

Gaze estimation using EEG signals for HCI in augmented and virtual reality headsets

Juan Manuel Fernandez Montenegro and Vasileios Argyriou
Kingston University, London

J.Fernandez@kingston.ac.uk, Vasileios.Argyriou@kingston.ac.uk

Abstract

Augmented and virtual reality have evolved significantly over the last few years providing new ways of entertainment and interaction with the environment. Although many systems and solutions are currently available, still there is much left unsettled and some technologies are missing from many VR/AR devices, such as foveated rendering and HCI. In this paper, a novel approach for coarse gaze estimation using EEG sensors with applications in items selection for HCI or foveated rendering for VR/AR devices is proposed. The suggested method requires only few electroencephalogram sensors that can be easily added to the current virtual and augmented reality headsets. A supervised machine learning approach was suggested utilising novel features, based on quaternions allowing gaze estimation. Experiments were performed to evaluate the proposed method and a new dataset was designed and captured. Finally, the introduced learning framework was compared with other similar techniques demonstrating further the gain of the proposed descriptors.

1. Introduction

Eye movements are essential in order to gather information and move through the visual world. They are linked to personality and reflect cognitive processes and human emotional states. Eye-gaze tracking (EGT) is the process of measuring either the point of gaze or the eye motion in relation to the head. Nowadays, there are many applications of eye and gaze tracking, including human-computer interaction (HCI), supporting disabled people, alerting drivers, diagnosing visual disorders, marketing and the understanding of human mental state. Regarding the applications in Virtual and Augmented Reality (VR/AR) eye tracking is essential for HCI (menu or item selection) and foveated rendering, which can improve significantly both the performance in terms of rendering speed and quality of experience. This topic of research is separated into two main areas:

eye detection and gaze tracking. Eye detection tries to locate the eyes of a human shown in an image or a video sequence, whereas gaze tracking estimates where a person is looking in 3D space. Gaze is tracked using a device that analyses eye movements (eye tracker), estimating the eyes position and tracking their movement over time to determine the 3D line of sight.

There are different methods for eye detection and gaze tracking that can be classified into two categories: sensor and computer vision based techniques [2, 7]. Sensor based approaches measure electric potentials utilising the electrooculogram (EOG) or the electroencephalogram (EEG). The eye acts as a dipole, considering the cornea as the positive pole and the retina as the negative one. About the EOG sensors, they are located around the eyes, with the electric potential field to be steady when the eye is at its normal state. If the retina approaches one sensor and the cornea the opposite one, a change in the EOG signal is produced that is used to track the eye movements. EEG sensors are placed on the head scalp and they record brain signals and artifacts, such as EOG and other muscle movements. Most EEG studies that are not focused on gaze tracking, try to remove these artifacts, in order to work only with brain signals. On the other hand, current methods for eye tracking use EEG signals and their EOG artifacts to detect saccade and pursuit movements. These methods use techniques such as Independent Component Analysis (ICA) to remove artifacts, extract certain features, (e.g. amplitude, spectral power) and use them with classifiers as k-Nearest Neighbour (kNN), [23, 16, 17, 20, 15, 26].

Computer vision approaches (video oculography, VOG) use cameras, in order to detect and track eyes over time. Using the obtained information about the eyes' location and the head pose, the gaze direction is estimated. Eye detection is influenced by the available eye model, the illumination conditions, viewing angle and several other parameters, [9]. Other techniques are based on their geometric and photometric properties:

shape-based, feature-based, appearance-based and hybrid. Some of these techniques employ active infrared (IR) illumination, since it improves pupil detection [21], with gaze tracking techniques to relate image data and gaze direction. Additionally, other eye movements that can be detected are the fixations and saccades. These techniques require head and pupil position estimation in order to track gaze accurately. Most of these approaches require hardware configurations to obtain head pose invariance with most of them to be feature or appearance based, [10, 24, 25, 19, 22, 18, 12].

