International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 26, No. 4 (2012) 1250008 (18 pages)

DOI: 10.1142/S0218001412500085

STATISTICAL PREDICTION OF EMOTIONAL STATES BY PHYSIOLOGICAL SIGNALS WITH MANOVA AND MACHINE LEARNING

TUNG-HUNG CHUEH

Green Energy and Environment Research Laboratories
Industrial Technology Research Institute
Chutung, Hsinchu 310, Taiwan, R. O. C.

TAI-BEEN CHEN

Department of Medical Imaging and Radiological Sciences
I-Shou University Kaohsiung 824, Taiwan, R. O. C.

HENRY HORNG-SHING LU

Institute of Statistics, National Chiao Tung University
Hsinchu 300, Taiwan, R. O. C.
hslu@stat.nctu.edu.tw

SHAN-SHAN JU, TEH-HO TAO and JIUNN-HAUR SHAW

Center for Measurement Standards
Industrial Technology Research Institute
Hsinchu 300, Taiwan, R. O. C.

Received 20 April 2009
Accepted 7 June 2012
Published 31 August 2012

For the importance of communication between human and machine interface, it would be valuable to develop an implement which has the ability to recognize emotional states. In this paper, we proposed an approach which can deal with the daily dependence and personal dependence in the data of multiple subjects and samples. 30 features were extracted from the physiological signals of subject for three states of emotion. The physiological signals measured were: electrocardiogram (ECG), skin temperature (SKT) and galvanic skin response (GSR). After removing the daily dependence and personal dependence by the statistical technique of MANOVA, six machine learning methods including Bayesian network learning, naive Bayesian classification, SVM, decision tree of C4.5, Logistic model and K-nearest-neighbor (KNN) were implemented to differentiate the emotional states. The results showed that Logistic model gives the best classification accuracy and the statistical technique of MANOVA can significantly improve the performance of all six machine learning methods in emotion recognition system.

Keywords: Emotion recognition; physiological signals; machine learning; daily effect; MANOVA.

1. Introduction

Research efforts between human and machine interfaces have developed over decades. With the progress of science and technology, the capability and functionality of automatic machines have rapidly progressed and improved. Even so, there are still challenges to design machines that can recognize human emotional states correctly. Essentially, a change in human emotional states will influence a lot of external behaviors and physiological characteristics, including facial expression, intonation of speech, gesture, posture, eye expression, blood pressure, heart beat, skin resistance and so forth. Hence, it is an important research topic to use these characteristics to detect the states of human emotion.

In the field of human and machine interaction, it would be valuable to develop an instrument capable of recognizing a person's emotional status. Emotion recognition has become a critical investigation in emotional intelligence and can be applied in many systems. In 1999, Ark *et al.*[2] at the laboratory of IBM established a mouse that can distinguish a user's affective states with 75% accuracy. A robot with the ability to recognize and determine the underlying emotion of a person can interact with humans using signals in human speech and facial expression.[3,17] Moreover, other applications such as driving safety, training and telemedicine also can implement an emotion recognition system to benefit users.[18]

In previous research about developing an emotion recognition system, features of facial expressions are most commonly used as the determinant attribute and have successfully obtained fairly high rates in emotion recognition.[6,8,9,25,26] Besides, there are also studies employing signals of speech and vocal intonations to recognize states of emotion.[7,19] Combining facial and voice expression has also been used in distinguishing affective emotional states recently.[4] However, these two characteristics are sometimes hardly recorded if the subject is moving. Therefore, recognizing emotion using physiological signals, which can be recorded for a moveable subject, is a critical study.

In the study of affective physiological states, Picard *et al.*[21] at MIT Media Laboratory have tried to differentiate eight different emotions of a single person using physiological characteristics recorded every day over six weeks, resulting in an 81% overall classification accuracy rate by using a hybrid method involving sequential floating forward search and Fisher projection. For handling the physiological signals with short-term segments, Kim *et al.*[14] proposed an algorithm to detect emotional statuses based on their experimental psychosomatic responses for multiple subjects and got the correct classification rate of 78.4% by the machine learning method of support vector machine (SVM). Nasoz *et al.*[18] employed three classification methods to discriminate six different emotional states from physiological signals collected via noninvasive technologies. Rani *et al.*[23] have applied four different classification methods to determine affective states from physiological signals and have made comparisons of these methods.

Among emotion recognition studies, there are typically two approaches: one against one[21] and one against all.[14] For the one against one approach, we can collect the labeled psychosomatic signals of a single subject on multiple observations and learn a trainer model out of the same person so that we can decipher the unknown emotional states of that person as a test of his (her) physiological signals. Though it has the benefit of removing the inter-subject difference for subject-based learning, this approach can only recognize one subject's emotion. Alternatively, we can measure the physiological signals of emotion from multiple subjects and learn a trainer model out of them. Hence, we can distinguish other people's emotion status using this system. In practice, this user-independent system is believed to be more convenient in the field of emotional recognition studies. However, the assumption of independence between physiological signals and subjects is not reasonable nor practical.

Furthermore, daily physiological signals can vary even for the same state of emotion. The daily effect could be removed using the statistical technique of multivariate analysis of variance (MANOVA). Then, typical machine learning methods could be applied to discriminate and predict the emotional state. Hence, the purpose of this work is to advance the improvement of emotion recognition by eliminating inter-subject differences and removing the daily effects by MANOVA with statistical machine learning.

