EEG Biometrics: On the Use of Occipital Cortex Based Features from Visual Evoked Potentials

Igor Barros Barbosa[1], Kenneth Vilhelmsen[2], Audrey van der Meer[2], Ruud van der Weel[2] and Theoharis Theoharis[1]

[1]Department of Computer and Information Science,
[2]Department of Psychology,
Norwegian University of Science and Technology

Abstract

The potential of using Electro-Encephalo-Gram (EEG) data as a biometric identifier is studied. This is the first study that assesses looming stimuli for the creation of biometrically useful Visual Evoked Potentials (VEP), i.e. EEG responses due to visual stimuli. A novel method for the detection of VEP responses with minimal expert interaction is introduced. The EEG data, segmented based on the VEP, are used to create a reliable feature vector. In contrast to previous studies, we provide a publicly available evaluation dataset based on infants which is therefore not biased due to unhealthy individuals. Only data from the occipital cortex are used (i.e. about 3 of the many possible electrode positions in the scalp), making the potential EEG biometric capture devices relatively simpler.

1 Introduction

The quest for reliable biometric identifiers has dominated biometrics research over the last decades. Spoofing is a major concern and in this respect many biometrics fail and have to resort to liveness testing, which further complicates matters. One biometric identifier that is potentially very hard to spoof is the activity of a person's brain. The question, of course, is how reliable a biometric identifier such brain data would be and it is this question that the present paper addresses by providing new results on an *unbiased* dataset of brain activity, which will be made publicly available.

Most work on biometric recognition from brain activity concentrates on Visual Evoked Potentials (VEPs) because of the relatively clear response and the fact that a publicly available dataset is available [3, 18]. However, this dataset is considered biased because the have been acquired from alcoholic individuals [5]. We addressed this issue by creating a new dataset based on electroencephalographic (EEG) recordings from infants' VEPs. Infants have considerably thinner skulls than adults and hardly any hair. This can explain why we are able to identify looming-related brain electrical responses on a trial-by-trial basis in the raw data from the EEG recordings in infants. Another aspect of our

This paper was presented at the NIK-2015 conference; see http://www.nik.no/.

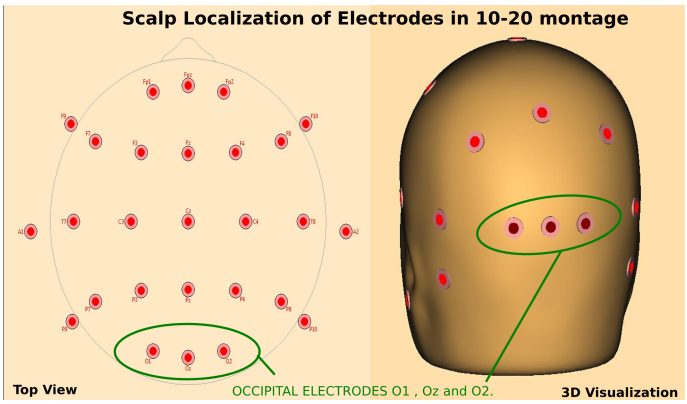

Figure 1: Head-drawing (nose up) and 3D head-model showing scalp localization of the standard 27 electrodes, with green circle showing electrodes O1, Oz, and O2 used in the current study. These electrodes capture EEG activity of the human visual cortex.

method is that it uses just 3 channels (O1, Oz, and O2 as shown in Fig. 1), since we know that VEPs reflect a response in the occipital area, elicited by visual stimuli; this would practically allow the acquisition device for such a biometric to be a far simpler version of the cumbersome tens of electrodes that are usually used to acquire a general brain EEG.

When a subject is exposed to specific external sensory stimulation, it is possible to register on EEG recordings an evoked response. Visual evoked potentials (VEPs) are electrical potentials, recorded by EEG, reflecting activity evoked by visual stimuli. In EEG research these VEPs are elicited by any type of visual stimuli, and are often found over visual areas (the occipital cortex). Looming stimuli in infants will elicit a negative component in the visual cortex.

In this work we show the potential usage of visual evoked responses in biometrics, without the averaging of EEG activity. More specifically, we show how visual evoked potentials (VEP) evoked by looming stimuli can be employed for biometrics.

Related Work

Biometric recognition based on brain electrical activity is a developing field of research. A recent review of existing work on the use of EEG for person recognition is [5]. As the review shows, a substantial part of the efforts to use EEG in biometrics focus on Event Related Potentials (ERP); VEP is the only ERP biometric analysed so far.