Regarding the above video-oculography methods, it can be argued that they are not suitable for VR/AR due to the size and the amount of the acquisition devices required to detect eyes and gaze position. The need of cameras and light sources limits their applicability on AR/VR headsets. Therefore, in this work, an approach that uses the EEG signals to detect a coarse gaze direction is proposed. The proposed method is focused only on the 4 frontal sensors, since they are affected the most by EOG artifacts. Also, our aim is the reduction of the amount of sensors, in order to improve the usability of virtual reality (VR) and augmented reality (AR) devices. Additionally, the coarse gaze direction estimation can improve significantly the overall HCI in AR/VR devices allowing interaction with the eyes especially for item or menu interactions. The proposed approach is based on supervised learning techniques introducing novel feature descriptors based on a simplified quaternion representation. In the proposed descriptor, the data provided by the 4 frontal EEG sensors are combined into a single quaternion, and used as input to machine learning techniques trying to identify the areas where the eyes are focused on. The remainder of this paper is organized as follows: Section 2 analyses the proposed methodology and in section 3 details on the evaluation process and the obtained results are presented. Finally, conclusion remarks are given.

2 Methodology

The proposed method for gaze estimation is based on a supervised learning framework that incorporates a training and a testing stage, (see figure 1). This section analyses the suggested methodology, and provides details for the data acquisition process. Also, all the pre-processing steps are analysed and the proposed novel quaternion based descriptor is presented.

2.1 Data collection

Initially, raw data was collected from EEG sensors placed on the scalp of the participants. During the acquisition process, the subjects were looking at different locations on a monitor in front of them. The Tobii Eye tracker was utilised to record both gaze information and

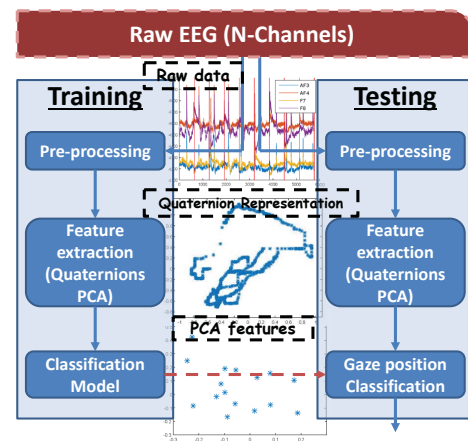


Figure 1. The proposed approach.

the EEG data synchronized using time-stamps. Regarding the setup, the subjects were seated in front of an eye tracker and a PC monitor, while wearing the EEG headset. The EEG device returns the amplified signal of the 14 sensors/channels, a value that indicates the quality of the signal, the gyroscope measurements, the battery level and the time when the capturing process started. The eye-tracker provides information that includes the coordinates of the eyes, fixations, saccades, areas of interest and time-stamps. The information about the location of the eyes is provided both in pixel and millimeter coordinates.

2.2 Pre-processing

Since the data is available, the obtained values are normalised for each channel in order to reduce the signal differences between the subjects. The average value is subtracted and the result is divided by their standard deviation. In addition, a median filter is used to reduce the noise produced by the electronic amplifier, the power line interference and any other external interferences.

2.3 Feature extraction for gaze classification

The training stage is oriented towards the extraction of features from the preprocessed data of the EEG channels to create a classification model. The selected EEG channels are combined into a novel quaternion representation and PCA is used to reduce the dimensionality of the obtained feature vector.

Quaternion Principal Component Analysis

It is true that a vector can be decomposed in linearly independent components, in the sense that they can be combined linearly to reconstruct the original vector. However, depending on the phenomenon that changes the vector, correlation between the components may exist from the statistical point of view (i.e. two uncorrelated variables are linearly independent but two linearly independent variables are not uncorrelated). If they are

independent our proposed descriptor does not provide any significant advantage, but if there is correlation this is considered. In most of the cases during the feature extraction process complex or hyper-complex features are generated but decomposed to be computed by a classifier. For example, normals and gradients in 2D/3D are features consisted of more than one element and this decomposition can imply a loss of information.

