Emotion Recognition From EEG Using Higher Order Crossings

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

***Abstract*—Electroencephalogram (EEG)-based emotion recognition is a relatively new field in the affective computing area with challenging issues regarding the induction of the emotional states and the extraction of the features in order to achieve optimum classification performance. In this paper, a novel emotion evocation and EEG-based feature extraction technique is presented. In particular, the mirror neuron system concept was adapted to efficiently foster emotion induction by the process of imitation. In addition, higher order crossings (HOC) analysis was employed for the feature extraction scheme and a robust classification method, namely HOC-emotion classifier (HOC-EC), was implemented testing four different classifiers [quadratic discriminant analysis (QDA), k-nearest neighbor, Mahalanobis distance, and support vector machines (SVMs)], in order to accomplish efficient emotion recognition. Through a series of facial expression image projection, EEG data have been collected by 16 healthy subjects using only 3 EEG channels, namely Fp1, Fp2, and a bipolar channel of F3 and F4 positions according to 10–20 system. Two scenarios were examined using EEG data from a single-channel and from combined-channels, respectively. Compared with other feature extraction methods, HOC-EC appears to outperform them, achieving a 62.3% (using QDA) and 83.33% (using SVM) classification accuracy for the single-channel and combined-channel cases, respectively, differentiating among the six basic emotions, i.e., *happiness*, *surprise*, *anger*, *fear*, *disgust*, and *sadness*. As the emotion class-set reduces its dimension, the HOC-EC converges toward maximum classification rate (100% for five or less emotions), justifying the efficiency of the proposed approach. This could facilitate the integration of HOC-EC in human machine interfaces, such as pervasive healthcare systems, enhancing their affective character and providing information about the user's emotional status (e.g., identifying user's emotion experiences, recurring affective states, time-dependent emotional trends).**

***Index Terms*—Electroencephalogram (EEG), emotion recognition, higher order crossings analysis, k-nearest neighbor (k-NN), Mahalanobis distance (MD), mirror neuron system, quadratic discriminant analysis, support vector machines (SVMs) .**

I. Introduction

DESPITE the difficulty to give a precise definition of emotion, it is globally accepted that everyday activities, such as social communication, decision-making, and fundamental adaptation tasks are highly influenced by people's moods and distinct emotional states. Moreover, it is evidenced [1], [2] that emotional intelligence, that is the ability to identify, assess, and manage emotions of one's self and of others, plays a crucial role in learning processes and particularly in the capability of extracting the information that is most important. Furthermore, a number of studies by neuroscientists, cognitive scientists, and psychologists have shown that emotion is of major importance in rational and intelligent thinking. More precisely, it has been shown [3] that patients without emotional brain functioning had severe impairments in everyday activities that require rational and intelligent behavior.

All the aforementioned findings have revealed the importance of not only emotional intelligence in achieving personal aspirations but also in accomplishing aggregate attainments via healthy social interaction; but what if interactions have to be between humans and machines? Nowadays, human–machine interactions (HMIs) are a vital aspect of our technologically invaded life. In order to make this collaboration more efficient and pragmatic, machines should be equipped with the ability to identify, understand, and integrate during the interaction phase the emotional needs of their human partner. There are circumstantial evidences from relative experiments [4] that human–human interaction (HHI) does not have significant differences from HMI. As a result, it is apparent that if HMI has to be as effective as HHI, an inevitable need to instill machines with emotional abilities emerges.

From a healthcare perspective, accounting for a patient's emotional state is essential in medical care [5]. In fact, not only may appropriate emotional state assessment be a key indicator of the patient's mental or physical health status, but the power of emotions themselves over the recovery process has also been documented [3]. In a tele-health scenario, clinicians should necessary focus to accurately assess patient's emotional status to enhance his/her care. Collection of physiological signals and mapping of the latter to emotional states can synthesize the patient's affective information for the healthcare provider [6]. Moreover, patients' models, not only of the physical characteristics of the patients, but also models of their emotions can be built, initiating a field of affective computing for tele-home health care applications [6]. Some examples in the latter include: the development of systems for monitoring and responding to human multimodal affect and emotions via multimedia and empathetic avatars; mapping of physiological signals to emotions and synthesizing the patient's affective information for the healthcare provider [6]; enhancement of behavioral patterns of daily activity using ontological models that embed the emotional factor, assisting diagnosis in conjunction to a person's health condition [7].

Research has already been conducted toward effective HMI by implementing methods in order to recognize emotions from face [8]–[10], voice [11]–[14], and data related to the

Manuscript received December 23, 2008; revised June 30, 2009. First published October 23, 2009; current version published March 17, 2010.

The authors are with the Signal Processing and Biomedical Technology Unit, Telecommunications Laboratory, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, GR 541 24 Thessaloniki, Greece (e-mail: ppetrant@auth.gr; leontios@auth.gr).

Digital Object Identifier 10.1109/TITB.2009.2034649

autonomous nervous system (galvanic skin response, heart rate, etc.) [15], [16]. Although, the performance of those emotion recognition techniques appear to be efficient there is a number of issues regarding their realization during the HMI interaction. For example, consider an HMI, where a human had to look directly to the camera all the time or the possibility for a machine to recognize an emotion through voice in a noisy environment. Autonomous nervous system signals can overcome the aforementioned problems, yet they are significantly influenced by other factors whose effect is similar with an emotion-derived one. For instance, perspiration due to physical activity rather than emotions would affect a galvanic skin response (GSR)-based emotion recognition system. Currently, few efforts have been initiated to recognize emotions with electroencephalogram (EEG)-based recognition systems by artificially eliciting emotional states. This approach eliminates the disadvantages of the previous emotion recognition techniques as central nervous system signals are barely influenced by the aspects marked previously.

Toward such direction, the most prominent methods employed statistical-based [17] and wavelet-based [18] analysis of EEG signals for feature extraction; combined with classification methods (such as support vector machine (SVM) [19], fuzzy k-means [20], and fuzzy c-means [21]) they have resulted in moderate emotion recognition percentages for up to three [18], four [22], and five [17] distinct emotions. Despite the classification potential of EEG for such kind of recognition reflected in the works reported in the literature [17], [18], [22]–[24], further research is needed in order to improve recognition rates and discover unknown aspects of emotion mechanisms performed in the human brain.

This paper describes the implementation of an EEG-based, user-independent emotion recognition system using a feature set drawn from higher order crossings (HOC) analysis [25]. The latter reveals the oscillatory pattern of the EEG signal providing a feature set that convey the emotion information to the classification space, where different types of classifiers use such emotion information to discriminate among different emotion states. The classification efficiency of the proposed approach, namely HOC-emotion classifier (HOC-EC), is outlined from the experimental results by the application of the HOC-EC to EEG-datasets 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 [26], which relates emotion induction with the process of imitation. Compared with previous approaches, the HOC-EC is quite attractive due to its superiority in the emotion recognition power for a combination of up to six distinct emotions.

The rest of this paper is structured as follows. Section II provides background information on the emotion elicitation theme and describes the proposed analysis. The experimental dataset used, alongside some realization issues are described in Section III. Section IV presents the HOC-EC performance through experimental results, discusses its efficiency to address the EEG-based classification problem and compares it with other works. Finally, Section V concludes the paper.

II. Method

A. Background

1) Emotions: Since different theories of how emotions arise and been expressed exist, there is a variety of emotion models used in affective computing [27]. Inspired by the Darwinian theory, Ekman *et al.* proposed the universality of six facial expressions [28], i.e., *happiness*, *surprise*, *anger*, *disgust*, *sadness*, and *fear*. On the other hand, psychologists do not present emotions as discrete states but rather as continuous ones and therefore demonstrate them in an n-dimensional (n-D) space; usually the 2-D valence/arousal space is adopted. 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. In this paper, the Ekman's six universal emotions [28] were adopted, with the latter consisting the class-set of the discrimination analysis. The selection of these emotions was based on the fact that this universal basis for emotional expression is considered a pancultural aspect of psychological functioning and is no longer debated in contemporary psychology [29].

2) Mirror Neuron System and Emotion Elicitation: Neurophysiological experiments have demonstrated that when individuals observe an action done by another individual seem to have the same brain activity, as if they did the corresponding action themselves [26]. Moreover, MNS has been connected with the ability of imitation, which also, among others, relates to recognition of emotions by others' face expressions and gestures [30]. Based on the MNS concept, the emotion elicitation process here was realized through Ekman's pictures [31] shown to subjects as a distinct emotion stimulus. According to MNS theory, the subjects' EEG signal is supposed to reflect the same or akin brain activity, as when they are really overcome by the same emotion.

