computer methods and programs in biomedicine 94 (2009) 48–57

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A new approach for EEG feature extraction in P300-based lie detection

Vahid Abootalebi[a,b,d,*], Mohammad Hassan Moradi[b], Mohammad Ali Khalilzadeh[c,d]

[a] Electrical Engineering Department, Yazd University, Yazd, Iran
[b] Biomedical Engineering Faculty, Amirkabir University of Technology, Tehran, Iran
[c] Islamic Azad University, Mashhad Branch, Iran
[d] Research Center of Intelligent Signal Processing, Tehran, Iran

ARTICLE INFO

Article history:
Received 15 July 2007
Received in revised form
10 June 2008
Accepted 6 October 2008

Keywords:
Psychophysiological detection of deception
Event-related potential (ERP)
P300
Lie detection
Pattern recognition
Feature selection

ABSTRACT

P300-based Guilty Knowledge Test (GKT) has been suggested as an alternative approach for conventional polygraphy. The purpose of this study was to extend a previously introduced pattern recognition method for the ERP assessment in this application. This extension was done by the further extending the feature set and also the employing a method for the selection of optimal features. For the evaluation of the method, several subjects went through the designed GKT paradigm and their respective brain signals were recorded. Next, a P300 detection approach based on some features and a statistical classifier was implemented. The optimal feature set was selected using a genetic algorithm from a primary feature set including some morphological, frequency and wavelet features and was used for the classification of the data. The rates of correct detection in guilty and innocent subjects were 86%, which was better than other previously used methods.

1. Introduction

The ability to detect deception has important legal, moral and clinical implications, and has recently received revived interest from the scientific community. Currently, the polygraphic tests are the most widely used technique for the quantitative discrimination between deceptive and truthful responses [1]. Polygraphy is relies on some measures of autonomic nervous system response such as respiration pattern, cardiovascular measures and electrodermal response (EDR).

By their very nature, polygraph measurements provide an extremely limited and indirect view of complex underlying brain processes. A researchable hypothesis is that by looking at brain function more directly, it might be possible to understand and ultimately detect deception [2,3]. Based on this hypothesis a number of neurophysiological signals have recently been investigated for the possible application to deception detection, including Functional Magnetic Resonance Imaging (fMRI) [4–10] and event related potentials (ERP) [2,11,12].

ERPs are recorded from the central nervous system and are considered to be affected by the recognition of important events, which is more cognitively determined activity than autonomic responses. An endogenous ERP, which has been extensively studied, is the P300 (P3) wave. It is seen in response to rare, meaningful stimuli often called "oddball"

* *Corresponding author at*: Electrical Engineering Department, Yazd University, Yazd, Iran. Tel.: +98 351 8210699; fax: +98 351 8210699.
E-mail addresses: abootalebi@yazduni.ac.ir (V. Abootalebi), mhmoradi@aut.ac.ir (M.H. Moradi), makhalilzadeh@mshdiau.ac.ir (M.A. Khalilzadeh).
0169-2607/$ – see front matter
doi:10.1016/j.cmpb.2008.10.001

stimuli [13]. For example, if a subject is viewing a Bernoulli (random) series of names, one every three seconds, and occasionally, one of these is the subject's name, a P3 wave is evoked in response to this rarely presented, recognized, meaningful (autobiographical) stimulus. P3 is a positive-going wave with a scalp amplitude distribution in which it is largest parietally (at Pz) and smallest frontally (Fz), taking intermediate values centrally (Cz). (Fz, Cz, and Pz are scalp sites along the midline of the head.) Its peak has a typical latency of 300–1000 ms from stimulus onset. The size or amplitude of P3 at a given recording site is inversely proportional to the rareness of presentation; in practice, probabilities <0.3 are typically used. The meaningfulness of the stimulus is also extremely influential in determining P3 size.

The P300-based GKT is a "Guilty Knowledge Test (GKT)" which utilizes P300 amplitude as an index of actual recognition of concealed information. GKT is a method of polygraph interrogation that facilitates psychophysiological detection of prior knowledge of crime details that would be known only by the suspect involved in the crime, and of course by authorities, police, etc. [14,15]. The GKT rests upon the assumption that familiar items will elicit different responses when presented in the context of a large number of homogeneous unfamiliar items.

In P300-based GKT, there are typically three kinds of stimuli presented to subjects:

(1) *Probes* (*P*), which are related to concealed information and are known only to the guilty person and authorities. In fact, the guilty subject is expected to know these stimuli, but the innocent one is not.
(2) *Irrelevants* (*I*), which are items unrelated to the criminal acts and thus unrecognized by all subjects (guilty or innocent).
(3) *Targets* (*T*), which are usually irrelevant items, but all subjects are asked to do a task (for example, pressing a button or count increasing) whenever they see the *T*, but not when they see a *P* or an *I*.

The number of irrelevant stimuli is many times greater than the numbers of the other two types; and therefore probes and targets are rare stimuli. The *T* stimuli force the subject to pay attention to items, because failure in responding to these stimuli suggests that the subject is not cooperating. Also, the *T* stimuli are rare and task relevant and thus evoke a P300 component that has been used in subsequent analysis of the probes as a typical P300 of the subject [2], although this assumption that the T-P300 is a classical rendition of standard P300 has been shown to be sometimes wrong [16].

The basic assumption in the P300-based GKT is that, if the subject has guilty knowledge of the probe stimuli, the infrequent nature of these items will cause them to elicit a P300 component like that for the *T* stimuli. However, if the person has no knowledge of the probe items, they will be perceived as belonging to the irrelevant stimulus set and thus elicit an ERP with only small or no P300 component.

There are conventionally two approaches in the analysis of signals and detection of deception in P300-based GKT. In the first – used by Rosenfeld et al. – the amplitude of P300 response in *P* and *I* items are compared [16]. In guilty subjects, one expects $P > I$ while in innocents *P* is another *I* and so no *P*–*I* difference is expected. Based on this theory, the *Bootstrapped amplitude difference* (BAD) method has been introduced and used by Rosenfeld.

