ECG Pattern Analysis for Emotion Detection

Foteini Agrafioti, *Member*, *IEEE*, Dimitrios Hatzinakos, *Senior Member*, *IEEE*, and Adam K. Anderson

Abstract—Emotion modeling and recognition has drawn extensive attention from disciplines such as psychology, cognitive science, and, lately, engineering. Although a significant amount of research has been done on behavioral modalities, less explored characteristics include the physiological signals. This work brings to the table the ECG signal and presents a thorough analysis of its psychological properties. The fact that this signal has been established as a biometric characteristic calls for subject-dependent emotion recognizers that capture the instantaneous variability of the signal from its homeostatic baseline. A solution based on the empirical mode decomposition is proposed for the detection of dynamically evolving emotion patterns on ECG. Classification features are based on the instantaneous frequency (Hilbert-Huang transform) and the local oscillation within every mode. Two experimental setups are presented for the elicitation of active arousal and passive arousal/valence. The results support the expectations for subject specificity, as well as demonstrating the feasibility of determining valence out of the ECG morphology (up to 89 percent for 44 subjects). In addition, this work differentiates for the first time between active and passive arousal, and advocates that there are higher chances of ECG reactivity to emotion when the induction method is active for the subject.

Index Terms—Electrocardiogram, emotion recognition, affective computing, arousal, valence, active stress, passive stress, bivariate empirical mode decomposition, intrinsic mode function, instantaneous frequency, oscillation.

1 Introduction

Human emotions are psychophysiological experiences that affect all aspects of our daily lives. Emotions are complex processes comprised of numerous components, including feelings, bodily changes, cognitive reactions, behavior, and thoughts. Various models have been proposed by considering the ways in which these components interact to give rise to emotions, but at the moment there is no single formulation that is universally acceptable. Modeling emotions is a very challenging problem that has drawn a great deal of interest from the emerging field of human-computer interaction. The objective is to design systems that can automatically identify emotional states, which would revolutionize applications in medicine, entertainment, education, safety, etc. The main difficulty in formulating these models lies in the fact that we must rely on visible manifestations of emotions to produce and verify them since the latent factors that generate emotions are unobservable.

The first step in modeling any phenomenon is data collection. We need to design experiments and institute methodologies that successfully induce emotions in a laboratory setting wherein we can record and collect psychological data. In quantifying psychological activity we are limited to the study of visible manifestations like facial expressions, gestures, vocal traits, etc. These modalities are popular in HCI since they use the same cues that humans rely upon to detect and recognize emotional states. Moreover, most human beings display similar manifestations in response to identical emotional stimuli, which allows for objective emotion annotation.

A major drawback of using behavioral modalities for emotion detection is the uncertainty that arises in the case of individuals who either are consciously regulating their emotional manifestations or are naturally suppressive. For instance, although facial expressions can be analyzed to determine emotions, there is no guarantee that an individual will express the corresponding cue, irrespective of whether they are experiencing a certain emotion. This has serious implications in some applications such as surveillance.

An interesting alternative to the use of behavioral modalities is the physiological signals (or biosignals) which constitute vital signs of the human body. Examples of this category include the electrocardiogram (ECG), electromyogram (EMG), electroencephalogram (EEG), galvanic skin response (GSR), blood volume pressure (BVP), heart rate (HR) or heart rate variability (HRV), temperature (T), respiration rate (RR). These signals have traditionally been used for clinical diagnostics, but there is significant evidence to suggest that they are sensitive to and may convey information about emotional states [1], [2], [3], [4], [5], [6]. One of the benefits of detecting emotions using physiological signals is that these are involuntary reactions of the body, and as such are very difficult to mask. Moreover, for the duration of time that the sensors are attached to the body, these signals are recorded continuously, enabling frequent emotional assessment. This is not the case with vocal features, for example, which can only be captured while the individual is speaking.

There are, however, many theoretical and practical challenges with regard to biosignal-based emotion detection. First, while the evidence suggests that physiological

- *F. Agrafioti and D. Hatzinakos are with the Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON M5S 3G4, Canada.*
E-mail: {foteini, dimitris}@comm.utoronto.ca.
- *A.K. Anderson is with the Department of Psychology, University of Toronto, 100 St. George Street, Toronto, Ontario M5S 3G3, Canada.*
E-mail: anderson@psych.utoronto.ca.

Manuscript received 28 Dec. 2010; revised 8 July 2011; accepted 30 July 2011; published online 18 Aug. 2011.
Recommended for acceptance by J.-M. Fellous.
For information on obtaining reprints of this article, please send e-mail to: taffc@computer.org, and reference IEEECS Log Number TAFFC-2010-12-0125.
Digital Object Identifier no. 10.1109/T-AFFC.2011.28.

1949-3045/12/$31.00 © 2012 IEEE Published by the IEEE Computer Society

signals are affected by emotions, the exact effects on the waveform patterns remain to be seen. For example, the heart rate increases under both fear and excitement, but whether we can differentiate between the two is as yet unknown. Second, there are open questions about the subject-specific nature of these effects.

Apart from the open theoretical questions, there are practical issues as well. The experimental protocols are far more complex than in behavioral emotion research, where the collection is facilitated by instructing volunteers to exhibit emotions. For biosignal-based experimentation more sophisticated practices are necessary to elicit truthful emotions in a laboratory setting. Furthermore, labeling physiological signals is subjective and, as such, very risky due to the difficulty in establishing the ground truth.

Another practical challenge relates to signal acquisition. The collection process is more invasive when compared to that for behavioral modalities since the sensors need to be in contact with the human body for the duration of a recording session. For this reason, it is important to minimize the number of data required for this task, i.e., for detection to rely on as few signals as possible.

Different physiological signals originate from different locations of the human body and may describe unrelated functions. For instance, the ECG and BVP signals are of cardiovascular origin, while the EMG relates to muscle electrical potential. It is important to investigate the dynamics of every signal in order to clearly establish its limits in assessing psychological activity. For instance, we anticipate that the experimental setup alone may induce emotional reactivity for one biosignal, but may not stimulate some other.

To this end, this work investigates the feasibility and limits of emotion detection using ECG signals alone. Previous works have employed ECG (the HR mostly) in conjunction with other biosignals. There are several open questions, such as understanding the psychophysiological rationale responsible for the formation of the signal, the subject specificity, the statistical limits to valence and arousal differentiation, and the type of experiments that can stimulate it. This work addresses these questions and aims to present a strong case for emotion detection using ECG signals.

The rest of this paper is organized as follows: Section 2 provides a overview of the related research on biosignal-based emotion detection, with particular interest in the works that employed the ECG. The psychophysiological basis for emotion detection from ECG signals is discussed in Section 3, along with a brief introduction to the challenges of signal processing for cognitive tasks. The proposed methodology for ECG feature extraction is presented in Section 4. Sections 5 and 6 report the experimental setup for signal collection and the performance of the discussed framework, respectively. Finally, we conclude in Section 7 with a summary of our findings and the perspectives for future work.

2 Related Research

There are numerous approaches in the literature for biosignal-based emotion detection, broadly classified here into two categories according to the way the emotional models are conceived and deployed.

The first category includes the *discrete emotional models* (DEM), where the objective is to recognize and label standard emotional states (for instance joy, sadness, or fear). These methods heavily rely on the assumption that an arbitrary affect state is well defined and completely distinguishable from the rest.

On the other hand, *affective dimensional models* (ADM) relax the conditions for discrete emotions and treat any affective state as a combination of two parameters, namely, *arousal* and *valence*.[1] Arousal is a measure of emotional stimulation (intensity) and it varies between low and high. Valence is a measure of pleasantness for the experienced emotion, ranging from very pleasant (positive) to very unpleasant (negative). These two variables form a 2D space (the AV plane), and in ADM models classification is carried out among predefined areas of the plane.

From an engineering perspective, both approaches have advantages and disadvantages. For instance, DEM models might be easily conceptualized and fitted to applications; however, every emotion needs to be absolutely defined and appropriately targeted by the experimental setup, which might be risky for neighboring emotions (for example, pleasantness and calmness). The remainder of this section provides a brief overview of some of these methodologies, with the goal of demonstrating the way that ECG has been previously deployed.

2.1 Discrete Emotional Models

Among the earliest efforts for emotion differentiation using physiological signals is the work of Ekman et al. [2]. At the time, this work opposed the dogma dictating that no activity of the nervous system is emotion specific. A total of six emotions were studied using facial expressions and physiological signals such as the HR, left and right hand temperature, skin resistance, and forearm muscle tension. Emotion induction was based on reliving experiences, which is still considered to be one of the most successful approaches to emotion elicitation. The statistical analysis was based on the *change scores* principle, while decision trees were used to examine emotion clustering. Despite the lack of sophisticated statistical tools, this work was a landmark to the establishment of physiological reactivity to emotion.

From a more clinical perspective, Sinha et al. [7], looked into the possibility of differentiating emotions through cardiovascular measurements. Six heart-related indexes were recorded, namely, HR, BVP, stroke volume, cardiac output, peripheral vascular resistance, and an index of myocardial contractility. These signals were collected under four emotional states: fear, anger, joy, and sadness. Once more, change scores were calculated for every cardiovascular measure, to demonstrate that the HR is very sensitive to anger as opposed to joy and sadness. This study focused specifically on the cardiovascular system to observe that there might be limitations in the detection of less arousing states. More information regarding cardiac specificity to emotion can be found in Section 3.2.

1. According to the circumplex model of affect, an emotion is defined based on arousal and valence. There are numerous variations of this model, the study of which is beyond the scope of this work.