According to Picard *et al.* [15], there are five factors that should be taken under consideration during the data collection stage of an experiment designed to elicit emotions.

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) *Laboratory setting versus real world:* Is subject in a laboratory 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?

Until now, the emotion evocation process in EEG-based emotion recognition experiments was depended on projection of photos that picturize event or situations which are claimed to elicit certain distinct emotions or emotional states that fall within the aforementioned valence/arousal space [32]. Although the last four factors above seem to be relatively controllable during such experiments the first one raises issues concerning the emotional impact of such pictures to the subjects. Aspects like personality, personal experiences, and particular subject's mood at the time that the experiment is conducted dramatically influence the way someone emotionally reacts in the view of that kind of images. As a result, as soon as the emotion is inducted by a stimulus outside the efforts of the subject, it is questioned how effectively the emotion is actually evoked 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 six basic emotions, i.e., *happiness*, *surprise*, *anger*, *fear*, *disgust*, and *sadness*; hence, according to the MNS theory [30], it is not expected from the subject to actually feel the emotion but express the same or akin EEG activity as s\he had really experienced it.

B. Previous Classification Features

For comparison reasons, two other feature extraction techniques for EEG-based emotion recognition were implemented and applied to the acquired signals. These methods, based on statistical values [17] and the wavelet transform [18], along with the construction of the relevant feature vectors ($\mathrm{FV_S}$), are further explained in the following.

1) Statistical-Based Features: A statistical feature vector ($\mathrm{FV_S}$) was proposed in [15] for physiological signals only. This $\mathrm{FV_S}$ was expanded in [17] by fusing the features of physiological signals with those derived from the respective EEG recordings. The statistical features used to form the proposed $\mathrm{FV_S}$ are defined as ($X(n)$, $n = 1, \ldots, N$ is the raw N-sample EEG signal) given in the following.

1) The mean of the raw signal

$$\mu_X = \frac{1}{N}\sum_{n=1}^{N} X(n). \tag{1}$$

2) The standard deviation of the raw signal

$$\sigma_X = \sqrt{\frac{1}{N}\sum_{n=1}^{N} (X(n) - \mu_X)^2}. \tag{2}$$

3) The mean of the absolute values of the first differences of the raw signal

$$\delta_X = \frac{1}{N-1}\sum_{n=1}^{N-1} |X(n+1) - X(n)|. \tag{3}$$

4) The mean of the absolute values of the first differences of the standardized signal

$$\overline{\delta}_X = \frac{1}{N-1}\sum_{n=1}^{N-1} \left|\overline{X}(n+1) - \overline{X}(n)\right| = \frac{\delta_X}{\sigma_X}. \tag{4}$$

5) The mean of the absolute values of the second differences of the raw signal

$$\gamma_X = \frac{1}{N-2}\sum_{n=1}^{N-2} |X(n+2) - X(n)|. \tag{5}$$

6) The mean of the absolute values of the second differences of the standardized signal

$$\overline{\gamma}_X = \frac{1}{N-2}\sum_{n=1}^{N-2} \left|\overline{X}(n+2) - \overline{X}(n)\right| = \frac{\gamma_X}{\sigma_X} \tag{6}$$

where $\overline{X}(n)$ is the standardized signal, i.e. $\overline{X}(n) = (X(n) - \mu_X)/\sigma_X$. The corresponding $\mathrm{FV_S}$ is defined as $\mathrm{FV_S} = [\mu_X, \sigma_X, \delta_X, \overline{\delta}_X, \gamma_X, \overline{\gamma}_X]$.

2) Wavelet-Based Features: Murugappan *et al.* [18] proposed a new methodology on the feature extraction issue from EEG signals in order to discriminate three distinct emotions, i.e., happiness, disgust, and fear. In particular, discrete wavelet transform [33] $C_X(l,n)$ for S scales using the Daubechies fourth-order orthonormal bases [34] was employed and the extracted wavelet coefficients at the lth scale that correspond to the alpha band (8–12 Hz) were used to estimate the wavelet energy and wavelet entropy, given by

$$\mathrm{ENG}_l = \sum_{n=1}^{2^{S-l}-1} |C_X(l,n)|^2, \quad N = 2^S, \quad 1 < l < S \tag{7}$$

$$\mathrm{ENT}_l = -\sum_{n=1}^{2^{S-l}-1} |C_X(l,n)|^2 \log(|C_X(l,n)|^2), \quad N = 2^S, \quad 1 < l < S. \tag{8}$$

The parameters of (7) and (8) were used as a feature vector, i.e., $\mathrm{FV_W} = [\mathrm{ENG}_l,\ \mathrm{ENT}_l]$.

C. Proposed HOC-Based Features

Almost all observed time series display local and global up and down movements as time progresses. This behavior, seen in a finite zero-mean series $\{Z_t\}, t = 1, \ldots, N$ oscillating about level zero can be expressed through the zero-crossing count. In general, when a filter is applied to a time series it changes its oscillation; hence its zero-crossing counts too. Under this perspective, the following iterative procedure could be assumed: apply a filter to the time series, and count the number of zero-crossings in the filtered time series; apply yet another filter to the original time series, and again observe the resulting zero-crossings, and so on—filter and count. The resulting zero-crossing counts are referred to as HOC [25]. 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.

Let ∇ be the backward difference operator defined by

$$\nabla Z_t \equiv Z_t - Z_{t-1}. \tag{9}$$

The difference operator ∇ is a high-pass filter. If we define the following sequence of high-pass filters

$$\Im_k \equiv \nabla^{k-1}, \quad k = 1, 2, 3, \ldots \tag{10}$$

with $\Im_1 \equiv \nabla^0$ being the identity filter, we can estimate the corresponding HOC, namely simple HOC [25], by

$$D_k = \mathrm{NZC}\{\Im_k(Z_t)\}, \quad k = 1, 2, 3, \ldots; \quad t = 1, \ldots, N \tag{11}$$

where NZC$\{\cdot\}$ denotes the estimation of the number of zero-crossings and

$$\nabla^{k-1} Z_t = \sum_{j=1}^{k} \binom{k-1}{j-1} (-1)^{j-1} Z_{t-j+1}$$

with

$$\binom{k-1}{j-1} = \frac{(k-1)!}{(j-1)!\,(k-j)!} \tag{12}$$

In practice, we only have finite time series and lose an observation with each difference. Hence, to avoid this effect we must index the data by moving to the right, i.e., for the evaluation of k simple HOC, the index $t = 1$ should be given to the kth or a later observation. For the estimation of the number of zero-crossings in (11), a binary time series $X_t(k)$ is initially constructed given by

$$X_t(k) = \begin{cases} 1, & \text{if } \Im_{\mathrm{k}}(Z_t) \geq 0 \\ 0, & \text{if } \Im_{\mathrm{k}}(Z_t) < 0 \end{cases}, \quad k = 1, 2, 3, \ldots; \quad t = 1, \ldots, N \tag{13}$$

and the desired simple HOC are then estimated by counting symbol changes in $X_1(k), \ldots, X_N(k)$, i.e.,

$$D_k = \sum_{t=2}^{N} [X_t(k) - X_{t-1}(k)]^2. \tag{14}$$

In finite data records, it holds $D_{k+1} \geq D_k - 1$ [25]. In addition, as k increases, the discrimination power of simple HOC diminishes, since different processes yield almost the same D_k [25].

In this paper, HOC are used to construct the feature vector ($\mathrm{FV_{HOC}}$), formed as follows:

$$\mathrm{FV_{HOC}} = [D_1, D_2, \ldots, D_L], \quad 1 < L \leq J \tag{15}$$

where J denotes the maximum order of the estimated HOC and L the HOC order up to they were used to form the $\mathrm{FV_{HOC}}$ (see following section for its selection). It must be noted that the use of the HOC term throughout the subsequent sections, actually refers to simple HOC. For an extended coverage of the HOC-based analysis of time series the reader is encouraged to consult Kedem's book [25].

D. Classification Processes and Setup

Four classification processes were employed in this paper. These are briefly described in the succeeding subsections followed by a description of the classification setup adopted.

