiers

The goal of this study was to build a non-biometric EEG-based authentication scheme using a 'password'. Though, as this 'password' is entered via a physiological measurement, false acceptance and false rejection errors may occur, like in biometric systems. Instead of using a general classifier, which is trained on all users, an individual classifier trained per user makes the system partially biometric. Such a classifier can adjust to individual characteristics of the P300 to increase the accuracy and thus reduce the error rates for the user. Rejection rates of impostors, even when knowing the secret, could be increased, which can be considered a benefit unless password sharing is a desired feature. The scores of our individual classifiers exceed the scores of the general classifier (see Fig. 4.12) and are expectedly lower than of the cross-validation, as the latter should be an overestimate of the real classifier performance. But the ICs fail to show a statistically significant advantage over the GC. The similar scores of the three classifier types may be due to the simple feature set that may be too coarse to capture slight individual differences of the P300 characteristics. More sophisticated signal processing and classification may allow using individual differences in the ERPs to create a more accurate individual classifier. Such a classifier might also show differences with other variables we tested, like stimulus type and whether a password was self selected.

4.3.4. Scores over Sessions and Experiment Runs

Preliminary experiments showed low P300 peaks and hence low classification scores. Low scores were reproduced in the first session this study. Half of the 32 experiment runs of the first session did not exceed an F_1 score of 0.25. But the mean of 0.33 in the first session increased to a more acceptable value of 0.48 in the third session (see Tab. 4.1). The survey on attentional control cannot explain

this increase of scores, as the reported values are very similar in all three sessions (see Fig. 4.5). The decreasing number of reported missed ('difficult') target images (see Tab. 4.6) could be an indication for a training effect. Generally the subjects reported that the task became easier over the sessions. Some reported that the task was very difficult in the beginning. From the second session on, two subjects experienced the experiment as boring. Training effects had been observed by Baykara et al. [8] in an auditory P300 BCI study with five sessions.

Another factor for the scores increasing over the sessions could be the high number of stimuli. Unlike the target images, which are prominently shown at the beginning of the experiment run, non-target stimuli are only presented in the bursts and therefore unfamiliar to the subject. With experiencing more bursts and sessions, non-target stimuli may become more familiar. Recognizing a novel, unfamiliar stimulus can elicit the P3a variant of the P300 response (see Sect. 2.1.3). As our classifier is not designed to distinguish between P3a and P3b, P300 responses are added to the averages of non-target stimuli, reducing the classifiability. This effect is expected to decrease with the decreasing novelty of the stimuli, increasing classifiability in later bursts and sessions.

As the longest intervals between sessions were few months, the study cannot prove long term permanency. It can be expected that at a certain number of sessions the scores will not improve any more. Scores may also deteriorate if the attention decreases as the task becomes a routine. This may be overcome by changing the image set.

Though subjects generally reported a significantly lower level of concentration in the second run of each session, there is no significant difference in the scores when comparing the two consecutive runs (see Figs. 4.10 and 4.11).

4.3.5. SVLO vs. Photos

Low classification accuracies in preliminary experiments with photo stimuli led to the hypothesis that an unlimited variety of natural photos as stimuli might affect the resulting ERPs due to varying difficulty of recognition and possibly different recognition processes. The discrimination difficulty is known to affect P300 latency and amplitude [83, 110].

Ruiz-Blondet et al. [94] hypothesized that natural photos as stimuli cause different ERPs among subjects as they may have different meanings to them. Their study of a biometric identification scheme included natural photos believed to be polarizing, for example publicly well known people or certain food. The study reached a very high accuracy on 50 subjects (see Sect. 3.1.1).

Strongly varying ERPs would reduce the accuracy of our binary P300 classifier, but SLVOs, which we expected to have very similar recognition difficulty and low variation in the meaning to the subject, did not show a significant difference in the classification scores compared to photo stimuli (see Tab. 4.1 and Fig. 4.6).

Though there is no significant difference in the scores, there might be a difference in usability resulting from individual perception of difficulty. Photos were perceived easier by more subjects than SVLO, but there were more reports of 'difficult' photos (see Sects. 4.2.3 and 4.2.7). To increase usability it can be considered to allow the user not only to choose the password, but a whole image set and the stimulus type. Allowing the user to compile a complete image set is not advisable, as a user might select two obviously different categories of images for the password and non-password set, leading to a weak password.

4.3.6. Security

The security of the proposed authentication system depends on the entropy of the password as well as the accuracy of the system in distinguishing targets from non-targets. Though we did not investigate the influence of target and non-target set sizes, a study on recognition memory by Standing et al. [104] found subjects could retain over 2000 photographic items, suggesting that the number of targets may be increased as long as the target rate stays low enough.

Exhaustive Search In this study, five out of a set of 100 images were used as target stimuli which allows $\binom{100}{5} = 75\,287\,520$ different selections. This corresponds to maximum security of 26.17 bit. This is much higher than the 11.17 bit (three out of 20) of the scheme proposed by Chen et al. [16]. The higher entropy is required for our non-biometric case, where it is the single security parameter. However, the security of 26.17 bit is only sufficiently secure against exhaustive search when the system is deployed in a way that prohibits concurrent login attempts. For a single user brute forcing a password would be infeasible due to the time required for the data acquisition. Even if the system could be optimized to five seconds (see Sect. 4.3.2), guessing the correct five images would take almost six years on average. To impede an exhaustive search, the time required can be extended by a waiting time between login attempts or accounts can be locked after a certain number of failed attempts. The above assumes a fully random selection of passwords and 100% classification accuracy. The latter assumption is not realistic, as even with a perfect classifier, due to noise in the data acquisition, misclassification may happen which leads to false acceptance and false rejection errors as in biometric systems (see Figs. 4.14 and 4.15). When operating at the EER, false rejection and false acceptance will cancel each other out, not influencing the chance of a successful login. But if the goal of the adversary is to find the correct password instead of successful login, the task becomes harder, as with a FRR above zero, testing the correct password does not guarantee a successful login. On the other hand, with a FAR above zero, a successful login does not mean the tested password is correct. But due to the design of the system, it is likely that the tested password is close.

As discussed in Sect. 4.3.3, using the individual classifier increases security by decreasing the error rates at the cost of generality, making the system partially biometric. The authentication scheme could be extended by other biometric features, like other properties of the ERPs.

For any type of classifier tested and also in case of only including 25 bursts, the FAR drops to zero about the threshold of 0.5, while the FRR only slightly increases, making this a useful setting when no false acceptance is tolerated. The unusual slopes of the error rates partially result from the relatively high entropy.

Shoulder Surfing Resistance In the classical shoulder surfing attack scenario, an adversary observes a victim while logging into a system. A password can be learned from the movement of the victim's fingers while typing. This may involve multiple observations as some characters may be hard to identify due to typing speed or viewing angle. We assume a stronger adversary who uses cameras to record logins from multiple angles which allows to reconstruct all related activity up to an arbitrary level of detail and an unlimited number of replays. Thus, the adversary is able to record the stream of images on the screen. Unless improperly implemented, for both targets and non-targets, the order of the images is randomized and timings are fixed, hence there is no method to distinguish targets from observing the stimulus alone. As the subject only needs to observe the stimuli and count the targets, no movements are required. With an absolutely static user, the adversary cannot gain any information about the secret.

But if the user does not comply with the authentication protocol, for example by counting the targets loudly or tapping a finger after a target, the secret would be leaked. Even by doing this occasionally or by any other observable activity that randomly follows a target with a distribution different from a non-target, the adversary may eventually learn the secret. There is also the risk that entropy is leaked by involuntary movements of the user, but to the best of our knowledge, there is no research published on external reactions to a recognition task.

One risk for observable reactions is caused by our design decision to use a fixed number of targets per burst. After having recognized all five targets in a burst, for the counting task, the user does not need to pay attention any more and might for example stop gazing at the fixation cross or start other behaviour like blinking. Two out of 26 subjects reported reduced attention to the task after having identified five targets.

If the adversary can identify the change of the state of the user that all targets have been shown in that burst, after multiple observations, the adversary can learn the secret by excluding all images displayed after the state change. To overcome this drawback of our system the number of targets displayed per burst needs to be randomized so that the user cannot be sure that no more targets are to be displayed. The randomization needs to be applied to the non-targets in the same way, so that targets cannot be identified by their distribution. This would be fulfilled by drawing a random image from the set with equal probability for all images in each trial. This would result in bursts where some images are missing and some are shown multiple times. But the setup might still be flawed especially for a low number of bursts: In the event of a very low or very high number of targets appearing, the error rates may be affected. For example a user that has a very weak P300 response on a certain target, will probably not be able to login when only this target is part of the burst. With too many targets, no P300 response will be generated, which impedes the system to identify the secret knowledge. In general, if there is a correlation between the success of login and the set of targets appearing in the login process, information may be leaked. The possibility of very deviant target rates could be overcome for example by creating a burst as proposed and add a small number of images from the set which are then displayed twice. In this case a rule to prevent the same image subsequently may be advisable, as otherwise the system may appear to the user hanging for a moment.

Other Attacks As with knowledge-based authentication schemes in general, our proposed system is susceptible to rubber-hose attacks. An adversary can coerce a user to reveal the password image set.

A shoulder surfing adversary able to access the EEG can also obtain secret from the EEG. Placing an additional EEG system on the subjects scalp is probably not well feasible in most situations. But the adversary could compromise the authentication system, or the EEG headset, or eavesdrop on the data sent from the headset to the system. As contemporary EEG headsets are not designed as security devices they might not have proper protection mechanisms and may offer side channels containing information from the EEG. Side channels might manifest in electromagnetic emanations. Although the EEG itself is an electrical signal, it is hard to receive signals remotely as the signals are very weak and of low frequency, ideally requiring large size antennas.

Knowing the secret, an adversary can log into the system when using a general P300 classifier. This is more difficult for a system that relies on more individual classifiers and thus being more biometric. The adversary can counter this with a presentation attack when possessing a complete set of authentication EEG readings of the victim. In principle, a device can be built that uses a camera to identify each image, to select the correct part of the data for playback and a wig that is able to shield the EEG of the

adversary and is able to replay the correctly ordered EEG data of the victim.

4.4. Conclusion

The study demonstrated a visual password authentication scheme based on a basic BCI. The usage of an EEG impedes shoulder surfing but requires multiple presentations of each image. To minimize the required time, an ISI of $200\,\mathrm{ms}$ was used. Due to the size of the image sets, which is important as protection against password guessing, the time required for the authentication process is still too high for a practical operation. Methods to further reduce the time were discussed.

Although the P300 response is known to be influenced by several parameters, the different experimental conditions compared in this study did not show significant differences in the performance of the authentication scheme, suggesting it may work with many types of images and even with reduced attention of the user during authentication.

The partial biometric approach using individual classifiers helped to decrease the error rates. Integrating more sophisticated biometrics may further improve the accuracy, thus helping to reduce the time required for the authentication process.

5. SSVEP Biometrics

To optimize the ITR, BCIs based on the SSVEP paradigm (see Sect. 2.2.2) often use stimulation frequencies individually adjusted for each subject to yield the maximum SNR. To achieve this, the experimental protocol is preceded by a 'sweep' stage, where different frequencies are tested, revealing the SNR for each tested frequency. These 'frequency responses' are obviously individual to some degree, but as BCI experiments are often not repeated after some time, it is not clear to which extent these individual frequency responses are subjected to change. There are only few works on SSVEP-based biometrics (see Sect. 3.1.2) and not all of those assess the permanency. Therefore, we decided to contribute by conducting five SSVEP 'sweep' experiments on a small number of subjects. The experiment was set up and conducted by a student within the scope of a project work. The project work also included creating MATLAB scripts to handle the data and analyse them to assess the applicability for biometrics. For this thesis, the data were analysed differently and reassessed.

5.1. Methods

5.1.1. Paradigm

The authentication scheme presented is based on individual responses of the subjects to a range of SSVEP stimuli. The user focuses on the center of a screen that shows a rectangle toggling between black and white at a certain frequency. Such a stimulation trial lasts for $5\,\mathrm{s}$ and is preceded by an ISI of $1.5\,\mathrm{s}$ to $2.5\,\mathrm{s}$ that shows a fixation dot on black background. The sequence of stimuli is depicted in Fig. 5.1. There are 13 stimulation frequencies: $15\,\mathrm{Hz}$, $17\,\mathrm{Hz}$, $19\,\mathrm{Hz}$, $21\,\mathrm{Hz}$, $23\,\mathrm{Hz}$, $27\,\mathrm{Hz}$, $29\,\mathrm{Hz}$, $31\,\mathrm{Hz}$, $33\,\mathrm{Hz}$, $35\,\mathrm{Hz}$, $37\,\mathrm{Hz}$, $39\,\mathrm{Hz}$ and $41\,\mathrm{Hz}$. For each frequency, ten trials are presented. These 130 trials yield an experiment session of approximately $15\,\mathrm{min}$. The order of the trials is randomized.

