Emotion Recognition from Brain Signals Using Hybrid Adaptive Filtering and Higher Order Crossings Analysis

Panagiotis C. Petrantonakis, *Student Member*, *IEEE*, and
Leontios J. Hadjileontiadis, *Member*, *IEEE*

Abstract—This paper aims at providing a new feature extraction method for a user-independent emotion recognition system, namely, HAF-HOC, from electroencephalograms (EEGs). A novel filtering procedure, namely, Hybrid Adaptive Filtering (HAF), for an efficient extraction of the emotion-related EEG-characteristics was developed by applying Genetic Algorithms to the Empirical Mode Decomposition-based representation of EEG signals. In addition, Higher Order Crossings (HOCs) analysis was employed for feature extraction realization from the HAF-filtered signals. The introduced HAF-HOC scheme incorporated four different classification methods to accomplish a robust emotion recognition performance. Through a series of facial-expression image projection, as a Mirror Neuron System-based emotion elicitation process, EEG data related to six basic emotions (*happiness*, *surprise*, *anger*, *fear*, *disgust*, and *sadness*) have been acquired from 16 healthy subjects using three EEG channels. Experimental results from the application of the HAF-HOC to the collected EEG data and comparison with previous approaches have shown that the HAF-HOC scheme clearly surpasses the latter in the field of emotion recognition from brain signals for the discrimination of up to six distinct emotions, providing higher classification rates up to 85.17 percent. The promising performance of the HAF-HOC surfaces the value of EEG signals within the endeavor of realizing more pragmatic, affective human-machine interfaces.

Index Terms—EEG, emotion recognition, EMD, genetic algorithms, higher order crossings analysis, hybrid adaptive filtering, mirror neuron system.

1 Introduction

Emotion is one of the fundamental factors that influence our intelligence and our daily enterprises that demand communication skills and constructive social interaction. As a result, it dramatically affects many aspects of everyday life. Considering the proliferation of machines in our commonness, it is essential that they can interact with people the same way people interact with each other if that Human-Machine Interaction (HMI) takes into account the human affective states. It is evidenced [1] that Human-Human Interaction (HHI) has no significant differences from an HMI regarding the social and communication perspectives of such an association. In particular, a variety of experiments have been conducted where, in an HHI, one of the humans is replaced by a computer and the results suggest that the fundamental HHI communication rules still hold. In order to accomplish such a pragmatic HMI, it becomes crucial to imbue machines with the ability to detect, recognize, and respond to human emotional states and reactions, that is, to outfit machines with Emotional Intelligence (EI) [2]. The latter is a target of Affective Computing (AC) [3], which is the branch of Artificial Intelligence (AI) that deals with the design of systems and devices that can detect, recognize, and process human emotion with the ultimate goal of implementing the aforementioned more reliable HMI.

Emotion Recognition (ER) is the first and one of the most important issues AC brings forward and plays a dominant role in the effort to incorporate computers, and generally machines, with the ability to interact with humans by expressing cues that postulate and demonstrate EI-related attitude. Successful ER will enable machines to recognize the affective state of the user and collect emotional data for processing in order to proceed toward the terminus of emotion-based HMI, the emotional-like response. Toward effective ER, a large variety of methods and devices have been implemented, mostly concerning ER from face [4], [5], [6], speech [7], [8], and signals from the autonomous nervous system (ANS), i.e., heart rate and galvanic skin response (GSR) [9], [10], [11].

A relatively new field in the ER area is the Electro-encephalogram (EEG)-based ER (EEG-ER), which overcomes some of the fundamental reliability issues that arise with ER from face, voice, or ANS-related signals. For instance, a facial expression recognition approach would be useless for people with the inability to express emotions via face, even if they really feel them, such as patients within the autism spectrum [12], or for situations of human social masking; for example, when smiling though feeling angry. Moreover, voice and ANS signals are vulnerable to "noise"

• *The authors are with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, University Campus, Building D, 6th floor, office 3, Thessaloniki GR-54124, Greece. E-mail: {ppetrant, leontios}@auth.gr.*

Manuscript received 8 Mar. 2010; revised 27 May 2010; accepted 26 July 2010; published online 17 Aug. 2010.
Recommended for acceptance by R. Calvo.
For information on obtaining reprints of this article, please send e-mail to: taffc@computer.org, and reference IEEECS Log Number TAFFC-2010-03-0017.
Digital Object Identifier no. 10.1109/T-AFFC.2010.7.

related to activity that does not derive from emotional experience, i.e., GSR signals are highly influenced by inspiration which may be caused from physical and not emotional activity. On the other hand, signals from the Central Nervous System (CNS), such as EEG, Magnetoencephalogram (MEG), Positron Emission Tomography (PET), or functional Magnetic Resonance Imaging (fMRI), are not influenced by the aforementioned factors as they capture the expression of emotional experience from its origin. Toward such a more reliable ER procedure, EEG appears to be the less intrusive and the one with the best time resolution than the other three (MEG, PET, and fMRI). Motivated by the latter, a number of EEG-ER research efforts have been proposed in the literature (see the preceding Section 2.3 for a thorough description).

In this work, we introduce a user-independent EEG-ER system for the classification of six typical emotions, i.e., *happiness*, *surprise*, *anger*, *fear*, *disgust*, and *sadness*. The contribution of the paper is in:

1. developing the novel Hybrid Adaptive Filtering (HAF) for efficiently isolating the emotion-related EEG-characteristics, by applying Genetic Algorithms [13] to the representation of the EEG signal on the Empirical Mode Decomposition [14] domain;
2. further extending the effectiveness of Higher Order Crossings (HOCs)-based feature vector, initially presented in [15], by structuring it via the oscillatory pattern of the HAF-outputted EEG signal, instead of the EEG signal itself (as used in [15]);
3. examining the potential of the intrinsic modes of the EEG signal to discriminate between different emotions; and
4. analyzing the classification power of individual EEG channels or combinations of them using the proposed feature extraction method.

The classification efficiency of the proposed approach, namely, HAF-HOC, is outlined from the experimental results by the application of the HAF-HOC to EEG-data sets acquired from 16 healthy volunteers under emotion evocation with visual inputs (pictures with emotion-related facial expressions). This setup was motivated by the Mirror Neuron System (MNS) concept [16], which relates emotion induction with the process of imitation. Compared to previous approaches, the HAF-HOC reveals its superiority in the EEG-ER field for the discrimination of up to six distinct emotions.

The rest of this paper is structured as follows: Section 2 provides an extended report for the related literature as far as EEG-based differentiation of emotions, MNS, and EEG-ER are concerned. The experiment protocols used for emotion elicitation and the data set construction issues are described in Section 3. Section 4 presents the proposed method regarding the HAF algorithm and the feature vector extraction technique based on HOC analysis, alongside some background issues concerning the theoretical aspects used for the proposed approach and the four classifiers used for the classification realization. Section 5 deals with the implementation issues of the proposed approach. Results from the application of the HAF-HOC method to the experimental data set are elaborately presented in Section 6.

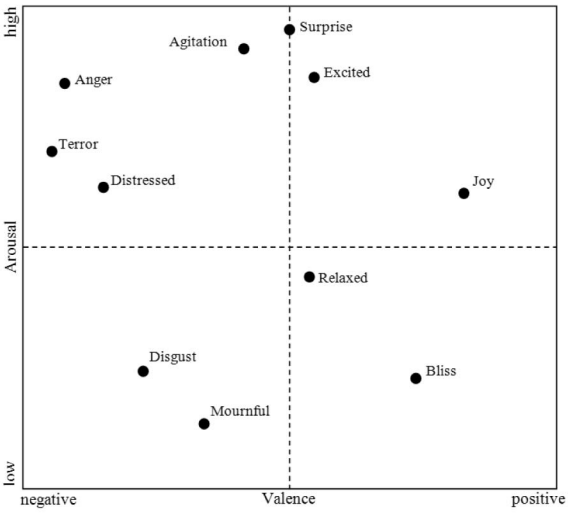

Fig. 1. 2D emotion model by valence and arousal.

Section 7 provides a discussion on issues referring to EEG-ER from HAF-HOC. Finally, Section 8 concludes the paper.

2 Related Research

2.1 Differentiation of Emotions from Brain Activity

The degree to which emotions are differentiated by unique patterns of physiological signals used to be or even still consists of a debate as old as the study of emotion itself. More particularly, there are two major positions to the subject. The first, supported by the cognitive scientists, based on Cannon's theory [17], claims that different emotions are associated with the same physiological patterns and it is not possible to differentiated between them. On the other hand, an alternative position that has become the most prominent is based on the theoretical writings of Darwin [18] and James [19] and supports the fact that different emotions are accompanied by discrete patterns of physiological activity. Furthermore, Ekman [20] has also reported that these patterns refer both to the ANS (e.g., heart rate, GSR, etc.) and the CNS (e.g., brain signals) and can differentiate the six primary emotions of *happiness*, *sadness*, *anger*, *fear*, *disgust*, and *surprise*. Specifically, Ekman et al. [21] have evidenced the existence of unique patterns of ANS activity that differentiates among the negative emotions of fear, anger, disgust, and sadness. Besides the research toward the ability of ANS activity to differentiate emotions, there has also been a great volume of research conducted for the same ability of CNS activity. In particular, relative psychophysiology literature has revealed the most prominent expression of emotion in brain signals, i.e., the asymmetry between the left and right brain hemispheres. Davidson et al. [22] developed a model that related this asymmetric behavior with emotions, with the latter analyzed in two main dimensions, i.e., *arousal* and *valence*. Valence stands for one's judgment about a situation as positive or negative and arousal spans from calmness to excitement, expressing the degree of one's excitation (see Fig. 1). According to that model, emotions are: 1) organized around approach-withdrawal tendencies and 2) differentially lateralized in the frontal region of the brain.

The left frontal area is involved in the experience of positive emotions, such as joy or happiness (the experience of positive affect facilitates and maintains approach behaviors), whereas the right frontal region is involved in the experience of negative emotions, such as fear or disgust (the experience of negative affect facilitates and maintains withdrawal behaviors). Furthermore, Davidson et al. [23] tried to differentiate the emotions of happiness and disgust with EEG signals captured from the left and right frontal, central, anterior temporal, and parietal regions (F3, F4, C3, C4, T3, T4, P3, P4 positions according to the 10-20 system [24]). The results revealed a more right-sided activation, as far as the power of *alpha* (8-12 Hz) band of the EEG signal is concerned, for the disgust condition for both the frontal and anterior temporal regions. Thus, the results enhanced the applicability of the aforementioned model and confirmed the evidenced extensive anatomical reciprocity of both regions with limbic circuits that have been directly implicated in the control of emotion [25]. Later, Davidson [26] examined the asymmetry concept in regard to the prefrontal cortex (PFC) based on data from the neuroscience literature on the heterogeneity of different sectors of the PFC. He envisaged the PFC role in affective expression in the brain and considered the use of more EEG signal bands other than *alpha*, i.e., *theta* (4-7 Hz), *beta* (13-30 Hz), and *gamma* (31-100 Hz). More recently, a variety of psychophysiological studies of emotion [27] have confirmed and adopted the model proposed by Davidson, expanding it by arguing that frontal EEG asymmetry may serve as both a moderator and a mediator of emotion and motivation-related constructs [28]. There, activity which is characterized in terms of decreased power in the *alpha* band was found to be associated with emotional states. Finally, Aftanas et al. [29] used affective pictures and Event-Related Desynchronization/Synchronization (ERD/ERS) analysis to study cortical activation during emotion processing. ERD refers to the desynchronization of a brain rhythm during a stimulus, which leads to a power decrease, whereas ERS refers to the synchronization of the rhythm and power increase [30]. In accordance with the asymmetry literature, they reported relatively greater right hemisphere ERS for negative emotional states and greater left hemisphere ERS for positive ones. All of the aforementioned studies reveal the potential of EEG signals as a means for developing an ER system based on the asymmetry concept expressed by frequency bands' power and ERD/ERS phenomenon.

2.2 Mirror Neuron System (MNS) and Affective Picture Processing

Neurophysiologic experiments have demonstrated that when individuals observe an action done by another individual, they seem to have the same or akin brain activity as if they did the corresponding action themselves. This activity is ascribed to the MNS [16], which is also connected with the ability of imitation that, among others, relates to recognition of emotions by others' facial expressions and gestures [31]. In particular, Lee et al. [32] employed facial electromyography (EMG) and fMRI to investigate the so-called Emotion Expression Interference (EEI), i.e., the contagion of emotion by observation of emotional facial expressions by others while people were observing clips with emotional facial expressions. They found a profound linkage of EEI with the neural substrates of the MNS and assumed that EEI may reflect the engagement of the MNS wherein the expression of discordant facial emotion overrides a prepotent automatic mimicry or mirror neuron response. In a previous work of imitating emotional facial expressions, Lee et al. [33] had again highlighted the involvement of both the motoric and emotional brain centers. Moreover, Carr et al. [34] used fMRI while subjects were either imitating or simply observing emotional facial expressions. According to their results, imitation and simple observation of emotions activated a largely similar network of brain areas, related to the MNS, and they concluded that people feel by a mechanism of action representation that allows empathy and modulates their emotional content. Finally, another, more specific study by Wicker et al. [35] investigated the neural mechanism that underlies the mechanism for feeling the emotion of disgust. An fMRI-based study was performed where participants inhaled odorants producing a strong feeling of disgust. Afterward, the same participants observed video clips showing the emotional facial expression of disgust. It was found that observing the faces and feeling disgust actually activated the same brain areas. They directly associated those findings with the MNS and concluded that this phenomenon, i.e., the activation of the neural representation of an emotion by simply observing that emotion, provides a unifying mechanism for understanding the behaviors of others. All of the aforementioned studies highlight the linkage between the brain activity and the observation of emotional facial expressions and establish the use of images that picturize emotional facial expression for the development of a novel emotion elicitation technique.

