An Improved P300-Based Brain-Computer Interface

Hilit Serby, Elad Yom-Tov, *Member, IEEE*, and Gideon F. Inbar, *Fellow, IEEE*

***Abstract*—A brain–computer interface (BCI) is a system for direct communication between brain and computer. The BCI developed in this work is based on a BCI described by Farwell and Donchin in 1988, which allows a subject to communicate one of 36 symbols presented on a 6 $\times$ 6 matrix. The system exploits the P300 component of event-related brain potentials (ERP) as a medium for communication. The processing methods distinguish this work from Donchin's work. In this work, independent component analysis (ICA) was used to separate the P300 source from the background noise. A matched filter was used together with averaging and threshold techniques for detecting the existence of P300s. The processing method was evaluated offline on data recorded from six healthy subjects. The method achieved a communication rate of 5.45 symbols/min with an accuracy of 92.1% compared to 4.8 symbols/min with an accuracy of 90% in Donchin's work. The online interface was tested with the same six subjects. The average communication rate achieved was 4.5 symbols/min with an accuracy of 79.5% as apposed to the 4.8 symbols/min with an accuracy of 56% in Donchin's work. The presented BCI achieves excellent performance compared to other existing BCIs, and allows a reasonable communication rate, while maintaining a low error rate.**

***Index Terms*—Brain–computer interface (BCI), independent component analysis (ICA), information bit rate, P300.**

I. Introduction

DURING the second half of the 20th century and especially during the last decade, a new capability has emerged by which the human brain can directly communicate with the environment. Unlike the common communication methods (e.g., speaking, writing, and indicating) that require motor ability, the new method of direct communication exploits the brain waves for communication and, therefore, does not require motor ability. Consequently, this new communication method is suitable for people that are incapable of any motor functions but still retain cognitive abilities; such people are locked in their own bodies and, therefore, cannot use any of the more traditional communication methods. Such a system is their only way to communicate with the outside world.

An application that exploits the electrical activity of the brain for communication is called a brain–computer interface (BCI). A BCI can be realized by either training the user to control his brain waves [1], [11], [12], [19], [20], [27], [39] or by exploiting natural responses of the brain to external stimuli [6], [13], [15]. The fact that a BCI that exploits natural responses does not require training of the user, along with the good performance such a BCI can achieve (as reported in [6]), motivated the use of such a BCI as a baseline for this work. The BCI reported in [6] achieved 20.1 b/min offline and 9.23 b/min online, compared to 2.65 b/min offline in [9], 5.64 b/min offline in [10], 3.46 b/min offline in [34], 2.35 b/min online in [25], 8.49 b/min online in [5], 9.475 b/min online in [12], and 10.88 b/min online in [18]. In our paper, the bit rates achieved are 23.75 b/min offline and 15.3 b/min online. Recent work includes an offline P300-BCI by Kaper *et al.* [23], [24], [26] that presents very promising results and an online BCI by Gao *et al.* [40] which achieved a transfer rate of 68 b/min, albeit on a single subject. A more detailed comparison between this work and other published works is given in Section IV.

Part of our lab work on electroencephalography (EEG) data [4], [7], [8], this work involves the development of a BCI that enables people to write text on a computer. The primary goal was to develop an interface that works online and facilitates reliable and rapid communication. An important advantage of the interface is the capability to adjust the communication rate to the user's ability while working.

The procedure used in this work for exploiting natural responses of the brain to external stimuli is a variation of a more general experiment known as the oddball paradigm. In the oddball experiment the subject is asked to distinguish between two stimuli, one common and one rare, by performing a mental count of one of the stimuli. In response to mentally counting the appearance of the rare stimulus, a typical potential is evoked in the brain. In the developed BCI, the responses were measured by noninvasive means, through three scalp electrodes ($\mathrm{C_Z}$, P_{Z} & $\mathrm{F_Z}$) referenced to the right mastoid. A 6 $\times$ 6 matrix that included all the alphabet letters as well as other useful symbols was presented to the user on a computer screen, see Fig. 1. The rows and columns of the matrix were intensified successively in a random order. At any given moment the user selected one of the letters or symbols that he wished to communicate, and maintained a mental count of the number of times the row and the column of the chosen symbol were intensified. In response to this counting, a potential was elicited in the brain. This response is known as a P300 wave, as first reported by Sutton *et al.* [35]. Detection of the responses and their timing in the measured signal made it possible to match the responses to one of the rows and one of the columns, and thus, the chosen symbol could be identified.

In addition to the desired response, the measured electrical activity from the scalp contains the ongoing activity of the brain; this activity is considered to be background noise by the interface. The power of this background noise is of the same order

Manuscript received January 12, 2003; revised October 12, 2004; accepted October 17, 2004.

H. Serby was with the Electrical Engineering Department, Technion-Israel Institute of Technology, Haifa 32000, Israel. She is now with the Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California at San Diego, La Jolla, CA 92093-0961 USA (e-mail: hilit@sccn.ucsd.edu).

E. Yom-Tov was with the Electrical Engineering Department, Technion-Israel Institute of Technology, Haifa 32000, Israel. He is currently with IBM's Haifa Research Laboratory, Haifa 31905, Israel.

G. F. Inbar is with the Electrical Engineering Department, Technion-Israel Institute of Technology, Haifa 32000, Israel.

Digital Object Identifier 10.1109/TNSRE.2004.841878

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
SPC	DOT	Y	Z	?	NEW LINE
1	2	3	4	5	EXIT

Fig. 1. Screen display as was shown to the subjects.

as that of the P300 waves. Other noise sources include EMG and the 50-Hz line noise. Thus, the main problem in developing a BCI based on P300 waves is to reliably discriminate the responses from the background noise. For this need, two signal processing methods were implemented. The first method was based on maximum likelihood (ML) classifier. The second method used independent component analysis (ICA) for blind source separation. The ICA was extensively used by S. Makeig and T-P Jung *et al.* for analysis and decomposition of P300 waves [33] as well as for the removal of eye artifacts from EEG data [37], [38]. The ICA method outperformed (on average) the ML method and, therefore, was chosen for online processing. The online interface adjusted itself to the user's performance. If the subject's responses were not clear enough, the system continued to present new trial repetitions. A decision was made only after a predefined threshold had been passed; this approach resembles the variable averaging technique used by Bayliss in her thesis [14]. We found that this flexibility is crucial to the achievable bit rate and proved to achieve substantially higher bit rates compared to the usage of any fixed number of repetitions. This important conclusion is discussed later in Section IV.

II. Methods

Six healthy subjects (three female), 27–33 years old, participated voluntarily in the study. All subjects were healthy university students with no known neurological problems. Each of the subjects gave consent to participate in the experiments.