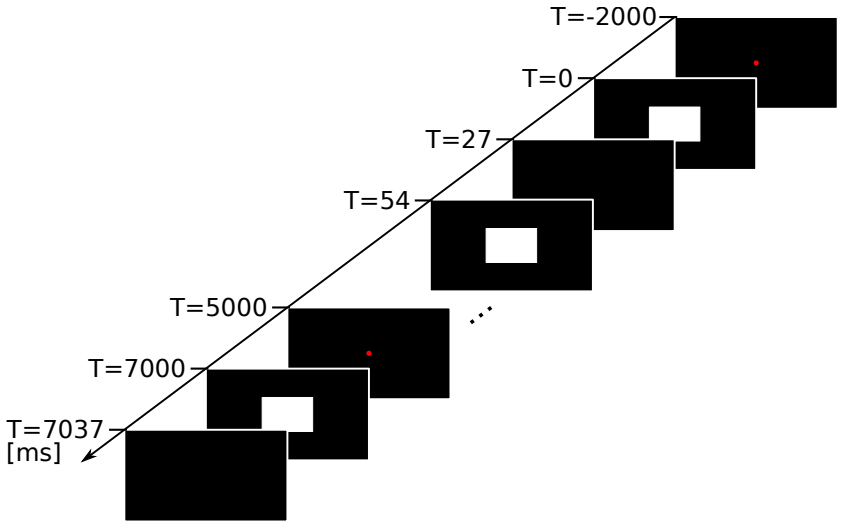


Figure 5.1.: Sequence of SSVEP stimuli with a $37\,\mathrm{Hz}$ trial of $5\,\mathrm{s}$ followed by $27\,\mathrm{Hz}$ stimuli.

5.1.2. Participants

Four male subjects, students or staff of Hamburg University of Technology, participated in the study. A fifth subject withdrew from the study after two sessions. The respective datasets were therefore excluded from analysis. The average distance between sessions was 13 days, between the first and the last session 53 days, varying from 21 to 72 days. The subjects were seated approximately $80\,\mathrm{cm}$ in front of the screen, keeping the stimulation rectangle of 600x400 pixels, with a vision angle of $11.5°$, in approximately the macular area. Subjects were instructed to focus on the stimulus and, to reduce artifacts, only to swallow or blink directly after each trial.

5.1.3. Experimental Setup

Stimuli were delivered by a Windows PC with a 24" $144\,\mathrm{Hz}$ gaming screen (AOC G2470PG) both supporting NVIDIA G-SYNC [73]. G-SYNC allows the graphics card to control the screen refresh to sync it to completely rendered frames. A custom made software using SFML [31] generated properly timed frames to achieve the required stimulation frequencies. For synchronizing the stimulation to the EEG, markers were sent via a virtual serial port to EMOTIV Testbench v1.5.1.2. An EMOTIV EPOC+ (see Sect. 2.1.5), configured to a sampling rate of $128\,\mathrm{Hz}$, was used for data acquisition.

5.1.4. Data Analysis

Data were analysed using MATLAB and the FieldTrip toolbox [74]. With the help of FieldTrip, raw EEG data were cut into trials of six seconds, consisting of the five seconds of SSVEP stimulation and the last second of the preceding break, which was later used as a baseline. Trials classified as containing eye artifacts by the z-value-based automatic artifact rejection of FieldTrip (channels AF3 and AF4, threshold 10) were excluded from further processing. The preprocessing stage included linear trend removal, rereferencing to common average, a $50\,\mathrm{Hz}$ notch filter and a bandpass of $3\,\mathrm{Hz}$ to $60\,\mathrm{Hz}$. The bandpass in principle allows to analyse the second harmonic for every stimulation frequency up to $29\,\mathrm{Hz}$, but for stimulation frequencies above $21\,\mathrm{Hz}$, the second harmonics will be attenuated due to the cutoff frequency of the EPOC+ of $45\,\mathrm{Hz}$ (see Sect. 2.1.5).

In subsequent data analysis only the occipital electrodes O1 and O2 were considered. For each trial, the average time-frequency representation (TFR) with respect to the power spectrum was computed using a hanning window with a fixed length of $0.5\,\mathrm{s}$. TFRs were baseline corrected using data of the ISI before each stimulus, range $-1\,\mathrm{s}$ to $-0.25\,\mathrm{s}$, by computing the relative change which from all samples of the stimulus $x_{t_{stim}}$ subtracts the mean of the baseline $\mu_{baseline}$ and divides the result by $\mu_{baseline}$.

$$\|x_{t_{stim}}\| = \frac{x_{t_{stim}} - \mu_{baseline}}{\mu_{baseline}} \tag{5.1}$$

The $0.25\,\mathrm{s}$ before stimulus onset were not considered for the baseline, as for this range the window of $0.5\,\mathrm{s}$ overlaps with the stimulus interval. Respectively, from the stimulation phase data, the first and last $0.25\,\mathrm{s}$ were disregarded yielding 36 TFR timebins for further processing.

Out of the TFRs, for each stimulus frequency, only the timebins closest to the $f/2$ subharmonic, the stimulus frequency and, if below $60\,\mathrm{Hz}$, the second harmonic were selected, three at most.

The average of the channels O1 and O2 was computed, halving the amount of data. These data were used to create boxplots and for statistical analysis. Statistical analysis only included the data at each stimulation frequency for each session, resulting in 65 data sets per subject, 260 in total. These data sets were tested for normality with the Anderson-Darling test. For each frequency, all five sessions'

data sets of a user were compared as groups in analysis of variance (ANOVA) to test for significant intra-subject variance. Inter-subject variance was tested pairwise by using ANOVA on two groups that contained data of all sessions of the respective subjects. For all hypothesis testing, the significance level α was set to 0.05.

To form a feature vector for the authentication scheme, medians over the timebins and over all trials of the same stimulation frequency in one experiment session were computed. Additionally, to be able to assess the effect of the trial length, medians over only the first 75%, 50%, and 25% of the timebins were computed.

For each experiment session a feature vector containing the median power values of all tested stimulation frequencies and their regarded (sub)harmonics were created, resulting in vectors of size 33. Feature vectors are normalized to a length of 1.

To obtain the error rates of the biometric authentication system, for each feature vector the self distance, which is the average euclidean distance to all other feature vectors of the same subject, and the cross distance with respect to all other subjects' feature vectors are computed. The varying threshold for the authentication decision is applied to the self distance and alternatively to the self distance divided by the cross distance.

5.2. Results

This section describes the results of the proposed authentication system in terms of the error rates and the statistics of the measurements which the feature vector of the authentication system is drawn from. The working principle of the approach is illustrated by detailed data of the $19\,\mathrm{Hz}$ stimulation as an example.

5.2.1. Descriptive Statistics of SSVEP responses

The boxplots in this subsection show the relative change of power at the respective frequencies for the stimulus at the specified frequency compared to no stimulus (see Sect. 5.1.4). Values are capped to ten. Outliers are represented by the '+' symbol. For the stimulus frequency of $19\,\mathrm{Hz}$ Fig. 5.2 shows the responses at the frequencies $9.5\,\mathrm{Hz}$, $19\,\mathrm{Hz}$ and $38\,\mathrm{Hz}$ of all sessions of all four subjects. At $9.5\,\mathrm{Hz}$ all subjects exhibit similar low values around zero. At $19\,\mathrm{Hz}$ and $38\,\mathrm{Hz}$, for all subjects, median values above zero are found. While at $19\,\mathrm{Hz}$ subject one and three have median values in the range of 1 to 1.5 and subject two and four have medians of approximately four and eight respectively, $38\,\mathrm{Hz}$ shows a different pattern: subjects one, two and four have median values below one, while subject three shows an average of about four.

Fig. 5.3 shows the SSVEP responses to $19\,\mathrm{Hz}$ stimulation for subject one broken down by the experiment sessions. For all (sub)harmonics, medians and deviations are very similar. An exception is the elevated median at $19\,\mathrm{Hz}$ for session two.

Fig. 5.4 shows the SSVEP responses to $19\,\mathrm{Hz}$ stimulation for subject two. While for $9.5\,\mathrm{Hz}$ and $38\,\mathrm{Hz}$ values are similar for all sessions, at $19\,\mathrm{Hz}$ the medians vary between approximately two and seven (sessions one and three respectively).

Fig. 5.5 shows the SSVEP responses to $19\,\mathrm{Hz}$ stimulation for subject three. The values for $9.5\,\mathrm{Hz}$ and $19\,\mathrm{Hz}$ are very similar for each session. For the second harmonic, $38\,\mathrm{Hz}$, where subject three shows higher values, there is also more variation. Mean values lie between about two and six (sessions three and one respectively).

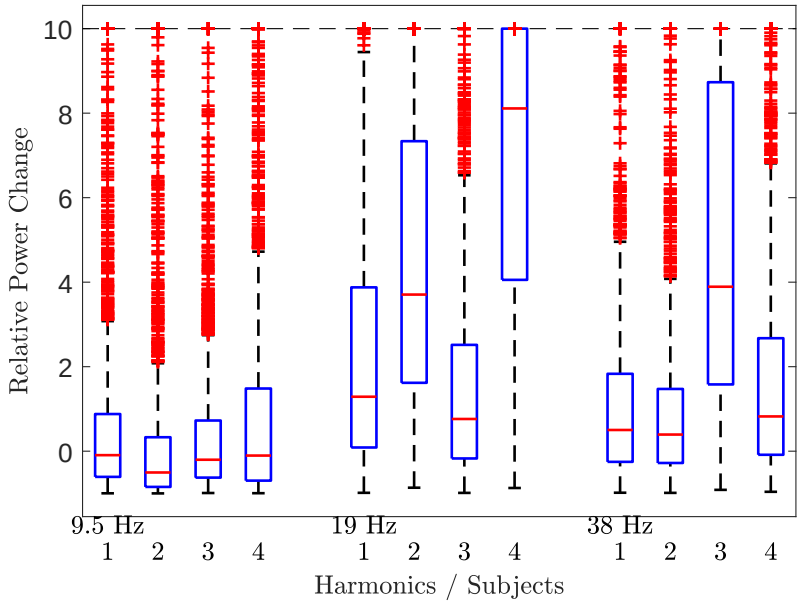


Figure 5.2.: SSVEP responses to 19 Hz stimuli by subject including data from all sessions.

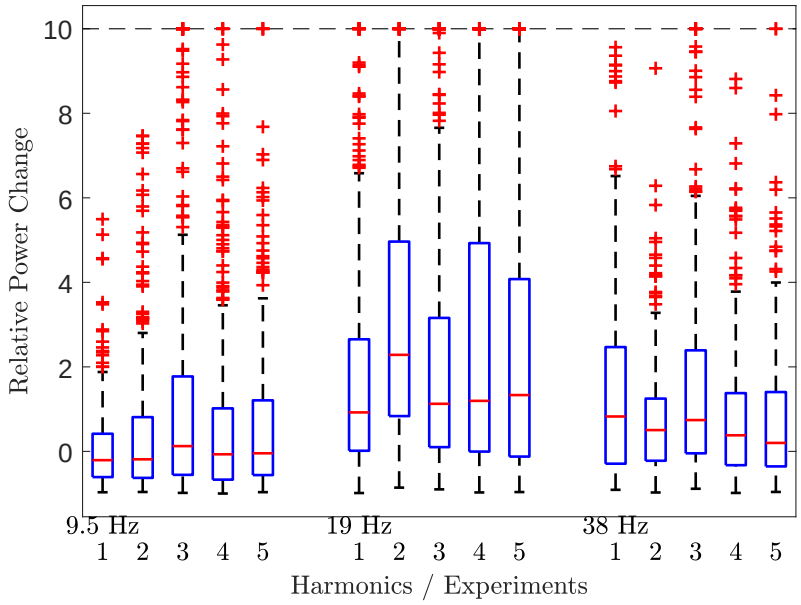


Figure 5.3.: Subject one SSVEP responses to 19 Hz stimuli.

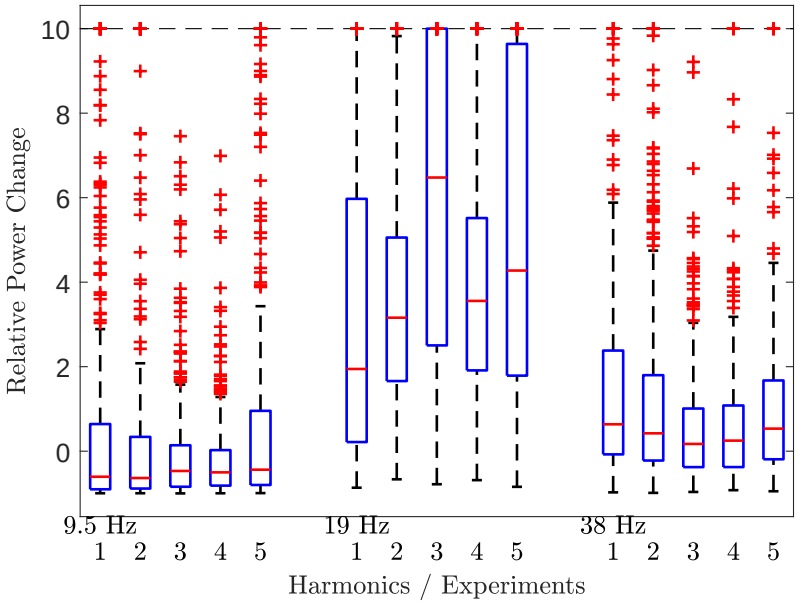


Figure 5.4.: Subject two SSVEP responses to 19 Hz stimuli.