To do so, vectorial features can be represented more precisely using a complex or hyper-complex representation [1, 14]. Since, in our case and many similar scenarios, vectorial features such as a location, speed, gradients or angles, are the primary source of information, a hyper-complex representation of these features is more efficient allowing better correlation between these channels [1, 14, 5]. The proposed method exploits the hyper-complex (quaternion) representation capturing the dependencies within the EEG sensors located on the sides of the head and the ones over the eyes, [13, 4].

In order to reduce the number of the selected hyper-complex features without increasing the complexity, quaternion PCA is applied. In more details, the quaternion representation was introduced in [3, 6] as a generalization of the complex numbers. A quaternion $q \in \mathcal{H}$ has four components:

$$q = q_r + q_i i + q_j j + q_k k \quad (1)$$

where $q_r, q_i, q_j, q_k \in \mathbb{R}$ and i, j , and k satisfy

$$\begin{aligned} i^2 &= j^2 = k^2 = -1, & ij &= -ji = k \\ jk &= -kj = i, & ki &= -ik = j \end{aligned} \quad (2)$$

Conjugation of quaternions denoted by H is analogous to conjugation of complex numbers elements and is defined as:

$$q^H = q_r - q_i i - q_j j - q_k k. \quad (3)$$

The square of the norm of a quaternion is defined as

$$\|q\|^2 = q_r^2 + q_i^2 + q_j^2 + q_k^2 = q^H q. \quad (4)$$

with $(q_1 q_2)^H = q_2^H q_1^H$ and the four components (q_r, q_i, q_j, q_k) to correspond to the available four EEG channels (see figure 2).

Let quaternion column vector $\mathbf{q} = [q_1, \dots, q_F]^T \in \mathcal{H}^F$ where T denotes simple transposition be the EEG values over time. The conjugate transpose of vector \mathbf{q} is denoted by \mathbf{q}^H . There is an isomorphy between a quaternion and a complex 2×2 matrix defined as

$$\mathbf{Q} = \begin{bmatrix} q_r + q_i i & q_j + q_k i \\ -q_j + q_k i & q_r - q_i i \end{bmatrix} \quad (5)$$

Let \mathbf{x}_l be the F -dimensional vector obtained by writing in lexico graphic ordering and form $\mathbf{X} =$

$[\mathbf{x}_1 | \dots | \mathbf{x}_N] \in \mathcal{H}^{F \times N}$. Also we denote by $\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$ and $\bar{\mathbf{X}}$ the sample mean and the centralized sample matrix \mathbf{X} , respectively. A projection vector is denoted by $\mathbf{u} \in \mathcal{H}^F$ and by $y_i = \mathbf{u}^H \mathbf{x}_i$ the projection of \mathbf{x}_i onto \mathbf{u} . We want to maximize the (sum of the) variances of the data assigned to a particular class (gaze location)

$$\begin{aligned} E(\mathbf{u}) &= \sum_{l=1}^N \|y_l - \bar{m}\|^2 = \sum_{l=1}^N \|\mathbf{u}^H (\mathbf{x}_l - \mathbf{m})\|^2 \\ &= \mathbf{u}^H \sum_{l=1}^N (\mathbf{x}_l - \mathbf{m})(\mathbf{x}_l - \mathbf{m})^H \mathbf{u} \\ &= \mathbf{u}^H \mathbf{S} \mathbf{u} \end{aligned} \quad (6)$$

where $\mathbf{S} = \bar{\mathbf{X}} \bar{\mathbf{X}}^H$. It can be easily proven that matrix \mathbf{S} is a quaternion Hermitian matrix i.e., $S_{ij} = S_{ji}^H$.

