Physiological Signals based Day-Dependence Analysis with Metric Multidimensional Scaling for Sentiment Classification in Wearable Sensors

Wei Wang, Xiaodan Huang, Jijun Zhao & Yanguang Shen

School of Information & Electrical Engineering, Hebei University of Engineering, Hebei Handan 056038, China
Email: wangwei83@hebeu.edu.cn

Abstract. Affective recognition has emerged in implicit human-computer interaction. Given the physiological signals in the process of affective recognition, the different positions where the physiological signal sensors are installed on the body, along with the daily habits and moods of human beings, influence the affective physiological signals. The scalar product matrix was calculated in this study, based on metric multidimensional scaling with a dissimilarity matrix. Subsequently, the matrix of individual attribute reconstructs was obtained using the principal component factor. The method proposed in this study eliminates day dependence, reduces the effect of time on the physiological signals of the affective state, and improves the accuracy of sentiment classification.

Keywords: day dependence; decision making; health management; physiological signal; psychology; sentiment classification.

1 Introduction

To realize natural human-computer interaction, more personalized services are required during the interactive process. Therefore, it is necessary to find an effective way of aligning the affection of a user with the body sensor network connected to the service. Artificial psychological and affective computing is a new research direction in the field of artificial intelligence [1]. Related studies have generally focused on simulating affective recognition, modeling and expression. Affective state recognition consists of two types, namely recognizing the affection of the user and the affection in content, such as light effect, context, and characters in a film and music and painting [2].

For the first type, studies use varied types of sensors to measure the affection of the user. Kim surveyed the amount of sweat, pulse, body temperature, and blood pressure of users, and determined their affective states after data fusions and fuzzy inferences based on the collected physiological information [3]. Nguyen used fewer sensors (i.e. only pulse and body temperature) to acquire affective physiological data [4]. Arapakis proposed a method using facial
expression [5]. However, there are also methods that indirectly measure affection with a questionnaire or an extension thereof. For instance, Gustavo González used an emotional quotient test table [6]. Ai Thanh Ho provided a group of different colors from which the user could select; he determined the affective state of the user based on the selection made [7]. The features of these affective state determination methods are compared in Table 1.

<table>
<thead>
<tr>
<th>Method</th>
<th>Portability</th>
<th>Continuity</th>
<th>Sensibility</th>
<th>Comfort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression</td>
<td>Poor</td>
<td>Poor</td>
<td>General</td>
<td>General</td>
</tr>
<tr>
<td>Voice</td>
<td>General</td>
<td>Poor</td>
<td>General</td>
<td>Good</td>
</tr>
<tr>
<td>Eye movement</td>
<td>General</td>
<td>General</td>
<td>General</td>
<td>Poor</td>
</tr>
<tr>
<td>Brainwave</td>
<td>Poor</td>
<td>Good</td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>Behavior</td>
<td>Poor</td>
<td>Poor</td>
<td>General</td>
<td>General</td>
</tr>
<tr>
<td>Physiological information</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>General</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>General</td>
<td>Poor</td>
<td>General</td>
<td>Poor</td>
</tr>
</tbody>
</table>

The method of using physiological information from the user to determine the user’s affective state is better than the other methods in terms of portability, signal continuity, sensibility, and comfort (Table 1). Therefore, studies on affection divisibility have attracted the interest of researchers in the field of human-computer interaction. Several psychologists have obtained different conclusions regarding the autonomic nervous system (ANS), such as the “detest” affection specified in a previous work [8]. These researchers also offer different views [9]. However, the existence of affection-specific ANS has gained increasing support from recent studies [8-15]. Some affective physiological data sample libraries have been established in the research on affection recognition based on physiological signals (Table 2). Although these libraries provide a foundation for comparison and analysis of affective physiological feature selection and sentiment classification, the data are easily influenced. Due to different positions where the physiological sensors are located on the body and the varied daily habits and moods of human beings, physiological samples from the same day correlate with one another. This condition is called day dependence. How strong the correlation is and how to eliminate it is worth studying. Consistent with the affective physiological database proposed in [16-18], the current study analyzes the day dependence and handling method of 20 groups of samples with 8 types of affections. In Table 2, EMG, ECG and EEG mean electromyogram, electrocardiogram and electroencephalogram respectively.
Table 2  Library of affective physiological samples.