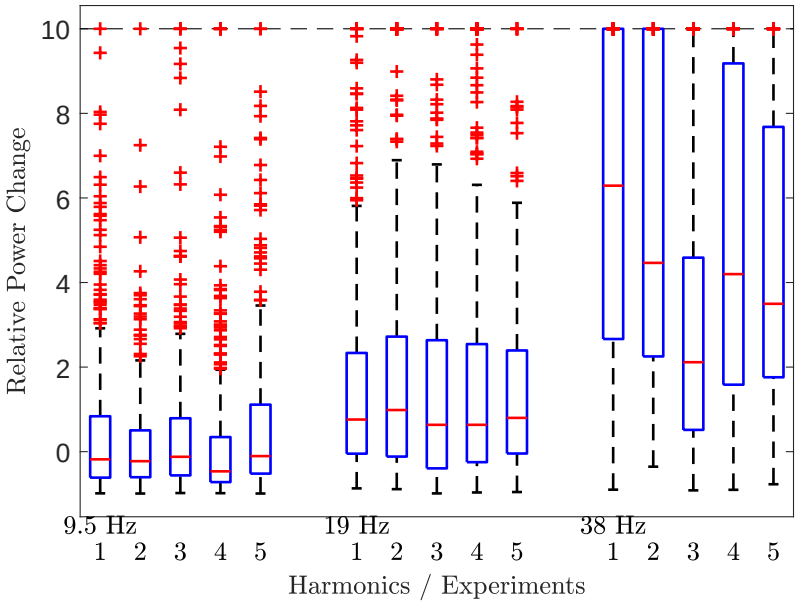


Figure 5.5.: Subject three SSVEP responses to 19 Hz stimuli.

Fig. 5.6 shows the SSVEP responses to 19 Hz stimulation for subject four. In all three frequencies the variation between the sessions is higher compared to the other subjects. Most notable at 38 Hz, where for sessions three to five the median values are just slightly above zero, while for session one and two they are at approximately two.

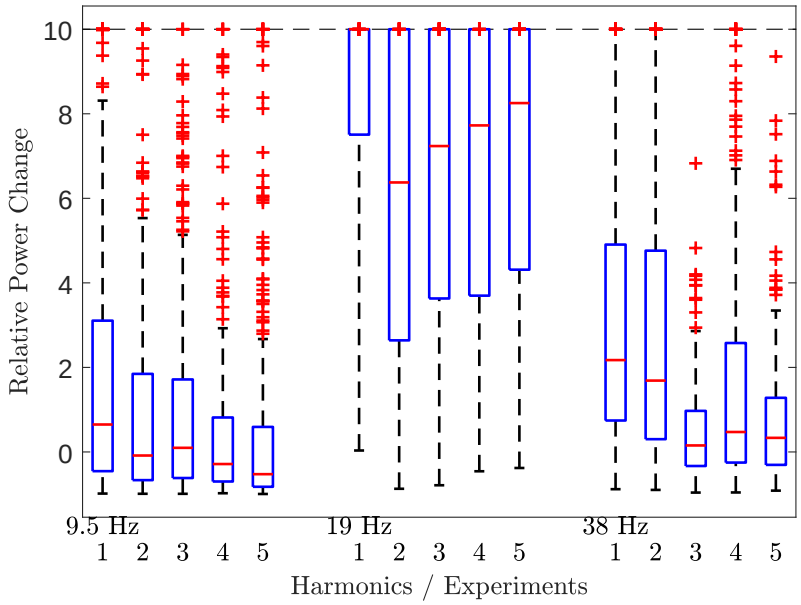


Figure 5.6.: Subject four SSVEP responses to 19 Hz stimuli

The SSVEP responses at stimulation frequencies over stimulation frequency are shown in Figs. 5.7 and 5.8. While subjects one and three have relatively low slopes, showing peaks at 33 Hz and 35 Hz respectively, subjects two and four show much higher peaks of about eleven and eight around 31 Hz and 19 Hz respectively. Fig. 5.8 shows the detailed statistical information as boxplots for all frequencies and subjects.

Significance Tests The Anderson-Darling test reached the significance level of $\alpha = 0.05$ for every dataset, hence the H_0 hypothesis that the samples originate from a population with a normal distribution was rejected. On average, 11.9 of the 13 tests per pair using ANOVA reached the significance level of $\alpha = 0.05$ and hence the H_0 hypothesis that the samples originate from populations with the same mean was rejected in these cases. The count of rejections by pair of subjects compared and the respective average F-values are shown in Tab. 5.1.

Table 5.1.: Pairwise comparison of user data sets using ANOVA for 13 stimulus frequencies

(a) Number of H_0 rejections

Sbj.	1	2	3	4
1	13	13	11	10
2	13	13	13	13
3	11	13	13	10
4	10	13	10	12

(b) Average F-values

Sbj.	1	2	3	4
1	19.5	462.1	28.6	149.1
2	462.1	33.1	361.5	549.9
3	28.6	361.5	24.3	159.4
4	149.1	549.9	159.4	18.7

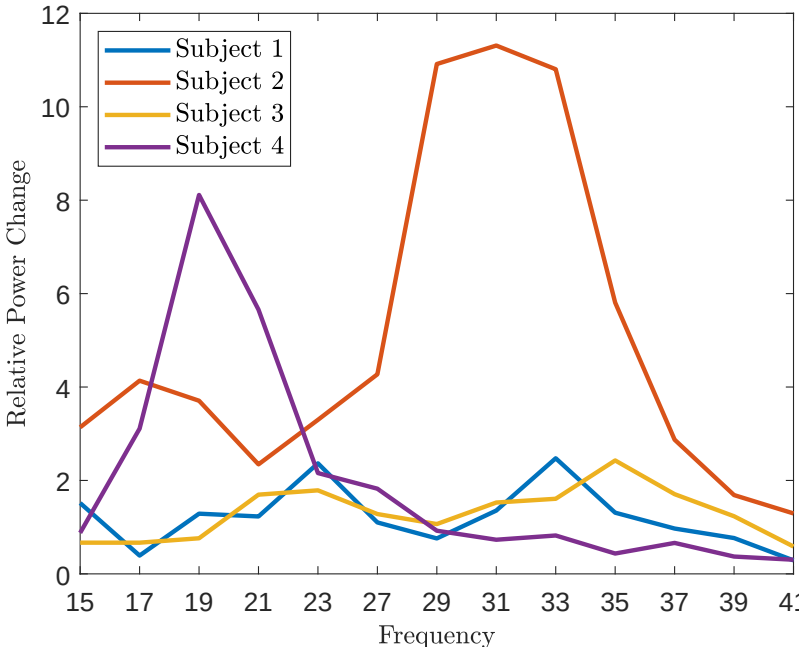


Figure 5.7.: SSVEP responses at stimulus frequency.

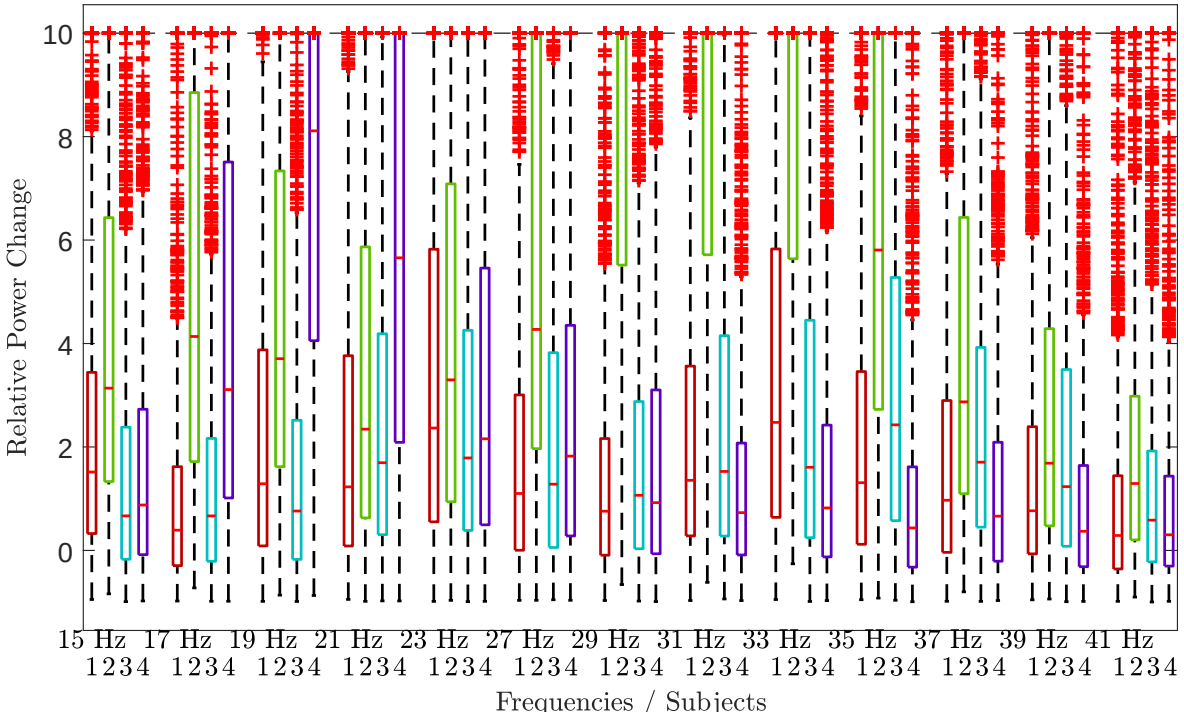


Figure 5.8.: SSVEP responses at stimulus frequency.

5.2.2. Error Rates of the Authentication System

The error rates of the simulated authentication system (see Sect. 5.1.4) are shown in this subsection. Error rates followed by 0.75, 0.5 and 0.25 refer to the error rates when only the first 75%, 50% or 25% of timebins are used respectively, which simulates a shorter trial length.

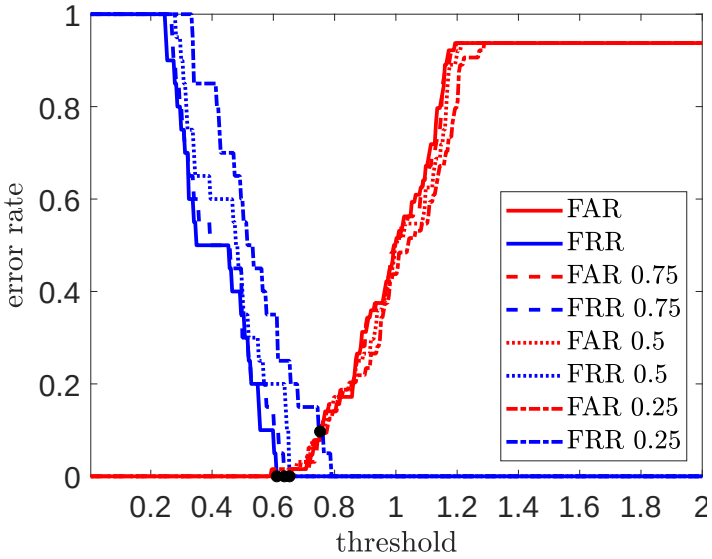


Figure 5.9.: Error rates for varying self distance threshold.

Fig. 5.9 shows the error rates over the varying threshold on the self distance. When regarding the full trial length, the FRR drops to zero at a threshold of about 0.6, where FAR is still at zero, i.e., the EER is zero. With slightly different thresholds, this is also true when only considering the first 75% or 50% of the timebins. With only using the first 25% of the timebins, FRR and FAR intersect at a threshold about 0.8, reaching an EER of 0.0969.

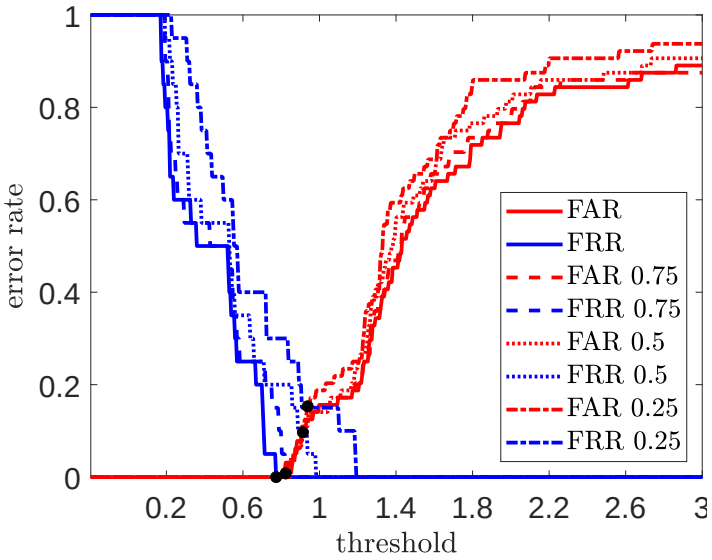


Figure 5.10.: Error rates for varying alternative threshold: $\frac{selfdistance}{crossdistance}$.

When using the alternative threshold $\frac{selfdistance}{crossdistance}$ (see Fig. 5.10) the error curves are more stretched out. When regarding all timebins, an EER of zero is achieved with a threshold of about 0.8. For 75% of

the timebins at a near threshold the EER is 0.0078. At thresholds about 0.9, with 50% and 25% of the timebins EERs of 0.0969 and 0.1531 are achieved respectively.

5.3. Discussion

At a first glance, the error rates of the simulated authentication system, achieving an EER of zero, seem very promising. But Figs. 5.9 and 5.10 show that the range of thresholds with a low EER is quite small, as the FRR and FAR rise quickly. With more than the four users of this study, the risk would increase that two subjects have too similar SSVEP responses, leading to a higher FAR unless the feature vector can be optimized to contain sufficient discriminant information.

