ABSTRACT
We present an overview of our Brain-computer interface (BCI) research, invasive as well as non-invasive, during the past four years. The invasive BCIs are based on local field and action potentials recorded with microelectrode arrays implanted in the visual cortex of the macaque monkey. The non-invasive BCIs are based on electroencephalogram (EEG) recorded from a human subject’s scalp. Several EEG paradigms were used to enable the subject to type text or to select icons on a computer screen, without having to rely on one’s fingers, gestures, or any other form of motor activity: the P300 event-related potential, the steady-state visual evoked potential, and the error related potential. We report on the status of our EEG BCI tests on healthy subjects as well as patients with severe communication disabilities, and our demonstrations to a broad audience to raise the public awareness of BCI.

Categories and Subject Descriptors
H.4 [Information Systems Applications]: Miscellaneous

General Terms
Design, Experimentation

Keywords
Brain-computer interface, local field potentials, action potentials, electroencephalogram

1. INTRODUCTION
Brain-Computer Interface (BCI) research has attracted a lot of attention in recent years. It allows a subject to communicate with the environment from brain signals directly, without speech or any other form of muscular activity. The results of BCI research, which is at the intersection of different scientific disciplines, is expected to substantially contribute to rehabilitation and restoration of lost communication abilities of patients suffering from amyotrophic lateral sclerosis, stroke, brain/spinal cord injury, cerebral palsy, muscular dystrophy, etc. The initial successes have attracted attention not only in the scientific community, but also in the popular media (see, for example, the movie "Surrogates" (2009), and episode 5.19 in the series “House MD” (2009)).

Any BCI consists of the following components: equipment for recording brain signals, a decoder, and an external device, usually a display on which visual feedback is given to the subject, or an actuator, such as a wheelchair. In general, BCIs are either invasive [1] or noninvasive [2]. The invasive BCIs rely on recordings made with electrodes implanted in the brain (e.g., local field potentials and action potentials) or with electrodes placed on the cortical surface (electro-corticogram), whereas the noninvasive ones mostly rely on electroencephalograms (EEGs) recorded from the subject’s scalp.

2. INVASIVE BCI RESEARCH
Invasive BCIs are usually based on electrodes implanted in the motor cortex, and use regressors and classifiers for decoding the recorded signals, e.g., for tracking the arm (or hand) position and for controlling prosthetic devices [3]. BCIs using implants in the visual cortex have received much less attention. To bridge this gap, we targeted our invasive research towards the decoding of recordings made in the monkey’s visual cortex.

The visual perception of objects is assumed to start with the detection of low-level features such as color, orientation, curvature, and so on. Orientation plays a particularly important role since it is implicated in the extraction of the contour and 2D shape of visual objects. There is a wealth of evidence that the visual cortex consists of orientation selective nerve cells or neurons [4]. The evidence served as the basis of population codes for the information representation in the brain [5]. The development of discriminative brain signal features and the application of powerful machine learning techniques could lead to the construction of a low-level visual decoder. Additionally to the decoding itself, BCI systems also generate feedback to the subject, which in turn, given the presence of perceptual learning, could lead to better discriminative brain signals to which the decoder should be adjusted.

In our research, we constructed invasive BCIs from local field potentials (LFPs), reflecting the summation of numerous overlapping electrical fields generated by hundreds of
neurons [6]) or from action potential (spikes), reflecting the activity of a single neuron or of a few neurons (multiunit activity). Both were recorded chronically with a Utah array consisting of a row and column with each electrode, and their activity was recorded using LFPs. The extracellularly recorded action potentials were lost after several days/weeks of recording due to cell expiration, inflammation, reactive gliosis and scarring. Since LFPs tended to be more stable, at least in our case, they were considered for studying the decoding performance of long-term brain implants.

For our LFP-based decoder, we used chronic recording made over several months in two monkeys. During a recording session, the monkey was shown one of two oriented gratings, one of which led to a juice reward and the other not. We focused on the LFP signal features, in the time interval starting from the onset of the stimulus and ending well before the time instant when the reward could be given (conditional training). We analyzed how these signal features depended on the presentation of a reward or a no-reward stimulus (and the background), and how they varied as a function of training time. One such signal feature was phase synchrony and we showed that it is significantly higher for the rewarded stimulus than for the unrewarded one [7]. After training the monkey, this difference is based on the reward, not the stimulus. The dependency of the synchrony on the electrode distance suggests the possible presence of propagating waves from one end of the array to the other. We then developed a new strategy for the single trial decoding of our invasive recordings [8]. It is based on three types of features applied to LFPs: wavelet-based features of LFPs from individual electrodes, and two types of spatial features: phase synchrony between electrode pairs and wave propagation [9]. These features led to an improved separability of the brain responses, to the two types of stimuli, during conditional training [10]. We were able to show that the two types of spatial features mentioned yielded a significant improvement in the single-trial decoding accuracy and, therefore, could be candidate LFP features also for other decoding applications. From a neuroscience perspective, the excellent single-trial performance we obtained in predicting the presence of a rewarded versus an unrewarded stimulus from LFP recordings shows that stimulus-reward association has a profound effect in area V4.