1) Quadratic Discriminant Analysis (QDA): QDA [35], [36] is based on the quadratic discriminant function

$$d_m(\mathrm{FV}) = -\frac{1}{2}\log|C_m| - \frac{1}{2}(\mathrm{FV} - \mu_m)^T C_m^{-1}(\mathrm{FV} - \mu_m) + \log p_m \tag{16}$$

where $m = 1,\ldots,M$ 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 dataset, the covariance matrices C_m, $m = 1, \ldots, M$ and the mean vectors μ_m, $m = 1, \ldots, M$ are calculated from the training part whereas each FV from the test set is classified into one of the M classes according to the following rule:

$$g(\mathrm{FV}) = \arg\max_m d_m(\mathrm{FV}) \tag{17}$$

where $g(\mathrm{FV})$ is the class the FV was assigned. The resulting decision boundaries of the quadratic discriminant analysis are quadratic equations in FV.

2) k-Nearest Neighbor (k-NN): k-NN algorithm [37] assumes all instances (FVs) correspond to points in the n-D space. The nearest neighbors of a given FV are defined in terms of the standard Euclidean distance. More precisely, let an arbitrary FV be

$$\mathrm{FV} = [x_1, x_2, \ldots, x_n]. \tag{18}$$

Then, the distance between two FVs, $\mathrm{FV}_1 = [x_1^1, x_2^1, \ldots, x_n^1]$ and $\mathrm{FV}_2 = [x_1^2, x_2^2, \ldots, x_n^2]$ is defined as $d(\mathrm{FV}_1, \mathrm{FV}_2)$, where

$$d(\mathrm{FV}_1, \mathrm{FV}_2) \equiv \sqrt{\sum_{i=1}^{n} (x_i^1 - x_i^2)^2}. \tag{19}$$

The k-NN algorithm is realized in two steps: 1) define the training set of FVs; 2) given a query FV_q to be classified, let $\mathrm{FV}_1, \ldots, \mathrm{FV}_k$ denote the k FVs from training sets that are nearest to FV_q; then, return the class that the majority of the k-NN (FVs) are from.

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 (n-D space) [38]. There is an MD for each case and each case is classified as belonging to the group for which the MD is minimum. The statistical distance or MD between two FVs, FV_1 and FV_2, as defined previously, in the n-D space, from the same distribution, which has a covariance matrix C is defined as

$$d_S(\mathrm{FV}_1, \mathrm{FV}_2) = \sqrt{(\mathrm{FV}_1 - \mathrm{FV}_2)^T C^{-1} (\mathrm{FV}_1 - \mathrm{FV}_2)}. \tag{20}$$

Obviously, the MD is the same as the Euclidean distance if the covariance matrix is the identity matrix.

4) Support Vector Machine: In the SVM classifier [19], a polynomial function is used as a kernel function that projects the data to high dimensional feature space

$$K(\mathrm{FV_{sv}}, \mathrm{FV}_q) = (\mathrm{FV_{sv}^T FV}_q + 1)^p \tag{21}$$

where $\mathrm{FV_{sv}}$ is the "support vector" and FV_q is the query FV. Originally the SVM is designed for the classification of two classes by finding the optimal hyperplane where the expected classification error of test samples is minimized. Among several approaches to apply a multiclass SVM classification process the one-versus-all method [39] was adopted here. Six SVMs that correspond to each of the six universal emotions were used. The ith SVM is trained with all of the training data in the ith class with positive labels, and all

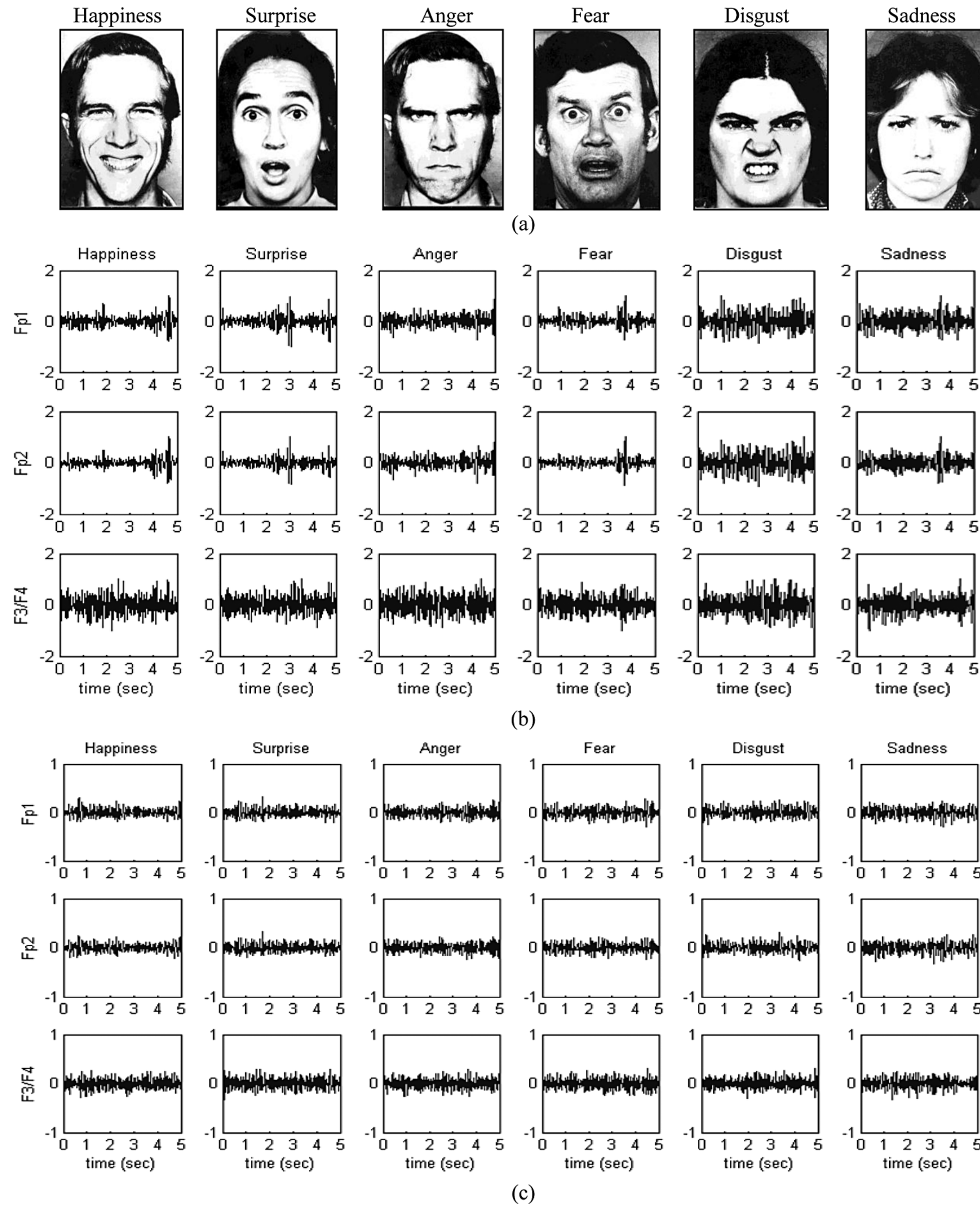

Fig. 1. (a) Visual stimulation drawn from the pictures of facial affect [31]. (b) EEG signals acquired from subject one corresponding to each emotion. (c) Averaged EEG signals corresponding to each emotion used for the HOC-EC analysis.

other training data with negative labels. During the emotion recognition procedure, the feature vector is simultaneously fed into all SVMs and the SVM that outputs a positive label is chosen and the class of the SVM indicates the recognition result.

5) Classification Setup: In order to test the classification efficiency of the HOC-based feature vector in an exhaustive way, we formed an iterative classification setup. The leave-n-out cross-validation technique [40] was adopted and the classification process was executed for 100 iterations. The classification results from each iteration were consequently averaged out resulting in the final mean classification rate, i.e., $\overline{C_r}$ (%). In all cases, 75% of the initial dataset was used to form the training set whereas the rest 25% was used as the test set. The training and test sets were drawn from different subjects, securing, thus, the independence between both sets.

III. Implementation Issues

A. Dataset Construction

In the line of this paper, a special experiment was designed in order to collect the EEG signals necessary for the HOC-EC analysis. In this experiment, 16 healthy volunteers participated; all were right-handed subjects (9 males and 7 females) in the age group of 19–32 years. Since they were aware of the purpose of the experiment, they tried to concentrate as much as possible during its realization.