Most EEG-related biometrics research has been using the publicly available dataset of [3, 18], which we shall refer to as Zhang-DDBB. This dataset consists of 125 subjects, out of which 77 were alcoholics, and was not created for biometrics research. *Palaniappan and Raveendram* started investigating the use of EEG VEPs for biometrics in [10]. In this work 61 electrodes were used to record brain activity after a visual stimulus and it uses information on the gamma frequency band to classify a dataset of 10 subjects which are not made publicly available. The method achieves a classification performance of 90.95%. The same research group published in 2007 [9] a follow-up work where a Multiple Signal Classification Algorithm computes dominant frequencies of EEGs. The dominant frequencies are the features used in classification. This time, a nearest neighbor approach achieved 97.61% classification rate on 102 subjects from Zhang-DDBB. The work of [8] uses 8 electrodes for VEP analysis, creating a large feature vector by an ensemble of voice processing techniques. This large feature vector is reduced using a correlation-based feature selector and a Support Vector Machine (SVM) is used for

classification. In VEP analysis the work achieves two different classification results. The first classification performance is 92.80% for a dataset of 20 subjects from Zhang-DDBB and the second is 61.70% for 122 subjects from Zhang-DDBB.

As mentioned above, the problem with the Zhang-DDBB dataset is that it mainly consists of data from unhealthy (alcoholic) subjects. To quote [5], *'the problem with this practice is that alcohol has been proven to affect various aspects of the EEG (Begleiter & Porjesz, 2006; Zietsch et al., 2007), including the so much used alpha rhythm (Vogel, 2000). Therefore, it is likely that such datasets were biased, overrating the systems' performances. In order to clarify this, a statistical study probing that the parameters used are not affect by the disorder, i.e. assuring they are uncorrelated, must be adjoined.'*

The work of [4] uses Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM). VEP recordings are done by 64 electrodes placed over the scalp. The work collected 1000 VEP trials for 20 subjects (non-public dataset). Classification performance shows that SVM outperforms LDA for VEP analysis. On a two-fold SVM it achieves 91.56% classification accuracy. The works of [8, 4, 9, 10] place electrodes over the entire scalp. All the electrodes are then used for biometric recognition.

The above works assume that all recorded activity on the scalp is a result of the visual stimulus input. This is in contrast to current VEP and ERP literature [7, 2] which focuses on identified responses, e.g. the N2, P300, N170. The brain activation area will vary according to the stimulus. Therefore, proper analysis of ERP responses should be carried out by investigating clearly identified areas in the brain. The brain is divided into different regions. Each region processes a different stimulus input and is responsible for different ongoing activities [2]. Thus, averaging whole brain activation across multiple areas will most likely result in a mix of signals. The research of [4] shows that occipital [1] electrodes produce the most discriminative signals. This finding seems to enforce the methodological points of [7, 2].

The works of [13, 14] do promising EEG VEP analyses while focusing solely on the occipital cortex. Both methods employ checkerboard patterns to generate VEP responses. The first work, [13], has a classification performance of 86.54% on a private dataset composed of 13 subjects. Classification is done by a Linear Discriminant Classifier (LDC). The classifier uses as features well-known ERP responses (P100, N75). The work of [14] focuses on the occipital 'Oz' electrode. The work proposes the use of wave landmarks and a two dimensional Gaussian kernel classifier and achieves 78% classification accuracy on a private dataset composed of 10 subjects.

Contributions

The first contribution is to assess whether a new stimulus can be used to generated VEPs with biometric capabilities. In this work we present the first use of looming stimuli in a biometric study.

Second, a novel method for the detection VEP responses with minimal expert interaction is introduced. EEG data, segmented based on the VEP, are used to create a reliable VEP feature extractor.

Third, we contribute to reproducible research by making our dataset publicly available at `http://www.idi.ntnu.no/grupper/vis/eeg/`. Note that our dataset is not biased by individuals with any known issue; instead it is based on infant subjects who should respond to the stimuli as purely as possible. We thus expect that our dataset (and its

[1] The occipital lobe is the brain region where the visual cortex is located.

planned future extension to hundreds of subjects) could become a standard in evaluation of biometric work using the EEG.

The rest of the paper is organized as follows. Section 2 presents the looming stimulus. Section 3 discusses the pre-processing steps that are useful for EEG analysis. Section 4 presents the segmentation and feature extraction methods of the proposed method. Section 5 discusses the classifier selection. Experimental results are given in Section 6 and conclusions in Section 7.

2 Looming Visual Evoked Potentials

Looming stimuli are defined as approaching (virtual) objects on a collision course with the observer. Such stimuli can be artificially created by displaying an object looming towards the observer, thus allowing for control over the variables and making it easier for the data (psychophysical or behavioural) to be recorded.

Literature shows that a negative VEP component is prevalent in electrodes O1, Oz, and O2 in the visual cortex in relation to a looming stimulus [16]. Brain activity following impending collision of an approaching object can be seen in the visual cortex using EEG [17].