We have also performed the spike-based decoding of the orientation of a visual grating displayed to a monkey, irrespective of the grating’s spatial frequency [11]. Since the cell responses are less strong in visual areas than in motor areas, the recordings contained spikes with smaller amplitudes compared to the noise level. By consequence, rather than LFPs, attempting to extract and assign spikes to individual neurons, feature selection was applied to extract the required information. Two types of feature selection procedures were compared, filter and wrapper. The wrapper is combined with a linear discriminant analysis (LDA) classifier, and the filter is followed by a radial-basis kernel support vector machine (SVM) classifier. In addition, since we have a multi-class classification problem, different methods for combining pairwise classifiers were compared. We showed that an LDA together with a wrapper was shown to yield the best decoding performance [11].

3. NONINVASIVE BCI RESEARCH

As for noninvasive BCI work, our aim is to develop a wireless, portable BCI system operating in real-time. We use the wireless EEG device developed by imec\(^1\), built around their ultra-low power 8-channel EEG amplifier chip [12]. It consists of two parts: an amplifier coupled with a wireless transmitter and a USB stick receiver. Each EEG channel is sampled at 12 bits per sample and 1000 Hz. We use an EEG-cap with large filling holes and sockets for active Ag/AgCl electrodes (ActiCap, Brain Products). The recordings are made with eight electrodes. The reference and ground electrodes are placed on the left and right mastoids respectively. A notch filter is always applied to remove the 50 Hz power-line interference.

Our BCI system is implemented in MATLAB as a client-server application and can run either in parallel MATLAB mode (as two labs) or on two MATLAB sessions started as separate applications. The server part is responsible for the EEG data acquisition, processing and classification. The client part is responsible for the BCI logic, the user interface and the rendering. The client-server communication is implemented using sockets and, due to a minimal data transfer rate (during recording only commands are sent from the server to the client), it can work over a regular network, allowing one to run the BCI application on two different computers. For an accurate (in terms of timing) visualization of the stimuli, we have used the Psychotoolbox\(^2\).

3.1 BCI based on P300

The first type of noninvasive BCI we concentrated on is based on the detection of the P300 component of Event-Related Potential (ERP: a stereotyped electrophysiological response to an internal or external stimulus [13]). This brain potential is elicited in the context of an oddball paradigm: when a subject perceives two types of events, one of which occurs only rarely, the rare event will elicit in the EEG an ERP with an enhanced positive-going component at a latency of about 300 ms (the P300) primary above parietal cortex.

A spelling system based on the detection of this ERP was introduced in 1988 by Farwell and Donchin [14], and it is one of the most studied BCIs. The work presented here deals with a similar application. The P300 Speller allows subjects to spell words by focusing on the desired characters shown in a matrix while the rows and columns of the matrix are consecutively and randomly intensified. The intensification of a row or column containing the target symbol will elicit a P300 potential and, by detecting this ERP component, the BCI is able to identify the target row and column and, thus, retrieves the symbol the subject has in mind.

Ideally, one sequence of intensifications of each row and column would suffice to identify the target symbol. Unfortunately, the low signal-to-noise ratio (SNR) of the ERP makes P300 almost undetectable in a single trial. It is, therefore, common practice to repeat several times the sequence of intensifications, in order to average the EEG responses and increase the SNR. Depending on the number of repetitions, this approach can lead to a dramatic increase of the time needed to communicate a symbol. It is thus important to work on robust and efficient feature extraction/selection and classification techniques to reduce required number of repetitions.