On the engineering side, Picard et al. [8] pioneered automatic detection by proposing an emotionally intelligent system that fuses information from four physiological signals (HR included). The feature space pictured the actual statistical properties of these signals, rather than the changing rate. The Sequential Floating Forward Search algorithm (SFFS) in conjunction with Fisher Projection (FP) was used for feature selection. An 81.25 percent detection rate was achieved for eight emotions. This work established the challenges of affective computing by emphasizing the fact that data collection for the purpose of emotion detection is a sensitive process compared to any standard experimental setup.

Music tracks were used by Kim and André [9] to induce pleasure, joy, anger, and sadness while four biosignals were recorded (ECG included). This work highlighted that subject-dependent detection is more accurate because people might be experiencing emotions in unique ways. Also, on the ground that arousal is better detectable than valence, an emotion-specific multilevel dichotomous classification algorithm was designed to gradually classify an input, using a number of intermediate binary classification steps. The overall recognition performance for four emotions was 95 percent (average for three subjects).

A more complete system was reported by Nasoz et al. [10], who looked into the problem of adapting the software interface according to a user's emotional state. A multimodal framework was evaluated based on the combination of facial, verbal, gesture, and physiological recordings (HR included). The Marquardt Backpropagation algorithm was used for classification, with a maximum performance of 83 percent.

Large-scale experiments with more than 100 subjects were reported by Kim et al. [11] for biosignal-based emotion recognition. With respect to cardiac reactivity, the authors suggested the spectrum of HRV for feature extraction. In a subject-independent mode, the combination of HRV and GSR achieved a performance of 61.8 percent for classification among four emotions.

2.2 Affective Dimensional Models

In the affective dimensional modeling field, Mandryk and Atkins [12] presented a two-stage emotion recognition system. In order to identify an emotion, classification was carried out among prequantized areas of the AV plane, while a fuzzy scheme was designed to establish the correspondence with the desired emotional states. The HR was measured among four other biosignals while the subjects were engaged with entertaining multimedia. The type of the stimulus that was used in this work is interesting with respect to cardiac reactivity because it constitutes an active stressor. The performance was measured as the *distance* between the detected arousal and valence values and the self-reported ones by the volunteers. The arousal error was within 3 percent, while valence was within 6 percent. On the other hand, a passive stimulus, the international Affective Picture System (IAPS) [13], was used by Haag et al. [14] to elicit emotions of varying arousal and valence. Standard deviation measurements were estimated for six biosignals (ECG included) and classified with a neural network. The overall classification rate was 96.58 percent for one subject.

Similarly, Jones and Troen [5] proposed a system for affective dimensional emotion recognition where the AV plane was divided into 25 equal areas, each defining an emotional region. The IAPS [13] picture set was used for emotion elicitation of 13 subjects. For classification, two multilayer perceptron neural networks were trained for arousal and valence, respectively. Performance was reported for various precision levels on the AV plane (class is correct if classified within a number of its neighbors). For a moderate classification rule, the system achieved a performance of 62 and 67 percent for the valence and arousal dimensions, respectively.

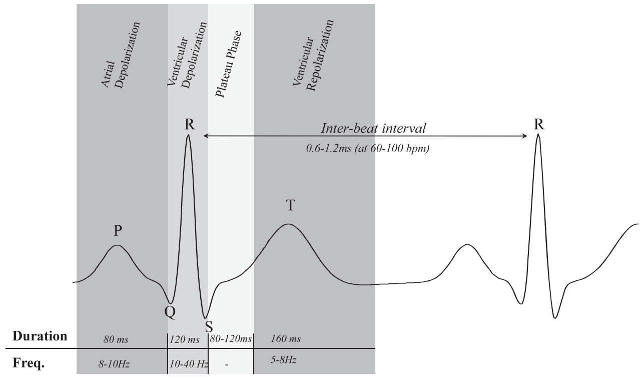

Fig. 1. Main components of an ECG heart beat. Each wave describes a phase of the cardiac cycle and has very distinct characteristics.

It is clear from the above works, that biosignal-based emotion detection has already been established. A variety of approaches have been reported depending upon the type of emotion elicitation, the targeted emotions, the specificity to subjects, and the envisioned applications. Despite the significance of the reported accuracy, prior systems rely on the fusion of a number of physiological measurements. This makes most applications impractical due to subject obtrusiveness. Furthermore, each of the previously examined biosignals has a unique psychophysiological origin, the study of which can deepen our understanding of emotion.

This work is interested in a cardiac signal, the ECG, which, despite the wealth of cardiac information that it carries, has been primarily studied on a heart rate basis only. The major novelty lies in the analysis of the dynamic properties of this signal's waveform using a discrete emotional model.

3 The Psychphysiological Rationale

3.1 The Physiology of ECG

The electrocardiogram (ECG) signal reflects the variation of the cardiac electrical potential over time. The ECG signal is recorded by attaching electrodes on the surface of the body using the standard *12 lead ECG system* [15].

The ECG signal is quasi periodic, constituted of heartbeat repetitions. Each beat is composed of three waves: the `P` wave, `QRS` complex, and the `T` wave (depicted in Fig. 1). These waves represent the sequential depolarization and repolarization of the cardiac muscle and they are usually referred to as *fiducial points*. Every healthy pulse is initiated by a number of pacemaker cells located at the sinoatrial (SA) node (upper left atrium). The electrical impulse travels through the left and right atria, causing them to depolarize. It is then collected and delayed by the atrioventricular (AV) node, which acts like a gateway to the ventricles. In essence, the ECG signal is a summary of the various phases of a cardiac cycle.

The P wave usually has positive polarity and a duration of approximately 80 ms [15]. The QRS duration varies between 80-120 ms. Finally, the T wave is usually observed about 300 ms after the QRS complex. However, its exact position depends on the heart rate, appearing closer to the QRS complex at rapid rhythms [15].

The signal-processing algorithms used to study ECG signals utilize its spectral characteristics. A healthy P wave is considered to contribute to the low-frequency components ($\approx< 10$ Hz). The QRS complex contributes to the higher frequency components (typically 10-40 Hz) in the spectrum due to its steep slopes.

3.2 Cardiovascular Reactivity to Emotion

The nerve-endings of the Autonomic Nervous System (ANS) within the cardiac muscle play a major role in the cardiac output because they affect the rhythm at which the muscle pumps blood. The fibers of the sympathetic system run along the atria and the ventricles and, when activated, stimulate the cardiac muscle to increase the heart rate. On the other hand, the parasympathetic system reduces the cardiac workload. Specifically, in the presence of a mental stressor, the sympathetic system dominates the parasympathetic, resulting in the following reactivity effects [16]:

1. *Automaticity*. The intrinsic impulse firing (automaticity) of the pacemaker cells increases, which translates directly to an increased heart rate.
2. *Contractility*. During every contraction the fibers of the heart shorten more compared to the case during homeostasis, thereby increasing the force of contraction.
3. *Conduction rate*. The natural pacemaker, the SA node, is forced to conduct faster.
4. *Excitability*. During sympathetic stimulation, the person has increased perceptiveness to internal and external stimuli, which increases the irritability of the cardiac muscle and possibly lead to ectopic beats.
5. *Dilation of coronary blood vessels*. The diameter of the coronary blood vessels increases, although the systemic arteries constrict.

Depending on the intensity of a particular emotion, the sympathetic system is stimulated to prepare the body for *vigorous* activity. Apart from the well-established cardiac reaction to ANS, there is a significant amount of work in the medical field that investigates the exact properties of the ANS innervation. For instance, it is not clear whether the sympathetic and parasympathetic branches of ANS affect the heart in the same way, i.e., whether there exist independent ANS drives on the cardiac muscle or not.

Shouldice et al. [17] provided evidence that there is separate ANS innervation to the two pacemakers (SA and VA nodes). By analyzing PP and PR intervals in transitions from supine to stand postures and vice versa, the authors showed a decoupling of the SA and VA modulation. This suggests that the cardiac reactivity may contain more detail about the stimulus than just the intensity (arousal). Another interesting finding in [17] is that there is large intersubject autonomic innervation variability, meaning that one cannot generalize how the excitability of the SA and VA nodes is affected over a population. Their work leads one to suspect the specificity of emotion with regard to particular people.

Furthermore, signals such as the HRV and BVP have been associated with the development of coronary artery disease. The cardiac reactivity to stress has been studied in relation to feelings such as anxiety, hostility, or challenge [18], [19]. Despite the dependence of the HRV time series on posture (standing/supine), increased vagal modulation of the cardiovascular system was statistically correlated with the aforementioned emotions. Similarly, the difficulty levels of memory tasks were shown to impact cardiac reactivity [20]. Furthermore, Blascovich et al. [21] were able to differentiate challenge from threat using cardiac traits such as the HR, ventricular contractility, and cardiac output.

As already discussed, the most widely studied trait of cardiovascular activity is the HR or HRV [22], [23], [24]. Based on the effects of sympathetic stimulation on the cardiac muscle, the HR is the most natural choice for arousal detection using comparison of sympathetic and parasympathetic frequency bands of the time series [24]. However, it is highly dependent on the position of the body during monitoring [18], [25].

Emotion specific reactivity has been observed on the ECG signal itself, without collapsing the embedded information to arbitrary interval measurements (HRV). In fact, Andrassy et al. [26] observed QT prolongation (typical for stress) without significant R–R interval changes.

Folino et al. [27] recorded measurements of features such as the duration of the QRS complex, low amplitude waves at the terminal portions of the complex, and the root mean square voltage of the QRS while the subject was performing various mental arithmetic tasks. They observed that with increasing levels of difficulty, the energy of the QRS complex increased significantly, while its duration was reduced (not a result of increasing HR).

The T wave amplitude was cross examined with mental stress in [28], [29]. An interesting observation brought to light by these works *was that manifestations of mental stress on ECG depends on whether the stressor is active or passive*. Contrary to the popular belief that mental stress response is generic, no significant effects were observed during passive tasks (i.e., stressors without the active involvement of the subject).