The EEG was recorded at the $\mathrm{C_Z}$, P_Z & $\mathrm{F_Z}$ sites using the International 10–20 system Electro-Cap, referenced to an external right mastoid electrode, the right elbow used as ground. All electrodes were Ag/AgCl surface electrodes with a 6-mm diameter. None of the impedances between the cap electrodes and the reference electrode exceeded 5 kΩ. The three EEG channels were amplified using a 10 000 gain amplifier with a 0.01–40 Hz bandpass, and sampled at 250 Hz using a PCI-DASK 9118HG 12-b A/D sampling card. The data was recorded at a regular room.

During the experiment, subjects sat at a distance of 70 cm from the computer screen. The subjects viewed a 6×6 matrix display as shown in Fig. 1. The matrix contained 36 black characters on a white background. The experiment paradigm was implemented with a Visual Basic (VB) program. The experiment included intensifications of the matrix rows and columns. During intensification, the background of the intensified row or column turned black. Each row or column intensification lasted 125 ms, with no delay between successive intensifications. Two subjects participated in two additional experiments. In these experiments, a delay of 375 ms between the end of one intensification and the beginning of the successive intensification was added. Comparing the performance of the two subjects in both conditions indicated that the lowest overall durations for a correct decision were achieved when no breaks between successive intensifications existed.

The experiments were divided into two parts, an offline part and an online part. Each of the subjects participated in an offline experiment and later on in an online experiment. Two different signal processing methods were implemented on the offline data using Matlab. The method that showed superior performance, over the data from all the offline experiments, was chosen for online processing. The online BCI was developed only after all offline experiments were performed. The experiment paradigm was implemented with a VB program that also managed the sampling card, while Matlab was called from VB for online signal processing.

A. Offline Procedure

During the experiment, the matrix rows were successively intensified for a period of 125 ms each in random order. After the intensification of all six rows ended, a similar procedure of column intensification began. The overall time duration of such a trial was 1.5 s. The offline experiment was divided into ten blocks. Each block contained either 44 or 88 such trials, and lasted 66 or 132 s, respectively.

Before each experiment block began, the subject selected a target symbol from the matrix characters and informed us of the selection. The subjects were instructed to maintain a mental count of the number of times the row or column of the selected symbol were intensified. (For example, if the selected symbol was c, after 5 intensifications of the first row and 4 intensifications of the third column, the subject's count should be 9). According to Farwell and Donchin [22], such a paradigm will elicit a P300 only when a target's row (or column) intensification occurs.

The subjects were instructed to avoid blinking during the experiments. Later analysis of the offline raw data, over all subjects, revealed that less than 3% of the trials were contaminated with eye artifacts. To avoid biasing the decisions, maximum and minimum values for the features to be classified (the output of the matched filter) were set. No eye fixation instructions were given to the subjects.

Direct communication between the brain and the computer is achieved by applying signal processing methods to isolate the desired response (the P300) from the ongoing activity of the brain, and matching it with the row or column intensification that elicited it.

The determination of the target the subject chose is made in two steps: 1) deciding in which row the target should be; and 2) deciding in which column the target should be.

In order to identify which row (column) intensification elicited a P300, the data was divided into segments. Each sub segment that represented one row or column's intensification, included all data samples between 100 and 600 ms posterior to the beginning of the intensification. Each processing segment was comprised of six subsegments that correspond to the intensifications of the six different rows (columns) (resulting in a total duration of 1125 ms per segment).

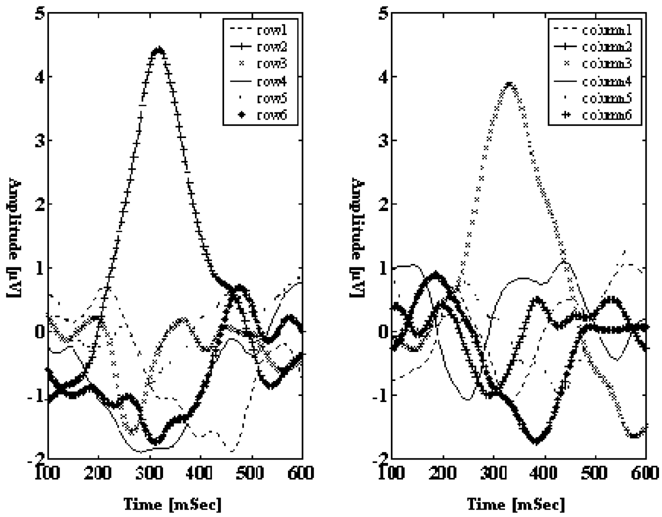

Fig. 2. Graphs presenting the averaged responses measured for the intensification of the six rows and for the six columns in the matrix during one experiment block (88 repetitions) of one of the subjects. For each of the rows and columns the responses were measured from 100 to 600 ms after the beginning of the specific row or column intensification.

At the end of each experiment block, graphs of the average responses to the different rows (columns) in that block were presented to the subject. Samples of such graphs are represented in Fig. 2.

B. Online Procedure

The online experiment was designed so that it would adapt to the subject's performance. The idea is that when given a significant response of the subject (for example when the P300 magnitude is high) or a response that can be easily isolated (the P300 can be easily discriminated from the ongoing EEG), less trial repetitions are needed in order to make a decision. The adaptive nature of the system allowed the number of repetitions until a decision on the row (column) is made to vary from one decision to the next. Therefore, the total elapsed time till a decision was made varied accordingly. This allowed maintaining a desired probability of success, while also allowing different subjects to reach different communication rates based on the quality of the signals they produced. Moreover, the communication rate a single subject achieved changed from one symbol to the next, while maintaining the probability of error.

The P300's potential intensity, shape, and latency vary from experiment to experiment for the same subject. These variations in the P300 are due to numerous causes, such as the subject's alertness and concentration, minor differences in exact electrode positions or in the distance and angle between the subject and the computer's screen. In order to reduce these variations, two additional offline experiment blocks were carried out before performing the online experiment. These two blocks were used to extract the current features of the subject's P300 potential, to be exploited in the online signal processing.

For convenience, the online experiment was divided into experiment blocks, in each of which five matrix symbols were communicated. Prior to each experiment block, the subject chose a set of five characters and informed us of his choice. As in the offline experiments the subjects were instructed to avoid blinking.

In order to take advantage of the fact that selection of a communicated symbol consists of two independent decisions (i.e., row and column), the online row and column intensification procedure was designed differently than the offline procedure. Instead of alternating between row and column intensifications, column intensifications began only after a decision on the row was made. This allowed maintaining the desired probability of success in both row and column decisions while adapting the number of required repetitions to the subject's performance in each part. For instance if the subject's performance for a specific communicated symbol was better for the row, a decision on the row could be reached after five intensifications of that row, even if a decision on the column could only be reached after 12 intensifications of the column. After a decision on a symbol was made (i.e., the decisions on both row and column were made) the system halted for 2 s to allow the subject to locate the next communicated symbol in the matrix. Fig. 3 illustrates the structure of an experiment block.