The example of $19\,\mathrm{Hz}$ (see Fig. 5.2) shows that using additional harmonics of the stimulus frequency contributes to the discriminant information. While the data of subject three show only a small SSVEP effect at the stimulation frequency, unlike the other subjects, a strong effect at the second harmonic ($38\,\mathrm{Hz}$) can be seen, making the response of subject three very distinct at this point. Unfortunately it cannot be expected that adding an arbitrary number of (sub)harmonics may help, as the SSVEP effect is typically strongest for the first harmonic and the nearest harmonics. Also, higher harmonics may be out of the frequency range measurable by EEG. Low frequency subharmonics on the other hand may be more affected by the pink noise floor of the EEG (see Sect. 2.1).

Increasing the number of tested stimulus frequencies to gain more discriminant information is probably also limited. Figs. 5.7 and 5.8 show frequencies neighbouring ones with a strong SSVEP effect, for example $31\,\mathrm{Hz}$ of subject two, also reach elevated levels. Very strong differences between close frequencies cannot be expected.

5.3.1. Permanency

The main focus of this study was to assess the permanency of the SSVEP responses. As for each subject five sessions were recorded and taken into account for the simulated authentication system, which is able to achieve an EER of zero, the differences of the feature vectors between the sessions must be lower than between the subjects. Inspecting the example of $19\,\mathrm{Hz}$ for each subject (see Figs. 5.2 to 5.6) supports this. But it also shows that measurements may vary between the sessions. The boxplot of subject three (see Fig. 5.5) shows very similar distributions for all sessions at $19\,\mathrm{Hz}$, but varying medians between 2 and 6 for $38\,\mathrm{Hz}$. Subject two (Fig. 5.4) shows a higher variation at $19\,\mathrm{Hz}$ and more similar distributions for $38\,\mathrm{Hz}$ so it seems unlikely that the order of the harmonic generally plays a role in the stability. The statistical tests in most cases show a significant difference when comparing data sets of different users but also when comparing different sessions of the same user. Although this outcome may not be reliable due to the failed normality tests, it indicates that between measurements the SSVEP responses can vary strongly. As the average F-values are lower for the intra-subject tests (see diagonal entries of Tab. 5.2b) it may be concluded that with a feature space of sufficient size it should be possible to differentiate users.

5.3.2. Performance and Usability

The proposed authentication system mainly suffers from the broad distributions of the relative power change values. By using the median for the feature vector sufficiently stable data can be extracted. As it has been tested for 75%, 50%, and 25% (see Figs. 5.9 and 5.10), regarding a shorter portion of the trials degrades the error rates of the simulated authentication system. Such an effect can also be expected when reducing the number of trials, hence reducing the time for the experiment session, will increase

error rates. By going to 25% and optimizing time windows and reducing the ISI the experiment session could probably be reduced from $15\,\mathrm{min}$ to $3\,\mathrm{min}$, which is still too long for a practical authentication system, though the usability score (see Ch. 3 and Eq. 3.1) would increase from 0.0015 at 100% or 0.0053 at 25% without optimization to 0.0066 (for comparison see Tab. 3.3). As long as the system is only intended for biometric authentication and not biometric identification, the protocol could further be optimized by only presenting the stimulus frequencies to an identified user that from the enrollment phase are known to hold the most discriminant information.

5.4. Conclusion

The experiment could verify that SSVEP contain discriminant information that appears to be stable over repeated measurements. The measurements used to derive the feature vector for biometric authentication consist of very broadly distributed values which requires to accept samples quite distant from the template. Hence, with a higher number of users, false positives are expected. Without further improvement the proposed SSVEP-based biometric authentication system is not practically usable. This is also due to the session length of $15\,\mathrm{min}$.

6. CIBA: Continuous Interruption-free Brain Authentication

The practical applicability of authentication schemes relying on EEG suffers from the time needed to mount an EEG device and to record sufficient data. This becomes less important in a scenario, where an EEG headset is already worn for a different purpose. While the headset is worn the EEG readings could be used by a continuous authentication scheme, as long as it does not interrupt the user's actual task. In this chapter, we describe such a scenario and conceptualize CIBA by discussing two possible approaches.

6.1. CIBA Scenario

Recent studies propose using EEG to determine mental parameters like awareness, stress and workload in safety critical scenarios like air traffic control [51, 122, 20]. Safety critical work places usually also have elevated security requirements. An adversary able to perform such a job may cause serious harm. If such a job requires wearing an EEG headset, the stream of EEG data can in principle be analysed for discriminant information, creating a continuous authentication scheme. If an adversary for example takes over the work station of an air traffic controller, the EEG-based system would cause an alarm or log out the adversary even when the mental parameters are in the normal range.

Due to the critical tasks of the users, such a continuous authentication system should not impair the performance of the users in any way. This excludes most EEG-based authentication schemes, as they require the user to perform a task for authentication (see Sect. 3.1). Usually also the presentation of stimuli is involved, which may, even without a task, distract the user from the actual work. CIBA tries to tackle this by using subliminal stimuli without an associated task for authentication. This requires to find the optimal distance to the perceptual threshold where the stimuli entail sufficient discriminant information in the EEG and the performance of the user is least degraded.

Subliminal Stimuli Stimuli below a perceptual threshold are called subliminal, stimuli above supraliminal. A perceptual threshold is specific to a certain kind of stimulus and is usually defined by the ability to detect or to recognize 50% of stimuli. For example, the flicker fusion threshold or critical flicker frequency (CFF), is the frequency where a flicker stimulus starts being perceived as a constant light source. The perceptual threshold of a stimulus can be shifted by masking. In visual masking, the perceptibility of the target stimulus is reduced by presenting another visual stimulus, the mask. The mask can be temporally aligned with the target, presented before (forward masking) or after (backward masking) the target. Masks do not necessarily need to overlap with the target position.

6.2. CIBA SSVEP

The SSVEP paradigm has a property beneficial to the CIBA concept: it only requires spatial attention to the stimulus, i.e., the user is not required to mentally focus on the stimulus itself; it is sufficient if the

stimulus originates from an attended area, allowing the user for example to read a text on a flickering background [67]. This should enable a range of basic office tasks and may even be suitable for more complex work with a PC.

Although background SSVEP stimuli do not require the attention of the user and hence do not interrupt the workflow, the presence of the stimuli might still decrease the user's performance. Flicker in the frequency range used for SSVEP may cause discomfort, strain, and even epileptic seizures. This can be mitigated by optimizing duty cycle and stimulus frequency. Lee et al. [50] showed for a $13.16\,\mathrm{Hz}$ stimulus that discomfort can be decreased by increasing the duty-cycle. Won et al. [117] tested longer duty-cycles with frequencies of $26\,\mathrm{Hz}$ to $34.7\,\mathrm{Hz}$ which did not decrease discomfort but decreased classifier performance. But this frequency caused less discomfort compared to lower frequencies $6\,\mathrm{Hz}$ to $14.9\,\mathrm{Hz}$. If higher frequencies generally decrease discomfort, going to the limits of SSVEP should be considered as the SSVEP paradigm can work at frequencies above the flicker fusion threshold [40, 98]. In a study with flicker stimuli masked by noise, Smout et al. [101] showed that stimuli do not need to be consciously perceived for the SSVEP paradigm to work and confirmed the requirement of spatial attention. As a noise signal mask may still be distracting to the user, we propose to achieve subliminal stimuli by using frequencies above the CFF. As the CFF is higher in the periphery of the visual field [105], we propose to limit the stimulus to an area around the point of gaze to keep stimuli subliminal. A sequence of such SSVEP stimuli is shown in Fig. 6.1. The sequence depicted contains an ISI between two SSVEP trials and an area of a sharp circle used for the stimuli. As these two properties may not be optimal for subliminal stimuli, it can be considered to use a continuous SSVEP that slowly adjusts frequencies and a blurred border of the circle for the stimuli that becomes more like the surroundings in the periphery.

Another property of SSVEP beneficial to the CIBA concept is that the signals are relatively immune to artifacts.

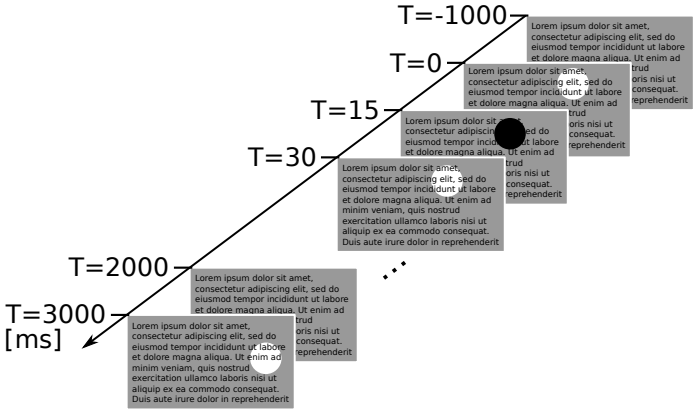


Figure 6.1.: Sequence of $66.67\,\mathrm{Hz}$ SSVEP stimuli for trial of $2\,\mathrm{s}$. Trials are separated by an ISI of $1\,\mathrm{s}$. The stimuli are limited to an area around the point of gaze.

6.3. CIBA ERP

Common ERP-based biometrics approaches are not suitable for the CIBA concept as they require the users to pay attention to (most commonly visual) stimuli, distracting the user from the normal workflow. Even without the need to focus on stimuli, users may be distracted just by perceiving stimuli. Thus, an ERP-based scheme for CIBA ideally is without a task and presents stimuli in a non-distracting way, which may be achieved by using subliminal stimuli which the user does not consciously perceive. Stimuli can

be made subliminal by masking or using a very short stimulus presentation interval (see Sect. 6.1). Stimuli can be masked by aligning them temporally and spatially with other events or selecting stimuli similar to the background.

The P300 of the ERP is discussed in research on neural correlates of consciousness (NCC) [28]. These studies typically present stimuli at sensory threshold, for example by backward masking, and show a decrease of the P300 amplitude under this condition [96, 84, 121]. Pitts et al. [84] confirmed that the P300 strongly depends on a relevant task, which would contradict the CIBA requirements. In a VEP study, Andreassi et al. [4] found decreased amplitudes for backward masked stimuli.

The amplitudes could also be decreased because the user is focusing on a work task. Allison et al. [2] showed this for the P300 amplitude of an auditory counting task which decreased with the difficulty of a game played simultaneously.

These studies indicate that lower ERP amplitudes are expected under the conditions needed for CIBA. Though this does not necessarily mean that discriminant information in the ERP is reduced, a decreased SNR is to be expected. Nevertheless, there are studies reporting detection of the P300 component in subliminal experiments. In an oddball paradigm experiment, Bernat et al. [10] presented stimuli for only $1\,\mathrm{ms}$, ensuring subliminal presentation. For the rare case of the two pictures of words used, an increased P300 component was found.

An ERP experiment aiming to probe if a person is known to the subject was conducted by Frank et al. [29]. The subjects were instructed to watch a movie, where occasionally, for intervals of $13.3\,\mathrm{ms}$, blurred faces and rarely known faces were overlaid, creating an oddball paradigm. Most subjects reported having perceived overlaid images to different degrees, seven of 22 were able to name the depicted person. In most cases, the appearance of the known face could be detected from the EEG. As no non-blurred unknown faces were shown in the experiment, the effect might be solely based on general face detection, not recognition of a known person. Detection whether a person is known to a subject can be used as a knowledge factor for authentication.

Meijer et al. [63, 64] conducted non-subliminal studies aiming to identify known faces in a concealed information test (CIT). With pictures of persons of lesser importance to the subject and not a part of the subject's task, detection of known faces did not succeed. It was concluded that viewing a familiar face may suffice to elicit a P300, but the P300 increases with stronger familiarity or a relevant task. The experiments also indicate that P300 may be elicited by mere recognition of autobiographical information.

Information being autobiographical or faces being known depends on the individual person, i.e., they contain discriminant information. Both lead to a P300 component which is relatively strong under normal conditions, making such stimuli promising candidates for the CIBA concept. Using familiarity of stimuli as a feature requires to investigate whether unfamiliar stimuli may become familiar when repeatedly presented in a subliminal way and thus alter the familiarity feature. If such a process can be effectively modelled, the templates can be adapted accordingly when using the system. The approach presented in Ch. 4 does not seem appropriate for CIBA even if the stimuli were presented subliminally, as the user must identify the password images for the counting task. But as with many authentication runs the password images may become increasingly familiar to a user, even a subliminal presentation supposably elicits a P300, which would allow using the password image set for CIBA.

Other types of stimuli used in ERP authentication schemes (see Sect. 3.1.1) should also be tested for discriminant information under the condition of subliminal and task-free presentation.

In a P300 BCI study, Brunner et al. [15] showed increased accuracy for gaze directed to the target

stimulus. By using an eye tracker, stimuli could be placed in the foveal area to make sure they are seen. This may help accuracy in the CIBAERP concept but it may also render the stimuli less subliminal.

A sequence of gaze dependent SVLO stimuli is shown in Fig. 6.2.