Apart from mental stress, cardiovascular responses have been examined for other particular emotions. Sinha et al. [7] experimented with five emotional states, namely, joy, sadness, fear, anger, and neutral. The HR, BVP, stroke volume, preejection period, and cardiac output were measured, employing imagery tasks for emotion elicitation. Pairwise comparison performance (for every two emotions) was reported for every measurement, with the HR being among the signals that exhibit the least discriminative power.

In summary, there is significant evidence in the literature that cardiovascular reactivity can differentiate not only the intensity of the stimulus but also the valence, depending on how the stimulus is presented. Although, in the affect research (Section 2), ECG is primarily used for HR estimation, we argue that it is not used to its full extend as this signal is still one of the most illustrative and detailed records of cardiac activity. Furthermore, ECG exhibits idiosyncratic patterns due to the unique anatomy of people's cardiac muscles. Along these lines, in this work we investigate the dynamics of the ECG signal and their association with emotion specific ANS activity, from a subject specific point of view.

3.3 Signal Processing for Emotion Detection

There are two main challenges that arise with signal processing of physiological signals when targeting emotion pattern classification. First, emotion-specific patterns are not well defined for biosignals. Second, it is typically uncertain whether emotions were manifested at all. For ECG signals, despite the reports on cardiovascular reaction to emotion wherein there are inconclusive findings about the various waveform patterns, the majority of previous works agree on the subject-specific nature of the emotion manifestation.

We argue that for ECG signals, universal emotion detectors run the risk of being inaccurate for the following two reasons:

1. ECG has been established as a biometric characteristic [30], [31], which means that by default it carries subject-specific information. The appearance of a particular heart beat depends on a number of factors, such as the geometry and orientation of the cardiac muscle, the conductivity of various areas of the heart, the activation order [32], and the subject's habitus or gender [33], [34], [35]. This *physiological aspect* of ECG formed a strong basis for the investigation and establishment of its biometric properties (namely, its use in discriminating individuals for recognition applications) [30], [31].
2. Different people experience emotions in different ways. Even if we ignore the biometric aspect of the ECG signal, there is large ANS innervation variability in a population [17].

For these reasons, emotional patterns are herein detected as variations from the typical appearance of an individual's ECG signal. We perform feature extraction in the time domain for two reasons—the time-domain signals have been reported to be more resourceful with regard to emotion specific features [36], and it is risky to impose stationarity and linearity restrictions (necessary for Fourier transform) on ECG [37]. Furthermore, since we cannot predict the effects of emotional processes on the ECG signal, there is a risk of missing the dynamic changes due to emotion if we use methodologies that rely on predefined bases.

We are more interested in the way properties of the ECG signal evolve over multiple heart-beats when a person experiences different emotions. The local properties of any particular pulse are of little interest. In this regard, the Empirical Mode Decomposition (EMD) [37] is a powerful tool since it is dynamic, data driven, and examines the signal holistically to highlight underlying trends. Without prior assumptions on the properties of the signal, EMD adapts to the embedded oscillatory activity and decomposes ECG into a number of intrinsic modes. It is our belief that understanding the intrinsic modes that are hidden in the cardiac oscillatory activity is an essential first step in any attempt to classify psychological states.

4 METHODOLOGY

The proposed methodology for ECG feature extraction is based on the Empirical Mode Decomposition. EMD is adaptive and the basis of the decomposition is self-defined, which makes it suitable for the analysis of complex underlying phenomena. In addition, since EMD operates in the time domain it favors the exploration of silently evolving trends such as low arousal emotions. However, it turns out that adaptivity is a mixed blessing. There are two important problems with regard to *uniqueness* and *mode mixing* that arise as a consequence of the adaptive nature of EMD. Evaluating the EMD algorithm on two instances of the same signal may result in incomparable decompositions. Moreover, each mode is not captured in a single decomposition—similar modes are present at various decomposition levels of the same signal. In our work, we address these issues by making use of a bivariate extension of EMD (BEMD) [38]. BEMD acts on two signals $x_I(t)$ and $x_S(t)$ (as opposed to EMD which acts on a single signal). In our work, the second signal ($x_S(t)$) is a synthetic ECG waveform, standardized over all emotional states and subjects, and designed to act as a decomposition guide. The proposed framework is comprised of three independent steps:

1. ECG synthesis, wherein a signal $x_S(t)$ is designed to be synchronous with a particular input $x_I(t)$.
2. Estimation of the oscillatory modes, called *Intrinsic Mode Functions* (IMF) ($d(t)$), of the input signal via decomposition using the BEMD.
3. Extraction of features associated with the instantaneous frequency and the local oscillation of the IMFs and classification among predefined affect states.

4.1 Step 1: ECG Synthesis

The objective of this step is to design a synthetic signal $x_S(t)$ whose properties are similar to a real ECG signal and whose main waves are synchronized with the waves of the input signal $x_I(t)$. In order to synchronize the main waves of the synthetic signal, we must first estimate the location of the `P`, `QRS`, and `T` waves throughout $x_I(t)$. It is important to note that ECG delineation is a difficult task due to the large variability of ECG signals. There is no universally accepted rule to guide the localization of fiducial points on heart beats. The *QRS* complex is detected using the algorithm described in [39]. The surrounding waves are localized using empirical rules—the `P` wave's healthy duration is approximately 120 msec, and the `T` wave extends about 300 msec after the *QRS* complex [15]. Once the fiducial points of the input signal have been localized, the synthetic ECG signal can be generated using the dynamic generation model described by McSharry et al. [40]. The model generates a trajectory in a 3D state space with coordinates `(x,y,z)`. Quasi periodicity of the ECG is reflected by the movement of the trajectory around an attracting limit cycle of unit radius in the `(x,y)` plane, at varying speed. A completion of one cycle along this circle is equivalent to a heart beat completion or one `R-R` interval. By adjusting the speed of the trajectory, quasi periodicity is achieved. Interbeat variation in the ECG signal is reproduced by the motion of the trajectory in the `z` direction. The location of the fiducial points of the input

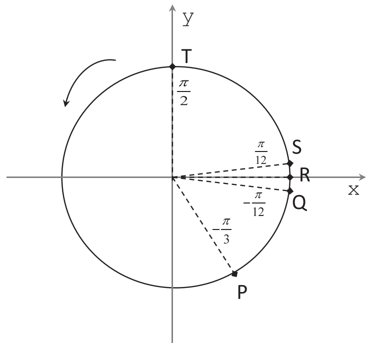

Fig. 2. Two-dimensional trajectory movement and P, Q, R, S, T typical locations.

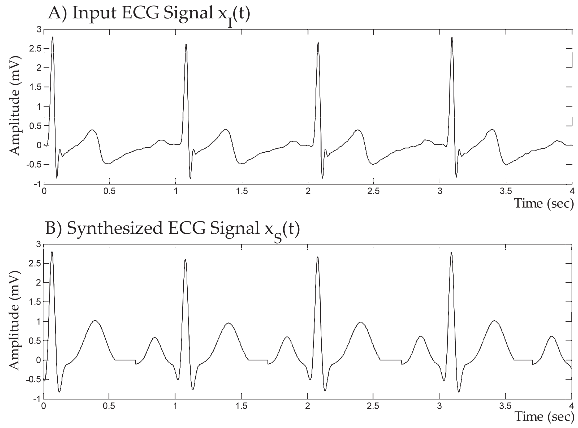

Fig. 3. A real and a synthetic ECG. The two signals are synchronized but $x_S(t)$ has no anatomical uniqueness or psychological variability.

signal define special events on the unit circle indexed by the angles $\theta_P, \theta_Q, \theta_R, \theta_S, \theta_T$ corresponding to points P, Q, R, S, T, respectively. These distinct points on the ECG are described in the model by events corresponding to the negative and positive attractors/repellors in the z direction. Each time the trajectory reaches one such point, it is repelled away from the unit circle along the z axis and then attracted back, creating a wave. The dynamical equations of motion along this trajectory are given by a set of three ordinary differential equations [40]. The authors in [40] provided typical values for the angular location of the fiducial points (Fig. 2) based on visual inspection of healthy ECG waveforms. The duration of one circle is set to a maximum of 1 second (rest condition, 60 bpm). Given the distances d_{ij} (in seconds) between the fiducial points i and j of $x_I(t)$, the points can be located along the unit circle using the following equations (wherein the location of a point is specified in radians): $\theta_P = -2\pi\frac{d_{PR}}{d_{RR'}}, \theta_Q = -2\pi\frac{d_{QR}}{d_{RR'}}, \theta_S = 2\pi\frac{d_{RS}}{d_{RR'}}, \theta_T = 2\pi\frac{d_{RT}}{d_{RR'}}$, where $d_{RR'}$ is the length of the R-R interval in seconds and $\theta_R = 0$. These angles are applied on the differential equations in [40] to obtain $x_S(t)$. The generated signal is an idealized, robust, noise-free representation of ECG (see Fig. 3 for an example of a synthetic ECG signal).

4.2 Step 2: Signal Decomposition

Huang et al. [37] proposed the Empirical Mode Decomposition as a way to empirically decompose a nonstationary, nonlinear signal into a number of IMFs, each of which represents a distinct oscillatory activity. An IMF is a function satisfying certain explicit properties [37]:

1. The number of extrema and the number of zero crossings must be equal or differ at most by one.
2. The mean of the envelopes defined by the maxima and the minima is zero for every sample.