The online experiment included successive row (column) intensification trials. This could result in a situation where the last intensified row (column) in a trial is the same as the first intensified row (column) in the successive trial (which results in one intensification that is twice as long), ignoring the subject's reaction (P300) to the latter. In order to avoid this, the random intensification order was constrained not to allow successive intensifications of the same row (column).

Similar to the data processing in the offline procedure, the data was divided into segments, as described in the offline procedure. Each segment contained six (partially overlapping) subsegments; each subsegment included the data samples acquired for row (column) intensifications (samples from 100 to 600 ms posterior to the beginning of the given intensification). After all the data for a full segment had been collected, the data was transferred from the VB program to a Matlab script for online processing. The Matlab script processed the data and checked whether the information gathered resulted in passing the threshold, then notified the Visual Basic program if a decision was reached or additional data collection was needed. The threshold was designed to maintain a specific probability of success so that this BCI could be used as a stand-alone system. The minimum number of trials required to reach a decision with the required level of certainty was 3. In order to ensure a reasonable experiment block time, a decision was made after the 25th trial even if the threshold hadn't been reached.

At the end of each experiment block, the five symbols that were recognized by the BCI were displayed to the subject. As a result of the fast paced nature of this experiment, it seemed that displaying the chosen symbol immediately after a decision is made would only confuse the subjects that were busy in locating the next symbol (in the 2-s break that was given).

C. Processing Methods

1) ML Method: In this method the data was sampled from channels C_Z and P_Z. All the data was low-pass filtered (cutoff frequency—6 Hz) and the dc was removed. Feature extraction was performed using a matched filter (see [32]), which was uniquely constructed for each subject. Until this stage, data processing was performed on each of the channels separately. The features from both channels, for each data segment, were classified using an ML classifier [31], also unique for each of the

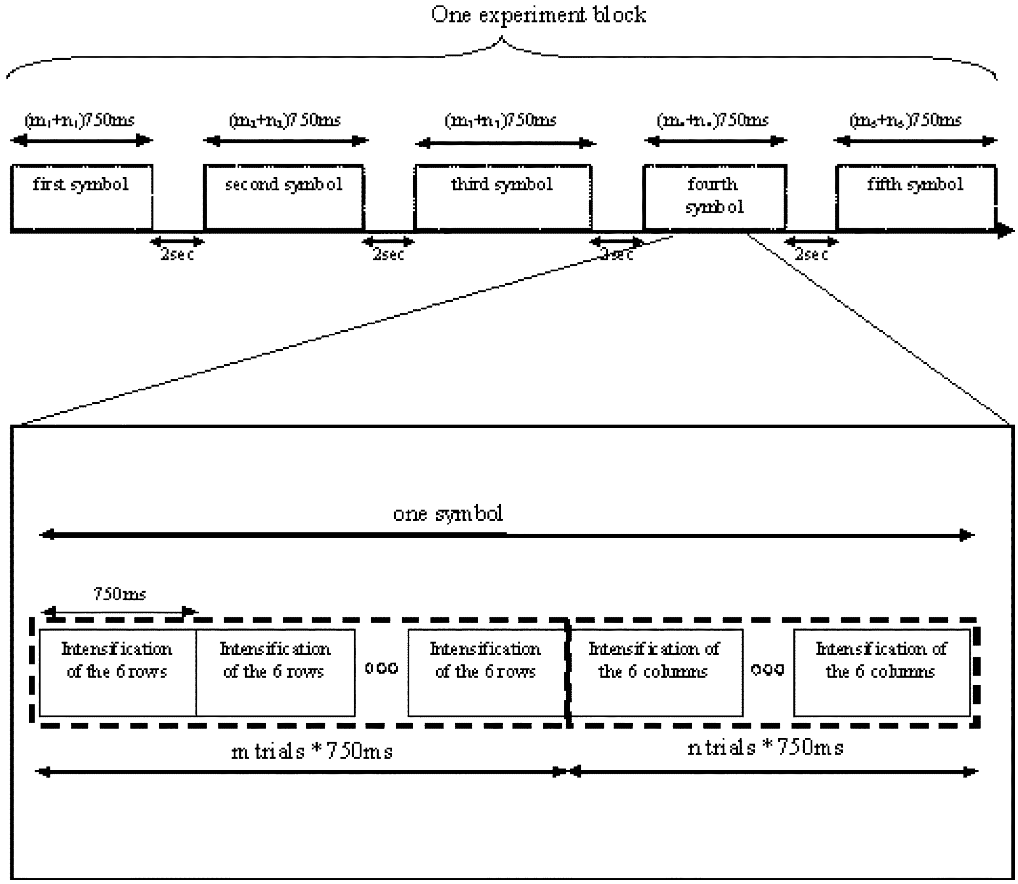

Fig. 3. Diagram of one experiment block. The block includes communication of five matrix symbols. A 2-s break between successive symbols is presented. The communication of one symbol includes m row intensification trials and n column intensification trials, when m and n depend on the subject's performance and, therefore, may vary from symbol to symbol.

subjects. The ML classifier was designed to separate the features into two classes, one class representing data that included P300 (data that resulted from intensifications of the target row or column), and the other class representing data that did not include P300 (data that resulted from intensifications of nontarget rows or columns). The *a priori* knowledge that only one of the six features should be classified to the first class was exploited to set thresholds for decision making. A set of threshold values that depended on the average number of trials was set. The same set of thresholds was used for all subjects. At the end of every trial, a decision was made if the corresponding threshold was reached; otherwise, an additional trial was performed.

2) ICA Method: In this method, the data from all three channels was processed. The ICA algorithm, a blind source separation technique (for algorithm details see [16] and [17]) was used to distinguish the P300s sources from the sources of the background noise (all signals other than P300 were considered background noise). As in the ML method, the data was low-pass filtered (cutoff frequency—6 Hz) and the dc was removed. In order to achieve better signal-to-noise ratio (SNR) the data was filtered using a matched filter, which was suited to the ICA source and designed uniquely for each subject. Uniform thresholds for all the subjects were used to decide on the target's row and column. The thresholds were based on the difference between the highest feature to the next highest feature, and depended on the number of trials averaged so far. The thresholds enabled the number of repetitions to vary from one decision to the other. The different stages of the processing algorithm are depicted in Fig. 4.

III. Results

In order to implement an online BCI that adapts to the user's performance, prior data acquisition and processing are required. Therefore, the BCI design consisted of two stages. The first stage included the development of the offline system, carrying out experiments, and using different methods to process the data. The data processing methods were aimed at achieving high accuracy levels, maintaining error probability of 5%–10%. The processing methods and the required thresholds for the online BCI were based on the results from the offline experiments.

