Automatic ECG-Based Emotion Recognition in Music Listening

Yu-Liang Hsu, *Member, IEEE*, Jeen-Shing Wang, *Member, IEEE*, Wei-Chun Chiang, and Chien-Han Hung

Abstract—This paper presents an automatic ECG-based emotion recognition algorithm for human emotion recognition. First, we adopt a musical induction method to induce participants' real emotional states and collect their ECG signals without any deliberate laboratory setting. Afterward, we develop an automatic ECG-based emotion recognition algorithm to recognize human emotions elicited by listening to music. Physiological ECG features extracted from the time-, and frequency-domain, and nonlinear analyses of ECG signals are used to find emotion-relevant features and to correlate them with emotional states. Subsequently, we develop a sequential forward floating selection-kernel-based class separability-based (SFFS-KBCS-based) feature selection algorithm and utilize the generalized discriminant analysis (GDA) to effectively select significant ECG features associated with emotions and to reduce the dimensions of the selected features, respectively. Positive/negative valence, high/low arousal, and four types of emotions (joy, tension, sadness, and peacefulness) are recognized using least squares support vector machine (LS-SVM) recognizers. The results show that the correct classification rates for positive/negative valence, high/low arousal, and four types of emotion classification tasks are 82.78, 72.91, and 61.52 percent, respectively.

Index Terms—Electrocardiogram, emotion recognition, music, machine learning

1 Introduction

With rapid advancements in technology, advanced and user-friendly human-computer interaction (HCI) systems should become capable of considering human affective states in interactions to promote mutual sympathy between humans and machines. In human communication, the expression and understanding of emotions can help to achieve mutual sympathy. To develop an emotion-based HCI, we need to equip machines with the abilities to understand and identify human emotions without any input devices to translate users' intentions. However, emotion is a psychological and physiological expression and is associated with mood, temperament, personality, disposition, and motivation, which are produced by cognitive processes, subjective feelings, motivational tendencies, behavioral reactions, and physiological arousal. Therefore, developing an effective emotion recognition system for identifying various emotions is an interesting and challenging topic.

Recently, many studies have focused on developing effective emotion recognition systems that identify implicit features of human communication, such as speech, facial expressions, gestures, or physiological measurements, in different experimental settings [1], [2], [3], [4], [5], [6], [7], [8]. However, the features of the abovementioned audiovisual emotion channels are not adequate for obtaining emotion classification results, because human can disguise their emotions by artifacts of human social masking. For example, people have a "poker face," or they may not express their emotions via intuitive human body language when they are angry. Similarly, using physiological measurements (biosignals), including electromyogram (EMG), electroencephalography (EEG), electrocardiography (ECG), galvanic skin resistance (GSR), skin temperature (ST), skin conductivity (SC), respiration (RSP), body expression, and blood oxygen saturation (OXY), for emotion classification has some limitations [4], [9], [10], [11], [12]. First, physiological patterns cannot be mapped onto specific emotional states uniquely because emotions could be influenced by time, context, space, and culture, while physiological patterns may differ from user to user and from situation to situation. Second, recorded biosignals usually include motion artifacts caused by electrodes' movement on the skin surface. Third, determination of the "ground truth" of the biosignals is an arduous task because we can only observe signal flows or trends; however, we perceive or experience emotions intuitively. On the other hand, physiological measurements offer numerous benefits for researches to develop an emotion-based HCI. First, various biosignals for users' affective states can be gathered continuously, and these can truly reflect human emotions in daily life through the autonomous nervous system (ANS). Second, physiological ANS activity could overcome possible artifacts of human social masking because that cannot be easily triggered by any conscious control, and some of that are not culturally specific. Therefore, many

- *Y.-L. Hsu is with the Department of Automatic Control Engineering, Feng Chia University, Taichung 40724, Taiwan. E-mail: hsuyl@fcu.edu.tw.*
- *J.-S. Wang, W.-C. Chiang, and C.-H. Hung are with the Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan. E-mail: jeenshin@mail.ncku.edu.tw, outlawbo@hotmail.com, chienhan.hung@gmail.com.*

Manuscript received 6 Apr. 2016; revised 28 Aug. 2017; accepted 7 Nov. 2017. Date of publication 18 Dec. 2017; date of current version 28 Feb. 2020. (Corresponding author: Jeen-Shing Wang.)
Recommended for acceptance by A. Potamianos.
Digital Object Identifier no. 10.1109/TAFFC.2017.2781732

researchers have used the fusion of the physiological measurements to recognize human emotions effectively because these reflect the involuntary reactions of the human body [4], [10], [11], [13]. However, using too many biosignals to recognize human emotions is not suitable for practical applications, and it may hinder subjects during daily life activities [14]. Therefore, it is important to develop a reliable emotion recognition system for an emotion-based HCI that uses suitable physiological channels and shows acceptable recognition abilities and robustness against any artifacts caused by human movement or human social masking.

In this paper, we use only one ECG channel to classify four types of emotions by an emotion recognition system. The novelty of this paper is the proposed automatic ECG-based emotion recognition algorithm. This algorithm consists of a sequential forward floating selection-kernel-based class separability-based (SFFS-KBCS-based) feature selection algorithm, the generalized discriminant analysis (GDA) feature reduction method, and least squares support vector machine (LS-SVM) classifiers for effective recognition of music-induced emotions by using physiological changes of only one ECG channel. For this purpose, we design an accurate experiment for collecting participants' ECG signals when listening to music and then develop an automatic emotion recognition algorithm using only ECG signals for detecting R-waves, generating significant emotion-relevant ECG features, and recognizing various human emotions effectively. After extracting ECG features from the time-, frequency-domain, and nonlinear analyses of ECG signals, we select appropriate features from a total of 34 normalized features by using the SFFS-based search strategy combined with the KBCS-based selection criterion. Subsequently, we utilize the GDA to effectively reduce the dimension of the significant ECG features associated with emotions. Finally, we use the LS-SVM classifiers to recognize arousal-valence emotion stages.

This paper is organized as follows. Section 2 presents a brief overview of related research on automatic emotion recognition systems based on ECG signals when listening to music. The experimental setup and protocol is presented in Section 3. Section 4 introduces the proposed automatic ECG-based emotion recognition algorithm. Next, Sections 5 and 6 present the experimental results and discussion, respectively. Finally, conclusions are presented in the last chapter.

2 Related Research

2.1 Dimensional Emotion Models

It is difficult to judge or model human emotions because people express their emotions differently based on such factors as their cognitive process and subjective feeling. Over the past several decades, many researchers have devoted to develop diverse emotion models for modeling human emotions [15], [16]. Among various emotion models, the discrete and affective dimensional models are two common approaches to model emotions, and they are not exclusive of each other. In the discrete models, humans must choose a prescribed list of word labels to label emotions in discrete categories for indicating their current emotion, for example, joy, tension, sadness, anger, fear, etc [10], [15]. However, the stimuli may elicit blended emotions that cannot be adequately expressed in words because the meanings of the chosen words are culturally dependent [4]. Therefore, the discrete models require more than one word to describe mixed emotions. In the affective dimensional models, humans need to scale emotions in multiple dimensions for categorizing emotions. Recently, two common scales used in emotion classification are valence and arousal [4], [14], [17], [18]. All emotions can be mapped onto the valance and arousal axes in the two-dimensional (2D) emotion plane. For example, joy has positive valence and high arousal, whereas sadness has negative valence and low arousal.

2.2 ECG Signal and Emotion

From the literature review, we found that emotion is systematically elicited by subjective feelings, physiological arousal, motivational tendencies, cognitive processes, and behavioral reactions. From the viewpoint of physiological arousal, it is difficult to find out an ANS differentiation of emotions because the ANS may be influenced easily by several factors such as attention, social interaction, appraisal, and orientation. However, recently reported studies have shown that ANS activity comprising the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) is viewed as an important component of the emotion response [19]. In addition, heart rate (HR) and heart rate variability (HRV), that are the variation over time of the period between successive heartbeats, are the common ECG features extracted from the ECG signals for emotion recognition [2], [6], [10], [13], [14], [20], [21]. Recent research studies have shown that music can actually produce specific physiological reactions of change in HR and HRV that are associated with different emotions [4], [22], [23], [24], [25], [26].

2.3 Boundary Conditions for Finding Relationship between Emotion and ECG Signals

This study focuses on the relationship between emotion and ECG signals. We summarize some possible factors that affect the emotion classification results by using ECG signals as follows:

1. Not enough time for some participants to reach a neutral state or emotional states elicited by the musical excerpts during the baseline and music listening stages.
2. The selected stimuli music does not have enough intensity to elicit emotions for the participants in the emotion classification tasks.
3. The difficulty of subject-independent classification is the intricate variety of nonemotional individual contexts among the participants, rather than an individual ECG specificity in emotion [4].
4. The participants cannot faithfully report their emotional state on the GEMS-9 questionnaire because of the unconcentrated condition during music listening or the misunderstanding of the meaning of the GEMS-9.
5. There is no one-to-one relation between emotion and physiological changes in ECG-based features: Feeling changes may occur without concomitant autonomic changes in ECG-based features and vice versa [19].

The abovementioned factors directly or indirectly affect the determinations of the ground truths for the classifiers,

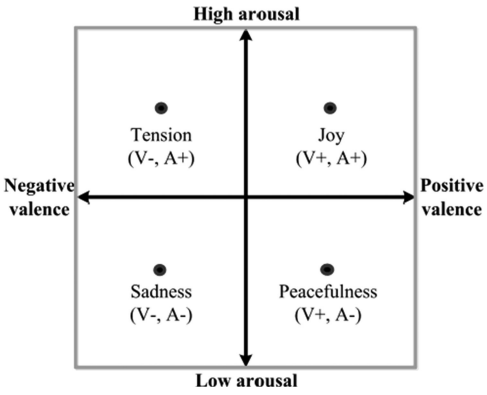


Fig. 1. Target emotions in a 2D emotion model comprising valence and arousal axes are Joy, Tension, Sadness, and Peacefulness.

and they further deteriorate the classification rates and trouble the model selection of classifiers.

3 Experimental Setup

3.1 Participants

A total of 61 healthy participants selected from a course of analysis of scientific evidence in healing music opened at National Cheng Kung University, aged between 17 and 36 years old (20.4 ± 2.4), participated in this study as partial fulfillment of the course requirement.

3.2 Materials and Setup

The experiment was performed in a laboratory environment with controlled temperature. ECG signals were recorded using a NeXus-10 instrument with Biotrace+ software on a dedicated recording PC (Intel® Core Processor i5-2400, 3.30 GHz). ECG signals (lead II) were recorded with a 2048 Hz sampling rate using three electrodes. Participants could watch their ECG signals on a 42-inch screen. Stimuli selected from the participants themselves (elective music) and from music experts (assigned music) were presented using a dedicated stimulus PC manipulated by the experimenter. The elective music was selected from participants themselves, which can truly induce their four types of emotions corresponding to the four quadrants in the 2D emotion model shown in Fig. 1 [16], [17]. The assigned music was selected from music experts according to specific musical features including tempo, mode, dynamics, timbre, rhythm, tonality, and harmonic progression [27]. For example, the songs chosen to represent joy are characterized by fast tempo, major mode, consonance, high loudness, high pitch, large pitch range, ascending melody, regular rhythm, many harmonics/bright timbre, and staccato articulation. In contrast, the songs chosen to represent sadness are characterized by slow tempo, minor mode, dissonance, soft loudness, low pitch, narrow pitch range, descending melody, firm rhythm, few harmonics/soft timbre, and legato articulation. In addition, the songs chosen to represent tension are characterized by dissonance, high loudness, rhythmic complexity, harmonic complexity, ascending melody, and increased note density. Finally, the songs chosen to represent peacefulness are characterized by slow tempo, consonance, soft loudness, narrow pitch range, regular rhythm, and legato articulation. For presenting the stimuli, Windows Media Player software and Yamaha Corporation speakers were used, and the music volume was set to the maximum volume of Windows Media Player and 75 percent loudness level of the computer speaker volume. However, each participant was asked before the experiment whether the volume was comfortable.