In order to find K projections $\mathbf{U} = [\mathbf{u}_1 | \dots | \mathbf{u}_K] \in \mathcal{H}^{F \times K}$ we may generalize $E(\mathbf{U})$:

$$\begin{aligned} \mathbf{U}_o &= \arg \max_{\mathbf{U} \in \mathbb{H}^{F \times p}} E(\mathbf{U}) \\ &= \arg \max_{\mathbf{U} \in \mathbb{H}^{F \times p}} \text{tr}[\mathbf{U}^H \mathbf{S} \mathbf{U}] \\ \text{s.t. } &\mathbf{U}^H \mathbf{U} = \mathbf{I}. \end{aligned} \quad (7)$$

We aim at solving the above noted problem by using the isomorphic complex form that can be reformulated as

$$\begin{aligned} \tilde{\mathbf{U}}_o &= \arg \max_{\tilde{\mathbf{U}}} \text{tr}[\tilde{\mathbf{U}}^H \tilde{\mathbf{S}} \tilde{\mathbf{U}}] \\ \text{s.t. } &\tilde{\mathbf{U}}^H \tilde{\mathbf{U}} = \mathbf{I}. \end{aligned} \quad (8)$$

Since \mathbf{S} is a quaternion Hermitian matrix, $\tilde{\mathbf{S}}$ is a complex Hermitian. Also, given that $\tilde{\mathbf{S}}$ is a positive semidefinite Hermitian matrix (i.e., it has only non-negative eigenvalues) the solution $\tilde{\mathbf{U}}_0$ is given by the p eigenvectors of $\tilde{\mathbf{S}}$ that correspond to p largest eigenvalues. We want an efficient algorithm for performing eigenanalysis to $\tilde{\mathbf{S}}$, which is a complex $2F \times 2F$ matrix and can be written as $\tilde{\mathbf{S}} = \tilde{\mathbf{X}} \tilde{\mathbf{X}}^H$ where $\tilde{\mathbf{X}} \in \mathbb{C}^{2n \times F}$ and needs $O((2F)^3)$ time.

In general, given a quaternion Hermitian matrix \mathbf{A} then it has n nonnegative real eigenvalues (due to the non-commutative multiplication property of quaternions, there exists two kinds of its eigenvalue; in this paper we are interested only on the left eigenvalues) $\mathbf{l} = [\sigma_1, \dots, \sigma_n]$. Let $\tilde{\mathbf{A}}$ be its complex form

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_r + i\mathbf{A}_i & \mathbf{A}_j + i\mathbf{A}_k \\ -\mathbf{A}_j + i\mathbf{A}_k & \mathbf{A}_r - i\mathbf{A}_i \end{bmatrix}$$

then the eigenvalues of $\mathbf{l}_{2n} = [\sigma_1, \sigma_1, \dots, \sigma_n, \sigma_n]$. Representing $\mathbf{A} = \mathbf{B} \mathbf{B}^H$, where \mathbf{B} is a quaternion matrix, and considering $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ to be the complex forms of matrices \mathbf{A} and \mathbf{B} , respectively, then, $\tilde{\mathbf{A}}$ will be given by $\tilde{\mathbf{A}} = \tilde{\mathbf{B}} \tilde{\mathbf{B}}^H$. So, based on this analysis, we can write $\tilde{\mathbf{S}} = \tilde{\mathbf{X}} \tilde{\mathbf{X}}^H$. Also by defining matrices \mathbf{A} and \mathbf{B} such that $\mathbf{A} = \mathbf{\Gamma} \mathbf{\Gamma}^H$ and $\mathbf{B} = \mathbf{\Gamma}^H \mathbf{\Gamma}$ with $\mathbf{\Gamma} \in \mathbb{C}^{m \times r}$, and considering \mathbf{U}_A and \mathbf{U}_B to be the eigenvectors corresponding to the non-zero eigenvalues $\mathbf{\Lambda}_A$ and $\mathbf{\Lambda}_B$ of \mathbf{A}

and \mathbf{B} , respectively, we finally obtain $\mathbf{\Lambda}_A = \mathbf{\Lambda}_B$ and $\mathbf{U}_A = \mathbf{\Gamma} \mathbf{U}_B \mathbf{\Lambda}_A^{-\frac{1}{2}}$.