These rules naturally force one mode of oscillatory activity in the IMF, since between two successive extrema no riding waves are allowed. The algorithm for the detection and extraction of IMFs is adaptive and iterative. Once an IMF is found, it is removed from the signal and the algorithm iterates on the residual in order to find more oscillatory modes. Fast oscillations are detected first. Given a signal $x(t)$, EMD operates as follows:

1. Detect local maxima $x_{max}(i)$ and minima $x_{min}(j)$ of $x(t)$.
2. Interpolate among $x_{max}(i)$ to get an upper envelope $x_{up}(t)$ and $x_{low}(t)$ for minima, respectively.
3. Compute the average of envelopes $m(t) = \frac{x_{up}(t)+x_{low}(t)}{2}$.
4. Subtract from signal $u(t) = x(t) - m(t)$.
5. Iterate for $x(t) = u(t)$.

This is a *sifting* process, which is terminated when $u(t)$ meets the IMF criteria. If it does, $u(t)$ will be describing a underlying oscillation of $x(t)$, referred to herein as $d(t)$. EMD continues with sifting on the residual $r(t) = x(t) - d(t)$. The original signal can then be expressed as

$$x(t) = \sum_{i=1}^{N-1} d_i(t) + r(t), \tag{1}$$

where $d_i(t)$ denotes the ith IMF extracted from the signal $x(t)$ and $r(t)$ is the final residual. Note that by definition, $r(t)$ is not an IMF. Fig. 4 shows an example of the resulting IMFs when EMD is applied on a synthetic signal $x_S(t)$ which is an idealized, noise-free waveform. It has been observed [36], [41], [42], [43] that the first few IMFs carry the quasi-periodicity property of ECG. Since every heart beat has three distinct waves (Fig. 1) in the absence of noise, the first IMF is expected exhibit three oscillatory components (*tricomponent*), primarily characterizing the behavior of the QRS complex. This is because the QRS complex contributes to the highest frequencies of the ECG. The first IMF, as shown in Fig. 4, depicts the fastest oscillating component of the signal. Once this is removed, the second and third IMFs exhibit *bicomponent* and *monocompoent* oscillations, respectively. As the IMF order increases, the strength of the oscillation decreases. However, for noise-free ECG, IMFs of order higher than three are almost zero and will be ignored hereafter. As stated before, the EMD encounters problems with regard to *uniqueness*. The number and type of IMFs generated by EMD are uncertain, even for signals with similar statistics. For instance, the decomposition in Fig. 4 would result in more IMFs and in stronger oscillations if there was high frequency noise in the signal. This restricts the utility of EMD as it renders comparisons among different (but real) ECG signals meaningless. Predetermining the number of IMFs (by forcing decomposition to stop) defeats the purpose of EMD as the analysis will no longer be

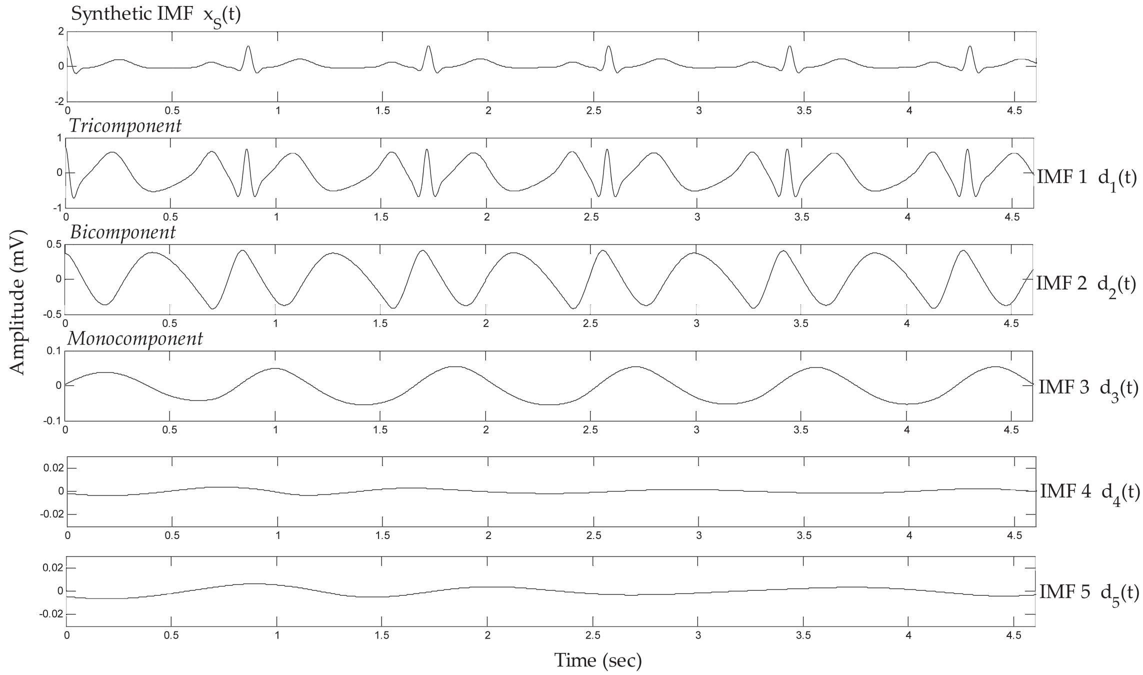

Fig. 4. EMD analysis for a standardized synthetic ECG signal. IMFs of order higher than three do not exhibit oscillatory activity. The first IMF has three oscillatory components, the second has two, and the third has one.

adaptive, nor will the IMFs have physical meaning. We address this problem by decomposing a real ECG signal $x_I(t)$ together with a synthetic signal $x_S(t)$ in a module that allows the latter to act as the *rule* of decomposition by determining which type of IMFs are important from the input signal. This is done using the BEMD [38] algorithm on the pair of signals $x_I(t)$ and $x_S(t)$. The difference between EMD and BEMD can be visualized as follows: Whereas EMD-sifting builds envelopes around $x(t)$, BEMD builds 3D cubes that surround a complex function $x_c(t)$. The analysis is performed simultaneously for the real and imaginary components of $x_c(t)$, and results in the same number of IMFs for both:

$$x_c(t) = \sum_{i=1}^{N-1} d_{c_i}(t) + r_c(t), \tag{2}$$

where $d_{c_i}(t)$ denotes a complex IMF and $r_c(t)$ the complex residual. Low order $d_{c_i}(t)$ describe fast rotating components, while the opposite is the case for higher order complex IMFs. Because of the consistency in the analysis of the real and the imaginary parts, BEMD has been suggested for signal separation (detrending) in filtering applications [44], [45]. For our purpose, we form a complex signal using the input ECG signal $x_I(t)$ and the synthetic signal $x_S(t)$ as the real and imaginary parts, respectively,

$$x_c(t) = x_I(t) + jx_S(t). \tag{3}$$

Applying BEMD on $x_c(t)$ we get

$$x_c(t) = \sum_{i=1}^{N} Re\{d_i(t)\} + j\sum_{i=1}^{N} Im\{d_i(t)\}, \tag{4}$$

where the residual has been included in the summation for simplicity. Similarly to the univariate case, Fig. 5 shows an example of five IMFs acquired from a BEMD analysis of two heart beats (one real and one synthetic). Since the synthetic ECG has an idealized waveform, the presence of oscillatory activity on the imaginary side guarantees that the corresponding mode is present on the real side. When the real ECG is contaminated with high-frequency noise, low order IMFs on the real part will exhibit strong (but physiologically meaningless) oscillations while almost zero activity will exist on the imaginary side, which makes them easily detectable.

4.3 Step 3: Feature Extraction

The IMFs are time domain signals carrying information about oscillation activity. Comparisons in the time domain are not straightforward since the IMFs $d_i(t)$ have to be aligned with similar modes from other ECG recordings. Therefore, it is important to design features that summarize the oscillatory activity within every IMF. In this work, we use two types of features—the *Hilbert instantaneous frequency* and a measure of *local oscillation*.

4.3.1 Instantaneous Frequency

The Hilbert transform is typically used in conjunction with the EMD because it provides accurate estimates of the instantaneous frequency for monochromatic signals like the IMFs. The transform is defined as the convolution of a signal with $h(t) = \frac{1}{\pi t}$. For each of the IMFs $d_i(t)$, the Hilbert transform is applied as follows:

$$\mathcal{H}[d_i(t)] = \frac{1}{\pi} P.V. \int_{-\infty}^{+\infty} \frac{d_i(\tau)}{t-\tau}\, \mathrm{d}\tau, \tag{5}$$

where P.V. indicates the Cauchy principal value. We can define the following analytical signals:

$$z_i(t) = d_i(t) + j\mathcal{H}[d_i(t)], \tag{6}$$

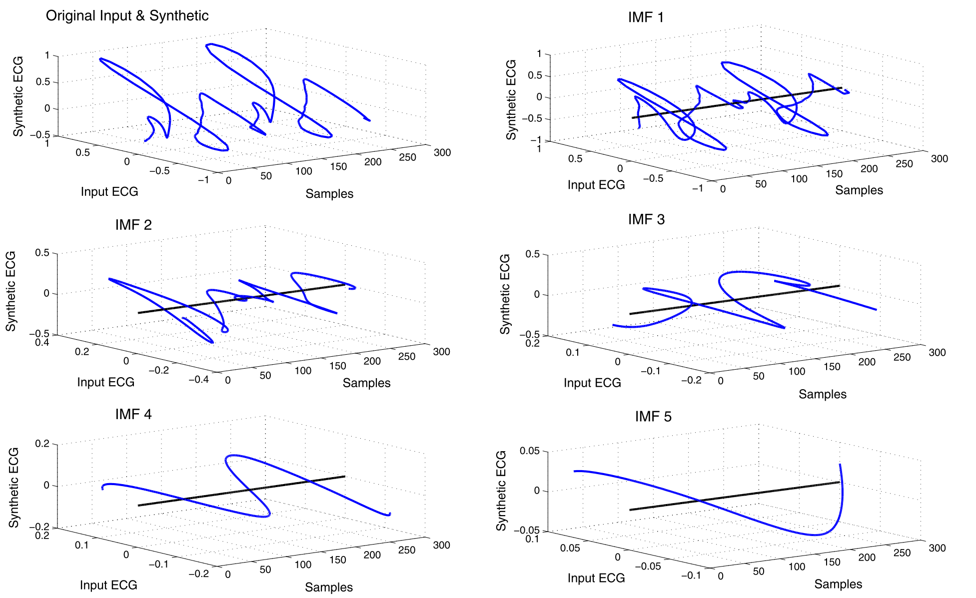

Fig. 5. BEMD example on a complex ECG signal, formed using a real ECG segment and a synthetic one. Low order IMFs show fast rotating components.

which can be rewritten as

$$z_i(t) = y_i(t)e^{j\theta_i(t)}, \tag{7}$$

where $y_i(t)$ is the magnitude and $\theta_i(t) = \arctan(\frac{\mathcal{H}[d_i(t)]}{d_i(t)})$ is the phase of the complex $z_i(t)$. For IMF i the instantaneous frequency can then be computed as

$$f_i(t) = \frac{1}{2\pi}\frac{\mathrm{d}\theta_i(t)}{\mathrm{d}t}. \tag{8}$$

Essentially, $f_i(t)$ is a measure of changeability within every IMF.