Physiological signals including skin temperature variation (SKT), galvanic skin response (GSR) and electrocardiogram (ECG) were implemented in this study. These physiological signals can be measured conveniently without any annoying sensors attached on the face or scalp. The subjects would induce three different emotional statuses by themselves: anger, joy and neutral. Besides, we would use the techniques of MANOVA and six different classification methods to discriminate various states of emotion.

This report is organized as follows. We present the procedures of data collection and features extraction from the measured physiological signals in Sec. 2. In Sec. 3, we discuss the problem of day-effects and remove daily effects using MANOVA. Then, we consider six classification methods in Sec. 4: Bayesian network learning, naive Bayesian classification, SVM, decision tree of C4.5, logistic model and K-nearest neighbor (KNN). Finally, we report the analysis results and discuss future works.

2. Data Preprocessing

2.1. *Data collections*

In the research of emotion recognition, the collection of physiological signals plays a important role for next analysis. In this study, the database of physiological signals and corresponding emotional states were collected and obtained from the Center for Measurement Standards of the Industrial Technology Research Institute (ITRI) in Taiwan.

The first group included two subjects, Jane and Alice; they are both female and in their twenties. Every morning between 8:30 am and 10:00 am, they were invited to our laboratory. They were asked to feel a neutral emotion for 200 s first, followed by an emotion of anger for at least 120 s and finish with a emotion of joy for at least 120 s. Meanwhile, those physiological signals were measured and recorded by MP100 system in BIOPAC (http://www.biopac.com). Regarding the approach in eliciting emotion, the method we used is similar to the efforts pioneered by Picard *et al.*[21] with a slight modification. The methodology is subject-elicited instead of event-elicited, open-recording and emotion-purpose. To prevent differences caused by different external stimulations on different days, we do not rely on any auxiliaries to arouse the emotions of subjects. The subjects were simply asked to feel an emotion without any assistance such as movies, voices or any other outer stimulus; namely, we do not employ a rigorous Clynes protocol as Picard *et al.* did. Data gathered from 11 days were used in this study. The default sampling rates were 256 points in 1 s for each state of emotion. An example of every emotional state is given in Figs. 1–3.

For the second dataset, the subjects we used were adults: five men and five women aged from 20 to 30 years. Every morning the subjects were invited to our laboratory with controlled temperature and humidity. At the first practice, the subjects were in a dark place without any voice or music for eliciting a neutral mood within 4 min. Then, to begin the negative emotion eliciting stage, the subjects received eight different pictures with negative expressions, and each picture was broadcasted for 30 s. At the same time, the subjects were asked to feel the negative emotion under the stimulus of pictures. Then, the positive emotion eliciting stage was implemented

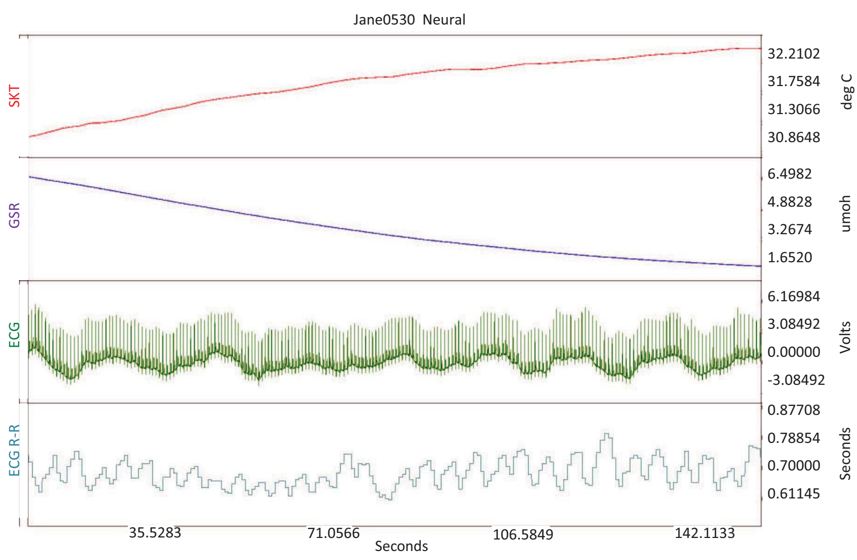

Fig. 1. Three physiological signals were recorded when the subject was asked to feel neutral. From top to bottom: skin temperature variation (SKT), galvanic skin response (GSR) and electrocardiogram (ECG). The physiological signals were sampling at 256 samples for every second and the measured times were 200 s.