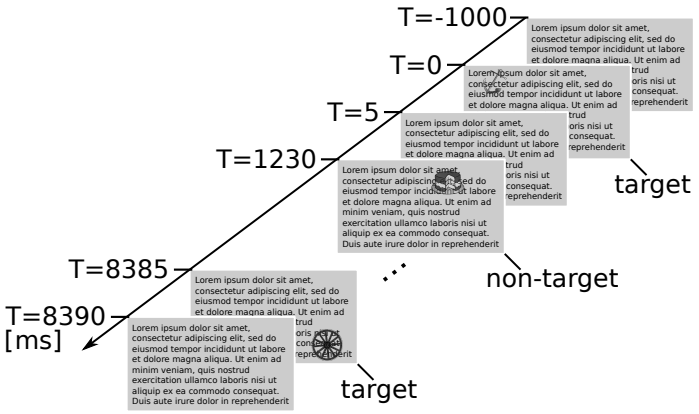


Figure 6.2.: Sequence of SVLO stimuli overlaid at the point of gaze for $5\,\mathrm{ms}$. Trials are separated by an ISI of $1\,\mathrm{s}$

6.4. Considerations for an Implementation

Both CIBA paradigms presented require a common basic setup to deliver stimuli and measure and analyse the EEG. We assume the user to be seated in front of a screen that is used for the actual task and also the stimulus presentation. Instead of a screen, a HMD could be used, or alternatively an optical head-mounted display (OHMD) could deliver the stimuli, which would also allow the user to do work which is not based on viewing a screen.

The display used for delivering the stimuli must be able to refresh quickly and precisely. The refresh rate must at least be twice the desired frequency of the SSVEP stimuli. The shortest duration for presenting a stimulus for an ERP is the inverse of the refresh rate. A liquid-crystal display (LCD) with a maximum refresh rate of $60\,\mathrm{Hz}$ would allow no more than $30\,\mathrm{Hz}$ SSVEP and durations of $16.7\,\mathrm{ms}$. Also, with a fixed refresh rate, only frequencies that are divisors of the refresh rate can be precisely presented. Contemporary gaming monitors offer refresh rates of $144\,\mathrm{Hz}$ and adaptive synchronization technology like FreeSync, which allows to dynamically control the exact point in time for the screen to refresh, enabling proper presentation of arbitrary frequencies in a specified range. Therefore, we consider a screen with adaptive synchronization and a refresh rate of $200\,\mathrm{Hz}$ for stimulus presentation in the CIBA setup.

To allow arbitrary work on the screen, we propose to control the stimulus overlay by a separate application. This application would run a transparent window on the top layer of the desktop environment allowing input (i.e., mouse clicks) to pass to the layer below. Screen synchronization would be tied to this application. In case of ERP stimuli, these can simply be rendered to an arbitrary position inside the application window for a short period of time. The sequence of such stimuli is shown in Fig. 6.2. For SSVEP stimuli, the transparency of the application window needs to be dynamically adjusted. In the area of the stimuli, the window would switch between fully transparent and non-transparent (black) to create stimuli without fully blocking the screen content. With a stimulus frequency above the CFF, this area will appear dimmed. For a uniform appearance, the area without stimuli would also be dimmed by setting an appropriate static transparency level. The sequence of the transparency switching process

is shown in Fig. 6.1.

Both paradigms require that the stimuli are actually displayed in the field of view of the user. With the help of an eye tracker, stimuli can be dynamically placed with regard to the gaze, allowing for example to place stimuli in the foveal area or at a certain distance. The eye tracker can also help in identifying eye artifacts (see Sect. 2.1.2). For this reason the eye tracking information is fed to the EEG system. Other sensors could be added to detect other motion artifacts. Artifacts are expected to happen frequently, as instructing the user to refrain from creating artifacts is not suitable as it may impede the performance in the actual tasks.

The EEG system used for data acquisition in the CIBA setup should have a high spatial resolution and, especially for SSVEP, also a high temporal resolution. Spatial information may help to obtain discriminant information in the ERP. To record high frequency SSVEPs of about $75\,\mathrm{Hz}$ and possibly their second harmonics, a sampling rate of at least $300\,\mathrm{Hz}$ is required. Thus, the Emotiv Epoc, which was used for the experiments conducted for this thesis would not suffice due to the limited bandwidth of $0.2\,\mathrm{Hz}$ to $45\,\mathrm{Hz}$ (see Sect. 2.1.5). Research grade EEG systems like the Biosemi Active Two fulfill the requirements.

The three stages of processing EEG data, artifact rejection, feature extraction, and classification need to be tuned to the operating environment, which may differ between an experimental and real world setup. Feature extraction and classification also vary with the paradigm and the actual discriminant features to be found. The stimulus controller application may retrieve user specific stimuli from the user database to maximize discriminant information. In principle, trials of both CIBA paradigms can be used in one session by taking turns or also concurrently. The stimulus controller will also insert ISIs to get the EEG baseline and randomize stimulus parameters like timing if necessary.

The components of the CIBA setup are depicted in Fig. 6.3.

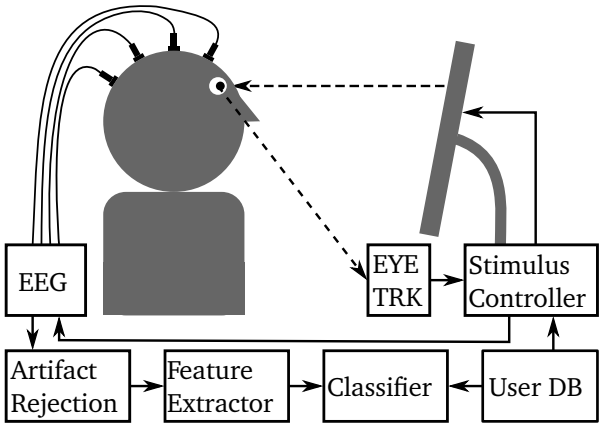


Figure 6.3.: The setup for CIBA consists of a stimulus controller presenting stimuli on the screen at the point of gaze, which is obtained from the eye tracker (EYE TRK). Stimulus information consisting of timing, position and type of stimulus is added to the EEG. Appropriate stimuli are selected from the user database. Artifacts are removed from the EEG before feature extraction. For the authentication decision the classifier compares the extracted features to templates stored in the user database [32].

6.5. Discussion

Both approaches presented for the CIBA concept rely on well-established paradigms that are used in BCIs, but are rather uncommon in EEG-based biometrics (see Sect. 3.1). Hence, it is not clear if suffi-

cient discriminant information can be revealed. The requirement to use the paradigms without a task for the user and with subliminal stimuli will probably lead to weaker signals which may impede gathering discriminant information. The SNR can generally be improved by averaging over more samples, requiring more or longer trials. But extending the sliding window of EEG data used for continuous authentication increases the response time for deauthentication, decreasing the security of the system. Alternatively, it can be investigated whether supraliminal stimuli can be used in a way minimizing the distraction. Even for the subliminal case, it needs be made sure that subconsciously processed stimuli do not significantly reduce the users' performance in the regular work.

CIBA SSVEP relies on flicker above the CFF. Before LCDs became standard, people were exposed to flicker of frequencies $50\,\mathrm{Hz}$ to $100\,\mathrm{Hz}$ from cathode-ray tube (CRT) TVs and computer screens. Negative effects of CRTs using lower refresh rates on the performance of visual tasks were found by Bridgeman et al. [13] for $60\,\mathrm{Hz}$ and $500\,\mathrm{Hz}$, and by Menozzi et al. [65] for $48\,\mathrm{Hz}$, $60\,\mathrm{Hz}$ and $75\,\mathrm{Hz}$. But in both studies no statistical significance was found. Ziefle et al. [126] found increased visual performance when using an LCD compared to a CRT at $100\,\mathrm{Hz}$. The LCD also reduced discomfort due to eye strain and was preferred by 18 of 24 subjects. These difference may also be caused by technical differences between LCDCRT other than the flicker of CRTs, hence the findings may not apply to flicker stimuli presented on LCDs as used for CIBA SSVEP.

In a study comparing lighting by fluorescent tubes flickering at $100\,\mathrm{Hz}$ and $64\,000\,\mathrm{Hz}$, Knez [47] found affective and cognitive effects, including increased problem solving performance for the high frequency flicker, although $100\,\mathrm{Hz}$ flicker can already be considered above the CFF and hence subliminal. Negative effects of subliminal stimuli have also been found by Tsushima et al. [109] who used moving dots as a distractor from a task. Dependent on the ratio of coherent motion, the direction was above or below the threshold of perception, which was found at 5%. While for clearly supraliminal and clearly subliminal motion ratio the performance of the task was not degraded, it degraded for stimuli below but close to the perceptual threshold. As in the CIBA concept stimuli would also be presented close to the threshold, there is a risk that the performance of the user degrades, both for the SSVEP and the ERP approach.

Besides users' performance, it needs to be studied how (subliminal) continuous authentication affects the subject's well-being. Both, performance and well-being, might not only be affected by the presentation of stimuli itself, but also by the subject knowing to be probed continuously.

Both approaches seem suitable for the CIBA concept, have similar risks and require a similar technical setup. At this point of literature research it is not possible to answer the question which of the two approaches is more promising. In SSVEP the stimulus is straightforward and discriminant features need to be found in the SSVEP responses. These may not be limited to mere signal power of stimulus frequency and harmonics as used in Ch. 5. The ERP approach allows a huge variety of different stimuli which may reveal very different amounts of discriminant information. The P300-based authentication system presented in Ch. 4 showed little dependence on the stimuli, but the design was intended only to distinguish between target and non-target stimuli not between subtle differences in responses to certain stimuli. With making the stimuli subliminal and eliminating the task, these differences may become more important. Therefore we propose studying the following types of ERP stimuli under the CIBA paradigm:

- Polarizing pictures of food and faces as used by Ruiz et al. [94]
- Password image sets as used in Ch. 4
- Autobiographical stimuli

- Known faces

6.6. Conclusion

Both approaches for CIBA use well known paradigms widely used in BCI research. To meet the requirement of CIBA not to reduce the performance of the user concurrently working on an important task, the parameters of these paradigms need to be pushed to their limits, where they are expected to be less effective. It is not likely that they can be used in a way that does not harm the performance of the user at all. Thus, a trade off between the performance of the authentication system and of the user is expected. Finding an optimum may require extensive studies, as weak effects are expected, requiring many samples for reliable statistics. If parameters yielding features with sufficient discriminant information are found, processing needs to be optimized. It can be expected that real world applications will also require advances in data acquisition systems.

7. Conclusion

EEG readings can be used for authentication schemes relying on biometrics or the voluntary communication of a shared secret. The main disadvantage of EEG-based authentication remains the time required to capture the data which typically ranges from several seconds to several minutes. In addition, mounting the EEG electrodes takes several minutes. Consumer grade EEG headsets using dry or semi-dry electrodes allow faster mounting but typically offer lower signal quality, which needs to be compensated by acquiring more data. Improved systems for recording electrical brain activity would be beneficial for BCIs and biometric schemes. Magnetoencephalography (MEG) and electrocorticography (ECoG) are currently not available as mobile consumer grade devices. Therefore, all experiments presented in this thesis have been conducted with an Emotiv EPOC+ headset that uses semi-dry electrodes. As authentication schemes proposed in this work rely on electricial brain activity, the schemes will profit from improved acquisition devices.

We proposed a knowledge based visual authentication scheme (see Ch. 4) which uses password images presented in a RSVP oddball paradigm to process many trials in a short period of time. Each image is classified whether it is from the password set based on the P300 component of the ERP. When using the GC, the system is a specialized BCI for entering passwords. As such, the ability to login after a period of time depends on the user's ability to recall the password, which may be supported by using visual stimuli. It could be shown that the scheme achieves permanency even when using ICs, which partially renders it a biometric scheme as it trained considering the average individual differences of the P300 component. This could be extended by including characteristics of the P300 that depend on certain stimuli. The scheme could also be extended by integrating discriminant information which is independent of the P300.

The responses to flicker stimuli, SSVEPs, vary between individuals. Evaluating an experiment with five sessions (see Ch. 5), considering the stimulus frequency and harmonics, we could show the permanency of the signals, which allows to use SSVEP for biometrics. Such a scheme could be extended by including multiple simultaneous stimuli or certain sequences of stimuli.

Due to the time needed for mounting an EEG headset, EEG-based authentication schemes are most suitable for practical applications when an EEG headset is already worn. We conceptualized CIBA (see Ch. 6) and discussed how SSVEP and ERP paradigms can be used for continuous authentication without distracting the user from regular work. Flicker stimuli become subliminal at higher frequency but may still entail SSVEPs. Stimuli for ERPs can be displayed shorter or masked to become subliminal. For both subliminal paradigms a decreased SNR must be expected, which may be compensated by longer time windows. The decreased SNR will require lengthy experimental studies to verify the efficacy of the CIBA system. The CIBA criteria could also be met by an authentication scheme that instead of subliminal stimuli uses biometric features of the EEG which are task-independent.

A universal task-independent authentication scheme needs to be based on features which are usually present in the EEG, making it easier for an adversary to obtain a sample. Generally, the ease of creating facsimiles of biometric samples is important for the security of a biometric scheme. This is often considered difficult for EEG-based schemes, but current research lacks the experimental evaluation of

facsimile creation and presentation attacks.

Acronyms

ADC analog-to-digital converter.

ANOVA analysis of variance.

AR autoregressive.

BCI brain-computer interface.

CFF critical flicker frequency.

CIBA continuous interruption-free brain authentication.

CIT concealed information test.

CNN convolutional neural network.

CRR correct recognition rate.

CRT cathode-ray tube.

CV cross-validation.

ECG electrocardiogram.

ECG electrocardiography.

ECoG electrocorticography.

EEG electroencephalogram.

EEG electroencephalograph.

EEG electroencephalography.

EER equal error rate.

ERP event-related potential.

FAR false acceptance rate.

FRR false rejection rate.

GC general classifier.

HMD head-mounted display.

IC individual classifier.

ISI interstimulus interval.

ITR information transfer rate.

k-NN k-nearest neighbors.

