Exploring Affective Peripheral Patterns Based on Body Surface Potentials with Covariance

Wei Wu[1], Yao Pi[1], Xianbin Zhang[1], Lin Xu[1], and Wanqing Wu[1]

[1]Affiliation not available

December 22, 2023

Abstract

Affective patterns based on physiological signals reflect bodily changes linked to specific emotional states. Previous studies on the cardiac electrical signal, a key peripheral physiological signal, were limited by the measurement density of single-lead ECG signal, focusing solely on temporal pattern analysis but ignoring topographic pattern analysis that can reflect the body's emotional response. Our research advances affective peripheral pattern studies by innovatively using body surface potentials to comprehensively monitor cardiac electrical activity with increased measurement density. To tackle the challenge of extracting spatial and temporal features from multi-channel body surface potentials, we establish a dynamic correlation among these diverse channel signals through covariance matrices. Our hypothesis is that the dynamic inter-channel relationship provides a valuable source of insights into emotional clues. Experimental results demonstrate that the extracted spatial and temporal features effectively capture topographic and temporal patterns from cardiac electrical signals, and achieve excellent performance in classification tasks simultaneously. Our finding reveals affective patterns based on body surface potentials for the first time, offering novel insights into affective peripheral patterns analysis.

Exploring Affective Peripheral Patterns Based on Body Surface Potentials with Covariance

Wei Wu, Yao Pi, Xianbin Zhang, *Student Member, IEEE*, Lin Xu, and Wanqing Wu*, *Member, IEEE*

***Abstract*—Affective patterns based on physiological signals reflect bodily changes linked to specific emotional states. Previous studies on the cardiac electrical signal, a key peripheral physiological signal, were limited by the measurement density of single-lead ECG signal, focusing solely on temporal pattern analysis but ignoring topographic pattern analysis that can reflect the body's emotional response. Our research advances affective peripheral pattern studies by innovatively using body surface potentials to comprehensively monitor cardiac electrical activity with increased measurement density. To tackle the challenge of extracting spatial and temporal features from multi-channel body surface potentials, we establish a dynamic correlation among these diverse channel signals through covariance matrices. Our hypothesis is that the dynamic inter-channel relationship provides a valuable source of insights into emotional clues. Experimental results demonstrate that the extracted spatial and temporal features effectively capture topographic and temporal patterns from cardiac electrical signals, and achieve excellent performance in classification tasks simultaneously. Our finding reveals affective patterns based on body surface potentials for the first time, offering novel insights into affective peripheral patterns analysis.**

***Index Terms*—Affective peripheral patterns, body surface potentials, covariance matrix, emotional response, affective computing**

I. INTRODUCTION

EMOTIONS exert a constant influence on humans, holding a significant place within human consciousness. They play a pivotal role in a wide spectrum of human intellectual activities, including perception, decision-making, logical reasoning, social interactions, and communication [1] [2] [3] [4]. Affective computing, as a research field, aspires to develop systems that can perceive, recognize, and comprehend human emotions, responding intelligently, sensitively, and naturally to this emotional states [3]. The investigation of affective patterns is a longstanding research focus of affective computing. In 2014, a groundbreaking study utilizing a topographical self-report tool showed that different emotional states correspond to distinct topography of bodily sensations [5], unveiling the topographic affective patterns of subjective sensations.

Physiological changes provide evidence of affective patterns from an objective perspective. Emotions trigger a cascade of alterations in the body's physiological mechanisms, which are discernible through various physiological signals [6]. EEG signals, as reflections of central nervous system (CNS) activities, have traditionally been the central focus of prior studies [7] [8]. In 2019, Zheng *et al.* explored stable electroencephalogram (EEG) temporal patterns for emotion recognition through machine learning approach and developed a public dataset called SEED [9]. Meanwhile, research evidence has indicated that fundamental human emotions not only engage distinct neural circuits within the cerebral cortex but are also accompanied by specific peripheral autonomic responses [10] [11]. A most recent study by Shui *et al.* provides the latest evidence, they demonstrated the electrodermal representation of emotions in the body and indicate the potential of affective computing through bodily electrodermal signals [12].

The heart is recognized as the peripheral organ most intimately linked with emotion [13]. Both Traditional Chinese Medicine (TCM) [14] and contemporary medical research [15] [16] have confirmed this perspective. Additionally, the "embodied emotion" theory proposes that emotions are perceptions of bodily physiological changes, such as rapid heart rates, which often precede feelings like fear and anxiety [17] [18]. Moreover, interoception, as the focus of "embodied emotion" theory, further emphasizes the heart's significant role [19] [20] [21] [22]. Karl Deisseroth's recent study also reinforced this concept, demonstrating that an increased heart rate itself can induce anxiety or fear, as shown by experiments using noninvasive optogenetic pacemakers for precise control of cardiac rhythms in mice [23].

Although the heart is closely related to affective states, previous research mainly focuses on RR interval changes during emotional experiences, not on the heart's comprehensive physiological patterns through multi-channel bioelectric signals. Heart rate variability (HRV) is recognized for offering insights into emotional processing and the heart's adaptive capabilities [13]. Evelina De Longis *et al.* (2020) proposed that the temporal dependency of negative emotions at work might correlate with lower HRV, indicative of reduced adaptability to changing environments [24]. However, the impact of emotions on the heart extends beyond changes in RR intervals. A recent review highlighted that mental and emotional stress can adversely affect the cardiac system, potentially increasing the risk of developing or triggering atrial fibrillation [25]. Negative emotions affect electrical activity, causing increased

This work was supported in part by the National Key R&D Program of China (2023YFE0110200) and the Guangzhou Science and Technology Planning Project (202003000040). Wanqing Wu is the corresponding author.

Wei Wu, Yao Pi, and Xianbin Zhang are with the School of Biomedical Engineering, Sun Yat-sen University, Shenzhen 518107, China. E-mail: {wuwei79, piyao, zhangxb55}@mail2.sysu.edu.cn.

Lin Xu is with the Department of Geriatric Cardiology, PLA General Hospital of the Southern Theatre Command, Guangzhou 510010, China, and also with The First School of Clinical Medicine, Southern Medical University, Guangzhou 510010, China. E-mail: xxgnk_xlin@126.com.

Wanqing Wu is with the School of Biomedical Engineering, Sun Yat-sen University, Shenzhen 518107, China, and also with the Guangdong Provincial Key Laboratory of Sensor Technology and Biomedical Instrument, Shenzhen 518107, China. E-mail: wuwanqing@mail.sysu.edu.cn.

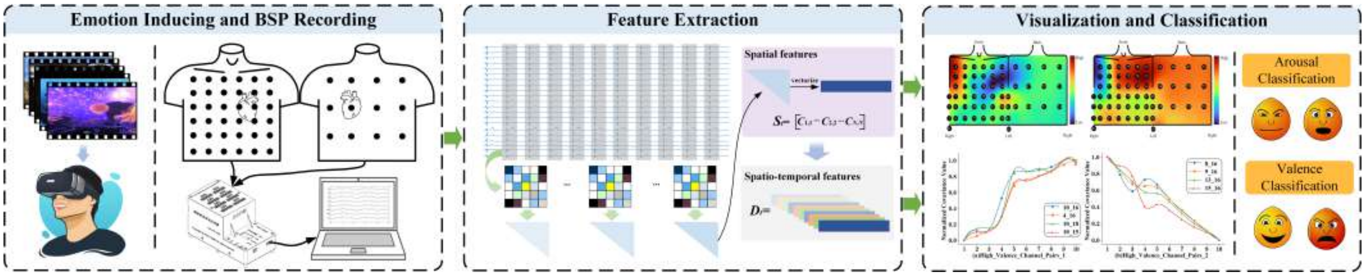


Fig. 1. Our research framework comprises three parts: data acquisition, feature extraction, and pattern visualization, alongside emotion classification.

heart rate, shortened PR and QT intervals, and lengthened QTc interval [26]. These findings indicate a comprehensive effect of emotions on the heart's electrical activity.

Consequently, to further understand the complex relationship between the heart and emotions, developing more comprehensive research methodologies is essential [27]. Regarding the heart's emotional response patterns, merely analyzing affective temporal patterns is insufficient. The affective topographic patterns, with their capacity to visualize the body's emotional responses, offer a valuable complement and novel insight. In fact, discovering the topographic patterns presents numerous potential benefits. The enhancement could lead to a more thorough understanding of the peripheral nervous system's electrical activities in emotional processing [28]. Additionally, it may offer new perspectives into potential autonomic electrophysiological markers [29]. Furthermore, these findings may provide sensitive data on central nervous activation, complementing the study of complex psychological processes [30].