The second approach, introduced by Farwell and Donchin [2], is based on the expectation that in a guilty person, the *P* and *T* stimuli should evoke similar P300 responses, whereas in an innocent subject, *P* responses will look more like *I* responses. Thus, in this method we called here *Bootstrapped correlation difference* (BCD), the cross correlation of *P* and *T* waveforms is compared with that of *P* and *I*. In guilty subjects, the *P*–*T* correlation is expected to exceed the *P*–*I* correlation and the opposite is expected in innocents.

Previously, we introduced a new approach for ERP assessment in a P300-based GKT [17]. This approach was a pattern recognition system based on some wavelet features and a statistical classifier. We implemented these three approaches and compared their performance in our previous study [17]. The purpose of the present study was to extend even further our pattern recognition system using new features. The new feature set consists of morphological and also frequency features in addition to previous wavelet features. Hence, first a feature set consisting of three groups of features were defined and extracted from the raw data. Following the feature extraction, the best features were selected from the primary feature set and were used for the classification of the data.

In this paper, we first explain the protocol of our experiments and the characteristics of the collected data. The definition of the used features and the method of feature selection will be described afterwards. The algorithm for the classification and detection of the subjects will be explained in the next section. In the end, the performance of this novel method will be compared with the previous wavelet classifier and also with the BAD and BCD methods.

2. Methods

2.1. Subjects

Sixty-two subjects (59 males, 3 females) participated in the study. They were generally undergraduate or postgraduate students and all had normal or corrected vision.

2.2. Data acquisition

The electroencephalogram (EEG) was recorded using Ag/AgCl electrodes placed at the Fz (Frontal), Cz (Central) and Pz (Parietal) sites (10–20 international system). All sites were referenced to linked mastoids. Only the results from Pz will be reported here. Electrooculogram (EOG) was recorded from sub- and supraorbital electrodes (above and below the right eye). The subjects were grounded at the forehead. Brain electrical activities were amplified and digitized at a rate of 256 samples per second. Digitized data were subsequently analyzed offline using MATLAB software. Prior to data analysis, all data were digitally filtered in the 0.3–30 Hz range. This is the frequency range which is used typically in P300-based GKT studies [16].

2.3. Procedure

All subjects were trained and then performed a mock crime scenario. After a training phase about the protocol, a box – containing a jewel – was given to the subject. The examiner left the room and permitted the subject to perform his role in the scenario. In this step the subject could choose and implement one of two possible roles (guilty/innocent). The guilty subject opened the box, saw the jewel precisely and imagined that he/she has stolen the jewel. The subject was asked to memorize the jewel details, so he/she can act as the actual robber. For reassurance, the subject was asked to write the detail of the object on a piece of paper and deliver it after the test. The innocent subject has never opened the box and thus has no information about the object. This protocol is a single-probe protocol similar to protocols which are usually used by Rosenfeld [16].

Then the examiner returned to the room and executed the P300-based GKT for the knowledge of the scenario. After the attachment of electrodes and starting the recording, stimuli consisting of single pictures of five peaces of jewelry, were presented randomly and successively on a monitor to the subject. Each item remained 1.1 s on the screen, of which 1 s were used for processing. The inter-stimulus interval was 2 s. During the test, pictures of five items (for example: ring, wrist watch, necklace, earring and hairpin) were randomly presented one at a time (each with 30 iterations). These objects contained one target, one probe and three irrelevants. The probe was the object in the box. The target was another object, which had been previously presented to the subject in the training phase at the start of the protocol.

One push button was given to each hand of the subject, right click as "YES" and left click as "NO". The subjects were asked to reply to the items by pressing one of the two buttons; "YES" for familiar items and "NO" for unknowns. This task had already been practiced during training phase. The subjects also were instructed to try to hide their information about the probe item in the guilty condition and to behave mentally as innocent. Any covert action was permitted, providing that it did not disarrange the totality of the test and was not observable to the examiner either by visual inspection of the subject or from the recorded signal. Hence, actions such as voluntary blinking or body movement detected by the examiner were not permitted. Indeed, all subjects were supposed to press "YES" for targets and "NO" for probes and irrelevants. Thus, guilties and innocents, both replied honestly to targets and irrelevants; but for probes, the guilty subjects replied falsely while the innocents answered honestly. Subjects with more than 10% error in correct clicking for all stimuli were removed from analysis.

Each subject participated in two experiments in which the boxes and all displayed jewels were different. The jewel in the box was a gold coin in the first experiment and a hairpin in the second experiment. Subjects were instructed to select a guilty role in one experiment and an innocent role in the other, by their choice. Signals from these two experiments were analyzed independently. Therefore with 62 subjects, a total number of 124 tests were performed. 32 subjects chose the innocent role and 30 subjects chose the guilty role in their first experiment. A few test results were removed due to misdoing of protocol or inappropriately recorded signals. Finally, 59 guilty cases and 51 innocent cases were chosen to be used in the following investigations.

2.4. Data analysis

After the filtering of signals, each continuous record was separated to single sweeps according to the known times of stimulus presentation. The length of each sweep is 1000 ms which contains 256 samples of signal. EOG data were checked for blink artifact by visual inspection and sweeps with blink artifacts were removed. Then the pattern recognition method including feature extraction, feature selection and classification were applied on the signals and the detection rates of them were assessed. It should be noted that, in all description of P300 follow, only the results at site Pz was noted, since Pz is the site where P300 is usually reported to be maximal and therefore the analytic procedure (below) were performed on Pz data only.

Similar to commonly used pattern recognition systems, in our processing method, first some suitable features are extracted from each single trial. These features will be those which contain suitable information about the inspected phenomena. After the feature extraction step, a classifier is used for the separation of ERPs from guilty subjects and those of innocents.