3.3 Experimental Protocol

Each participant signed a consent form and filled out a questionnaire before participating in this experiment. Next, the electrodes of the ECG recording instrument were placed on the participants and the recorded ECG signals should be checked by the experimenter to avoid interference resulting from incorrect placement of the electrodes or a destructive ECG recording instrument. Then, the participants were given instructions to comprehend the experimental protocol and the meanings of the scale used for self-assessment. When the instructions were clear to the participants, the experiment was started. The experiment involved the following six stages:

1. Baseline stage (5 min): The experimenter instructed the participants to press the "Start Recording" button on the screen to start ECG signal recording. After ECG signal recording, the experimenter played the music of "Whale Music" to let each participant's emotion reach a neutral state (near the origin of the 2D emotion model in Fig. 1) during this stage.
2. Elective music listening stage (15 min): At the beginning of this music listening stage, the participants pressed the enter key on the keyboard and entered "1" to make a marker for setting a start point of the ECG signal recording in this music listening stage, and then the experimenter played the elective music that expresses only one target emotion. After listening to the elective music, the participants pressed the enter key on the keyboard and entered "2" to make a marker for setting an end point of the ECG signal recording in this music listening stage.
3. First self-assessment stage (3 min): The participants should fill out the self-assessment within 3 min to reflect their current emotion after listening to the elective music.
4. Assigned music listening stage (15 min): At the beginning of this music listening stage, the participants pressed the enter key on the keyboard and entered "3" to make a marker for setting a start point of the ECG signal recording in this music listening stage, and then the experimenter played the assigned music with the same emotion as the elective music. After listening to the assigned music, the participants pressed the enter key on the keyboard and entered "4" to make a marker for setting an end point of the ECG signal recording in this music listening stage.
5. Second self-assessment stage (3 min): The participants should fill out the self-assessment within 3 min to reflect their current emotion after listening to the assigned music.
6. Recovery stage (5 min): The participants pressed the enter key on the keyboard and entered "5" to make a

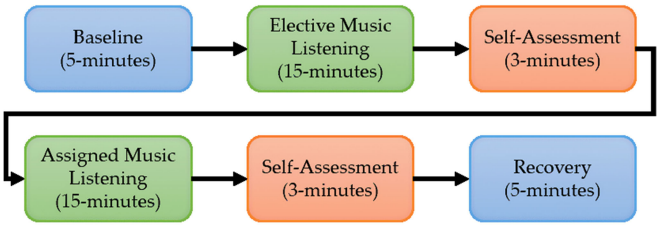


Fig. 2. Block diagram of experimental procedures.

marker for setting a start point of the ECG signal recording in this stage, and the experimenter played the music of "Yellow Fantasy" used to let each participant recover their emotion to a neutral state during this stage.

The abovementioned markers were used to separate the ECG signal recording within each experimental stage in this experiment. Fig. 2 shows a block diagram of the experimental procedures.

3.4 Participant Self-Assessment

After listening to the music in the two music listening stages, the participants performed self-assessments to reflect their current emotion after listening to the elective and assigned music, respectively. In the self-assessment stages, the participants were asked to indicate how strongly they had experienced the corresponding feeling during stimulus presentation for each of the nine Geneva Emotional Music Scale (GEMS-9) emotion categories. Each of the nine emotion labels (Wonder, Transcendence, Power, Tenderness, Nostalgia, Peacefulness, Joy, Sadness, and Tension) was scaled from 1 to 5 (1 = *not at all*, 5 = *very much*).

3.5 Dataset for Evaluation

The aims of the proposed automatic ECG-based emotion recognition algorithm are to discriminate the positive/negative valence, high/low arousal, and four target emotions (including joy, tension, sadness, and peacefulness). Toward this end, the participants' ratings labeled in the GEMS-9 during the experiment are used as the ground truths (the ideal outputs for the LS-SVM classifiers) for the classification tasks, and we set two criteria to identify the strongly elicited emotion patterns in music listening. These are described as follows.

1. We assume that the last 5-min ECG segment in the music listening stages could reflect the feeling of the participants' ratings.
2. Regarding the GEMS-9, we only focus on the last four emotion ratings, such as Peacefulness, Joy, Sadness, and Tension, because this study aims to classify these four emotions and their categories on the valence-arousal space shown in Fig. 1. According to the ratings, we selected the last 5-min ECG segments of these participants from the self-collected database when the rating score of only one of the four emotions was four or five and this emotion is the same as that of the elicited music.

As mentioned earlier, a total of 395 strongly elicited emotion samples (1-min ECG segment) distributed in the four emotions (including 105 for Joy, 55 for Tension, 40 for Sadness, and 195 for Peacefulness) were collected from 61 participants. The ECG segments distributed in the positive/negative valence are 300 and 95, respectively. The ECG segments distributed in the high/low arousal are 160 and 235, respectively. Fig. 3 shows the ECG signals collected from different participants for four emotions. The ECG database used in this paper is available and accessible through a web-interface.[1]

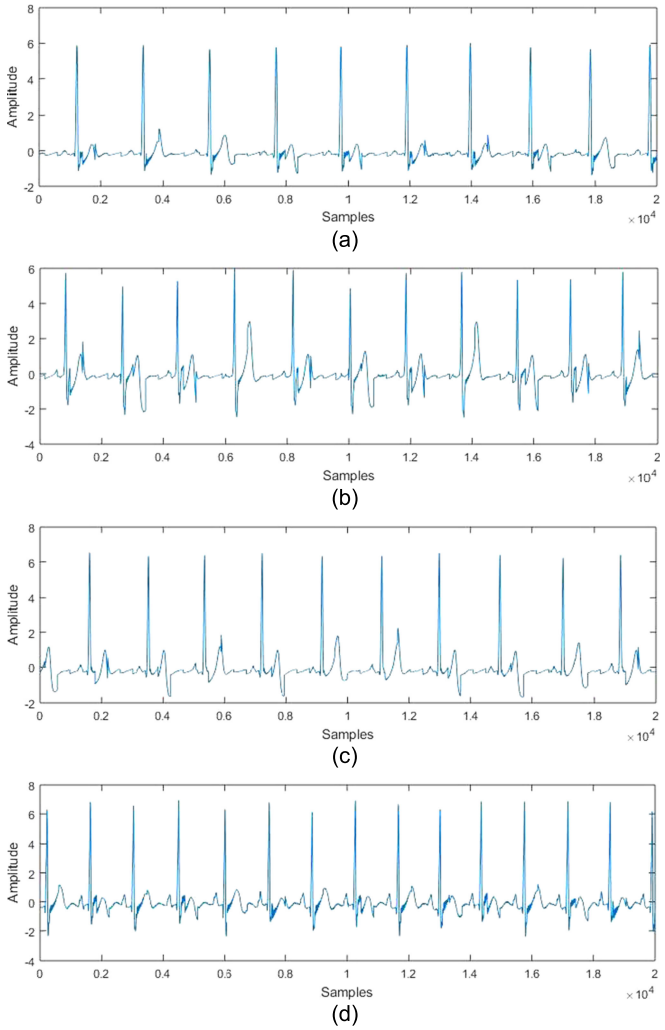


Fig. 3. ECG signals collected from different participants for four emotions. (a) Joy (b) Tension (c) Sadness (d) Peacefulness.

4 Methodology

Fig. 4 shows the block diagram of the proposed automatic ECG-based emotion recognition algorithm. We now briefly describe each procedure of the proposed automatic ECG-based emotion recognition algorithm as follows: (1) Signal preprocessing: Remove baseline wander and reduce signal amplitude biases by using the median filters and Z-score normalization method, respectively. (2) R-wave detection: Use the R-waves detected by the QRS detection algorithm proposed by Pan and Tompkins [28] to derive the R-R time intervals (RR intervals) for further extracting ECG-based features. (3) Windowing: Window the ECG recording to classify emotions based on each 1-min epoch. (4) Noisy epoch rejection: Reject incorrect epochs that include body

1. http://nemedataset.ddns.net/index/

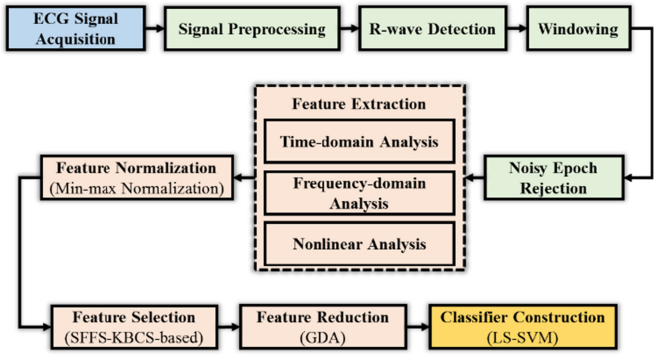


Fig. 4. Block diagram of automatic ECG-based emotion recognition algorithm.

artifact noise and incomplete ECG recording. (5) Feature extraction: Obtain a total of 34 features extracted from the time-, frequency-domain, and nonlinear analyses of ECG signals. (6) Feature normalization: First, obtain the degrees of the physiological changes of the features between the conditions of the baseline stage and the music listening stage. Subsequently, reduce the effect of the value's range from the degrees of the physiological changes of the features through the min-max normalization method. (7) Feature selection: Select the appropriate features out of a total of 34 normalized features by using the proposed SFFS-KBCS-based algorithm. (8) Feature reduction: Reduce the dimensions of the significant features from the selected features using the GDA. (9) Classifier construction: Discriminate emotions by using the LS-SVMs with one-against-one strategy based on the reduced features. Now, we introduce each procedure of the proposed automatic ECG-based emotion recognition algorithm in detail in the following sections.

4.1 Signal Preprocessing

In the ECG signal preprocessing procedure, two interference signals are removed for the following reasons: (1) Baseline wander caused by the subject's movement or breathing might mislead the ECG annotation from the accurate identification of the ECG features [29]. (2) ECG signal amplitude biases generated by individuals and instrumental differences cause extreme amplitude scaling. The baseline wander removal and Z-score normalization procedures are used to remove the abovementioned interference signals, respectively. (1) Baseline wander removal: According to [30], a median filter of 200-ms width is used to remove the QRS complexes and P-waves of the original ECG signals. Subsequently, a median filter of 600-ms width is used to remove the T-waves of the abovementioned resulting signal. Once we obtain the signals resulting from the second filter operation, the baseline wander signal of the original ECG signals can be obtained. Then, the baseline wander signal is subtracted from the original ECG signals. (2) Z-score normalization method: Because of instrumental and human differences, the signal amplitude biases of the waveforms of the ECG signals are inconsistent. Therefore, the abovementioned differences of each waveform of each ECG signal are reduced by the Z-score normalization method [31]. This procedure results in a normalized ECG signal with zero mean and unity standard deviation.