Thus according to the above, in a classification problem, such as the current one for gaze location, we may represent the quaternion Hermitian matrix (descriptor) providing a subspace analysis method in the quaternion domain. So, assuming we have a quaternion matrix P with dimension $m \times n$, we consider n to be the total number of the captured data and m the number of the actual hyper-complex features. A quaternion PCA of P as it was analysed above seeks a solution that contains r ($r < m, n$) linearly independent quaternion eigenvectors in the columns of Q ($m \times r$) such that $P = QA$ where the rows of A ($r \times n$) contain the r quaternion principal component (QPC) series. As a result a solid representation of the selected quaternion features is obtained, while the computational complexity is low.

Model creation and classification

The dimensionality of the quaternion feature vector was reduced using Q-PCA, and the final model is created using KNN classifiers. The k-nearest neighbors (KNN) algorithm finds the k-nearest neighbors among the training data, and they are used to weight the category candidates. The performance of this algorithm depends on two factors: the similarity function and the k value (e.g. if k is too large, big classes will overwhelm the small ones). The approaches that can be used are shown in the equations below.

$$y(\mathbf{d}_i) = \arg \max_{\mathbf{k}} \sum_{\mathbf{x}_j \in \text{kNN}} \|y(\mathbf{x}_j, \mathbf{c}_k)\| \quad (9)$$

$$y(\mathbf{d}_i) = \arg \max_{\mathbf{k}} \sum_{\mathbf{x}_j \in \text{kNN}} \|s(\mathbf{d}_i, \mathbf{x}_j) y(\mathbf{x}_j, \mathbf{c}_k)\| \quad (10)$$

where \mathbf{d}_i is the test data, with \mathbf{x}_j to belong in class \mathbf{c}_k , and $s(\mathbf{d}_i, \mathbf{x}_j)$ represents the similarity function for \mathbf{d}_i and \mathbf{x}_j . Also, the most commonly used similarity functions are Euclidean, City block, Cosine and Correlation distances.

2.4 Testing

The objective of the testing stage is the classification of new raw data. Therefore, new incoming raw data is preprocessed, transformed into quaternions and the features are extracted using Q-PCA, following the same pre-processing steps as in the training stage. Once the features are extracted, the model created during the training is used to classify the new input data and determine the gaze location.

3 Evaluation

3.1 Acquisition process and dataset

Nine healthy participants without severe visual impairment were selected. The population of participants is formed of 5 females and 4 males aged between 27

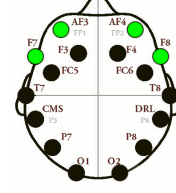


Figure 2. Only 4 of the 14 channels (AF3, AF4, F7 and F8) are used and combined in a single quaternion representation.

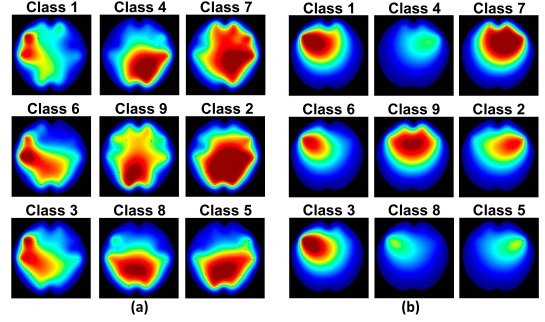


Figure 3. Examples of the EEG signals for each gaze location using (left) all the 14 channels and (right) only the 4 of them.