Besides feature extraction and classification, feature selection is also a key issue in pattern recognition. It reduces the size of the feature set and though reduces the computation time. Moreover, the reduction of the feature size prevents overfitting of the classifier and enhances the generalization of the trained classifier. Next, we will explain our approaches for extraction, evaluation, selection and classification of the features.

2.4.1. Feature extraction

Three groups of features were defined and evaluated in this study. These features have shown good performances in similar studies and hence, were believed to be useful as well for this application.

2.4.1.1. Morphological features. First group of features contains 17 morphologic features. These features were previously used by Kalatzis et al. in discriminating depressive patients from healthy controls using the P600 component of ERP signal [18]. These features are defined and calculated as follows (due to the characteristics and typical delay of cognitive components of ERP, the time interval was confined between 400 ms and 800 ms after stimulus):

(1) Latency (LAT, $t_{s_{\max}}$)—the ERP's latency time, i.e. the time where the maximum signal value appears:

$$t_{s_{\max}} = \{t | s(t) = s_{\max}\}, \tag{1}$$

where $s(t)$ is the ERP single trial during 400–800 ms after stimulus and $s_{\max}$ is the maximum signal value in this time interval.

(2) Amplitude (AMP, s_{max})—the maximum signal value:

$$s_{max} = \max\{s(t)\} \tag{2}$$

(3) Latency/amplitude ratio (LAR, $t_{s_{max}}/s_{max}$).
(4) Absolute amplitude (AAMP, $|s_{max}|$).
(5) Absolute latency/amplitude ratio (ALAR, $|t_{s_{max}}/s_{max}|$).
(6) Positive area (PAR, A_p)—the sum of the positive signal values:

$$A_p = \sum_{t=400\,\text{ms}}^{800\,\text{ms}} 0.5(s(t) + |s(t)|) \tag{3}$$

(7) Negative area (NAR, A_n)—the sum of the negative signal values:

$$A_n = \sum_{t=400\,\text{ms}}^{800\,\text{ms}} 0.5(s(t) - |s(t)|) \tag{4}$$

(8) Total area (TAR, A_{pn}):

$$A_{pn} = A_p + A_n \tag{5}$$

(9) Absolute total area (ATAR, $|A_{pn}|$).
(10) Total absolute area (TAAR, $A_{p|n|}$):

$$A_{p|n|} = A_p + |A_n| \tag{6}$$

(11) Average absolute signal slope (AASS, $|\bar{\dot{s}}|$):

$$|\bar{\dot{s}}| = \frac{1}{n} \sum_{t=400\,\text{ms}}^{800\,\text{ms}-\tau} \frac{1}{\tau}|s(t+\tau) - s(t)| \tag{7}$$

where τ is the sampling interval of the signal, n the number of samples of the digital signal, and $s(t)$ the signal value of the tth sample.

(12) Peak-to-peak (PP, pp):

$$pp = s_{max} - s_{min} \tag{8}$$

where s_{max} and s_{min} are the maximum and the minimum signal values, respectively:

$$s_{max} = \max\{s(t)\}, \qquad s_{min} = \min\{s(t)\} \tag{9}$$

(13) Peak-to-peak time window (PPT, t_{pp}):

$$t_{pp} = t_{s_{max}} - t_{s_{min}} \tag{10}$$

(14) Peak-to-peak slope (PPS, $\dot{s}_{pp}$):

$$\dot{s}_{pp} = \frac{pp}{t_{pp}} \tag{11}$$

(15) Zero crossings (ZC, n_{ZC})—the number of times t that $s(t) = 0$, in peak-to-peak time window:

$$n_{ZC} = \sum_{t=t_{s_{min}}}^{t_{s_{max}}} \delta_s \tag{12}$$

where $\delta_s = 1$ if $s(t) = 0$, 0 otherwise.

(16) Zero crossings density (ZCD, d_{ZC})—zero crossings per time unit, in peak-to-peak time window:

$$d_{ZC} = \frac{n_{ZC}}{t_{pp}} \tag{13}$$

where n_{ZC} are the zero crossings and t_{pp} is the peak-to-peak time window.

(17) Slope sign alterations (SSA, n_{sa})—the number of slope sign alterations of two adjacent points of the ERP signal:

$$n_{sa} = \sum_{t=400\,\text{ms}+\tau}^{800\,\text{ms}-\tau} 0.5 \left| \frac{s(t-\tau) - s(t)}{|s(t-\tau) - s(t)|} + \frac{s(t+\tau) - s(t)}{|s(t+\tau) - s(t)|} \right| \tag{14}$$

where τ is the sampling interval of the signal ($\tau = 3.9$ ms, for the sampling rate of 256 Hz).

2.4.1.2. *Frequency features*. The second group of defined features are three frequency characteristics of the signals. These features are *mode frequency*, *median frequency* and *mean frequency*, which are described and calculated as follows:

(1) Mode frequency:
f_{mode} is the frequency with the most energy content in the signal spectrum, so the maximum amplitude in the power spectrum density of the signal is at this frequency:

$$S(f_{mode}) = \operatorname*{Max}_f\{S(f)\} \tag{15}$$

S is the power spectral density of signal and f is frequency.

(2) Median frequency:
Median frequency (f_{median}) separates the power spectrum into two equal energy areas and is calculated from the following equation:

$$\int_0^{f_{median}} S(f)\,df = \int_{f_{median}}^{\infty} S(f)\,df \tag{16}$$

(3) Mean frequency:
Mean frequency (f_{mean}) represents the centroid of the spectrum and is calculated from the weighted averaging of the frequencies in the power spectral density of signal:

$$f_{mean} = \frac{\int_0^\infty f \cdot S(f)\,df}{\int_0^\infty S(f)\,df} \tag{17}$$

2.4.1.3. *Wavelet features*. Discrete wavelet coefficients were also used as a group of features which were extracted from the single trials. The idea of using of wavelet coefficients as features and their advantages were explained in detail in our previous paper [17].