At first, a suitable interface had to be implemented for the automated projection of the emotion-related pictures and the self-assessment of the emotion picturized each time. Pictures of facial affect by Ekman and Friesen [31], showing people expressing the six basic emotions [see Fig. 1(a)], 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 s after a 5-s black screen period, a 5-second period in which countdown frames were demonstrated and an 1-s projection of a cross shape in the middle of the screen to attract the sight of the subject. The same 16-s procedure was repeated for every 1 of the 60 pictures. The subjects were advised to stay as still as possible (to comprehend the emotion rather to mimic the facial expression) and not to blink during the image projection phase. This could assist the acquisition of noiseless EEG signals, as far as transient signals (spikes due to eyes blinking and possible facial muscle movements) are concerned. 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 naught emotional content. After the completion of the projection phase, the subject had to proceed with the self-assessment phase, where s/he could see the pictures previously projected and mark them according to the emotion each picture present according to her/his opinion. Apart from the six categories of emotions, s/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 I.

The EEG signals from each subject were recorded during the projection phase [see Fig. 1(b)]. In order to reduce the number of the EEG channels as much as possible and implement, an emotion recognition method that would result in a more user-friendly environment in the future, the signals were acquired from four positions only, according to the 10–20 system [41]. These positions were Fp1, Fp2, F3, and F4; Fp1 and Fp2 positions were recorded as monopole channels (channels 1 and 2, respectively), whereas the F3 and F4 positions as a dipole (channel 3), resulting in a three-EEG channel set (see Fig. 2). The ground was placed in the right earlobe. Psycho-physiological research has shown that left frontal lobe and right frontal lobe appear to have certain activity during the experience of negative or positive emotions. Particularly, left frontal lobe exhibit a more intense activity while the subject experiences a positive affect whereas right frontal lobe shows the same activity during a negatively originated emotion [42]. Those studies along with the evidenced role of prefrontal cortex in affective reactions and particularly in emotion regulation and conscious experience [43] justify the selection of the Fp1, Fp2, F3, and F4 positions to collect the signals intended to be processed.

After the acquisition part, the signals were filtered and cut in order to isolate the recordings during the picture projection period only. 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. In particular, an increased alpha band activity joined with a decreased one in the beta band indicates a cortical inactivation whereas the opposite indicates an active prefrontal cortex [42]. Finally, after the filtering, the EEG signals were segmented into 5 s segments corresponding to the duration of each picture projection and the set of the signals for each emotion, that is 10 signals/emotion, were averaged [see Fig. 1(c)] for noise

TABLE I
SELF-ASSESSMENT CLASSIFICATION ACCURACY (%) OF THE SIX BASIC EMOTIONS PER SUBJECT

Subjects	*Happiness*	*Surprise*	*Anger*	*Fear*	*Disgust*	*Sadness*
1	100	80	80	80	80	70
2	100	100	100	90	90	100
3	100	100	100	100	90	90
4	100	100	100	90	70	100
5	100	100	100	90	90	50
6	100	100	100	70	90	90
7	90	80	100	70	80	60
8	100	100	100	90	70	90
9	100	100	90	50	80	70
10	100	70	70	50	50	60
11	100	100	100	100	100	90
12	100	90	100	90	70	100
13	100	100	100	80	50	80
14	100	100	90	60	60	60
15	100	90	100	70	60	70
16	100	90	100	20	90	80
Mean	99.375	93.750	95.625	75	76.250	78.750

Mean: Mean classification accuracy (%) of each emotion.

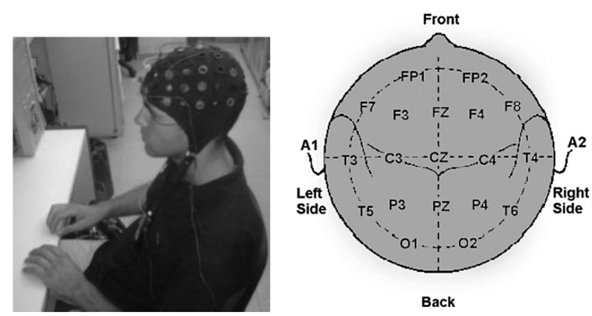

Fig. 2. Fp1, Fp2, F3, and F4 positions used for the EEG acquisition according to the 10–20 system [41].

elimination (e.g., artifact effect). Consequently, a new set of 16 EEG signals for each emotion lasting 5 s each was constructed in order to evaluate the HOC-EC on a user-independent basis.

B. Realization

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

The proposed analysis was realized on a PC (Intel Core 2 CPU 8400@3GHz) using MATLAB R2008 a (The Mathworks, Inc., Natick, MA). The filter used to isolate the alpha and beta bands was a tenth 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 $\mathrm{FV_{HOC}}$ [see (15)] was extracted from the whole 5 s signal within a range of order $K = 1, \ldots, 50 (= J)$ 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 channel, for each combination of channels and for each classification method (see subsequent section for the justification of the L selection). For a

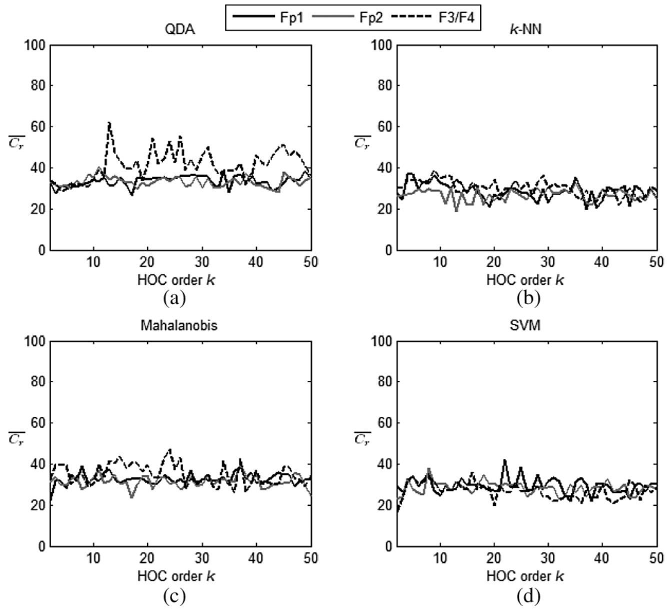

Fig. 3. $\overline{C_r}$ values for the single-channel case for all classification methods versus the k order of the HOC analysis.

spherical evaluation of the recognition power of the HOC-EC, 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 and for all classifiers. In all cases, averaged EEG signals from 12 and 4 subjects were used as training and test sets, respectively. In QDA, equal prior probabilities [p_m in (16)] 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 (21) was set as $p = 5$.

IV. Results and Discussion

A. Single-Channel Case

For the single-channel case, the EEG data from each channel were subjected to the HOC-based analysis and for each HOC order k, the $\overline{C_r}$ was estimated using the four aforementioned classifiers, i.e., QDA, k-NN, MD, and SVM, in order to identify the L value that provides with the maximum classification efficiency and defines the $\mathrm{FV_{HOC}}$, i.e., $\mathrm{FV_{HOC}} = [D_1, D_2, \ldots, D_L]$. In fact, this L value was selected as the one that leads to the highest $\overline{C_r}$ for the class-set of six distinct emotions. Fig. 3 depicts the $\overline{C_r}$ for all classifiers (see Fig. 3(a)–(d) correspond to QDA, k-NN, MD, and SVM, respectively) for $k = 2, \ldots, 50$ per channel. As it is clear from each subplot, for all classifiers but SVM L is defined from channel 3 (F3/F4), as there, the $\overline{C_r}$ gets its highest value. For the SVM case, L is derived from channel 1 (Fp1); the selected L values are [13 (QDA), 9 (k-NN), 24 (MD), 22 (SVM)] corresponding to $\overline{C_r}$ values of {62.3% (QDA), 38.67% (k-NN), 46.94% (MD), 41.94% (SVM)}. From an overall perspective of Fig. 3, it is noticed that, according to the classifier used, there is a severe or moderate dependence of the classification accuracy with the selection of the HOC order k. For example, channel 3 in the QDA case exhibits a noticeable variation in $\overline{C_r}$ for $k > 12$; on the contrary, this is not the case for rest classifiers. In a similar manner, channel 1 exhibits a moderate variation in the SVM case for $k > 21$, although in the case of the rest classifiers it shows a rather constant behavior. The constructed $\mathrm{FV_{HOC}}$ based on the selection of the L value per classifier was used throughout the channels, ensuring the highest classification efficiency and the minimum size of the $\mathrm{FV_{HOC}}$.