\(^1\)http://www.imec.be\n
\(^2\)http://psychotoolbox.org
Our P300 Speller research was mostly concentrated on the development and selection of powerful machine learning algorithms to arrive at a robust, yet fast P300 detection. We developed a feature extraction approach [15] based on an algorithm proposed in [16] and a feature selection approach [17] based on the group methods of data handling (GMDH) [18]. Our on-line comparison results on healthy subjects suggested that, in general, feature extraction performs better for EEG signal amplitudes as features [19]. To boost the classification rate, we also considered time-frequency features and common spatial pattern features as steps prior to classification [20], and we tried to optimize the time- and frequency range of the P300 BCI speller [21]. Results obtained with healthy subjects suggested that not only the P300 has its impact to the classification performance, but also the early ERP components, starting from about 200 ms (the necessity of taking into account these early components, in more occipital sites, was also confirmed in a patient study in [23]). As to the frequency range, the interval [0.1;10] Hz should be kept for further classification. Different linear [22] together with nonlinear classifiers [23] were also compared during an on-line testing of the P300 speller on (partially) disabled subjects. We found that the Bayesian linear discriminant analysis (BLDA) [24] together with the nonlinear SVM (nSVM) yielded the better classification accuracy. When taking into account the geometry of the spelling matrix, we showed that the BLDA classifier, for the case of the rows, and the BLDA together with nSVM, for the case of the columns yield, in general, the smallest standard deviations between the actually typed and the intended characters (while mistyping mostly occurs either in the same row or column), suggesting that those classifiers lead to a more concentrated distribution of mistakes [23]. We also observed that the vertical standard deviation of mistakes is in general smaller than the horizontal one, particularly for the most accurate classifiers, especially when more than 5-6 intensification sequences were taken for averaging. Our results with (partially) disabled subjects suggested that motor aphasia deteriorates the applicability of P300 leading to what one might call a BCI illiteracy [23].

### 3.2 BCI based on SSVEP

The steady-state visual evoked potential (SSVEP) relies on the psychophysiological properties of EEG brain responses recorded from the occipital area during the periodic presentation of identical visual stimuli (flickering stimuli). When the periodic presentation is at a sufficiently high rate (> 6 Hz), the individual transient visual responses overlap and become a steady state signal: the signal resonates at the stimulus rate and its multipliers [13]. This means that, when the subject is looking at stimuli flickering at frequency $f_i$, we can detect $f_1$, $2f_i$, $3f_i$, ... in the Fourier transform of the EEG signal recorded from the occipital pole. Since the amplitude of a typical EEG signal decreases as $1/f$ in the spectral domain, the higher harmonics become less prominent. Furthermore, the fundamental harmonic $f_i$ is embedded into other on-going brain activity and (recording) noise. Thus, when considering a small recording interval it is quite likely to detect an (irrelevant) increase in the amplitude at frequency $f_i$. Finally, when building a BCI based on SSVEP, not one but several frequencies need to be simultaneously displayed, one for each command the subject wants to issue. Therefore, the detection problem becomes more complex since now one of several possible flickering frequencies $f_i$ need to be detected from the EEG recordings.

For our SSVEP BCI system, we used signals from above occipital cortex. The recorded EEG signals underwent spatial filtering according to the minimum noise energy combination [25] and classified with the use of a (weighted) sum of SNRs in signals after filtering [26]. As a primary application, a spelling device was developed, where the subject could type symbols of his/her choice by directing the gaze onto character (or group of characters) flickering at a particular frequency [27]. A comparison of the synchronous mode (for a fixed duration of stimulation and a sequential signal processing and decoding) and the asynchronous mode of operation (were all system components work in parallel: signal recording, processing and decoding are performed during the stimulation phase) revealed the superior performance of the latter in terms of the information transfer rate (ITR) (number of bits conveyed per minute). As a second application, “The Maze” game was developed, in which case the subject needs to navigate an avatar in a maze-like environment [28]. This game is targeted for a broad audience, to demonstrate what BCI can do. It was tested by 53 persons during the i-brain & senses event. In order to make the BCI game more attractive, we did not use our conventional EEG system, but utilized a commercial EEG gaming headset EPOC from Emotiv. This enabled us to reduce the EEG preparation time, since the EPOC does not require conductive gel. However, as we needed to access brain regions other than the ones the EPOC was designed for, we had to reverse the EPOC (rotate 180° axially). Feedback from the players was collected anonymously. The results revealed that 66% of the players reported a good control over the game, while only 3.8% reported no control at all. This shows the potential applicability of EEG-based BCI in gaming, since also 85% reported that they had fun with it (and no one reported to dislike it). The ability to control the game was strongly correlated (0.7031) with the ability to concentrate during the game, indicating the necessity to be involved and to keep a proper level of attention going to flickering stimuli. As a third application, we combined object detection with BCI with the aim to develop a new way of subject-environment interaction [29]. This application enables the subject to select an object from a restricted set of them with the use of the SSVEP paradigm; the objects were captured with a camera and presented on the screen. Such a system could be used in hospitals to enhance the interaction of bedridden, temporarily immobilized patients with objects present in their rooms.

Apart from frequency domain, we also developed time domain based decoders [30], since they have been conjectured to yield a better performance for the SSVEP paradigm [31]. Rather than using FFT and frequency domain features, time domain SSVEP classifier utilize a stimuli-locked averaging of the EEG response (with lengths equal to integer multiples of the periods of the stimulation frequencies used), and further classify the time-amplitude feature scores by means of linear classifiers [30]. This method enables one to obtain more precise SSVEP estimates when using computer based displays. (Note that such displays do not have a refresh rate that is exactly equal to, e.g., 60 Hz, but rather a slightly lower one, with the refresh rate usually ranging from 60 to 120 Hz.)