4.2 R-Wave Detection

The RR intervals acquired from the normalized ECG signals are crucial indices for extracting the ECG based features. The more accurate the R-waves detection are, the more accurate ECG-based features we can obtain. In this paper, the R-waves of the normalized ECG signals are detected based on the QRS detection algorithm proposed by Pan and Tompkins [28] to derive the R-R time intervals (RR intervals), and then the amplitude and occurrence time of the R-waves are obtained. Subsequently, the RR intervals are obtained by calculating the time difference between successive R-waves detected in the normalized ECG signals.

4.3 Windowing

According to [19], the most widely used duration of physiological variables is 1 min; therefore, we use 1 min as the window size for classifying the emotions during music listening. Hence, in this paper, the ECG signals and its R-waves are segmented into a series of successive 1-min windows.

4.4 Incorrect Epoch Rejection

Before extracting features from each ECG epoch (1-min window), we must ensure that the signal of each ECG epoch is a complete observation. The two types of incorrect epoch rejection conditions include: (1) Incomplete/noisy ECG data caused by electrodes not well attached to subject's skin. (2) Sudden change in the RR intervals caused by the ECG data with the electrode motion artifact noise or body movement noise that cannot be filtered. As mentioned earlier, we should reject incorrect epochs by the conditions of no change in the ECG data and sudden change in the RR intervals. Therefore, the incorrect epochs are excluded from the subsequent procedures.

4.5 Feature Extraction

In the feature extraction procedure, a total of 34 features listed in Table 1 can be extracted from the time-, frequency-domain, and nonlinear analyses of ECG signals in each 1-min epoch [20], [32], [33].

4.5.1 Time-Domain Analysis

In the time-domain analysis, we calculated a total of 12 features. (1) HRV related parameters: The standard deviation of RR intervals (*SDNN*), root mean square of differences between adjacent RR intervals (*RMSSD*), number of successive RR intervals that differ more than 50 ms (*NN50*), percentage of successive RR intervals that differ more than 50 ms (*pNN50*), and standard deviation of differences between adjacent RR intervals (*SDSD*). The relative equations for calculating the time-domain HRV related parameters are shown as follows.

$$SDNN = \sqrt{\sum_{i=1}^{n} (R_i - \bar{R})^2 / n}, \tag{1}$$

$$RMSSD = \sqrt{\sum_{i=2}^{n} (R_i - R_{i-1})^2 / (n-1)}, \tag{2}$$

$$SDSD = \sqrt{\sum_{i=2}^{n} (R_i - R_{i-1} - RMSSD^2)^2 / (n-1)}, \tag{3}$$

TABLE 1
ECG Feature Sets Generated in This Study

Time-domain analysis	HRV	SDNN, RMSSD, NN50, pNN50, SDSD
	HR	BPM
	RR interval	Median-RRI, IQR-RRI, MAD-RRI, Diff-RRI, CV-RRI, Range
Frequency-domain analysis	HRV	VLF, LF, HL, TP, LF/HF, LF_{norm}, HF_{norm}, pVLF, pLF, pHF, VLF_{fr}, LF_{fr}, HF_{fr}
Nonlinear analysis	EDR	RSP_{rate}, Coherence
	Poincaré plot	SD_1, SD_2, SD_{12}
	Nonlinear dynamics	ApEn, SampleEn
	Autocorrelation	ACF_{coef}, ACF_{freq}

where R_i is the ith RR interval, $\bar{R}$ is the average of the RR intervals, and n is the number of the RR intervals. (2) HR related parameter: The number of R-waves within one epoch divided by 1 min (*BPM*). (3) RR interval related parameters: The median value of RR intervals (*Median-RRI*), interquartile range of RR intervals (*IQR-RRI*), mean absolute deviation of RR intervals (*MAD-RRI*), mean of the difference between adjacent RR intervals (*Diff-RRI*), coefficient of variation of RR intervals (*CV-RRI*), and difference between the maximum and the minimum RR interval (*Range*). The relative equations for calculating the RR interval related parameters are shown as follows.

$$MAD\_RRI = \frac{\mathbf{1}}{n}\sum_{i=\mathbf{1}}^{n}\left|R_i - \bar{R}\right|, \tag{4}$$

$$Diff\_RRI = \frac{\mathbf{1}}{n-\mathbf{1}}\sum_{i=\mathbf{2}}^{n}\left|R_i - R_{i-\mathbf{1}}\right|, \tag{5}$$

$$CV\_RRI = \frac{\mathbf{1}}{\bar{R}}\sqrt{\sum_{i=\mathbf{1}}^{n}\left(R_i - \bar{R}\right)^{\mathbf{2}}/n}. \tag{6}$$

4.5.2 Frequency-Domain Analysis

In the frequency-domain analysis, a total of 13 HRV related parameters are calculated at certain frequency bands. The RR interval signal need to be resampled and interpolated to transform them into a series of regularly resampled signals and to prevent the generation of additional harmonic components [34]. After resampling and interpolating, the power spectral density (PSD) of the resampled RR intervals is calculated by using the fast Fourier transform (FFT)-based method. The PSD analysis could be used to calculate the power of specific frequency ranges and the peak frequencies for three different frequency bands: very-low-frequency range (VLF) (0.0033–0.04 Hz), low-frequency range (LF) (0.04–0.15 Hz), and high-frequency range (HF) (0.15–0.4 Hz). Concerning the frequency-domain analysis, we calculated the following features: power calculated in the VLF, LF, and HF bands (*VLF*, *LF*, and *HF*), total power in the full frequency range (*TP*), ratio of power calculated within the LF band to that calculated within the HF band (*LF*/*HF*), LF power normalized to the sum of the LF and HF power (*LFnorm*), HF power normalized to the sum of the LF and HF power (*HFnorm*), VLF power expressed as percentage of the total power (*pVLF*), LF power expressed as percentage of the total power (*pLF*), HF power expressed as percentage of the total power (*pHF*), frequency of the highest peak in the VLF band (VLF_{fr}), frequency of the highest peak in the LF band (LF_{fr}), and frequency of the highest peak in the HF band (HF_{fr}). The relative equations for calculating the frequency-domain HRV related parameters are summarized as follows.

$$TP = VLF + LF + HF. \tag{7}$$

$$LF_{norm} = LF/(TP - VLF) = LF/(LF + HF). \tag{8}$$

$$HF_{norm} = HF/(TP - VLF) = HF/(LF + HF). \tag{9}$$

$$pVLF = (VLF/TP) \times \mathbf{100}. \tag{10}$$

$$pLF = (LF/TP) \times \mathbf{100}. \tag{11}$$

$$pHF = (HF/TP) \times \mathbf{100}. \tag{12}$$

4.5.3 Nonlinear Analysis

In the nonlinear analysis, a total of 9 features, including 2 ECG-derived respiration (EDR) related parameters, 3 Poincaré plot related parameters, 2 nonlinear dynamics related parameters, and 2 autocorrelation related parameters, are described as follows:

A. EDR related parameters. Before we extract two EDR-related features (*RSPrate* and *Coherence*), we should first obtain the EDR signal. The detailed procedure for extracting the EDR signal can be referred from [35], [36]. Subsequently, we extract two features: respiratory rate (*RSPrate*) and coherence between final EDR signal and RR intervals (*Coherence*). To estimate the respiratory rate (*RSPrate*), the final EDR signal is first subtracted by its mean to remove the DC component. Then, the PSD analysis of the EDR signal is applied to obtain the respiratory rate from the frequency of the maximum peak in the low-frequency band (0.1–1 Hz) multiplied by 60 [35], [36]. In addition, according to [37], the PSD of the EDR signal in the low-frequency band, 0–0.4 Hz, is similar to the PSD of the RR intervals when humans have positive-valence emotion. Therefore, we can calculate the coherence between the final EDR signal and the RR intervals in the low-frequency band (0–0.4 Hz), which is a function of the auto spectral density of the final EDR signal (G_{xx}) and the RR intervals (G_{yy}), and the cross spectral density (G_{xy}) of the EDR signal and the RR intervals ($Coherence = |G_{xy}|^2/(G_{xx} \times G_{yy})$).

B. Poincaré plot related parameters. The Poincaré plot analysis measures the quantitative beat-to-beat correlation between adjacent RR intervals. According to the ellipse fitting process introduced in [38], [39], SD_1 represents the standard deviation of the instantaneous beat-to-beat RR interval variability, SD_2 represents the standard deviation of the continuous long-term beat-to-beat RR interval variability, and SD_{12} is the ratio of SD_1 to SD_2.

C. Nonlinear dynamics related parameters. In the nonlinear dynamics analysis, we extract two features *ApEn* and *SampleEn* by calculating the approximate entropy and sample entropy, respectively. The approximate entropy and sample entropy are similar methods that quantify the randomness or predictability of RR interval dynamics. They are scale-invariant and model-independent; however, there is some

computational difference. They all assign a nonnegative number to a series of RR intervals that have larger values with more complexity or irregularity in the data [32], [40].

1) *ApEn*: A parameter quantifies the amount of regularity or predictability of the RR intervals. There are two user-specified parameters in the *ApEn* measure: a run length m and a tolerance window r (m and r used in this study are 1 and 0.25 times the standard deviation of the RR intervals). Initially, given N data points from a time series of RR intervals {$rr(1)$, $rr(2)$, ... , $rr(N)$}. *ApEn* is computed according to:

Step 1: Obtain a sequence of vectors $\mathbf{rr}(i) = [rr(i), rr(i+1), \ldots, rr(i+m-1)]$, $i = 1, 2, \ldots, N-m+1$. These vectors represent m successive rr values, starting with the ith point.

Step 2: Calculate the maximum absolute difference between the respective scalar components of $rr(i)$ and $rr(j)$ to obtain the distance between $\mathbf{rr}(i)$ and $\mathbf{rr}(j)$.

$$d[\mathbf{rr}(i), \mathbf{rr}(j)] = \max_{k=0,1,\ldots,m-1}(|rr(i+k) - rr(j+k)|). \quad (13)$$

Step 3: Calculate the distance between a given vector $\mathbf{rr}(i)$ and the other vectors $\mathbf{rr}(j)$ for $j = 1, 2, \ldots, N-m+1$. If $d[\mathbf{rr}(i), \mathbf{rr}(j)] \leq r$ is true, then set $N^m(i) = N^m(i) + 1$, where $i = 1, 2, \ldots, N-m+1$. Then, within tolerance r, we can obtain the $C_r^m(i)$ values that measure the regularity of patterns similar to a given segment of length m for $i = 1, 2, \ldots, N-m+1$.

$$C_r^m(i) = N^m(i)/(N - m + \mathbf{1}). \quad (14)$$

Step 4: Obtain the average frequency of all m-point patterns in the sequence close to each other, $\phi^m(r)$, by computing the average of the natural logarithm of each $C_r^m(i)$.

$$\phi^m(r) = \sum_{i=1}^{N-m+1} \ln C_r^m(i)/(N - m + \mathbf{1}). \quad (15)$$

Step 5: *ApEn* is obtained as follows:

$$ApEn(m, r) = \phi^m(r) - \phi^{m+1}(r). \quad (16)$$

2) *SampleEn:* A parameter quantifies the amount of regularity or predictability of the RR intervals. The run length m and tolerance window r are set the same as *ApEn* in this study. Initially, given N data points from a time series of RR intervals {$rr(1)$, $rr(2)$, ... , $rr(N)$}. *SampleEn* is calculated as follows:

Step 1 and Step 2: The two procedures are the same as those of *ApEn*.