Infants presented with a looming stimulus display VEP components in occipital areas approximately 800 ms before loom hit [16].

In this work 16 infants were exposed to looming stimuli. The EEG recordings are then processed for VEP segmentation and feature extraction. For more details on the dataset see Section 6.

3 Processing of Raw EEG

This Section will present magnitude frequency responses and the reasoning behind the selection of preprocessing filters for EEG recordings.

Digital EEG recordings are preprocessed and filtered differently depending on acquisition procedures and purpose. The main ambition of this work is to assess the biometric capabilities of VEPs. Thus, signal preprocessing needs to keep VEPs intact while removing noise and unwanted data.

The first issue to address is signal corruption caused by utility/mains frequency. It is common for EEG recordings to be dominated by the frequency of AC currents, F_{AC}, which is either $50Hz$ or $60Hz$ depending on region. This signal corruption is visible in the unprocessed EEG readings of Fig. 2. Therefore, the filtering of F_{AC} and its harmonics is vital for subsequent data analysis.

Another issue is artifacts generated by bio-electric flowing potentials [15]. Small body movements, breathing and other similar behaviours can distort EEGs. These distortions create an amplitude modulation in low frequencies. As a consequence, a high pass filter (HPF) with a small cutoff frequency becomes essential.

Literature shows that VEPs have a concise spectrum [17, 16, 1]. This means that information related to the signal is not spread over all frequencies. One would ideally like to filter out frequencies carrying unwanted data. Thus, a low pass filter (LPF) must be employed in addition to the HPF. This filter needs to have a well selected cutoff frequency so as not to impair VEPs.

To further specify filters and respective cutoff frequencies, the present work assessed VEP characteristics. Literature shows [17, 1, 16] that looming VEPs are found between 1.8Hz and 60Hz. To further restrain operational frequencies, we computed the time

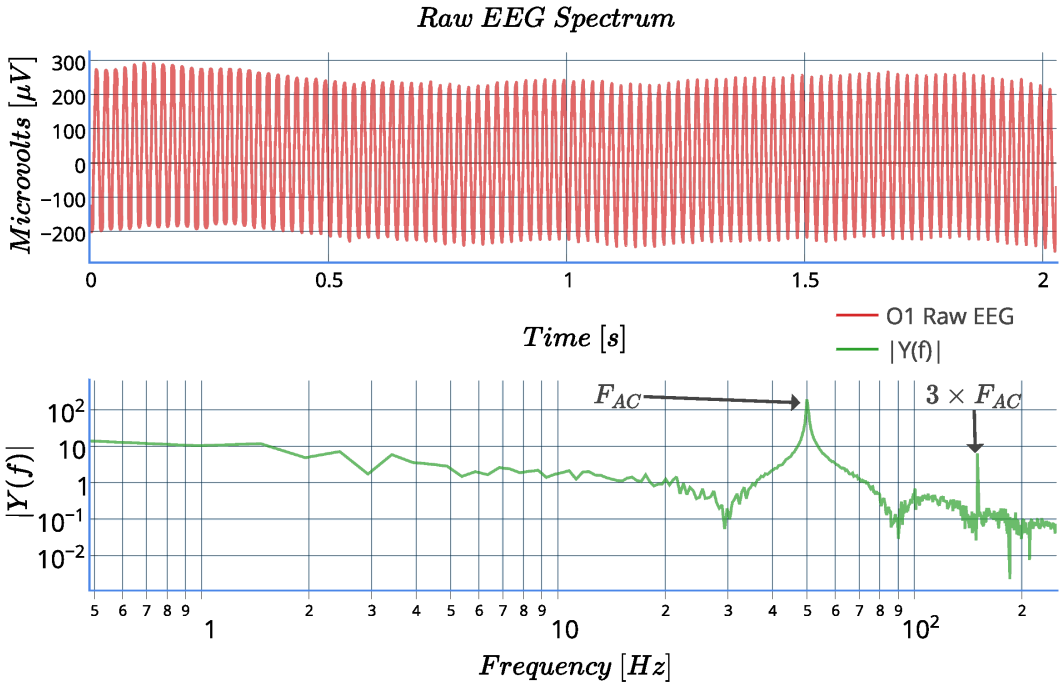

Figure 2: A sample of Raw EEG readings for Electrode 'O1' and its spectrum analysis. The EEG readings are dominated by F_{AC} as shown in both the time and frequency domains.

between peaks and valleys on typical VEP responses of [17]. This showed that VEP frequencies do not exceed 20Hz. Consequently, all harmonic contributions of F_{AC} would be filtered out by a LPF at 20Hz, implying no need for any further filters. However, given the high noise amplitude at 50Hz, a filter at F_{AC} is also needed, to minimise the effect of noise on VEP analysis.