2.3 Approaches to EEG-Based Emotion Recognition

In the line of EEG-ER, considerably little research has been conducted compared to emotion recognition from face, voice, and ANS-based signals. In particular, the majority of the related studies are user-dependent, confined to feature extraction methods relevant with alpha and beta bands and their characteristics (e.g., power and peaks), and induce emotions with pictures, audios, or clips that are assumed to elicit certain affective states within the 2D valence/arousal space. More specifically, Choppin [36] analyzed EEG signals and used neural networks to classify them in six emotions based on emotional valence and arousal with a 64 percent success rate. Takahashi [37], by using statistical feature vectors that Picard et al. [9] used for emotion recognition from physiological signals, conducted a user-independent emotion recognition study from either physiological (pulse and GSR) and EEG (Fp1 and Fp2 channels according to 10-20 system) signals, their fusion included. With data from 12 subjects, a 41.68 percent success rate accomplished only from the EEG signals and 41.72 percent with the fusion of the physiological and EEG signals for the discrimination of five distinct emotions, i.e., joy, anger, sadness, fear, and relax. Other studies conducted experiments for emotion elicitation with the participation of four [38], five [39], [40], and 10 [41] subjects and tried to classify EEG and/or physiological

signals in emotion or affective states (classes within a specific range in the arousal and/or valence axis) with moderate results, i.e., success classification rates less than 60 percent for discrimination between three and five classes. Another more prominent feature extraction method for EEG-ER was conducted by Murugappan et al. [42], which resulted in a satisfactory clustering within the emotions of disgust, happiness, and fear with fuzzy c-means [43] and fuzzy k-means [44] as the classification methods. The feature vector extraction method was based on Daubechies [45] order 4 wavelet transform and referred to the energy and entropy characteristics of the signal. For the emotion elicitation experiment, clips with specified emotional content were used to induct emotions from six subjects whose EEG activity was recorded from 64 channels. Finally, more recent works in the field [46], [47] have tried to implement EEG-ER applications in the field of learning and robotics, respectively, in an effort to construct a brain wave sensing intelligent tutoring system [46] or an emotional humanoid robot [47].

Despite the classification potential of EEG for such kinds of recognition reflected in the works reported in the literature, further research is needed in order to improve recognition rates and develop more effective feature extraction methods. An initial effort toward a more robust EEG-ER approach has been proposed by the authors in a recent work [15] as a first attempt to incorporate HOC analysis [48] in the feature extraction process for an effective EEG-ER system, namely, HOC-Emotion Classifier (HOC-EC). The classification efficiency achieved by the HOC-EC was substantially improved (62.3 percent using only one EEG channel and 83.33 percent for fusion of three distinct channels) compared to the previous EEG-ER approaches, and paved the way for further consideration of the HOC potential in the field of EEG-ER. This was the motivation for deepening further and proposing in this paper an extended feature extraction framework combining novel HAF with HOC, resulting in the HAF-HOC scheme (see Section 4).

3 Experimental Protocol and Data Set Construction

3.1 Emotion Elicitation

Until now, the most prominent emotion induction technique for emotion recognition (either for physiological or EEG signals) was predominantly based on pictures, sounds, or clips that were assumed to elicit certain emotional states, basically in the valence/arousal space. Two of the databases most frequently used for that purposes are the International Affective Picture System (IAPS) database [49] and the International Affective Digitized Sounds (IADSs) [50] database that include pictures and sounds labeled with valence and arousal values by subjects during large-scale experiments. More specifically, the images included in the IAPS database mostly picturize situations that evoke negative or positive feelings with low or high intensity.

How effectively the desired emotion is inducted is a significant query that is relatively difficult to answer. Picard et al. [9] stressed five important factors that should be taken under consideration during an emotion elicitation experiment:

1. *Subject elicited versus event elicited*: Does subject purposefully elicit emotion or is it elicited by a stimulus or situation outside the subject's efforts?
2. *Lab setting versus real world:* Is subject in a lab or in a special room that is not their usual environment?
3. *Expression versus feeling*: Is the emphasis on external expression or on internal feeling?
4. *Open recordings versus hidden recordings*: Does subject know that anything is being recorded?
5. *Emotion-purpose versus other-purpose*: Does subject know that the experiment is about emotion?

The last four factors above are quite controllable during an experimental phase; the first one, however, raises issues concerning the emotional impact to the subjects by the situations included in IAPS database. Aspects like personality, personal experiences, and a particular subject's mood at the time that the experiment is conducted dramatically influence the way someone reacts emotionally in view of that kind of image. As a result, as soon as the emotion is induced by a stimulus outside the efforts of the subject, how effectively the emotion is actually evoked is questioned, and consequently, how representative the EEG activity is with regard to a particular emotion. This is addressed by replacing those pictures with others, showing humans expressing the basic emotions. In this paper, Ekman's six universal emotions [51] were adopted, i.e., *happiness*, *surprise*, *anger*, *fear*, *disgust*, and *sadness*, with the latter consisting of the class set of the discrimination analysis. The adoption of images with humans expressing emotions as emotion elicitation stimulus was supported by the MNS theory and the related literature (see Section 2.2); in fact, it is not expected that the subject will actually feel the emotion but will express the same or akin EEG activity as if she/he had really experienced it. In this way, there was no need to employ any body language and facial-based monitoring procedure or manually annotated affect labels for the subjects' emotion identification by external judges [52].

In the line of this work, a special experiment was designed in order to collect the EEG signals necessary for the HAF-HOC analysis. In this experiment, 16 healthy volunteers participated; all were right-handed subjects (nine males and seven females) in the age group of 19-32 yrs. At first, a suitable interface was implemented for the automated projection of the emotion-related pictures and the self-assessment of the emotion picturized each time. Pictures of Facial Affect (POFA) by Ekman and Friesen [53] showing people expressing the six basic emotions (see Fig. 4) were subsequently previewed, separated by black and counting down frames to accomplish a relaxation phase before the projection of the new picture. More specifically, 60 pictures (10 pictures per emotion) were randomly projected for 5 seconds after a 5-second black screen period, a 5-second period in which countdown frames were demonstrated, and a 1-second projection of a cross shape in the middle of the screen to attract the sight of the subject. The same 16-second procedure was repeated for every 1 of the 60 pictures. An illustrated version of the experiment protocol is shown in Fig. 2. The subjects were advised to stay as still as possible (to comprehend or simply observe the emotion rather than to mimic the facial expression) and not to blink or move, as much as possible, during the image projection phase to

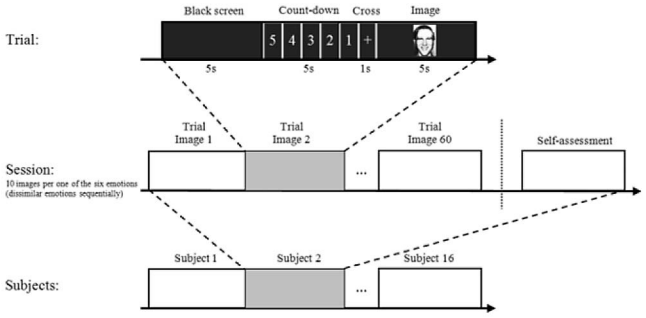

Fig. 2. Illustration of the experimental protocol for emotion elicitation.

reduce the appearance of relevant artifacts in the EEG recordings (such as from facial muscles). The black screen phase was employed to offer time for the subjects to relax, whereas the countdown phase was used to serve as an emotional-reset tool due to its lack of emotional content. After the completion of the projection phase, the subject entered into a self-assessment phase where she/he could preview the pictures previously projected and categorize them according to the emotion each picture present according to her/his opinion. Apart from the six categories of emotions, she/he could choose another option, namely, "other," if none of the previous categories were suitable for her/him. The results of the self-assessment phase are tabulated in Table 1. It should be noted that the final categorization of the signals was conducted according to the norms provided with the POFA database. This was adopted for three major reasons. First, POFA is extensively studied for its reliability to effectively picturize the aforementioned six basic emotions. Second, Table 1 shows an actually confined misclassification, proportional with those obtained by the reliability studies previously mentioned. Finally and most importantly, the MNS concept adopted here relies mostly on subconscious brain functioning. As Rizzolati and Craighero coin [16]: "Mirror neurons represent the neural basis of a mechanism that creates a direct link between the sender of a message and its receiver. Due to this mechanism, actions done by other individuals become messages that are understood by an observer *without any cognitive mediation*." Thus, the categorization of the signals with regard to the self-assessment phase, during which the subjects were consciously choosing emotion category for each picture, would circumvent the whole MNS approach as an emotion elicitation technique.

TABLE 1
Self-Assessment Classification Accuracy (in Percent) of the Six Basic Emotions, Confusion Table Including "Other" Emotion

Emotions	*Happiness*	*Surprise*	*Anger*	*Fear*	*Disgust*	*Sadness*	*Other*
Happiness	**99.37**	0	0	0	0	0	0.63
Surprise	1.87	**93.75**	0	1.88	0	0	2.50
Anger	0	0	**95.63**	0	0	0	4.38
Fear	0	20.63	1.25	**75**	1.88	0	1.25
Disgust	0	0	15	0	**76.25**	0	8.75
Sadness	0	2.50	0	1.25	3.75	**78.75**	13.75

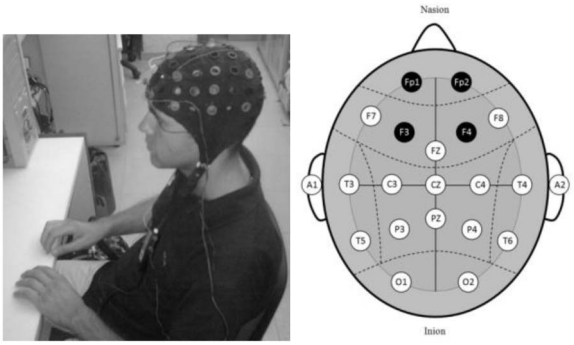

Fig. 3. The Fp1, Fp2, F3, and F4 positions used for the EEG acquisition according to the 10-20 system [24].

Many of the previous studies considered the original emotions and the "null" emotion situation, i.e., neutral emotional state (although neutral and emotional seem to be conflicting terms). In this paper, the neutral emotion as a different emotion class was rejected since the adopted MNS concept upon which the whole emotion elicitation experiment was built includes the essence of mimicry; hence, the induction of an emotional state with "null" emotional meaning, according to the MNS concept, would have no meaningful explanation.

3.2 EEG-Signal Recordings

The EEG signals from each subject were recorded during the projection phase. The EEG signals were acquired from Fp1, Fp2, F3, and F4 positions, according to the 10-20 system. The small number of EEG channels selected was an effort to implement an emotion recognition method which would result in a more user-friendly environment in the future and to acquire signals from the brain areas according to the emotion expression in the brain, based on the asymmetry concept, as described in Section 2.1 [23], [25], [26], [27], [28], [29], [30]. The Fp1 and Fp2 positions were recorded as monopole channels (channels 1 and 2, respectively), whereas the F3 and F4 positions were dipoles (channel 3), resulting in a 3-EEG channel set (see Fig. 3). The ground was placed in the right earlobe.