A. Offline Results

All six subjects performed ten experiment blocks. Each block consisted of 44 or 88 trials each. (Each trial consisted intensification of all rows and columns of the matrix, which resulted in 12 intensifications per trial). Thresholds were set to achieve the aim of accuracy levels over 90%. The number of trials needed to reach a decision, varied according to the threshold requirements; whenever a threshold was reached a decision was made. After the thresholds were set, the performances achieved by both processing methods were compared on the basis of the accuracy achieved and the average symbol rate (determined by the average number of trials needed to reach a decision and the duration of each trial).

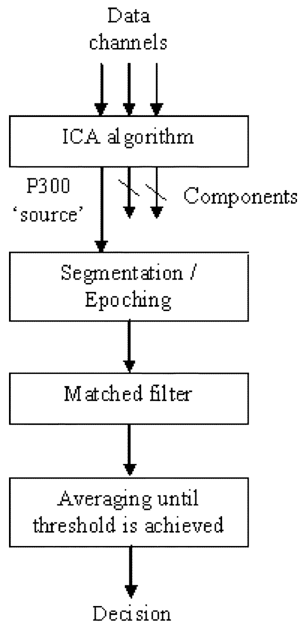

Fig. 4. Flowchart of the different stages of the ICA processing method. The three recorded data channels, from electrodes C_Z P_Z & F_Z, were the input of the ICA algorithm. The ICA algorithm returned three independent sources, one was associated with the P300 response and the other two were omitted. The ICA P300 source was segmented into overlapping segments, from 100 to 600 ms posterior to the beginning of any given intensification. Each such segment was passed through a matched filter, to give one feature that represents the maximum correlation between the segment and an average P300 template. Those six features (associated with the six rows or columns) were averaged over several intensifications until a threshold was reached. A decision on which row or column was selected by the user was based on the averaged features.

The results of the offline experiments revealed that for most subjects, the decision on the target's row was reached, on average, in fewer trials than the decision on the target's column. This might occurred since row intensifications preceded column intensifications, but was not further investigated. The average rates and the matching error percentages for each of the subjects for the ML and ICA processing methods are shown in Table I and Table II, respectively. Table III compares the averages for all subjects of the two processing methods.

Both processing methods achieved the aim of over 90% accuracy. The ML method achieved an accuracy of just above 90% with a communication rate of 4.19 symbols/min. The ICA processing method outperformed the ML method with an accuracy of 92.12% with a communication rate of 5.45 symbols/min. Note that if we inserted a break between successive intensifications allowing only one intensification to occur every 500 ms, then in order to obtain a rate of 5 symbols/min (less than the achieved 5.45 symbols/min) we had to rely on an average of no more than two repetitions per decision

B. Online Results

In order to maintain one compatible interface for all subjects (avoiding the need to calibrate it to a specific subject), one processing method had to be chosen. From the performance summary presented in Tables I and II, the difference between the error rates of the two methods is not significant, but the communication rates of the ICA method are significantly better ($p<0.0083$, using the one-tailed Wilcoxon matched pair test). As mentioned earlier, for most subjects the decision on the target's row was reached on average, in fewer trials than the decision on the target's column. The online system was designed taking the previous results from the offline experiments into account. The ICA processing method was chosen for the online signal processing and the decision on the row was completely independent of the decision on the column for each of the block symbols (i.e., column intensification began only after a decision on the row was reached).

The same six subjects that participated in the offline experiments took part in the online experiments. The online experiment was divided into blocks. Each experiment block included the transmission of a five-character word in which a 2-s break was given between successive characters to allow the subject to locate the next character of the communicated word. The experiment block began with row trials that continued until a threshold

TABLE I
AVERAGED OFFLINE RESULTS FOR A SYMBOL OBTAINED BY THE FIRST PROCESSING METHOD

Subject	Average rate (symbol/min)	Error percentage
1	5.43	3.6%
2	4.99	7.3%
3	3.9	4.2%
4	4.43	5.7%
5	3.2	13.5%
6	3.21	25.6%
Average	**4.19**	**9.98%**

TABLE II
AVERAGED OFFLINE RESULTS FOR A SYMBOL OBTAINED BY THE SECOND PROCESSING METHOD

Subject	Average rate (symbol/min)	Error percentage
1	7.03	3.3%
2	7.74	8.6%
3	5.35	15.3%
4	5.46	12.8%
5	2.89	0%
6	4.21	7.3%
Average	**5.45**	**7.88%**

TABLE III
COMPARISON BETWEEN THE TWO PROCESSING METHODS IN THE OFFLINE EXPERIMENTS AVERAGED OVER ALL SUBJECTS

	Rate (symbols/min)	Success percentage
First Method	4.19 (s.d. 0.84)	90.02% (s.d. 7.7%)
Second Method	5.45 (s.d. 1.62)	92.12% (s.d. 5.21%)

was reached. The number of trials until a decision was made depended on the subject's performance and was determined automatically during the experiment. Therefore, the average number of trials might have been different for different communicated symbols. Immediately after a decision on the row was reached, column trials began. The decision on the column was independent of the decision on the rows. (For example, assume it took 7 row trials for the system to reach a decision on the target's row. The number of column trials for the same target symbol could be 4 if the subject's performance on the column trials allowed making a decision). The number of row (column) trials per communicated symbol was limited to a minimum of 3 and a maximum of 25. All subjects were informed regarding the changes that had been made in the online procedure.

The results of the online experiment are summarized in Table IV. The average communication rate achieved was 4.51 symbols/min with a symbol accuracy of 79.5%. The 2-s break between consecutive characters was omitted in rate calculation. For all subjects except subject 3, the interface demonstrated reliable communication with reasonable communication rates. Omitting subject 3, the average accuracy level is 86.2% and the communication rate is 4.54 symbols/min.