As per the above analysis, three filters are implemented. The three filters are designed as Infinite Impulse Response (IIR), allowing us to create digital counterparts of well established analogs filters [12]. The filter design starts with a first order Butterworth HPF with a cutoff frequency of 1.6Hz; this is a safe cutoff frequency as it is below the $1.8Hz$ lowest VEP frequency reported in the literature. The second filter used is a fourth order Butterworth LPF. This time, the selected cutoff frequency is 20Hz as estimated above. The third filter is a sixth order notch filter. This is a band rejection Butterworth. The rejected band ranges from $45Hz$ to $55Hz$. The magnitude responses of the three designed filters are shown in Fig. 3.

The resulting filtered EEG as well as its spectral analysis are shown in Fig. 4. The processed signal is used for feature extraction, as demonstrated in Section 4.

4 Segmentation and Feature Extraction

Research shows that looming stimuli create distinctive VEP responses. An example of such responses is shown in Fig. 5, which shows VEP responses on 'O1', 'Oz' and 'O2'.

Unfortunately, VEP responses do not always take place consistently. Their timing occurrence varies even within subjects. It is even common for a subject to produce multiple VEPs for a single looming stimulus. We would ideally like to segment out VEP responses before feature extraction, in order to increase the reliability of the latter. In this

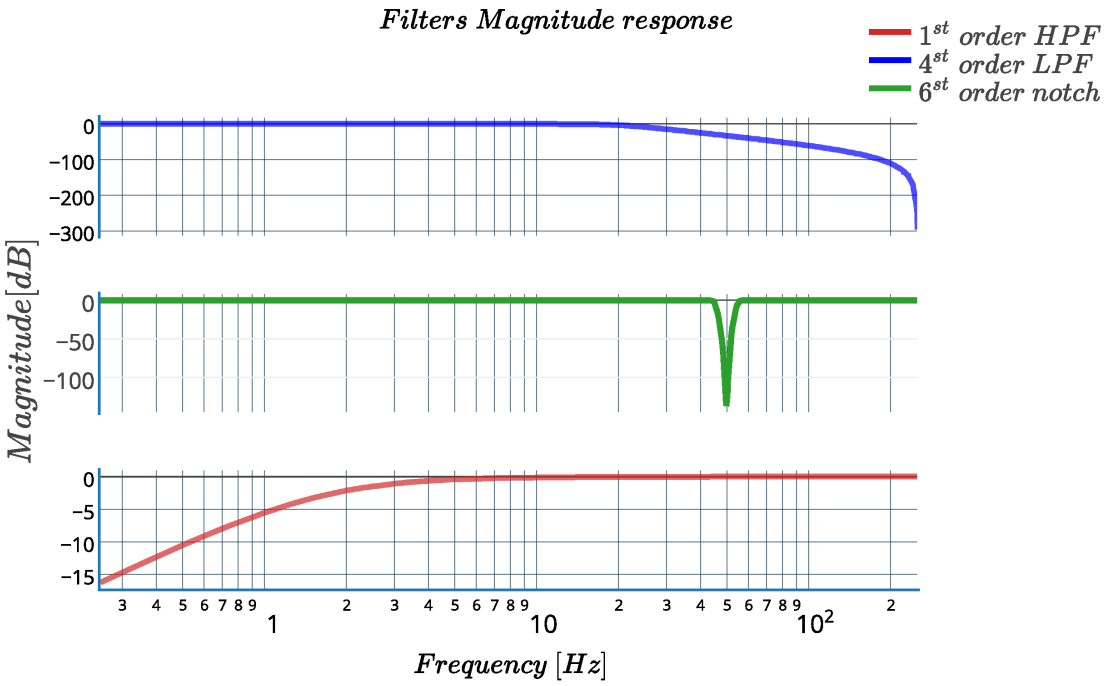

Figure 3: A combination of three filters is used to filter out noise and artifacts from Raw EEG recordings. The magnitude responses of such filters are shown here.