Step 3: Calculate the distance between a given vector $\mathbf{rr}(i)$ and the other vectors $\mathbf{rr}(j)$ for $j = 1, 2, \ldots, N-m+1$, where $j \neq i$. If $d[\mathbf{rr}(i), \mathbf{rr}(j)] \leq r$ is true, then set $N'^m(i) = N'^m(i) + 1$, where $i = 1, 2, \ldots, N-m+1$. Then, within a tolerance r, we can obtain the $C_r'^m(i)$ values that measure the regularity of patterns similar to a given segment of length m for $i = 1, 2, \ldots, N-m+1$.

$$C_r'^m(i) = N'^m(i)/(N - m + 1). \quad (17)$$

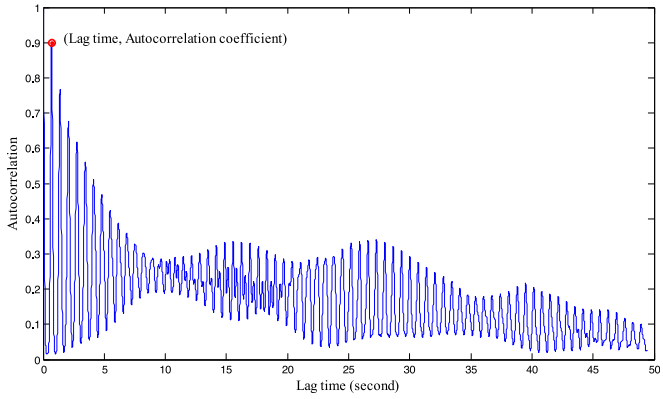


Fig. 5. Autocorrelation of QRS complexes. (*ACF_{coef} is 0.8982 and its lag time is 0.69 s, then ACF_{freq} is the reciprocal of the lag time, 1.4492* Hz.)

Step 4: Obtain the average frequency of all m-point patterns in the sequence close to each other, $\phi'^m(r)$, by computing the average of each $C_r'^m(i)$.

$$\phi'^m(r) = \sum_{i=1}^{N-m+1} C'^m{}_r(i)/(N - m + \mathbf{1}). \quad (18)$$

Step 5: *SampleEn* is obtained as follows:

$$\mathrm{SampleEn}(m, r) = -\ln\left[\phi'^m(r) - \phi'^{m+1}(r)\right]. \quad (19)$$

D. Autocorrelation related parameters. Nonlinear features such as the maximum autocorrelation coefficient (ACF_{coef}) and reciprocal of the lag time (ACF_{freq}) is obtained from the autocorrelation of the QRS complexes, which is used to quantify the similarity of the QRS complexes corresponding to the time, as shown in Fig. 5.

4.6 Feature Normalization

To obtain the degrees of physiological changes in the ECG features between the conditions of the baseline stage and the music listening stage and to reduce the effect of the value's range from the degrees of the physiological changes of the ECG features, we first calculate the degrees of physiological changes in each subject for each extracted feature as follows:

$$u_{ij} = (x_{ij} - b_{ij})/b_{ij}, \quad (20)$$

where x_{ij} represents the ith feature extracted from the jth subject during the music listening stage, b_{ij} represents the ith feature extracted from the jth subject during the baseline stage, and u_{ij} is the degree of physiological change in the feature. Then, the 34 degrees of the physiological changes of the ECG features obtained in (20) are mapped in the new range $[-1, 1]$ by the min-max normalization method.

4.7 Feature Selection

From the feature extraction and feature normalization procedures, the normalized feature vectors representing the degrees of physiological changes between the conditions of the baseline stage and the music listening stage are high-dimensional vectors. Feature selection or feature reduction is the most important procedure for any ECG-related analysis, and it must be applied to reduce the dimensionality of normalzied feature vectors for constructing effective

classifiers. In other words, feature selection or feature reduction can reduce the computational complexity and increase the classification accuracy.

In this paper, the proposed feature selection method combines a selection criterion with a search strategy for further selecting the appropriate features out of a total of 34 normalized features. This study develops an SFFS-KBCS-based feature selection algorithm that uses the SFFS method as the search strategy and the KBCS method as the selection criterion. Herein, we describe these methods as follows.

4.7.1 KBCS-Based Selection Criterion

The KBCS method utilized in this paper was originally developed by Wang [41]. Let $(\mathbf{x},\ y) \in (\mathbb{R}^d \times Y)$ represent a sample, where $\mathbb{R}^d$ stands for a d-dimensional feature space, Y denotes the set of class labels, and the size of Y is the number of classes c. This method projects the samples onto a kernel space $\mathcal{K}$, $\mathbf{m}_i^{\phi}$ is the mean vector for the ith class in the kernel space $\mathcal{K}$, n_i is the number of samples in the ith class, $\mathbf{m}^{\phi}$ is the mean vector for all classes in the kernel space $\mathcal{K}$, $\mathbf{S}_B^{\phi}$ denotes the between-class scatter matrix in the kernel space $\mathcal{K}$, $\mathbf{S}_W^{\phi}$ denotes the within-class scatter matrix in the kernel space $\mathcal{K}$, and $\mathbf{S}_T^{\phi}$ denotes the total scatter matrix in the kernel space $\mathcal{K}$. Let $\phi(\cdot)$ be a possibly nonlinear feature mapping from the feature space $\mathbb{R}^d$ to a kernel space $\mathcal{K}$: ϕ: $\mathbb{R}^d \rightarrow \mathcal{K}$, $\mathbf{x} \rightarrow \phi(\mathbf{x})$, $\mathbf{K}$ denotes a kernel matrix with $\{\mathbf{K}\}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, where $k(\mathbf{x}_i, \mathbf{x}_j)$ is a kernel function, and $\mathbf{K}_{A,B}$ is a kernel matrix with the constraints of $\mathbf{x}_i \in \mathcal{A}$ and $\mathbf{x}_j \in \mathcal{B}$. The operator $\mathrm{Sum}(\cdot)$ denotes the summation of all elements in a matrix, and trace($\mathbf{A}$) is the trace of a square matrix $\mathbf{A}$. The followings are the relative equations:

$$\begin{aligned}\mathbf{trace}(\mathbf{S}_B^{\phi}) &= \mathbf{trace}\Big[\textstyle\sum_{i=\mathbf{1}}^{c} n_i(\mathbf{m}_i^{\phi} - \mathbf{m}^{\phi})(\mathbf{m}_i^{\phi} - \mathbf{m}^{\phi})^T\Big] \\ &= \textstyle\sum_{i=\mathbf{1}}^{c} \big[\mathbf{Sum}(\mathbf{K}_{D_i,D_i})/n_i - \mathbf{Sum}(\mathbf{K}_{D,D})/n\big],\end{aligned} \tag{21}$$

$$\begin{aligned}\mathbf{trace}(\mathbf{S}_W^{\phi}) &= \mathbf{trace}\Big[\textstyle\sum_{i=\mathbf{1}}^{c}\sum_{j=\mathbf{1}}^{n_i} (\phi(\mathbf{x}_{ij}) - \mathbf{m}_i^{\phi})(\phi(\mathbf{x}_{ij}) - \mathbf{m}_i^{\phi})^T\Big] \\ &= \mathbf{trace}(\mathbf{K}_{D,D}) - \textstyle\sum_{i=\mathbf{1}}^{c} \mathbf{Sum}(\mathbf{K}_{D_i,D_i})/n_i,\end{aligned} \tag{22}$$

$$\begin{aligned}\mathbf{trace}(\mathbf{S}_T^{\phi}) &= \mathbf{trace}(\mathbf{S}_B^{\phi}) + \mathbf{trace}(\mathbf{S}_W^{\phi}) \\ &= \mathbf{trace}(\mathbf{K}_{D,D}) - \mathbf{Sum}(\mathbf{K}_{D,D})/n,\end{aligned} \tag{23}$$

where $\mathbf{x}_{ij}$ represents the jth sample of the ith class. The class separability in the kernel space is represented as

$$J^{\phi} = \mathbf{trace}(\mathbf{S}_B^{\phi})/\mathbf{trace}(\mathbf{S}_T^{\phi}). \tag{24}$$

When using the normalized kernel or the Gaussian radial basis function (RBF) kernel (stationary kernel), the class separability can obtain the lower bound as

$$J_l^{\phi} \triangleq \mathbf{trace}(S_B^{\phi}). \tag{25}$$

In this study, we use the Gaussian RBF kernel $k(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{exp}(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2)$, where the Gaussian width σ is the kernel parameter, and J_l^{ϕ} is proposed as a criterion function for feature selection [41].

4.7.2 SFFS-Based Search Strategy

The SFFS is a well-known suboptimal feature selection method that combines sequential forward selection (SFS) and sequential backward selection (SBS) to reduce the nesting effect. The SFS is a bottom-up feature selection method, whereas the SBS is a top-down feature selection method. Suppose that we want to choose a p-dimensional feature subset from the original feature set. SFS starts from an empty feature set and sequentially adds one feature from the original feature set, thereby resulting in the best selection criterion. On the other hand, SBS starts from the original feature set and sequentially deletes one feature from the original feature set, thereby resulting in the best selection criterion. The SFFS process consists of three steps: inclusion, conditional exclusion, and continuation of conditional exclusion [42].

First, suppose $F_k = \{f_i : 1 \le i \le k\}$ is a selected feature set in which k features have already been selected from the original feature set $Y = \{y_i :\ 1 \le i \le n\}$, where n is the number of total features.

Step 1: [*Inclusion*] Select feature f_{k+1} by using the basis SFS method from the available set $\{Y - F_k\}$ to form a feature set F_{k+1}, that is, the feature f_{k+1} is the most significant feature in the available set $\{Y - F_k\}$, which is then added to F_k. Therefore, $F_{k+1} = \{F_k,\ f_{k+1}\}$.

Step 2: [*Conditional exclusion*] Find the least significant feature f_j in the feature set F_{k+1}. If f_{k+1} is the least significant feature in the feature set F_{k+1}, that is,

$$J_l^{\phi}(F_{k+\mathbf{1}} - f_{k+\mathbf{1}}) = J_l^{\phi}(F_k) \ge J_l^{\phi}(F_{k+\mathbf{1}} - f_j), \forall j = 1, 2, ..., k, \tag{26}$$

where $J_l^{\phi}(\cdot)$ denotes the feature selection criterion function obtained from the KBCS. It means that the best feature combination in F_k is $\{F_{k+1} - f_{k+1}\}$. Then, set $k = k + 1$ and return to Step 1. If $f_j \ne f_{k+1}$ is the least significant feature in the feature set F_{k+1}, that is,

$$J_l^{\phi}(F_{k+\mathbf{1}} - f_j) \ge J_l^{\phi}(F_k), \forall j = 1, 2, \ldots, k. \tag{27}$$

Then, exclude f_j from the feature set F_{k+1} to form a new feature set F'_k, that is,

$$F'_k = F_{k+\mathbf{1}} - f_j. \tag{28}$$

Note that $J_l^{\phi}(F'_k) > J_l^{\phi}(F_k)$. If $k = 2$, then feature set $F_k = F'_k$ and $J_l^{\phi}(F_k) = J_l^{\phi}(F'_k)$ and return to Step 1, else go to Step 3.

Step 3: [*Continuation of conditional exclusion*] Find the least significant feature f_s in the feature set F'_k. If $J_l^{\phi}(F'_k - f_s) \le J_l^{\phi}(F_{k-1})$, then feature set $F_k = F'_k$, $J_l^{\phi}(F_k) = J_l^{\phi}(F'_k)$ and return to Step 1. If $J_l^{\phi}(F'_k - f_s) > J_l^{\phi}(F_{k-1})$, then exclude f_s from the feature set F'_k to form a newly reduced feature set F'_{k-1}, that is,

$$F'_{k-\mathbf{1}} = F'_k - f_s, \tag{29}$$

and set $k = k - 1$. If $k = 2$, then feature set $F_k = F'_k$, $J_l^{\phi}(F_k) = J_l^{\phi}(F'_k)$ and return to Step 1, else repeat Step 3.