In this study, to explore affective peripheral patterns, particularly topographic patterns, we incorporate body surface potential signals to enhance the density of cardiac electrical activity measurements. We hypothesize that the relationships between different body surface potential (BSP) channels are closely linked to emotional states, and we use BSP's covariance matrix to represent these inter-channel relationships. In feature extraction, we utilize the matrix's upper triangular elements to form spatial features, which, when combined from ten successive heartbeats, generate spatio-temporal features. In the experiment, we rigorously evaluate the effectiveness of the spatial features in classifying emotion valence and arousal, using various feature dimensionality reduction techniques and classifiers, and visualize these features on the body surface as a topographic pattern. To effectively harness temporal information from emotional experiences, we feed spatial-temporal features into a Recurrent Neural Network (RNN) model. The analysis of feature trajectories reveals emotionally significant temporal patterns within the BSP.

In summary, the contributions of this work encompass the following key points:

- We establish a linear transformation between the BSP covariance matrix and the heart surface potentials (HSP) covariance matrix through rigorous proof. This implies that extracting features from the BSP's covariance matrix also indirectly involves extracting features from the HSP's covariance matrix.
- As each BSP channel records the cardiac electrical activity of specific heart regions, we flatten the BSP covariance matrix for each heartbeat to construct 1-D spatial features. Experimental results indicate that different emotions' topographic patterns of this feature are significantly distinct, and they exhibit an excellent ability to distinguish between valence and arousal.
- To capture the temporal dynamics in the emotional process, we introduce a 2-D spatio-temporal features construction method and create an RNN model for the spatio-temporal fusion modeling of the BSP covariance matrix. The emotion classification results demonstrate its effectiveness and evidence from the temporal patterns validates the plausibility of the approach.

The rest of this paper is organized as follows: Section II describes the experimental setup and data preprocessing for the emotion-evoking experiment. Section III outlines the proposed methodology. Section IV discusses the experimental results. Section V provides the conclusion.

II. MATERIALS

A. Participants

A total of 28 healthy participants, ranging in age from 19 to 24 (M = 20.97, SD = 1.53), were enrolled in this experiment. All participants, either undergraduate or graduate students from Sun Yat-sen University, had no history of chronic diseases, heart diseases, neurological conditions, or psychological disorders. Additionally, they have not used any medication for mental health conditions in the last six months, and have not previously taken part in any emotion-evoking experiments. Each participant is instructed to abstain from drugs, alcohol, and caffeine the day preceding the experiment, and to refrain from strenuous exercise for 48 hours prior to the experiment.

B. Experiment Protocol

The experiment is carried out in a dedicated room, maintaining optimal conditions such as a comfortable temperature, humidity, subdued lighting, and complete silence. To capture the BSP, we employ a self-developed BSP acquisition system created by our research group. This system is capable of simultaneously collecting data from up to 128 channels of BSP and storing the acquired data. For the electrode configuration during the BSP acquisition, a total of 54 electrodes are strategically placed across the entire thorax. Among these,

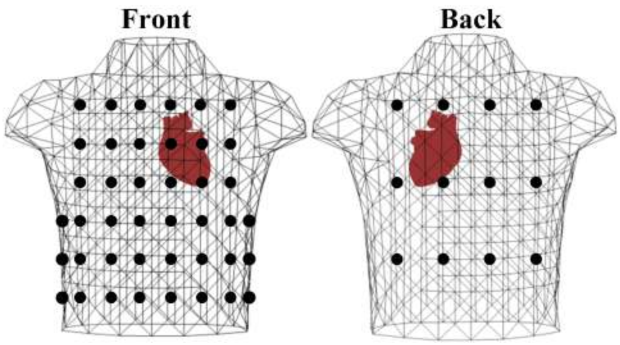


Fig. 2. The electrode configuration of the BSP applied in this experiment.

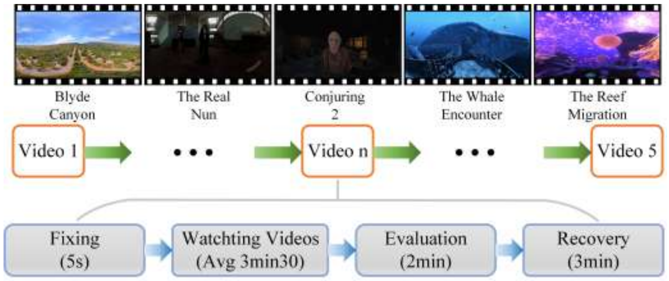


Fig. 3. The experimental procedure. The top of the figure shows the sample screenshot of VR videos the participants watched. They are Blyde Canyon(No.1), The Real Nun(No.2), Conjuring 2(No.3), The Whale Encounter(No.4), The Reef Migration(No.5) from left to right.

36 electrodes are evenly distributed on the anterior chest, 12 electrodes are evenly positioned on the back, and 6 electrodes are situated on the lateral side of the body. Additionally, two extra electrodes, namely the Right-Leg Drive and the Reference electrode, were used (refer to Fig. 2).

The emotion-inducing device used in this experiment is the Oculus Quest 2 VR glasses. VR devices have been demonstrated to provide participants with an immersive audio-visual experience in emotion-inducing experiments. To ensure the credibility and effectiveness of the experimental materials, we select materials from two distinct sources. The first source is a previous study conducted by Li Ming *et al.* in 2022. Their research scientifically evaluated the use of VR as an active method for eliciting emotions and established that VR has significant advantages in evoking emotions related to arousal and valence [32]. We select four video materials from this study for our experiment. The second source is the Stanford Immersive VR Video Public Database [31], from which we choose one video material for our experiment. The details of the selected VR videos for the experiment can be found in TABLE I.

The experimental setup involves two computers. One is equipped with an Intel Core i9 processor, a GTX 1660Ti GPU, Windows 10, and 32 GB of RAM. This computer is connected to the VR display device and is responsible for controlling video playback and facilitating the participants' subjective evaluation of their emotional experience using the Self-assessment Manikin (SAM) [33]. The other computer featured an Intel Core i5 processor, Windows 11, and 16 GB of RAM. This computer is dedicated to real-time monitoring and the storage of BSP recordings.

Prior to the start of the experiment, every participant is given an informed consent form to read and understand. This form outlines the purpose, procedure, and potential risks associated with the experiment. Participants are required to sign the form as an indication of their agreement to participate. Participants are informed they could withdraw from the experiment at any time without any consequences. Following this, the experimenter provides instructions to the participants on how to complete the Self-assessment Manikin (SAM) scale and communicates any precautions to be observed during the experiment. Subsequently, the experimenter places the electrodes for the participants and ensures the quality of the signal. Once all preparations are in place, the experiment starts, and the physiological signals are recorded.

TABLE I
INFORMATION OF THE VR VIDEOS

NO.	Name	Duration	Source
1	Blyde Canyon	2min37s	[31]
2	The Real Nun	2min54s	[32]
3	Conjuring 2	3min8s	[32]
4	The Whale Encounter	1min30s	[32]
5	The Reef Migration	6min50s	[32]

The experiment comprises five trials, with each trial involving the participants watching one video and performing subjective self-assessment once. Each trial for every participant included three steps: firstly, focusing on a central cross that appeared on the screen for 5 seconds; secondly, viewing the video corresponding to the trial; and finally, rating their emotional state using the SAM. A 3-minute break was provided between each trial to allow participants to rest and recover, preventing mental fatigue. The entire experiment lasted approximately 1.5 hours. An overview of the procedure is depicted in Fig. 3.

III. METHODOLOGY

A. *Signal Preprocessing*

Our self-developed device captures body surface potential signals at a sample rate of 250 Hz. The quality of signals significantly influences signal analysis results, so preprocessing signals is crucial for further analysis.

