Gender impact on the identification based on EEG

Jinghai Yin, Zhendong Mu*
Institute of Information Technology, Jiangxi University of Technology, Nanchang 330098, China
Received 1 July 2014, www.cmnt.lv

Abstract

To study the genders impact on identification, this paper analysis the electroencephalograph (EEG) of eight male subjects and seven female subjects. In order to reduce the noise signal interference, the high pass and low pass were used to cut extra frequencies, and in order to prominent the feature signal, the power spectrum method was used to convert the time domain signal to frequency domain, then Fisher distance was used to extract the feature. All EEG signal was acquired by neuroscan, and the EEG signal was evoked by VEP method used subjects photo. The experiment was divided into three models: all subjects were the same sex, added some opposite sex, added some stranger. The analysis results show, to model 1, the correct recognition rate for male subjects, average is 88.50, and this for female, average is 92.51%; the false recognition rate for male subjects, average is 30.84%, and this for female, average is 27.67%, this result indicates VEP can be used as identification tool, the results of model 2 and model 3 show weather opposite sex or stranger should affect the correct recognition rate, but to male subject, the opposite sex effect is greater than stranger, to female, the result were reversed. The results also show noise photo affected female lower than male.

Keywords: Genders Impact, EEG, gender, Fisher Distance

1 Introduction

EEG brain of external things to make a biological reaction process to produce electricity, and when received outside special stimulation, the brain will produce features with this stimulation, the use of this nature, by brain researchers electrical signals to the brain as a bridge launched a variety of studies, such as the use of EEG studies BCI system control peripheral devices, the use of EEG studies of various diseases, the use of EEG to study various psychological phenomena, etc. With the development of a variety of camouflage technology, the traditional identification technology increasingly difficult to do today's information security applications, looking for new biometric technology becomes increasingly important.

EEG advantages due to its difficult to counterfeit, and therefore the use of EEG as an identification tool for many a research investigator in the study of electrical signals in the brain, evoked potentials, especially the visual evoked potential cut is easy to implement because of the stability characteristics which is widely used in the identification same study, the researchers also used a lot of evoked potentials identification study, more successful in these studies. Paranjape et al Acquisition subjects wide open, EP signal is generated when the eyes closed, use AR model parameters as recognizable features [1]. In a recent research results, Riera et al chose the two electrodes, collecting subjects wide open, EP eyes closed when the signal is analyzed using autoregressive, Fourier transform, mutual information, etc. multi-feature fusion achieved 87.5% to 98.1% recognition rate[2]. When there is a specific visual stimulus will produce visual evoked potential (Visual Evoked Potential, VEP), which is a good entry point, Singhal et al use VEP peak there are individual differences in characteristics, setting relevant peak potential matching algorithm identifying subjects also achieved good results[3]. In study Biel, etc., then the design of the subjects viewed a picture, the choice of the three electrodes, using VEP as a signal source analysis, implementation verification using multivariate statistical methods. Palaniappan use the definition of an automatic identification method using Davies-Bouldin index information to select the largest electrode, the black and white pictures of common objects as visual stimuli, using a neural network classification [4, 5]. Touyama et al study is unfolding from the perspective of ERP[6], they take advantage of the characteristics of the P300, P300 evoked that will only appear when the target stimulus, the experiment they let 9 photos randomly, participants selected targets stimulation, selected different subjects of different target stimulus, they choose a password that is [7]. Wardzinski et al while the use of linear and nonlinear methods for EEG signal analysis, the model parameters as features, and based on Mahalanobis distance classifier, the recognition rate of 88% [8, 9]. Poulos et al uses of wave EEG signals were identified, using the AR model for feature extraction, the final recognition rate between 72 -84%[10,11,12,13].

Based on existing research, we use visual evoked potential as a tool for identity recognition research, the first of the collected EEG filtering, select a specific wavelength as the signal source, and then on the EEG power spectrum estimation, calculation different band EEG power spectrum, and finally the use of distance between the characteristics of different subjects for feature selection, feature matching validated by the final

* Corresponding author e-mail: 418623577@qq.com
results achieved recognition rate 80%-96% [14, 15], the results show that the use of visual evoked potentials can be very good for identity recognition research.