<table>
<thead>
<tr>
<th>Researcher</th>
<th>Inducing method</th>
<th>Affection types</th>
<th>Physiological signals</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picard, et al. [18]</td>
<td>Image</td>
<td>Anger, hate, grief, platonic love, romantic love, joy and reverence</td>
<td>EMG, pulse, skin conductance, respiratory</td>
<td>20 groups of a single participant</td>
</tr>
<tr>
<td>Kim, et al. [19]</td>
<td>Audio and video</td>
<td>Sadness, depression, anger and surprise</td>
<td>ECG, pulse rate, skin conductance, ECG, skin temperature, skin conductance ECG, skin temperature, skin conductance, respiratory</td>
<td>175 groups of 125 participants and 50 participants</td>
</tr>
<tr>
<td>Li, et al. [20]</td>
<td>Movie clips</td>
<td>Fear, happiness, easily</td>
<td>ECG, skin temperature, skin conductance, respiratory, EMG, skin conductance, respiratory</td>
<td>55 participants</td>
</tr>
<tr>
<td>Kim, et al. [21]</td>
<td>Music</td>
<td>Happy, anger, sadness, joy</td>
<td>Skin conductance, heart rate, pulse, ECG, respiratory, facial electromyography and frontal lobe EEG</td>
<td>90 groups of three participants</td>
</tr>
<tr>
<td>Liu, et al. [17]</td>
<td>Movie clips</td>
<td>Happy, surprise, disgust, sadness, anger and fear</td>
<td></td>
<td>300 participants</td>
</tr>
</tbody>
</table>

2  Affective Physiological Samples

The affection database used in this study, DataSet I, was established by Picard, et al. [18] and contains recorded physiological data of 20 days. The samples include 4 types of data: EMG, pulse, skin conductance, and respiration. The affective states have 8 classes: Calm (N), Anger (A), Hate (H), Grief (G), Platonic love (P), Romantic love (L), Joy (J), and Reverence (R). The signals were sampled in 100 seconds. The length of the time series was $N = 2001$. Thus, the number of daily physiological data is $8 \times 4 \times 2001 = 64032$ in total. The organization of affection database DataSet I is shown in Figure 1.

DataSet I contains the first data set generated as part of the MIT Affective Computing Group’s research. The database consists of measurements of 4 physiological signals and 8 affective states, taken once a day in a session lasting around 25 minutes, for recordings from an individual over 20 days, while trying to keep every other aspect of the measurements constant. Based on skin-surface sensing, the data were obtained from a real body sensor network, to be specific, from EMG, pulse, skin conductance, respiration sensors used in medical devices. The data were recorded when showing images to the participant in order to experimentally manifest the participant’s changing affections.
3 Analysis of Day Dependence

Physiological signals are used in the affective state discriminating process. The training samples and selected features directly affect the classifier in feature-based classification methods. Considering the deployment of sensors in a body sensor network and the daily lives of the participants, training samples are highly correlated in the same day and week among similar affective states. This influences the classifying process. Therefore, after selecting a certain feature set, day-dependence analysis for affective physiological samples is one of the key factors identified in training the affective state classifier. The current study illustrates a general day-dependence analysis method based on dimensionality reduction algorithms using the dataset proposed in [18].

There are many algorithms for dimensionality reduction, such as principal components analysis (PCA), independent component analysis (ICA), and the multidimensional scaling method (MDS). These are all classical dimensionality reduction approaches, which are similar in expressing observed data with a smaller number of dimensions. In addition, they only characterize linear subspaces in the data. Each new data is a linear combination of the observed data, reflecting the combined effect of the observed data. The effectiveness of PCA is limited by its global linearity and Gauss distribution [22]. The ICA method transforms the observed data into a linear combination of non-Gaussian distribution data. It differs from PCA in the way that it identifies and models the subspace. Statistical independence is an important hypothesis while using ICA [23]. MDS is closely related to PCA. However, with MDS, the distances of the transformed data in the low-dimensional space are consistent with the similarity of the observed data in the high-dimensional space. MDS attempts to preserve pair-wise distances [24]. Therefore, the MDS method is a suitable approach to solve the problem presented in this study. The similarity or dissimilarity of the affective physiological signals of the samples in the low-dimensional space
reveals the underlying structures of the samples, which are then used to analyze day dependence [25]. The MDS method was employed in this study because it can reduce day dependence.