3.3 Preprocessing and Data Set Construction

After the acquisition part, the EEG signals were subjected to filtering. In particular, a bandpass Butterworth filter was used in order to retain only the frequencies within the *alpha* (8-12 Hz) and *beta* (13-30 Hz) bands in order to exploit the mutual relation regarding the prefrontal cortical activation or inactivation (see Section 2.1). Then, the EEG signals were segmented into 5 sec segments corresponding to the duration of each picture projection and the set of the signals for each emotion, that is, 10 signals per emotion per subject per channel, was averaged to further reveal the ERD/ERS characteristics [54]. Consequently, a new set of 16 EEG signals for each emotion lasting 5 sec each was constructed in order to evaluate the HAF-HOC performance on a user-independent basis. As far as the noise embedded in the recordings is concerned, such as superimposed artifacts from various sources, it has been reported [55], [56], [57] that they are effectively eliminated by appropriate bandpass filtering. More particularly, the

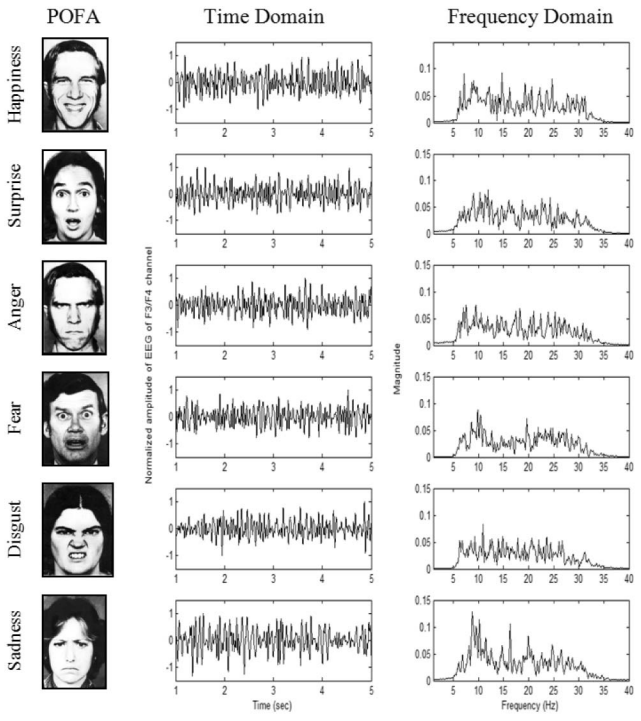

Fig. 4. Averaged EEG signal in the time domain, corresponding to each one of the six basic emotions for channel 3, along with the corresponding frequency content, estimated using 2,048-point FFT.

influence of eye movement/blinking is most dominant below 4 Hz; heart-functioning causes artifacts around 1.2 Hz, whereas muscle artifacts affect the EEG spectrum above 30 Hz. Nonphysiological artifacts caused by power lines are clearly above 30 Hz (50-60 Hz). Consequently, by extracting the *alpha* and *beta* frequency bands only from the acquired EEG recordings, much of the noise influence is circumvented.

Fig. 4 depicts an example of the averaged EEG signal in the time domain, corresponding to each one of the six basic emotions for channel 3, along with the corresponding frequency content, estimated using 2,048-point Fast Fourier transform (FFT).

4 Methodology

The overall structure of the HAF-HOC scheme, along with a working example, is depicted in Fig. 5. After the preprocessing stage for signal segmentation and denoising, the EEG signal is input to the HAF section of the HAF-HOC scheme. The role of HAF is to isolate the emotion-related EEG characteristics, facilitating the task of the feature vector extraction that follows. To achieve this, HAF incorporates GA that acts upon the representation of the EEG signal on the EMD space. In the latter, the EEG signal is decomposed through an iterative sifting process [14] into a series of Intrinsic Mode Functions (IMFs) that correspond to different oscillatory modes of the EEG signal (see the working example). During the sifting procedure, the oscillation mode is qualified as an IMF according to two conditions: 1) For the entire data set, the number of extrema and the number of zero crossings are either equal or differ at most by one, and 2) the mean value of the envelope defined by

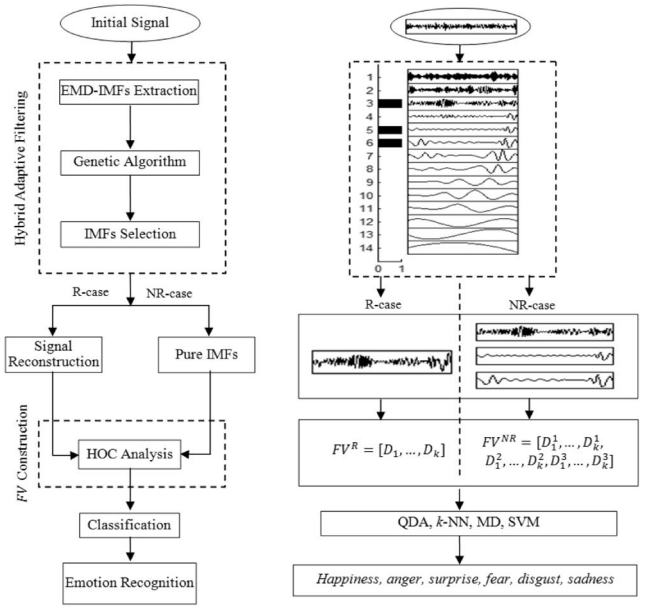

Fig. 5. The proposed HAF-HOC scheme, along with a working example.

the local maxima and the mean value of the envelope defined by the local minima are zero at any point [14]. In the end of the EMD process, the data series can be decomposed into M intrinsic mode functions and a residue. GA then, by using energy or fractal dimension-based fitness function, selects the optimum IMFs that relate the most with the emotion-related EEG characteristics (IMF_3, IMF_5, and IMF_6 in the working example of Fig. 5). The HAF output consists of the selected IMFs that either could be combined through a reconstruction process, i.e., addition across the selected IMFs, to produce a reconstructed EEG signal (R-case) or used directly, without employing any reconstruction process (NR-case) (see the corresponding signals in the working example of Fig. 5). Under both scenarios (i.e., the R and NR-case), the HAF output is then used as input to the HOC section of the HAF-HOC scheme. There, HOC-based analysis is performed, resulting in the efficient extraction of the feature vector FV^{HOC} (see the preceding section for a detailed description), corresponding to the R and NR-case (FV^R and FV^{NR}, respectively, in the working example of Fig. 5). The latter is then forwarded to the next stage of the HAF-HOC scheme, i.e., the classification process, where four classification algorithms are adopted (Quadratic Discriminant Analysis (QDA), k-Nearest Neighbor (k-NN), Mahalanobis Distance (MD), and Support Vector Machines (SVMs)), as there is no single best classification algorithm, that is, no one-size-fits-all. A thorough description of each step of the HAF-HOC scheme follows.

4.1 Hybrid Adaptive Filtering (HAF)

In order to follow the EEG signal characteristics and focus upon the ones that mostly relate to the emotion information, an HAF approach was developed. As *hybrid* declares, HAF combines two processing tools, i.e., EMD and a simple GA optimization concept, in order to construct a filtering process adapted to specific characteristics of the filtered signal. Particularly, Flaudrin et al. [58] qualified EMD as an adaptive filter bank due to its functionality to decompose a

signal in modes and residuals, which are intuitively interpreted as a "spectral" representation. Moreover, he further argues that a consequent result of the above is that selection of modes corresponds to an automatic and adaptive (signal-dependent) time-variant filtering.

In order to exploit the aforementioned capability of the EMD algorithm, a new GA-based approach was developed for the optimized selection of modes that correspond to a specific feature of a signal. The GA used here was based on the realization strategy proposed in [13], including a series of basic steps, i.e., initialization (initial population definition), pair selection, cross-over, mutation, elitist strategy (randomly removing one string from the current population and replacing it with the string that obtained the maximum fitness function value in the previous generation), and termination. As shown in Fig. 5, the initial EEG signal is decomposed into the corresponding IMFs, according to the EMD sifting process [14]. Subsequently, the GA is applied to the extracted IMFs (see the implementation details in Section 5) and those which are selected are used either to reconstruct the new signal (R-case), i.e., the filtered EEG signal, or provide separate signals (NR-case) that represent specific modes of oscillation coexisting in the initial EEG signal.

One of the most important modules of the filtering procedure described above is the fitness function (FF) that is selected for the GA, as this is the major criterion according to which the filtering is conducted. In this work, two FFs were used, one Energy-based (EFF) and one Fractal Dimension-based (FDFF).

4.1.1 Energy-Based Fitness Function (EFF)

A preliminary study of the EFF was attempted in [59], where the potential of the whole HAF concept was evidenced, in a low-scale classification process. The aim of using EFF was to conduct a filtering procedure by selecting the IMFs which embed the majority of the energy of the signal. As described in Section 2.1, emotion differentiation through brain activity is highly related to the power spectrum of specific subbands (e.g., *alpha* band) of the EEG signal. Thus, an IMF selection procedure using an energy-based criterion, directly related to the power spectrum of the EEG signal, would result in a filtered signal with boosted information related to emotion expression in EEG signals. This fact is also justified by the experimental results (see Section 6). The formula used for the EFF is expressed by

$$f(S) = \frac{\sum_{\{S|s_r=1\}} E\{c_r(n)^2\}}{\sum_{i=1}^{M} E\{c_i(n)^2\}}, \quad n = 1, \ldots, N, \tag{1}$$

where S is the string of 1s and 0s, $S_r = 1$ is the set of the elements of S with value 1, and $c(n)$ represents an IMF. According to (1), it is obvious that, during the GA selection phase, the strings constituted of ones (starting from the initial population) that correspond to more energy-loaded IMFs are more likely to give off springs to the next generation. As a result, a bunch of IMFs that are more likely to embrace the majority of the initial signal energy is finally selected; in this way, an energy-based filtering is accomplished.

4.1.2 Fractal Dimension-Based Fitness Function (FDFF)

Apart from the EFF-based FF of the HAF, a second approach was attempted by employing Fractal Dimension (FD) estimation in the FF structure. In general, the FD can be considered as a relative measure of the number of basic building blocks that form a pattern [60]. In that way, the FD could be used as a measure of the signal complexity. Calculating the FD of waveforms has the advantage of fast computation. It consists of estimating the dimension of a time-varying signal (waveform) directly in the time domain, which allows significant savings in program runtime. The aim of the resulting FDFF was to capture the variations in the complexity of the EEG signal. As has been proven in [61], the dynamical complexity of cortical networks, measured by means of the FD, reflects the degree of synergism between neurons and relates to the concepts of ERD/ERS characteristics. On the one hand, synchronization, corresponding to high neural synergism and low complexity, could reflect a resting state of cortical networks. On the other hand, desynchronization, corresponding to low neural synergism and high complexity, could correspond to active information processing in the cortex. To this end, by taking into account the capability of FD to express the complexity of the EEG signal, and consequently, to monitor the ERD/ERS phenomenon, an FD-based filtering procedure would act as a boosting procedure of the information of the initial EEG signal related to the emotion expression in it. To this end, an FDFF was used in order to select the optimum IMFs from an FD-based perspective.

Here, the Petrosian method [62] was used for the estimation of the FD. This method has a preprocessing step where a new, binary, sequence is generated from the initial signal. As described in [62], there are actually five different variations of the Petrosian method, depending on the way the binary sequence is extracted. In this paper, the a approach (as denoted in [62]) was adopted. According to the a-Petrosian method, the binary sequence is generated by assigning one or zero when the waveform value is greater or lower than the mean value of the signal, respectively. The FD is calculated as

$$FD = \frac{log_{10}N}{log_{10}N + log_{10}\left(\frac{N}{N+0.4N_{\Delta}}\right)}, \tag{2}$$

where N is the length of the binary sequence (number of points) and N_{Δ} is the number of dissimilar pairs in the binary sequence generated. The reason that the a-Petrosian algorithm was chosen instead of other methods for FD estimation (e.g., Higuchi [63], Katz [60]) is its relatively low computational burden and its relation to a binary sequence that is the same as the one extracted for the computation of the HOC sequence (see Section 4.2.1).

The formula used for the FDFF is expressed by

$$f(S) = \sum_{\{S|s_r=1\}} FD\{c_r(n)\}, \quad n = 1, \ldots, N. \tag{3}$$

4.2 HOC Analysis

Until now, it has been clear that emotion differentiation from brain signals is highly related to the power spectrum of EEG signals in certain brain locations and specific frequency bands (mostly in *alpha* and *beta* bands). Moreover, this differentiation is expressed by the asymmetry concept as far as ERD/ERS neuronal characteristics are concerned. These facts facilitate the development of a feature vector-extraction technique that considers the spectrum-related attitude of the signals and it is highly

dependent on the dominancy (power) of a certain frequency in a specific subband of the whole frequency spectrum. HOC-based analysis provides such a perspective by rigorously analyzing the signal in the time domain and without employing spectral transforms, as described in the following sections.

4.2.1 HOC-Based Perspective

Consider a finite zero-mean series $\{Z_n\}$, $n = 1, \ldots, N$. In general, its oscillating behavior about the zero level can be expressed by the zero-crossings count. If a filter applies to a time series, it generally changes its oscillation attitude, and consequently, its zero-crossings count. Under this perspective, an iterative procedure regarding the sequential application of a filter to a time series and the counting of the corresponding zero-crossings can be assumed, i.e., repeatedly filter and count (each time the filtering order k increases). The resulting zero-crossings counts are referred to as HOC [48]. When a specific sequence of filters is applied to a time series, the corresponding sequence of zero-crossing counts is obtained, resulting in the so-called HOC sequence. Many different types of HOC sequences can be constructed by appropriate filter design, according to the desired spectral and discrimination analysis. For extended coverage of the HOC sequence extraction, the reader is encouraged to consult the previous work of the authors [15].