In this study, we must eliminate two interfering signals: (1) Noise originating from the power supply line, typically at a frequency of 50 Hz or its multiples, as well as noise from EMG signals, can overpower the relatively weak BSP signals and disrupt the measurement process (refer to Fig. 4(a)). (2) Baseline wander, which can result from a participant's movement or respiration, may lead to incorrect signal labeling (as depicted in Fig. 4(a)), subsequently impacting the accurate identification of features [34]. To eliminate these noises and prepare the signals for further analysis, we employ a two-step approach:

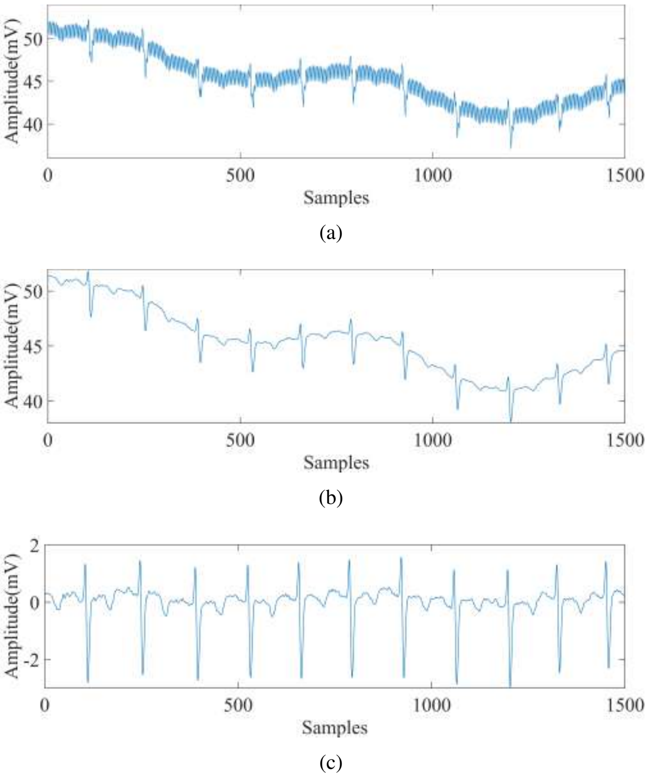


Fig. 4. The preprocessing for one channel BSP signal. (a) Raw signal (b) Filtered signal (c) Pure signal after preprocessing

1) Digital Filtering: In this step, we use the Filter Designer Toolbox in MATLAB [35] to design two notch filters with center frequencies, f_C, set at 50 Hz and 100 Hz, respectively. Additionally, we design a low-pass filter with a cut-off frequency, f_L, at 45 Hz. Subsequently, we apply these filters to the BSP signals using a Zero-phase filter, achieved through the “filtfilt” function in MATLAB. The first two filters are responsible for eliminating power line interference, while the third filter serves to remove high-frequency EMG components not within the BSP signal’s frequency band (see Fig. 4(b)).
2) Baseline Wander Removal: Following the approach detailed in [36], we remove the QRS complexes, P waves, and T waves from the BSP signal. This is accomplished by using a mean filter, which is implemented through the “movmean” function in MATLAB, to obtain the baseline wander signal. Subsequently, we subtract the baseline wander component from the original signals (see Fig. 4(c)). This process yields clean and pure signals suitable for subsequent feature extraction.

B. R-Wave Detection

Cardiac electrical activity exhibits a quasi-periodic nature, marked by the occurrence of each heartbeat, which serves as the fundamental unit of observation. Consequently, feature extraction in this paper is conducted on a per-heartbeat basis. The accuracy of this feature extraction process is highly dependent on precise heartbeat segmentation. In this study, we utilize an R-wave detection algorithm to segment the BSP signal into individual heartbeat records. The R-wave position is identified using the algorithm proposed by Pan and Tompkins, which determines the time of occurrence of the R-wave as well as its amplitude [37]. In this paper, each heartbeat is composed of 171 sampling points, including 70 points before the R-wave and 100 points after the R-wave.

C. Incorrect Signal Rejection

Prior to feature extraction, it is crucial to ensure that the signals used are complete and accurate observations. Two types of signals need to be excluded from consideration: (1) Incomplete or Noisy BSP Signals: These can result from inadequate electrode-skin contact during the experiment due to participant sweating or movement. (2) Sudden Changes in RR Intervals: Such changes may be incorrectly marked by the R-Wave detection algorithm due to the presence of spike noise that cannot be effectively filtered. Both of these types of recordings can introduce contamination to the dataset. To address this issue, we perform manual checks on all BSP signals to ensure that incorrect signals are excluded from further analysis.

D. Feature Extraction

The BSP signal in this work encompasses up to 54 channels, indicating that the raw signal possesses high dimensionality and contains redundant information. Analyzing the raw signal directly is a challenging task. Consequently, feature extraction is employed to transform the raw data into features. In this work, we utilize the covariance matrix of the BSP for feature extraction, the covariance matrix of the BSP signal serves as a representation of the BSP’s inter-channel relationship and is used to characterize the second-order statistical features of the BSP. The detailed procedures for feature extraction are elaborated below.

1) BSP Covariance Matrices: The covariance between the two channel BSP is calculated based on

$$\operatorname{Cov}(x, y)=\frac{1}{N} \sum_{i=1}^{N}\left(x_{i}-E(x)\right)\left(y_{i}-E(y)\right), \tag{1}$$

where x and y represent the signals from two channel, N denotes the sample points of the segment.

In this work, to make the calculations simple, we let the signals be zero-centered, so the covariance is simplified as

$$\operatorname{Cov}(x, y)=\frac{1}{N} \sum_{i=1}^{N}\left(x_{i} y_{i}\right). \tag{2}$$

For a zero-centered multi-channel BSP time series $X(t)$, the covariance matrix is given as

$$M=\frac{1}{K-1} X X^{T}, X \in R^{N \times K}, \tag{3}$$

where N represents the number of selected BSP channels, and $K=171$ indicates the number of sample points in one heartbeat.

The BSP covariance matrices are calculated to capture the inter-channel relationship of the BSP. These matrices indirectly reflect the inter-channel relationship between the heart surface potentials (HSP).

The relationship between BSP $y(s,t)$ and HSP $u(s,t)$ can be formulated as

$$\boldsymbol{y}(\boldsymbol{s},t) = \boldsymbol{A}\boldsymbol{u}(\boldsymbol{s},t), \tag{4}$$

where t and s denotes the temporal and spatial coordinates respectively, and $\boldsymbol{A}$ represents the forward transfer matrix. Inversely, the relationship between HSP $u(s,t)$ and BSP $y(s,t)$ can be formulated as

$$\boldsymbol{u}(\boldsymbol{s},t) = \boldsymbol{W^T}\boldsymbol{y}(\boldsymbol{s},t), \tag{5}$$

where $\boldsymbol{W^T}$ denotes the inverse transfer matrix. $\boldsymbol{A}$ and $\boldsymbol{W^T}$ should subject to

$$\boldsymbol{A}\boldsymbol{W^T} = \boldsymbol{I}, \tag{6}$$

$\boldsymbol{I}$ is an identity matrix. Based on Eq. 3, we performed a transformation of Eq. 4, leading to the derivation of the relationship between the BSP covariance matrixces and the HSP covariance matrices, as demonstrated in Eq. 7,

$$\Sigma_{\boldsymbol{y}(\boldsymbol{s},t)} = \boldsymbol{A}\Sigma_{\boldsymbol{u}(\boldsymbol{s},t)}\boldsymbol{A^T}, \tag{7}$$

where Σ indicates the covariance matrices. By applying the constraint described in Eq. 6, we deduce Eq. 8 as follows.

$$\Sigma_{\boldsymbol{y}(\boldsymbol{s},t)}\boldsymbol{W} = \boldsymbol{A}\Sigma_{\boldsymbol{u}(\boldsymbol{s},t)} \tag{8}$$

This approach establishes a quantitative relationship between the BSP covariance matrices and the HSP covariance matrices. It demonstrates that by conducting feature extraction on the BSP covariance matrices, we can indirectly extract features from the HSP covariance matrices. As a result, this study computes the covariance matrix of heartbeats, which encapsulates both spatial and temporal features of the electrical signals originating from various locations on the heart surface.

2) Spatial and Temporal Features: In terms of spatial feature extraction, we make the assumption that different emotional states give rise to distinct distributions of covariance matrices. Consequently, elements derived from these covariance matrices can, to some degree, differentiate between emotional states. As depicted in Fig. 5, considering the symmetry of the covariance matrices, we exclusively extract the upper triangular elements of the covariance matrices for each 171-point heartbeat. These elements are then organized into a 1-D vector denoted as S_t, with a dimension of $N(N+1)/2$. These vectors serve as representations of the spatial features of the BSP.