LCD liquid-crystal display.

LDA linear discriminant analysis.

MEG magnetoencephalography.

MSE mean squared error.

NCC neural correlates of consciousness.

OHMD optical head-mounted display.

PII personally identifiable information.

PSD power spectral density.

RSVP rapid serial visual presentation.

SNR signal-to-noise ratio.

SSEP steady state evoked potential.

SSVEP steady state visually evoked potential.

SVLO Snodgrass and Vanderwart like object.

TFR time-frequency representation.

VEP visually evoked potential.

Glossary

accuracy accuracy of a binary classifier is the ratio of correctly classified samples to the total number of samples classified. $accuracy = \frac{TP+TN}{TP+TN+FP+FN}$. 29

brain-computer interface input device usable without activation of efferent nerves. i, 1, 65

concealed information test detecting concealed crime-relevant knowledge by measuring reactions to related stimuli. 57, 65

correct recognition rate accuracy of a classifier, ratio of correctly classified samples to the total number of samples classified. 19, 22, 65

critical flicker frequency; also flicker fusion frequency frequency threshold above which flicker becomes imperceptible. 55, 65

efferent nerves nerves carrying information from the brain to peripheral effector organs. 67

electrocardiogram graph of the electrical activity of the heart acquired by electrocardiography. 1, 65

electrocardiography measurment of the electrical activity of the heart. 1, 65

electrocorticography electroencephalography with electrodes placed on the brain (intracranial). 63, 65

electroencephalogram graph of the electrical activity of the brain acquired by electroencephalography. i, 1, 5, 65

electroencephalograph device for recording the electroencephalogram. 1, 65

electroencephalography measurment of the electrical activity of the brain on the scalp. i, 1, 5, 19, 65

equal error rate the error rate of a biometric system using a threshold at the intersection of false acceptance rate and false rejection rate. i, 16, 35, 65, 77

event-related potential EEG response to a certain event. i, 2, 65

head-mounted display device worn on the head with a small display in front of one or each eye. 24, 65

interstimulus interval time between stimulus offset and onset of the following stimulus. 19, 66

magnetoencephalography measurement of magnetic fields produced by electrical activity of the brain. 63, 66

optical head-mounted display device worn on the head capable of projecting images in the user's view. 58, 66

P300 cognitive ERP component 300 ms after a rare and relevant stimulus. 11, 13, 19, 25, 29, 39, 42, 57, 60, 75

precision in binary classification, precision is ratio of true positives (TP) among all samples classified as positive which includes false positives (FP). If there are no false positives, precision reaches the maximum of 1. $precision = \frac{TP}{TP+FP}$. 29

sensitivity in binary classification, sensitivity is ratio of true positives (TP) among all samples of the class positive which includes false negatives (FN). If there are no false negatives, sensitivity reaches the maximum of 1. $sensitivity = \frac{TP}{TP+FN}$. 29

steady state evoked potential EEG response to a periodic stimulus. 66

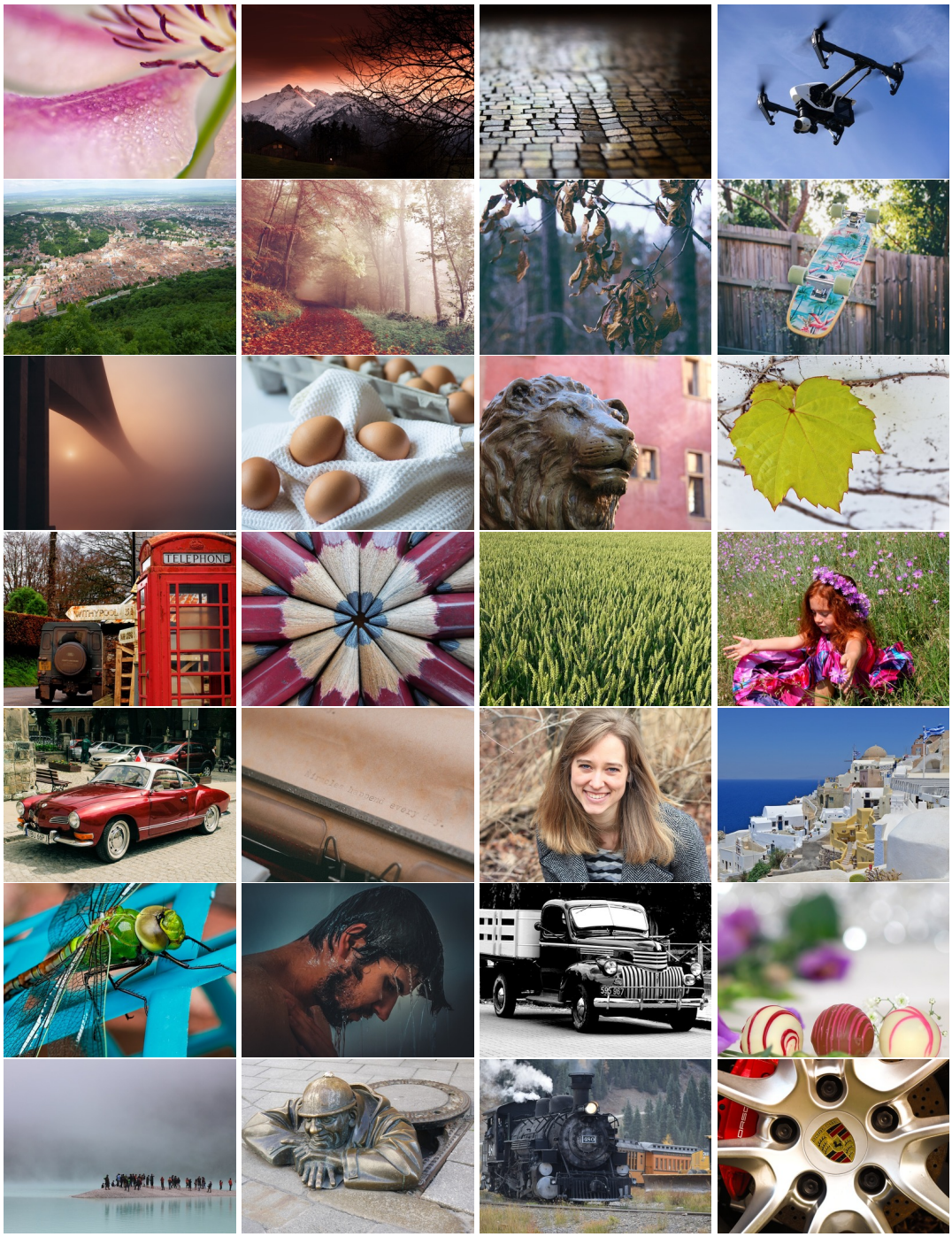
steady state visually evoked potential EEG response to a visual flicker stimulus. i, 3, 66

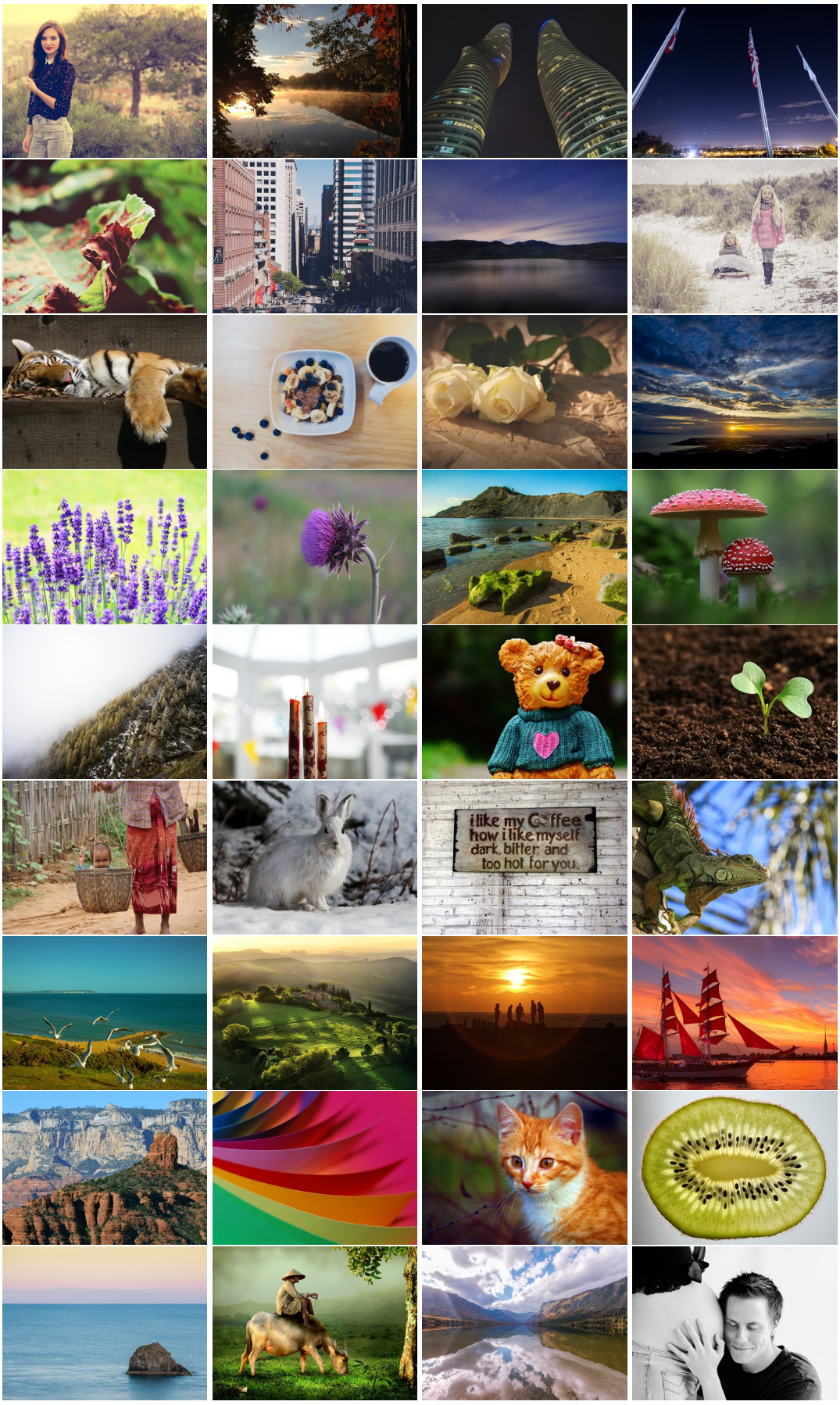
time-frequency representation representation of a signal by spectra over time. 46, 66

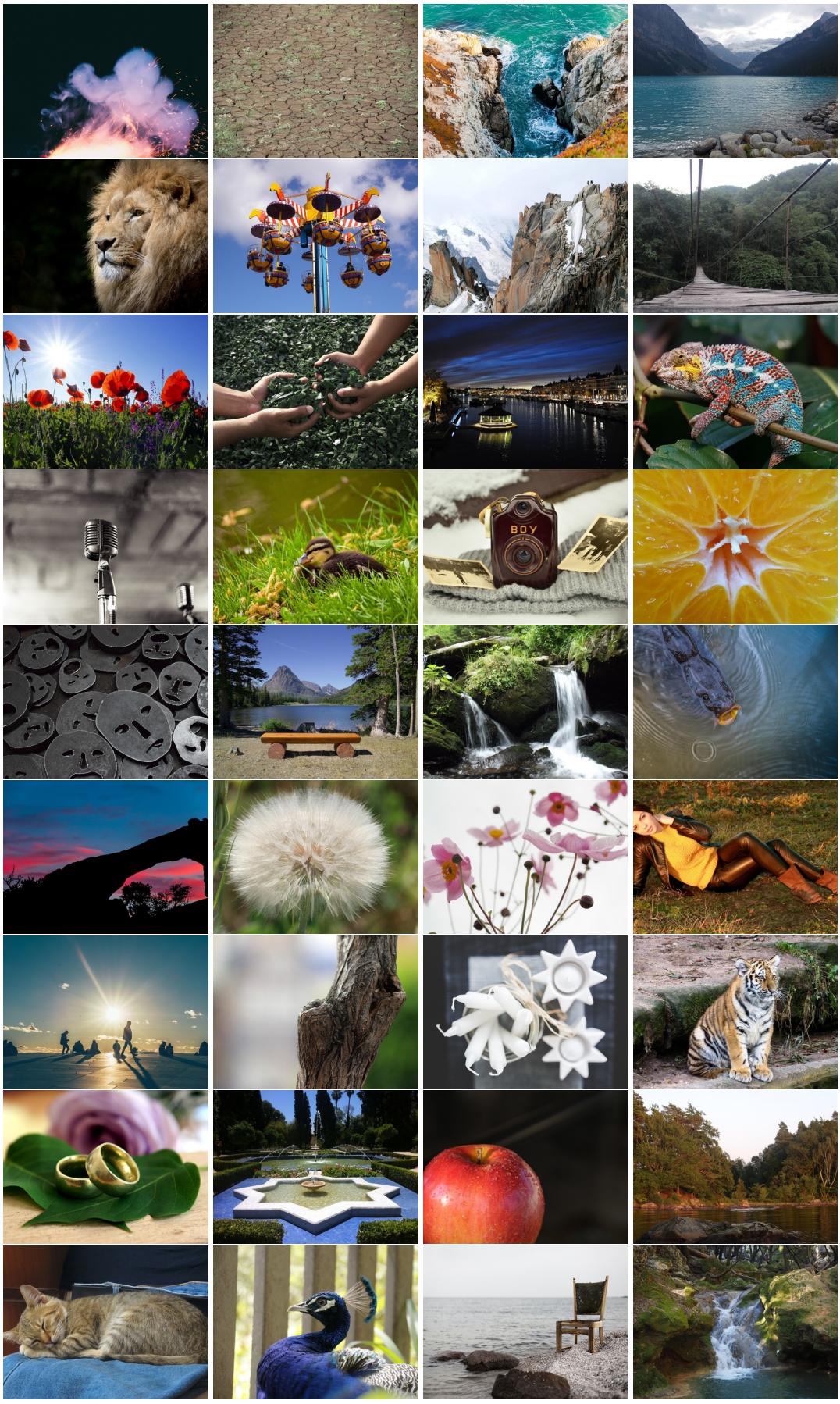
visually evoked potential EEG response to a visual stimulus. 19, 66

A. P300 Visual Authentication Stimuli

A.1. Photos




i like my Coffee
how i like myself
dark, bitter, and
too hot for you.