and 59 years old of whom 4 wore glasses or contact lenses during the test. Non-intrusive devices for EEG and eye tracker data collection were utilised. The EEG signal was collected using Emotiv EPOC headset. This EEG headset is formed of 16 sensors; 2 of them are reference sensors, therefore the brain signal is collected from the remaining 14 sensors/channels at 128Hz. Nevertheless, since our aim is to combine our method with wearable equipment such as Virtual Reality and Augmented Reality devices, only the information of the 4 sensors located closer to the eyes are used (see figure 2). The eye-tracker data is captured using the Tobii Eye tracker, a device located below the monitor and tracks the pupil position returning gaze data at 60Hz. These data include time-stamps of each sample, which allows the synchronization with the EEG signal in order to be used also as ground truth. The duration of the test is 45 seconds; 9 positions are shown using a white area with a duration of approximately 5 seconds each. Their EEG signal and gaze data is recorded during that period.

The proposed approach uses the 4 channels located on the frontal lobe (AF3, AF4, F7 and F8). AF3 and AF4 are affected by vertical eye movements; and F7 (left) and F8 (right) by horizontal eye movements, (see Figure 3). Also these signals/data were recorded using a 128 frame rate EEG recorder.

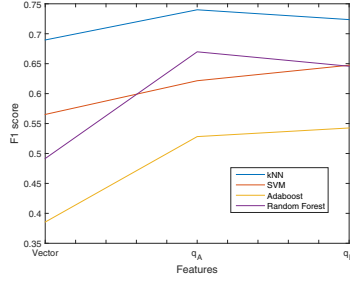


Figure 4. The average F1 scores of the classes.

3.2 Features

After the data is preprocessed to reduce the noise, the proposed novel quaternion descriptors are estimated. The 4 channels are combined in a single quaternion, with the four components (q_r, q_i, q_j, q_k) to correspond to the sensors located on the sides of the head (F7 and F8) and ones over the eyes (AF3 and AF4), respectively. Horizontal movements produce opposite-going voltage traces in F7 and F8 (e.g. when the eye is moving to the left F7 amplitude is increased and F8 is decreased), [11]. The changes in the AF3 and AF4 sensors are proportional, when it comes to vertical eye movements, [8]. The data is split in two different sets: 67% for training and 33% for testing. The separation of the data is done following the Leave-persons out protocol, that is, a set of people is removed to obtain the minimum test set that contains instances of all classes. Since the number of subjects is small, k-fold cross validation is used to limit overfitting problems. The proposed method uses 9-fold cross validation, with the testing data to be taken from a different set of users on each fold. Quaternion PCA is used to reduce the features by 97% and the number of dimensions was selected according to the minimum necessary to obtain the best results.

3.3 Results

The classification accuracy of the proposed approach was evaluated using three different metrics: precision (P), recall (R) and F1 score. Precision is the number of true positives (TP) divided by the amount of true positives plus false positives (FP) $P = \frac{TP}{TP+FP}$, and recall is the number of true positives divided by the amount of true positives plus false negatives (FN) $R = \frac{TP}{TP+FN}$. While, the F1 score is a measure of the tests' accuracy based on the precision and recall and it is defined as $F1 = 2 * \frac{P * R}{P + R}$.

Four different classification methods were used in our comparative study: k-nearest neighbors (k-NN), Support Vector Machines (SVM), AdaBoost and Random Forests. k-NN is a multi class classification

method that assigns new unclassified samples to the class to which the majority of its k nearest neighbors belong. It assumes that the k nearest neighbors of a test sample are located at roughly the same distance from it (this approach uses cosine distance). SVM classifier is a machine learning algorithm that maps the input features into a higher dimensional feature space. A linear decision surface is then constructed in this high-dimensional-feature space so that the margin between the surface and the nearest point is maximized. Since SVM is a binary classification method, it is transformed to a multi-class classifier according to the one vs all relation. AdaBoost classifier combines weak classifiers to create a strong one. The weak classifier has to solve a sequence of learning problems and they are weighted according to their results. The final strong classifier is a weighted combination of the weak classifiers. AdaBoostM2 is a multi-class AdaBoost version, where each weak learner is associated to one class. Random forests is a machine learning classifier that uses a set of tree predictors and weights their output in order to perform a prediction. A tree predictor is a classifier that recursively partitions a data set into smaller subdivisions according to a set of tests defined on each branch of the tree. At the end of the tree, there are the final nodes that are linked with the label of the classes.