Additionally, given the dynamic and ever-changing nature of emotional experiences [38], we posit that the covariance matrices derived from successive heartbeats encapsulate the temporal relationships between the electrical activities of various locations on the heart. This temporal information can effectively discriminate between emotional states. As depicted in Fig. 5, in this paper, we extract spatio-temporal features from ten consecutive heartbeats within a single trial. The covariance matrix for the tth position in a sequence of ten consecutive heartbeats is defined as:

$$M(t) = \frac{1}{K-1}X(t)X^T(t), t \in 1,2,3,...,9,10. \tag{9}$$

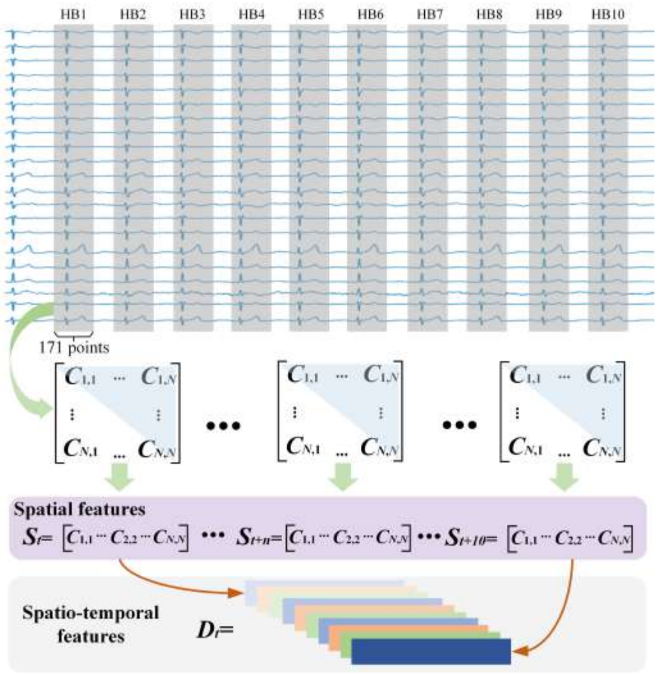


Fig. 5. Illustration of feature extraction scheme. HB1 to HB10 represent ten consecutive heartbeats BSP recordings. In the covariance matrices, $C_{i,j}$ denotes the covariance between the ith channel and jth channel. For spatial features, BSP recordings of each heartbeat form a 1-D vector S_t consisting of unique elements in covariance matrices. For spatio-temporal features, ten stacked consecutive 1-D vectors S_t form D_t.

Also, each covariance matrix is reshaped into a 1-D vector S_t. Consequently, the spatio-temporal features D_t for one trial are defined as a set of these vectors:

$$\{S_t, S_{t+1}, S_{t+2}, S_{t+3}, S_{t+4}, S_{t+5}, S_{t+6}, S_{t+7}, S_{t+8}, S_{t+9}\}.$$

E. Feature Dimensionality Reduction

In this study, we collect 54 channels of BSP signals for each participant. As previously mentioned, the dimension of spatial features is denoted as $N(N+1)/2$. When N is set to 54, the dimension of spatial features extends to 1485, making it infeasible to analyze all features, especially for feature assessments or classification tasks. Therefore, in this study, we employ three well-known approaches to address this issue: singular value decomposition (SVD), principal component analysis (PCA), and the minimal redundancy maximal relevance (MRMR) algorithm [39]. These methods help reduce the dimensionality of the features and enhance the effectiveness of feature analysis and classification tasks.

PCA reduces dimensionality by mapping high-dimensional data into a lower-dimensional space using linear transformations. On the other hand, SVD accomplishes sparse representation through matrix decomposition. The singular value decomposition of an $m \times n$ real matrix A is described by

$$\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V^T}, \tag{10}$$

where the diagonal entries of Σ represent the singular values of matrix A, and the first $p = \min(m,n)$ columns of U and V serve as the left- and right-singular vectors corresponding to these singular values. In this work, SVD is applied to the covariance matrices M to extract a 54-dimensional set of singular values.

While PCA and SVD are effective at reducing dimensionality, they tend to lose the interpretable information of the original features, such as signal channels and frequency. Therefore, in this study, we chose the MRMR algorithm to select a subset of the original features, maintaining the signal's original domain information. The MRMR algorithm selects features based on a minimal-redundancy and maximal-relevance criterion. Max-Relevance criterion seeks to identify feature subsets that maximize the average mutual information between feature x_d and class c:

$$\max D(S, c), D = \frac{1}{|S|} \sum_{x_d \in S} I(x_d; c). \tag{11}$$

Here, S represents the candidate feature subset. However, selecting features based solely on Max-Relevance can lead to redundancy. When two features are highly dependent, removing one of them wouldn't significantly impact class discrimination. Therefore, Min-Redundancy is introduced to ensure the selection of mutually exclusive features:

$$\min R(S), R = \frac{1}{|S|^2} \sum_{x_{di}, x_{dj} \in S} I(x_{di}, x_{dj}). \tag{12}$$

The MRMR algorithm combines these two constraints, resulting in the following objective function:

$$\max\ \delta(D, R), \delta(D, R) = D - R. \tag{13}$$

In this paper, we employ an incremental search method to identify the near-optimal set of K features.

F. Classification

The obtained feature sets after dimensionality reduction are subsequently used as inputs for classifiers to assess their validity. The following machine learning classifiers are employed: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Adaptive Boosting (Adaboost), and Multi-layer Perceptron (MLP). To fine-tune the hyperparameters of these classifiers, a cross-validation grid search is conducted. The specific hyperparameters explored for each classifier are as follows:

- For SVM, a grid with an RBF kernel is employed, and combinations of C values within the range [0.1, 1, 10, 100, 1000] and gamma values within the range [0.1, 0.01, 0.001, 0.0001] are examined.
- For KNN, two grids are explored: one with "uniform" weights and n_neighbors values from 1 to 50, and another with "distance" weights and n_neighbors values also ranging from 1 to 50.
- For Adaboost, a grid with decision trees as the base estimator is used, and the following hyperparameters are considered: estimator_max_depth in the range of 1 to 30, n_estimators in the range of 30 to 60, and learning_rate within the range [0.001, 0.005, 0.01, 0.05, 0.1, 0.5].
- For MLP, a grid with 4 hidden layers, each consisting of 30 neurons, is employed. Hyperparameters explored include activation options ($tanh$ or $relu$), alpha values within the range [0.0001, 0.001, 0.01, 0.05, 0.1, 1], and learning_rate options ("constant" or "adaptive").

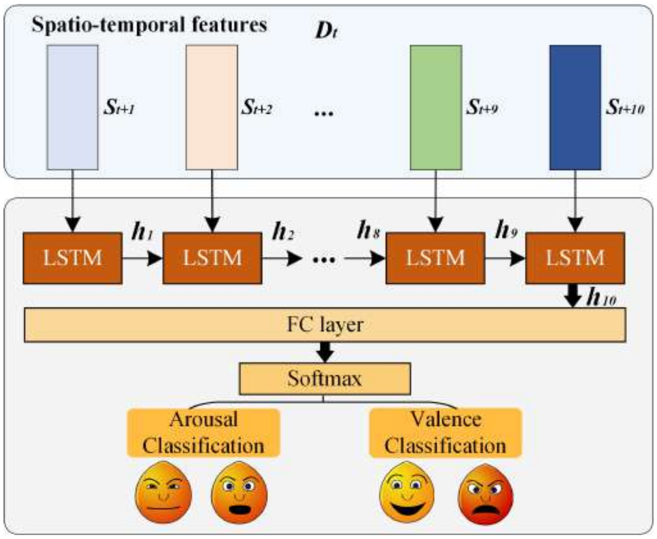


Fig. 6. The framework of our model: the covariance matrices of ten consecutive heartbeats are flattened and then input to a single-layer RNN network to extract temporal dependencies of spatio-temporal features. The classification results are derived from these temporal dependencies by sequentially passing them through a fully connected layer and a softmax layer.

The entire process is implemented using the scikit-learn toolbox [40].

To address the modeling of spatio-temporal features, which comprise stacked 1-dimensional covariance matrices from ten consecutive heartbeats, it is crucial to establish a model that can effectively learn spatial and temporal dynamics simultaneously. The modeling mechanism should be capable of exploring the temporal relationships between different channels present in the spatio-temporal features. Recurrent Neural Network (RNN) is a powerful model designed for end-to-end learning and has demonstrated notable success in modeling temporal sequences, including natural language and speech signals.