4.3.2 Local Oscillation

Modeling the type of oscillation within an IMF is a difficult problem due to the empirical nature of EMD. We herein propose features that are extracted as measures of the oscillation time scale, rather than the velocity or frequency of alternation among extrema. For every IMF $d_i(t)$, let $\mathbf{u}_i$ and $\mathbf{v}_i$ be the time instances of the maxima and minima, respectively. By definition (of IMFs), there is one zero crossing between every pair of consecutive extrema and the objective is to define the rate of interchange. We define $\rho_i(t)$, a function that describes local oscillation, which is estimated by parsing an IMF as follows:

1. For an element $\mathbf{u}_i^k$ of $\mathbf{u}_i$, examine $d_i(t)$ in the interval $[d_i(u_i^k), d_i(u_i^{k+1})]$ (between two consecutive maxima).
2. Compute *max-to-min* and *min-to-max* transition times: $a = v_i^k - u_i^k$, $b = u_i^{k+1} - v_i^k$.
3. $y_i(t) = \min(a, b)$, $u_i^k \leq t \leq u_i^{k+1}$.

The local oscillation $\rho_i(t)$ is then computed from $y_i(t)$ with normalization across all IMFs as follows:

$$\rho_i(t) = \frac{1 - y_i(t)}{A}, \tag{9}$$

where A is the maximum of all $\rho_i(t)$, i.e., local oscillation is normalized across all IMFs. The higher the values in $\rho_i(t)$, the faster the local extrema interchange. Fig. 6 shows $\rho_i(t)$ for the first three IMFs of a synthetic ECG. The oscillation histogram shows that the first IMF has three major components, the second has two, and the third just one. For a real ECG signal, this distinction is unclear as mode mixing is a very common phenomenon for EMD, i.e., different types of oscillatory patterns appear throughout one IMF. To address this problem, classification is carried out only among the *dominant* local oscillations and frequencies. The feature vector is the concatenation of these measures for the first three IMFs, which exhibit quasi periodicity. More specific, since the first IMF is a tricomponent signal, the three most frequent local oscillations, referred to as $\hat{\rho}_1^1$, $\hat{\rho}_1^2$, and $\hat{\rho}_1^3$ (see Fig. 6G), will participate in the feature vector. Similarly, three frequency measures are considered for the first IMF ($\hat{f}_1^1$, $\hat{f}_1^2$, and $\hat{f}_1^3$). Overall, the feature vector is formed as follows:

$$x = \left[\hat{\rho}_1^1, \hat{\rho}_1^2, \hat{\rho}_1^3, \hat{\rho}_2^1, \hat{\rho}_2^2, \hat{\rho}_3^1, \hat{f}_1^1, \hat{f}_1^2, \hat{f}_1^3, \hat{f}_2^1, \hat{f}_2^2, \hat{f}_3^1\right]. \tag{10}$$

Classification is carried out using linear discriminants among a number of predetermined classes (see Section 6). The assumption in this work, to be verified experimentally, is that both the local oscillations and the instantaneous frequencies carry emotion specific information.

5 Experimental Protocol

The proposed methodology was evaluated on data from two experiments that were conducted at the Affect and Cognition Laboratory, of the University of Toronto. In the first experiment, visual stimuli (passive stressors) were used to induce positive and negative emotions for valence and passive arousal performance evaluation. The second experiment attempted to induce active mental stress (*active arousal*) using a video game.

5.1 Experiment I. ECG Reactivity to Valence and Passive Arousal

The International Affective Picture System (IAPS) [13] was used in this experiment as a passive stimulus for emotion

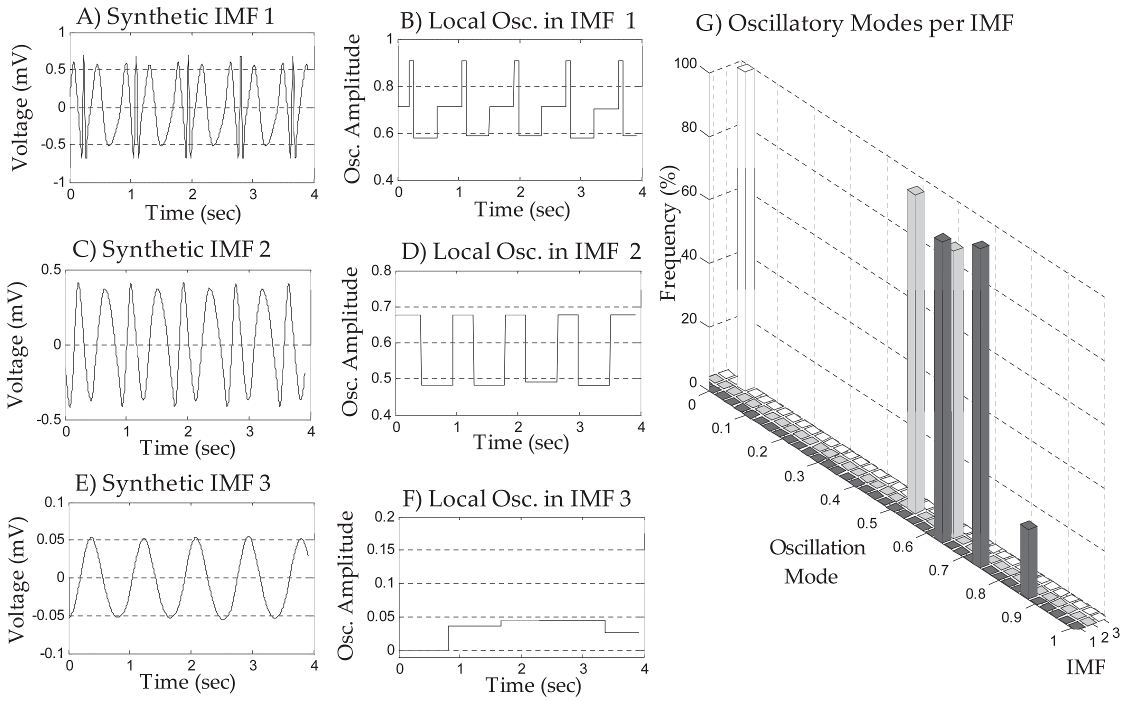

Fig. 6. Local oscillation for a synthetic ECG. (A), (C), (E) First three IMFs of the synthetic signal. (B), (D), (F) $\rho_i(t)$ for the previous IMFs. (G) Dominant oscillations for the three IMFs. With increasing order of IMF, the strength of the oscillation (time scale) decreases.

elicitation. Pictures from this photoset induced the following emotional conditions: *gore*, *fear*, *disgust*, *excitement*, *erotica*, and *neutral*. These conditions cover the valence spectrum from negative (gore, fear, disgust) to positive (erotica, excitement). Also, with respect to the arousal, the gore and erotica stimuli are considered to have induced the most highly arousing states. An emotional state was achieved by displaying five images of the same emotional target for 8 seconds each. In total, 44 volunteers participated in this experiment. Recordings from 12 individuals were discarded due to noise and artifacts. Lead I ECG was collected at 1 KHz with the BIOPAC MP 150 system. Five pictures of the same emotional condition (emotion batch) were displayed in random order. Every batch was repeated twice with a different set of pictures. Between batches the subject was asked to perform a simple arithmetic task that would bring the psychological activity to baseline. Upon completion of the experiment, the volunteers were instructed to look at the pictures again and report, subjectively, valence and arousal on a scale between 1 and 9 (with 1 indicating low arousal or negative valence). These self-reports were used in order to validate the success of the experiment for each picture that was displayed. The purpose was to identify those ECG readings for which the targeted emotional condition failed to be induced. Any such reading was discarded as an outlier. With this procedure only 2 percent of the ECG segments were discarded.

5.2 Experiment II. ECG Reactivity to Active Arousal

In this experiment, a commercial video game was used to elicit active mental arousal. The game was designed to present increasing difficulty. The goal was to have the player gradually immersed in the game, by increasingly concentrating in order to meet the game requirements. The subjects got motivated with deception by letting them know that the purpose of the experiment is to measure game completion time. In total 43 volunteers participated in this study. Data from one person were discarded due to noise.

Depending on the familiarity of the subject with game playing, the duration of the experiment varied between 20-45 minutes. During the game, ECG was monitored using Hidalgo's Equivital sensor, which is portable and wireless. Unobtrusiveness was very important for the subjects to naturally immerse to the game. ECG was recorded from the chest, and digitized at 256 Hz. Because of the chaotic nature of the game and the unforeseeable order of events that can take place, arousal annotation was self-determined. For this reason, a video of the player's facial expressions was captured during the game (synchronous to ECG). Upon game completion, the subjects were asked to watch a playback video of the game and their facial expressions while continuously reporting arousal using FEELTRACE [46]. The valence axis of FEELTRACE was intentionally ignored as a video game experience is too complex to accurately infer instantaneous valence.