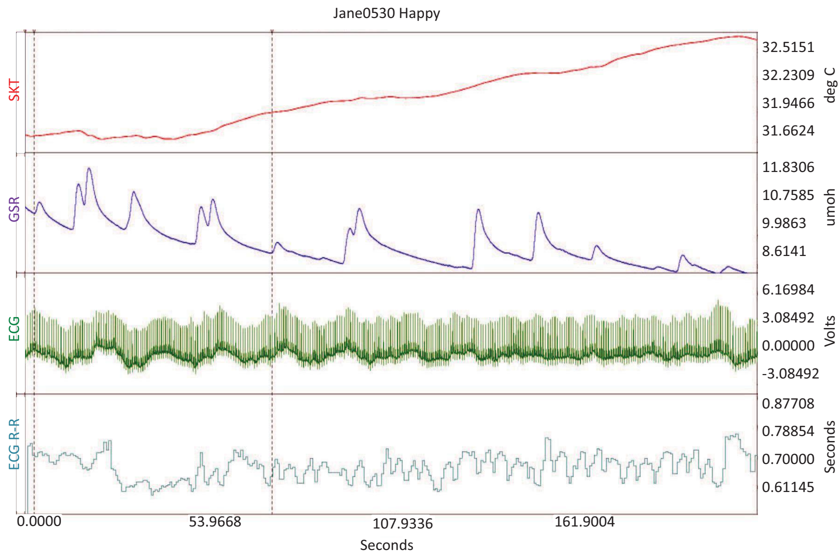

Fig. 2. Three physiological signals were recorded when the subject was asked to feel joyful. The physiological signals were sampling at 256 samples for every second and the measured times were 120 s.

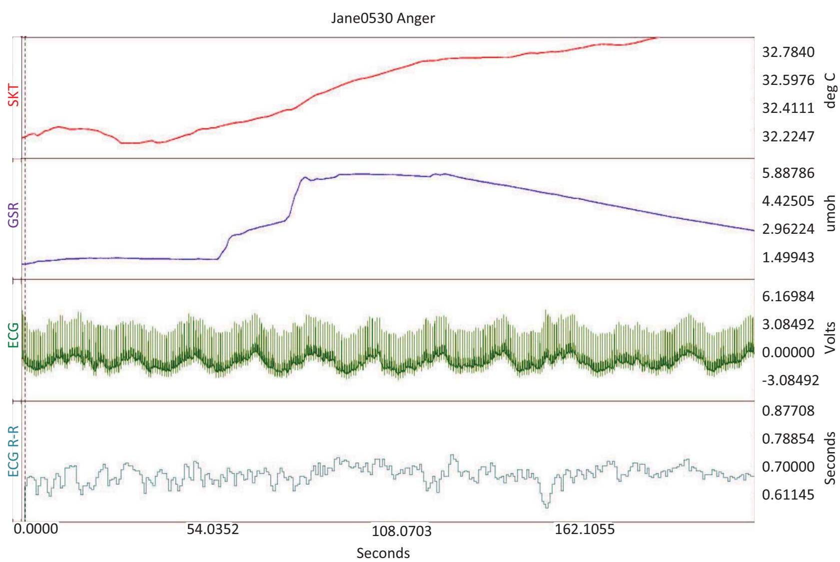

Fig. 3. Three physiological signals were recorded when the subject was asked to feel angry. The physiological signals were sampling at 256 samples for every second and the measured times were 120 s.

using the same protocol. In the meantime, the physiological signals of the subjects were also measured and recorded by a MP100 system in BIOPAC over the whole experiment. For every subject, the data we gather are from using different pictures over seven days.

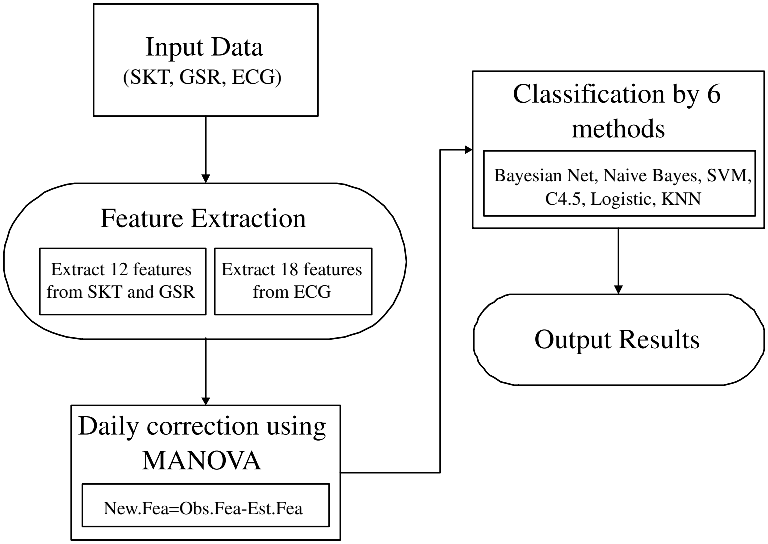

Fig. 4. Flow chart in the study of emotion recognition.

After gathering good affective data, the next step was the extraction of representative features from physiological signals. In this study, we would extract six features from the collected SKT data, six features from the GSR data and 18 features from the ECG data. Then, the daily dependence and personal dependence would be corrected by the statistical technique of MANOVA. Finally, the methodology of leave-one-out cross-validation was performed to evaluate the prediction accuracies of six classifiers. The flow chart of the proposed emotion recognition system is given in Fig. 4.

2.2. *Features extraction*

Much research has shown significant correlation between physiological signals and emotional status. However, unlike vision or speech recognition, physiological signals in different emotion statuses are not easy to be distinguished by a person immediately. Hence, it is very important to extract representative features that characterize main patterns from the raw physiological signals for classification pattern. For completeness, we would consider most of the features proposed from other literature.[14,21]

For the physiological signals of GSR and SKT, we would use the same features as Picard *et al.*[21] Those six statistic features were the mean, standard deviation, mean of the absolute values of the first difference, standard deviation of the first difference, mean of the absolute values of the second difference and standard deviation of the second difference of the sequence.