As depicted in Fig. 6, RNN networks are equipped to extract temporal relationships in the input data. In this work, we employ ten Long Short-Term Memory (LSTM) cells to construct a single-layer RNN network. This network processes the consecutive covariance matrices, generating a sequence of hidden variables, denoted as $[h_1, h_2, h_3, ..., h_{10}]$, where each h_t represents the hidden state at step t. Consequently, the RNN network captures the long-term dependencies present in the time series. The RNN network's final output, h_{10}, is fed into a subsequent fully connected (FC) layer, as shown in Fig. 6.

Following the FC layer, a softmax layer outputs the probability for each category,

$$P = Softmax(W \times h_{10} + b), \quad P \in R^z. \tag{14}$$

Here, $W \in R^{z \times l}$ represents the FC layer's weight matrix and $b \in R^z$ is the bias vector, where z denotes the number of classes in the model and l indicates the FC layer's size.

IV. RESULTS

A. Results of Subjective SAM Ratings

In the classification session, labels for each segment are required, and these labels are assigned based on the subjective ratings from the participants. In this work, segments with ratings greater than 5 are labeled as “high arousal” or “high valence”, while segments with ratings less than 5 are labeled as “low arousal” or “low valence”. The average ratings for the five video stimuli are presented in TABLE II.

TABLE II
PARTICIPANTS’ RATING OF VR VIDEOS

NO.	Name	Arousal (M±SD)	Valence (M±SD)
1	Blyde Canyon	3.8 ± 1.8	6.3 ± 1.2
2	The Real Nun	6.9 ± 1.4	2.8 ± 1.5
3	Conjuring 2	6.7 ± 1.4	3.0 ± 1.3
4	The Whale Encounter	4.2 ± 2.0	5.2 ± 1.2
5	The Reef Migration	4.0 ± 1.8	5.8 ± 1.3

B. Results of Topographic Patterns and Classification with Spatial Features

As previously mentioned, we employ a combination of four classical machine learning models and three feature dimensionality reduction algorithms to evaluate the effectiveness of spatial features. For the construction of spatial features, we utilize the BSP signals from all 54 channels, i.e., N is set to 54. Specifically, SVD is applied to extract a 1-D singular value vector from a 54×54 covariance matrix, resulting in 54 features after dimensionality reduction. To ensure fair comparisons, we have also reduced the original 1485-dimensional features to 54 dimensions using PCA and MRMR.

The dataset is split into a training set and a test set at a 4:1 ratio. The test set remains completely independent of the parameter optimization and training process. It is solely used to evaluate the model performance with the optimal combination of parameters. We assess the model using four machine learning metrics: accuracy (Acc.), F1 score (F1.), precision (Pre.), and sensitivity (Sen.). These metrics provide a comprehensive evaluation of the model’s performance.

The classification results for arousal based on spatial features are presented in TABLE III. Out of the twelve combinations of classifiers and dimensionality reduction methods tested, the PCA method consistently outperforms both the SVD and MRMR methods for each classifier. The best-performing combination of classifier and dimensionality reduction method is SVM with PCA, which results in an accuracy of 0.989, F1 score of 0.988, precision of 0.983, and sensitivity of 0.991.

For valence classification based on spatial features, the outcomes are shown in TABLE IV. Similar to the arousal classification, the PCA method consistently outperforms SVD and MRMR methods for each classifier. The SVM combined with PCA emerged as the most effective classifier and dimensionality reduction method, achieving an accuracy of 0.986, F1 score of 0.983, precision of 0.987, and sensitivity of 0.980.

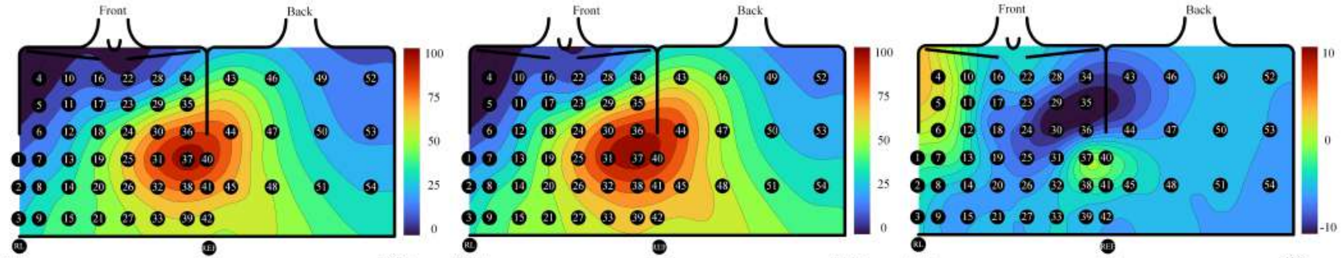
To further investigate the topographic patterns of spatial features, we average the covariance matrices across four emotional states (high arousal, low arousal, high valence, and low valence). We then aggregate the values by summing each row of the covariance matrix for each type, resulting in 54 values. These values are depicted on the body surface as a heat map in Fig. 7. Notably, due to the quasi-periodic nature of cardiac activity, there is no discernible difference in the distribution pattern of features (Fig. 7(a) and Fig. 7(b), Fig. 7(d) and Fig. 7(e)). However, a significant difference emerges in the covariance amplitude. Specifically, the characteristic amplitudes of high arousal in nearly all channels are lower than those of low arousal (Fig. 7(c)), while the characteristic amplitudes of high valence surpass those of low arousal (Fig.

TABLE III
CLASSIFICATION PERFORMANCE COMPARISON OF SPATIAL FEATURES IN AROUSAL CLASSIFICATION TASK

Arousal Classification Results					
Classifier	**Dimension Reduction Method**	**Acc.**	**F1.**	**Pre.**	**Sen.**
SVM	SVD	0.865	0.848	0.850	0.846
	PCA	**0.989**	**0.988**	**0.983**	**0.991**
	MRMR	0.696	0.656	0.659	0.652
KNN	SVD	0.854	0.832	0.848	0.817
	PCA	**0.976**	**0.973**	**0.972**	**0.975**
	MRMR	0.964	0.959	0.960	0.959
Adaboost	SVD	0.881	0.864	0.871	0.858
	PCA	**0.975**	**0.972**	**0.970**	**0.974**
	MRMR	0.961	0.957	0.954	0.960
MLP	SVD	0.868	0.858	0.823	0.895
	PCA	**0.981**	**0.979**	**0.977**	**0.981**
	MRMR	0.926	0.917	0.910	0.924

TABLE IV
CLASSIFICATION PERFORMANCE COMPARISON OF SPATIAL FEATURES IN VALENCE CLASSIFICATION TASK

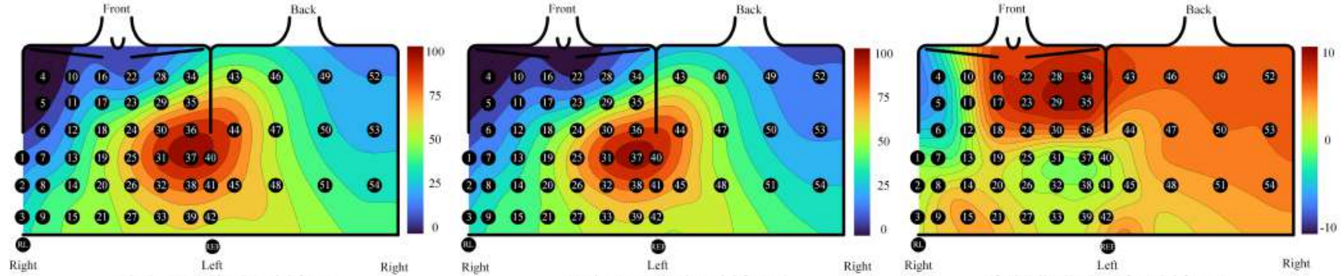
Valence Classification Results					
Classifier	**Dimension Reduction Method**	**Acc.**	**F1.**	**Pre.**	**Sen.**
SVM	SVD	0.874	0.850	0.842	0.858
	PCA	**0.986**	**0.983**	**0.987**	**0.980**
	MRMR	0.682	0.581	0.645	0.529
KNN	SVD	0.858	0.831	0.824	0.838
	PCA	**0.975**	**0.970**	**0.976**	**0.964**
	MRMR	0.956	0.947	0.956	0.938
Adaboost	SVD	0.884	0.857	0.879	0.837
	PCA	**0.972**	**0.966**	**0.970**	**0.961**
	MRMR	0.956	0.947	0.956	0.938
MLP	SVD	0.886	0.862	0.863	0.862
	PCA	**0.977**	**0.973**	**0.970**	**0.976**
	MRMR	0.921	0.905	0.906	0.904



(a)The normalized spatial feature distribution of high arousal

(b)The normalized spatial feature distribution of low arousal

(c)The distribution of spatial feature differences between high and low arousal



(d)The normalized spatial feature distribution of high valence

(e)The normalized spatial feature distribution of low valence

(f)The distribution of spatial feature differences between high and low valence

Fig. 7. topographic patterns of body surface potential signals obtained by visualizing spatial features. (a) and (b) are the topographic patterns of high arousal and low arousal respectively, and (c) is the topographic patterns of the difference between high arousal and low arousal features; (d) and (e) are topographic patterns of high valence and low valence respectively, and (c) is the topographic patterns of the difference between high valence and low valence features;

7(f)).