The SFFS-based search strategy is initialized by setting $k = 0$ and $F_0 = \emptyset$, and the SFS method is used until a feature set of cardinality 2 is obtained (it means until the two

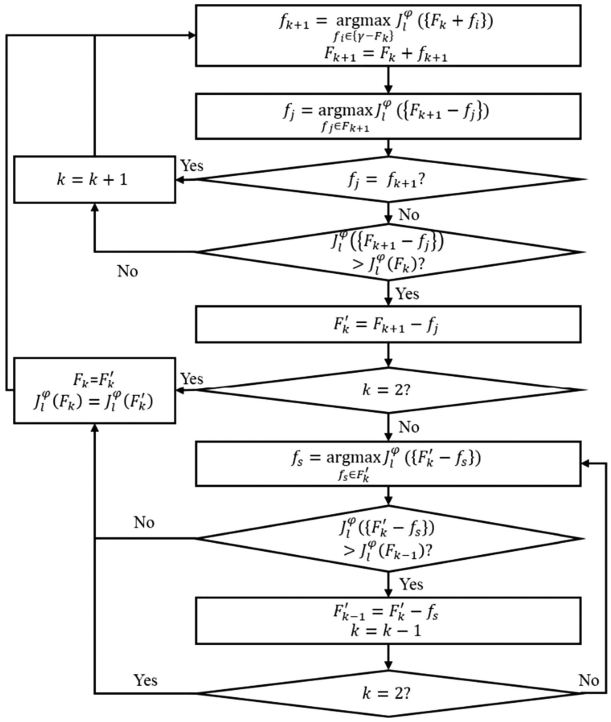


Fig. 6. Flowchart of SFFS-KBCS-based feature selection algorithm.

most significant features are included). Then, the process continues with Step 1. Finally, we combine the KBCS-based search criterion with the SFFS-based search strategy to form the proposed SFFS-KBCS-based feature selection algorithm. Fig. 6 shows the flowchart of the SFFS-KBCS-based feature selection algorithm.

4.8 Feature Reduction

To reduce the dimensions of the selected features obtained from the proposed SFFS-KBCS-based feature selection algorithm, the GDA is utilized in this study [43]. The GDA is a feature reduction approach that is used for dealing with nonlinear discriminant problems using a kernel function operator. The GDA aims to use linear properties via a possibly nonlinear feature mapping function $\phi(\cdot)$ and then find a projection matrix $\mathbf{T}$, which can maximize the ratio of the between-class scatter matrix $\mathbf{S}_B^{\phi}$ to the within-class scatter matrix $\mathbf{S}_W^{\phi}$ both in the kernel space, to map the original selected feature set $\mathbf{f}_i \in \mathbb{R}^p$ in a p-dimensional space onto another smaller feature set $\mathbf{g}_i \in \mathbb{R}^q$ in a q-dimensional space. Let $(\mathbf{f},\ u) \in (\mathbb{R}^p \times U)$ represent a sample, where $\mathbb{R}^p$ denotes a p-dimensional feature space, U denotes the set of class labels, and the size of U is the number of classes c. This method firstly projects the samples from the original selected feature space $\mathbb{R}^p$ onto a kernel space $\mathcal{K}$ via a nonlinear feature mapping function $\phi(\cdot)$: ϕ: $\mathbb{R}^p \rightarrow \mathcal{K}$, $\mathbf{f} \rightarrow \phi(\mathbf{f})$. We summarize the GDA method as follows:

Step 1: Compute both a within-class scatter matrix $\mathbf{S}_W^{\phi}$ and a between-class scatter matrix $\mathbf{S}_B^{\phi}$ in the kernel space by the following equations:

$$\mathbf{S}_W^{\phi} = \sum_{i=1}^{c} \sum_{j=1}^{n_i} (\phi(\mathbf{f}_{ij}) - \mathbf{m}_i^{\phi})(\phi(\mathbf{f}_{ij}) - \mathbf{m}_i^{\phi})^T, \tag{30}$$

$$\mathbf{S}_B^{\phi} = \sum_{i=1}^{c} n_i (\mathbf{m}_i^{\phi} - \mathbf{m}^{\phi})(\mathbf{m}_i^{\phi} - \mathbf{m}^{\phi})^T, \tag{31}$$

where $\mathbf{f}_{ij}$ is the jth sample of the ith class, $\mathbf{m}_i^{\phi}$ is the mean vector for the ith class in the kernal space $\mathcal{K}$, n_i is the number of samples in the ith class, and $\mathbf{m}^{\phi}$ is the mean vector for all classes in the kernel space $\mathcal{K}$.

Step 2: Find transformation matrix $\mathbf{T} \in \mathcal{K}$ by maximizing the following equation in the kernel space $\mathcal{K}$:

$$J(\mathbf{T}) = \left|\mathbf{T}^{\mathrm{T}} \mathbf{S}_B^{\phi} \mathbf{T}\right| / \left|\mathbf{T}^{\mathrm{T}} \mathbf{S}_W^{\phi} \mathbf{T}\right|, \tag{32}$$

where $\mathbf{T}^{\mathrm{T}}\mathbf{S}_B^{\phi}\mathbf{T}$ and $\mathbf{T}^{\mathrm{T}}\mathbf{S}_W^{\phi}\mathbf{T}$ are the new between-class scatter matrix and new within-class scatter matrix in the new feature space, respectively. In general, $\mathbf{T} = [\mathbf{t}_1,\ \mathbf{t}_2,\ \ldots, \mathbf{t}_q]$ is obtained by applying an eigenvalue (λ) decomposition of $\mathbf{S}_W^{\phi-1}\mathbf{S}_B^{\phi}$ via the following equation when the inverse of $\mathbf{S}_W^{\phi}$ exists

$$\mathbf{S}_B^{\phi}\mathbf{T} = \lambda \mathbf{S}_W^{\phi}\mathbf{T}. \tag{33}$$

Explicitly computing the mapping functions $\phi(\mathbf{f}_i)$ and then performing the process is computationally complex. Thus, the feature data is implicitly embedded via dot products and by using a kernel trick in which a kernel function $k(\mathbf{f}_i, \mathbf{f}_j) = \langle \phi(\mathbf{f}_i), \phi(\mathbf{f}_j) \rangle$ is used to replace the dot product in the new feature space. In this study, we use the Gaussian RBF kernel function $k(\mathbf{f}_i, \mathbf{f}_j) = \mathbf{exp}(-\|\mathbf{f}_i - \mathbf{f}_j\|^2 / \mathbf{2\sigma^2})$, where the Gaussian width σ is the kernel parameter. Subsequently, $\mathbf{t}_k$ is calculated using (34), which is a linear combination of all samples in $\mathcal{K}$.

$$\mathbf{t}_k = \sum_{i=1}^{c} \sum_{j=1}^{n_i} \alpha_{ij}\ \phi(\mathbf{f}_{ij}), \tag{34}$$

where α_{ij} are real coefficients corresponding to $\phi(\mathbf{f}_{ij})$, and they are obtained by solving

$$\lambda = \alpha^{\mathrm{T}}\mathbf{KDK}\alpha / (\alpha^{\mathrm{T}}\mathbf{KK}\alpha), \tag{35}$$

where $\alpha^{\mathrm{T}} = [\alpha_1^{\mathrm{T}},\ \alpha_2^{\mathrm{T}},\ \ldots,\ \alpha_c^{\mathrm{T}}]$ is a vector of coefficients with $\alpha_i^{\mathrm{T}} = [\alpha_{i1},\ \alpha_{i2},\ \ldots,\ \alpha_{in_i}]$, and $\mathbf{K}$ is an N by N symmetric kernel matrix defined on the class elements as follows:

$$\mathbf{K} = (\mathbf{K}_{xy})_{x=1,\ldots c; y=1,\ldots,c}, \tag{36}$$

where N is the number of all training samples, $(\mathbf{K}_{xy}) = (k(\mathbf{f}_{xi}, \mathbf{f}_{yj}))_{i=1,\ldots n_i; j=1,\ldots,n_j}$. The matrix $\mathbf{D}$ is a block diagonal matrix that is given as

$$\mathbf{D} = (\mathbf{D}_i)_{i=1,\ldots c}, \tag{37}$$

where $\mathbf{D}_i \in \mathbb{R}^{n_i \times n_i}$ with all terms equal to $1/n_i$. Finally, by solving the eigenvalue problem, we can obtain the eigenvector $\boldsymbol{\alpha}$ that defines the projection matrix $\mathbf{T} \in \mathcal{K}$.

Step 3: Finally, a new feature vector $\mathbf{g}_i = [g_1, g_2, \ldots, g_q]^{\mathrm{T}}$ is obtained from the original selected feature $\mathbf{f}_i$ by the following equation:

$$\begin{aligned} g_k = \mathbf{t}_k^{\mathrm{T}}\phi(\mathbf{f}_i) &= \sum_{l=1}^{c} \sum_{j=1}^{n_l} \alpha_{lj} \phi(\mathbf{f}_{lj})^{\mathrm{T}} \phi(\mathbf{f}_i) \\ &= \sum_{l=1}^{c} \sum_{j=1}^{n_l} \alpha_{lj}\ k(\mathbf{f}_{lj}, \mathbf{f}_i). \end{aligned} \tag{38}$$

After the significant feature set is determined through the feature reduction method, we integrate the SFFS-KBCS+GDA-based features as the input features for the following classifier.

4.9 Classifier Construction

An LS-SVM is a binary classifier that relies on a nonlinear mapping of the training set to a higher-dimensional space, wherein the transformed data is well-separated by a separating hyperplane. Assume a training set of N data points, $\{\mathbf{x}_i, y_i\}_{i=1}^{N}$, where $\mathbf{x}_i = [x_{i1}, \ldots, x_{iq}] \in \mathbb{R}^q$ is the ith input feature vector (q is the number of dimensions of the reduced features obtained by the GDA) and $y_i \in \{-1,\ 1\}$ is the ith output label. The LS-SVMs are used to perform classification tasks using the following decision functions:

$$y(\mathbf{x}) = \mathbf{sign}\Big[\sum\nolimits_{i=\mathbf{1}}^{N} \alpha_i\ y_i k(\mathbf{x},\mathbf{x}_i) + b\Big], \tag{39}$$

where α_i are Lagrange multipliers (which are either positive or negative) and b is a real constant. Subsequently, we utilize the Gaussian RBF function for the kernel function k in this study, which is described as follows:

$$k(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{exp}(-\|\mathbf{x}_i - \mathbf{x}_j\|^{\mathbf{2}}/\mathbf{2}\sigma^{\mathbf{2}}), \tag{40}$$

where the Gaussian width σ is the kernel parameter ($\sigma > 0$). The decision functions are constructed as follows:

$$y_i\big[\mathbf{w}^{\mathbf{T}}\boldsymbol{\phi}(\mathbf{x}_i) + b\big] \geq \mathbf{1}, i = \mathbf{1}, \ldots, N, \tag{41}$$

where $\phi(\cdot)$ is a nonlinear mapping function to map the input space onto a higher dimensional space. However, $\phi(\cdot)$ is not explicitly constructed since the possibility of a separating hyperplane does not exist in the higher dimensional space. Therefore, slack variables $\xi = (\xi_1, \ldots, \xi_N)$ introduced as follows are utilized to solve this misclassification problem as follows:

$$\begin{cases} y_i\big[\mathbf{w}^{\mathbf{T}}\boldsymbol{\phi}(\mathbf{x}_i) + b\big] \geq \mathbf{1} - \xi_i, i = \mathbf{1}, \ldots, N \\ \xi_i \geq \mathbf{0}, i = \mathbf{1}, \ldots, N \end{cases} \tag{42}$$