The physiological signals of ECG had been calibrated to heart rate variability (HRV) with baseline correction and their R peaks detection. Then, six statistic features were considered as well. In addition, 12 features were extracted from the power spectrum transformation, where the range of the high-frequency (HF) was set as $0.15 \sim 0.40$ MHz, the median-frequency (MF) was set as $0.08 \sim 0.15$ MHz and the low-frequency (LF) was set as $0.04 \sim 0.08$ MHz. In this study, the twelve features we selected were LF, MF, HF, TOTAL (LF+MF+HF), LF/TOTAL, MF/TOTAL, HF/TOTAL, LF/HF, MF/HF, (LF+MF)/HF, (LF+MF)/TOTAL and median of HRV.

3. Daily Correction

3.1. *The problem of day-effects*

There are many external stimuli, such as temperature and humidity, which can affect a person's physiological signals. In addition, a person's diet and sleep patterns can also cause variations in physiology. Hence, a person could have a different expression of the same physiological signal on different days even when he experiences the same emotion. Although we have made an effort to control these annoying factors, there are still some factors, such as hormones or a person's baseline mood, that are not controllable. Therefore, we must remove the day-effects for the emotion recognition study.

In a previous study, Picard proposed some methods to handle the problem of daily variations. Suppose we let the notation D and F as the number of experimental days and the number of features, respectively. In the method of day matrix for handling day-dependence, the method Picard proposed have to enlarge the original $D \times F$ matrix as $D \times (F + D - 1)$ matrix. Hence, if the experimental days are long, we must have a large amount of training day data, and consequently the computational overhead would be increased. Even though another method of baseline matrix for handling day-dependence would have avoided the above defect, the state of neutral emotion would be used as the baseline. Hence, we have to lose the opportunity to recognize the neutral emotion, and our number of states of emotion would be reduced.

Besides, in most previous studies of affective status from multiple subjects, the emotion recognition system treated the subjects and physiological signals independently over the same emotional status. However, because of people with different characteristics such as sex, age, weight and so forth, the physiological signals of different subjects would have different expressions even though they are experiencing the same emotion. Hence, it is necessary to develop an algorithm or method that can compensate the personal variations and day-to-day variations.

3.2. *MANOVA*

Since the problem of personal variations and daily variations would significantly influence the pattern classification in the system of emotion detection, we must remove the day-effects and person-effects for the emotion recognition study. In this project, we use the technique of MANOVA, which can be used even on a large number of experimental days; in the meanwhile, it doesnot have to reduce the number of states of emotion. After getting those 30 features from physiological signal of ECG, SKT and GSR, we would transform the features by the statistical technique MANOVA to remove the day-effects and person-effects. The MANOVA in this study is expressed as Eq. (1).

$$Z_{ijkl} = \mu_i + \tau_{ij} + \tau_{ik} + \tau_{ijk} + e_{ijkl}, \tag{1}$$

where $i = 1, 2, \ldots, I, j = 1, 2, \ldots, J, k = 1, 2, \ldots, K,$ and $l = 1, 2, \ldots, L$.

The notation of Z_{ijkl} represents the value of ith feature measured in jth subjects, kth days and lth sample. For the first database, the value of I is 30, J is 1, K is 11 and L is 3. In the second database, the value of I is 30, J is 10, K is 7 and L is 3. We let $Z_{jkl} = (Z_{1jkl}, Z_{2jkl}, \ldots, Z_{Ijkl})^{\mathrm{T}}$, $\mu = (\mu_1, \mu_2, \ldots, \mu_I)^{\mathrm{T}}$, $\tau_j = (\tau_{1j}, \tau_{2j}, \ldots, \tau_{Ij})^{\mathrm{T}}$, $\tau_k = (\tau_{1k}, \tau_{2k}, \ldots, \tau_{Ik})^{\mathrm{T}}$, $\tau_{jk} = (\tau_{1jk}, \tau_{2jk}, \ldots, \tau_{Ijk})^{\mathrm{T}}$, and $e_{jkl} = (e_{1jkl}, e_{2jkl}, \ldots, e_{Ijkl})^{\mathrm{T}}$. Equation (1) can be re-expressed as Eq. (2)

$$Z_{jkl} = \mu + \tau_j + \tau_k + \tau_{jk} + e_{jkl}, \tag{2}$$

where $j = 1, 2, \ldots, J, k = 1, 2, \ldots, K$ and $l = 1, 2, \ldots, L$.