The above described HOC, computed in a finitely long real-valued time series, can be viewed as a measure of its oscillation [48]; the more pronounced the oscillation, the higher the expected number of zero-crossings is and vice versa. This relation in the oscillatory pattern can also be viewed through the correlation concept. The autocorrelation and, in particular, the first order autocorrelation ρ_1, i.e.,

$$\rho_1 = \rho(n_1, n_2) = \frac{R(n_1, n_2)}{\sqrt{R(n_1, n_1)}\sqrt{R(n_2, n_2)}}, \quad n_2 - n_1 = 1, \quad (4)$$

(where $R(.,.)$ is the covariance or the autocovariance function), can also serve as a measure of oscillation [48]. It is expected that graphs of sequences of random variables are rather "smooth" or slowly varying when the correlation between neighboring random variables is high (i.e., close to 1) and display rapid oscillation when the same correlation lies close to -1. Considering that both number of zero-crossings and first order autocorrelation are measures of oscillation, it becomes obvious that there must be a direct relationship between them and, in particular, an inverse relationship, reflecting that slowly varying time series are associated with higher correlations, yet lower zero-crossing rates. Profoundly, the situation is inversed for rapidly oscillating series. Consequently and most importantly, the relationship between the zero-crossing rates and autocorrelation implies a formal connection between the zero-crossings and the spectrum, a fact that is exploited in this paper for the feature extraction case in order to construct a feature vector highly related to the emotion expression in the brain. Specifically, it has been proven [48] that the relationship between first order autocorrelation ρ_1 and the number of zero-crossings D is governed by the *cosine formula*, i.e.,

$$\rho_1 = cos\left(\frac{\pi E[D]}{N-1}\right), \quad (5)$$

where $E[D]$ is the mean value of D and N is the size (samples) of the time series. Obviously, the cosine formula also exhibits the inverse relationship concept previously described. In addition, the cosine formula also elucidates the connection between the expected number of zero-crossings and the spectrum. From the definition of ρ_1 and by the Wiener-Khintchine theorem, Papoulis [64] follows the relationship:

$$cos\left(\frac{\pi E[D]}{N-1}\right) = \frac{\int_{-\pi}^{\pi} \cos(\omega) dF(\omega)}{\int_{-\pi}^{\pi} dF(\omega)}, \quad (6)$$

where F is a uniquely defined bounded monotone non-decreasing real function such that $F(-\pi) = 0$; $F(\omega)$ is the spectral distribution function. When F is continuous in the origin, a symmetry assumption would imply that (6) can be expressed in terms of positive frequencies, i.e.,

$$cos\left(\frac{\pi E[D]}{N-1}\right) = \frac{\int_{0}^{\pi} \cos(\omega) dF(\omega)}{\int_{0}^{\pi} dF(\omega)}. \quad (7)$$

An immediate consequence of (7) is that the normalized expected number of zero-crossings $\bar{\gamma} = \pi E[D]/(N-1)$ is a weighted average of the spectral mass distributed over the interval $[0, \pi]$. This point can be vividly illustrated by two extreme cases, i.e., a flat-spectrum process and a process with a spectrum including a dominant frequency. For the former case, a white noise process is concerned. Then, the spectrum is continuous with a flat density equal to a constant over the interval $[0, \pi]$. In that case, (7) gives

$$cos(\bar{\gamma}) = 0, \quad (8)$$

and because of the monotonicity of cosine function over $[0, \pi]$:

$$\bar{\gamma} = \frac{\pi}{2}. \quad (9)$$

Due to the absence of any dominant region, $\bar{\gamma}$ falls in the middle of $[0, \pi]$. On the other hand, supposing a certain frequency band becomes dominant, that is, carries more power than other frequency bands, $\bar{\gamma}$ will tend to land there. In the extreme case where the process is a pure sinusoid with frequency ω_0, $\bar{\gamma}$ will equal ω_0 [48]. Concluding, when a certain frequency becomes dominant, it attracts the expected normalized number of zero-crossings, and the latter will land at or near that frequency. This tendency of $\bar{\gamma}$ to admit values in a dominant spectral band, when it exists, is referred to as the *dominant frequency principle* [48]. For an extended coverage of the relationship between the zero-crossings and spectrum, the reader is encouraged to consult Kedem's book [48].

4.2.2 HOC-Based Feature Vector

The HOC-based FV extraction method was implemented in two different approaches, following the R-case and NR-case scenarios. Here, we present the general concept of the FV extraction; the implementation of FV for the R and NR-case is thoroughly described in Section 5.3. From a general perspective, the previously described HOC-based analysis is used to construct the feature vector (FV^{HOC}), formed as follows:

$$FV^{HOC} = [D_1, D_2, \ldots, D_L], 1 < L \leq J, \quad (10)$$

where J denotes the maximum order of the estimated HOC and L the HOC order up to which they were used to form

the FV^{HOC} (see subsequent section for the selection of L). Moreover, an important advantage of using HOC is that the number of zero-crossings is not only referred to the initial signal but also to signals that consist the output of high-pass filtering of the initial. As a consequence, the dominant frequency principle is applied to a set of subbands of the initial signal, detecting a respective set of dominant regions/frequencies. This is one of the major advantages of the HOC as a feature vector extraction technique, and in conjunction with its implicit computation, it becomes a robust and efficient feature vector for emotion recognition from EEG signals.

4.3 Classification

For an extensive evaluation of the classification performance of the HAF-HOC scheme, the widely used classifiers, i.e., QDA [65], k-NN [66], MD [67], and SVM [68], were adopted, similarly to [15], epitomized below.

4.3.1 Quadratic Discriminant Analysis (QDA)

QDA is based on the quadratic discriminant function:

$$d_q(FV) = -\frac{1}{2}log|C_q| - \frac{1}{2}(FV - \mu_q)^T C_q^{-1}(FV - \mu_q) + logp_q, \tag{11}$$

where $q = 1, \ldots, Q$ corresponds to the number of classes, C is the covariance matrix for each class, μ is a vector with the mean values of each variable consisting the FV, and p is the prior probability for each class. Assuming training and test sets as proportions of the initial data set, the covariance matrices C_q, $q = 1, \ldots, Q$ and the mean vectors μ_q, $q = 1, \ldots, Q$, are calculated from the training part, whereas each FV from the test set is classified into one of the Q classes according to the following rule:

$$g(FV) = arg\ max_q\ d_q(FV), \tag{12}$$

where $g(FV)$ is the class the FV was assigned.

4.3.2 k-Nearest Neighbor (k-NN)

The k-NN algorithm assumes that all instances (FV_S) correspond to points in the n-dimensional (nD) space. The nearest neighbors of a given FV are defined in terms of the standard euclidean distance. The k-NN algorithm is realized in two steps: 1) Define the training set of FV_S and 2) given a query FV_q to be classified, let $FV_1, \ldots, FV_k$ denote the k FV_S from training sets that are nearest toFV_q; then, return the class that the majority of the k nearest neighbors (FV_S) are from. The selection among different metrics (e.g., euclidean, Manhattan, or city block) is generally dictated by computational concerns [69]. Nevertheless, for the case where there is a great difference in the range of the data along different axes in a multidimensional data space, a transformation (scaling) of the latter is necessary; this, however, was not noticed in the examined problem; thus, the euclidean distance was selected as the most simplified metric.

4.3.3 Mahalanobis Distance (MD)

The discriminant analysis used in this approach is based on the MD criterion. MD is the distance between a case (FV) and the centroid for each group in attribute space (nD space). There is an MD for each case and each case is classified as belonging to the group for which the MD is minimum. The MD between two FV_S, FV_1 and FV_2, assuming the same distribution and a corresponding covariance matrix C, is defined as

$$d_s(FV_1, FV_2) = \sqrt{(FV_1 - FV_2)^T C^{-1}(FV_1 - FV_2)}. \tag{13}$$

4.3.4 Support Vector Machines (SVMs)

In the SVM classifier, a polynomial function is used as a kernel function that projects the data into high-dimensional feature space, i.e.,

$$K(FV_{sv}, FV_q) = \left(FV_{sv}^T \cdot FV_q + 1\right)^p, \tag{14}$$

where FV_{sv} is the "support vector" and FV_q is the query FV. Among several available approaches to realize a multiclass SVM classification process, the *one-versus-all* method was adopted here (see the subsequent section for more details on implementation issues).

5 Implementation Issues

5.1 EEG Recordings

EEG recordings were conducted using the g.MOBIlab (g.tec medical and electrical engineering, Guger Technologies, www.gtec.at) portable biosignal acquisition system (four EEG bipolar channels, passive electrodes, Filters: 0.5-30 Hz, Sensitivity: 100μV, Data acquisition: A/D converter with 16-bit resolution and sampling frequency of 256 Hz; Data transfer: wireless, Bluetooth "Class I" technology).

5.2 Hybrid Adaptive Filtering Realization

In order to implement the IMF selection via HAF, the half data set was used. As described in Section 3, 16 subjects participated in the emotion elicitation experiment and their EEG recordings, after preprocessing, resulted in a data set consisted of 16 vectors for each emotion per channel, i.e., 6 (emotions) $\times$ 16 = 96 vectors per channel in total. For those 16 vectors per emotion and per channel, the data referring to eight randomly selected subjects were used during the HAF procedure for the IMF selection. All of these 48 vectors per channel were used without considering the emotion class or the user they belong. The only consideration was to select the IMFs that, in general, are more representative of what the FF used was implying. For each generation of the GA, one FF value was assigned to each string of the population initially created. This FF value was calculated according to the whole data set of the 48 vectors per channel, i.e., for each string, the FF value was the mean of the set of values that come from each one of the 48 vectors per channel. For each one of the latter, the IMF set was extracted via EMD analysis and the quantity that the FF was realizing was calculated only for those IMFs where the corresponding positions of the string examined were set to one (see (1) and (3)). With the aforementioned procedure, the IMFs that were more likely to exhibit the attitude that the FF implied were selected beyond the emotion class or user. This process was repeated for each one of the three channels; the IMFs selected for each FF method (EFF and FDFF) are depicted cumulatively for the three channels in Fig. 6. The selected IMFs per channel and FF method were found to be: $\{(3,5,6)_{|\text{channel 1}}$;

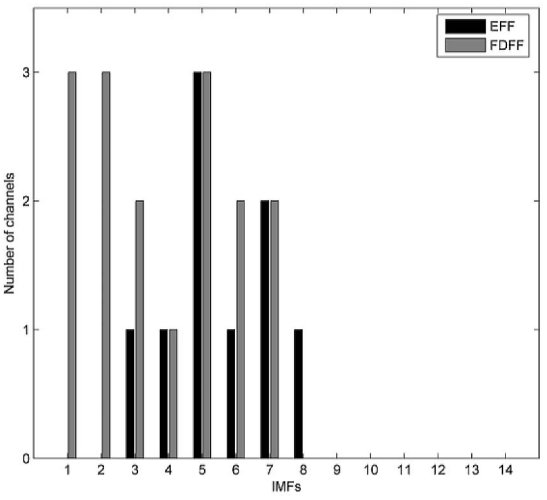

Fig. 6. Histogram of selected IMFs after the GA process.

$(4,\ 5,\ 7)_{|\text{channel 2}}; (5,\ 7,\ 8)_{|\text{channel 3}}; \}|_{\text{EFF}}$ and $\{(1-7)_{|\text{channel 1}}; (1,2,3,5,7)_{|\text{channel 2}}; (1,2,5,6)_{|\text{channel 3}}; \}|_{\text{FDFF}}$.

For the GA, the initial population was composed of 20 strings, the number of generations, which was also used as the stopping condition of the GA, was set to 10, and mutation probabilities used were $P_m(1 \rightarrow 0) = 0.01$ and $P_m(0 \rightarrow 1) = 0.001$, in order to exclude any tendency from HAF to select the majority of the IMFs.

5.3 Data Preprocessing and Feature Vector Construction

The filter used to isolate the *alpha* and *beta* bands was a 10th order IIR Butterworth bandpass filter. The filtered and averaged EEG signals were subjected to a zero-mean process by subtracting their mean value in order to apply any FV extraction method.

The FV^{HOC}(see (10)) was extracted from the whole 5 sec signal within a range of HOC order $k = 1, \ldots, 30\ (= J)$ in order to find the optimum $k\ (= L)$ that would result in the maximum classification rate of the six basic emotions; this was set separately for each individual channel, for each combination of channels, for each feature extraction approach, and for each classification method (see the subsequent section for the justification of the L selection). Thus, the following FVs were constructed:

- $FV_I^{R,HOC} = [D_1,\ D_2, \ldots, D_L]$ and $FV_I^{NR,HOC} = [D_1^1, \ldots, D_k^1, \ldots, D_1^v, \ldots, D_k^v]$, where v is the number of the selected IMFs (see, for example, Fig. 5) for the individual channel case and for the R and NR-case, respectively, and
- $$FV_C^{R,HOC} = \left[D_1^i, D_2^i, \ldots, D_L^i, D_1^j, D_2^j, \ \ldots\ , D_L^j, D_1^s, D_2^s, \ldots, D_L^s\right], i = 1,2; j = 2,3; s = 0,3, i \neq j, j \neq s,$$

  where i, j, s denote the channel number that participates to the combination; D_l^0, $l = 1, 2, \ldots, L$, corresponds to a null vector element. The $FV_C^{R,HOC}$ consisted of subfeature vectors that were extracted from a certain channel through a reconstruction phase according to the IMFs that were selected for the particular channel (see Section 5.2) that contributes to the combination.