C. Results of Channel Selection By MRMR

In this study, the MRMR algorithm does not transform the original features but ranks them based on their relevance and redundancy. The top-ranked features, according to the MRMR algorithm, are expected to be more informative for the classification task. To evaluate this, we incrementally include the top 60 features, ranked by the MRMR method, one by one into the feature set and assess the dynamic performance of the four classifiers. This analysis aims to investigate how the classifier's performance changes as more of the top-ranked features are added. The analysis results are expected to offer insights into the correlation between feature ranking and classification performance, aiding in identifying the optimal number of features for effective emotion recognition.

In Fig. 9, for both arousal and valence classification tasks, the performance of KNN, Adaboost, and MLP exhibits rapid improvement as the feature dimension increases from 1 to approximately 5. Beyond this point, the rate of performance enhancement slows down, with MLP's performance showing notable fluctuations. In the case of SVM, its performance in the arousal classification task gradually improves, eventually stabilizing at 70%, as the feature dimension increases. In the valence classification task, SVM initially sees minimal improvement until the feature dimension reaches 44, after which it experiences a 9% performance gain from 44 to 53, but then plateaus at around 69%. These findings align with our hypothesis that classification performance initially improves rapidly with increasing feature dimensions, but the rate of improvement diminishes after reaching a critical point, estimated to be around 5 in this study.

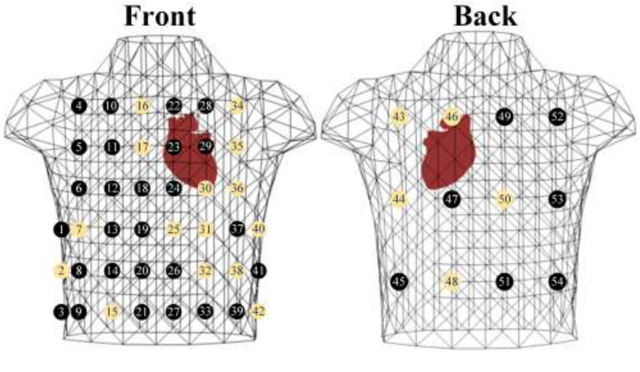
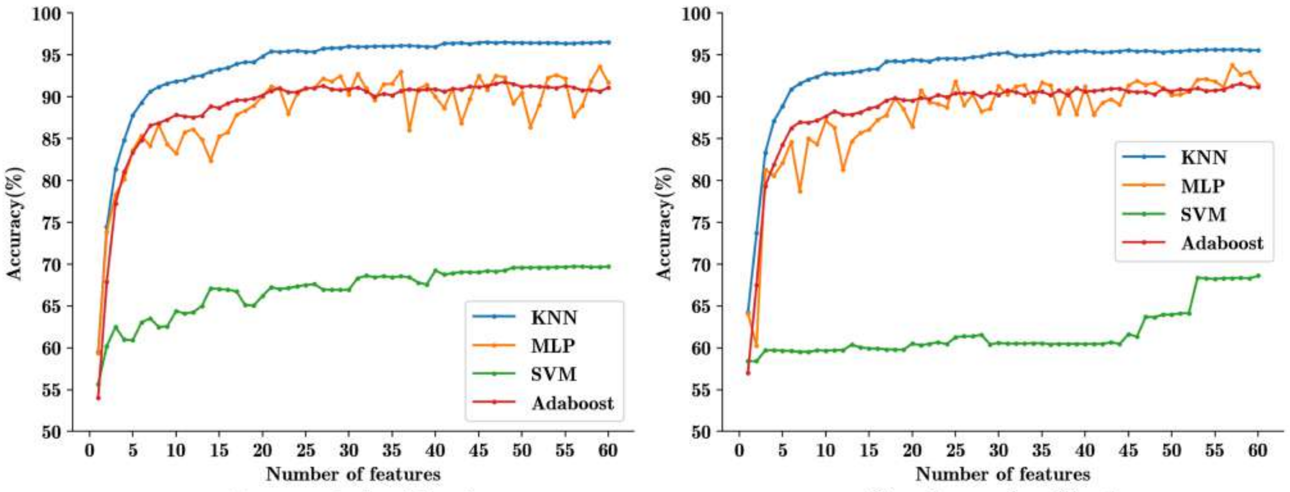


Fig. 8. The top 20 ranked channels' positions on electrode configuration map.

TABLE V
TOP 6 RANKED CHANNEL PAIRS IN THE AROUSAL/VALENCE CLASSIFICATION TASK

Rank	#1	#2	#3	#4	#5	#6
Arousal	[34 54]	[36 38]	[17 46]	[31 50]	[15 30]	[2 36]
Valence	[32 34]	[7 25]	[36 43]	[16 35]	[40 42]	[30 31]

Furthermore, as the features selected by the MRMR method have a one-to-one correspondence with the elements of the covariance matrix, which, in turn, correspond one-to-one with the channel pairs, we can assess channel importance based on MRMR-selected features, facilitating channel selection. Channel selection holds significant value; fewer channels result in less data. If a reduced number of channels can still capture the data patterns of a larger set, it unquestionably enhances analysis efficiency. Moreover, fewer channels are more suitable for practical applications.



(a) arousal classification

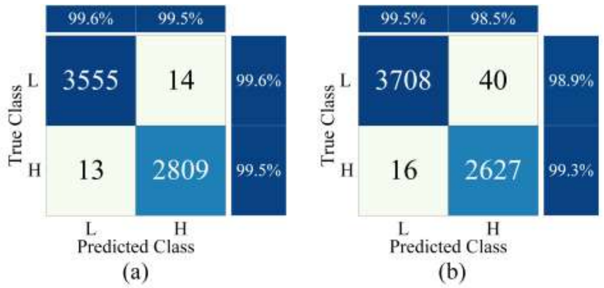
(b) valence classification

Fig. 9. The dynamic classification performance of top 60 spatial features in arousal and valence classification task.

In TABLE V, we present the top six ranked channel pairs obtained through the MRMR approach in the arousal/valence classification task. After removing duplicates from these 12 channel pairs, a total of 20 unique channels are left, which are identified and labeled on the electrode configuration map (highlighted in dark blue font against a light yellow background) in Fig. 8.

D. Results of Temporal Patterns and Classification with Spatio-temporal Features

In order to assess the effectiveness of spatio-temporal features from a reduced number of channels in distinguishing valence and arousal, we construct an RNN model to analyze the spatio-temporal features of the previously selected top 20 channels. For the construction of spatio-temporal features, we segment the BSP recording using a sliding window with a 10-heartbeat window and a 1-heartbeat step. We ultimately obtain a feature set with dimensions $31952 \times 10 \times 210$, where 31952 represents the sample count, 10 denotes the ten consecutive heartbeat features, and $20 \times (20 + 1)/2 = 210$ corresponds to the number of covariance values derived from the top 20 channels. When partitioning the test and validation sets, we maintain the same ratio as in the spatial classification experiment. Additionally, to prevent model overfitting, we apply the early stopping technique during the training process.



(a)

(b)

Fig. 10. Confusion matrix of (a) arousal and (b) valence classification with spatial-temporal features.

As demonstrated in Fig. 10, the proposed model achieves an accuracy of 99.6% for low arousal classification, 99.5% for high arousal classification, 99.5% for low valence classification, and 99.3% for high valence classification. This indicates that it is feasible to classify arousal and valence using only the top-ranked 20 channels BSP screened by MRMR. Furthermore, the classification accuracies obtained by constructing spatio-temporal features and employing an RNN model are slightly superior to the best results achieved by classifiers using spatial features solely from all BSP channels.