In order to compare the effects of active and passive arousal, we asked the same volunteers to participate in both experiments to acquire subject-specific arousal data for both scenarios. In total, signals from 31 volunteers were common and eligible for processing.

6 Experimental Performance

6.1 Valence Classification

A first step to ECG signal processing is noise filtering. In this setup, a butterworth bandpass filter with cutoffs at 0.5 and 40 Hz was used. The order of the filter was set to 4 based on empirical results. Further processing of the signal is slightly

TABLE 1
List of Classification Experiments Performed

5-class: Distinct emotions		**Experiment A** *class 1*: Erotica, *class 2*: Excitement, *class 3*: Fear, *class 4*: Disgust, *class 5*: Gore
Valence Differentiation	Arousal Dependent	**Experiment B** *class 1*: Erotica *class 2*: Gore
		Experiment C *class 1*: Erotica *class 2*: Disgust
	Arousal Independent	**Experiment D** *class 1*: Erotica, Excitement *class 2*: Fear, Gore, Disgust
Within Valence Differentiation	Within Positive	**Experiment E** *class 1*: Erotica *class 2*: Excitement
	Within Negative	**Experiment F** *class 1*: Fear *class 2*: Disgust
		Experiment G *class 1*: Gore *class 2*: Disgust
		Experiment H *class 1*: Gore *class 2*: Fear
Arousal Differentiation		**Experiment I** *class 1*: Erotica, Gore *class 2*: Neutral

different for the two experiments. In the IAPS case, the signal is divided into segments corresponding to the particular emotional conditions of the experiment (erotica, excitement, disgust, fear, gore, neutral). Every segment was further subdivided into a number of windows of 10 seconds length. This guarantees that only one emotional condition was experienced during this time frame. Practically, the goal was to classify a state in every 10 seconds, but this duration is flexible to the needs of a particular application environment. Out of the examined conditions, erotica and excitement fall under positive valence, with erotica being of higher arousal. On the negative valence side, gore exhibits the strongest arousal. A variety of experiments were conducted as listed in Table 1. The goal was to demonstrate that ECG can differentiate among or within the same valence conditions, when this operation is subject dependent. Synthetic ECG signals where generated and participated in BEMD analysis for each of the 10 seconds segments. The instantaneous frequency and the local oscillations of the first three IMFs were used as classification features. Since pictures of the same condition were displayed twice, data from the first batch were used for training and from the second one for testing. Classification was performed with linear discriminants, and with leave-one-out cross validation. More sophisticated classifiers can be applied as well, but this is beyond the scope of this work.

Fig. 7 shows per subject performance for experiments A, B, C, and D. It can be observed that among highly aroused conditions (gore and erotica), valence can be differentiated for most subjects, but not for all. For certain individuals erotica is more distinguishable from disgust than gore. As it will be clarified in the subsequent discussion, this observation is related to the intensity with which every participant perceived the experiment. For individuals that reported stronger immersion, experiment C indicates higher classification performance.

The performance for all experiments in Table 1 is graphically depicted in Fig. 8. It is important to note that for each case there are individuals which score considerably high, but it is difficult to generalize especially for the within valence classification cases. However, valence differentiation (in experiment B) is clearly feasible as performance ranges from 52 to 89 percent with a standard deviation of 11.16. An interesting finding is also with regard to classification between two negative valence but high arousal states (gore and fear). As can be seen from Fig. 8 ECG can differentiate within negative emotions with higher probability than within positive. However, this might be attributed to the fact that negative emotions were better induced with visual stimuli.

Subject independent experiments were also performed and the performance is herein reported for reference. For experiment B, of valence differentiation, the classification rate is 46.56 percent. Similarly, for experiment H, of within negative valence classification, the performance is 53.07 percent. These rates correspond to random picks for the two class problems and support the expectations for subject-specificity of the ECG waveform.

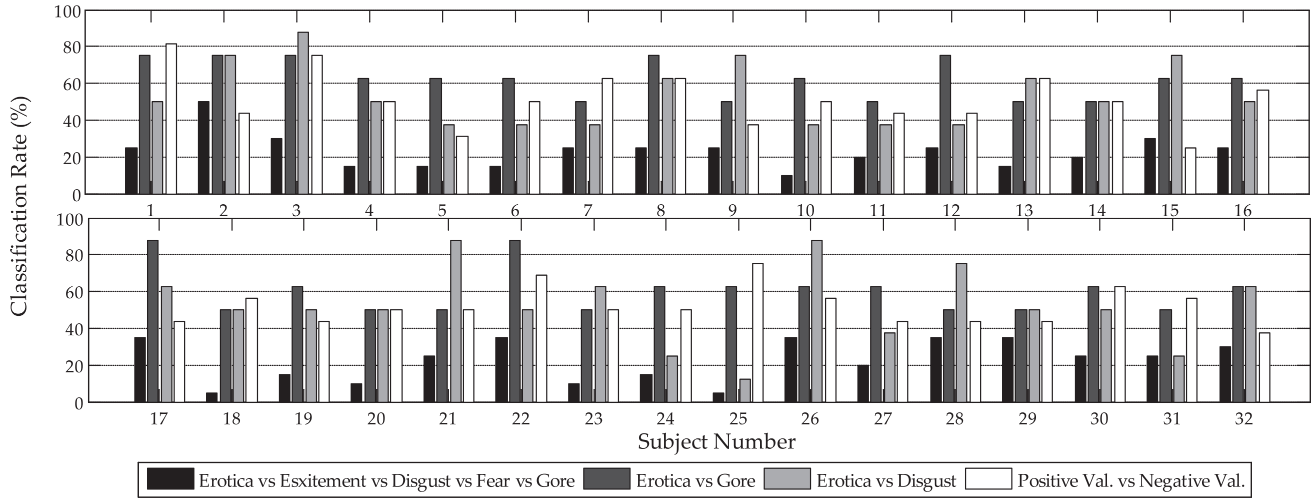

Fig. 7. Per subject classification performance for each of the 32 individuals in the database. While classification among five classes is promising for certain individuals, valence separation is feasible with respect or irrespective of arousal (erotica versus gore or erotica versus disgust).

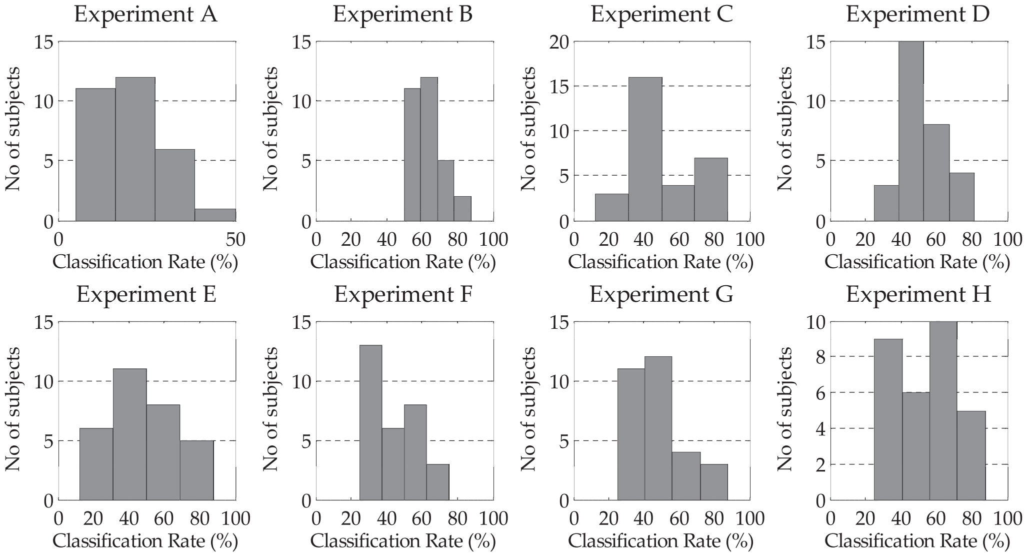

Fig. 8. Classification performance for all subjects in experiments A-H.

6.2 Arousal Classification

For the data of the active arousal experiment, a similar windowing approach was adopted (10 seconds segments). Since the self-assessment was done using FEELTRACE [46], labeling was continuous for the duration of the video game. On FEELTRACE, arousal takes values in $[-10, +10]$ with -10 indicating very low emotional intensity (boredom) and $+10$ the opposite. Depending on the desired precision, arousal was quantized to a number of levels, and the average within every 10 seconds window was used to characterize it. The standard deviation was also computed for every segment to ensure the emotional consistency within the 10 seconds time frame. Two classification problems were examined, i.e., a three class problem of discrimination among *low* ($[-10, -6]$), *medium* ($(-6, +6)$), and *high* arousal ($[+6, +10]$), and a two class problem of discrimination between *low* ($[-10, 0)$) and *high* ($[0, +10]$) arousal. Because for a game experience it is difficult to report instantaneous valence information, the volunteers were asked to focus their reporting to the arousal dimension. However, continuous self-reporting is ambiguous, first because determining the actual onset of excitement or boredom can be misleading, and second because even if the onset of a cognitive reaction is accurately determined, there might be cardiac latency, especially for transitions between high-to-low arousal. For example, when an interesting event in the game is followed by a monotonous situation (for instance because the subject got lost in the map), the progression between high-to-low arousal is gradual and it is thus difficult for the player to assess the exact onset of boredom. Perfect data annotation is, of course, of great importance, and this was expected to add to the overall classification error. However, the immersion that video games provide, in conjunction with the user being actively engaged to the task (active stressor), proved beneficial for the problem in hand.