TABLE IV
BCI RESULTS AS OBTAINED FROM ONLINE EXPERIMENTS

Subject	Average rate (symbol/min)	Success percentage
1	3.74	90%
2	6.87	92.5%
3	4.37	46%
4	4.53	90%
5	3.11	86.7%
6	4.44	72%
Average	**4.51 (s.d. 1.27)**	**79.5% (s.d. 18%)**

TABLE V
COMPARISON BETWEEN SEVERAL REPRESENTATIVE INTERFACES

Paper (first author name)	Online system	Rate [bits/minute]	Number of electrodes	Training time	Number of subjects (out of)
Donchin[6]	-	20.1	6	-	10 able
Donchin [6]	√	9.23	6	-	5(10) able
Babiloni [9]	-	2.65	9	-	5 able
Cincotti [10]	-	5.64	4	-	13 able
Levine [34]	-	3.46	16 - 126*	-	17 patients
Birbaumer [25]	√	2.35	5	Month	3 disable
Pfurtscheller [11]	√	6.3	-	7 sessions	3 able
Pfurtscheller [12]	√	9.48	2	Few weeks	4 patients
Wolpaw [18]	√	10.88	2	2 month	4(60) able
McFarland [5]	√	8.49	2-6	Few months	8 (2 disable)
Kaper [23]	-	47.26	10	-	8 able
This work	-	23.75	3	-	6 able

IV. DISCUSSION AND CONCLUSION

The reported results strengthen Donchin and Farwell [6], [22] findings that an online P300-based BCI is feasible. In a P300-based BCI data acquisition is performed by noninvasive means, with only a few scalp electrodes, unlike some other BCIs that require invasive means [2], [28]–[30], [34]. While some BCIs require intensive training over several months [18]–[20], [25], the P300 BCI requires no training. Simplicity and ease-of-use are obvious advantages of a P300-based BCI.

Besides the previous advantages, the P300 BCI seems to perform well compared to other BCIs (see Table V). The great variety of BCI's today, as presented in the first and second international meeting devoted to BCI [21], [36], requires a measure to evaluate all BCIs on a common basis. The following example of two different interfaces helps understand why a uniform performance measure is needed. Assume one interface that allows the subject to choose one symbol out of 36, as does the interface described in this work. The other interface may be a binary interface that allows the subject to choose one of two different objects. For the latter, in order to allow the same variety of symbols as in the 36-symbol interface, more than five binary decisions are needed to communicate a single symbol. For the simplicity of the comparison assume the number of possible symbols in the first interface is 32. A selection from the possible 32 symbols will be communicated with no error in the binary interface only if all five binary decisions are correct. Consequently, even if the binary decisions are 95% accurate, the accuracy per symbol will be only 77.4%. If the binary decisions are only 90% accurate the accuracy per symbol drops to 59%. Since a BCI is a communication system, the best measure to compare the performance of the different interfaces is the bit rate. The bit rate takes into account the number of possible selections (its importance was illustrated in the earlier example), the accuracy of the communication, and the time required to communicate each symbol. The bit rate R measures the achievable information rate per unit time, given the decision accuracy and duration. The number of achievable bits per decision is given by (see [21])

$$B = \log_2 N + P \log_2 P + (1-P) \log_2 \left[\frac{(1-P)}{(N-1)} \right]$$

where N is the number of possible selections (all selections are equally probable) and P is the accuracy probability. The bit rate R in bits/minute is given by $\mathrm{R} = BM$ where B is the number of bits per decision, and M is the average number of decisions per minute.

Evaluation of the bit rates achieved by some of the BCIs that were presented in the first international meeting devoted to BCI revealed the superiority of P300 BCI [6]. Table V compares between several representative works and this work. The table presents whether the system is online or offline, the systems' bit rate calculated using the previous equations, the number of electrodes, the duration of the training if any, and the number of subjects (out of the number of subjects that initially participated in the experiments). The electrodes in [34] (denoted with asterisk) are invasive electrodes, and their number was changed from patient to patient. In the work by McFarland [5], six subjects were able-bodied and the other two of the eight subjects were disabled. Besides the high rates Donchin's work achieved, it had another advantage: no user training was required. Thus, efforts to improve the P300-based BCI, in order to achieve better

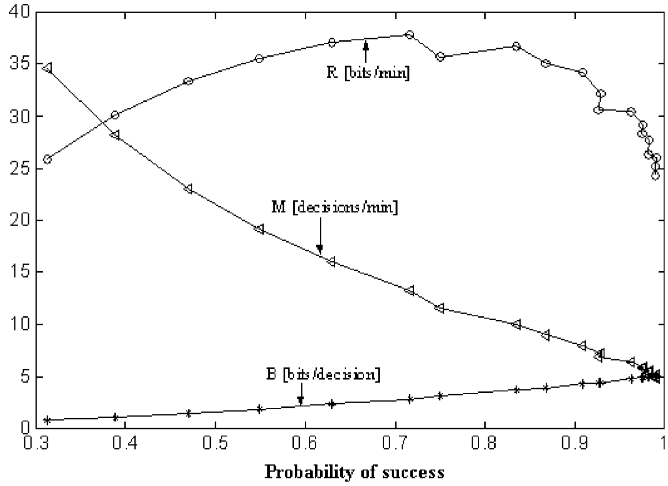

Fig. 5. Number of bits per decision—B, the number of the decisions per minute—M, and the resulted bit rate—R, versus the probability of success, based on the results of one subject.

rates and a real online BCI seemed to be a natural step that this study aimed to achieve. The work by Keeper *et al.* [23], [24], [26] is another offline P300 BCI which demonstrates the advantages of P300 BCI. In Kaper's work, the average transfer rate of 47.26 b/min corresponds to an accuracy of 44%, which is not feasible in a real BCI system. A more acceptable accuracy of 90% corresponds to an average transfer rate of 29.63 b/min, which is still a very impressive transfer rate. The most distinct difference between the three offline P300 BCIs is the signal processing methods that were used. Only Donchin's work [6] and this work present an online BCI. In both of these studies the number of possible selections is $N = 36$. The accuracy probability P and the amount of decisions per minute M varied from the offline to the online systems. In this study the offline system achieved $P = 0.92$ and $M = 5.45$ decisions/min resulting in a rate of $R = 23.75$ b/min. In [6], the offline system that was tested on ten able-bodied subjects, achieved $P = 0.9$ and $M = 4.8$ decisions/min resulting in a rate of $R = 20.1$ b/min. (The other reported case in [6] where $P = 0.95$ and $M = 4.3$ decisions/min results in a lower rate of $R = 19.9$ b/min). The offline results in this study show a slight improvement compared to the offline system described in [6]. In [6], it is mentioned that due to technical considerations a pause of 1 s after each 1.5-s trial was forced. Omitting this pause from the bit-rate calculations clearly prevents comparing both interfaces on the same basis. Applying this pause helps discriminate the P300 response, since intertrial interference is completely avoided. If we take the 1-s pauses into account the actual bit rate for the interface described in [6] is $R = 12.1$ b/min. Obviously the comparable rate achieved in [6], if no technical consideration were needed, could be more than 12.1 b/min, but probably less than 20.1 b/min.