The value μ is an overall mean value, the value τ_j represents the jth personal effect, the value τ_k represents the kth daily effect, and the value τ_{jk} represents the interact effect of daily and personal factor with the constraints that $\sum_{j=1}^{J} \tau_j = 0$, $\sum_{k=1}^{K} \tau_k = 0$ and $\sum_{j=1}^{J} \sum_{k=1}^{K} \tau_{jk} = 0$. The I-dimensional error vector $e_{jkl} = (e_{1jkl}, e_{2jkl}, \ldots, e_{Ijkl})^{\mathrm{T}}$ follows an I-dimensional multivariate distribution with a zero mean vector and a positive definite matrix Σ. Hence, the least squared estimates of $\hat{\mu}$, $\hat{\tau}_j$, $\hat{\tau}_k$ and $\hat{\tau}_{jk}$ are $\bar{Z}$, $\bar{Z}_j - \bar{Z}$, $\bar{Z}_k - \bar{Z}$ and $\bar{Z}_{jk} - \bar{Z}_j - \bar{Z}_k + \bar{Z}$ respectively, where $\bar{Z} = \frac{1}{JKL} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{l=1}^{L} Z_{jkl}$, $\bar{Z}_j = \frac{1}{KL} \sum_{k=1}^{K} \sum_{l=1}^{L} Z_{jkl}$, $\bar{Z}_k = \frac{1}{JL} \sum_{j=1}^{J} \sum_{l=1}^{L} Z_{jkl}$, $\bar{Z}_{jk} = \frac{1}{L} \sum_{l=1}^{L} Z_{jkl}$. Therefore the estimate $Z_{jkl} - \hat{\tau}_j - \hat{\tau}_k - \hat{\tau}_{jk} = Z_{jkl} - \bar{Z}_{jk} + \bar{Z}$ can be used to represent data after correction and we will use $X_{ijkl} = Z_{ijkl} - \bar{Z}_{ijk} + \bar{Z}_i$ as our attribute in the following classification methods.

For the comparison of two classification results, we treat the result of discrimination as a Bernoulli trail for every sample. Then, two sample t-tests could be applied in testing the difference between the classifiers. In this study, we use the p-value of the statistical improvement to compare the results of classification with and without daily and personal correction by MANOVA.

4. Pattern Classification

Tools of machine learning could be applied to discriminate the emotional states by the physiological signals. After daily and personal correction, we used the estimator $X_{ijkl} = Z_{ijkl} - \bar{Z}_{ijk} + \bar{Z}_i$ as our attribute for pattern classification of the emotional state Y_{jkl}, which represented the emotional state in jth subject, on the kth day, and for the lth sample. We let the variable Y represent the emotional status and the variable X_i represent the value of ith feature after removing the daily and personal correction. Six selected classifiers were tested for their performance and accuracy using the method of leave-one-out cross-validation. All of these six classification methods were performed by the software Weka (http://www.cs.waikato.ac.nz/ml/weka), and all of the classifiers used the default option in Weka. Further investigation of other options for classifiers in Weka could be studied in the future. The methods of classifiers were described as below.

4.1. *Bayesian network*

A Bayesian network, also called Bayes nets, is a directed acyclic graph (DAG) which consists of two components. The first component G comprises vertices corresponding to a set of variables $V = \{V_1, V_2, \ldots, V_N\}$ and a set of directed edges between variables with the Markov properties. The second component θ is attached the potential table $P(V_i|U_{V_i})$, for each variable V_i in V with the corresponding parents nodes U_{V_i}.[11,20] Given the structure G and the parameter θ, the joint probability distribution can be written as Eq. (3):

$$P(V) = \prod_{i=1}^{N} P(V_i|U_{V_i}). \tag{3}$$

For the purpose of learning to take place in a Bayesian network, we have to reconstruct the network structure and the field values. In this study, we apply the hill climbing algorithm and simple estimator to reconstruct the network and estimate the parameters. After getting the network structure, we used junction tree methods which can convert our DAG to a tree by clustering variables.[15] Then an efficient algorithm using belief propagation can be applied for our inference. In our study, we would use the estimator $X_1, X_2, \ldots, X_I$ and Y as the prediction variables $V = \{V_1, V_2, \ldots, V_{I+1}\}$ and calculate the conditional distribution of Y given the observation $X_1, X_2, \ldots, X_I$ in the constructed Bayesian network structure.

4.2. *Naive Bayesian*

A naive Bayesian classifier is a simple approach based on the Bayes' theorem. The network structure is illustrated in Fig. 5. There are two assumptions in the naive Bayesian classifier as follows[12]: (i) Given the class attribute (Y), the predictive attributes ($X_1, X_2, \ldots, X_I$) are independent. (ii) There were no other attributes affecting the prediction process. By the Bayes' theorem,

$$P(Y = y|X = x) = \frac{P(Y = y)P(X = x|Y = y)}{P(X = x)}. \tag{4}$$

We can predict the class attribute by finding y that maximizes $P(Y = y|X = x)$ in Eq. (4) given the predictive attributes x. As the predictive attributes $(X_1, X_2, \ldots, X_I)$

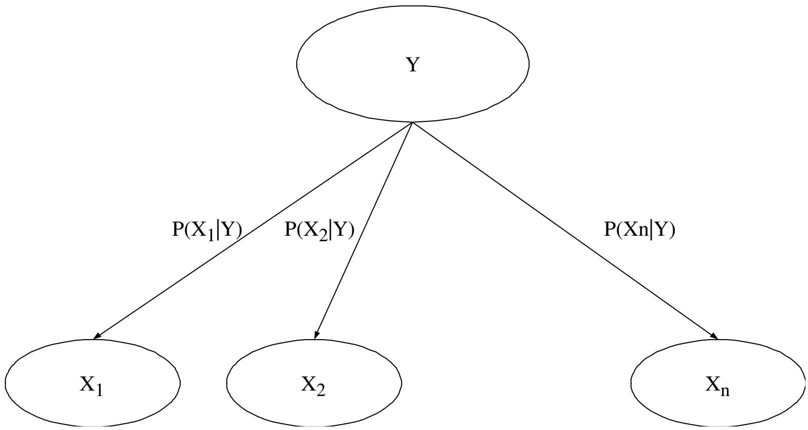

Fig. 5. The network structure of the naive Bayesian classifier.

are assumed to be conditionally independent, we have

$$P(X = x|Y = y) = \prod_{i=1}^{I} P(X_i = x_i|Y = y). \tag{5}$$

For the numeric attributes, we would assume that X_i is distributed as $N(\mu_{iy}, \sigma_{iy}^2)$ given the class $Y = y$ for every $i = 1, 2, \ldots, I$. Hence, we can estimate the parameters by the maximum likelihood estimates for each class.