For the combined channel case, the NR-case was not considered as it would lead to a lengthy FV and the whole classification process would be "contaminated" by the *curse of dimensionality*. For instance, for the extreme case of the combination of all three channels and $k = 30$, the resulted FV would have a length of 630 features (for 7 out of 14 IMFs selected, as is the case for channel 1 and FDFF).

5.4 Classification Setup

The classification procedure realized in the remaining data from the eight subjects whose EEG recordings did not contribute to the IMF selection, i.e., were not used for the HAF implementation. For a spherical evaluation of the recognition power of the HAF-HOC, the classification process was conducted for all possible combinations of the six basic emotions in groups of five, four, three, and two emotions, respectively, for all channels separately, for all combinations of channels, for the two feature vector cases, and for all classifiers. In all cases, the averaged EEG signals from five and three subjects were used as training and test sets, respectively. In QDA, equal prior probabilities (p_q in (11)) were assumed for each emotion class. After testing, 3-NN was used as classifier for the k-NN case, whereas the SVM kernel function parameter p in (14) was set as $p = 5$. Regarding SVM, according to the *one-versus-all* method, six SVMs that correspond to each of the six universal emotions were used. The ith SVM was trained with all of the training data in the ith class with positive labels and all other training data with negative labels. During the emotion recognition procedure, the FV was simultaneously fed into all SVMs and the SVM that output a positive label was chosen and the class of the SVM indicated the recognition result.

In order to test the classification efficiency of the HOC-based feature vector in an exhaustive way, iterative classification setup was formed. The *leave-n-out* cross-validation technique was adopted and the classification process was executed for 56 iterations in order to evaluate all possible combination of training and testing sets (i.e., $\binom{8}{3} = \frac{8!}{3!(8-3)!} = 56$). The classification results from each iteration were consequently averaged out, resulting in the final mean classification rate, i.e., $\overline{C}$ (in percent).

6 Results

The performance of the proposed HAF-HOC scheme is evaluated through the experimental results derived from the application of the introduced EEG-ER method to the experimental data set described in Section 3.3. To this end, results from every individual channel, i.e., channel 1 (Fp1), channel 2 (Fp2), and channel 3 (F3/F4), and all possible combinations of them, i.e.,

$$\text{CB1} = \{\text{channel 1(Fp1)}, \text{channel 2(Fp2)}\},$$
$$\text{CB2} = \{\text{channel 1(Fp1)}, \text{channel 3(F3/F4)}\},$$
$$\text{CB3} = \{\text{channel 2(Fp2)}, \text{channel 3(F3/F4)}\}, \text{ and}$$
$$\text{CB4} = \{\text{channel 1(Fp1)}, \text{channel 2(Fp2)}, \text{channel 3(F3/F4)}\},$$

under both HAF-HOC implementation scenarios (i.e., the R and NR-case) are presented in this section.

6.1 Individual-Channel Case

For the individual-channel case, the $\overline{C_I^R}$ and $\overline{C_I^{NR}}$ (in percent) values were calculated for the R and NR-case, respectively, for the HOC order k (up to $J = 30$) and for the four aforementioned classifiers, i.e., QDA, k-NN, MD, and SVM. In this way, the identification of the L value (optimum HOC order) that provides the maximum classification efficiency and defines that $FV_I^{R,HOC}$ and $FV_I^{NR,HOC}$ is feasible.

6.1.1 Reconstruction Case

For the R-case, the $\overline{C_I^R}$ values for the class set of six distinct emotions are depicted in Fig. 7a for all classifiers, all individual channels, for the two FFs used (EFF (black line) and FDFF (gray line)), and for $k = 2, \ldots, 30$. Clearly, the application of the QDA classification method to channel 2 data derives the highest $\overline{C_I^R}$ value for $L = 24$. Channel 2 also provides with the best performance for the Mahalanobis classifier, whereas channel 1 and channel 3 outperform for 3-*NN* and SVM classifiers, respectively. Particularly, the selected L values are {24 (QDA), 4 (k-NN), 6 (MD), 4 (SVM)} corresponding to $\overline{C_I^R}$ values of {73.86 percent (QDA), 41.14 percent (k-NN), 64.98 percent (MD), 39.64 percent (SVM)}. From an overall perspective of Fig. 7a, it is evidenced the existence of a relative dependence between k order and $\overline{C_I^R}$ value, mostly seen (as $\overline{C_I^R}$ variation) in the QDA and Mahalanobis classification methods. This fact is realized to all three channels, explaining the need for the selection of the best L value according to the classification performance.

Table 2 tabulates the confusion matrix for the class set of the six emotions, i.e., *happiness* (H), *anger* (A), *fear* (F), *surprise* (S), *disgust* (D), and *sadness* (s). The classification results refer to the overall best classification performance for the R-case, following the format of {classification method, channel, HOC k order, FF} that corresponds to {QDA, channel 2, 24, FDFF}. The distribution of the classification results across the six emotions tabulated in Table 2 displays the correct classification (percentages in bold in diagonal) and the misclassification (percentages out of diagonal) rate, respectively. For comparison reasons, the classification results when using FV_I^{HOC-EC}, FV_I^S, and FV_I^W, corresponding to the HOC-EC [15], S-EC [37], and W-EC [42] approaches, respectively (see the Appendix), are also included in this table, presented in the format of (percent/percent/percent), i.e., (percent from HOC-EC/percent from S-EC/percent from W-EC), respectively. The tabulated percentages correspond to the $\overline{C_I^R}$ derived when averaging the classification rates across the 56 iterations. The S-EC and W-EC approaches were selected as baseline ones, as they were the same ones used in HOC-EC [15]; hence, a direct comparison could be carried out with the proposed HAF-HOC approach.

Finally, for a spherical representation of HAF-HOC efficiency as EEG-ER technique, the maximum $\overline{C_I^R}$, i.e., $\overline{C_I^{R,max}}$, derived from all class sets consisting of emotion groups (i.e., groups of six, five, four, three, and two emotions) for all three single channels, for all classifiers, and for the two FFs used was estimated; these results are depicted in Fig. 7b. From the latter, it is noticed that the highest $\overline{C_I^{R,max}}$ values among all of the classification

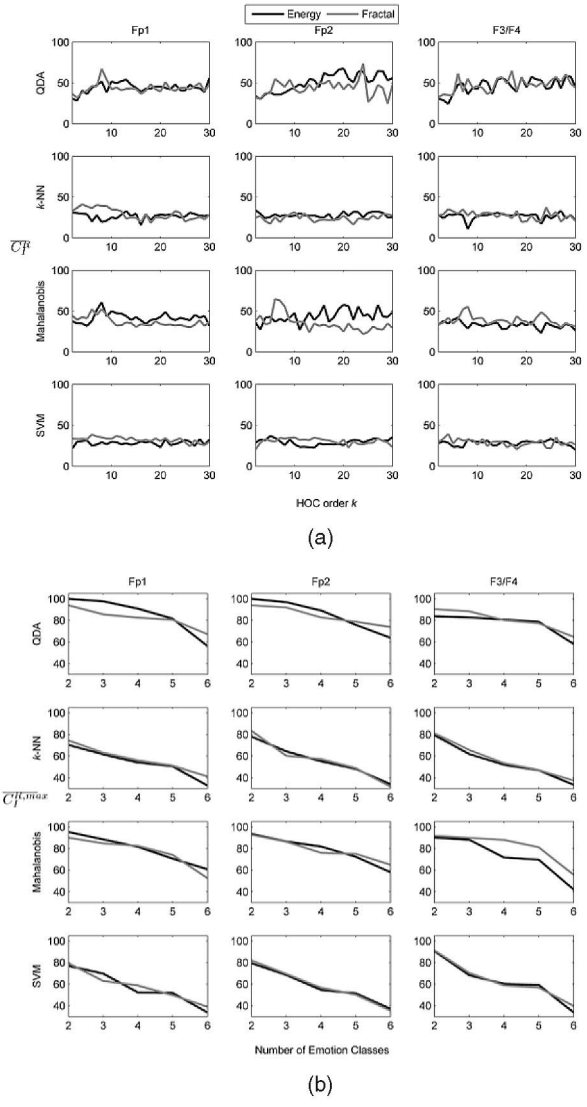

Fig. 7. (a) $\overline{C_I^R}$ values using both EFF and FDFF, for the individual-channel case for all classification methods versus the k order of the HOC analysis. (b) $\overline{C_I^{R,max}}$ values for the proposed method (HAF-HOC) using both EFF and FDFF, for the individual-channel case, for all classification methods and all possible combinations of emotion classes.

methods for different emotion class sets of five, four, three, and two emotions are exhibited by the QDA classification algorithm for channel 1, that is, {81.58 percent, 90.87 percent, 97.78 percent, 100 percent} for class sets {(H,S,A,F,s), (H,S,A,D), (H,D,S), (A,F)}. Also, a 100 percent classification accuracy is accomplished for the class set {(S,D)} for channel 2 and the QDA classification method.

6.1.2 No-Reconstruction Case

The procedure followed in the R-case was also adopted for the NR-case. More specifically, the $\overline{C_I^{NR}}$ values for the class set of six distinct emotions are depicted in Fig. 8a for all classifiers, all individual channels, for the two FFs used (EFF (black line) and FDFF (gray line)), and for $k = 2, \ldots, 30$. Fig. 8a clearly shows that, as in the R-case, channel 2 depicts the best performance regarding the value of $\overline{C_I^{NR}}$. Nevertheless, the most efficient classification

TABLE 2
$\overline{C_I^R}$ Values (in Percent) of Each Emotion of the Proposed HAF-HOC (FDFF) Compared with HOC-EC, S-EC, and W-EC Using the QDA Classifier for Channel 2

Emotions	*Happiness*	*Surprise*	*Anger*	*Fear*	*Disgust*	*Sadness*
Happiness	**71.34**	0	0	0	20	8.66
	(54.17/35/46.4)	(8.33/25/21.42)	(0/5/10.74)	(12.50/0/3.57)	(12.50/15/8.92)	(12.50/20/8.92)
Surprise	0	**71.34**	14.33	14.33	0	0
	(0/22.73/17.86)	**(75/36.36/35.71)**	(0/4.55/17.86)	(0/15.91/10.71)	(0/9.09/12.50)	(25/11.36/5.36)
Anger	0	0	**77.33**	0	11.56	11.11
	(7.14/16.67/25)	(0/12.50/15.62)	**(50/27.08/25)**	(25/8.33/9.38)	(17.86/12.50/9.38)	(0/22.92/15.62)
Fear	0	11.43	0	**82.67**	2.79	3.11
	(3.57/16.67/25)	(0/25/16.67)	(21.43/4.16/0)	**(53.57/29.17/41.67)**	(7.14/0/8.33)	(14.29/25/8.33)
Disgust	0	0	10.20	6.02	**78.67**	5.11
	(0/20.83/0)	(0/16.67/0)	(0/4.33/0)	(12.50/8.17/0)	**(62.50/29.17/50)**	(25/20.83/50)
Sadness	13.17	0	8.33	0	16.67	**61.83**
	(10.72/7.14/0)	(0/7.14/0)	(3.57/0/0)	(7.14/35.71/75)	(0/21.43/0)	**(78.57/28.58/25)**

All $\overline{C_I^R}$ values are derived from a 56-iteration cross validation process. The format (%/%/%) corresponds to the $\overline{C_I^R}$ values derived from HOC-EC, S-EC and W-EC, respectively, i.e., (HOC-EC/S-EC/W-EC).

method now appears to be the SVM classifier for $L = 12$. Moreover, channel 2 also provides the best performance for the rest of the classifiers, i.e., QDA, 3-*NN*, and MD classifiers. Particularly, the selected L values are {22 (QDA), 2 (k-NN), 22 (MD), 12 (SVM)} corresponding to $\overline{C_I^{NR}}$ values of {70.29 percent (QDA), 41.51 percent (k-NN), 41.64 percent (MD), 77.66 percent (SVM)}.

In Table 3, the confusion matrix for the class set of the six basic emotions is depicted for the overall best classification performance for the NR-case, i.e., {SVM, channel 2, 12, FDFF}. Again, for comparison reasons, the classification results when using the HOC-EC, S-EC, and W-EC approaches (see the Appendix) are also included in the same format as Table 2. Finally, the maximum $\overline{C_I^{NR}}$, i.e., $C_I^{NR,max}$, derived from all class sets consisting of emotion groups (i.e., groups of six, five, four, three, and two emotions) for all three single channels, for all classifiers, and for the two FFs used was estimated; these results are depicted in Fig. 8b. According to the results depicted in Fig. 8b, the highest $C_I^{NR,max}$ values among all of the classification methods for different emotions class sets of five, four, three, and two emotions correspond to {88.15 percent, 96.67 percent, 100 percent, 100 percent} for the class sets of $\{(\mathrm{H,S,A,D,s})|_{\text{channel 1}}, (\mathrm{H,F,D,s})|_{\text{channel 2}}, [(\mathrm{F,D,s})], \text{and } (\mathrm{D,s})]|_{\text{channel 1-3}}$, corresponding to classifiers {QDA, SVM, SVM}, respectively.