The online system in this study adjusts to the user's performance; the more significant the P300s are, the less repetitions the user encounters and less time is required to communicate each symbol. The online system was tested on the same six subjects and achieved $P = 0.795$ and $M = 4.5$ decisions/min, thus achieving a rate of $R = 15.3$ b/min. Unlike this BCI, in [6], the number of trial repetitions had been fixed prior to the experiment and was set to the number of repetitions the subject needed for achieving 90% accuracy according to its offline results. The average number of repetitions in the online experiment is not mentioned, and only five out of the ten subjects performed the experiment. Therefore, it is impossible to evaluate the exact communication rate the online experiments in [6] achieved. Assuming that the five subjects that participated in the online experiment achieved an average accuracy of $P = 0.56$ possess the same average of $M = 4.8$ decisions/min (the average for all ten subjects participating in the offline experiments), the rate this interface achieved was in the range $5.54 < R < 9.23$ b/min. While the online interface of this study achieves almost 80% accuracy and therefore facilitates recognizable and understandable communication on which a stand-alone system can be based, the BCI described in [6] achieves only 56% accuracy, not allowing understandable communication without implementing necessary modifications.

As mentioned earlier, the bit rate, R, is a function of B, the average number of bits per decision, and M, the average number of decisions per minute. B is a function of N, the number of possible selections, and P, the accuracy probability. Since N is a fixed number (36), the bit rate R is a function of P and M solely. Since M and P are inversely correlated, the optimum bit rate R can be calculated from the equations above. M, B, and R as a function of P, as measured for one of the subjects, are presented in Fig. 5. Fig. 5 demonstrates the fact that the bit rate, R, is not at its maximum when P equals 1. R reaches its maximum at some arbitrary point that depends jointly on M and B.

In order to achieve the maximum bit rate, the BCI should have been designed for lower success probabilities. However, the main goal of this study was not to achieve maximum bit rate, but to develop a BCI that reaches high success probability along with reasonable communication rates and can be used as is. Therefore, the importance of reliable communication outweighs the need for fast communication rate. Thus, this BCI was designed to reach a goal of 90% accuracy, assuming this would allow understandable communication with no need for further coding and processing. To attain this goal, the signal processing method used in the online BCI included three stages that increase the SNR. The first stage used an ICA algorithm for source separation, the second stage included filtering the processed data using a matched filter, and the third stage included averaging trials till a threshold was reached. The cost of averaging over more and more trials is a decrease in the decision rate.

In order to estimate the contribution of the ICA algorithm to the discrimination of the P300 from the background noise, we calculated the ratio between the average energy of trials that were expected to contain P300 and the average energy of the rest of the trials. The ratio was averaged over all offline data of all subjects and was compared to the same ratio when the ICA algorithm was not used. Instead of the ICA data, the mean of the three channels' data was taken. The comparison results showed an increase of 20% in the ratio when using the ICA algorithm. This significant increase in the discrimination ability enabled an improvement in the BCI performance. In the work of Makeig *et al.* [33] and Jung *et al.* [37], [38] the ICA algorithm was used on ERP data recorded from more than 30 electrodes. Increasing the number of electrodes used by the BCI might improve its discrimination ability, for example by exploiting the spatial distribution of the P300 wave and incorporate it into the

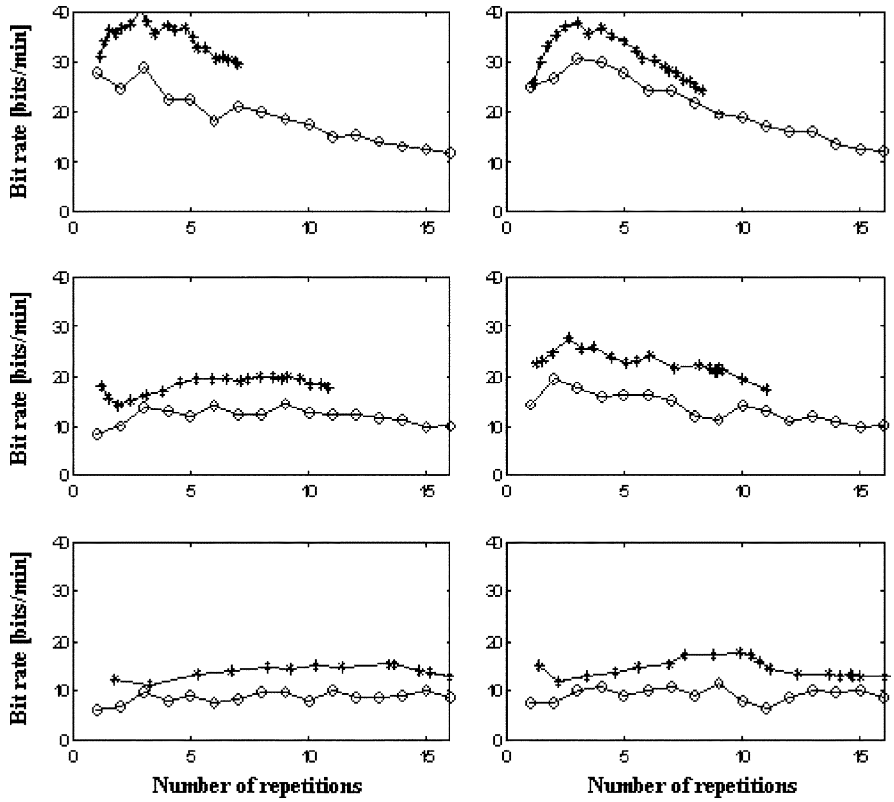

Fig. 6. Bit rate R in bits/minute, as a function of the number of trial repetitions. The lines denoted by circles represent a fixed number of iterations while the ones with stars represent the case when the number of repetitions depends on the thresholds. The six graphs display the results from the offline experiment for each of the six subjects.