According to the structural risk minimization principle of statistical learning theory, the risk bound of the LS-SVM is minimized by solving the following optimization problem:

$$\begin{aligned} &\min_{\mathbf{w},b,e} J(\mathbf{w}, b, e) = (1/2)\mathbf{w}^{\mathbf{T}}\mathbf{w} + (\gamma/\mathbf{2})\sum\nolimits_{i=\mathbf{1}}^{N} e_i^{\mathbf{2}}, \\ &\text{subject to: } y_i\big[\mathbf{w}^{\mathbf{T}}\boldsymbol{\phi}(\mathbf{x}_i) + b\big] = \mathbf{1} - e_i, i = \mathbf{1}, \ldots, N. \end{aligned} \tag{43}$$

The Lagrange method is used to solve this problem as follows:

$$\begin{aligned} L(\mathbf{w}, b, e;\alpha)y_i = {} & J(\mathbf{w}, b, e) \\ & - \sum_{i=\mathbf{1}}^{N} \alpha_i\big\{y_i\big[\mathbf{w}^{\mathbf{T}}\boldsymbol{\phi}(\mathbf{x}_i) + b\big] - \mathbf{1} + e_i\big\}. \end{aligned} \tag{44}$$

The Karush-Kuhn-Tucker (KKT) conditions and Mercer's theorem are then applied to solve the equation [44]. More detailed information on the LS-SVM can be referred in [44], [45], and [46].

Because LS-SVM is a binary classifier, a one-against-one strategy is utilized for multiclass classification [45]. In this study, we use the LS-SVMs with one-against-one strategy to classify participants' emotions. For classifying the target emotions in this study, the one-against-one method constructs $m(m-1)/2$ binary classifiers, where a max-wins voting strategy is utilized if m classes need to be classified. If the output of each LS-SVM classifier is one of the two classes, then the assigned class is increased by one vote. Finally, the classification result is represented as the class with the largest votes. If two classes have an identical number of votes, we simply select the one with the smaller label, that is, if Joy and Peacefulness, which are labeled as "1" and "4", respectively, have the same number of votes, we predict that the output is Joy. In the one-against-one method for multiclass support vector machines, if a given sample is classified into two classes with the same number of votes, it is assigned into one class randomly [46] or the class with the smaller index [45]. Our selection is based on the method proposed in [45]. In this study, the output of the classifier is represented as the label of the two types of valence (i.e., negative and positive valence are labeled as "1" and "2"), the two types of arousal (i.e., low and high arousal are labeled as "1" and "2"), and the four types of emotions (i.e., Joy, Tension, Sadness, and Peacefulness are labeled as "1", "2", "3", and "4", respectively).

5 Results

The classification performances of the proposed automatic ECG-based emotion recognition algorithm were validated using a total number of 395 ECG samples collected from 61 participants (including 105 for Joy, 55 for Tension, 40 for Sadness, and 195 for Peacefulness). The ECG segments distributed in positive/negative valence are 300 and 95, respectively, and those distributed in high/low arousal are 160 and 235, respectively. In addition, we compared the classification performances of the LS-SVM classifier between the proposed method and four other feature reduction methods, such as SFFS-KBCS+GDA, SFFS-KBCS+PCA, SFFS-KBCS+LDA, SFFS-KBCS+PCA+LDA, and SFFS-KBCS+PCA+GDA, once the optimal dimensions of each of the feature reduction schemes were estimated. To investigate the robustness of our proposed algorithm, we evaluated the classification performances of the combinations of the feature selection method and feature reduction methods by 2-fold, 10-fold, leave-one-subject-out (LOSO), and leave-one-out (LOO) cross-validation strategies. Four common measures of accuracy (Acc), specificity (Sp), sensitivity (Se), and correct classification rate (CCR) were used to evaluate the performance of the proposed classification scheme.

5.1 Valence Classification

To obtain an optimal feature subset selected by the SFFS-KBCS-based feature selection method for the PCA, LDA, GDA, PCA+LDA, and PCA+GDA feature reduction methods from the 34 features, we varied the number of features from 1 to 34. The LS-SVM classifiers between the aforementioned five feature reduction methods were verified by LOO cross-validation for different numbers of features. A comparison of the CCRs using different numbers of features through the different feature reduction methods with the LS-SVM classifiers is shown in Fig. 7; the best CCR achieved a performance of 82.78 percent when 32 dimensions were selected by the SFFS-KBCS for the GDA method. The overall CCRs of SFFS-KBCS+PCA+LS-SVM, SFFS-KBCS+LDA+LS-SVM, SFFS-KBCS+GDA+LS-SVM, SFFS-KBCS+PCA+LDA+LS-SVM, and SFFS-KBCS+PCA+GDA+LS-SVM were 78.73, 71.14, 82.78, 69.87, and

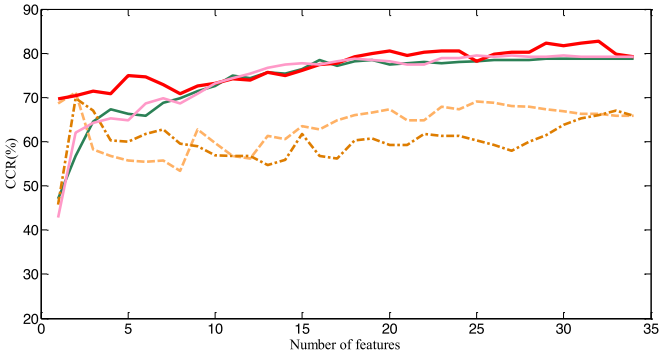


Fig. 7. CCRs versus number of features selected by SFFS-KBCS with LS-SVM classifier between different feature reduction methods by LOO cross-validation in the positive/negative valence classification task. (Green: PCA. Orange: LDA. Red: GDA. Brown: PCA + LDA. Pink: PCA + GDA.).

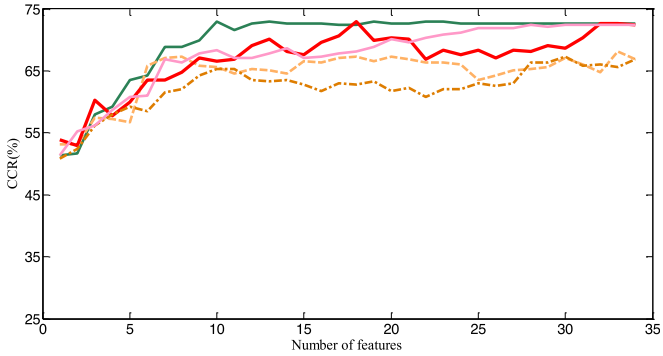


Fig. 8. CCRs versus number of features selected by SFFS-KBCS with LS-SVM classifier between different feature reduction methods by LOO cross-validation in the high/low arousal classification task. (Green: PCA. Orange: LDA. Red: GDA. Brown: PCA+LDA. Pink: PCA+GDA.).

79.49 percent, respectively. The classification performance comparisons of the proposed feature reduction methods with the LS-SVM classifiers by the 2-fold, 10-fold, LOSO, and LOO cross-validation strategies are summarized in Table 2. Obviously, these results demonstrate that the proposed SFFS-KBCS+GDA+LS-SVM scheme outperforms other classification schemes for classifying positive/negative valence emotions.

5.2 Arousal Classification

To obtain an optimal feature subset selected by the SFFS-KBCS-based feature selection method for the PCA, LDA, GDA, PCA+LDA, and PCA+GDA feature reduction methods from the 34 features, we varied the number of features from 1 to 34. A comparison of the CCRs using different numbers of features through the different feature reduction methods with the LS-SVM classifier verified by LOO cross-validation is shown in Fig. 8; the best CCR was 72.91 percent when 18 dimensions were selected by the SFFS-KBCS for the GDA method. The overall CCRs of SFFS-KBCS+PCA+LS-SVM, SFFS-KBCS+LDA+LS-SVM, SFFS-KBCS+GDA+LS-SVM, SFFS-KBCS+PCA+LDA+LS-SVM, and SFFS-KBCS+PCA+GDA+LS-SVM were 72.91, 68.10, 72.91, 67.34, and 72.41 percent, respectively. The classification performance comparisons of the proposed feature reduction methods with the LS-SVM classifiers by the 2-fold, 10-fold, LOSO, and LOO cross-validation strategies are summarized in Table 3. These results demonstrate that the SFFS-KBCS+GDA+LS-SVM classification scheme can obtain the best classification performance in the high/low arousal classification task.

5.3 Four Types of Emotion Classification

Similar to the aforementioned emotion classification tasks, we varied the number of features from 1 to 34 for obtaining an optimal feature subset for PCA, LDA, GDA, PCA+LDA, and PCA+GDA feature reduction methods from the 34 features. The LS-SVM classifiers between the aforementioned five feature reduction methods were also verified by LOO cross-validation for different numbers of features. A comparison of the CCRs using different numbers of features through the different feature reduction methods with the LS-SVM classifiers is shown in Fig. 9; the best CCR was 61.52 percent when 24 dimensions were selected by the SFFS-KBCS for the GDA method. The average accuracy of SFFS-KBCS+PCA+LS-SVM, SFFS-KBCS+LDA+LS-SVM, SFFS-KBCS+GDA+LS-SVM, SFFS-KBCS+PCA+LDA +LS-SVM, and SFFS-KBCS+PCA+GDA+LS-SVM was 78.73, 75.19, 80.76, 75.19, and 80.63 percent, respectively. The overall CCRs of the aforementioned five classification schemes were 57.47, 50.38, 61.52, 50.38, and 61.27 percent, respectively. The classification performance comparisons of the proposed feature reduction methods with the LS-SVM classifiers by the four cross-validation strategies are summarized in Table 4. These results demonstrate the ability of the SFFS-KBCS+GDA+LS-SVM classification scheme to classify the four types of emotions effectively.