6.2 Combined-Channel Case

The evaluation of HAF-HOC in emotion recognition for the combined-channel data followed the same procedure as in individual-channel data. It should be remembered that for the combined-channel case, the R-case was only considered, to avoid the *curse of dimensionality* effect to the classification process, due to the overlength of the resulting FV_C^{NR}.

In accordance with the individual-channel case, the $\overline{C_C^R}$ values for the class set of six distinct emotions are depicted in Fig. 9a for all classifiers, all individual channels, for the two FFs used (EFF (black line) and FDFF (gray line)) and for $k = 2, \ldots, 30$. From Fig. 9a, the highest $\overline{C_C^R}$ values for each classification method are derived, corresponding to percentages of {77.28 percent (QDA), 41.84 percent (k-NN),

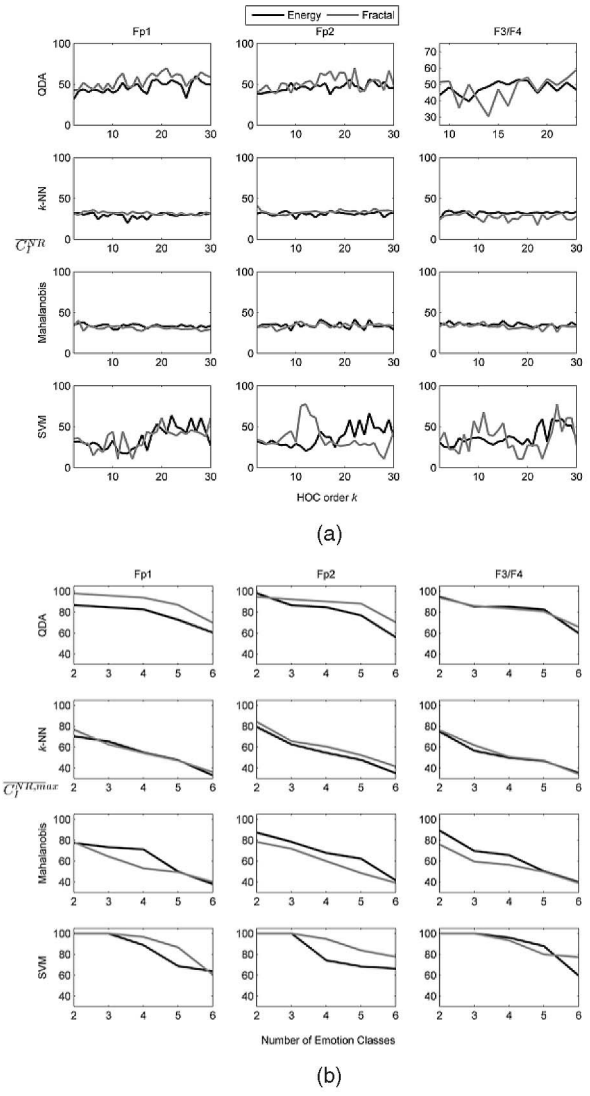

Fig. 8. (a) $\overline{C_I^{NR}}$ values using both EFF and FDFF, for the individual-channel case for all classification methods versus the k order of the HOC analysis. (b) $C_I^{NR,max}$ values for the proposed method (HAF-HOC) using both EFF and FDFF, for the individual-channel case, for all classification methods and all possible combinations of emotion classes.

TABLE 3
$\overline{C_I^{NR}}$ Values (in Percent) of Each Emotion of the Proposed HAF-HOC (FDFF) Compared with HOC-EC, S-EC, and W-EC Using the SVM Classifier for Channel 2

Emotions	*Happiness*	*Surprise*	*Anger*	*Fear*	*Disgust*	*Sadness*
Happiness	**81.74**	12.33	5.93	0	0	0 (12.50/20/8.92)
	(54.17/35/46.4)	(8.33/25/21.42)	(0/5/10.74)	(12.50/0/3.57)	(12.50/15/8.92)	
Surprise	10.06	**74.05**	6.13	0	9.76	0
	(0/22.73/17.86)	**(75/36.36/35.71)**	(0/4.55/17.86)	(0/15.91/10.71)	(0/9.09/12.50)	(25/11.36/5.36)
Anger	13.56	0	**86.44**	0	0	0
	(7.14/16.67/25)	(0/12.50/15.62)	**(50/27.08/25)**	(25/8.33/9.38)	(17.86/12.50/9.38)	(0/22.92/15.62)
Fear	0	10.9	0	**84.21**	4.89	0
	(3.57/16.67/25)	(0/25/16.67)	(21.43/4.16/0)	**(53.57/29.17/41.67)**	(7.14/0/8.33)	(14.29/25/8.33)
Disgust	0	0	13.18	5.93	**80.89**	0
	(0/20.83/0)	(0/16.67/0)	(0/4.33/0)	(12.50/8.17/0)	**(62.50/29.17/50)**	(25/20.83/50)
Sadness	10.33	13.67	7.57	9.76	0	**58.67**
	(10.72/7.14/0)	(0/7.14/0)	(3.57/0/0)	(7.14/35.71/75)	(0/21.43/0)	**(78.57/28.58/25)**

All $\overline{C_I^{NR}}$ values are derived from a 56-iteration cross validation process. The format (%/%/%) corresponds to the $\overline{C_I^{NR}}$ values derived from HOC-EC, S-EC and W-EC, respectively, i.e., (HOC-EC/S-EC/W-EC).

60.29 percent (MD), 85.17 percent (SVM)} for respective {16 (QDA), 8 (k-NN), 6 (MD), 23 (SVM)} L values and for data referring to {CB1 (QDA), CB4 (k-NN), CB1 (MD), CB4 (SVM)} channel combinations. The best classification performance is accomplished for the SVM classifier with $L = 23$ and for the channel combination CB4.

Table 4 depicts the confusion matrix of the six basic emotions for the best case of {SVM, 23, CB4}. The classification results when using FV_C^{HOC-EC}, FV_C^S, and FV_C^W are also included, as in previous tables (Tables 2 and 3). Moreover, the maximum $\overline{C_C^R}$, i.e., $\overline{C_C^{R,max}}$, derived from all class sets consisting of emotion groups (i.e., groups of six, five, four, three, and two emotions) for all four channel combinations, for all classifiers, and for the two FFs used was estimated; these results are depicted in Fig. 9b. From this figure, the highest $\overline{C_C^{R,max}}$ values among all of the classification methods for different emotions class sets of five, four, three, and two emotions correspond to {96.67 percent, 100 percent, 100 percent, 100 percent} for the class sets of {(H,S,A,D,s), (H,S,A,D), (H,S,A), [(H,S), (H,A), (S,A)]}, respectively. All of these $\overline{C_C^{R,max}}$ values refer to the SVM classification method and channel combination CB4.

6.3 Overall Performance

In order to evaluate the HAF-HOC scheme from an overall performance perspective, Table 5 was constructed. In particular, Table 5 presents the best $\overline{C_I^{max}}$ and $\overline{C_C^{max}}$ values for the hardest case of differentiating among the six emotions, derived using the HAF-HOC, HOC-EC, S-EC, and W-EC methods, both for the individual and combined-channel cases. The corresponding channel number and channel combination along with the R and NR-case approaches and the relevant FF and classifier types are also given.

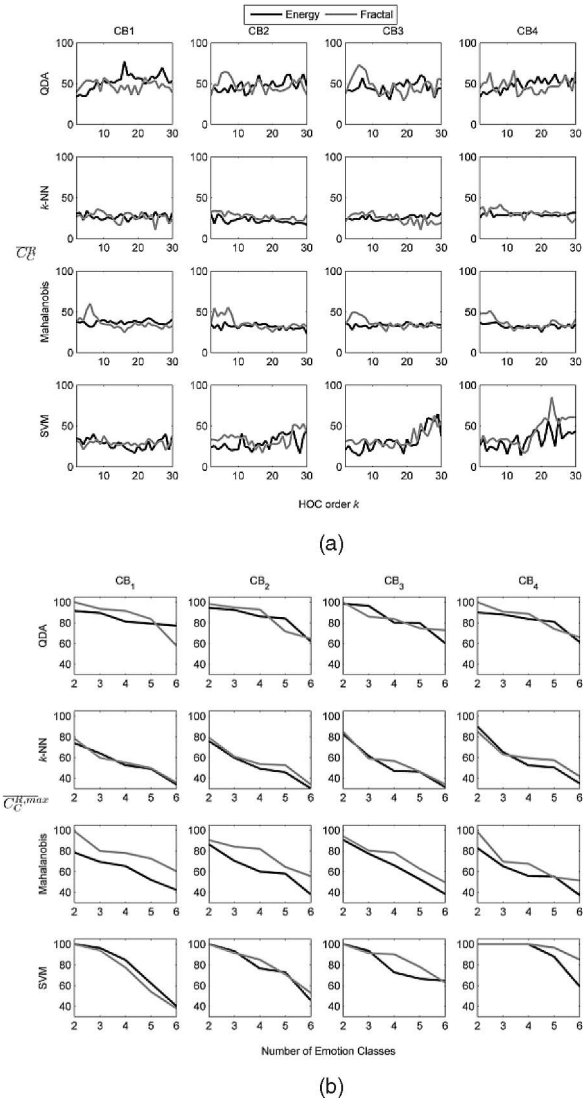

Fig. 9. (a) $\overline{C_C^R}$ values using both EFF and FDFF, for the combined-channel case for all classification methods versus the k order of the HOC analysis. (b) $\overline{C_C^{R,max}}$ values for the proposed method (HAF-HOC) using both EFF and FDFF, for the combined-channel case for all classification methods and all possible combinations of emotion classes.

7 Discussion

The presented results have shown the ability of the proposed HAF-HOC scheme to achieve a user-independent overall classification of 85.17 percent for six emotions (see

TABLE 4
$\overline{C_C^R}$ Values (in Percent) of Each Emotion of the Proposed HAF-HOC (FDFF) Compared with HOC-EC, S-EC, and W-EC Using the SVM Classifier for the Combined Channel CB4

Emotions	*Happiness*	*Surprise*	*Anger*	*Fear*	*Disgust*	*Sadness*
Happiness	**100**	0	0	0	0	0
	(100/29.17/25)	(0/29.17/10)	(0/6.25/15)	(0/10.42/15)	(0/14.58/20)	(0/10.42/15)
Surprise	0	**72.33**	0	27.67	0	0
	(0/50/10)	**(25/25/30)**	(0/12.5/10)	(75/0/35)	(0/12.5/10)	(0/0/5)
Anger	0	0	**96.67**	0	3.33	0
	(0/25/25)	(0/25/20)	**(100/25/25)**	(0/0/0)	(0/25/15)	(0/0/15)
Fear	0	13.42	0	**79.22**	7.36	0
	(0/25/12.5)	(0/12.5/12.5)	(25/0/37.5)	**(75/37.5/37.5)**	(0/0/0)	(0/25/0)
Disgust	0	0	3.89	0	**96.11**	0
	(0/0/25)	(0/12.5/0)	(0/25/0)	(0/12.5/25)	**(100/25/25)**	(0/25/25)
Sadness	0	0	16.67	8.33	8.33	**66.67**
	(0/47.5/37.5)	(0/15/6.25)	(0/12.5/18.75)	(0/12.5/6.25)	(0/12.5/6.25)	**(100/0/25)**

All $\overline{C_C^R}$ values are derived from a 56-iteration cross validation process. The format (%/%/%) corresponds to the $\overline{C_C^R}$ values derived from HOC-EC, S-EC and W-EC, respectively, i.e., (HOC-EC/S-EC/W-EC).

Table 5), using the EEG potentials from three combined channels stemming from emotion stimulus. It should be noted that the user-independency is not affected by the implied dependency on the selection of the channel (individual or combinations), reconstructed or not EEG signals, classifier type, and HOC order of the HAF-HOC scheme; as in all of these scenarios, the EEG signals from *all* users were employed for the estimation of the optimum performance scenarios. Moreover, these results support the view that the emotion information hidden in the EEG signal during emotion elicitation could be efficiently surfaced by appropriate processing and used as a means for creating more effective HMI systems.

As the experimental results have shown (see Section 6), the HAF-HOC (FDFF-based) implementation exhibited (on average) higher classification performance than the HAF-HOC (EFF-based) one. This potentially implies that the nonlinear dynamic behavior of the EEG signal, as reflected in the corresponding FD at the EMD domain, is more sensitive than the energy distribution in capturing the variation in the structural characteristics of the EEG signal (through the variation in its complexity) due to the emotion domination. This increased sensitivity of the estimated FD incorporated into the FDFF of the GA part plays the role of a "fine-tuning" procedure embedded within the HAF module, contributing to the enhancement of the overall performance of the HAF-HOC scheme.