4.3. *Support vector machine*

SVM[24] is a popular classification method used by a lot of research works currently being conducted in the field of emotion recognition.[5,14] Suppose $\{(x_1^*, y_1^*), (x_2^*, y_2^*), \ldots, (x_n^*, y_n^*)\}$ is the training set, where y_i^* is 1 or -1, denoting whether x_i^* belongs to one of two classes. In SVM, it is aimed to minimize the cost function $\frac{1}{2}w^{\mathrm{T}}w + C \times \sum_{i=1}^{n} \xi_i$ under the constraints $y_i^*(w^{\mathrm{T}}x_i^* + b) \geq 1 - \xi_i$ for $i = 1, 2, \ldots, n$. By using the Lagrange multiplier method, the original problem can be transformed as optimizing $\alpha_i's$ in Eq. (6).

$$\begin{aligned} &\arg\max_{\alpha} Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\alpha_j y_i^* y_j^* x_i^{*\mathrm{T}} x_i^* \quad s.t.\ 0 \leq \alpha_i \leq C\,\forall i; \\ &\sum_{i=1}^{n} \alpha_i y_i^* = 0. \end{aligned} \tag{6}$$

After obtaining α_i, we can apply the following decision function for prediction using the new predictive attribute of $x_{\text{new}}^* : f(x_{\text{new}}^*) = \text{sign}(\sum_{i=1}^{n} y_i^* \alpha_i K(x_{\text{new}}, x_i^*) + b)$, where $K()$ is the kernel function. In this study, we use the Gaussian kernel and the sequential minimal optimization (SMO) algorithm.[13] Besides, because our case has multiple classes (three emotional statuses), we used the approach of pairwise classification by the one-against-one approach in the SVM classification method.

4.4. *Decision tree of C4.5*

Decision tree is also a common method used in classification.[10] C4.5 is a hierarchical data structure using the divide-and-conquer strategy to growing decision trees.[22] In decision trees, each decision node using a test function to partition original data D into subsets $D_1, D_2, \ldots, D_n$. Suppose the set D consists of C numbers of classes and $p(D, j)$ denotes the proportion of cases in D that belongs to the jth class. We can define the information gain by a test T with m outcomes as Eq. (7):

$$\mathrm{Gain}(D, T) = \mathrm{Info}(D) - \sum_{i=1}^{m} \frac{|D_i|}{|D|} \times \mathrm{Info}(D_i), \tag{7}$$

where $\mathrm{Info}(D) = -\sum_{j=1}^{C} p(D, j) \times \log(p(D, j))$ and it can reach its maximal when there is one case left in each subset D_i. The split information is defined as Eq. (8):

$$\mathrm{Split}(D, T) = -\sum_{i=1}^{m} \frac{|D_i|}{|D|} \times \log\left(\frac{|D_i|}{|D|}\right). \tag{8}$$

For every possible test, the ratio of its information gain over its split information is assessed and the test with maximum gain ratio is selected.

4.5. *Logistic model*

Logistic regression is a classical method to model category data for classification.[16] Suppose there are n samples with c classes and I attributes. The parameter matrix B is calculated as an $I \times (c-1)$ matrix. The probability that the ith sample, given the value of x_i^*, in the jth class but not in the last cth class is shown in Eq. (9).

$$P_j(x_i^*) = \frac{\exp(x_i^* B_j)}{\sum_{k=1}^{c-1} \exp(x_i^* B_k) + 1}, \quad \text{where } j = 1, 2, \ldots, c-1. \tag{9}$$

The probability that the ith sample, given the value of x_i^*, in the last cth class is shown in Eq. (10).

$$P_c(x_i^*) = 1 - \sum_{k=1}^{c-1} P_k(x_i^*) = \frac{1}{\sum_{k=1}^{c-1} \exp(x_i^* B_k) + 1}. \tag{10}$$

The log-likelihood l of the data (K, X) under this model is shown in Eq. (11).

$$l(\beta) = \sum_{i=1}^{n} \left\{ \sum_{k=1}^{c-1} K_{ik}^* \ln(P_k(x_i^*)) + \left(1 - \sum_{k=1}^{c-1} K_{ik}^*\right) \ln\left(1 - \sum_{k=1}^{c-1} P_k(x_i^*)\right) \right\}. \tag{11}$$

The indicator variable $K_{ij}^* = 1$ if the ith sample belongs to the jth class, where $j \neq c$. Otherwise, $K_{ij}^* = 0$ if the ith sample belongs to the last cth class. The parameter matrix B can be estimated by the maximize likelihood estimates of the likelihood function, $l(\beta)$.