---

1. Ghent, Belgium, 18–19 March, 2011
2. [http://www.emotiv.com](http://www.emotiv.com)
smaller one.) Indeed, it allows for a more accurate detection using shorter recording intervals: frequencies that can not be distinguished with FFT, because of the limited spectral resolution of a short recording interval, can now become distinguishable.

While the previously discussed SSVEP BCI systems achieve a reasonable ITR, since they rely on a computer screen (monitor) for the visual stimulation, they have some limitations: the stimulation frequencies are related to the refresh rate of the computer screen [32] (they could not be higher than half the screen’s refresh rate), and restricted to specific (subject-dependent) frequency bands, in order to obtain a good SSVEP response, the harmonics of some stimulation frequencies could interfere with one another (and their harmonics), leading to a deterioration of the decoding performance [32], and so on. These restrictions limit the number of targets that can be displayed simultaneously in a SSVEP-based BCI. In order to have more targets, the phase of the SSVEP could be used as an alternative [33]: even a single frequency could be used but with different phase lags, for encoding different targets. Thus, by extracting phase information from the Fourier transform of the EEG signal, and by comparing it to the phase of some reference signal(s), one can detect the target the subject is looking at. We have developed classifiers for decoding phase information in SSVEP using a multilayer feedforward neural network based on multivalued logic [34] and using a fuzzy logic controller [35]. Both of these classifiers incorporate circularity in the input (phase shift) and output (classes), and are able to deal with multichannel data. When comparing the results, the proposed decoding methods perform better for a larger amount of target classes and for a sufficient length of the decoding window.

3.3 ErrPs in the Context of BCI

Like any communication system, a BCI is prone to mistakes. One could attempt to incorporate the correction of those mistakes into the BCI system. This can be done, for example, in the case of the speller, by including the backspace character into the spelling matrix. But a more elegant way to cope with mistakes is to detect the so-called Error-related Potentials (ErrPs). They were suggested to be generated in the anterior cingulate cortex with a spatial distribution over the fronto-central regions of the scalp, and related to the subject’s perception of an error [36]. Thus, whenever the system detects an ErrP in the EEG recordings, evoked in response to the subject realizing an error has occurred, the system is in a position to correct the error (using some error correction algorithm), at least in principle.

We have investigated the ErrP when recording online with the P300 BCI speller. We considered 6 healthy subjects [37] and found that, on average, the difference between ErrPs and potentials occurring when correctly typing characters has a statistically significant negative peak at 320 ms followed by a positive one at 450 ms. This was most prominently the case for electrodes Fz and FCz. This suggests that these time instances could be important for ErrP detection.

In order to evaluate the integration of ErrPs into a P300 BCI spelling system, we performed a study with several longer recording sessions so as to collect a sufficient amount of ErrPs. (Note that, in a properly working BCI systems, errors should occur only infrequently.) We evaluated the single-trial decoding of the ErrP in the terms of the ErrP detection accuracy, as well as the sensitivity (i.e., the proportion of ErrP segments that are correctly identified) and the specificity (the proportion of non-ErrP segments that are correctly identified) [38]. The latter measure is particularly important for a BCI system, since we have to minimize the occurrence of accidentally generated mistakes as a result of the automatic “correction” of faithfully typed symbols, while also keeping the correct detection of mistyped symbols high (but having in mind that their correction potentially can also be done with the use of the back space symbol). We have found that the linear SVM has very high values for the specificity and relatively lower values for the sensitivity, while the LDA shows a more balanced performance for the specificity and the sensitivity. But for the global accuracy, the linear SVM seems to outperform the LDA.

To show the potential of an ErrP-based correction strategy for a P300 Speller, we evaluated the efficiency of correcting the mistyped symbols by replacing it with the second most probable symbols (in terms of the classifier ranking) [38]. We found that, for all tested subjects, the SVM outperforms the LDA. But only for 2 out of our 3 subjects the SVM-based ErrP detection is beneficial (the resulting typing accuracy was higher than without ErrP correction), after using around 100 and 250 typed symbols for training.

4. ACKNOWLEDGMENTS

NVM is supported by the research grant GOA 10/019, NC is supported by IST-2007-217077, AC and AR are supported by IWT doctoral grants, MvV is supported by IUAP P6/29, MMVH is supported by PFV/10/008, CREA/07/027, IST-2007-217077, G.0588.09, IUAP P6/29, GOA 10/019.

5. REFERENCES