6 Discussion

6.1 Comparisons of Proposed Method with Other Existing Approaches for MAHNOB-HCI Database

In this section, we compared the classification performances of the proposed SFFS-KBCS+GDA+LS-SVM scheme with

TABLE 2
Classification Performance Comparisons of Proposed Feature Reduction Methods with LS-SVM Classifiers by Cross-Validation in Positive/Negative Valence Classification Task

K-fold / Reduction Method	2-fold				10 fold				LOSO				LOO			
	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)
SFFS-KBCS+PCA	72.15	81.00	44.21	72.15	77.72	84.33	56.84	77.72	75.95	83.67	51.58	75.95	78.73	86.67	56.84	78.73
SFFS-KBCS+LDA	70.38	82.67	31.58	70.38	69.37	67.33	75.79	69.37	69.37	82.00	29.47	69.37	71.14	81.33	38.95	71.14
SFFS-KBCS+GDA	**77.22**	**85.67**	**50.53**	**77.22**	**82.28**	**87.33**	**66.32**	**82.28**	**76.20**	**84.67**	**49.47**	**76.20**	**82.78**	**88.00**	**66.32**	**82.78**
SFFS-KBCS+PCA+LDA	64.56	67.00	56.84	64.56	67.09	68.33	63.16	67.09	65.32	76.67	29.47	65.32	69.87	84.33	24.21	69.87
SFFS-KBCS+PCA+GDA	75.19	86.33	40.00	75.19	78.99	85.67	57.89	78.99	75.95	84.33	49.47	75.95	79.49	85.33	61.05	79.49

TABLE 3
Classification Performance Comparisons of Proposed Feature Reduction Methods with LS-SVM Classifiers by Cross-Validation in High/Low Arousal Classification Task

K-fold / Reduction Method	2 fold				10-fold				LOSO				LOO			
	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)
SFFS-KBCS+PCA	68.86	58.75	75.74	68.86	72.66	64.38	78.30	72.66	53.42	43.75	60.00	53.42	72.91	66.25	77.45	72.91
SFFS-KBCS+LDA	64.81	65.63	64.26	64.81	69.37	63.13	73.62	69.37	61.77	63.13	60.58	61.77	68.10	61.88	72.34	68.10
SFFS-KBCS+GDA	**70.89**	**68.13**	**72.77**	**70.89**	**73.67**	**70.00**	**76.17**	**73.67**	**57.22**	**51.88**	**60.85**	**57.22**	**72.91**	**72.50**	**73.19**	**72.91**
SFFS-KBCS+PCA+LDA	63.04	48.13	73.19	63.04	70.13	63.75	74.47	70.13	59.24	41.88	71.06	59.24	67.34	61.25	71.49	67.34
SFFS-KBCS+PCA+GDA	67.85	59.38	73.62	67.85	73.67	65.00	79.57	73.67	53.16	46.88	57.45	53.16	72.41	61.88	79.57	72.41

TABLE 4
Classification Performance Comparisons of Proposed Feature Reduction Methods with LS-SVM Classifiers by Cross-Validation in Four Emotions Classification Task

K-fold / Reduction Method	2-fold				10-fold				LOSO				LOO			
	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)	Acc (%)	Sp (%)	Se (%)	CCR (%)
SFFS-KBCS+PCA	77.47	83.32	52.88	54.94	77.85	83.76	53.31	55.70	65.57	74.08	25.20	31.14	78.73	84.29	54.86	57.47
SFFS-KBCS+LDA	74.43	82.54	50.80	48.86	75.95	83.30	53.48	51.90	68.99	77.53	37.46	37.97	75.19	82.76	53.14	50.38
SFFS-KBCS+GDA	**78.86**	**84.24**	**52.44**	**57.72**	**80.51**	**85.45**	**59.10**	**61.01**	**68.10**	**78.01**	**38.63**	**36.20**	**80.76**	**85.69**	**61.17**	**61.52**
SFFS-KBCS+PCA+LDA	74.43	81.13	46.82	48.86	75.70	83.08	53.00	51.39	66.96	76.11	30.89	33.92	75.19	82.04	50.40	50.38
SFFS-KBCS+PCA+GDA	78.10	82.89	48.59	56.20	79.75	84.08	54.02	59.49	65.95	73.42	22.85	31.90	80.63	84.47	54.69	61.27

the existing ANOVA+SVM scheme presented in [8] by using only the ECG signals and the peripheral biosignals (ECG, GSR, RSP, and ST) in the MAHNOB-HCI database, respectively. The MAHNOB-HCI database developed by Soleymani et al. is an open database for emotion analysis. In this database, the EEG signal, peripheral biosignals (ECG, GSR, RSP, and ST), face and body videos, eye gaze, and audio were collected from 27 subjects (11 males and 16 females, age: 26.06±4.39 years). The emotions of each subject were elicited through 20 emotional video clips, and subjects also presented a self-report for the five questions, including emotional label, arousal, valence, dominance, and predictability. According to [8], the affective states are divided into three levels on the valence and arousal axes based on the subject's emotional labels. The user-independent classification rate was validated by the LOSO cross-validation strategy. The performance comparisons of our proposed scheme using only ECG signals and the ANOVA+SVM scheme using the peripheral biosignals are summarized in Table 5. The results show that the classification performance of the proposed SFFS-KBCS+GDA+LS-SVM scheme is similar to that of the existing ANOVA+SVM scheme even if the former used only ECG signals collected from [8] instead of the four peripheral biosignals. This validates that the proposed scheme can be served as an effective tool for classifying positive/negative valence and high/low arousal elicited through the video clips by using only ECG signals.

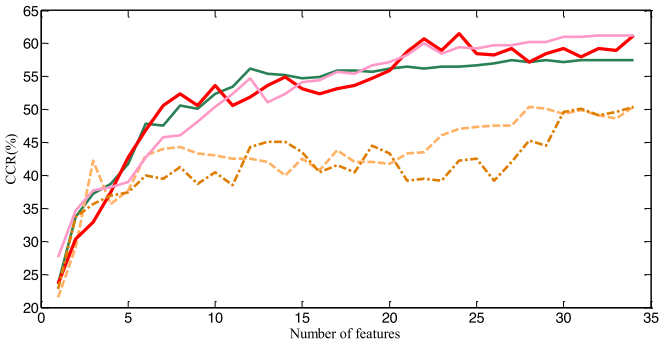


Fig. 9. CCRs versus number of features selected by SFFS-KBCS with LS-SVM classifier between different feature reduction methods by LOO cross-validation in the four emotions classification task. (Green: PCA. Orange: LDA. Red: GDA. Brown: PCA+LDA. Pink: PCA+GDA.).

6.2 Comparisons of Proposed Method with Other Existing Approaches Using Biosignals

This study aimed to develop an automatic ECG-based emotion recognition algorithm that can classify positive/negative valence, high/low arousal, and four types of emotions (positive valance/high arousal (Joy), negative valance/high arousal (Tension), negative valance/low arousal (Sadness), and positive valance/low arousal (Peacefulness)) effectively, respectively, using only ECG signals when listening to music. In this section, we compared the existing classification schemes with our proposed scheme for the positive/negative valence, high/low arousal, and multi-emotion classification tasks, respectively. The performance comparisons of our proposed scheme and the existing schemes for the valence, arousal, and multi-emotion classification tasks are summarized in Tables 6, 7, and 8, respectively.

We used the proposed SFFS-KBCS+GDA+LS-SVM scheme to classify positive/neutral/negative valance using the ECG data collected from [8]; as shown in Table 6, the classification performance of the proposed SFFS-KBCS+GDA+LS-SVM scheme and the ANOVA+SVM scheme [8] is similar. In addition, the CCR of the proposed SFFS-KBCS+GDA+LS-SVM scheme is approximately 20 percent higher than that of the LDA+Gaussian naïve Bayes scheme with BVP, EOG, EMG, GSR, ST, and RSP biosignals [5]. Thus, the proposed SFFS-KBCS+GDA+LS-SVM scheme is appropriate

TABLE 5
CCRs Obtained from Different Classification Schemes Using Different Signals for MAHNOB-HCI Database

Signals	Induction Method	Structure	Valence (%)	Arousal (%)
ECG	Video Clips	SFFS-KBCS+GDA+LS-SVM	44.1	49.2
ECG, GSR, ST, RSP	Video Clips	ANOVA+SVM	45.5	46.2

TABLE 6
Classification Performance Comparisons of Proposed Classification Scheme with Some Existing Schemes for Positive/Negative Valence Classification Task

Author	Signals	No. of Subjects	Emotions	Induction Method	Classification Scheme	CCR
Proposed method	ECG	61	Positive/Negative Valence	Music	SFFS-KBCS+	82.78%
		27	Positive/Neutral/Negative Valence	Video Clips	GDA+LS-SVM	44.10%
Koelstra et al. [5]	BVP, EOG, EMG, GSR, ST, RSP	32	Positive/Negative Valence	Music Video Clips	LDA+Gaussian naïve Bayes	62.70%
Soleymani et al. [8]	ECG, GSR, RSP, SC	27	Positive/Neutral/Negative Valence	Video Clips	ANOVA+SVM	45.50%

TABLE 7
Classification Performance Comparisons of Proposed Classification Scheme with Some Existing Schemes for High/Low Arousal Classification Task

Author	Signals	No. of Subjects	Emotions	Induction Method	Classification Scheme	CCR
Proposed method	ECG	61	High/Low Arousal	Music	SFFS-KBCS+GDA+LS-SVM	72.91%
		27	Positive/Neutral/Negative Valence	Video Clips		49.20%
Koelstra et al. [5]	BVP, EOG, EMG, GSR, ST, RSP	32	High/Low Arousal	Music Video Clips	LDA+Gaussian naïve Bayes	57.00%
Soleymani et al. [8]	ECG, GSR, RSP, SC	27	Excited/Medium/Calm Arousal	Video Clips	ANOVA+SVM	46.20%

TABLE 8
Classification Performance Comparisons of Proposed Classification Scheme with Some Existing Schemes for Multi-Emotion Classification Task

Author	Signals	No. of Subjects	Emotions	Induction Method	Classification Scheme	CCR
Proposed method	ECG	61	Joy, Tension, Sadness, Peacefulness	Music	SFFS-KBCS+ GDA+LS-SVM	61.52% (4 emotions)
Kim et al. [2]	ECG, ST, SC	50	Sad, Anger, Stress, Surprise	Multimodal (Audio, Visual and Cognitive stimuli)	SVM	78.40% (3 emotions)
Rainville et al. [6]	ECG, RSP	43	Fear, Anger, Sadness, Happiness	Recall their personal emotional episode	PCA+Heuristic decision tree	61.80% (4 emotions) 65.30% (4 emotions)
Rigas et al. [7]	EMG, ECG, GSR, RSP	9	Happiness, Disgust, Fear	Picture viewing (IAPS)	K-nearest neighbors	62.70% (3 emotions)
Kim and André [4]	ECG, SC, EMG, RSP	3	Joy, Anger, Sad, Pleasure	Music	pLDA+EMDC	69.70% (4 emotions)
Wen et al. [10]	OXY, GSR, ECG	101	Amusement, Anger, Grief, Fear, Baseline	Video Clips	Random forests classifier	74.00% (5 emotions)
Gu et al. [1]	ECG, BVP, GSR, EMG, RSP	28	Positive & High arousal, Negative & High arousal, Positive & Low arousal, Negative & Low arousal	Picture viewing (IAPS)	K-nearest neighbors	50.30% (4 emotions)

for evaluating positive/negative valence emotion classification performance by using only ECG signals.

According to Table 7, the CCRs obtained using the proposed SFFS-KBCS+GDA+LS-SVM scheme for the high/low arousal classification task were 72.91 and 49.20 percent when the participants' emotions were elicited by music and video clips, respectively. In addition, the overall CCR of the proposed SFFS-KBCS+GDA+LS-SVM scheme was better than that of the ANOVA+SVM scheme [8] and the LDA+Gaussian naïve Bayes scheme [5] by more than 3.00 and 15.91 percent, respectively. Therefore, the results indicate that the proposed SFFS-KBCS+GDA+LS-SVM scheme is the best combination for the high/low arousal emotion classification task when using only ECG signals.

The performance comparisons of our proposed scheme and six existing schemes for the multi-emotion classification task are summarized in Table 8. The results obviously show that the performance of the proposed SFFS-KBCS+GDA+LS-SVM scheme using only ECG signals is similar to that of other schemes using multi biosignals. This validates that the proposed SFFS-KBCS+GDA+LS-SVM scheme can achieve performance similar to that of other existing schemes even when using fewer biosignals. Furthermore, we find that the CCRs of the proposed SFFS-KBCS+GDA+LS-SVM scheme deteriorated from 82.78-72.91 to 61.52 percent when the number of classified emotion categories was increased from two to four. In other words, the number of emotion categories is an influential factor that deteriorates the performance of the feature reduction method and the LS-SVM classifier when the proposed scheme uses only ECG signals.