2 Visual evoked potentials and experimental models

Visual evoked potential (Visual Evoked Potential, EVP) is occipital cortex electrical response to visual stimuli occurred District, on behalf of the retina receive stimulation through visual pathway conduction to the occipital cortex caused by potential changes. Visual evoked potential (VEP) in respect of its stimulation divided into non graphical stimulus ( flash stimulation ) and graphic stimuli into two categories, the current use of the latter in particular Othello plate grid, grating pattern image. Their black and white compositions at a rate conversion constitute a valid alternative stimulation induced electrical activity called pattern reversal visual evoked potentials (PRVEP). With visual stimulation in the occipital scalp recorded VEP main representative of the central field of vision visual impulses 6° ~ 12°, and by the lateral geniculate body projected onto the occipital pole electrical activity from the state after the split rear and pillow. Today, visual evoked potentials, is the most successful application of visual evoked potentials for detection of retinal ganglion cells and optic nerve diseases such as central as the passage of optic neuritis, multiple sclerosis, cancer and other oppressed optic neuropathy disease ; graphics VEP can be used for an objective determination of visual acuity. In the visual evoked potential studies often use ERP component analysis, the main object of study for the N75, P100, N145 and other ingredients.

As used herein, the basis of the different subjects of the experiment is to familiarize yourself with the different responses favorite pictures and other pictures reactions, resulting in different brain wave signal is mainly based on, for example, a photo of yourself and a photo of a different reaction to the others collected in accordance with the visual EEG evoked potentials data . Experiments in Jiangxi Institute of Technology Institute of Information Technology BCI laboratory, Jiangxi University of Science subjects in school, subjects were placed in a quiet shielded room, sitting on soft chairs without armrests experiment during the experiment, subjects in accordance with the test requirements, looking at the computer screen in front of the related operations, 15 experimenters were divided into three groups of five people, the experiencer experimental models include the following three modes : experimental background for the next gay familiar experiments have become familiar with the background to participate in sex case experiments and experimental background in the case of a stranger to participate in the experiment. Under each group mode, the experiencer was asked to do attention be collected under their own photos EEG circumstances and concerns of others cases, each experiment, the picture display 750ms, then the interval 1s.

EEG acquisition is the use of lead Neuroscan amplifier 40 through scan4.3 software for acquiring, using the right of way of the mastoid reference electrode as the reference electrode, the use of 1000Hz sampling rate, the use of 200Hz low-pass band collection, 0.05Hz 50Hz high-pass and notch.

3 Method

3.1 FILTER

Intermediate signals received useful signal selection process is called filtering, "received signal" corresponds to the observed random process, the "useful signal" corresponds to the random process is estimated. E.g., aircraft tracking radar, the measured position data of the aircraft, the other containing the measurement errors and random noise, how to use them to estimate as accurately as possible the position of the aircraft at each moment, velocity, acceleration, and to predict future aircraft position, is a filtering and prediction problems. Such problems in electronics, aerospace science, control engineering and other scientific and technical sectors are abounding.

Filtering is carried out in accordance with, or just on some sampling points can be divided into a continuous-time filter and a discrete-time filtering on the entire time. The former set time parameters T will be desirable half real axlet{0,∞} or the real axis; latter T desirable non-negative integers {0, 1, 2, 3…} or integers {…..-2, -1, 0, 1, 2…}.

Set \( X = \{ X, t \in T, t \in T \} \) is limited, that is to say:

\[
C(H_j | r) = \sum_{i=1}^{M-1} C_{ij} P(H_i | r),
\]

where X is a process to be estimated, it can not be directly observed; Y is the observed process, which contains some information of X.

Use \( y' = \{ y_s : s \in T, s \leq t \} \) to represents the observed data with the time t until the whole. If you can find a function f(x) about a variable in y making it reach the minimum mean square error of \( E \mid X, f(y)' \), \( X_t - f(y)' \) is said to the optimal filtering \( X_t \).

In order to facilitate the application and narrative, sometimes also above definition to classify more detail. Let \( \tau \) t is a real number or integer determined, and the process is considered to be estimated as

\[
(X_{\tau+t}, t \in T), X_{\tau+t} = \tilde{E}(X_{\tau+t} | y).
\]

Or

\[
\tilde{E}(X_{\tau+t} | y'), X_{\tau+t} = X_{\tau+t} - X_{\tau+t} D_{\tau+t}.
\]

According to \( \tau = 0, \tau > 0, \tau < 0 \), are called optimal filtering, (\( \tau \)-step) prediction or extrapolation, (\( \tau \) step)
smoothing or interpolation. \( X_{i+r} \), \( D_{i+r} \) corresponding to each error and the mean square error, and collectively these problems for the filter problem.