For every subject, the first 10 min of data were used for training and the remaining for testing. Every feature vector was the concatenation of instantaneous frequency and local oscillation measures. Classification was performed with linear discriminants, and it was subject dependent.

For the three class problem (*low*, *medium*, and *high* arousal), the average classification rate over all subjects is 35.36 percent. For the two class problem (*low* and *high* arousal), the average rate over all subjects is 76.19 percent. The mean, however, is not very descriptive for the population of 42 individuals, where classification is subject dependent. Fig. 9 shows the probability mass function for the two class problem. For approximately 50 percent of the subjects, a detection rate between 90 and 100 percent was achieved.

6.3 Active versus Passive Arousal

Inspired by the reports of varying cardiac reactivity due to passive or active stressors [28], [29], [47], a performance comparison between arousal classification in Experiments I and II was performed. In total there are 31 subjects who participated in both experiments, and for whom artifact-free data are available. Performances were compared for the two class problems of discrimination between *high* and *low* arousal. In Experiment I, this was achieved by distinguishing neutral reactions from the union of gore and erotica conditions.

Fig. 10 lists the classification performance per subject, for both passive (Experiment I) and active (Experiment II) arousal. The average rate is 52.41 and 78.43 percent,

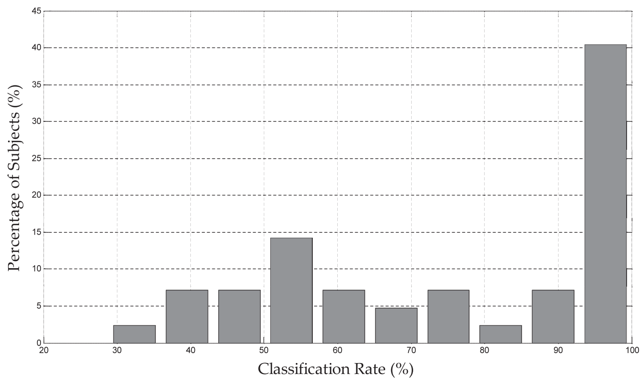

Fig. 9. Arousal detection performance for 42 subjects. The average rate is 76.19 percent and the standard deviation 23.06.

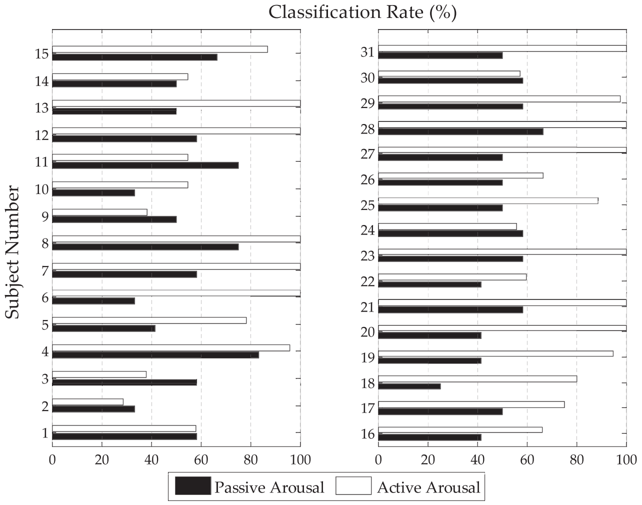

Fig. 10. Subject specific arousal detection for the two experiments. The average rate for all subjects under passive arousal is 52.41 percent. Similarly, for active arousal the average is 78.43 percent.

respectively. For the majority of the subjects, active arousal detection outperforms passive. For the few subjects for whom this is not the case, the reason might be that visual stimuli were of equal immersion as the game.

7 Conclusion

In emotion research, it is very important to collect meaningful data. It is very difficult to design an experimental setup that can induce the same emotion in every subject, especially if the same stimuli are used across all subjects. Different characters, varying moods, and the inability to accurately self-report an emotional experience may significantly affect the outcome of a study. In particular, when internal modalities like the ECG are examined, data labeling is very subjective and can only be verified by the participants themselves.

Despite the difficulties, establishing emotions from internal manifestations of the body is worthwhile for human computer interaction systems. For the naive user, hiding emotions with regard to cardiac reactions is difficult, while, for behaviorally suppressive individuals, physiological patterns can provide hints of emotion. The obtrusive nature of the acquisition, however, poses restrictions on the number of sensors that can be worn and subsequently to the signals that can be collected. The majority of the approaches in the literature rely on fusion of various physiological sources for emotion detection. Although such treatment presents significant performance benefits, they have limited practicality. Following the findings of this study, future treatments will attempt to establish the limits of emotion assessment for other biosignals as well. Although ECG has been extensively employed in affect research, little attention has been paid so far to the waveform patterns under emotional activity. A typical approach is to extract the *R-R* intervals for the computation of the HRV and treat the remaining signal as redundant. However, given that the ANS has endings in each of the four chambers of the heart, it is expected that ECG will exhibit emotion specific patterns. However, for making any analysis meaningful, one has to take into consideration the specificity of the signal to the particular subjects due to both the biometric aspect of the signal and the specificity of the ANS innervation.

This work proposed subject-dependent ECG signal decomposition using the BEMD. Local oscillation and instantaneous frequency features were used for the detection of emotional conditions as their combination fully describes the oscillatory activity of every IMF. It was concluded that adaptive, data-driven analyses are suitable for emotion modeling because of the unforeseeable patterns that may arise within every individual. In this work, the original EMD algorithm was implemented for demonstration of its full potential. Faster and more efficient implementations of this algorithm may also be considered [48], [49].

With the employment of oscillation data as the feature space, it was shown that ECG waveform reactivity depends highly on the activeness of the emotional experience. Data from visual stimuli inspection resulted in 26.02 percent lower arousal detection performance (on average for 31 volunteers) than playing a video game. This finding is in agreement with prior reports [28], [29], [47] on cardiovascular differentiation between active and passive stress. Valence could also be differentiated, especially for high arousal states, with an average of 62 percent and a standard deviation of 11.16. This work statistically demonstrates the feasibility of using the ECG for emotion detection. Since the signal has a physiological (relative to the heart's anatomy) and psychological aspect (herein established), it is interesting to investigate statistical methods that can separate the two. This would be beneficial both for biometric recognition (where ANS introduces variability) and for emotion detection.

References

[1] R.W. Picard, *Affective Computing*. MIT Press, July 2000.

[2] P. Ekman, R.W. Levenson, and W.V. Friesen, "Autonomic Nervous System Activity Distinguishes among Emotions," *Science*, vol. 221, pp. 1208-1210, Sept. 1983.

[3] F. Hönig, A. Batliner, and E. Nöth, "Real-Time Recognition of the Affective User State with Physiological Signals," *Proc. Doctoral Consortium Conf. Affective Computing and Intelligent Interaction*, pp. 1-8, 2007.

[4] J. Anttonen and V. Surakka, "Emotions and Heart Rate While Sitting on a Chair," *Proc. SIGCHI Conf. Human Factors in Computing Systems*, pp. 491-499. 2005,

[5] C.M. Jones and T. Troen, "Biometric Valence and Arousal Recognition," *Proc. 19th Australasian Conf. Computer-Human Interaction*, pp. 191-194, 2007.

[6] R.L. Mandryk, T.W. Inkpen, and K.M. Calvert, "Using Psychophysiological Techniques to Measure User Experience with Entertainment Technologies," *Behaviour and Information Technology*, vol. 25, no. 2, pp. 141-158, Mar.-Apr. 2006.

[7] R. Sinha, W.R. Lovallo, and O.A. Parsons, "Cardiovascular Differentiation of Emotions," *Psychosomatic Medicine*, vol. 54, pp. 422-435, 1992.

[8] 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.

[9] J. Kim and E. André, "Emotion Recognition Based on Physiological Changes in Music Listening," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2067-2083, Dec. 2008.

[10] F. Nasoz, K. Alvarez, C.L. Lisetti, and N. Finkelstein, "Emotion Recognition from Physiological Signals Using Wireless Sensors for Presence Technologies," *Int'l J. Cognition, Technology, and Work*, special issue on presence, vol. 6, no. 1, pp. 4-14, Feb. 2004.