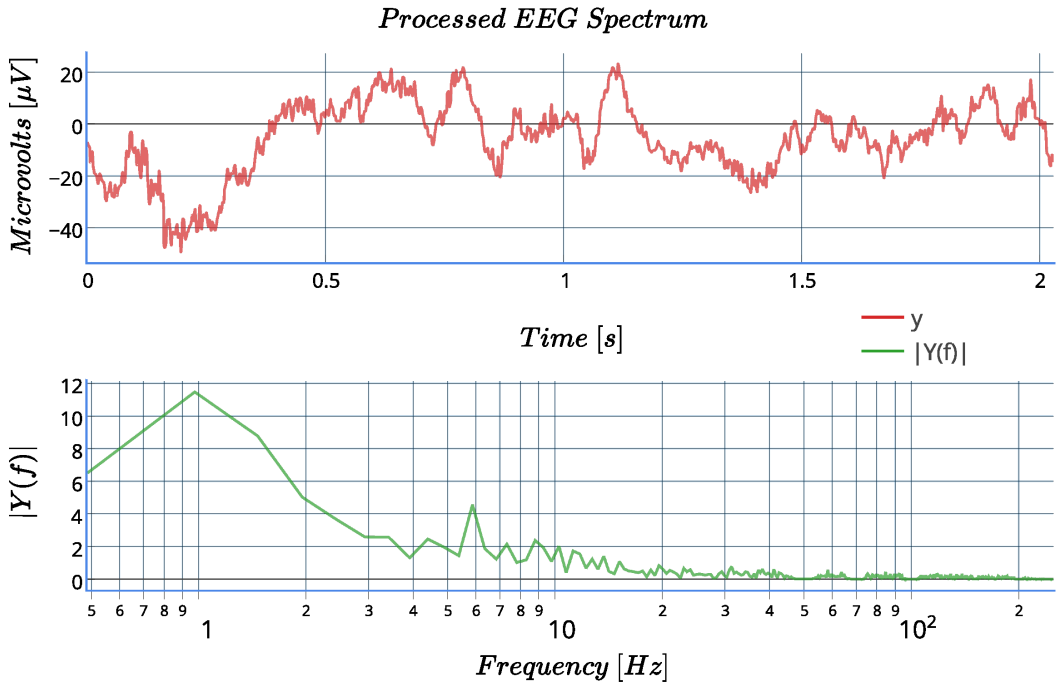

Figure 4: A sample of filtered EEG data for Electrode 'O1' and its spectrum analysis, after the application of three filters to remove noise and artifacts from the raw signal.

section we present a reliable VEP segmentation technique followed by an inexpensive and dependable feature extraction method.

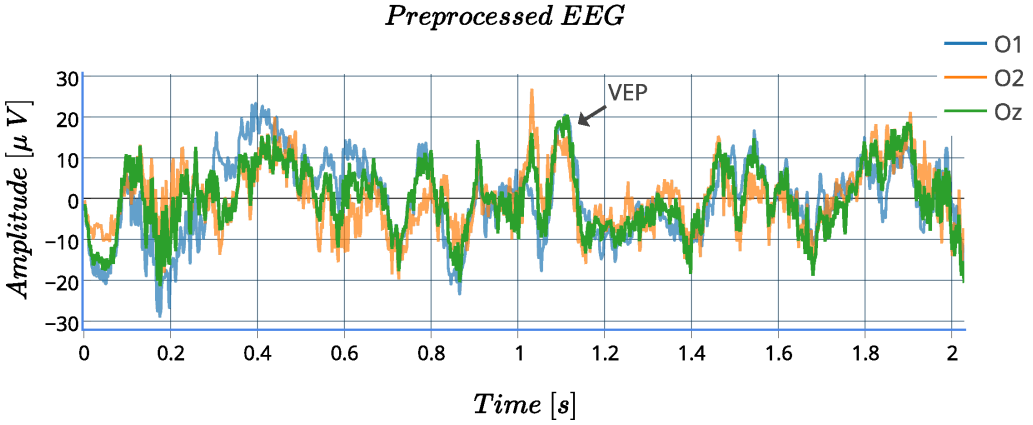

Figure 5: A sample of EEG looming VEP. The graph show readings for electrodes 'O1', 'Oz', and 'O2'. These are the electrodes over the occipital cortex.

VEP Segmentation

The potential segmentation algorithm should first list VEP candidates. This is arduous because a single trial can have multiple VEP responses, or none at all, for a given stimulus.

Our detection technique explores characteristics of typical VEP responses. As one can notice in Fig. 5, sharp transitions delineate VEPs. Thus, we can select potential candidates by searching for high frequencies. For this we propose a uni-dimensional edge detector to detect the sharp transitions. This edge detector is instantiated by Eq. 1:

$$\mathcal{E}(x) = \left[\frac{\delta}{\delta x_{ma}}\Phi_{ma}(x, 30)\right]^3 \tag{1}$$

where $\Phi_{ma}(x, N)$ is a moving average of the input EEG signal x with N elements and x_{ma} is the output of the moving average function. Edge detectors are sensitive to noise in input signals. Thus, we first apply the moving average filter Φ_{ma} to the input. This above edge detector can be approximated in the discrete case by the difference Eq. 2:

$$\mathcal{E}[k] = (x_{ma}[k+1] - x_{ma}[k])^3 \tag{2}$$

In this discrete version, the computed edge $\mathcal{E}$ is one element smaller than x. This small difference can be ignored in the following steps. We next compute a search vector $\mathcal{S}$ which will allow us to circumvent issues common to looming stimuli. An initial $\mathcal{S}$ is computed by Eq.3:

$$\mathcal{S}(\mathcal{E}_{O1} + \mathcal{E}_{O2} + \mathcal{E}_{Oz}) = \mathcal{G}\left(\left|\frac{\mathcal{E}_{O1} + \mathcal{E}_{O2} + \mathcal{E}_{Oz}}{3}\right|, \sigma, N\right) \tag{3}$$

where $\mathcal{E}$ represents a computed edge in the electrode given in the respective index. $\mathcal{G}(x, \sigma, N)$ represents the convolution of a Gaussian filter with the signal x with standard deviation σ and N represents the size of the convolution filter. For this work a tentative $sigma = 5$ and $N = 15$ were employed.

No looming related activity happens in the initial 500ms. We thus replace the first half second of the search vector by its average value. Finally the search vector $\mathcal{S}$ is normalized to the interval $[0, 1]$. An example of a resulting search vector is shown in Fig. 6.

We can now look for VEP candidates by searching for peaks (local maxima) in $\mathcal{S}$. Peak candidates need to have a minimal distance of $200ms$ to each other and we only

accept peaks higher than 0.95; if these criteria are fulfilled the VEP candidates are treated as VEP indicators.

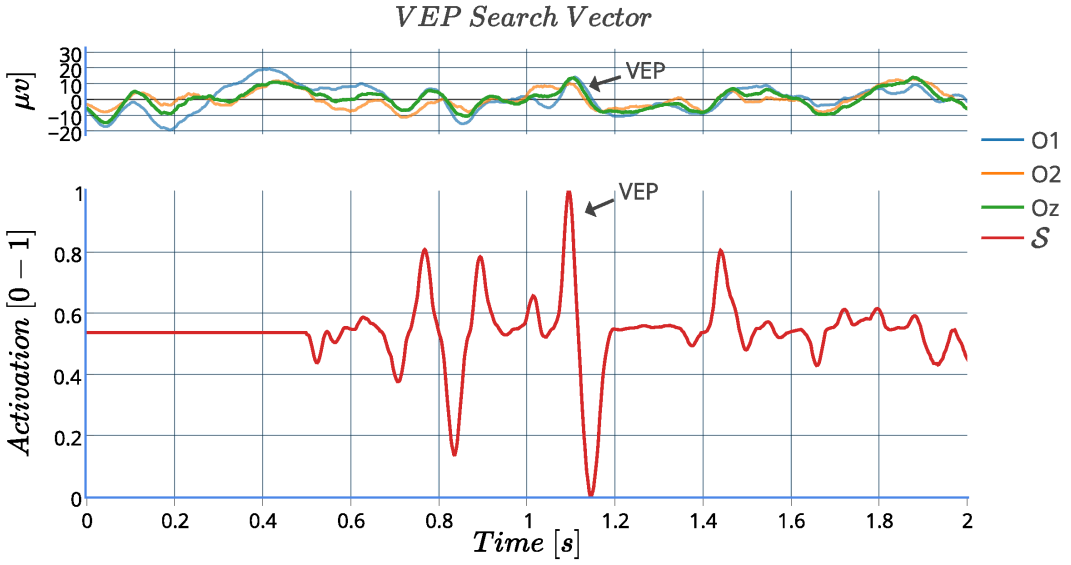

Figure 6: The first graph shows the EEG readings. These readings were processed with a moving average filter. The second graph shows the search vector. In $\mathcal{S}$ the peaks indicate VEP candidates.

With a list of VEP indicators it is possible to segment and extract features. Segmentation extracts a 200ms long signal sample. The segmented signal starts 80ms before the VEP indicator and ends 120ms after it. Thus, all information from the VEP is represented in the segmented signal. This segmented data is next used for feature extraction.

Feature Extraction

For feature extraction we create a signal 400ms long. Two slices of 200ms compose this signal. The first slice is the preprocessed EEG recording and the second is the computed edge detector $\mathcal{E}$. For concise results we normalise to the range $[0, 1]$ for each 200ms slice independently. Fig. 7 shows the two slices that compose the feature vector.

Using the above signals, the feature extraction process is computationally inexpensive, while it concisely inherits information on VEPs. The result is an efficient and robust feature extractor.