Additionally, to intuitively examine the temporal patterns from features, we concatenate flattened covariance matrices from 10 heartbeats in a single trial to form a 10×210 matrix, Q. Each row of Q is a 210-dimensional 1-D vector representing the flattened covariance matrices of individual heartbeats, and each column depicts the temporal covariance variation from a pair of BSP channels in a single trial. We calculated Q_{HV}, to represent the average of trials in high valence states, and Q_{LV} for trials in low valence states, without loss of generality. Likewise, Q_{HA} denotes the average of all trials in high arousal states, and Q_{LA}, those in low arousal states. To determine if these variations hold discriminative information, we use the Pearson correlation coefficient (PCC) to measure the correlation between columns.

In this work, we exclusively examine the comparison of each 4 channel pairs. We randomly select combinations of 4 channel pairs and compute both their intra-class PCC and inter-class PCC. For instance, we choose four columns from Q_{HV} and Q_{LV} that share the same channel pairs denoted as f. Subsequently, we compute the values of Γ, Λ, and Ω, each with its respective definition:

$$\Gamma = \phi(c_i^h, c_j^h), i, j \in f, i \neq j, \tag{15}$$

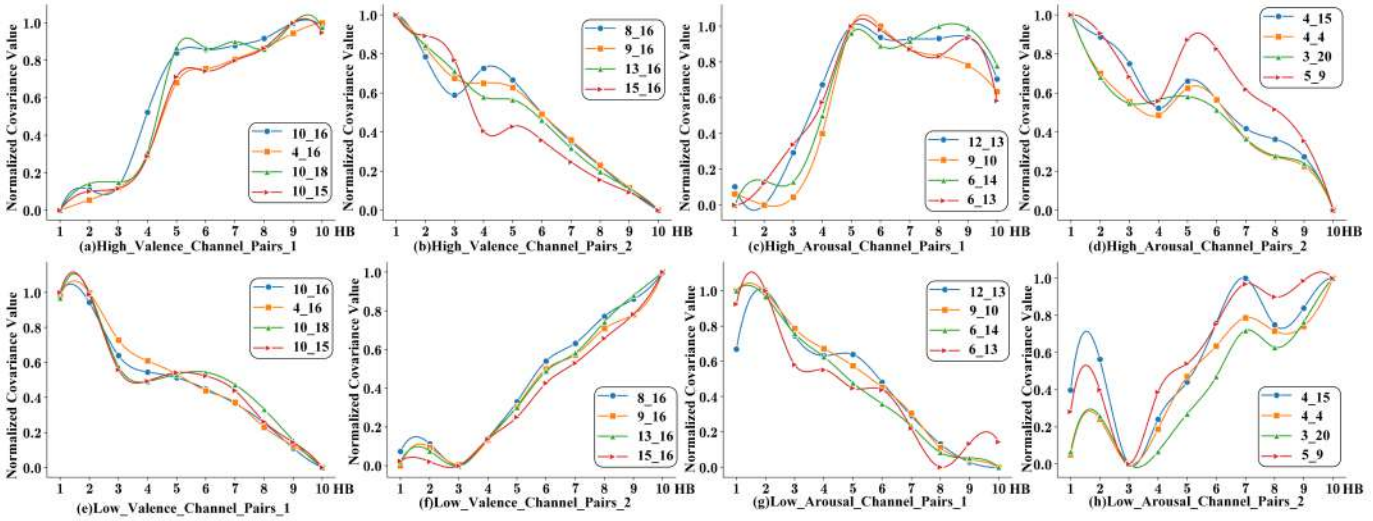


Fig. 11. Trajectories of normalized covariance values over time for 4 combinations of channel pairs. In each graph (a)-(h), the horizontal axis corresponds to ten consecutive heartbeats, and the vertical axis represents the normalized covariance values of the channel pairs. Each column contains two graphs that illustrate the trajectories of the same combination of channel pairs in two distinct states. For example, (a) and (e) compare the trajectories of 4 channel pairs (10_16, 4_16, 10_18, 10_15) between the high-valence and low-valence.

$$\Lambda = \phi(c_i^l, c_j^l), i, j \in f, i \neq j, \tag{16}$$

$$\Omega = \phi(c_k^h, c_k^l), k \in f. \tag{17}$$

Here, c_i^h represents the ith column in Q_{HV}, c_j^l represents the jth column in Q_{LV}, and $\phi(\cdot,\cdot)$ represents the PCC.

It's evident that Γ and Λ quantify intra-class similarity, and Ω quantifies inter-class similarity. Given the goal of our experiment, which is to explore whether spatio-temporal features contain distinctive temporal patterns, we seek channel pairs characterized by greater intra-class similarity and simultaneously less inter-class similarity. To accomplish this, we randomly select four combinations of channel pairs that meet the requirement that $\Gamma > 0.9$, $\Lambda > 0.9$, and $\Omega < 0.8$ simultaneously, two of them for comparing valence and two for comparing arousal. Fig.11 visually displays the time-varying covariance values of each combination of channel pairs across the two classes. To emphasize the temporal dynamics of the covariance values, we normalize them to a 0-1 scale in the figure, facilitating comparisons.

Temporal patterns visualization reveals that temporal trajectories are remarkably consistent within classes, yet they exhibit substantial distinctions between classes. This strongly indicates that the time-varying information within the covariance values of channel pairs encapsulates distinguishable emotional content. It also underscores the success of our spatio-temporal feature extraction and the rationality of constructing the RNN model.

V. CONCLUSION

In this study, we propose an affective peripheral patterns construction approach, with a specific focus on body surface potential's covariance features. We create a dataset with BSP recordings from 28 individuals, collected during exposure to 5 virtual reality (VR) videos. Following preprocessing, we compute the covariance matrix of BSP for each segment of heartbeats. To confirm the feasibility of using extracted spatial features in emotion classification tasks, we evaluate their performance in valence and arousal classification tasks, employing 12 different combinations of feature dimensionality reduction techniques and classifiers. Our results demonstrate the superiority of the SVM+PCA approach, achieving an impressive 98.6% accuracy in valence classification and a noteworthy 98.9% accuracy in arousal classification. Additionally, we project spatial features onto the body surface to create topographic patterns, revealing distinct patterns across valence and arousal dimensions.

Furthermore, considering the constantly changing nature of emotions, we extract spatio-temporal features of BSP and develop an RNN network to achieve fusion modeling across the spatial and temporal domains of the BSP's covariance features. It is noteworthy that emotion classification yields excellent results. Furthermore, the temporal patterns extracted from BSP reveal distinct variations across different arousal and valence dimensions. In conclusion, our findings indicate a correlation between emotional experiences and affective peripheral patterns derived from the inter-channel covariance of body surface potentials. This offers a novel perspective on affective computing grounded in physiological signals.

Although our work innovatively explores affective patterns through body surface potentials, it still has its limitations. This study uses five emotion-evoking materials, limiting the fine categorization of emotions. Our findings validate the method in terms of valence and arousal dimension respectively, but future research should include a wider range of emotional categories to explore affective patterns more comprehensively. Significantly, our research has provided new insights into exploring peripheral emotional patterns, potentially leading to more comprehensive studies on the synergistic changes in multimodal physiological signals' affective patterns.

REFERENCES

[1] A. R. Damasio and S. Sutherland, "Descartes' error: Emotion, reason and the human brain," *Nature*, vol. 372, no. 6503, pp. 287–287, 1994.

[2] J. Panksepp, *Affective neuroscience: The foundations of human and animal emotions*. Oxford university press, 2004.

[3] R. W. Picard, *Affective computing*. MIT press, 2000.

[4] J. S. Lerner, Y. Li, P. Valdesolo, and K. S. Kassam, "Emotion and decision making," *Annual review of psychology*, vol. 66, pp. 799–823, 2015.

[5] L. Nummenmaa, E. Glerean, R. Hari, and J. K. Hietanen, "Bodily maps of emotions," *Proceedings of the National Academy of Sciences*, vol. 111, no. 2, pp. 646–651, 2014.

[6] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, 2019.

[7] N. S. Suhaimi, J. Mountstephens, J. Teo *et al.*, "Eeg-based emotion recognition: A state-of-the-art review of current trends and opportunities," *Computational intelligence and neuroscience*, vol. 2020, 2020.

[8] E. H. Houssein, A. Hammad, and A. A. Ali, "Human emotion recognition from eeg-based brain–computer interface using machine learning: a comprehensive review," *Neural Computing and Applications*, vol. 34, no. 15, pp. 12 527–12 557, 2022.