In present study, a multi-resolution decomposition [19] was performed by applying a decimated discrete wavelet. Quadratic B-Spline functions were used here as mother wavelets due to their similarity with the evoked responses. These functions have been used in similar studies, and their efficiency for ERP decomposition and detection has been reported [20–23].

For the extraction of wavelet features, each single trial was decomposed into five-octaves using the wavelet transform. Six sets of coefficients (including residual scale) within the following frequency bands were obtained: 0–4 Hz, 4–8 Hz, 8–16 Hz, 16–32 Hz, 32–64 Hz and 64–128 Hz. The coefficients in each set are concerned with sequential time bands between 0 ms and 1000 ms. Because of previous filtering of the signal, coefficients within 30–128 Hz and 0–0.3 Hz ranges had not any useful information. But other coefficients represent the signal information in four frequency bands: (A) 0.3–4 Hz, (B) 4–8 Hz, (C) 8–16 Hz and (D) 16–30 Hz. Fig. 1 shows the decomposition and reconstruction of a single trial ERP into five octaves, using the quadratic B-Spline wavelet. As shown in Fig. 1(a), each octave contained the number of coefficients necessary to provide the relevant time resolution for that frequency range. The interpolation of coefficients by quadratic spline functions was used to reconstruct the analyzed signal within the particular band in Fig. 1(b). In the present study, 8 coefficients of band A, 8 coefficients of band B and 16 coefficients of band C were obtained for the post-stimulus epoch. The wavelet coefficients are designated with a letter corresponding to each frequency band ((A) for 0.3–4 Hz, (B) for 4–8 Hz, and (C) for 8–16 Hz) together with the coefficient number and two other numbers in brackets representing the approximate time window in millisecond (for example, A5 (450–580)).

2.4.2. *Feature evaluation*

The defined features were extracted from all probe sweeps of all subjects. Then a statistical analysis was applied for evaluating the fitness of each feature in discriminating between guilty and innocent subjects. For this reason, all 52 extracted features for the probe sweeps of all subjects were saved in a file and fed into the SPSS program. There were totally 2552 probe sweeps (1371 single sweep in guilty subjects + 1181 sweeps in innocent subjects). To determine the most relevant features differing between probes from guilty subjects (G-Probes) and probes from innocents (I-Probes), the features were subjected to independent sample *t*-test, with the subject type (guilty/innocent) as the grouping variable.

2.4.3. *Feature selection*

In any classification task, there is a possibility that some of the extracted features might be redundant. These features can increase the cost and running time of the system, and decrease its generalization performance. In this way, the selection of the best discriminative features plays an important role when constructing classifiers.

In this study, we employed a method using a genetic algorithm (GA) to identify the best subset of features for classification. GAs [24] are a class of robust problem-solving techniques based on a population of solutions, which evolve through successive generations by means of the application of three genetic operators: selection, crossover, and mutation [24]. GAs are suited to perform a search in huge search spaces, where other methods (local or gradient searches) cannot provide good results. The feature selection problem is one of these cases.

GA is based on genetic processes of biological organism. Over many generations, natural populations evolve according to the principles of natural selection and "survival of the fittest" [24]. GA requires a fitness or objective function, which provides a measure of performance of the population individuals. The fitness function must be relatively fast since GA incurs the cost of evaluating all the population in each generation.

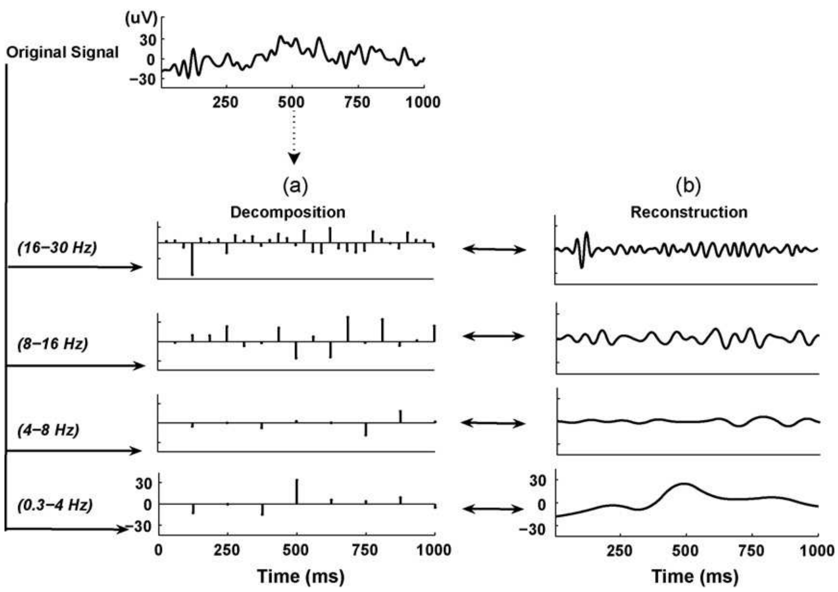

Fig. 1 – Wavelet decomposition of a single trial ERP: (a) the wavelet coefficients in the A (0.3–4 Hz), B (4–8 Hz), C (8–16 Hz) and D (16–30 Hz) bands. (b) The reconstructed signal in these bands.

Table 1 – GA parameters for feature selection.

Coding of genes	Binary coding
Population size	20
No. of genes	52
Reproduction	Tournament selection
Crossover	Two-point crossover
Crossover rate	0.5
Mutation	Random mutation
Mutation rate	0.01

In this study a binary representation of chromosomes was used for feature selection, where a 1 in the ith position of the binary vector means that the feature i is considered within the subset of selected features and a 0 in the jth position of binary vector means that feature j is not considered within the subset of selected features. Also, the detection rate of the classifier (which is explained in next section) was used as the fitness function. By running the GA over the features, the best subset of features for the classification of guilty subjects vs. innocent ones was selected. Table 1 gives the GA parameters used.