5 Matching / Classification

For classification we use a simple feed-forward Neural Network. A one-hidden-layer Multi Layer Perceptron (MLP) is trained with the standard back-propagation algorithm [6]. This work employs stochastic gradient descent with mini batches for training, ensuring a faster training time. The MLP uses the sigmoid function for neuron activations and soft-max for the output neurons.

The EEG recording apparatus had a sampling frequency of 500Hz. Therefore, the 400ms feature vector is in fact a 200 dimension vector. We thus create a neural network with 200 input neurons. These neurons are fully connected to a hidden layer with 15 neurons. These 15 neurons are connect to the output layer. The number of output neurons

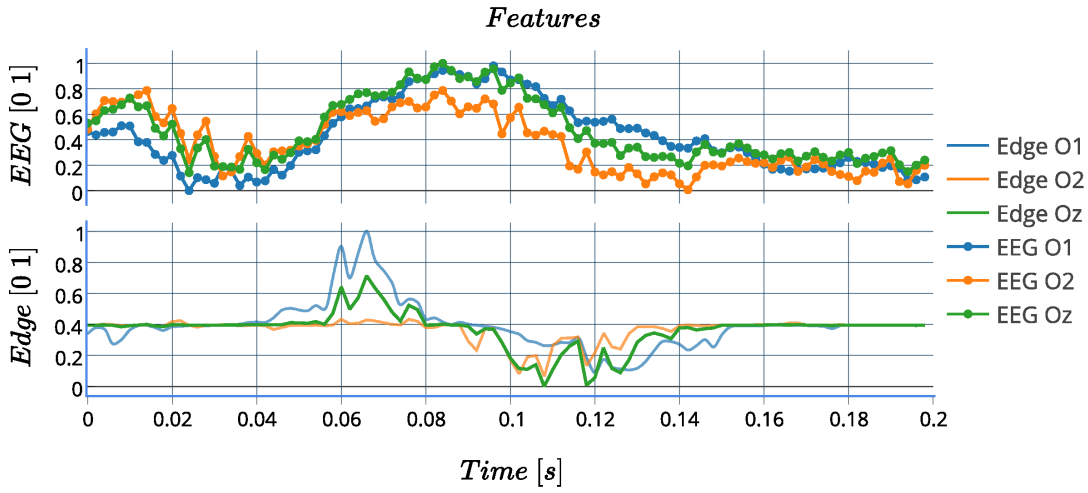

Figure 7: The first graph shows normalized EEG readings; these 200ms signals are the result of the proposed VEP detection and segmentation. The second graph shows edge features computed over the EEG signals. Together these two 200ms slices compose a Feature Vector. The Feature Vector can be computed for any electrode. Here we use electrodes over the visual cortex: 'O1','Oz', and 'Oz'.

is the same as the number of subjects in the dataset. Thus, for this experiment we use 16 output neurons. The MLP was implemented with the help of a publicly available library [11].

As a single subject generates many feature vectors, a final classification is done by majority voting the MLP outputs.

6 Experimental Results

This Section presents the experimental results as well as details of the acquired dataset.

Dataset and Acquisition Procedure

The data were initially acquired from 16 infants. The infants would arrive with their parent(s) and be familiarised with the surroundings. Meanwhile, the parent(s) would be informed about the experiment and signed a consent form. The EEG net was prepared in electrolyte solution to ensure good impedance; the recording equipment was prepared before the participant arrived. The net would be applied and the infants would be seated in front of the screen; a parent would always be with the infant in the test room. The experiment would be aborted if an infant became too fussy or started crying.

Infants were presented with a black approaching rotating circle consisting of four smaller circles $\frac{1}{3}$ of the size (red, green, blue, yellow) rotating within the black circle on a white background. The loom rotated with a constant angular velocity of $300°/s$ and started at a virtual distance of 43.1 m giving a visual angle of 5° (6cm diameter); it then grew to a maximum size of 131° (350cm diameter). The loom would approach the infants at constant accelerations at three different speeds with a duration of 2 seconds ($-21.1\frac{m}{s^2}$), 3 seconds ($-9.4\frac{m}{s^2}$), and 4 seconds ($-5.3\frac{m}{s^2}$). The looming object would move the same distance in all three cases, as well as in reverse.

The order of the presented stimuli was randomly generated. The only constraint was that consecutive stimuli should be different. There was a break of 1.5 seconds between

two stimuli. In this work looming with two second durations are used for classification. We only employ information from the occipital electrodes 'O1', 'O2' and 'Oz'.

In the acquired dataset we have a varying number of looming stimuli with 2 second duration. Some infants were presented up to 20 stimuli and some just 7. The number of trials depended on the willingness of each baby.