Table II presents the classification results from the HOC-EC for the hardest classification problem, i.e., when the class-set consists of six distinct emotions, derived from the HOC-based analysis of the EEG data from channel 3. The distribution of the classification results across the six emotions tabulated in Table II 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 $\mathrm{FV_S}$ and $\mathrm{FV_W}$ (corresponding to S-EC and W-EC approaches, respectively) are also included in this table, presented in the format of (%/%), i.e., (% from S-EC/% from W-EC), respectively. The percentages correspond to the $\overline{C_r}$ derived when averaging the classification rates across the 100 iterations. As Table II shows, in all cases, the HOC-EC outperforms the other two methods, exhibiting a $\overline{C_r}$ of 62.3%, whereas S-EC and W-EC provide 30.89% and 37.3% $\overline{C_r}$ values, respectively. From the six basic emotions, anger seems to be the most difficult to distinguish, as all three methods have exhibited the lowest mean classification rates (see Table II, fourth column-fourth row). Nevertheless, the HOC-EC has provided almost doubled $\overline{C_r}$, compared with the other two methods (50% versus 27.08/25%). The classification results presented in Table II are derived using the QDA classifier, as this was the one that provided with the highest classification rate for the proposed HOC-EC scheme amongst the four examined classification processes.

As presented in Table I, the $\overline{C_r}$ performed by the subjects themselves has revealed a difficulty in distinguishing fear, disgust and sadness, exhibiting a $\overline{C_r}$ less than 79%. In the other three emotions, i.e., happiness, surprise and anger, the $\overline{C_r}$ was found greater than 93%. Comparing these ‘‘human” classification results with the ones presented in Table II, it is clear that all three algorithms come short of the human ability to classify the basic six emotions. HOC-EC, however, is the only one that comes closer to the human $\overline{C_r}$ values, with the other two methods significantly lacking behind, achieving almost the same $\overline{C_r}$ for sadness (78.57% from HOC-EC versus 78.75% from the subjects).

From Table II, it is also clear that the HOC-EC presents the lowest misclassification distribution, as it exhibits, in most of cases, very low or zero mean misclassification rates. The W-EC approach follows, leaving behind the S-EC one, which shows a tendency to assign nonzero mean misclassification rates for all emotion pairs.

For a spherical presentation of the HOC-EC ability to distinguish between distinct emotions, the maximum $\overline{C_r}$, i.e., $\overline{C_r^{\max}}$, derived from all class-sets consisting of emotion groups (i.e., groups of six, five, four, three, and two emotions) for all three

TABLE II

$\overline{C_r}$ VALUES (%) OF EACH EMOTION OF THE PROPOSED HOC-EC COMPARED WITH S-EC AND W-EC USING THE QDA CLASSIFIER FOR THE SINGLE-CHANNEL CASE (CHANNEL 3)

Emotions	*Happiness*	*Surprise*	*Anger*	*Fear*	*Disgust*	*Sadness*
Happiness	**54.17 (35/46.43)**	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 (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	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	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 (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	10.72 (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 $\overline{C_r}$ values are derived from a 100-iteration cross validation process [40]. The formate (%%) corresponds to the $\overline{C_r}$ values derived from S-EC and W-EC, respectively, i.e., (S-EC/W-EC).

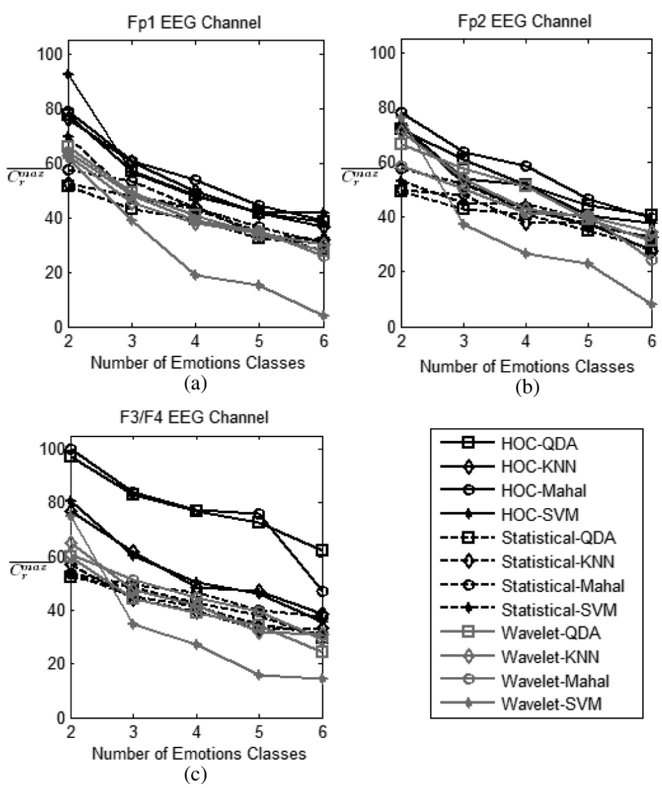

Fig. 4. $\overline{C_r^{\max}}$ values for the proposed method (HOC-EC) against the S-EC and W-EC, for the single-channel case for all classification methods and all possible combinations of emotion classes.

single channels and for all classifiers was estimated; Fig. 4 depicts the derived $\overline{C_r^{\max}}$ values. In particular, Fig. 4(a)–(c) correspond to the derived results for channels Fp1, Fp2, and F3/F4, respectively. In all parts of Fig. 4, the resulted classification rates from HOC-EC, S-EC, and W-EC are plotted with solid black line, dashed black line, and solid gray line, respectively; the four classifiers are denoted with square (□)-QDA, diamond (◊)-k-NN, circle (o)-MD, and asterisk ($*$)-SVM. As it is clear from this figure, channel 3 [F3/F4-Fig. 4(c)] provides the highest $\overline{C_r^{\max}}$ (almost in all emotion classes) compared to the other two channels [Fig. 4(a) and (b)]. Moreover, HOC-EC outperforms the other two feature extraction methods in all cases of channels exhibiting higher $\overline{C_r^{\max}}$ values for {six, five, four, three, two} emotion groups compared with the ones from the S-EC and W-EC. As Fig. 4 demonstrates, this holds almost independently from the classifier type; an exception occurs for the case of channel 2 [see Fig. 4(b)], where W-EC provides same or slightly higher $\overline{C_r^{\max}}$ values than the HOC-EC in all cases of the classifier types except the case of the MD classifier. It is noteworthy that the SVM classifier performs poorly for the case of W-EC when the number of emotions class-set is increased (>2), independently from the channel; nevertheless, this performance is inversed for the class-set with two emotions [especially for channels 2 and 3, respectively, see Fig. 4(b) and (c)]. This figure additionally shows that the HOC-EC steadily converges to the human's classification rate when the number of emotions class-set is of reduced size; on the contrary, the W-EC and S-EC seem to exhibit slower convergence for all classifiers except SVM, reflecting a lower response to the reduction of the emotion class-set from six to two emotions.

B. Combined-Channel Case

The procedure followed in the single-channel case was also adopted for the combined-channel scenario. In particular, four channel combinations were produced, i.e., CB1 = {channel 1 (Fp1), channel 2 (Fp2)}, CB2 = {channel 1 (Fp1), channel 3 (F3/F4)}, CB3 = {channel 2 (Fp2), channel 3 (F3/F4)}, and CB4 = {channel 1 (Fp1), channel 2 (Fp2), channel 3 (F3/F4)}; for each one, EEG data from the combined channels were subjected to the HOC-based analysis and for each HOC order k, the mean classification rate, denoted as $\overline{C_r^c}$, was estimated using again the QDA, k-NN, MD, and SVM classifiers. This way, the L value that provides with the maximum classification efficiency and defines the combined feature vector, $\mathrm{FV}^c_{\mathrm{HOC}}$, i.e., for each classifier was identified, accordingly. The $\mathrm{FV}^c_{\mathrm{HOC}}$ was structured as

$$\begin{aligned}&\mathrm{FV}^c_{\mathrm{HOC}}\\&= [D_1^i, D_2^i, \ldots, D_L^i, D_1^j, D_2^j, \ldots, D_L^j, D_1^s, D_2^s, \ldots, D_L^s],\\&\quad i = 1,2; \quad j = 2,3; \quad s = 0,3, \quad i \neq j, \quad j \neq s \qquad (22)\end{aligned}$$

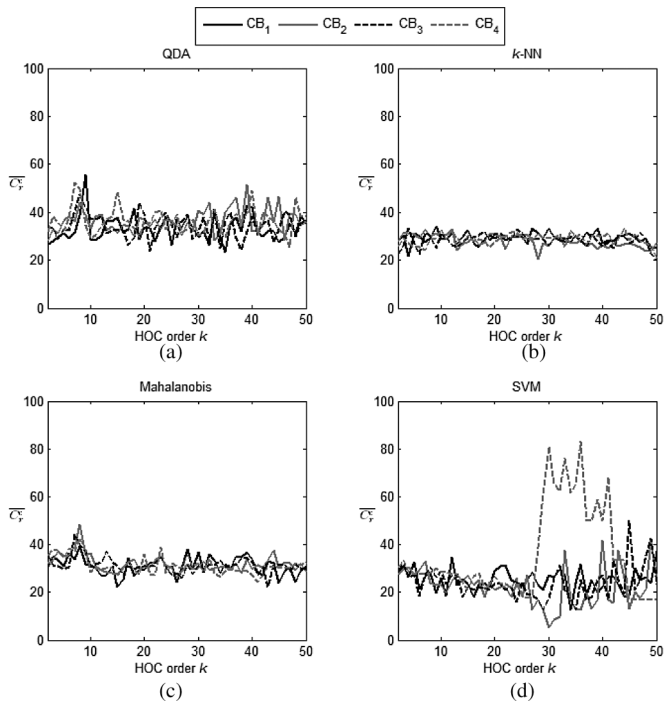

Fig. 5. $\overline{C_r^c}$ values for the combined-channel case for all classification methods versus the k order of the HOC analysis.