2.4.4. *Classification*

Features of probe stimuli were subjected to the linear discriminant analysis (LDA) [25,26]. The aim of LDA (also known as Fisher's LDA) is to use hyperplanes to separate the data representing the different classes. For a two-class problem, the class of a feature vector depends on which side of the hyperplane the vector is. This technique has a very low computational requirement which makes it suitable for many pattern recognition problems. Moreover this classifier is simple to use and generally provides good results. Consequently, LDA has been used with success in a great number of ERP and EEG processing researches such as motor imagery based Brain–Computer Interface (BCI) [27], P300 speller [28], asynchronous BCI [29] and P300 detection [23].

In this study, the performance of LDA algorithm was estimated using a leave-one-out jackknifing method. For each subject, the probe sweeps of other subjects with their real labels (G-probe/I-probe) were used as training data, and then the trained classifier applied on the probe sweeps of the given subject. Then the number of sweeps labeled as G-probe were counted and saved as N_G. N_G is an index of guilt for each subject and if it was greater than a predefined threshold (N_{th}), a guilty decision was made.

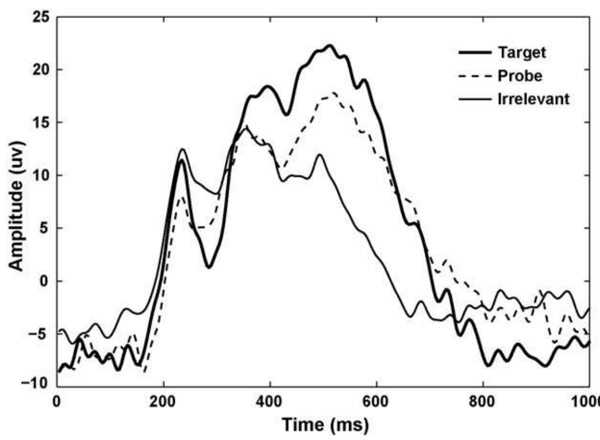

Fig. 2 – Superimposed grand averaged *P*, *T* and *I* responses in guilty group.

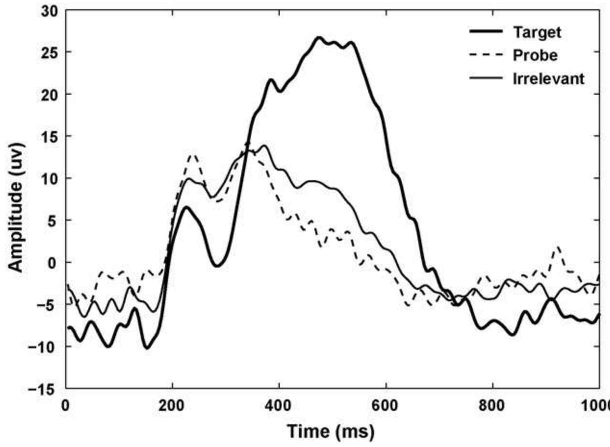

Fig. 3 – Superimposed grand averaged *P*, *T* and *I* responses in innocent group.

3. Results

3.1. *Time domain analysis*

Fig. 2 below shows grand averages in guilty group for superimposed probe, target and irrelevant responses. Also, Fig. 3 shows superimposed ERPs to *P*, *T*, and *I* items in the innocent group. The responses were as predicted. As can be seen in these figures, a large P300 was elicited by the target stimuli but not by the irrelevant stimuli. The probes elicited a P300 when they were relevant to subject's "crime", i.e. in guilty condition. A very small P300, if any, was elicited by probes when the subject was "innocent".

3.2. *Feature evaluation*

To determine the significance level of the difference of each feature's average in the two groups (guilty/innocent), the t-test was applied on the data. According to the results of the test, 29 of 52 defined features have significant difference between the two groups (p-value < 0.05). The results for the 12 most significant features are presented in Table 2. This table displays the t-values and significance levels (p-values) measured for the difference between guilties' and innocents' corresponding features.

A most remarkable group effect is observed in some area-based features (PAR and TAR) from morphological group and also in partially predominant fifth and forth A-band coefficients (A (500–625) and A (375–500)) from the wavelet group. These features were significantly larger for guilty subjects. Also all three frequency features were significantly larger for the innocent subjects.

Table 2 – Results of comparing probe response between guilty and innocent group using statistical *t*-test on features.

	Feature no.	Feature name	*t*-Value	*p*-Value
1	8	TAR	−15.55	2.10E−23
2	25	A (500–625)	−15.2	2.10E−23
3	7	NAR	−13.75	2.20E−23
4	6	PAR	−13.39	2.10E−23
5	24	A (375–500)	−11.61	2.10E−23
6	2	AMP	−7.71	1.40E−14
7	4	AAMP	−7.66	2.10E−14
8	19	f_median	6.7	2.30E−11
9	22	A (125–250)	5.72	1.10E−08
10	18	f_mode	5.43	5.90E−08
11	20	f_mean	5.4	7.10E−08
12	26	A (625–750)	−5.16	2.50E−07

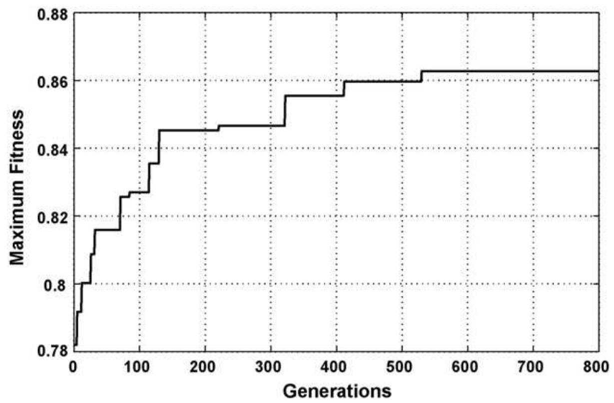

Fig. 4 – GA performance; the *x*-axis is the GA generations, and the *y*-axis the fitness function of the best Chromosome in each generation.