BOY

A.2. Snodgrass and Vanderwart Like Objects

List of Figures

List of Tables

Bibliography

[1] Shohei Akiyama, Takamasa Shimada, and Tadanori Fukami. Personal identification by steady-state visual evoked potentials based on voting process by Mahalanobis distance. *International Journal of Innovative Computing, Information and Control*, 16(5):1801–1810, 2020.

[2] Brendan Z Allison and John Polich. Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological psychology*, 77(3):277–283, 2008.

[3] Setare Amiri, Reza Fazel-Rezai, and Vahid Asadpour. A Review of Hybrid Brain-Computer Interface Systems. *Advances in Human-Computer Interaction*, 2013.

[4] JL Andreassi, JJ De Simone, and BW Mellers. Amplitude changes in the visual evoked cortical potential with backward masking. *Electroencephalography and Clinical Neurophysiology*, 41(4):387–398, 1976.

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

[6] Asanagi. 21 electrodes of international 10-20 system for EEG. `https://commons.wikimedia.org/wiki/File:21_electrodes_of_International_10-20_system_for_EEG.svg`, 2010. [Online; accessed 16-May-2020].

[7] Garima Bajwa and Ram Dantu. Neurokey: Towards a new paradigm of cancelable biometrics-based key generation using electroencephalograms. *Computers & Security*, 62:95–113, 2016.

[8] E Baykara, CA Ruf, C Fioravanti, I Käthner, N Simon, SC Kleih, A Kübler, and S Halder. Effects of training and motivation on auditory P300 brain–computer interface performance. *Clinical Neurophysiology*, 127(1):379–387, 2016.

[9] Hans Berger. Über das Elektrenkephalogramm des Menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1):527–570, 1929.

[10] Edward Bernat, Howard Shevrin, and Michael Snodgrass. Subliminal visual oddball stimuli evoke a P300 component. *Clinical neurophysiology*, 112(1):159–171, 2001.

[11] Amir Jalaly Bidgoly, Hamed Jalaly Bidgoly, and Zeynab Arezoumand. A survey on methods and challenges in EEG based authentication. *Computers & Security*, 93:101788, 2020.

[12] David H. Brainard. The psychophysics toolbox. *Spatial Vision*, 10(4):433–436, 1997.

[13] Bruce Bridgeman, Michael J Montegut, and Jeff Sykes. High refresh rate and oculomotor adaptation facilitate reading from video displays. *Spatial Vision*, 10(4):305–322, 1997.

[14] Katharine Brigham and BVK Vijaya Kumar. Subject Identification from Electroencephalogram (EEG) Signals During Imagined Speech. In *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8. IEEE, 2010.

[15] Peter Brunner, S Joshi, Samuel Briskin, Jonathan R Wolpaw, Horst Bischof, and Gerwin Schalk. Does the 'P300' speller depend on eye gaze? *Journal of neural engineering*, 7(5):056013, 2010.

[16] Yiyu Chen, Ayalneh Dessalegn Atnafu, Isabella Schlattner, Wendimagegn Tariku Weldtsadik, Myung-Cheol Roh, Hyoung Joong Kim, Seong-Whan Lee, Benjamin Blankertz, and Siamac Fazli. A High-Security EEG-Based Login System with RSVP Stimuli and Dry Electrodes. *IEEE Transactions on Information Forensics and Security*, 11(12):2635–2647, 2016.

[17] R. Das, E. Maiorana, D. La Rocca, and P. Campisi. EEG Biometrics for User Recognition Using Visually Evoked Potentials. In *International Conference of the Biometrics Special Interest Group*, pages 1–8, Sept 2015.

[18] Rig Das, Emanuele Maiorana, and Patrizio Campisi. EEG Biometrics Using Visual Stimuli: A Longitudinal Study. *IEEE Signal Processing Letters*, 23(3):341–345, 2016.

[19] Rig Das, Emanuele Maiorana, and Patrizio Campisi. Visually Evoked Potential for EEG Biometrics using Convolutional Neural Network. In *2017 25th European Signal Processing Conference (EUSIPCO)*, pages 951–955. IEEE, 2017.

[20] Essam Debie, Raul Fernandez Rojas, Justin Fidock, Michael Barlow, Kathryn Kasmarik, Sreenatha Anavatti, Matthew Garratt, and Hussein A Abbass. Multimodal Fusion for Objective Assessment of Cognitive Workload: A Review. *IEEE Transactions on Cybernetics*, 2019.

[21] Tamer Demiralp, Ahmet Ademoglu, Martin Schürmann, Canan Basar-Eroglu, and Erol Basar. Detection of P300 waves in single trials by the wavelet transform (WT). *Brain and language*, 66(1):108–128, 1999.

[22] Yevgeniy Dodis, Leonid Reyzin, and Adam Smith. Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data. In *International conference on the theory and applications of cryptographic techniques*, pages 523–540. Springer, 2004.

[23] Günter Edlinger, Brendan Z Allison, and Christoph Guger. How Many People Can Use a BCI System? In *Clinical systems neuroscience*, pages 33–66. Springer, 2015.

[24] EMOTIV Inc. EMOTIV EPOC Technical Specifications. `https://emotiv.gitbook.io/epoc-user-manual/introduction-1/technical_specifications`, 2019. [Online; accessed 19-September-2020].

[25] EMOTIV Inc. EMOTIV EPOC+. `https://www.emotiv.com/epoc/`, 2020. [Online; accessed 19-September-2020].

[26] Owen Falzon, Rosanne Zerafa, Tracey Camilleri, and Kenneth P Camilleri. EEG-Based Biometry Using Steady State Visual Evoked Potentials. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4159–4162. IEEE, 2017.

[27] Tobias Fiebig, Jan Krissler, and Ronny Hänsch. Security Impact of High Resolution Smartphone Cameras. In *8th USENIX Workshop on Offensive Technologies (WOOT 14)*, 2014.

[28] Jona Förster, Mika Koivisto, and Antti Revonsuo. ERP and MEG correlates of visual consciousness: The second decade. *Consciousness and cognition*, 80:102917, 2020.

[29] Mario Frank, Tiffany Hwu, Sakshi Jain, Robert T Knight, Ivan Martinovic, Prateek Mittal, Daniele Perito, Ivo Sluganovic, and Dawn Song. Using EEG-based BCI devices to subliminally probe for private information. In *Proceedings of the 2017 on Workshop on Privacy in the Electronic Society*, pages 133–136, 2017.

[30] Tadanori Fukami, Yuto Abe, Takamasa Shimada, and Bunnoshin Ishikawa. Authentication system preventing unauthorized access of a third person based on steady state visual evoked potentials. *Int. J. of Innovative Computing Information and Control*, 14(6):2091–2100, 2018.

[31] Laurent Gomila. SFML. `http://www.sfml-dev.org`. [Online; accessed 08-November-2020].

[32] Florian Gondesen and Dieter Gollmann. CIBA: Continuous Interruption-free Brain Authentication. In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 4: BIOSIGNALS*, pages 314–319. INSTICC, SciTePress, 2021.

[33] Florian Gondesen, Matthias Marx, and Dieter Gollmann. *EEG-Based Biometrics*, pages 287–318. Springer International Publishing, Cham, 2019.

[34] Florian Gondesen, Matthias Marx, and Ann-Christine Kycler. A Shoulder-Surfing Resistant Image-Based Authentication Scheme with a Brain-Computer Interface. In *2019 International Conference on Cyberworlds (CW)*, pages 336–343. IEEE, 2019.

[35] Henry Gray. *Anatomy of the Human Body*. Lea & Febiger, 1918.

[36] Christoph Guger, Brendan Z Allison, Bernhard Großwindhager, Robert Prückl, Christoph Hintermüller, Christoph Kapeller, Markus Bruckner, Gunther Krausz, and Günter Edlinger. How many people could use an SSVEP BCI? *Frontiers in neuroscience*, 6:169, 2012.

[37] Qiong Gui, Maria V Ruiz-Blondet, Sarah Laszlo, and Zhanpeng Jin. A Survey on Brain Biometrics. *ACM Computing Surveys (CSUR)*, 51(6):1–38, 2019.

[38] Qiong Gui, Wei Yang, Zhanpeng Jin, Maria V Ruiz-Blondet, and Sarah Laszlo. A Residual Feature-based Replay Attack Detection Approach for Brainprint Biometric Systems. In *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6. IEEE, 2016.

[39] Rajbir Singh Harshit, Kavitha P Thomas, KG Smitha, and AP Vinod. Online Electroencephalogram (EEG) based Biometric Authentication using Visual and Audio Stimuli. In *IEEE EMBS Conference on Biomedical Engineering and Sciences*, pages 454–459, 2016.

[40] Sophie K Herbst, Amir Homayoun Javadi, Elke van der Meer, and Niko A Busch. How Long Depends on How Fast—Perceived Flicker Dilates Subjective Duration. *PloS one*, 8(10), 2013.

[41] Christoph S Herrmann. Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, 137(3-4):346–353, 2001.

[42] Michael Inzlicht, Ian McGregor, Jacob B Hirsh, and Kyle Nash. Neural Markers of Religious Conviction. *Psychological science*, 20(3):385–392, 2009.

[43] Ben H Jansen, Anand Allam, Prashant Kota, Kathleen Lachance, Ayokunle Osho, and Karthik Sundaresan. An exploratory study of factors affecting single trial P300 detection. *IEEE Transactions on Biomedical Engineering*, 51(6):975–978, 2004.

[44] Isuru Jayarathne, Michael Cohen, and Senaka Amarakeerthi. Survey of EEG-Based Biometric Authentication. In *2017 IEEE 8th International Conference on Awareness Science and Technology (iCAST)*, pages 324–329. IEEE, 2017.

[45] Xiaoxuan Jia and Adam Kohn. Gamma Rhythms in the Brain. *PLoS Biology*, 9(4), 2011.

[46] Mario Kleiner, David Brainard, Denis Pelli, Allen Ingling, Richard Murray, and Christopher Broussard. What's new in Psychtoolbox-3. *Perception*, 36(14):1, 2007.

[47] Igor Knez. Affective and cognitive reactions to subliminal flicker from fluorescent lighting. *Consciousness and cognition*, 26:97–104, 2014.

[48] Daria La Rocca, Patrizio Campisi, and Gaetano Scarano. EEG biometrics for individual recognition in resting state with closed eyes. In *2012 BIOSIG - Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG)*, pages 1–12. IEEE, 2012.

[49] Daria La Rocca, Patrizio Campisi, and Gaetano Scarano. Stable EEG features for biometric recognition in resting state conditions. In *International Joint Conference on Biomedical Engineering Systems and Technologies*, pages 313–330. Springer, 2013.

[50] Po-Lei Lee, Chia-Lung Yeh, John Yung-Sung Cheng, Chia-Yen Yang, and Gong-Yau Lan. An SSVEP-Based BCI Using High Duty-Cycle Visual Flicker. *IEEE Transactions on Biomedical Engineering*, 58(12):3350–3359, 2011.

[51] Wei Lun Lim, Yisi Liu, Salem Chandrasekaran Harihara Subramaniam, Serene Hui Ping Liew, Gopala Krishnan, Olga Sourina, Dimitrios Konovessis, Hock Eng Ang, and Lipo Wang. EEG-Based Mental Workload and Stress Monitoring of Crew Members in Maritime Virtual Simulator. In *Transactions on Computational Science XXXII*, pages 15–28. Springer, 2018.

[52] Feng Lin, Kun Woo Cho, Chen Song, Wenyao Xu, and Zhanpeng Jin. Brain Password: A Secure and Truly Cancelable Brain Biometrics for Smart Headwear. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, pages 296–309, 2018.

[53] E. Maiorana, D. La Rocca, and P. Campisi. EEG-based biometric recognition using EigenBrains. In *IEEE International Conference on Multimedia Expo Workshops*, pages 1–6, 2015.

[54] Emanuele Maiorana. Learning deep features for task-independent EEG-based biometric verification. *Pattern Recognition Letters*, 143:122–129, 2021.

[55] Emanuele Maiorana and Patrizio Campisi. Longitudinal Evaluation of EEG-Based Biometric Recognition. *IEEE Transactions on Information Forensics and Security*, 13(5):1123–1138, 2017.

[56] Emanuele Maiorana, Daria La Rocca, and Patrizio Campisi. On the Permanence of EEG Signals for Biometric Recognition. *IEEE Transactions on Information Forensics and Security*, 11(1):163–175, 2015.