The outcomes presented in table 1 are the averaged results of the 9-fold process. The results shown in the table present the results for each classification method, using the obtained features in a vector representation or the proposed one combined in a single quaternion. Regarding the novel quaternion representation two combinations of the four channels were utilised and evaluated altering channel 2 and 3, resulting in the following cases $q_A = (q_r, q_i, q_j, q_k)$ and $q_B = (q_r, -q_j, q_i, q_k)$. The improvement can be seen by observing the results using the proposed quaternion representation over all the classifiers. Furthermore, in figure 4 compares all the classifiers in terms of F1 score over the vector and quaternion representation of the descriptors and it can be seen the improvement that is obtained using the proposed quaternion representation.

4 Conclusion

Virtual and Augmented reality will change the way humans interact with the environment. Since eye movements are important in human computer interaction, the proposed approach provides a novel gaze estimation system using EEG signal for selection based tasks and interactions or foveated rendering. Also, since the VR and AR headset devices have specific requirements in terms of size and weight, the proposed system does not affect the overall design. Our method is using four EEG sensors located on the front of the head without re-

Table 1. Results for all the classifiers.

Feat.	Method	Precision	Recall	F1
Vect.	kNN	0.7500	0.6914	0.6896
	SVM	0.5838	0.6049	0.5652
	Adab	0.4185	0.4568	0.3855
	RF	0.5594	0.5432	0.4916
q_A	kNN	0.7901	0.7407	<u>0.7400</u>
	SVM	0.6321	0.6667	0.6215
	Adab	0.5691	0.5432	0.5282
	RF	0.7031	0.6790	<u>0.6698</u>
q_B	kNN	0.7741	0.7284	0.7236
	SVM	0.6821	0.6667	<u>0.6474</u>
	Adab	0.5679	0.5556	<u>0.5426</u>
	RF	0.7080	0.6543	0.6457

quiring external cameras. Additionally a novel quaternion feature representation was suggested to classify gaze positions improving significantly the overall accuracy of the available classifiers. Experiments were performed based on a cross validation approach demonstrating improvements in the classification precision.