4.6. *K-nearest neighbor (KNN)*

The k-nearest neighbor (KNN) algorithm is one of the classical classification methods that have wide applications.[1] KNN compares the similarity between testing data and every training data. Then it uses the top k similarity categories of training data to decide the category of the testing data by a weighted vote. For any testing data of H and training data of $\{G_1, G_2, \ldots, G_n\}$, we would classify the category of H as Eq. (12).

$$C(H) = \arg\max_m \sum_{G_i \in S} \mathrm{Sim}(H, G_i) I(G_i, C_m). \tag{12}$$

The notation of $\mathrm{Sim}(H, G_i)$ is the similarity measure of H and G_i. The set $S = \{\tilde{G}_1, \tilde{G}_2, \ldots, \tilde{G}_k\}$ is the data set closed to the testing point H, and the notation of $I(G_i, C_m) \in \{0, 1\}$ indicates whether G_i belongs to C_m. If there are tie cases in the classification, we will use the group with a minimal index as the corresponding category of testing data. In this study, we would use the Euclidean distance as the similarity measure and choose the number of nearest neighbors $k = 3$.

5. Empirical Results

Due to the existence of day-dependency of the features from various physiological signals, some of the features would have a large discrepancy even though they are in the same emotional state. Besides, some of the feature expression would be quite similar even for different emotional states. However, after removing the daily effects by MANOVA, the scatter of the features in the same state would be more tight. Therefore, it would become much more differentiable for distinct states of emotion. In order to display the feature discrimination by different emotions, we use the dimension reduction approach of principal component analysis in the thirty features. Then, the largest two principal components were used as the two dimensions for the scatter plots without and with daily correction in Figs. 6 and 7, respectively. According to the scatter distribution of Figs. 6 and 7, we found that the statistical technique of MANOVA would be helpful for emotion classification and recognition.

After getting the features from physiological signals and removing daily effects by MANOVA, six classification methods would be applied as discussed before. For the first database, the classification results without daily correction and with daily correction by leave-one-out cross-validation in three emotional statuses by six classification methods are listed in Tables 1 and 2. From the classification result of subject Alice, the highest correct recognition rate is 90.91% using the classification method of logistic model. As for the classification result of subject Jane, the highest correct recognition rate is 93.94% by the same classification method of logistic model.

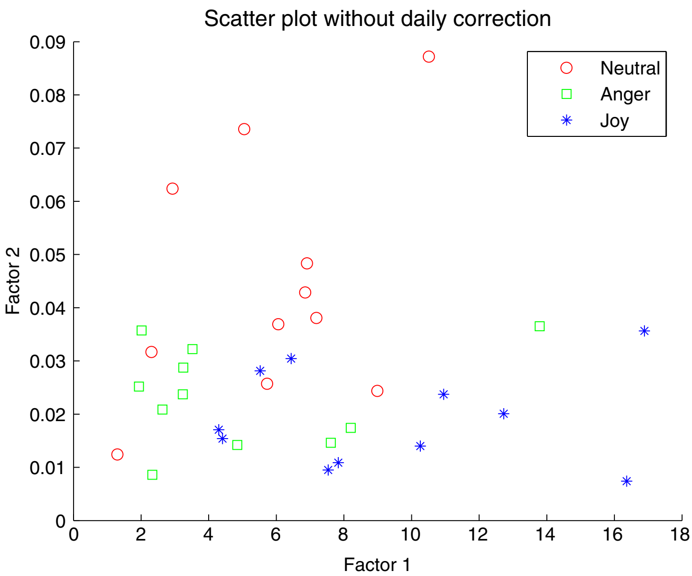

Fig. 6. The scatter plot of three statuses of emotion without daily correction.

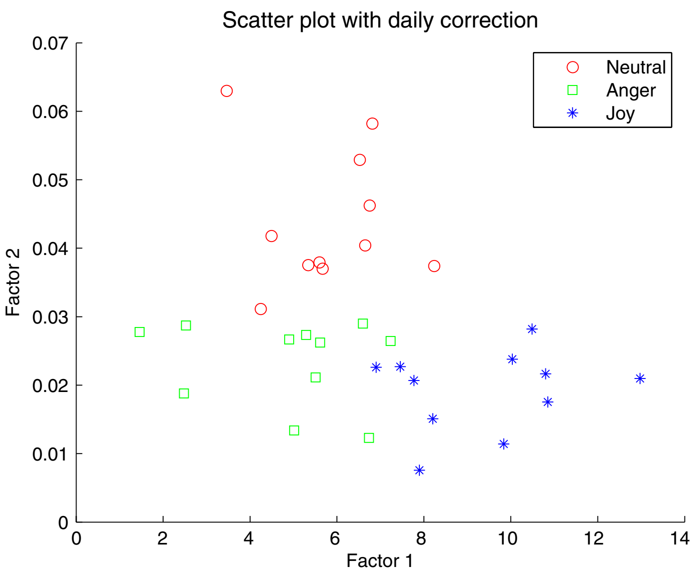

Fig. 7. The scatter plot of three statuses of emotion with daily correction.

The leave-one-out cross-validation is used to evaluate the prediction accuracy. For evaluating the effect of MANOVA for daily correction, we use the pairwise t-test. The results show that the p-values are all significant in most classification methods, except for the subject Jane with the classification method of C4.5.

Table 1. Classification of three emotional statuses by the physiological signals of subject Alice.