3.3. *Feature selection*

As explained before, a GA method was used for the selection of the best subset of features. Fig. 4 shows the maximum measure of the fitness function in each generation during the running of the GA. After producing 800 generations, the best Chromosome was obtained and its corresponding feature set was determined. This optimal feature set included 2 features from the morphological group, 2 features from the frequency group and 17 features from the wavelet group. These features are listed in Table 3.

Table 3 – Selected features from GA method.

Feature type	Selected features
Morphological	ALAR, SSA
Frequency	f_median, f_mean
Wavelet	A (0–125), A (250–375), A (375–500), A (500–625), A (875–1000), B (500–625), B (625–750), B (750–875), C (0–62), C (63–125), C (188–250), C (313–375), C (375–437), C (500–562), C (625–687), C (813–875), C (938–1000)

3.4. *Classification*

The selected 21 features were then used in the design of the classifier and in the analysis of the data as explained before. Fig. 5(a) displays the N_G parameter for all subjects. Solid circles belong to guilty subjects and empty circles to innocents. It was expected that solid circles would have large positive measures and in contrast empty circles would have small measures. If this happened, a suitable threshold could easily discriminate the groups. But in real state, the groups have overlaps and each selected threshold results in some misclassification.

For better inspection of the overlaps, histograms of these measures in two groups are drawn in Fig. 5(b). Histograms are drawn as percents in total number of the corresponding group. Statistical analysis reveals that N_G is significantly greater in the guilty group ($t = 8.3$, $p < 0.001$).

Also cumulative histograms are displayed in Fig. 5(c). The cumulative histograms can be used to estimate the effect of selected threshold on false/true detection rate. For example, a threshold level at 50% results in detection 67% of guilty subjects, but 8% of innocent subjects also detected as guilty. Obviously, a lower threshold will raise the percent of guilty true detection, but the certainty of the results will decrease and further number of innocents will be wrongly detected as guilty. If the importance of guilty detection and innocent detection are equal, concerning to Fig. 5(c), the best threshold that can give the best detection in the two groups, is near 45%. This threshold results in 86% correct detection of total subjects.

The performance of the method at all possible threshold may also be evaluated by its error-plot. In an error-plot diagram, the guilty against the innocent detection errors are plotted for all possible thresholds. An error-plot curve shows that by changing the threshold, we have a tradeoff between the guilty and innocent detection errors. An error-plot is also a two-dimensional depiction of classifier performance. To compare classifiers we may want to reduce error-plot performance to a single scalar value representing expected performance. A common method is to calculate the area under the error-plot curve which is roughly inversely proportional to the performance of the corresponding method. Fig. 6 displays the error-plot of the present implemented method.

4. Discussion

P300-based GKT, as a new method for the psychophysiological detection of concealed information, was tested on innocent and guilty subjects who were concealing information regarding a mock crime committed as a part of the experiment.

Figs. 2 and 3 indicate that P300 amplitude to probe stimuli in guilty subjects, are larger than those in innocents, which is consistent with our expectation and also with the previous studies of GKT by P300 [2,11,12].

The main purpose of this study was to evaluate the performance of a new classifying method in a P300-based GKT. This new method is an extension of our previously studied method [17].

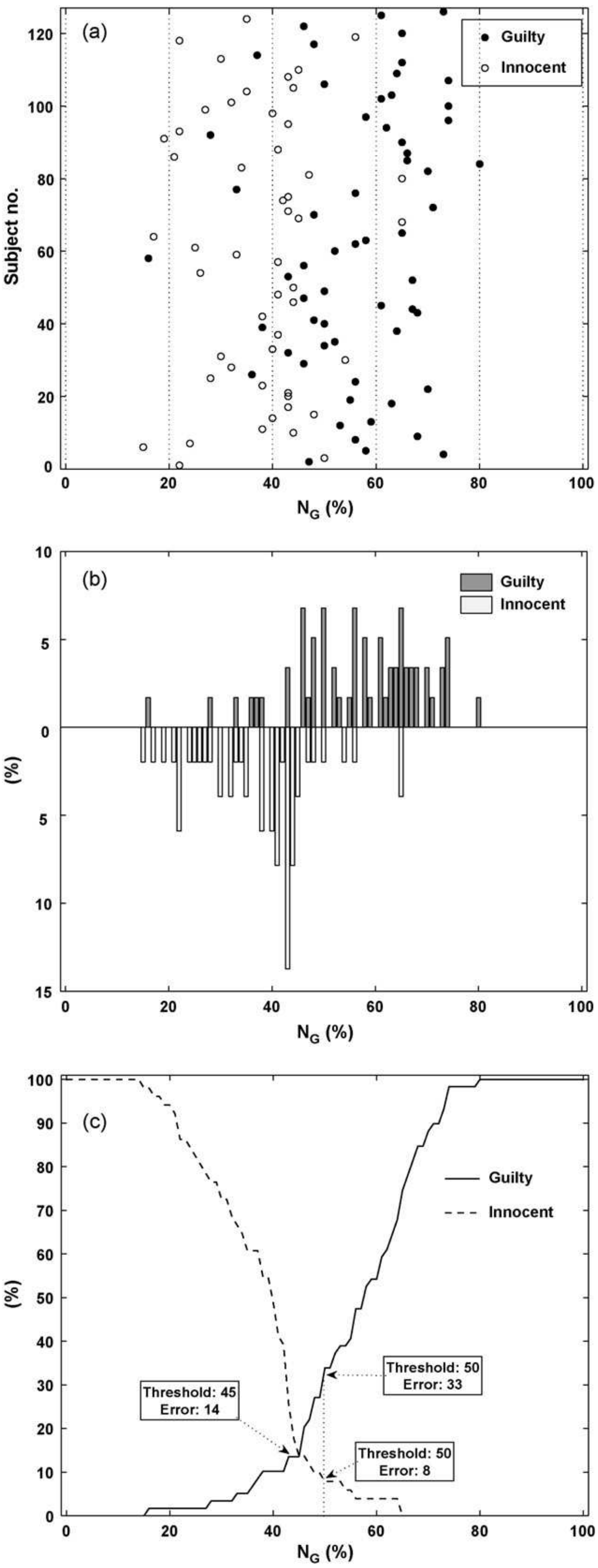

Fig. 5 – Classification results: (a) N_G parameter for all subjects, (b) superimposed histograms of N_G measures in guilty and innocent group and (c) superimposed cumulative histograms of N_G measures in guilty and innocent group.