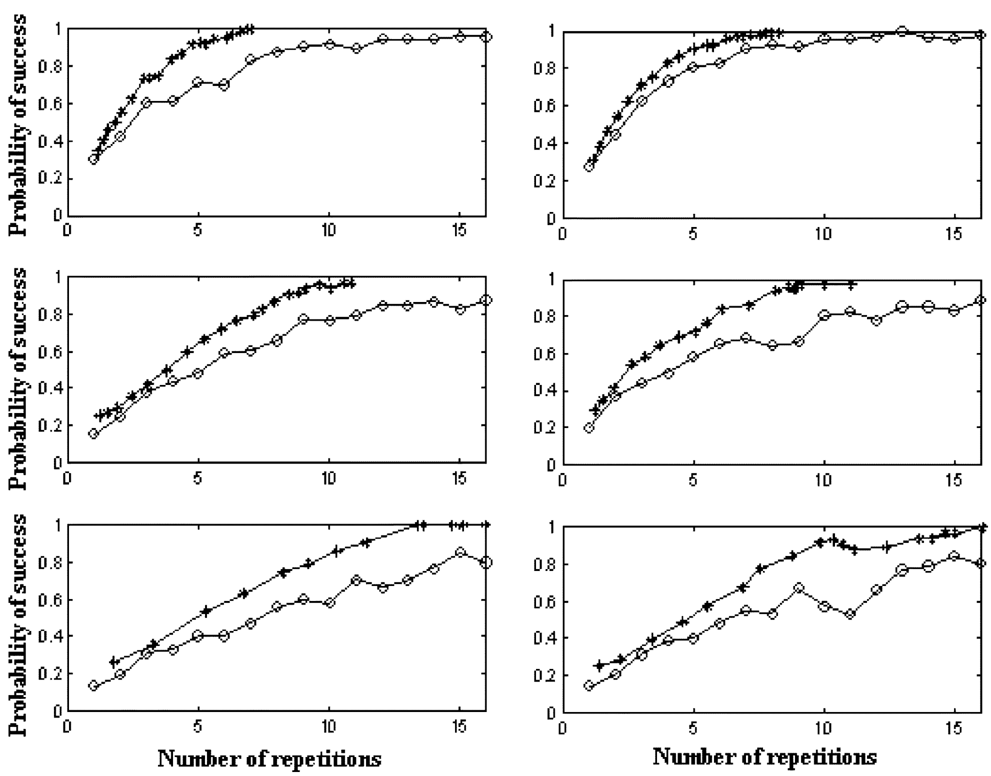

Fig. 7. The probability of success, Ps, as a function of the number of trial repetitions. The lines denoted by circles represent a fixed number of iterations while the ones with stars represent the case when the number of repetitions depends on the thresholds. The six graphs display the results from the offline experiment for each of the six subjects.

decision making. Unlike the work of Makeig *et al.*, that aimed to decompose the P300 into separate components, this BCI only aimed to distinguish between trials containing P300 and other trials. Therefore, this BCI could be settled for a lower number of electrodes and thereby making the interface less cumbersome. Since the more electrodes used the more robust the ICA algorithm becomes, this option should be considered, especially in noisy environments.

There are two different approaches to averaging over trials for making a decision. The straight forward way is to average over a constant number of trials before reaching a decision. This number of trials should be adjusted to the subject's average performance in order to achieve the desired accuracy. The other approach notes that the subject's performance varies rapidly. When the subject's performance is at its best, allowing reaching a decision based on fewer trials will not decrease the accuracy. On the other hand, when the subject's performance is poor, maintaining the desired accuracy may require many trials. Thresholds enable reaching a decision based on performance (gained certainty) rather than a fixed number of trials. A similar approach of using thresholds to determine whether to average the current trial with a subsequent trial or not was used by Bayliss [14]. Applying thresholds to the system, results in varying number of trials before making a decision. We compared the performance achieved with both approaches. For the first we calculated the probability of success P and the bit rate R as a function of the number of trials. For the second, we calculated P, R and the average number of trials as a function of the threshold. R and P as a function of the number of trials are presented in Figs. 6 and 7, respectively. The results show that the use of thresholds results in inherently better performance ($P<0.0001$ when comparing the probability of success of the two approaches for the same number of repetitions). For any given number of repetitions the threshold approach outperforms the fixed approach for all subjects. The average maximum bit rate in the threshold technique is 29% higher than the average maximum bit rate when the number of repetitions is fixed.

To further improve the performance of P300-based BCIs, additional work should be considered, whether in speeding up the communication rate or in increasing the accuracy level. Embedding software for spell check and word completion in the BCI as in [22], [25] should be considered, or perhaps using one of the symbols in the matrix as a delete function. One can also take advantage of the different occurrence probabilities of letters as well as letter pairs in the language (e.g., allowing the intensification probability to differ from letter to letter to match the occurrence probability in the language as used in Huffman codes). Other paradigms along with other kinds of displays should be tested too, see the work by Allison *et al.* [3]. And above all, using different or additional signal processing methods to overcome discrimination mistakes. Since this BCI is designated for locked-in patients, before seeking further improvement, one should test this kind of BCI on these patients. Unfortunately, this research involved only healthy subjects.

Acknowledgment

The authors would like to thank Prof. H. Pratt for his useful advice and remarks and for his assistance in this study.

References

[1] A. Kostov and M. Polak, "Parallel man-machine training in development of EEG-based cursor control," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 203–205, June 2000.

[2] B. Graimann, J. E. Huggins, S. P. Levine, and G. Pfurtscheller, "Toward a direct brain interface based on human subdural recordings and wavelet-packet analysis," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 954–962, June 2004.

[3] B. Z. Allison and J. A. Pineda, "ERP's evoked by different matrix sizes: implications for a brain computer interface (BCI) system," *IEEE Trans. Neural. Syst. Rehab. Eng.*, vol. 11, no. 2, pp. 110–113, June 2003.

[4] D. H. Lange, H. Pratt, and G. F. Inbar, "Modeling and estimation of single evoked brain potential components," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 791–799, Sep. 1997.

[5] D. J. McFarland, W. A. Sarnacki, and J. R. Wolpaw, "Brain-computer interface (BCI) operation: Optimizing information transfer rates," *Biol. Psych.*, vol. 63, pp. 237–251, 2003.

[6] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a P300-based brain-computer interface," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 174–179, June 2000.

[7] E. Yom-Tov, A. Grossman, and G. F. Inbar, "Movement-related potentials during the performance of a motor task 1: The effect of learning and force on the movement-related potentials," *Biol. Cybern.*, vol. 85, pp. 395–399, 2001.

[8] ——, "Movement-related potentials during the performance of a motor task 2: Cerebral areas activated during learning of the task," *Biol. Cybern.*, vol. 85, pp. 387–394, 2001.

[9] F. Babiloni, F. Cincotti, L. Lazzarini, and M. G. Marciani, "Linear classification of low-resolution EEG patterns produced by imagined hand movements," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 186–188, June 2000.

[10] F. Cincotti, D. Mattia, C. Babiloni, F. Carducci, S. Salinari, L. Bianchi, M. G. Marciani, and F. Babiloni, "The use of EEG modifications due to motor imagery for brain-computer interfaces," *IEEE Trans. Neural. Syst. Rehab. Eng.*, vol. 11, no. 2, pp. 131–133, June 2003.

[11] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlögl, B. Obermaier, and M. Pregenzer, "Current trends in Graz brain–computer interface (BCI) research," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 216–219, June 2000.

[12] G. Pfurtscheller, C. Neuper, G. R. Müller, B. Obermaier, G. Krausz, A. Schlögl, R. Scherer, B. Graimann, C. Keinrath, D. Skliris, M. Wörtz, G. Supp, and C. Schrank, "Graz-BCI: State of the art and clinical applications," *IEEE Trans. Neural. Syst. Rehab. Eng.*, vol. 11, no. 2, pp. 177–180, June 2003.

[13] J. B. Polikoff, H. T. Bunnell, and W. J. Borkowski, "Toward a p300-based computer interface," in *Proc. RESNA Annu. Conf. RESNAPRESS* , Arlington, VA, 1995, pp. 178–180.