TABLE 5
Best $\overline{C_I^{\max}}$ and $\overline{C_C^{\max}}$ Values (in Percent) for the Six Basic Emotions of the Proposed HAF-HOC Compared with HOC-EC, S-EC, and W-EC

Method	Analyzed Case	
	Individual-Channel Case	Combined-Channels Case
	$\overline{C_I^{max}}$(%)	$\overline{C_C^{max}}$(%)
HAF-HOC (FDFF)	77.66 (SVM, Fp2, NR)	85.17 (SVM, CB_4, R)
HAF-HOC (EFF)	67.89 (QDA, Fp2,R)	77.28 (QDA, CB_1, R)
HOC-EC	62.30 (QDA, F3/F4)	83.33 (SVM, CB_4)
S-EC	37.50 (MD, F3/F4)	44.90 (MD, CB_2)
W-EC	34.60 (*3*-NN, Fp2)	32.70 (QDA, CB_3)

The corresponding classifiers, channels (or channel combinations), and R- or NR-case are given in parentheses.

As was initially examined in the realization of HOC-EC [15] and thoroughly presented here, the HOC sequence provides an effective domain for feature extraction in order to construct an efficient EEG-ER system. The contribution of the HAF part, however, in fostering the HOC-based FV potential is significant, especially in the individual-channel case; this is clearly deduced from the results tabulated in Table 5. According to the latter, there is an improvement of 15.36 and 1.84 percent for the individual and combined-channel cases, respectively, when moving from the HOC-EC to the HAF-HOC (FDFF-based) scheme. The latter clearly surpasses the S-EC and W-EC (see the Appendix) feature extraction techniques, exhibiting an almost doubled (on average) classification performance.

The aforementioned application scenarios presuppose adaptation of the HAF-HOC scheme in terms of its optimization setup as far as the best HOC order for each classification scenario is concerned. Although trained HAF-HOC requires short (5 sec) epochs of the EEG signal to come up with classification outcome, the training process is the most time-consuming part of the whole approach due to the dependence of the HOC order on different channels and different channel combinations. Nevertheless, the classification part assumes that a given HOC order k is executed in a satisfactory execution time by means of real-time applications, especially for small k values, i.e., $k < 20$. For example, in a PC Intel Core2Quad 9650 (3 GHz), the estimation of the FV^{HOC} for a 5 s-EEG signal requires an execution time less than 0.5 sec for $k < 20$, according to an experimentally found execution curve of $10^{-3}k^{1.88} + 0.04,\ 2 \leq k \leq 30$. Toward that, dedicated hardware (with embedded filtering modules and processing power), along with enhancement of EEG electrodes from their current simplified version to a wirelessly miniaturized form, could facilitate integration of the HAF-HOC within AC expert systems, almost transparently to the user.

Apart from the increase in the classification performance, an additional improvement of the HAF-HOC over the

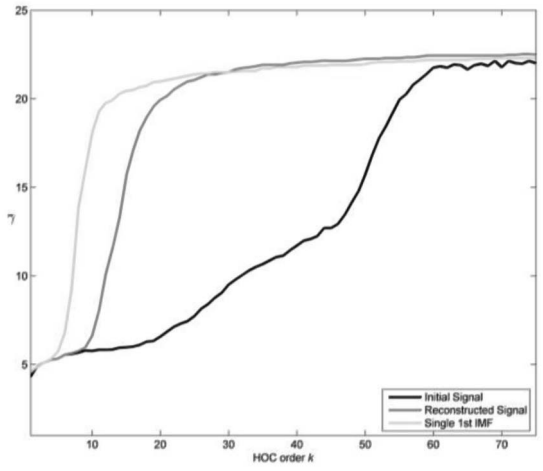

Fig. 10. $\overline{\gamma}$ values versus the k order of the HOC analysis for the initial EEG signal acquired from channel 2 and corresponding to the emotion of anger (black line), for the reconstructed EEG signal using 1, 2, 3, 5, and 7 IMFs (gray line) and for the single first IMF signal (light gray line).

HOC-EC scheme was the reduction in the FV length, adding in the realization flexibility of the HAF-HOC. As was explained in Section 5.3, the length of the FV depends on the selection of the L order, with $L \leq J$. For the case of HAF-HOC, J was set equal to 30, whereas for the case of HOC-EC, the value of $J = 50$ was used [15]. This reduction in the upper bound of the k order range was adopted by examining the way the HOC sequence monitors the dominant frequencies existing in the analyzed EEG spectra (see Fig. 4). In particular, as presented in Section 4.2.1, $\overline{\gamma}$ provides a weighted average of the spectral mass distributed over the interval $[0, f_s/2]$; hence, by plotting $\overline{\gamma}$ versus k one could identify the sensitivity of the HOC sequence to follow the distribution of the spectral mass, thus the dominant frequencies in the spectral representation of the analyzed EEG signals, under different emotion scenarios. An example of the relationship between $\overline{\gamma}$ and k, for the HOC analysis of the initial EEG signal acquired from channel 2 and the corresponding to emotion of *anger* (black line; as used in HOC-EC), the reconstructed EEG signal using 1, 2, 3, 5, 7 IMFs (gray line, as used in HAF-HOC (R-case)), and the single first IMF signal (light gray line, as used in HAF-HOC (part of the NR-case)), is given in Fig. 10. From the latter, it is noticed that under both HAF-HOC realization scenarios, i.e., the R and NR-case, the adoption of $k > 30$ does not offer any additional information, as $\overline{\gamma}$ convergences to a constant value; this is clearly achieved for higher k values ($k > 55$) for the case of HOC-EC. This is a direct consequence of the HAF functioning, as it only keeps those IMFs of the EMD-based EEG analysis that are the most effective in the representation of the emotion information, under both the R and NR-case (see also Fig. 6); this IMFs selection isolates the corresponding dominant modes, accelerating the convergence of $\overline{\gamma}$ to a constant value with the increase of k, hence reducing the effective value of J, and as a result, the time-consuming issues discuses earlier in this section. It should be noted that a similar relationship between $\overline{\gamma}$ and k as the one shown in Fig. 10 was also noticed in all examined cases (channels/emotions).

8 Conclusion

A new method, namely, HAF-HOC, for emotion recognition using three-channel EEG signals recorded under an emotion elicitation protocol has been presented in this paper. The proposed scheme combined HAF with HOCs analysis to initially process the EEG signals for enhancing the underlying emotion information, by incorporating EMD and GA-based techniques and then applying a feature extraction analysis that resulted in a HOC-based feature vector with increased classification potential. The adopted emotion elicitation technique was based on the Mirror Neuron System as a means to achieve more effective representation of the emotion stimulus to the evoked brain potentials. EEG data from individual-channel and combined-channel cases from 16 subjects were subjected to HAF-HOC analysis and four different classifiers (QDA, k-NN, MD, and SVM) were used to provide with the classification results under a user-independent scenario for the classification of six basic emotions, i.e., *happiness*, *surprise*, *anger*, *fear*, *disgust*, and *sadness*. For the individual-channel case, the best results were obtained by the QDA (77.66 percent mean classification rate), whereas for the combined-channel case, the best results were obtained using SVM (85.17 percent mean classification rate), for the hardest case of differentiating among the six basic emotions. Extended classification tests were implemented concerning all possible combinations of emotion classes, showing the effectiveness of the proposed method, which gradually achieved maximum (100 percent) classification rates as soon as the number of the classes of emotions was reduced from six to two. The comparison of the introduced HAF-HOC method with other relevant feature extraction and emotion recognition methods (i.e., HOC-EC, S-EC, and W-EC) overwhelmed its robustness and consistency to more effectively discriminate emotions from EEG signals.

Appendix

The statistical features [37] used to form the proposed FVs are defined as:

1. the mean of the raw signal (μ_X),
2. the standard deviation of the raw signal (σ_X),
3. the mean of the absolute values of the first differences of the raw signal (δ_X),
4. the mean of the absolute values of the first differences of the standardized signal ($\overline{\delta}_X$),
5. the mean of the absolute values of the second differences of the raw signal (γ_X), and
6. the mean of the absolute values of the second differences of the standardized signal ($\overline{\gamma}_X$).

The corresponding FVs for the individual and combined-channel cases are defined as

$$FV_I^S = [\mu_X, \sigma_X, \delta_X, \overline{\delta}_X, \gamma_X, \overline{\gamma}_X], \tag{15}$$

$$FV_C^S = [FV_I^{S,i}, FV_I^{S,j}, FV_I^{S,s}], \quad \begin{array}{l} i = 1, 2; j = 2, 3; \\ s = 0, 3, i \neq j, \quad j, \neq s, \end{array} \tag{16}$$

where i, j, s denote the channel number that participates to the combination; $FV_I^{S,0}$ corresponds to a null vector element.

The wavelet-oriented FV [42] is based on the wavelet energy (ENG_l) and wavelet entropy (ENT_l). These para-

meters were used as an FV, defined for the individual and combined-channel cases as

$$FV_I^W = [ENG_l, ENG_l], \tag{17}$$

$$FV_C^W = \left[FV_I^{W,i}, FV_I^{W,j}, FV_I^{W,s}\right], \quad i = 1,2; \quad j = 2,3; \quad s = 0,3, i \neq j, \quad j, \neq s, \tag{18}$$

where i, j, s denote the channel number that participates to the combination; $FV_I^{W,0}$ corresponds to a null vector element.

In the HOC-EC [15], the HOC sequence up to the 50th order was used to form the FV, for the individual- and combined-channel cases, i.e.,

$$FV_I^{HOC-EC} = [D_1, D_2, \ldots, D_L], 1 < L \leq 50. \tag{19}$$

$$FV_C^{HOC-EC} = \left[FV_I^{HOC-EC,i}, FV_I^{HOC-EC,j}, FV_I^{HOC-EC,s}\right], \quad i = 1,2; j = 2,3; s = 0,3, i \neq j, j \neq s, \tag{20}$$

where i, j, s denote the channel number that participates to the combination; $FV_I^{HOC-EC,0}$corresponds to a null vector element.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to all 16 subjects who participated in the experiment.