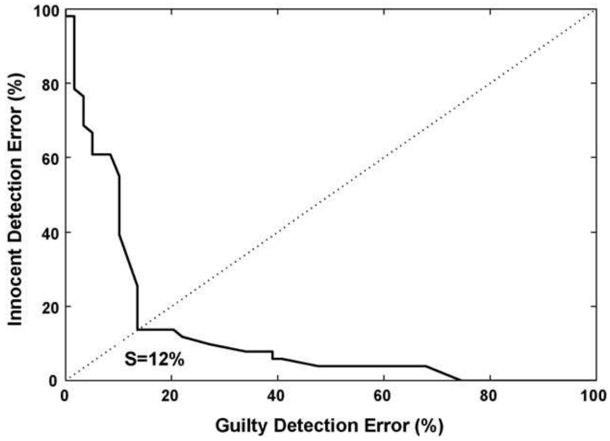

Fig. 6 – Error-plot of the implemented method.

In our present method, first a feature set consisting of three groups of features were defined and extracted from the raw data. Then the features corresponding to the probe stimuli were compared between guilty and innocent subjects, by statistical analysis. The most significant differences in the fifth and forth A-band coefficients (A (375–500) and A (500–625)), is in accordance with the characteristics of P300. These coefficients belong to the time domain of P300 component (Figs. 2 and 3) and therefore this result is in accordance to our expectation that probe stimuli result in larger P300 in guilty subjects than innocents. Also the significant increase in PAR and TAR features and the significant decrease in all frequency features for the guilty subjects are in accordance with the appearance of a positive peak with low frequency content (P300) in the brain response.

After feature evaluation, the optimal subset of the feature set was selected using a genetic algorithm. Then the selected features were used for the classification of subjects using the linear discriminant analysis. Concerning Fig. 5(c), by assuming equal importance for guilt and innocence detection, the method can detect 86% of the subjects correctly. The detection rates of the previous implemented methods [17] on the same data are displayed in Table 4. According to these measures, our recent method performs slightly better than BAD, BCD and the Wavelet classifier.

The methods may also be compared according to their error-plots. Fig. 7 displays the error-plots of the methods. The areas under the curves are reported in Table 4. According to these results, the present implemented method has the min-

Table 4 – Comparison of the present method with three previous implemented methods.

Method	Detection rate	Area under error-plot curve
BAD	74%	17%
BCD	80%	14%
Classifier (wavelet features)	79%	18%
Classifier (selected features)[a]	86%	12%

[a] Present method.

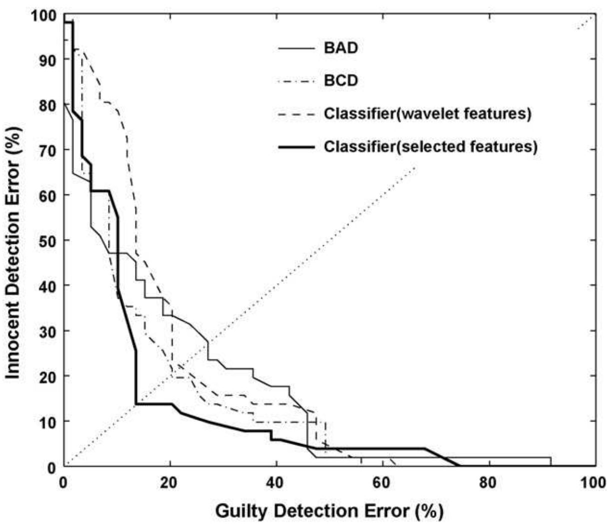

Fig. 7 – Superimposed error-plots of the implemented methods.

imum area among all the other methods (i.e. BAD, BCD and Wavelet Classifier) and thus has the best performance from this point of view.

The detection accuracies of applied methods are lower than what some other groups have reported [16]. However, it must be noted that the implemented methods were applied on different datasets. The obtained final accuracy depends on the subject type, used protocol, and the processing algorithm. So, the processing methods can be compared fairly only when the subjects and protocols are similar.

It should also be noted that such as conventional polygraphy, P300-based GKT is potentially vulnerable to countermeasures. (Countermeasures are covert actions taken by the subjects so as prevent detection by the GKT [30].) This study did not focus on any specific countermeasure and the subjects were permitted to use any covert action, providing that it did not disarrange the totality of the test and was not observable to the examiner. This matter is extensively investigated by some researchers [16,31]. The result of such researches can help in the design of more reliable paradigms for P300-based lie detection.

Conflict of interest

None declared.

Acknowledgements

This research was supported by the Research Center of Intelligent Signal Processing (RCISP). The authors would like to thank them and also all those who took part in this study.

REFERENCES

[1] National Research Council, The Polygraph and Lie Detection, National Academies Press, Washington, DC, 2002.

[2] L.A. Farwel, E. Donchin, The truth will out: interrogative polygraphy (lie detection) with event-related potentials, Psychophysiology 28 (1991) 531–547.

[3] J.P. Rosenfeld, J. Ellwanger, J. Sweet, Detecting simulated amnesia with event-related brain potentials, International Journal of Psychophysiology 19 (1995) 1–11.

[4] G. Ganis, S.M. Kosslyn, S. Stose, W.L. Thompson, D.A. Yurgelun-Todd, Neural correlates of different types of deception: an fMRI investigation, Cerebral Cortex 13 (2003) 830–836.