According to brain wave frequency characteristics, this paper 1Hz high pass and 55Hz low-pass method of combining the signal source is filtered, collected EEG first by 1Hz high-pass filter to reduce direct impact, and then conduct a 55Hz low-pass filter, so You can select 1-55hz useful EEG.

3.2 POWER SPECTRUM CALCULATION

The original EEG is a signal domain, the features are hidden between the various noise data, through the original EEG feature extraction speed and accuracy will be affected, in order to improve the speed of analysis and analysis accuracy, can be converted to the original EEG power spectrum, the frequency analysis.

A simple way to estimate the power spectrum of the stochastic process is a direct request sampling DFT, and then takes the results of the square of the amplitude. Such a method is called period gram

A length \( L \) of the signal \( x_L[n] \) is estimated period gram PSD as follow

\[
P_{xx}(f) = \frac{|X_L(f)|^2}{f L},
\]

where \( X_L(f) \) is defined using the FFT matlab inside without normalization coefficients, so dividing \( L \).

\[
X_L(f) = \sum_{n=0}^{L-1} x_L[n]e^{-2\pi jfn/L}.
\]

The actual calculation of \( X_L(f) \) may be performed using FFT, and only on a limited frequency. Applications in practice are most period gram PSD estimation to calculate the N-point:

\[
P_{xx}(f) = \frac{|X_L(f)|^2}{f L}, f_k = \frac{fk}{N}, k = 1,2,3...,N - 1
\]

\[
X_L(f_k) = \sum_{n=0}^{L-1} x_L[n]e^{-2\pi jfn/N}.
\]

3.3 FEATURE EXTRACTION METHODS

EEG feature extraction in order to extract useful brain signals, thereby to classify the feature extraction method to identify subjects used herein are described below:

Step1. Common average: In this paper, we use Hjort derivation to reduce interference from the neighbouring electrode, The Hjort derivation \( C_i \) is calculated as

\[
C_i = c_i - \frac{1}{4} \sum_{j \in S} s c_j,
\]

where \( c_i \) is the reading of the centre electrode \( s c \), with \( i=1...30 \) and \( j \) is the set of indices corresponding to the eight electrodes surrounding electrode \( c_i \).

Step2. Filter: EEG signal acquisition band from 0.05Hz to 200Hz, in order to extract features, we filter EEG signal band from 0.05Hz to 50Hz.

Step3. AR conversion: time-domain EEG data disorganized EEG in order to better highlight the characteristics of EEG signal, we use AR model to convert the time domain signals into frequency domain, and extract the feature from the frequency domain signals.

Step4. Calculate add of Fisher’ distance: Calculated the Fisher distance between two classes. Fisher distance \( F_{i,j} \) is calculated as

\[
F_{i,j} = \frac{(\mu_i - \mu_j)^2}{\sigma_i^2 + \sigma_j^2}.
\]

where \( F_{i,j} \) is Fisher’ distance between the subject I and the subject j. \( \mu \) and \( \sigma \) are the mean and the standard deviation of the feature they correspond to.

The add of Fisher’ distance \( F_i \) is calculated as

\[
F_i = \sum_{j \neq i} F_{i,j}.
\]

Step5. Sort the Fisher’s distance adds: The Fisher distance was often used to denote differences between classes in classification research. The bigger the Fisher distance was the more notable the difference was. So we sort the Fisher’s distance on descending order.

Step6. Feature extraction: Through analysis the five subjects Fisher’s distance, we select three thousand as feature range, all feature are selected from this range.

Step7. Get the classifier: Use BP neural network to calculate the feature then can get these five subjects classifier. To the test sample of these subjects, if the result is in credibility, that is to say, this sample is right

In this paper, the correct recognition rate, error recognition rate, false recognition rate and credibility used to analysis the results, the following were defined them:

1. Correct recognition rate: the correct recognition rate is defined as the rate of correct recognizes the iden-
2. Error recognition rate: the error recognition rate is defined as the rate of incorrect recognizes the identity of subjects through use subjects' classification to test their own samples.
3. False recognition rate: the false recognition rate is defined as the rate of recognize the identity of subjects is himself through use subjects' classification to test others samples.
4. Credibility. The credibility is defined as acceptable error range (in this paper, the credibility is [-0.02 0.02]).