REFERENCES

[1] B. Reeves and C. Nass, *The Media Equation. How People Treat Computers, Television, and New Media Like Real People and Places.* CSLI, Cambridge Univ. Press, 1996.
[2] T. Bradberry and J. Greaves, *Emotional Intelligence 2.0.* Publishers Group West, 2009.
[3] R.W. Picard, *Affective Computing.* MIT Press, 1997.
[4] I. Cohen, A. Garg, and T.S. Huang, "Emotion Recognition from Facial Expressions Using Multilevel HMM," *Proc. Neural Information Processing Systems Workshop Affective Computing,* http://www.ifp.uiuc.edu/~ashutosh/papers/NIPS_emotion.pdf, 2000.
[5] F. Bourel, C.C. Chibelushi, and A.A. Low, "Robust Facial Expression Recognition Using a State-Based Model of Spatially-Localized Facial Dynamics," *Proc. Fifth IEEE Int'l Conf. Automatic Face and Gesture Recognition,* pp. 106-111, 2002.
[6] J.J. Lien, T. Kanade, J.F. Cohn, and C. Li, "Automated Facial Expression Recognition Based on FACS Action Units," *Proc. Third IEEE Conf. Automatic Face and Gesture Recognition,* pp. 390-395, 1998.
[7] B. Schuller, S. Reiter, R. Mueller, M. Al-Hames, and G. Rigoll, "Speaker Independent Speech Emotion Recognition by Ensemble Classification," *Proc. Sixth Int'l Conf. Multimedia and Expo,* pp. 864-867, 2005.
[8] C. Busso, Z. Deng, S. Yildirim, M. Buut, C.M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of Emotion Recognition Using Facial Expressions, Speech and Multimodal Information," *Proc. Sixth Int'l Conf. Multimodal Interfaces,* pp. 205-211, 2004.
[9] R.W. Picard, E. Vyzas, and J. Healey, "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 10, pp. 1175-1191, Oct. 2001.
[10] F. Nasoz, C.L. Lisetti, K. Alvarez, and N. Finkelstein, "Emotion Recognition from Physiological Signals for User Modeling of Affect," *Proc. Ninth Int'l Conf. User Model,* http://www.eurecom.fr/util/publidownload.en.htm?id=1806, June 2003.
[11] C.L. Lisetti and F. Nasoz, "Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals," *EURASIP J. Applied Signal Processing,* vol. 11, pp. 1672-1687, 2004.
[12] D.N. McIntosh, A. Reichmann-Decker, P. Winkielman, and J.L. Wilbarger, "When the Social Mirror Breaks: Deficits in Automatic, But Not Voluntary, Mimicry of Emotional Facial Expressions in Autism," *Developmental Science,* vol. 9, pp. 295-302, 2006.
[13] D.E. Goldberg, *Genetic Algorithms in Search, Optimization & Machine Learning.* Addisson-Wesley, 1989.
[14] N. Huang, Z. Shen, S. Long, M. Wu, H.H. Shih, N.C. Zheng, N.C. Yen, C. Tung, and H. Liu, "The Empirical Mode Decomposition and Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis," *Proc. Royal Soc. London A,* vol. 454, pp. 903-995, 1998.
[15] P.C. Petrantonakis and L.J. Hadjileontiadis, "Emotion Recognition from EEG Using Higher Order Crossings," *IEEE Trans. Information Technology in Biomedicine,* vol. 14, no. 2, pp. 186-197, Mar. 2010.
[16] G. Rizzolatti and L. Craighero, "The Mirror-Neuron System," *Ann. Rev. Neuroscience,* vol. 27, pp. 169-192, 2004.
[17] W.B. Cannon, "The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory," *Am. J. Psychology,* vol. 39, pp. 106-124, 1927.
[18] C. Darwin, *The Expression of Emotion in Man and Animals.* Philosophical Library (Original work published in 1872), 1955.
[19] W. James, *The Principles of Psychology.* Holt, Rinehart and Winston, 1890.
[20] P. Ekman, "Expression and the Nature of Emotion," *Approaches to Emotion,* K. Scherer and P. Ekman, eds., Erlbaum, 1984.
[21] P. Ekman, R.W. Levenson, and W.V. Friesen, "Emotions Differ in Autonomic Nervous System Activity," *Science,* vol. 221, pp. 1208-1210, 1983.
[22] R.J. Davidson, G.E. Schwartz, C. Saron, J. Bennett, and D.J. Goleman, "Frontal versus Parietal EEG Asymmetry During Positive and Negative Affect," *Psychophysiology,* vol. 16, pp. 202-203, 1979.
[23] R.J. Davidson, P. Ekman, C.D. Saron, J.A. Senulis, and W.V. Friesen, "Approach-Withdrawal and Cerebral Asymmetry: Emotional Expression and Brain Physiology," *J. Personality and Social Psychology,* vol. 58, pp. 330-341, 1990.
[24] H. Jasper, "The Ten-Twenty Electrode System of the International Federation," *Electroencephalography and Clinical Neurophysiology,* vol. 39, pp. 371-375, 1958.
[25] W.J.H. Nauta, "The Problem of the Frontal Lobe: A Reinterpretation," *J. Psychiatric Research,* vol. 8, pp. 167-187, 1971.
[26] R.J. Davidson, "What Does the Prefrontal Cortex 'Do' in Affect: Perspectives on Frontal EEG Asymmetry Research," *Biological Psychology,* vol. 67, pp. 219-233, 2004.
[27] D. Hagemann, E. Naumann, A. Lurken, G. Becker, S. Maier, and D. Bartussek, "EEG Asymmetry, Dispositional Mood and Personality," *Personality and Individual Differences,* vol. 27, pp. 541-568, 1999.
[28] J.A. Coan and J.J.B. Allen, "Frontal EEG Asymmetry as a Moderator and Mediator of Emotion," *Biological Psychology,* vol. 67, pp. 7-49, 2004.
[29] L.I. Aftanas, A.A. Varlamov, S.V. Pavlov, V.P. Makhnev, and N.V. Reva, "Affective Picture Processing: Event-Related Synchronization within Individually Defined Human Theta Band Is Modulated by Valence Dimension," *Neuroscience Letters,* vol. 303, pp. 115-118, 2001.
[30] G. Pfurtscheller and F.H.L. da Silva, "Event-Related EEG/MEG Synchronization and Desynchronization: Basic Principles," *Clinical Neurophysiology,* vol. 110, pp. 1842-1857, 1999.
[31] E. Oztop, M. Kawato, and M. Arbib, "Mirror Neurons and Imitation: A Computationally Guided Review," *Neural Networks,* vol. 19, pp. 254-271, 2006.
[32] T.W. Lee, R.J. Dolan, and H.D. Critchley, "Controlling Emotional Expression: Behavioral and Neural Correlates of Nonimitative Emotional Responses," *Cerebral Cortex,* vol. 8, pp. 104-113, 2008.
[33] T.W. Lee, O. Josephs, R.J. Dolan, and H.D. Critchley, "Imitating Expressions: Emotion-Specific Neural Substrates in Facial Mimicry," *Social Cognitive and Affective Neuroscience,* vol. 1, pp. 122-135, 2006.
[34] L. Carr, M. Iacoboni, M.C. Dubeau, J.C. Mazziotta, and G.L. Lenzi, "Neural Mechanisms of Empathy in Humans: A Relay from Neural Systems for Imitation to Limbic Areas," *Proc. Nat'l Academy of Sciences,* vol. 100, pp. 5497-5502, 2003.
[35] B. Wicker, C. Keysers, J. Plailly, J.P. Royet, V. Gallese, and G. Rizzolatti, "Both of Us Disgusted in My Insula: The Common Neural Basis of Seeing and Feeling Disgust," *Neuron,* vol. 40, pp. 655-664, 2003.

[36] A. Choppin, "EEG-Based Human Interface for Disabled Individuals: Emotion Expression with Neural Networks," master's thesis, Tokyo Inst. of Technology, 2000.
[37] K. Takahashi, "Remarks on Emotion Recognition from Bio-Potential Signals," *Proc. Second Int'l Conf. Autonomous Robots and Agents,* pp. 186-191, 2004.
[38] G. Chanel, J. Kronegg, D. Grandjean, and T. Pun, "Emotion Assessment: Arousal Evaluation Using EEG and Peripheral Physiological Signals," technical report, Univ. of Geneva, 2005.
[39] D.O. Bos,, "EEG-Based Emotion Recognition: The Influence of Visual and Auditory Stimuli," http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Oude_Bos-Danny.pdf, 2006.
[40] Z. Khalili and M. Moradi, "Emotion Detection Using Brain and Peripheral Signals," *Proc. Biomedical Eng. Conf.,* pp. 1-4, 2008.
[41] R. Horlings, D. Datcu, and L.J.M. Rothkrantz, "Emotion Recognition Using Brain Activity," *Proc. Int'l Conf. Computer Systems and Technologies,* pp. 1-6, 2008.
[42] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry, "Lifting Scheme for Human Emotion Recognition Using EEG," *Proc. Int'l Symp. Information Technology,* pp. 1-7, 2008.
[43] K.G. Srinivasa, K.R. Venugopal, and L.M. Patnaik, "Feature Extraction Using Fuzzy C-Means Clustering for Data Mining Systems," *Int'l J. Computer Science and Network Security,* vol. 6, pp. 230-236, 2006.
[44] J.J. De Gruijter and A.B. McBratney, "A Modified Fuzzy K Means for Predictive Classification," *Classification and Related Methods of Data Analysis,* H.H. Bock, ed., pp. 97-104, Elsevier Science, 1988.
[45] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Comm. Pure and Applied Math.,* vol. 41, pp. 909-996, 1988.
[46] A. Heraz and C. Frasson, "Predicting the Three Major Dimensions of the Learner's Emotions from Brainwaves," *Int'l J. Computer Science,* vol. 2, no. 3, pp. 187-193, 2008.
[47] K. Schaaff and T. Schultz, "Towards an EEEG-Based Emotion Recognizer for Humanoid Robots," *Proc. 18th IEEE Int'l Symp. Robot and Human Interactive Comm.,* pp. 792-796, 2009.
[48] B. Kedem, *Time Series Analysis by Higher Order Crossings.* IEEE Press, 1994.
[49] P.J. Lang, M.M. Bradley, and B.N. Cuthbert, "International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual," Technical Report A-8, Univ. of Florida, 2008.
[50] M.M. Bradley and P.J. Lang, "The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective Ratings of Sounds and Instruction Manual," Technical Report B-3, Univ. of Florida, 2007.
[51] P. Ekman et al., "Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion," *J. Personality and Social Psychology,* vol. 53, no. 4, pp. 712-717, 1987.
[52] S. D'Mello, T. Jackson, S. Craig, B. Morgan, P. Chipman, H. White, N. Person, B. Kort, R. el Kaliouby, R.W. Picard, and A. Graesser, "AutoTutor Detects and Responds to Learners Affective and Cognitive States," *Proc. Workshop Emotional and Cognitive Issues at the Int'l Conf. Intelligent Tutoring Systems,* pp. 31-43, 2008.
[53] P. Ekman and W.V. Friesen, "Pictures of Facial Affect," *Human Interaction Laboratory,* Univ. of California Medical Center, 1976.
[54] B. Graimann, J.E. Huggins, S.P. Levine, and G. Pfurtscheller, "Visualization of Significant ERD/ERS Patterns in Multichannel EEG and ECoG Data," *Clinical Neurophysiology,* vol. 113, pp. 43-47, 2002.
[55] K. Coburn and M. Moreno, "Facts and Artifacts in Brain Electrical Activity Mapping," *Brain Topography,* vol. 1, pp. 37-45, 1988.
[56] D.O. Olguin, "Adaptive Digital Filtering Algorithms for the Elimination of Power Line Interference in Electroencephalographic Signals," master's thesis, Inst. Tecnologico y de Estudios Superiores de Monterrey, 2005.
[57] M. Fatourechi, A. Bashashati, R.K. Ward, and G.E. Birch, "EMG and EOG Artifacts in Brain Computer Interface Systems: A Survey," *Clinical Neurophysiology,* vol. 118, pp. 480-494, 2007.
[58] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical Mode Decomposition as a Filter Bank," *IEEE Signal Processing Letters,* vol. 11, no. 2, pp. 112-114, Feb. 2004.
[59] P.C. Petrantonakis and L.J. Hadjileontiadis, "EEG-Based Emotion Recognition Using Hybrid Filtering and Higher Order Crossings," *Proc. Int'l Conf. Affective Computing and Intelligent Interaction,* pp. 147-152, 2009.
[60] M. Katz, "Fractals and the Analysis of Waveforms," *Computers in Biology and Medicine,* vol. 18, pp. 145-156, 1988.
[61] A. Accardo, M. Affinito, M. Carrozzi, and F. Bouquet, "Use of Fractal Dimension for the Analysis of Electroencephalographic Time Series," *Biological Cybernetics,* vol. 77, pp. 339-350, 1997.
[62] A. Petrosian, "Kolmogorov Complexity of Finite Sequences and Recognition of Different Preictal EEG Patterns," *Proc. IEEE Symp. Computer-Based Medical Systems,* pp. 212-217, 1995.
[63] T. Higuchi, "Approach to an Irregular Time Series on the Basis of the Fractal Theory," *Physica D,* vol. 31, pp. 277-283, 1988.
[64] A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* third ed. McGraw-Hill, 1991.
[65] W.J. Krzanowski, *Principles of Multivariate Analysis.* Oxford Univ. Press, 1988.
[66] T. Mitchell, *Machine Learning.* McGraw-Hill, 1997.
[67] P.C. Mahalanobis, "On the Generalized Distance in Statistics," *Proc. Nat'l Inst. of Science of India,* vol. 2, pp. 49-55, 1936.
[68] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods.* Cambridge Univ. Press, 2000.
[69] R. Duda, P. Hart, and D. Stork, *Pattern Classification.* John Wiley & Sons, 2001.

Panagiotis C. Petrantonakis received the Diploma degree in electrical and computer engineering in 2007 from Aristotle University of Thessaloniki (AUTH), Greece, where he is currently working toward the PhD degree, affiliated with the Signal Processing and Biomedical Technology Unit of the Telecommunications Laboratory. His current research interests lie in the area of advanced signal processing techniques, nonlinear transforms, and affective computing. He is a member of the Technical Chamber of Greece and a student member of the IEEE and the IEEE Computer Society.

Leontios J. Hadjileontiadis received the Diploma degree in electrical engineering and the PhD degree in electrical and computer engineering from Aristotle University of Thessaloniki, Greece, in 1989 and 1997, and the PhD degree in music composition from the University of York, United Kingdom, in 2004. He is currently a professor in composition at the State Conservatory of Thessaloniki, Greece. In December 1999, he joined as a faculty member in the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, where he is currently an associate professor, working on lung sounds, heart sounds, bowel sounds, ECG data compression, seismic data analysis, and crack detection in the Signal Processing and Biomedical Technology Unit of the Telecommunications Laboratory. His research interests are in higher order statistics, alpha-stable distributions, higher order zero crossings, wavelets, polyspectra, fractals, neurofuzzy modeling for medical, and mobile and digital signal processing applications. He is a member of the Technical Chamber of Greece, the IEEE, the Higher-Order Statistics Society, the International Lung Sounds Association, and the American College of Chest Physicians. He was the recipient of the Second Award at the Best Paper Competition of the Ninth Panhellenic Medical Conference on Thorax Diseases '97, Thessaloniki. He was also an open finalist at the Student paper Competition (Whitaker Foundation) of the IEEE EMBS '97, Chicago, Illinois, a finalist at the Student Paper Competition (in memory of Dick Poortvliet) of the MEDICON '98, Lemesos, Cyprus, and the recipient of the Young Scientist Award of the 24th International Lung Sounds Conference '99, Marburg, Germany. In 2004, 2005, and 2007, he organized and served as a mentor to three to five student teams that have ranked as third, second, and seventh worldwide, respectively, at the Imagine Cup Competition (Microsoft), Sao Paulo, Brazil (2004)/Yokohama, Japan (2005)/Seoul, Korea (2007), with projects involving technology-based solutions for people with disabilities.