[11] K. Kim, S. Bang, and S. Kim, "Emotion Recognition System Using Short-Term Monitoring of Physiological Signals," *Medical and Biological Eng. and Computing,* vol. 42, no. 3, pp. 419-427, May 2004.
[12] R.L. Mandryk and M.S. Atkins, "A Fuzzy Physiological Approach for Continuously Modeling Emotion during Interaction with Play Technologies," *Int'l J. Human Computer Studies,* vol. 65, no. 4, pp. 329-347, 2007.
[13] P.J. Lang, M.M. Bradley, and B.N. Cuthbert, "Int. Affective Picture System (IAPS): Instruction Manual and Affective Ratings," Technical Report A-5, 2001.
[14] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion Recognition Using Bio-Sensors: First Steps towards an Automatic System," *Affective Dialogue Systems,* vol. 3068, pp. 36-48, 2004.
[15] L. Sornmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications.* Elsevier, 2005.
[16] J.T. Catalano, *Guide to ECG Analysis.* J.B. Lippincott, 1993.
[17] R. Shouldice, C. Heneghan, P. Nolan, and P.G. Nolan, "PR and PP ECG Intervals as Indicators of Autonomic Nervous Innervation of the Cardiac Sinoatrial and Atrioventricular Nodes," *Proc. First IEEE Int'l Conf. Neural Eng.,* pp. 261-264, Mar. 2003.
[18] R. Virtanen, A. Jula, J.K. Salminen, L.M. Voipio-Pulkki, H. Helenius, T. Kuusela, and J. Airaksinen, "Anxiety and Hostility Are Associated with Reduced Baroreflex Sensitivity and Increased Beat-to-Beat Blood Pressure Variability," *Psychosomatic Medicine,* vol. 65, pp. 751-756, 2003.
[19] S. Booth-Kewley and H.S. Friedman, "Psychological Predictors of Heart Disease: A Quantitative Review," *Psychological Bull.,* vol. 101, no. 3, pp. 343-362, 1987.
[20] G.H.E. Gendolla, M. Richter, and A. Friedrich, "Task Difficulty Effects on Cardiac Activity," *Psychophysiology,* no. 45, pp. 869-875, 2008.
[21] J. Blascovich, M.D. Seery, C.A. Mugridge, R.K. Norris, and M. Weisbuch, "Predicting Athletic Performance from Cardiovascular Indexes of Challenge and Threat," *J. Experimental Social Psychology,* vol. 40, no. 5, pp. 683-688, 2004.
[22] H. Ue, I. Masuda, Y. Yoshitake, Y. Inazumi, and T. Moritani, "Assessment of Cardiac Autonomic Nervous Activities by Means of ECG R-R Interval Power Spectral Analysis and Cardiac Depolarization-Repolarization Process," *Annals of Noninvasive Electrocardiology,* vol. 5, no. 4, pp. 336-345, 2000.
[23] W.L. Wasmund, E.C. Westerholm, D.E. Watenpaugh, S.L. Wasmund, and M.L. Smith, "Interactive Effects of Mental and Physical Stress on Cardiovascular Control," *J. Applied Physiology,* vol. 92, pp. 1828-1834, 2002.
[24] R. Pramila, J. Sims, R. Brackin, and N. Sarkar, "Online Stress Detection Using Psychophysiological Signals for Implicit Human-Robot Cooperation," *Robotica,* vol. 20, no. 6, pp. 673-685, 2002.
[25] M. Dambacher, W. Eichinger, K. Theisen, and A.W. Frey, "RT and Systolic Blood Pressure Variability after Sympathetic Stimulation during Positive Tilt in Healthy Volunteers," *Proc. Computers in Cardiology,* pp. 573-576, Sept. 1994.
[26] G. Andrassy, A. Szabo, G. Ferencz, Z. Trummer, E. Simon, and A. Tahy, "Mental Stress May Induce QT-Interval Prolongation and T-Wave Notching," *Annals of Noninvasive Electrocardiology,* vol. 12, no. 3, pp. 251-259, 2007.
[27] A.F. Folino, G. Buja, P. Turrini, L. Oselladore, and A. Nava, "The Effects of Sympathetic Stimulation Induced by Mental Stress on Signal Averages Electrocardiogram," *Int'l J. Cardiology,* vol. 48, pp. 279-285, 1995.
[28] A. Szabo, "The Combined Effects of Orthostatic and Mental Stress on Heart Rate, T-Wave Amplitude, and Pulse Transit Time," *European J. Applied Physiology,* vol. 67, no. 6, pp. 540-544, 1993.
[29] H. Scher, J.J. Furedy, and R.J. Heslegrave, "Phasic T-Wave Amplitude and Heart Rate Changes as Indices of Mental Effort and Task Incentive," *Psychophysiology,* vol. 21, no. 3, pp. 326-333, 1984.
[30] F. Agrafioti, F.M. Bui, and D. Hatzinakos, "Medical Biometrics: The Perils of Ignoring Time Dependency," *Proc. IEEE Third Int'l Conf. Biometrics: Theory, Applications, and Systems,* pp. 1-6, Sept. 2009.
[31] S.A. Israel, J.M. Irvine, A. Cheng, M.D. Wiederhold, and B.K. Wiederhold, "ECG to Identify Individuals," *Pattern Recognition,* vol. 38, no. 1, pp. 133-142, 2005.
[32] R. Hoekema, G. Uijen, and A. van Oosterom, "Geometrical Aspect of the Interindividual Variaility of Multilead ECG Recordings," *IEEE Trans. Biomedical Eng.,* vol. 48, no. 5, pp. 551-559, May 2001.
[33] L.S. Green, R.L. Lux, C.W. Hawsand, R. Williams, S. Hunt, and M. Burgess, "Effects of Age, Sex, and Body Habitus on QRS and ST-T Potential Maps of 1100 Normal Subjects," *Circulation,* vol. 85, pp. 244-253, 1985.
[34] G. Kozmann, R.L. Lux, and L.S. Green, "Sources of Variability in Normal Body Surface Potential Maps," *Circulation,* vol. 17, pp. 1077-1083, 1989.
[35] B.P. Simon and C. Eswaran, "An ECG Classifier Designed Using Modified Decision Based Neural Network," *Computers and Biomedical Research,* vol. 30, pp. 257-272, 1997.
[36] C. Zong and M. Chetouani, "Hilbert-Huang Transform Based Physiological Signals Analysis for Emotion Recognition," *Proc. IEEE Int'l Symp. Signal Processing and Information Technology,* pp. 334-339, Dec. 2009.
[37] N.E. Huang, Z. Shen, R.R. Long, M.L. Wu, Q. Zheng, N.C. Yen, and C.C. Tung, "The Empirical Mode Decomposition and Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis," *Proc. Royal Soc. London,* vol. 454, pp. 903-995, 1998.
[38] G. Rilling, P. Flandrin, P. Gonalves, and J.M. Lilly, "Bivariate Empirical Mode Decomposition," *IEEE Signal Processing Letters,* vol. 14, no. 12, pp. 936-939, Dec. 2007.
[39] J. Pan and W.J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomedical Eng.,* vol. 32, no. 3, pp. 230-236, Mar. 1985.
[40] P.E. McSharry, G.D. Clifford, L. Tarassenko, and L.A. Smith, "A Dynamical Model for Generating Synthetic Electrocardiogram Signals," *IEEE Trans. Biomedical Eng.,* vol. 50, no. 3, pp. 289-294, Mar. 2003.
[41] M. Blanco-Velasco, B. Weng, and K.E. Barner, "ECG Signal Denoising and Baseline Wander Correction Based on the Empirical Mode Decomposition," *Computers in Biology and Medicine,* vol. 38, pp. 1-13, Jan. 2008.
[42] A. Arafat and K. Hasan, "Automatic Detection of ECG Wave Boundaries Using Empirical Mode Decomposition," *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing,* pp. 461-464, Apr. 2009.
[43] A.J. Nimunkar and W.J. Tompkins, "R-Peak Detection and Signal Averaging for Simulated Stress ECG Using EMD," *Proc. 29th Ann. IEEE Int'l Conf. Eng. in Medicine and Biology Soc.,* pp. 1261-1264, Aug. 2007.
[44] P. Flandrin, P. Goncalves, and G. Rilling, "Detrending and Denoising with Empirical Mode Decompositions," *Proc. XII EUSIPCO,* Sept. 2004.
[45] M.K.L. Molla, T. Tanaka, T.M. Rutkowski, and A. CiChocki, "Separation of EOG Artifacts from EEG Signals Using Bivatiate EMD," *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing,* pp. 562-565, Mar. 2010.
[46] R. Cowie, E. Douglas-Cowie, S. Savvidou, E. McMahon, M. Sawey, and M. Schroder, "FEELTRACE: An Instrument for Recording Perceived Emotion in Real Time," *Proc. ISCA Tutorial and Research Workshop Speech and Emotion,* pp. 19-24, 2000,
[47] H. Rau, "Responses of the T-Wave Amplitude as a Function of Active and Passive Tasks and Beta-Adrenergic Blockade," *Psychophysiology,* vol. 28, no. 2, pp. 231-239, 1991.
[48] S.M.A. Bhuiyan, R.R. Adhami, and J.F. Khan, "A Novel Approach of Fast and Adaptive Bidimensional Empirical Mode Decomposition," *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Processing,* pp. 1313-1316, Apr. 2008.
[49] S.M.A. Bhuiyan, R.R. Adhami, and J.F. Khan, "Fast and Adaptive Bidimensional Empirical Mode Decomposition Using Order-Statistics Filter Based Envelope Estimation," *EURASIP J. Advances in Signal Processing,* vol. 2008, pp. 1-18, 2008.

Foteini Agrafioti received the diploma degree in electronics and computer engineering from the Technical University of Crete, Greece, in 2006. She then received the MASc degree in electrical engineering, from the University of Toronto, Canada, in 2007. She is currently working toward the PhD degree in the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto. Her research interests include biometric systems, multimedia processing, and affective computing. She is a member of the Biometrics Security Laboratory, the IEEE, and the IEEE Signal Processing Society.

Dimitrios Hatzinakos received the diploma degree from the University of Thessaloniki, Greece, in 1983, the MASc degree from the University of Ottawa, Canada, in 1986, and the PhD degree from Northeastern University, Boston, Massachusetts, in 1990, all in electrical engineering. In September 1990, he joined the Department of Electrical and Computer Engineering, University of Toronto, where he now holds the rank of professor with tenure. He is the holder of the Bell Canada Chair in Multimedia, and the director of the Identity, Privacy, and Security Institute (IPSI) as well as the Biometrics Security Laboratory at the University of Toronto. His research interests and activities include the areas of Digital Signal Processing, wireless communications, multimedia security, and biometric systems. He is a senior member of the IEEE.

Adam K. Anderson received the BA degree from Vassar College in cognitive science in 1991, the PhD degree in experimental psychology from Yale University in 2000, and completed postdoctoral training at Stanford University in 2003. In 2003, he joined the Department of Psychology at the University of Toronto, where he is the Canada Research Chair in Cognitive and Affective Neuroscience, an associate professor, and a research associate at the Rotman Research Institute. His research focuses on the psychological neural mechanisms of emotion-cognition interactions and he is the recipient of the 2009 American Psychological Society Distinguished Early Career Contributions Award and 2010 Cognitive Neuroscience Society Young Investigator Award. He presently serves as an associate editor for the journal *Emotion*.