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 L value was selected as the one that leads to the highest $\overline{C_r^c}$ for the class-set of six distinct emotions. The $\mathrm{FV_S}$ and $\mathrm{FV_W}$ for the combined-channel case were constructed as

$$\mathrm{FV}_\mathrm{S}^c = [\mathrm{FV}_\mathrm{S}^i, \mathrm{FV}_\mathrm{S}^j, \mathrm{FV}_\mathrm{S}^s], \quad i = 1, 2; \quad j = 2, 3; \quad s = 0, 3, \quad i \neq j, \quad j \neq s \tag{23}$$

$$\mathrm{FV}_\mathrm{W}^c = [\mathrm{FV}_\mathrm{W}^i, \mathrm{FV}_\mathrm{W}^j, \mathrm{FV}_\mathrm{W}^s], \quad i = 1, 2; \quad j = 2, 3; \quad s = 0, 3, \quad i \neq j, \quad j \neq s \tag{24}$$

where i, j, and s denote the channel number that participates to the combination; $\mathrm{FV}_{\mathrm{S,W}}^0$ corresponds to a null vector element.

Fig. 5 depicts the $\overline{C_r^c}$ for all classifiers (Fig. 5(a)–(d) correspond to QDA, k-NN, MD, and SVM, respectively) for $k = 2, \ldots, 50$ per channel combination. Based on the results of Fig. 5, the selected L values were {9 (QDA/CB1), 9 (k-NN/CB1), 8 (MD/CB2), 36 (SVM/ CB4)} corresponding to $\overline{C_r^c}$ values of {55.48% (QDA), 33.98% (k-NN), 48.21% (MD), 83.33% (SVM)}. From an overall perspective of Fig. 5, there is not any noticeable dependence of the classification accuracy with the selection of the HOC order k and classifier type, except for the case of SVM/CB4 [see Fig. 5(d)]. In the latter, when $k \in [28, 42]$ the $\overline{C_r^c}$ exhibits significant increase, reaching its highest value (83.33%) at $k = 36$. Similar to the $\mathrm{FV_{HOC}}$ case, the constructed $\mathrm{FV}_\mathrm{HOC}^C$ based on the selection of the L value per classifier was used throughout the channel combinations, ensuring the highest classification efficiency and the minimum size of the $\mathrm{FV}_\mathrm{HOC}^C$.

Table III presents the $\overline{C_r^c}$ classification results from the HOC-EC for the hardest classification problem, i.e., when the class-set consists of six distinct emotions, derived from the HOC-based analysis of the EEG data from CB4 using the SVM classifier. The structure of Table III is the same as in the case of Table II, including the classification results when using S-EC and W-EC. The percentages correspond to the $\overline{C_r^c}$ derived when averaging the classification rates across the 100 iterations. As Table III shows, in all cases, the HOC-EC outperforms the other two methods, exhibiting a $\overline{C_r^c}$ of 83.33%, whereas the S-EC and the W-EC provide 23.61% and 27.92% $\overline{C_r^c}$ values, respectively. From the six basic emotions, HOC-EC faces difficulty to distinguish surprise from fear, exhibiting the lowest $\overline{C_r^c}$ (25%, see Table III, second column-second row). Nevertheless, in all other cases, the HOC-EC has provided almost four times higher $\overline{C_r^c}$ values compared with the other two examined methods.

Comparing the $\overline{C_r}$ performed by the subjects themselves (see Table I) and the $\overline{C_r}$ derived from (channel 3/QDA, see Table II) with $\overline{C_r^c}$ derived from (CB4/SVM, see Table III), it is clear that HOC-EC applied to the EEG data from the combination of all channels using SVM as a classifier approaches the human ability to classify the basic six emotions and, even, in most cases surpasses it (greater for happiness, anger, disgust, sadness; same for fear; significantly smaller for surprise).

As in the single-channel case (see Table II), Table III shows that the HOC-EC presents the lowest misclassification distribution, as it exhibits, in most of cases, very low or zero mean misclassification rates (except for the case of surprise). The W-EC follows, leaving behind the S-EC, which tends to assign nonzero mean misclassification rates for all emotion pairs.

Fig. 6 follows the same concept of Fig. 4, depicting the maximum $\overline{C_r^c}$, i.e., $\overline{C_r^{c,\max}}$, derived from all class-sets consisting of emotion groups for all three combined-channels and for all classifiers. In particular, Fig. 6(a)–(d) correspond to the derived results for CB1, CB2, CB3, and CB4, respectively. In all parts in Fig. 6, the resulted classification rates from HOC-EC, S-EC, and W-EC are plotted with solid black line, dashed black line, and solid gray line, respectively; the four classifiers are denoted with square (□)-QDA, diamond (◊)-k-NN, circle (o)-MD, and asterisk (∗)-SVM.

As it is clear from this figure, CB4 [see Fig. 6(d)] provides the highest $\overline{C_r^{c,\max}}$ in all emotion classes compared with the other three channel-combinations [see Fig. 6(a)–(c)]. Moreover, HOC-EC outperforms the other two feature methods in all cases of channel combinations, exhibiting higher $\overline{C_r^{\max}}$ values for {six, five, four, three, two} emotion groups compared with the ones from FV_S^c and FV_W^c. As Fig. 6 demonstrates, this holds for all classifier types except k-NN, which results in HOC-EC-based $\overline{C_r^{c,\max}}$ values smaller than the ones from the FV_S^c. The poor performance of the SVM classifier seen in the single-channel case of the $\mathrm{FV_W}$ (see Fig. 4), when the number of emotions class-set was increased (>2), has been replaced by a more enhanced one, probably due to the increased size of FV_W^c compared with $\mathrm{FV_W}$. This effect of the increased FV size is more evident in the exceptional performance of HOC-EC for the case of (SVM/CB4) [see Fig. 6(d)], where using an $\mathrm{FV}_\mathrm{HOC}^C$ of 108 features $(3 \times 36 (= L))$ it results in

TABLE III
$\overline{C_r^c}$ VALUES (%) OF EACH EMOTION OF THE PROPOSED HOC-EC COMPARED WITH S-EC AND W-EC USING THE SVM CLASSIFIER FOR THE COMBINED-CHANNEL CASE (CHANNEL COMBINATION CB_4)

Emotions	*Happiness*	*Surprise*	*Anger*	*Fear*	*Disgust*	*Sadness*
Happiness	**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 (50/10)	**25** **(25/30)**	0 (12.5/10)	75 (0/35)	0 (12.5/10)	0 (0/5)
Anger	0 (25/25)	0 (25/20)	**100** **(25/25)**	0 (0/0)	0 (25/15)	0 (0/15)
Fear	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 (0/25)	0 (12.5/0)	0 (25/0)	0 (12.5/25)	**100** **(25/25)**	0 (25/25)
Sadness	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 $\overline{C_r^c}$ values are derived from a 100-iteration cross validation process [40]. The formate (%%) corresponds to the $\overline{C_r^c}$ values derived from S-EC and W-EC, respectively, i.e., (S-EC/W-EC).