4 Results and discussion

In this paper, experiments using three modes: all same group subject is same gender and familiar, some subject in group is different gender, some subjects in group is stranger. Participation has sex with strangers during the experiment, namely to study the impact of the opposite sex and strangers on the classification accuracy, each of which is divided into a group of five people to experiment, three groups were collected from 15 individuals of brain waves, which 8 males, 7 females, each subject collected EEG data 183 trial, excluding invalid trial, each subject was eventually carried out using 156 trial identification calculations, this paper is divided into these 156 trial characteristics extraction collection and validation collection, which features a collection of 50 extraction trial, the validation set 106 trial.

After the same background, the same group of subjects were EEG analysis, the feature set of the first to use the method described earlier for data conversion, data conversion After feature extraction, feature subjects extracted using the same data conversion method for EEG data conversion, and then select the appropriate feature set, classification calculated by feature matching, feature extraction for a set T, and a test set tc, electrical test matches z formula can be expressed in the brain as follows:

\[ z = \frac{1}{n} \sum_{i=1}^{n} |x_i - y_i| \]

where n is the number of features.

Classification accuracy of the eight male subjects show as following.

As described in Figure 1, the x axis indicates that the three experimental model, 1 is all same group subject is same gender and familiar, 2 is some subject in group is different gender, 3 is some subjects in group is stranger. The Y axis indicates that the value of correct recognition rates, and corresponding to the histogram of each model is eight male subjects.

As figure 1 shows, when input the test samples, the correct recognition rate of model 1 are 96.33%, 91.18%, 96.03%, 94.96%, 79.71%, 82.41%, 90.04%, 77.05%. The correct recognition rate of model 2 are 59.89%, 59.20%, 58.70%, 63.96%, 59.52%, 59.79%, 65.21%, 65.59%. The correct recognition rate of model 3 are 76.53%, 66.85%, 79.85%, 74.71%, 82.30%, 71.22%, 81.64%, 76.37%. As correct recognition rate result shows, if the same group subject all is male, the lowest result is 77.05%, and the highest result is 96.33%, that is to say, use EEG signal as tool of identification can recognize them, if test in model 2, the lowest result is 58.70%, the highest result is 65.59%, compare to model 1, the correct recognition rate is decreased, if test in model 3, the lower result is 66.85%, the highest result is 82.30%, compare to model 1, the correct recognition rate is decreased too. Figure 2 show the eight male subjects false recognition rates.
As described in Figure 2, the x axis indicates same as Figure 1; y axis indicates that the value of false recognition rates of eight male subjects.

As figure 2 shows, when input the test samples, the false recognition rate of model 1 are 29.31%, 27.20%, 26.45%, 30.98%, 35.48%, 35.36%, 34.17%, 27.75%. The false recognition rate of model 2 are 61.66%, 49.90%, 61.89%, 41.49%, 59.65%, 62.87%, 61.73%, 54.06%. The false recognition rate of model 3 are 41.79%, 43.13%, 37.86%, 40.56%, 40.95%, 36.91%, 43.86%, 38.67%.

As correct recognition rate result shows, if the same group subject all is male, the lowest result is 89.21%, and the highest result is 96.46%, that is to say, use EEG signal as tool of identification can recognize them, if test in model 2, the lowest result is 77.66%, the highest result is 86.88%, compare to model 1, the correct recognition rate is decreased, if test in model 3, the lower result is 81.11%, the highest result is 93.69%, compare to model 1, the correct recognition rate is decreased too.

As described in Figure 3, the x axis indicates same as Figure 1; y axis indicates that the value of correct recognition rates of seven female subjects.

As figure 3 shows, when input the test samples, the correct recognition rate of model 1 are 89.46%, 93.20%, 91.32%, 96.24%, 96.46%, 91.69%, 89.21%. The correct recognition rate of model 2 are 77.66%, 85.70%, 81.31%, 82.31%, 86.68%, 83.63%, 85.38%.. The correct recognition rate of model 3 are 81.11%, 90.78%, 86.43%, 93.69%, 84.96%, 91.74%, 83.92%.. As correct recognition rate result shows, if the same group subject all is male, the lowest result is 89.21%, and the highest result is 96.46%, that is to say, use EEG signal as tool of identification can recognize them, if test in model 2, the lowest result is 77.66%, the highest result is 86.88%, compare to model 1, the correct recognition rate is decreased, if test in model 3, the lower result is 81.11%, the highest result is 93.69%, compare to model 1, the correct recognition rate is decreased too.

As described in Figure 4, the x axis indicates same as Figure 1; y axis indicates that the value of false recognition rates of seven female subjects.