The literature in psychophysiology and related disciplines describes various methods of recognizing affective states based on facial muscle movement, heart rate changes, and so on [11]. Picard, et al. [18] proposed 6 static characteristics, including mean value $\mu_i$, standard deviation $\sigma_i$, mean value of the first-order differential absolute value $\delta_i$, its normalized value $\tilde{\delta}_i$, mean value of the second-order differential absolute value $\gamma_i$, and its normalized value $\tilde{\gamma}_i$ [18].

Suppose the $i^{th}$ $i \in \{1, 2, 3, 4\}$ type of physiological signal acquiring devices takes the $n^{th}$ $n \in \{1, 2, \ldots, N\}$ measured value $X_n^i$, and its normalized value is $\tilde{X}_n^i$. The 6 static features can be obtained using Formulas (1) to (6), given respectively by the following:

$$\mu_i = \frac{1}{N} \sum_{n=1}^{N} X_n^i$$  \hspace{1cm} (1)

$$\sigma_i = \left[ \frac{1}{N-1} \sum_{n=1}^{N} (X_n^i - \mu_i)^2 \right]^{1/2}$$  \hspace{1cm} (2)

$$\delta_i = \frac{1}{N-1} \sum_{n=1}^{N} |X_{n+1}^i - X_n^i|$$  \hspace{1cm} (3)

$$\tilde{\delta}_i = \frac{1}{N-1} \sum_{n=1}^{N} |\tilde{X}_{n+1}^i - \tilde{X}_n^i| = \frac{\delta_i}{\sigma_i}$$  \hspace{1cm} (4)

where $\tilde{X}_n^i = \frac{X_n^i - \mu_i}{\sigma_i}$.

$$\gamma_i = \frac{1}{N-2} \sum_{n=1}^{N-2} |X_{n+2}^i - X_n^i|$$  \hspace{1cm} (5)

$$\tilde{\gamma}_i = \frac{1}{N-2} \sum_{n=1}^{N-2} |\tilde{X}_{n+2}^i - \tilde{X}_n^i| = \frac{\gamma_i}{\sigma_i}$$  \hspace{1cm} (6)

Based on the abovementioned features, the time series acquired by the $i^{th}$ type of physiological signal acquiring devices can be described with $F_i = [\mu_i \quad \sigma_i \quad \delta_i \quad \tilde{\delta}_i \quad \gamma_i \quad \tilde{\gamma}_i]$. Thus, the feature vector of the $j^{th}$ $j \in \{1, 2, \ldots, 8\}$
class at the $k$th day is denoted as $\tilde{S}_j = [\tilde{F}_1 \tilde{F}_2 \tilde{F}_3 \tilde{F}_4] \in \mathbb{R}^{1 \times 24}$. This dataset contains 8 types of affective state data collected within 20 days. To discuss the difference among these data (i.e. day dependence), an affective state feature matrix $S$ is established with $\tilde{S}_j$. $S$ is composed of 8 types of affective eigenvectors at the $k$th day and $i$th type of affective eigenvectors in the remaining 19 days, and is given by:

$$S = [\tilde{S}_1 \tilde{S}_2 \ldots \tilde{S}_8 \tilde{S}_j \tilde{S}_j \tilde{S}_{j+1} \tilde{S}_{j+2} \ldots \tilde{S}_{20}]^T \in \mathbb{R}^{27 \times 24}.$$ 

A dissimilarity matrix $\Delta_{L \times L}$ with affective state feature matrix, and especially $L = 27$, based on multidimensional scaling theory [24], was created in this study. Here, $\sigma_{mn}$ in $\Delta$ can be calculated as follows:

$$\sigma_{mn} = d_{mn} = [(\tilde{S}_m - \tilde{S}_n)(\tilde{S}_m - \tilde{S}_n)^T]^{1/2} = \left[\sum_n (\tilde{S}_{mp} - \tilde{S}_{np})^2\right]^{1/2} \quad \forall m, n \in \{1, 2, \ldots, L\}, p \in \{1, 2, \ldots, 24\}$$