[57] Ran Manor and Amir B Geva. Convolutional Neural Network for Multi-Category Rapid Serial Visual Presentation BCI. *Frontiers in computational neuroscience*, 9:146, 2015.

[58] Chengsheng Mao, Bin Hu, Manman Wang, and P. Moore. EEG-based biometric identification using local probability centers. In *International Joint Conference on Neural Networks*, pages 1–8, 2015.

[59] Sebastien Marcel and José del R Millán. Person Authentication Using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):743–752, 2007.

[60] Ivan Martinovic, Doug Davies, Mario Frank, Daniele Perito, Tomas Ros, and Dawn Song. On the Feasibility of Side-Channel Attacks with Brain-Computer Interfaces. In *21st USENIX Security Symp*, 2012.

[61] MATLAB. *version 9.1 (R2016b)*. The MathWorks Inc., 2016.

[62] Erika McCallister, Timothy Grance, and Karen A Scarfone. Guide to Protecting the Confidentiality of Personally Identifiable Information (PII), 2010.

[63] Ewout H Meijer, Fren TY Smulders, Harald LGJ Merckelbach, and Ann G Wolf. The P300 is sensitive to concealed face recognition. *International Journal of Psychophysiology*, 66(3):231–237, 2007.

[64] Ewout H Meijer, Fren TY Smulders, and Ann Wolf. The Contribution of Mere Recognition to the P300 Effect in a Concealed Information Test. *Applied psychophysiology and biofeedback*, 34(3):221–226, 2009.

[65] M Menozzi, F Lang, U Naepflin, C Zeller, and H Krueger. CRT versus LCD: Effects of refresh rate, display technology and background luminance in visual performance. *Displays*, 22(3):79–85, 2001.

[66] George A Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[67] ST Morgan, JC Hansen, and SA Hillyard. Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences*, 93(10):4770–4774, 1996.

[68] Ron Morstyn, Frank H Duffy, and Robert W McCarley. Altered P300 topography in schizophrenia. *Archives of General Psychiatry*, 40(7):729–734, 1983.

[69] Isao Nakanishi and Masashi Hattori. Biometric Potential of Brain Waves Evoked by Invisible Visual Stimulation. In *2017 International Conference on Biometrics and Kansei Engineering (ICBAKE)*, pages 94–99. IEEE, 2017.

[70] Isao Nakanishi and Takuya Yoshikawa. Brain waves as unconscious biometrics towards continuous authentication-the effects of introducing PCA into feature extraction. In *2015 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pages 422–425. IEEE, 2015.

[71] Markus Näpflin, Marc Wildi, and Johannes Sarnthein. Test–retest reliability of resting EEG spectra validates a statistical signature of persons. *Clinical Neurophysiology*, 118(11):2519–2524, 2007.

[72] Raymond S. Nickerson. A note on long-term recognition memory for pictorial material. *Psychonomic Science*, 11(2):58, 1968.

[73] NVIDIA Corporation. Best Gaming Monitors and Displays | NVIDIA G-SYNC. `https://www.nvidia.com/en-us/geforce/products/g-sync-monitors/`. [Online; accessed 08-November-2020].

[74] Robert Oostenveld, Pascal Fries, Eric Maris, and Jan-Mathijs Schoffelen. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011.

[75] R Palaniappan and P Raveendran. Individual identification technique using visual evoked potential signals. *Electronics Letters*, 38(25):1634–1635, 2002.

[76] Ramaswamy Palaniappan. Method of identifying individuals using VEP signals and neural network. *IEE Proceedings-Science, Measurement and Technology*, 151(1):16–20, 2004.

[77] Ramaswamy Palaniappan. Electroencephalogram Signals from Imagined Activities: A Novel Biometric Identifier for a Small Population. In *Intelligent Data Engineering and Automated Learning*, pages 604–611. Springer, 2006.

[78] Ramaswamy Palaniappan. Two-stage biometric authentication method using thought activity brain waves. *International Journal of Neural Systems*, 18(01):59–66, 2008.

[79] Ramaswamy Palaniappan and Danilo P Mandic. Energy of Brain Potentials Evoked During Visual Stimulus: A New Biometric? In *International Conference on Artificial Neural Networks*, pages 735–740. Springer, 2005.

[80] Ramaswamy Palaniappan and Danilo P Mandic. Biometrics from Brain Electrical Activity: A Machine Learning Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4), 2007.

[81] R. B. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles. The electroencephalogram as a biometric. In *Canadian Conference on Electrical and Computer Engineering*, volume 2, pages 1363–1366 vol.2, 2001.

[82] Emanuela Piciucco, Emanuele Maiorana, Owen Falzon, Kenneth P Camilleri, and Patrizio Campisi. Steady-State Visual Evoked Potentials for EEG-Based Biometric Identification. In *2017 International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–5. IEEE, 2017.

[83] Terence W Picton. The P300 Wave of the Human Event-Related Potential. *Journal of Clinical Neurophysiology*, 9(4):456–479, 1992.

[84] Michael A Pitts, Jennifer Padwal, Daniel Fennelly, Antígona Martínez, and Steven A Hillyard. Gamma band activity and the P3 reflect post-perceptual processes, not visual awareness. *Neuroimage*, 101:337–350, 2014.

[85] Pixabay. Public domain imagery. `https://pixabay.com/`, 2019. [Online; accessed 28-January-2019].

[86] Mary C Potter. Rapid Serial Visual Presentation (RSVP): A Method for Studying Language Processing. *New methods in reading comprehension research*, 118:91–118, 1984.

[87] M Poulos, M Rangoussi, V Chrissikopoulos, and A Evangelou. Person identification based on parametric processing of the EEG. In *IEEE Electronics, Circuits and Systems*, volume 1, pages 283–286, 1999.

[88] Marios Poulos, Maria Rangoussi, and E Kafetzopoulos. Person identification via the EEG using computational geometry algorithms. In *European Signal Processing Conference*, pages 1–4. IEEE, 1998.

[89] Marcos Del Pozo-Banos, Jesús B. Alonso, Jaime R. Ticay-Rivas, and Carlos M. Travieso. Electroencephalogram subject identification: A review. *Expert Systems with Applications*, 41(15):6537 – 6554, 2014.

[90] Christian Rathgeb and Andreas Uhl. A survey on biometric cryptosystems and cancelable biometrics. *EURASIP Journal on Information Security*, 2011(1):1–25, 2011.

[91] Jane E Raymond, Kimron L Shapiro, and Karen M Arnell. Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18(3):849, 1992.

[92] Alejandro Riera, Aureli Soria-Frisch, Marco Caparrini, Carles Grau, and Giulio Ruffini. Unobtrusive Biometric System Based on Electroencephalogram Analysis. *EURASIP Journal on Advances in Signal Processing*, 2008:1–8, 2007.

[93] Bruno Rossion and Gilles Pourtois. Revisiting Snodgrass and Vanderwart's Object Pictorial Set: The Role of Surface Detail in Basic-Level Object Recognition. *Perception*, 33(2):217–236, 2004.

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

[95] Maria V Ruiz-Blondet, Zhanpeng Jin, and Sarah Laszlo. Permanence of the CEREBRE brain biometric protocol. *Pattern Recognition Letters*, 95:37–43, 2017.

[96] R Rutiku, M Martin, T Bachmann, and J Aru. Does the P300 reflect conscious perception or its consequences? *Neuroscience*, 298:180–189, 2015.

[97] Carolina Saavedra and Laurent Bougrain. Wavelet-based Semblance for P300 Single-trial Detection. In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing - Volume 1: BIOSIGNALS, (BIOSTEC 2013)*, pages 120–125. INSTICC, SciTePress, 2013.

[98] Takeshi Sakurada, Toshihiro Kawase, Tomoaki Komatsu, and Kenji Kansaku. Use of high-frequency visual stimuli above the critical flicker frequency in a SSVEP-based BMI. *Clinical Neurophysiology*, 126(10):1972–1978, 2015.

[99] Saeid Sanei and Jonathon A Chambers. *EEG Signal Processing*. John Wiley & Sons, 2013.

[100] Malte Sengelmann, Andreas K Engel, and Alexander Maye. Maximizing Information Transfer in SSVEP-Based Brain-Computer Interfaces. *IEEE Transactions on Biomedical Engineering*, 64(2):381–394, 2016.

[101] Cooper A Smout and Jason B Mattingley. Spatial Attention Enhances the Neural Representation of Invisible Signals Embedded in Noise. *Journal of cognitive neuroscience*, 30(8):1119–1129, 2018.

[102] Joan G Snodgrass and Mary Vanderwart. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2):174, 1980.

[103] Nancy K Squires, Kenneth C Squires, and Steven A Hillyard. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 38(4):387 – 401, 1975.

[104] Lionel Standing, Jerry Conezio, and Ralph Norman Haber. Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 19(2):73–74, 1970.

[105] Hans Strasburger, Ingo Rentschler, and Martin Jüttner. Peripheral vision and pattern recognition: A review. *Journal of vision*, 11(5):13–13, 2011.

[106] Fei Su, Huangling Zhou, Zhiyin Feng, and Junshui Ma. A biometric-based covert warning system using EEG. In *2012 5th IAPR International Conference on Biometrics (ICB)*, pages 342–347. IEEE, 2012.

[107] Mikhail Tokovarov, Monika Kaczorowska, and Małgorzata Plechawska-Wójcik. Towards Human Identification Based On SSVEP Response A Proof Of Concept Study. In *2017 International Conference on Electromagnetic Devices and Processes in Environment Protection with Seminar Applications of Superconductors (ELMECO & AoS)*, pages 1–4. IEEE, 2017.

[108] Sina Alexa Trautmann-Lengsfeld and Christoph Siegfried Herrmann. Virtually simulated social pressure influences early visual processing more in low compared to high autonomous participants. *Psychophysiology*, 51(2):124–135, 2014.

[109] Yoshiaki Tsushima, Yuka Sasaki, and Takeo Watanabe. Greater Disruption Due to Failure of Inhibitory Control on an Ambiguous Distractor. *Science*, 314(5806):1786–1788, 2006.

[110] Deirdre M Twomey, Peter R Murphy, Simon P Kelly, and Redmond G O'Connell. The classic P300 encodes a build-to-threshold decision variable. *European journal of neuroscience*, 42(1):1636–1643, 2015.

[111] Sven Uebelacker. *Investigations into Social Engineering Evidence for Security Research*. PhD thesis, 2023.

[112] Anthony Vance, Bonnie Brinton Anderson, C Brock Kirwan, and David Eargle. Using measures of risk perception to predict information security behavior: Insights from electroencephalography (EEG). *Journal of the Association for Information Systems*, 15(10):2, 2014.

[113] Min Wang, Hussein A Abbass, and Jiankun Hu. Continuous Authentication Using EEG and Face Images for Trusted Autonomous Systems. In *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pages 368–375. IEEE, 2016.

[114] Y. Wang and L. Najafizadeh. On the invariance of EEG-based signatures of individuality with application in biometric identification. In *International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4559–4562, 2016.

[115] Chun-Shu Wei, Masaki Nakanishi, Kuan-Jung Chiang, and Tzyy-Ping Jung. Exploring Human Variability in Steady-State Visual Evoked Potentials. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 474–479. IEEE, 2018.

[116] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.

[117] Dong-Ok Won, Han-Jeong Hwang, Sven Dähne, Klaus-Robert Müller, and Seong-Whan Lee. Effect of higher frequency on the classification of steady-state visual evoked potentials. *Journal of neural engineering*, 13(1):016014, 2015.

[118] Qunjian Wu, Bin Yan, Ying Zeng, Chi Zhang, and Li Tong. Anti-deception: Reliable EEG-based biometrics with real-time capability from the neural response of face rapid serial visual presentation. *Biomedical engineering online*, 17(1):1–16, 2018.

[119] Qunjian Wu, Ying Zeng, Zhimin Lin, Xiaojuan Wang, and Bin Yan. Real-time EEG-based Person Authentication System Using Face Rapid Serial Visual Presentation. In *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 564–567. IEEE, 2017.

[120] Su Yang and Farzin Deravi. On the Usability of Electroencephalographic Signals for Biometric Recognition: A Survey. *IEEE Transactions on Human-Machine Systems*, 47(6):958–969, 2017.

[121] Muwang Ye, Yong Lyu, Ben Sclodnick, and Hong-Jin Sun. The P3 Reflects Awareness and Can Be Modulated by Confidence. *Frontiers in neuroscience*, 13:510, 2019.

[122] Lee Guan Yeo, Haoqi Sun, Yisi Liu, Fitri Trapsilawati, Olga Sourina, Chun-Hsien Chen, Wolfgang Mueller-Wittig, and Wei Tech Ang. Mobile EEG-based situation awareness recognition for air traffic controllers. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3030–3035. IEEE, 2017.

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

[124] Ting Yu, Chun-Shu Wei, Kuan-Jung Chiang, Masaki Nakanishi, and Tzyy-Ping Jung. EEG-Based User Authentication Using a Convolutional Neural Network. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1011–1014. IEEE, 2019.

[125] Ying Zeng, Qunjian Wu, Kai Yang, Li Tong, Bin Yan, Jun Shu, and Dezhong Yao. EEG-Based Identity Authentication Framework Using Face Rapid Serial Visual Presentation with Optimized Channels. *Sensors*, 19(1):6, 2019.

[126] Martina Ziefle. Aging, visual performance and eyestrain in different screen technologies. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 45, pages 262–266. SAGE Publications Sage CA: Los Angeles, CA, 2001.