[9] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from eeg," *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 417–429, 2017.

[10] M. S. Zitouni, C. Y. Park, U. Lee, L. J. Hadjileontiadis, and A. Khandoker, "Lstm-modeling of emotion recognition using peripheral physiological signals in naturalistic conversations," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 912–923, 2022.

[11] E. F. Pace-Schott, M. C. Amole, T. Aue, M. Balconi, L. M. Bylsma, H. Critchley, H. A. Demaree, B. H. Friedman, A. E. K. Gooding, O. Gosseries *et al.*, "Physiological feelings," *Neuroscience & Biobehavioral Reviews*, vol. 103, pp. 267–304, 2019.

[12] X. Shui, R. Lin, Z. Luo, B. Lin, X. Mao, H. Li, R. Liu, and D. Zhang, "Bodily electrodermal representations for affective computing," *IEEE Transactions on Affective Computing*, 2023.

[13] U. Provenzani, "Emotional processing and heart activity," *Brain and Heart Dynamics*, pp. 213–227, 2020.

[14] X. Chi, S. Wang, Z. Baloch, H. Zhang, X. Li, Z. Zhang, H. Zhang, Z. Dong, Y. Lu, H. Yu *et al.*, "Research progress on classical traditional chinese medicine formula lily bulb and rehmannia decoction in the treatment of depression," *Biomedicine & Pharmacotherapy*, vol. 112, p. 108616, 2019.

[15] W. Li, Z. Zhang, and A. Song, "Physiological-signal-based emotion recognition: An odyssey from methodology to philosophy," *Measurement*, vol. 172, p. 108747, 2021.

[16] X. Liao, E. Chang, X. Tang, I. Watanabe, R. Zhang, H.-W. Jeong, R. H. Adams, and M. K. Jain, "Cardiac macrophages regulate isoproterenol-induced takotsubo-like cardiomyopathy," *JCI insight*, vol. 7, no. 3, 2022.

[17] W. James, "Discussion: The physical basis of emotion." *Psychological review*, vol. 1, no. 5, p. 516, 1894.

[18] P. M. Niedenthal, "Embodying emotion," *science*, vol. 316, no. 5827, pp. 1002–1005, 2007.

[19] S. Davey, J. Halberstadt, and E. Bell, "Where is emotional feeling felt in the body? an integrative review," *Plos one*, vol. 16, no. 12, p. e0261685, 2021.

[20] H. D. Critchley and S. N. Garfinkel, "Interoception and emotion," *Current opinion in psychology*, vol. 17, pp. 7–14, 2017.

[21] J. L. Hazelton, S. Fittipaldi, M. Fraile-Vazquez, M. Sourty, A. Legaz, A. L. Hudson, I. G. Cordero, P. C. Salamone, A. Yoris, A. Ibañez *et al.*, "Thinking versus feeling: How interoception and cognition influence emotion recognition in behavioural-variant frontotemporal dementia, alzheimer's disease, and parkinson's disease," *Cortex*, vol. 163, pp. 66–79, 2023.

[22] N. Parrinello, J. Napieralski, A. L. Gerlach, and A. Pohl, "Embodied feelings-a meta-analysis on the relation of emotion intensity perception and interoceptive accuracy," *Physiology & Behavior*, p. 113904, 2022.

[23] B. Hsueh, R. Chen, Y. Jo, D. Tang, M. Raffiee, Y. S. Kim, M. Inoue, S. Randles, C. Ramakrishnan, S. Patel *et al.*, "Cardiogenic control of affective behavioural state," *Nature*, vol. 615, no. 7951, pp. 292–299, 2023.

[24] E. De Longis, G. Alessandri, and C. Ottaviani, "Inertia of emotions and inertia of the heart: Physiological processes underlying inertia of negative emotions at work," *International Journal of Psychophysiology*, vol. 155, pp. 210–218, 2020.

[25] D. G. Leo, H. Ozdemir, D. A. Lane, G. Y. Lip, S. S. Keller, and R. Proietti, "At the heart of the matter: How mental stress and negative emotions affect atrial fibrillation," *Frontiers in Cardiovascular Medicine*, vol. 10, p. 1171647, 2023.

[26] A. Bhide, R. Durgaprasad, L. Kasala, V. Velam, and N. Hulikal, "Electrocardiographic changes during acute mental stress," *International Journal of Medical Science and Public Health*, vol. 5, no. 5, p. 835, 2016.

[27] W. Li, "Finding needles in a haystack: Recognizing emotions just from your heart," *IEEE Transactions on Affective Computing*, no. 01, pp. 1–1, 2021.

[28] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biological psychology*, vol. 84, no. 3, pp. 394–421, 2010.

[29] H. Sequeira, P. Hot, L. Silvert, and S. Delplanque, "Electrical autonomic correlates of emotion," *International journal of psychophysiology*, vol. 71, no. 1, pp. 50–56, 2009.

[30] E. Duffy, "The concept of energy mobilization." *Psychological Review*, vol. 58, no. 1, p. 30, 1951.

[31] B. J. Li, J. N. Bailenson, A. Pines, W. J. Greenleaf, and L. M. Williams, "A public database of immersive vr videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures," *Frontiers in psychology*, vol. 8, p. 2116, 2017.

[32] M. Li, J. Pan, Y. Gao, Y. Shen, F. Luo, J. Dai, A. Hao, and H. Qin, "Neurophysiological and subjective analysis of vr emotion induction paradigm," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 11, pp. 3832–3842, 2022.

[33] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[34] M. Mneimneh, E. Yaz, M. Johnson, and R. Povinelli, "An adaptive kalman filter for removing baseline wandering in ecg signals," in *2006 Computers in Cardiology*. IEEE, 2006, pp. 253–256.

[35] T. M. Inc., "Matlab version: 9.13.0 (r2022a)," Natick, Massachusetts, United States, 2022. [Online]. Available: https://www.mathworks.com

[36] W. Hao, Y. Chen, and Y. Xin, "Ecg baseline wander correction by mean-median filter and discrete wavelet transform," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 2712–2715.

[37] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE transactions on biomedical engineering*, no. 3, pp. 230–236, 1985.

[38] P. Kuppens and P. Verduyn, "Emotion dynamics," *Current Opinion in Psychology*, vol. 17, pp. 22–26, 2017.

[39] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.

[40] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

Wei Wu is currently working toward a master's degree with the School of Biomedical Engineering, Sun Yat-Sen University, China. His research interests include affective computing and physiological signal analysis.

Yao Pi is currently working toward a PhD degree with the School of Biomedical Engineering, Sun Yat-sen University, China. His research interests include affective computing and pattern recognition.

Xianbin Zhang received his master's degree in biomedical engineering from Sun Yat-sen University, Shenzhen, China, in 2021. He is currently pursuing a doctor of engineering degree in biomedical engineering from Sun Yat-sen University. His research interests include electrocardiogram monitoring, wearable systems, and collaborative intelligence.

Lin Xu obtained her bachelor's degree in clinical medicine from the Fourth Military Medical University, Xi'an, China, in 1999, and master's degree in internal medicine-cardiology from the Fourth Military Medical University, Xi'an, China, in 2003, and a Ph.D. from the Department of Clinical Medicine at Southern Medical University, Guangzhou, China, in 2006. Now, she is a medical discipline leader and department director in the General Hospital of the Southern Theatre Command. Her research interests are the identification and treatment techniques of cardiovascular diseases, cardiovascular rehabilitation, multi-modal medical image fusion, wearable sensing technology, and artificial intelligence-assisted prediction of cardiovascular diseases. She has totally published more than 90 papers in high-impact journals and relevant conferences and has secured over 6 major competitive research grants.

Wanqing Wu received his bachelor's degree in computer science and technology from Hunan Normal University, Hunan, China, in 2004, and master's degree in computer science from Chongqing University, Chongqing, China, in 2007, and a Ph.D. from the Pusan National University of Computer Engineering, Korea, in 2013. He is currently an Associate Professor at the School of Biomedical Engineering, Sun Yat-Sen University, Shenzhen, China. His research interests include wearable sensing technology, biomedical signal sensing and processing, body sensor networks, biofeedback, biometric security, the Internet of Things, and multimodal medical data fusion. He has published over 60 papers in renowned journals and conferences, mostly in the area of wearable sensor technology, wireless body sensor networks, and the Internet of Medical Things, and has also contributed to the peer review of 10 journals. He has secured over 15 major competitive research grants.