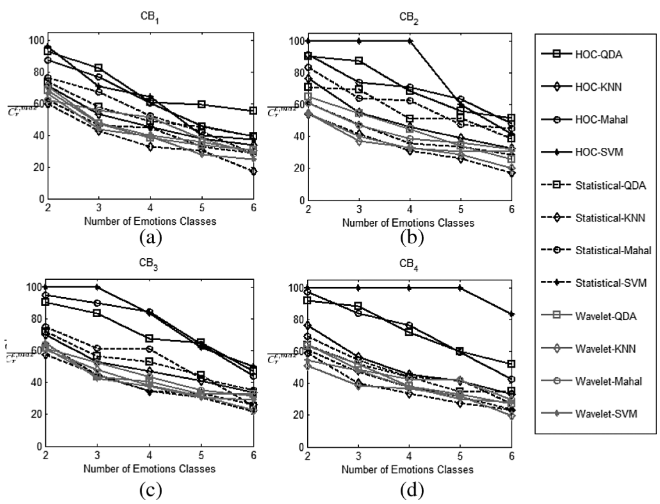

Fig. 6. $\overline{C_r^{c,\max}}$ values for the proposed method (HOC-EC) against the S-EC and W-EC, for the combined-channel case for all classification methods and all possible combinations of emotion classes.

$\overline{C_r^{c,\max}} = \{83.33\%, 100\%, 100\%, 100\%, 100\%\}$ for {six, five, four, three, two} emotion groups, respectively. Nevertheless, the increase in FV size should be handled with care, as it severely affects the computational burden of the classification process, especially when the issue of real-time implementation is of primary priority.

C. Overall Performance

For an overall performance perspective of the HOC-EC, Table IV was constructed. In particular, Table IV presents the best $\overline{C_r^{\max}}$ and $\overline{C_r^{c,\max}}$ values for the introduced HOC-EC and the two other examined methods, i.e., S-EC and W-EC, for the single- and combined-channel cases, respectively, when examining the hardest case of differentiating among the six emotions. The corresponding channel number and channel combination along with the relevant classifier types are also given. As Table IV tabulates, HOC-EC almost doubles the best $\overline{C_r^{\max}}$ and $\overline{C_r^{c,\max}}$ values from the S-EC and W-EC approaches, showing a significant enhancement in the EEG-based emotion detection process. Furthermore, Table IV shows that the use of more than one channels does not always imply increase in the classification performance; for example, when using three instead of one channels, the performance of HOC-EC is significantly increased (∼20%). Nevertheless, this increase is smaller (∼6%) for the case of S-EC approach when using two instead of one channels, and even a decrease is noticed (∼2%), as it happens for the case of the W-EC. The classification power of the examined classifiers seems to be affected by the selection of the FV (type and number of features). In fact, SVM shows high classification ability for extended-size FV (e.g., $\mathrm{FV}_{\mathrm{HOC}}^{C}$), whereas QDA seems to better handle the classification process for small-size FVs (e.g., $\mathrm{FV}_{\mathrm{HOC}}$, $\mathrm{FV}_{\mathrm{W}}^{c}$). Finally, MD exhibits the highest classification performance both for the case of FV_s and $\mathrm{FV}_{\mathrm{S}}^{c}$, respectively. The fact that the SVM provides with the more efficient classification is beneficial, as SVMs give a magnitude for recognition; using this magnitude a degree could be possibly associated with the emotion stimulus [14].

TABLE IV
BEST $\overline{C_r^{\max}}$ AND $\overline{C_r^{c,\max}}$ VALUES (%) FOR THE CASE OF SIX EMOTION CLASSES OF THE PROPOSED HOC-EC COMPARED WITH S-EC AND W-EC

Method	Analyzed Case	
	Single-Channel Case $\overline{C_r^{max}}$	Combined-Channels Case $\overline{C_r^{c,max}}$
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)

Corresponding classifiers and channels (or channel combinations) are shown in parentheses.

The presented results have shown the ability of the HOC-EC to better discriminate the EEG potentials stemming from emotion stimulus. Apart from the EEG-based perspective, speech and facial expressions are widely examined for the extraction of the underlying emotion, as there is a great deal of mutual information between vocal and facial expressions. Nevertheless, the body of work on detecting emotion in speech is quite limited. Currently, researchers are still debating what features influence the recognition of emotion in speech. There is also considerable uncertainty as to the best algorithm for classifying emotion, and which emotions to class together. Some promising approaches are reported in [11]–[14]. In particular, in [13] the k-NN and SVM classifiers were used to classify pairs of opposing emotions; they report a mean classification rate of 73.94% when using both k-NN and SVM. Moreover, in [14], the k-Means, SVM, and neural network classifiers were used to classify the emotions neutral, anger, happiness, and sadness; they report a mean classification rate of 74.28% using SVM. In addition, in [12], facial expressions, acoustic information and combination of them are examined for their emotion discrimination efficiency among the four aforementioned emotions using SVM; they report a mean classification rate of 70.9% when using acoustic information only, 85.1% when using facial information only, and 89.1% when facial expressions and acoustic information were fused at feature-level. These classification rates are clearly lower than the ones derived using the proposed HOC-EC, which refer to harder case of differentiating among six emotions simultaneously. This also holds for the case of emotion detection using data related to the autonomous nervous system [15], [16]; for example, in [16], the reported mean classification performance is 67.81% (using k-NN) and 62.27% (using discriminant function analysis) for the differentiation among sadness, anger, surprise, fear, frustration, and amusement.

It should be stressed out that the HOC-EC is a user independent approach; hence, it has high degree of generalization facilitating its transfer to many HMI environments without requiring an arduous adaptation or pre-analysis phase. In fact, the proposed HOC-EC scheme, contributes, as a basic research, to the area of affective computing and HMI; applications, however, in the field of healthcare are visible, such as, identifying patient's emotion experiences, recurring affective states, emotional trends over time, and providing appropriate feedback dynamically adapted to the patient's emotional states via a variety of appropriate means depending upon the given context (e.g., disorders within the autism spectrum and/or cerebral palsy). Adaptation of the HOC-EC method to a fully pervasive system would require additional effort (both in EEG acquisition and real-time implementation aspects). Nevertheless, advent technology could provide in the forthcoming future smaller wireless EEG sensors and dedicated hardware, fostering the HOC-EC integration within expert pervasive systems that could collectively assess health indicators and patterns for diagnostic, predictive and prescriptive purposes.

V. Conclusion

A new classification tool, namely HOC-EC, has been presented here for EEG-based emotion recognition. In particular, a new feature vector extraction method based on HOC analysis has been proposed and has revealed the potential of a robust emotion recognition method from EEG. Moreover, a novel emotion elicitation technique based on the Mirror Neuron System was proposed in order to achieve more effective representation of the emotion stimulus to the artificially evoked brain potentials. EEG data from single-channel and combined-channel cases were subjected to HOC-EC and four different classifiers (QDA, k-NN, MD, and SVM) were employed. For the single-channel case, the best results were obtained by the QDA (62.3% mean classification rate), whereas for the combined-channel case, the best results were obtained using SVM (83.33% mean classification rate), for the hardest case of differentiating among the six basic emotions, i.e. happiness, surprise, anger, fear, disgust, and sadness. Extended classification tests were implemented concerning all possible combinations of emotion classes showing the effectiveness of the proposed method, which achieved gradually maximum (100%) classification rates as soon as the number of the classes emotions was reduced from six to two. The comparison of the HOC-EC method with other feature extraction methods (S-EC and W-EC, respectively) overwhelmed its robustness and consistency to effectively discriminate emotions from EEG signals. Integration of HOC-EC in a pervasive healthcare system could contribute to improved HMI that enables advanced monitoring and interpretation of patient status and environment optimizing the whole medical assessment procedure.

Acknowledgment

The authors would like to thank all the 16 subjects participated in the experiment for their patience during the tedious EEG recording phase. Moreover, the authors would also like to appreciate the comments of the two anonymous reviewers that contributed to a more in-depth expansion of the paper upon the examined issues.

References

[1] P. SaloVey and J. D. Mayer, "Emotional intelligence," *Imagination, Cognition Pers.*, vol. 9, no. 3, pp. 185–211, 1990.

[2] D. Goleman, *Emotional Intelligence*. New York: Bantam Books, 1995.

[3] A. R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Gosset/Putnam, 1994.

[4] B. Reeves and C. Nass, *The Media Equation. How People Treat Computers, Television, and New Media like Real People and Places*. New York: CSLI, Cambridge Univ. Press, 1996.

[5] C. Doukas and I. Maglogiannis, "Intelligent pervasive healthcare systems," *Adv. Comput. Intell. Paradigms Healthcare-3*, vol. 107, pp. 95–115, 2008.

[6] C. Lisetti, F. Nasoz, C. LeRouge, O. Ozyer, and K. Alvarez, "Developing multimodal intelligent affective interfaces for tele-home health care," *Int. J. Human-Comput. Stud.*, vol. 59, pp. 245–255, 2003.