The inner product matrix $\Gamma$ is then calculated. Its element $r_{ij}$ is as follows:

$$r_{ij} = -0.5 \times (\sigma_{mn}^2 - \sigma_{mn}^2 - \sigma_{mn}^2 + \sigma_{mn}^2)$$

where

$$\sigma_{mn}^2 = \frac{1}{L} \sum_m \sigma_{mn}^2,$$

$$\sigma_{mn}^2 = \frac{1}{L} \sum_m \sigma_{mn}^2,$$

$$\sigma_{mn}^2 = \frac{1}{L} \sum_m \sigma_{mn}^2.$$

The reconstructed matrix in the low-dimensional space is designated as $\hat{\Omega}_{M \times M}$ ($M < 24$), which can be obtained through the affective state feature matrix. Correspondingly, its affective state dissimilarity matrix is recorded as $\hat{D}$. Based on the theory of metric multidimensional scaling, $\hat{D}$ is similar to $D$ in a sense. Then, we can assume that:

$$\Gamma = \hat{\Omega} \hat{\Omega}^T$$

The affective state feature reconstructed matrix in the low-dimensional space $\hat{\Omega}$ can be obtained using Eq. (9). Thus, the affective state feature can be presented in the low-dimensional space. At the same time, it is easier to obtain eigenvalues $\hat{\lambda}_j$, which correspond to the $j$th dimension of $\hat{\Omega}$, using the following expression:
\[ \lambda_j = \sum_i \hat{\psi}_{ij}^2 \]  

(10)

where \( \hat{\psi}_{ij} \) are elements of \( \hat{\Omega} \).

4 Experiment

4.1 Day Dependence Problem

Eight types of affective eigenvectors from the \( k \)th day and the \( i \)th types of affective eigenvectors in the remaining 19 days should be obtained to form a matrix given by:

\[ S = [\hat{\xi}^1_k \hat{\xi}^2_k \ldots \hat{\xi}^8_k \hat{\xi}^j_k \hat{\xi}^2_{k-1} \hat{\xi}^j_{k-1} \hat{\xi}^2_{k-2} \ldots \hat{\xi}^2_{24}] \in \mathbb{R}^{12 \times 24}. \]

The day dependence of samples measured by the body sensor network should be analyzed after reconstructing the affective state feature matrix in the low-dimensional space based on the abovementioned theory of metric multidimensional scaling. In this study, each affective eigenvector is reconstructed in the two-dimensional space (Figure 2) based on Kruskal’s stress requirements.

(a) Eight types of affective eigenvectors at day 1 and the Anger affective eigenvectors in the remaining 19 days.

(b) Eight types of affective eigenvectors at day 1 and the Hate affective eigenvectors in the remaining 19 days.

Figure 2 Restructuring graphs of affective eigenvectors (with day dependence).
Physiological Signals for Sentiment Classification

(c) Eight types of affective eigenvectors at day 1 and the Grief affective eigenvectors in the remaining 19 days.

(d) Eight types of affective eigenvectors at day 2 and the Anger affective eigenvectors in the remaining 19 days.

(e) Eight types of affective eigenvectors at day 3 and the Anger affective eigenvectors in the remaining 19 days.

(f) Eight types of affective eigenvectors at day 4 and the Anger affective eigenvectors of the remaining 19 days.

Figure 2 (continued) Restructuring graphs of affective eigenvectors (with day dependence).
The left parts of Figures 2(a) to (c) are the restructuring graphs of the affective state feature matrix $S$ in the two-dimensional space, which consist of 8 classes of affective eigenvectors at the first day and the emotional state eigenvectors of the Anger, Hate, and Grief affective states in the remaining 19 days. The left part of Figure 2(c) shows the reconstructed points of the 8 types of affection [Calm (N), Anger (A), Hate (H), Grief (G), Platonic love (P), Romantic love (L), Joy (J), and Reverence (R)] at day 1, located in the left half plane of the coordinate system; as can be seen, their distances are quite small. However, the reconstructed point distances of the Grief affection at the first day and in the remaining 19 days are larger, thus demonstrating that the affective eigenvectors at the same day have day dependence. In addition, the analyses found in the left parts of Figures 2(a) and (b) are similar.

The left parts of Figures 2(d) to (f) are the restructuring graphs of the affective state feature matrix $S$ in the two-dimensional space, which consist of 8 classes of affective eigenvectors at the $k^{th}$ ($k = 2, 3, 4$) day and the affective eigenvectors of the Anger affective state in the remaining 19 days. The left part of Figure 2(f) shows the reconstructed points of the 8 types of affections [Calm (N), Anger (A), Hate (H), Grief (G), Platonic love (P), Romantic love (L), Joy (J), and Reverence (R)] on day 4 located in the left half plane and near the x-axis of the coordinate system; as can be seen, their distances are quite small. However, the reconstructed point distances of the Anger affection at the fourth day and in the remaining 19 days are larger, thus demonstrating that the affective eigenvectors at the same day also have day dependence. In addition, the analyses in the left parts of Figures 2 (d) and (e) are similar.