References

- [1] T. Adali, P. Schreier, and L. Scharf. Complex-valued signal processing: The proper way to deal with impropriety. *IEEE Trans. Signal Processing (overview paper)*, 59(11):5101–5123, 2011.
- [2] A. Al-Rahayfeh and M. Faezipour. Eye tracking and head movement detection: A state-of-art survey. *Trans. Engin. in Health and Medicine, IEEE*, 1:210–212, 2013.
- [3] N. Bihan and S. Sangwine. Quaternion principal component analysis of color images. In *ICIP*, volume 1, pages 1–809–12, 2003.
- [4] J. Bonita, L. Ambolode, B. Rosenberg, C. Cellucci, T. Watanabe, P. Rapp, and A. Albano. Time domain measures of inter-channel eeg correlations: a comparison of linear, nonparametric and nonlinear measures. *Cognitive Neurodynamics*, 8(1):1–15, 2014.
- [5] Z. Chai, K.-K. Ma, and Z. Liu. Complex wavelet-based face recognition using independent component analysis. *Fifth Intern. Conf. on Intelligent Information Hiding and Multimedia Signal Proc.*, pages 832–835, 2009.
- [6] M. Chen, X. Meng, and Z. Wang. Quaternion fisher discriminant analysis for bimodal multi-feature fusion. *Chap Adv Intel Syst & Comp*, 370:479–487, 2015.
- [7] G. Gredebck and T. Falck-Ytter. Eye movements during action observation. *Perspect. on Psycholog. Science*, 10(5):591–598, 2015.
- [8] C. Guerrero-Mosquera and A. Vazquez. Automatic removal of ocular artifacts from eeg data using adaptive filtering and independent component analysis. *17th European Signal Proc. Conf.*, pages 2317–2321, 2009.
- [9] D. W. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *Pattern Analysis and Machine Intelligence, IEEE*, 32(3):478–500, 2010.
- [10] D. W. Hansen and A. E. Pece. Eye tracking in the wild. *CVIU*, 98(1):155–181, 2005.
- [11] O. Jensen, J. Gelfand, J. Kounios, and J. Lisman. Oscillations in the alpha band increase with memory load during retention in a short-term memory task. *Cerebral cortex*, 12(8):877–882, 2002.
- [12] C. C. Lai, S. W. Shih, and Y. P. Hung. Hybrid method for 3-d gaze tracking using glint and contour features. *Circuits and Systems for Video Technology, IEEE*, 25(1):24–37, 2015.
- [13] K. Li, G. Sun, B. Zhang, S. Wu, and G. Wu. Correlation between forehead eeg and sensorimotor area eeg in motor imagery task. In *Dependable, Autonomic and Secure Computing, 2009. DASC '09. Eighth IEEE International Conference on*, pages 430–435, Dec 2009.
- [14] X.-L. Li, T. Adali, and M. Anderson. Noncircular principal component analysis and its application to model selection. *IEEE Sig. Proc.*, 59(10):4516–4528i, 2011.
- [15] P. Majaranta and A. Bulling. Eye tracking and eye-based humancomputer interaction. *Adv. in Physiological Computing*, pages 39–65, 2014.
- [16] H. Manabe, M. Fukumoto, and T. Yagi. Direct gaze estimation based on nonlinearity of eeg. *Biomedical Engineering, IEEE*, 62(6):1553–1562, 2015.
- [17] G. R. Mller-Putz, R. Riedl, and S. C. Wriessnegger. Electroencephalography (eeg) as a research tool in the information systems discipline: Foundations, measurement, and applications. *Communications of the Association for Information Systems*, 37(1):46, 2015.
- [18] A. Plopski, C. Nitschke, K. Kiyokawa, D. Schmalstieg, and H. Takemura. Hybrid eye tracking: combining iris contour and corneal imaging. *25th on Artif. Reality and Telexistence*, pages 183–190, 2015.
- [19] T. Schneider, B. Schauerte, and R. Stiefelhagen. Manifold alignment for person independent appearance-based gaze estimation. *22nd Intern. Conf. on Pattern Recognition (ICPR)*, pages 1167–1172, 2014.
- [20] N. Steinhausen, R. Prance, and H. Prance. A three sensor eye tracking system based on electrooculography. *Sensors*, pages 1084–1087, 2014.
- [21] M. Stengel, S. Grogorick, E. Eisemann, M. and Eisemann, and M. A. Magnor. An affordable solution for binocular eye tracking and calibration in head-mounted displays. *23rd Annual ACM Conference on Multimedia Conference*, pages 15–24, 2015.
- [22] Y. Sugano, Y. Matsushita, and Y. Sato. Learning-by-synthesis for appearance-based 3d gaze estimation. *IEEE CVPR*, pages 1821–1828, 2014.
- [23] S. Valenzi, T. Islam, and A. Jurica, P. and Cichocki. Individual classification of emotions using eeg. *Journal of Biomedical Scien. and Engin.*, 7(8):604–620, 2014.
- [24] E. Wood and A. Bulling. Eyetab: Model-based gaze estimation on unmodified tablet computers. *Symp. on Eye Tracking Research and Applic.*, pages 207–210, 2014.
- [25] D. Yang, Q. Bai, Y. Zhang, R. Ji, and H. Zhao. Eye location based on hough transform and direct least square ellipse fitting. *Journal of Software*, 9(2):319–323, 2014.
- [26] S. Yang and F. Deravi. Novel hht-based features for biometric identification using eeg signals. In *ICPR 22nd Inter. Conf.*, pages 1922–1927, 2014.