As figure 4 shows, when input the test samples, the false recognition rate of model 1 are 23.22%, 22.04%, 27.99%, 35.32%, 32.42%, 18.73%, 33.95%.. The false recognition rate of model 2 are 30.48%, 36.23%, 27.08%, 45.44%, 38.15%, 44.87%, 27.76%. The false recognition rate of model 3 are 56.91%, 41.50%, 40.65%,39.63%, 55.31%, 49.00%, 51.80%..

Experimental results show that the time when the subject of the same sex, male subjects being able to confirm the correct recognition rate their average of 88.50%, and the average of female correct recognition is
92.51%, the test sample is added when the others after EEG sample misidentification rate male subjects with an average of 30.84%, the error recognition rate of female subjects with an average of 27.67%, the experimental data show that in their own photo as a reference, the female subjects than recognition rate is generally higher than male subjects male subjects, while female subjects to others EEG signals recognized as its own recognition rate is also lower than male subjects, VEP feature shows the subjects picture familiarity observed, the average correct recognition rate and error recognition rate results showed photos of female subjects for their degree of concern than male subjects, the correct identification rate data also show the use of visual evoked potentials can be used as identity identification tool.

In order to study the change of identification rate, this paper introduced opposite sex and strangers photo as noise. in the same experimental model and methods of analysis, compared to model 1, table 1 show the correct recognition rate of eight male change of add opposite sex and strangers experimental model. the classification accuracy decreased average: 0.2702, wherein the maximum change is 0.3733, min changes is 0.1146, when the stranger cases, the average correct classification rate dropped 0.1232, the biggest change is 0.2433, the smallest change is 0.0259.

The result of Table 1 show that male subject easer affected by opposite sex than stranger.

The same experiment for female, we can get the result table 2.

As table 2 shows, to the seven female subjects, if add opposite sex subject, the correct recognition rates would decrease, the average rate dropped 0.0924, the biggest change is 0.1393 and the smallest change is 0.0383. The result of Table 2 show that the opposite sex and strange also affected the female subjects focus.

Compare table 1 and table 2 results, gender impact on the experiment as show in Figure 5, to opposite sex, this effect is more than female, and to strange, this effect also male more than female, this VEP result show that female is focus on their own photo more than male, so the correct recognition rate effect to noise photo is smaller than male. Whether male or female, the correct recognition rate show use this way to identification is effective.

![Figure 5](image_url)

**FIGURE 5 Gender impact on correct recognition rate**

The correct recognition rate is affected by noise photo; weather false recognition rate would effect by noise too?

Table 3 and table 4 is male and female result of this effect.
To the study of VEP, if uses the method of overlap some trial data, the gender impact is not big, but to the study of identification, weather this impact is no big too? Through the eight male subjects and seven female subjects research results show that the noise of opposite sex and strangers subjects will be affected, but the impact of female subjects than male subjects suffered by little affected.

As show in Table 1 and Table 2, under model 2, that is to say if a opposite sex subject be added, the false recognition rate, to male subject will ascend, the average number is raise 0.2657, and under model, the false recognition rate is also ascend, this number average raise 0.0963, this result show, to male subject, the opposite sex is more attractive than stranger. The same analyses method to female, under this two models, the false recognition rate are all ascend, to model 1, this number raise 0.0805, and to model 2, this number raise 0.2016, that is to say, to female subjects, the stranger is more attractive than opposite sex, this result can describe in Figure 6.

**References**


[7] Rocca Daria La Patrizio Campisi, Jordi Sole-Casals 2013 EEG based user recognition using BUMP modelling *Biometrics Special Interest Group (BIOSIG), 2013 International Conference of the* IEEE

**Acknowledgments**

This work was supported in part by a grant from natural science foundation of Jiangxi Province [No 20122BAB201049] and IT projects of Jiangxi Office of Education [No. GJJ14765].
Yin Jinghai, Mu Zhendong


Authors

Jinghai Yin, born on January 28, 1977, China
Current position, grades: researcher at Jiangxi University of Technology, China.
University studies: M.E degree in computer science and technology from Nanchang University, China in 2005.
Scientific interests: information security, algorithm design and EEG analysis.

Zhendong Mu, born on November 1, 1975, China
Current position, grades: researcher at Hunan University of Commerce, China.
University studies: M.E degree in computer science and technology from Nanchang University, China in 2004.
Scientific interests: information security, algorithm design and EEG analysis.