7 Conclusion

In this paper, we present an automatic ECG-based emotion recognition algorithm consisting of the SFFS-KBCS+GDA feature reduction method and LS-SVM classifiers using only ECG signals for discriminating positive/negative valence, high/low arousal, and four types of emotions (joy, tension, sadness, and peacefulness) elicited by listening to music, respectively. A total of 34 features extracted from the time-, frequency-domain, and nonlinear analyses of ECG signals to provide discriminative information for the emotion recognition tasks. Subsequently, the degrees of the physiological changes in the aforementioned ECG features between the conditions of the baseline stage and the music listening stage were obtained for the proposed SFFS-KBCS+GDA+LS-SVM classification scheme. The overall CCRs of 82.78, 72.91, and 61.52 percent were obtained by the LOO cross-validation strategy for valence, arousal, and four emotion classes using the ECG features with the proposed SFFS-KBCS+GDA+LS-SVM classification scheme to classify a multi-subject database consisting of 395 strongly elicited affective samples from 61 participants. According to the aforementioned experimental results, the effectiveness of the proposed SFFS-KBCS+GDA+LS-SVM scheme has been validated successfully. Moreover, the CCRs are higher than or similar to those reported in the literatures reviewed in this paper when considering the different induction methods, emotion types, and number of subjects. In conclusion, from the aforementioned experimental results, we believe that the proposed automatic ECG-based emotion recognition algorithm can be considered effective for ECG-based emotion recognition tasks.

Acknowledgments

This work was supported by the Ministry of Science and Technology of the Republic of China, Taiwan, under Grant No. MOST 106-3011-E-006-002 and MOST 106-2221-E-035-004.

References

[1] Y. Gu, S. L. Tan, K. J. Wong, M. H. R. Ho, and L. Qu, "A biometric signature based system for improved emotion recognition using physiological responses from multiple subjects," in *Proc. 8th IEEE Int. Conf. Industrial Informat.*, 2010, pp. 61–66.

[2] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Med. Biol. Eng. Comput.*, vol. 42, pp. 419–427, 2004.

[3] D. Kulić and A. Croft, "Affective state estimation for human-robot interaction," *IEEE Trans. Robotics*, vol. 23, no. 5, pp. 991–1000, Oct. 2007.

[4] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.

[5] S. Koelstra, et al., "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan.-Mar. 2012.

[6] P. Rainville, A. Bechara, N. Naqvi, and A. R. Damasio, "Basic emotions are associated with distinct patterns of cardiorespiratory activity," *Int. J. Psychophysiology*, vol. 61, pp. 5–18, 2006.

[7] G. Rigas, C. D. Katsis, G. Ganiatsas, and D. I. Fotiadis, "A user independent, biosignal based, emotion recognition method," in *Proc. 11th Int. Conf. User Modeling*, 2007, pp. 314–318.

[8] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 42–55, Jan.-Mar. 2012.

[9] A. Kleinsmith and N. Bianchi-Berthouze, "Affective body expression perception and recognition: A survey," *IEEE Trans. Affect. Comput.*, vol. 4, no. 1, pp. 15–33, Jan.-Mar. 2013.

[10] W. Wen, G. Liu, N. Cheng, J. Wei, P. Shangguan, and W. Huang, "Emotion recognition based on multi-variant correlation of physiological signals," *IEEE Trans. Affect. Comput.*, vol. 5, no. 2, pp. 126–140, Apr.-Jun. 2014.

[11] K. Wac and C. Tsiourti, "Ambulatory assessment of affect: Survey of sensor systems for monitoring of autonomic nervous systems activation in emotion," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 251–272, Jul.-Sep. 2014.

[12] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 327–339, Jul.-Sep. 2014.

[13] M. Kusserow, O. Amft, and G. Tröster, "Modeling arousal phases in daily living using wearable sensors," *IEEE Trans. Affect. Comput.*, vol. 4, no. 1, pp. 93–105, Jan.-Mar. 2013.

[14] F. Agrafioti, D. Hatzinakos, and A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 102–115, Jan.-Mar. 2012.

[15] P. Ekman, "An argument for basic emotions," *Cognition Emotion*, vol. 6, pp. 169–200, 1992.

[16] J. A. Russell, "A circumplex model of affect," *J. Personality Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[17] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Develop. Psychopathology*, vol. 17, no. 3, pp. 715–734, 2005.

[18] M. D. van der Zwaag, J. H. Janssen, and J. H. D. M. Wesrerlink, "Directing physiology and mood through music: Validation of an affectice music player," *IEEE Trans. Affect. Comput.*, vol. 4, no. 1, pp. 57–68, Jan.-Mar. 2013.

[19] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biol. Psychology*, vol. 84, pp. 394–421, 2010.

[20] G. Valenza, A. Lanatà, and E. P. Scilingo, "The role of nonlinear dynamics in affective valence and arousal recognition," *IEEE Trans. Affect. Comput.*, vol. 3, no. 2, pp. 237–249, Apr.-Jun. 2012.

[21] M. Nardelli, G. Valenza, A. Greco, A. Lanata, and E. P. Scilingo, "Recognizing emotions induced by affective sounds through heart rate variability," *IEEE Trans. Affect. Comput.*, vol. 6, no. 4, pp. 385–394, Oct.-Dec. 2015.

[22] M. Orini, R. Bailón, R. Enk, S. Koelsch, L. Mainardi, and P. Laguna, "A method for continuously assessing the automatic response to music-induced emotions through HRV analysis," *Med. Biol. Eng. Comput.*, vol. 48, pp. 423–433, 2010.

[23] C. L. Krumhansl, "An exploratory study of musical emotions and psychophysiology," *Canadian J. Exp. Psychology*, vol. 51, pp. 336–352, 1997.

[24] A. L. Roque, et al., "The effects of auditory stimulation with music on heart rate variability in healthy women," *Clinics*, vol. 68, no. 7, pp. 960–967, 2013.

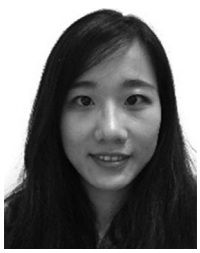
[25] M. Naji, M. Firoozabadi, and P. Azadfallah, "Classification of music-induced emotions based on information fusion of forehead biosignals and electrocardiogram," *Cognitive Comput.*, vol. 6, no. 2, pp. 241–252, 2014.

[26] F. M. Vanderlei, L. C. de Abreu, D. M. Carner, and V. E. Valenti, "Symbolic analysis of heart rate variability during exposure to musical auditory stimulation," *Alternative Therapies Health Med.*, vol. 22, no. 2, pp. 24–31, 2016.
[27] A. Gabrielsson and P. N. Juslin, "Emotion expression in music," in *Handbook of Affective Sciences*, R. J. Davidson, K. R. Scherer, and H. H. Goldsmith, Eds. Oxford, U. K.: Oxford Univ. Press, 2003, pp. 503–534.
[28] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230–236, 1985.
[29] M. Mneimneh, E. Yaz, M. Johnson, and R. Povinelli, "An adaptive kalman filter for removing baseline wandering in ECG signals," in *Proc. Comput. Cardiology*, 2006, pp. 253–256.
[30] P. de Chazal, C. Heneghan, E. Sheridan, R. Reilly, P. Nolan, and M. O'Malley, "Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnea," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 686–696, Jun. 2003.
[31] J. S. Wang, W. C. Chiang, Y. L. Hsu, and Y. T. C. Yang, "ECG arrhythmia classification using a probabilistic neural network with a feature reduction method," *Neurocomputing*, vol. 116, pp. 38–45, 2013.
[32] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: A review," *Med. Biol. Eng. Comput.*, vol. 44, pp. 1031–1051, 2006.
[33] Task Force of the European Society of Cardiology and North American Society of Pacing and Electrophysiology, "Heart rate variability: Standards of measurement, physiological interpretation and clinical use," *Eur. Heart J.*, vol. 17, pp. 354–381, 1996.
[34] J. P. Niskanen, M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen, "Software for advanced HRV analysis," *Comput. Methods Programs Biomed.*, vol. 76, pp. 73–81, 2004.
[35] S. B. Park, Y. S. Noh, S. J. Park, and H. R. Yoon, "An improved algorithm for respiration signal extraction from electrocardiogram measured by conductive textile electrodes using instantaneous frequency estimation," *Med. Biol. Eng. Comput.*, vol. 46, pp. 147–158, 2008.
[36] R. Bailón, L. Sörnmo, and P. Laguna, "A robust method for ECG-based estimation of the respiratory frequency during stress testing," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1273–1285, Jul. 2006.
[37] W. A. Tiller, R. McCraty, and M. Atkinson, "Cardiac coherence: A new, noninvasive measure of autonomic nervous system order," *Alternative Therapies*, vol. 2, no. 1, pp. 52–65, 1996.
[38] M. P. Tulppo, T. H. Mäkikallio, T. E. S. Takala, T. Seppänen, and H. V. Huikuri, "Quantitative beat-to-beat analysis of heart rate dynamics during exercise," *Amer. J. Physiology-Heart Circulatory Physiology*, vol. 271, pp. 244–252, 1996.
[39] G. D. Vito, S. D. R. Galloway, M. A. Nimmo, P. Maas, and J. J. V. McMurray, "Effects of central sympathetic inhibition on heart rate variability during steady-state exercise in healthy humans," *Clinical Physiology Functional Imaging*, vol. 22, pp. 32–38, 2002.
[40] J. McNames and M. Aboy, "Reliability and accuracy of heart rate variability metrics versus ECG segment duration," *Med. Biol. Eng. Comput.*, vol. 44, pp. 747–756, 2006.
[41] L. Wang, "Feature selection with kernel class separability," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 9, pp. 1534–1546, Sep. 2008.
[42] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, pp. 1119–1125, 1994.
[43] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Comput.*, vol. 12, no. 10, pp. 2385–2404, 2000.
[44] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Process. Lett.*, vol. 9, pp. 293–300, 1999.
[45] C. W. Hsu and C. J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425, Mar. 2002.
[46] B. Liu, Z. Hao, and E. C. C. Tsang, "Nesting one-against-one algorithm based on SVMs for pattern classification," *IEEE Trans. Neural Netw.*, vol. 19, no. 12, pp. 2044–2052, Dec. 2008.

Yu-Liang Hsu (M'17) received the BS degree in automatic control engineering from the Feng Chia University, Taichung, Taiwan, in 2004, and the MS and PhD degrees in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2007 and 2011, respectively. He is currently an assistant professor in the Department of Automatic Control Engineering, Feng Chia University. His research interests include computational intelligence, biomedical engineering, nonlinear system identification, and wearable intelligent technology. He is a member of the IEEE.

Jeen-Shing Wang (S'94-M'02) received the BS and MS degrees in electrical engineering from the University of Missouri, Columbia, in 1996 and 1997, respectively, and the PhD degree from Purdue University, West Lafayette, Indiana, in 2001. He is currently a distinguished professor in the Department of Electrical Engineering, National Cheng Kung University, Taiwan. His research interests include computational intelligence, werable system design, big data analysis, and optimization. He is a member of the IEEE.

Wei-Chun Chiang received the BS and PhD. degrees in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2007 and 2014, respectively. He is currently a postdoctoral fellow in the Department of Electrical Engineering, National Cheng Kung University. His research interests include computational intelligence and biomedical signal analysis.



Chien-Han Hung received the BS degree in engineering science from National Cheng Kung University, in 2011, and the MS degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2013. Her research interests include signal processing and affective computing.