[7] T. S. Barger, D. E. Brown, and M. Alwan, "Health-status monitoring through analysis of behavioral patterns," *IEEE Trans. Syst., Man Cybern.*, vol. 35, no. 1, pp. 22–27, Jan. 2005.

[8] I. Cohen, A. Garg, T. S. Huang. (2000)."Emotion recognition from facial expressions using multilevel HMM" presented at the Neural Inf. Process. Syst. (NIPS) Workshop Affective Comput., Colorado. [Online]. Available: http://www.ifp.uiuc.edu/∼ashutosh/ papers/NIPS_emotion.pdf

[9] F. Bourel, C. C. Chibelushi, and A. A. Low, "Robust facial expression recognition using a state-based model of spatially-localized facial dynamics," in *Proc. 5h IEEE Int. Conf. Automat. Face Gesture Recognit.*, 2002, pp. 106–111.

[10] J. J. Lien, T. Kanade, J. F. Cohn, and C. Li, "Automated facial expression recognition based on FACS action units," in *Proc. 3rd IEEE Conf. Automat. Face Gesture Recognit.*, 1998, pp. 390–395.

[11] B. Schuller, S. Reiter, R. Mueller, M. Al-Hames, and G. Rigoll, "Speaker independent speech emotion recognition by ensemble classification," in *Proc. 6th Int. Conf. Multimedia Expo. ICME 2005*, pp. 864–867.

[12] 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," in *Proc. 6th Int. Conf. Multimodal Interfaces (ICMI-04)*, 2004, pp. 205–211.

[13] R. Shah and M. Hewlett. (2007). Emotion detection from speech. [Online]. Available: http://www.stanford.edu/class/cs229/proj2007/ShahHewlett%20–20Emotion%20Detection%20 from%20Speech.pdf

[14] F. Yu, E. Chang, Y. Q. Xu, and H. Y. Shum, "Emotion detection from speech to enrich multimedia content," in *Proc. IEEE Pacific Rim Conf. Multimedia*, 2001, pp. 550–557.

[15] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[16] F. Nasoz, C. L. Lisetti, K. Alvarez, and N. Finkelstein. (2003, Jun. 22–26). "Emotion recognition from physiological signals for user modeling of affect," presented at the UM'2003, 9th Int. Conf. User Model,

Pittsburg, PA, [Online]. Available: http://www.eurecom. fr/util/ publidownload.en.htm?id = 1806.
[17] K. Takahashi, "Remarks on emotion recognition from bio-potential signals," in *Proc. of 2nd Int. Conf. Autonomous Robots Agents*, 2004, pp. 186–191.
[18] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry, "EEG feature extraction for classifying emotions using FCM and FKM," *J. Comput. Commun.*, vol. 1, pp. 21–25, 2007.
[19] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[20] J. J. De Gruijter and A. B. McBratney, "A modified fuzzy k means for predictive classification," in *Classification and Related Methods of Data Analysis*, H. H. Bock, Ed. Amsterdam, The Netherlands: Elsevier, 1988, pp. 97–104.
[21] K. G. Srinivasa, K. R. Venugopal, and L. M. Patnaik, "Feature extraction using fuzzy C-Means clustering for data mining systems," *Int. J. Comput. Sci. Netw. Security*, vol. 6, pp. 230–236, 2006.
[22] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry, "Lifting scheme for human emotion recognition using EEG," in *Proc. Int. Symp. Inf. Technol.*, 2008, pp. 1–7.
[23] D. O. Bos. (2006). EEG-based emotion recognition: The influence of visual and auditory stimuli. [Online]. Available: http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Oude_Bos-Danny.pdf
[24] R. Horlings, D. Datcu, and L. J. M. Rothkrantz, "Emotion recognition using brain activity," in *Proc. Int. Conf. Comput. Syst. Technol.*, 2008, pp. 1–6.
[25] B. Kedem, *Time Series Analysis by Higher Order Crossings*. Piscataway, NJ: IEEE Press, 1994.
[26] G. Rizzolatti and L. Craighero, "The mirror-neuron system," *Annu. Rev. Neurosci.*, vol. 27, pp. 169–192, 2004.
[27] R. W. Picard, *Affective Computing*. Boston, MA: MIT Press, 1997.
[28] P. Ekman *et al.*, "Universals and cultural differences in the judgments of facial expressions of emotion," *J. Pers. Soc. Psychol.*, vol. 53, pp. 712–717, 1987.
[29] C. Barr-Zisowitz, "Sadness-Is there such a thing," in *Handbook of Emotions*, 2nd ed, M. Lewis and J. M. Haviland-Jones, Eds. New York: The Guilford Press, 2000, ch. 38, pp. 607–622.
[30] E. Oztop, M. Kawato, and M. Arbib, "Mirror neurons and imitation: A computationally guided review," *Neural Netw.*, vol. 19, pp. 254–271, 2006.
[31] P. Ekman and W. V. Friesen, "Pictures of facial affect," in *Human Interaction Laboratory*. San Francisco, CA: Univ. California Medical Center, 1976.
[32] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (IAPS): Affective ratings of pictures and instruction manual," University of Florida, Gainesville, FL, Tech. Rep. A-8, 2008.
[33] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science Engineering*. Bristol, U.K.: Institute of Physics (IOP) Publishing Ltd, 2002.
[34] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.
[35] W. J. Krzanowski, *Principles of Multivariate Analysis*. Oxford, U.K.: Oxford Univ. Press, 1988.
[36] G. A. F. Seber, *Multivariate Observations*. Wiley Series in Probability and Statistics. New York: Wiley, 1984.
[37] T. Mitchell, *Machine Learning*. New York: McGraw-Hill, 1997.
[38] P. C. Mahalanobis, "On the generalized distance in statistics," *Proc. Nat. Inst. Sci. India*, vol. 2, pp. 49–55, 1936.
[39] O. Chapelle, P. Haffner, and V. Vapnik, "Support Vector Machines for histogram-based image classification," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 1055–1064, Sep. 1999.
[40] P. A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*. London, U.K.: Prentice-Hall, 1982.
[41] E. Niedermeyer and F. Lopes da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Baltimore, MD: Williams & Wilkins, 1993.
[42] C. P. Niemic, "Studies of emotion: A theoretical and empirical review of psychophysiological studies of emotion," *J. Undergraduate Res.*, vol. 1, pp. 15–18, 2002.
[43] R. J. Davidson, D. C. Jackson, and N. H. Kalin, "Emotion, plasticity, context, and regulation: Perspectives from affective neuroscience," *Psychol. Bull.*, vol. 126, pp. 890–909, 2000.

Panagiotis C. Petrantonakis (S'08) was born in Ierapetra, Crete, Greece, in 1984. He received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, in 2007. He is currently working toward the Ph.D. degree with the Signal Processing and Biomedical Technology Unit, Telecommunications Laboratory, Department of Electrical and Computer Engineering, AUTH.

His current research interests include advanced signal processing techniques, nonlinear transforms, and affective computing.

Mr. Petrantonakis is a member of the Technical Chamber of Greece.

Leontios J. Hadjileontiadis (S'87–M'98) was born in Kastoria, Greece, in 1966. He received the Diploma degree in electrical engineering and the Ph.D. degree in electrical and computer engineering both from Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, in 1989 and 1997, respectively. He also received the Ph.D. degree in music composition from the University of York, Heslington, U.K., in 2004.

In December 1999, he joined as a Faculty Member with the Signal Processing and Biomedical Technology Unit, Telecommunications Laboratory, Department of Electrical and Computer Engineering, AUTH, where he is currently an Associate Professor, engaged on lung sounds, heart sounds, bowel sounds, ECG data compression, seismic data analysis, and crack detection. He is also a Professor in music composition with the State Conservatory of Thessaloniki, Thessaloniki. His research interests include higher order statistics, alpha-stable distributions, higher order zero crossings, wavelets, polyspectra, fractals, neuro-fuzzy modeling for medical, mobile, and digital signal processing applications.

Dr. Hadjileontiadis is a member of the Technical Chamber of Greece, Higher Order Statistics Society, International Lung Sounds Association, and 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, IL, 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 at the 24th International Lung Sounds Conference'99, Marburg, Germany. In 2004, 2005, and 2007, he organized and served as a mentor to three five-student teams that have ranked as third, second, and seventh worldwide, respectively, at the Imagine Cup Competition (Microsoft), Sao Paulo, Brazil in 2004/Yokohama, Japan, in 2005/Seoul, Korea, in 2007, with projects involving technology-based solutions for people with disabilities.