[5] F.A. Kozel, L.J. Revell, J.P. Lorberbaum, A. Shastri, J.D. Elhai, M.D. Horner, A. Smith, Z. Nahas, D.E. Bohning, M.S. George, A pilot study of functional magnetic resonance imaging brain correlates of deception in healthy young men, The Journal of Neuropsychiatry and Clinical Neurosciences 16 (2004) 295–305.

[6] F.A. Kozel, T.M. Padgett, M.S. George, A replication study of the neural correlates of deception, Behavioral Neuroscience 118 (2004) 852–856.

[7] D.D. Langleben, J.W. Loughead, W.B. Bilker, K. Ruparel, A.R. Childress, S.I. Busch, R.C. Gur, Telling truth from lie in individual subjects with fast event-related fMRI, Human Brain Mapping 26 (2005) 262–272.

[8] T.M.C. Lee, H.L. Liu, L.H. Tan, C.C.H. Chan, S. Mahankali, C.M. Feng, J. Hou, P.T. Fox, J.H. Gao, Lie detection by functional magnetic resonance imaging, Human Brain Mapping 15 (2002) 157–164.

[9] J.M. Nunez, B.J. Casey, T. Egner, T. Hare, J. Hirsch, Intentional false responding shares neural substrates with response conflict and cognitive control, Neuroimage 25 (2005) 267–277.

[10] K.L. Phan, A. Magalhaes, T.J. Ziemlewicz, D.A. Fitzgerald, C. Green, W. Smith, Neural correlates of telling lies: a functional magnetic resonance imaging study at 4 Tesla, Academic Radiology 12 (2005) 164–172.

[11] J.P. Rosenfeld, Event-related potentials in the detection of deception, malingering, and false memories, in: M. Kleiner (Ed.), Handbook of Polygraph Testing, Academic Press, New York, NY, 2002, pp. 265–286.

[12] J. Allen, W.G. Iacono, K.D. Danielson, The identification of concealed memories using the event-related potential and implicit behavioral measures: a methodology for prediction in the face of individual differences, Psychophysiology 29 (1992) 504–522.

[13] J. Polich, P300 in clinical applications, in: E. NiederMeyer, F. Lopes da Silva (Eds.), Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 4th edn, Williams and Wilkins, Baltimore, MD, 2000, pp. 1073–1091.

[14] J.A. Matte, Forensic Psychophysiology Using the Polygraph: Scientific Truth Verification—Lie Detection, J.A.M. Publications, Williamsville, New York, 1996.

[15] G. Ben-Shakhar, E. Elaad, The Guilty Knowledge Test (GKT) as an application of psychophysiology: future prospects and obstacles, in: M. Kleiner (Ed.), Handbook of Polygraph Testing, Academic Press, New York, NY, 2002, pp. 87–102.

[16] J.P. Rosenfeld, M. Soskins, G. Bosh, A. Ryan, Simple, effective countermeasures to P300-based tests of detection of concealed information, Psychophysiology 41 (2004) 205–219.

[17] V. Abootalebi, M.H. Moradi, M.A. Khalilzadeh, A comparison of methods for ERP assessment in a P300-based GKT, International Journal of Psychophysiology 62 (2006) 309–320.

[18] I. Kalatzis, N. Piliouras, E. Ventouras, C.C. Papageorgiou, A.D. Rabavilas, D. Cavouras, Design and implementation of an SVM-based computer classification system for discriminating depressive patients from healthy controls using the P600 component of ERP signals, Computer Methods and Programs in Biomedicine 75 (2004) 11–22.

[19] S. Mallat, A Wavelet Tour of Signal Processing, 2nd edn, Academic Press, San Diego, 1999.
[20] R.Q. Quiroga, Obtaining single stimulus evoked potentials with wavelet denoising, Physica D 145 (2000) 278–292.
[21] A. Ademoglu, E. Micheli-Tzanakou, Y. Istefanopulos, Analysis of pattern reversal visual evoked potentials by spline wavelets, IEEE Transactions on Biomedical Engineering 44 (1997) 881–890.
[22] R.Q. Quiroga, O.W. Sakowitz, E. Basar, M. Schurmann, Wavelet transform in the analysis of the frequency composition of evoked potentials, Brain Research Protocols 8 (2001) 16–24.
[23] V. Abootalebi, M.H. Moradi, M.A. Khalilzadeh, Detection of the cognitive components of brain potentials using wavelet coefficients, Iranian Journal of Biomedical Engineering 1 (2004) 25–46 (Article in Persian).
[24] D.E. Goldberg, Genetic Algorithm in Search, Optimization and Machine Learning, Addison-Wesley, Reading, MA, 1989.
[25] K. Fukunaga, Statistical Pattern Recognition, 2nd edn, Academic Press, New York, 1990.
[26] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Recognition, 2nd edn, Wiley–Interscience, New York, 2001.
[27] G. Pfurtscheller, EEG event-related desynchronization (ERD) and event-related synchronization (ERS), in: E. Niedermeyer, F. Lopes da Silva (Eds.), Electroencephalography: Basic Principles, Clinical Applications and Related Fields, 4th edn, Williams and Wilkins, Baltimore, MD, 2000, pp. 958–967.
[28] V. Bostanov, BCI competition 2003—data sets ib and iib: feature extraction from event-related brain potentials with the continuous wavelet transform and the *t*-value scalogram, IEEE Transactions on Biomedical Engineering 51 (2004) 57–61.
[29] R. Scherer, G.R. Muller, C. Neuper, B. Graimann, G. Pfurtscheller, An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate, IEEE Transactions on Biomedical Engineering 51 (2004) 979–984.
[30] C.R. Honts, S.L. Amato, Countermeasures, in: M. Kleiner (Ed.), Handbook of Polygraph Testing, Academic Press, New York, NY, 2002, pp. 251–264.
[31] M. Sasaki, S. Hira, T. Matsuda, Effects of a mental countermeasure on the physiological detection of deception using event related brain potentials, Shinrigaku Kenkyu 72 (2001) 322–328 (Article in Japanese).