The right part of Figure 2 shows the linear fit scatter plot of the reconstruction process. The x-axis reveals the disparity of the samples, whereas the y-axis reflects the fitting distance. The fitting coefficient $R^2 = 0.999$ is good enough. Young’s S-stress, Kruskal’s stress, and the determination coefficient RSQ are summarized in Table 3, which shows the proportion of explaining by relative spatial distance is larger.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Young’s S-stress</th>
<th>Kruskal’s stress</th>
<th>Determination coefficient RSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2(a)</td>
<td>0.01889</td>
<td>0.07515</td>
<td>0.99082</td>
</tr>
<tr>
<td>2(b)</td>
<td>0.03033</td>
<td>0.04907</td>
<td>0.98887</td>
</tr>
<tr>
<td>2(c)</td>
<td>0.05533</td>
<td>0.06360</td>
<td>0.98169</td>
</tr>
<tr>
<td>2(d)</td>
<td>0.01551</td>
<td>0.07418</td>
<td>0.99152</td>
</tr>
<tr>
<td>2(e)</td>
<td>0.01840</td>
<td>0.06677</td>
<td>0.99232</td>
</tr>
<tr>
<td>2(f)</td>
<td>0.01279</td>
<td>0.05289</td>
<td>0.99498</td>
</tr>
</tbody>
</table>
Based on the abovementioned statement, the day dependence among the affections is derived from three aspects, namely, (1) different positions by which physiological signal sensors are installed on the body; (2) different daily habits of human beings, such as hand washing or walking, which affect physiological data; and (3) different everyday moods of human beings based on psychology.

4.2 Eliminating the Correlation Problem

Considering the main source of the day dependence of affective eigenvectors, a method to eliminate correlation is proposed in this article. With the calm affective state \( N \) as the baseline, this study only focuses on the changes between the calm state \( N \) and the rest of the affective states to handle the day dependence problem.

![Restructuring graphs of affective eigenvectors (without day dependence).](image)

(a) Eight types of affective eigenvectors at day 1 and the Anger affective eigenvectors in the remaining 19 days.

(b) Eight types of affective eigenvectors at day 1 and the Hate affective eigenvectors in the remaining 19 days.

**Figure 3** Restructuring graphs of affective eigenvectors (without day dependence).

The affective feature matrix \( S \) with relative values of affective eigenvectors \( \tilde{S}^i_j - \tilde{S}^i_k \) is constructed. The day dependence based on Eq. (7) to (10) should be analyzed while considering the theory of metric multidimensional scaling.
Affective eigenvectors without day dependence are reconstructed in the two-dimensional space (Figure 3).

Comparing the left part of Figure 3(a) with that of Figure 2(a) in order to restructure the points of the Anger affection with day dependence, it can be seen that the distances between day 1 and the remaining days are large. However, after eliminating the dependence, the distances become smaller. This result demonstrates that the handling method is effective in enhancing the affection class correlation. Similar results can be obtained by comparing the left part of Figure 3(b) with that of Figure 2(b).

The right part of Figure 3 shows the linear fit scatter plot of the reconstruction process; as can be seen, the fitting coefficient $R^2$ is closer to 1. Young’s S-stress, Kruskal’s stress, and the determination coefficient RSQ are summarized in Table 4.

<table>
<thead>
<tr>
<th>Figure</th>
<th>Young’s S-stress</th>
<th>Kruskal’s stress</th>
<th>Determination coefficient RSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>3(a)</td>
<td>0.03017</td>
<td>0.04326</td>
<td>0.99138</td>
</tr>
<tr>
<td>3(b)</td>
<td>0.04123</td>
<td>0.05415</td>
<td>0.98755</td>
</tr>
</tbody>
</table>

5 Conclusion

The different positions where the physiological signal sensors of the body sensor network are located, along with the different daily habits and moods of human beings, influence the training samples. Determining how large the effects are and how to handle them is worth studying. In this study, the affective physiological database DataSet I was considered. The effect called day dependence of 20 groups of samples that belong to 8 types of affections was analyzed. The results show that day dependence was widespread in the samples. One method to eliminate day dependence and verification of its effectiveness by metric multidimensional scaling was proposed and discussed in this study.

Acknowledgements

This work was supported by the Foundation for Young Scholars of Hebei Educational Committee (QN20131152).

References