Classification Results

The pipeline described in this paper is followed, i.e. pre-processing the EEG recordings as described in Section 3, segmentation of the information related to looming with a 2 second duration and feature extraction as described in Section 4. The Feature Vector is then used in the MLP Voting classifier described in Section 5 to obtain the classification results.

We use half of the available data for training, and the other half for testing, obtaining a recognition (classification) accuracy of 62.50%. This result is comparable to the state of the art occipital VEP EEG processing [13, 14]. Different from these previous works, the dataset assessed in this paper has more subjects and is composed of a much smaller number of trials, making it a harder evaluation scenario. This explains why the evaluation performance achieved here is smaller than [13, 14], which used private datasets. Table 1 shows classification performance as well as details of the assessed datasets.

Method	**Subjects**	**VEP trials**	**Acc'cy**	**N- Folds**
Proposed	16	7 to 20	62.50%	2
[13]	13	120	86.54%	4
[14]	10	200	78.00%	2

Table 1: Classification accuracy

7 Conclusion

The first attempt to assess whether looming stimuli can be used to generate occipital Visual Evoked Potential (VEP) in EEG signals with biometric capabilities, is presented. At the same time the first public dataset of EEG responses to looming stimuli is being made available. This dataset is based on infant subjects that should respond to the stimuli as purely as possible. We expect that this dataset and its planned extensions can become a standard for biometric studies of VEP.

We show that our VEP segmentation methodology, as well as our proposed feature extraction, can produce biometrically capable EEG based features. The achieved classification performance is comparable to the state of the art biometric studies using occipital VEPs. However, these results were achieved using a dataset with more subjects and with fewer samples (VEP trials) per subject.

Further work could address automatically learned features (e.g. with a deep learning approach), but this would require at least two orders of magnitude more data. Another possibility is to investigate other machine learning techniques capable of learning while using a small numper of samples (VEP Trials).

References

[1] S. B. Agyei, M. Holth, F. R. van der Weel, and A. L. H. van der Meer. Longitudinal study of perception of structured optic flow and random visual motion in infants using high-density EEG. *Developmental Science*, 2014.

[2] M. F. Bear, B. Connors, and M. Paradiso. *Neuroscience: Exploring the Brain*. Lippincott Williams & Wilkins, fourth international edition, March 2015.

[3] H. Begleiter. Eeg database data set. `http://archive.ics.uci.edu/ml/datasets/EEG+Database` . Accessed: 2015-03-23.

[4] K. Das, S. Zhang, B. Giesbrecht, and M. P. Eckstein. Using rapid visually evoked EEG activity for person identification. *IEEE Engineering in Medicine and Biology Society. Conference*, 2009:2490–2493, 2009.

[5] M. Del Pozo-Banos, J. B. Alonso, J. R. Ticay-Rivas, and C. M. Travieso. Electroencephalogram subject identification: A review. *Expert Systems with Applications*, 41:6537–6554, 2014.

[6] Y. Le Cun, D. Touresky, G. Hinton, and T. Sejnowski. A theoretical framework for back-propagation. In *The Connectionist Models Summer School*, volume 1, pages 21–28, 1988.

[7] S. J. Luck and E. S. Kappenman, editors. *The Oxford Handbook of Event-Related Potential Components (Oxford Library of Psychology)*. Oxford University Press, 1 edition, December 2011.

[8] P. Nguyen, D. Tran, X. Huang, and D. Sharma. A Proposed Feature Extraction Method for EEG-based Person Identification. In *The International Conference on Artificial Intelligence*, 2012.

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

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

[11] R. B. Palm. Prediction as a candidate for learning deep hierarchical models of data. Master's thesis, 2012.

[12] T. W. Parks and C. S. Burrus. *Digital Filter Design*. Wiley-Interscience, New York, NY, USA, 1987.

[13] A. Power, E. Lalor, and R. Reilly. Can visual evoked potentials be used in biometric identification? In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 5575–5578, Aug 2006.

[14] G. K. Singhal and P. Ramkumar. Person identification using evoked potentials and peak matching. *2007 Biometrics Symposium, BSYM*, 2007.

[15] M. Teplan. Fundamentals of eeg measurement. *Measurement Science Review*, 2(2):1–11, 2002.

[16] A. L. H. van der Meer, M. Svantesson, and F. R. van der Weel. Longitudinal study of looming in infants with high-density EEG. *Developmental Neuroscience*, 34(6):488–501, 2012.

[17] F. R. van der Weel and A. L. H. van der Meer. Seeing it coming: infants brain responses to looming danger. *Naturwissenschaften*, 96(12):1385–1391, 2009.

[18] X. L. Zhang, H. Begleiter, B. Porjesz, W. Wang, and A. Litke. Event related potentials during object recognition tasks. *Brain Research Bulletin*, 38:531–538, 1995.