[14] J. D. Bayliss, "A flexible brain-computer interface," Ph.D. dissertation, Computer Science Dept., Univ. Rochester, Rochester, NY, 2001.

[15] J. D. Bayliss and D. H. Ballard, "A virtual reality testbed for brain-computer interface research," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 188–190, June 2000.

[16] J. F. Cardoso, "High-order contrasts for independent component analysis," *Neural Comput.*, vol. 11, pp. 157–192, Jan. 1999.

[17] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proceedings-F*, vol. 140, pp. 362–370, Dec. 1993.

[18] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An EEG-based brain-computer interface for cursor control," *Electroencephalogr. Clin. Neurophysiol.*, vol. 78, pp. 252–259, 1991.

[19] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain-computer interface research at the Wadsworth center," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 222–226, June 2000.

[20] J. R. Wolpaw, D. J. McFarland, T. M. Vaughan, and G. Schalk, "The Wadsworth center brain-computer interface (BCI) research and development program," *IEEE Trans. Neural. Syst. Rehab. Eng.*, vol. 11, no. 2, pp. 204–207, June 2003.

[21] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: A review of the first international meeting," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 164–173, June 2000.

[22] L. A. Farwell and E. Donchin, "Talking of the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1988.

[23] M. Kaper and H. Ritter, "Generalizing to new subjects in brain-computer interfacing," in *Proc. 26th IEEE EMBS Annu. Int. Conf. (EMBC)*, Sep. 2004, pp. 1073–1076.
[24] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "BCI competition 2003-data set IIb: Support vector machines for the P300 speller paradigm," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1073–1076, June 2004.
[25] N. Birbaumer, A. Kübler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor, "The thought translation device (TTD) for completely paralyzed patients," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 190–193, June 2000.
[26] P. Meinicke, M. Kaper, F. Hoppe, M. Heumann, and H. Ritter, "Improving transfer rates in brain computer interfacing: A case study," *Neural Information Processing Systems (NIPS)*, pp. 1107–1114, 2002.
[27] P. Sykacek, S. Roberts, M. Stokes, E. Curran, M. Gibbs, and L. Pickup, "Probabilistic methods in BCI research," *IEEE Trans. Neural. Syst. Rehab. Eng.*, vol. 11, no. 2, pp. 192–195, June 2003.
[28] P. R. Kennedy and K. D. Adams, "A decision tree for brain-computer interface devices," *IEEE Trans. Neural. Rehab. Syst. Eng.*, vol. 11, no. 2, pp. 148–150, June 2003.
[29] P. R. Kennedy, R. A. E. Bakay, M. M. Moore, K. Adams, and J. Goldwaithe, "Direct control of a computer from the human central nervous system," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 198–202, June 2000.
[30] R. E. Isaacs, D. J. Weber, and A. B. Schwartz, "Work toward real-time control of a cortical neural prosthesis," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 196–198, June 2000.
[31] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2000, pp. 85–89.
[32] S. Haykin, *Communication Systems*, 2nd ed. New York: Wiley, 1983, pp. 530–534.
[33] S. Makeig, M. Westerfield, T. P. Jung, J. Covington, J. Townsend, T. J. Sejnowski, and E. Courchesne, "Functionally independent components of the late positive event-related potential during visual spatial attention," *J. Neurosci.*, vol. 19, pp. 2665–2680, Apr. 1999.
[34] S. P. Levine, J. E. Huggins, S. L. BeMent, R. K. Kushwaha, L. A. Schuh, M. M. Rohde, E. A. Passaro, D. A. Ross, K. V. Elisevich, and B. J. Smith, "A direct brain interface based on event-related potentials," *IEEE Trans. Rehab. Eng.*, vol. 8, no. 2, pp. 180–185, June 2000.
[35] S. Sutton, M. Braren, J. Zubin, and E. R. John, "Information delivery and the sensory evoked potential," *Science*, vol. 155, pp. 1436–1439, 1965.
[36] T. M. Vaughan, W. J. Heetderks, L. J. Trejo, W. Z. Rymer, M. Weinrich, M. M. Moore, A. Kubler, B. H. Dobkin, N. Birbaumer, E. Donchin, E. W. Wolpaw, and J. R. Wolpaw, "Brain-computer interface technology: A review of the second international meeting," *IEEE Trans. Neural. Syst. Eng.*, vol. 11, pp. 94–109, June 2003.
[37] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials," *Hum. Brain Map.*, vol. 14, pp. 168–185, 2001.
[38] ——, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clin. Neurophysiol.*, vol. 111, pp. 1745–1758, 2000.
[39] W. D. Penny, S. J. Roberts, E. A. Curran, and M. J. Stokes, "EEG-based communication: a pattern recognition approach," *IEEE Trans. Rehab. Eng.*, vol. 8, pp. 214–215, June 2000.
[40] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Trans. Neural. Syst. Rehab. Eng.*, vol. 11, no. 2, pp. 137–140, June 2003.

Hilit Serby received the B.Sc. and M.Sc. degrees, both in electrical engineering, from the Technion—Israel Institute of Technology, Haifa, Israel, in 1999 and 2002, respectively.

She is currently with the Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California at San Diego, La Jolla. Her research interests include signal processing and clustering with respect to brain functionality and dynamics.

Elad Yom-Tov (S'96–M'02) received the B.Sc. degree in electrical engineering from Tel-Aviv University, Tel-Aviv, Israel, in 1997, and the M.Sc. and Ph.D. degrees, both in electrical engineering, from the Technion—Israel Institute of Technology, Haifa, in 1999 and 2002, respectively.

He is currently with IBM's Haifa Research Laboratory, Haifa, where he works in the field of machine learning. He is the author (with D. Stork) of the *Computer Manual in MATLAB to Accompany Pattern Classification, Second Edition* (New York: Wiley, 2004). His research interests include machine learning and biomedical signal processing, with an emphasis on the processing of movement related potentials.

Gideon F. Inbar (S'63–M'64–SM'87–F'90) received the B.Sc. degree from the Technion—Israel Institute of Technology, Haifa, in 1959, the M.Sc. degree from Yale University, New Haven, CT, in 1963, and the Ph.D. degree from the University of California, Davis, in 1969, all in electrical engineering.

In 1970, he joined the Department of Electrical Engineering, the Technion, where he is a Professor Emeritus and holds the Otto Barth Chair in Biomedical Sciences. He is the Editor of *Signal Analysis and Pattern Recognition in Biomedical Engineering* (New York: Wiley, 1975) and Co-Editor of *Advances in Processing and Pattern Analysis of Biological Signals* (New York: Plenum, 1996). His major research interests are in the areas of brain functions and movement control, biological cybernetics, and biological signal processing.