	Alice		
Method	Without Daily Correction (%)	With Daily Correction (%)	p-Value of Improvement
Bayesian Network	45.45	81.82	0.001
Naive Bayesian	48.48	75.76	0.011
SVM	45.45	78.79	0.003
C4.5	45.45	78.79	0.003
Logistic model	57.58	90.91	0.001
KNN	39.40	75.76	0.001

Table 2. Classification of three emotional statuses by the physiological signals of subject Jane.

	Jane		
Method	Without Daily Correction (%)	With Daily Correction (%)	p-Value of Improvement
Bayesian Network	51.52	90.91	0.000
Naive Bayesian	48.48	78.79	0.005
SVM	69.70	84.85	0.073
C4.5	66.67	75.76	0.211
Logistic model	60.61	93.94	0.000
KNN	63.64	87.88	0.011

For second database, in order to see how the statistical technique of MANOVA influence classification, we compare the classification accuracy with and without removing the daily and subject dependence in Table 3. In addition, the p-value with statistical significance of the difference in classification between with and without the technique of MANOVA for every classification method is also attached in Table 3 as well. Among these six classification methods, the highest correct recognition rate is 74.76% using the classification method of logistic model and all of these classification method can significantly improve the overall accuracy rate after removing daily and personal effects by MANOVA.

Table 3. Classification of three emotion status by the physiological signals of 10 subjects and 7 times.

Method	Without Daily Correction (%)	With Daily Correction (%)	p-Value of Improvement
Bayesian network	49.05	64.76	0.001
Naive Bayes	48.57	65.71	0.000
SVM	45.24	70.48	0.000
C4.5	50.00	61.90	0.007
Logistic model	54.29	74.76	0.000
KNN	42.38	58.10	0.001

6. Conclusion and Discussion

In this study, we have compared six typical classifiers by their performances in emotion recognition using physiological signals with daily and personal correction by MANOVA. As the results mentioned above, these classification methods can be very useful to perform emotion recognition by using the physiological signals with daily and personal correction. In particular, we can successfully correct daily effects using the statistical techniques of MANOVA.

There are still challenges for future studies. For example, we could investigate and determine significant features using feature selection and dimensional reduction methods. In addition, more data collection could be performed in future studies to improve the accuracy. Real-time applications could be further investigated for the prediction of emotional states based on the physiological signals with daily correction. Further adjustments of parameters in classification methods could be investigated. These are interesting topics that we plan to study in the future based on the framework of the current research results.

Acknowledgments

We would like to thank the editor and reviewers for their constructive comments. This work was supported by the Center for Measurement Standards at the Industrial Technology Research Institute (ITRI) in Taiwan.

References

1. D. W. Aha, D. Kibler, and M. K. Albert, Instanced-based learning algorithms, *Mach. Learn.* **6**(1) (1991) 37–66.
2. W. S. Ark, D. C. Dryer and D. J. Lu, The emotion mouse, *Int. Conf. Human-Computer Interaction* (1999) pp. 818–823.
3. C. Breazeal and L. Aryananda, Recognition of affective communicative intent in robot directed speech, *Auton. Robot.* **12** (2002) 83–104.
4. C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann and S. Narayanan, Analysis of emotion recognition using facial expressions, speech and multinodal information, in *ICMI '04: Proc. the 6th Int. Conf. Multimodal Interfaces*, New York, NY, USA (ACM, 2004), pp. 205–211.
5. C. F. Chuang and F. Y. Shih, Recognizing facial action units using independent component analysis and support vector machine, *Pattern Recogn.* **39**(9) (2006) 1795–1799.
6. R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. Taylor, Emotion recognition in human-computer interaction, IEEE Signal Process. Mag. **18**(1) (2001) 32–80.
7. F. Dellaert, T. Polzin, and A. Waibel, Recognizing emotion in speech, *Int. Conf. Spoken Language Proceedings*, Vol. 3, (1996), pp. 1970–1973.
8. B. Fasel and J. Luettin, Automatic facial expression analysis: A survey, *Pattern Recogn.* **36**(1) (2003) 259–275.
9. T. Hu, L. C. De Silva, and K. Sengupta, A hybrid approach of NN and HMM for facial emotion classification, *Pattern Recogn. Lett.* **23** (11) (2002) 1303–1310.

10. E. Hunt, J. Marin, and P. Stone, *Experiments in Induction* (Academic Press, New York, 1966).
11. F. Jensen, *Bayesian Networks and Decision Graphs* (Springer, New York, 2001).
12. G. H. John and P. Langley, Estimating continuous distributions in Bayesian classifiers, in *Proc. 11th Conf. Uncertainty in Artificial Intelligence* (1995) 338–345.
13. S. Keerthi, S. Shevade, C. Bhattacharyya, and K. Murthy, Improvements to Platt's SMO algorithm for SVM classifier design, *Neural Comput.* **13**(3) (2001) 637–649.
14. K. Kim, S. Bang, and S. Kim, Emotion recognition system using short-term monitoring of physiological signals, *Med. Biol. Eng. Comput.* **42**(3) (2004) 419–427.
15. S. Lauritzen and D. Spiegelhalt, Local computations with probabilities on graphical structures and their application to expert systems, *J. Roy. Stat. Soc. S. B* **50** (1988) 157–224.
16. S. Le Cessie and J. Van Houwelingen, Ridge estimators in logistic regression, *Appl. Stat.* **41**(1) (1992) 191–201.
17. G. Littlewort, M. S. Bartlett, I. R. Fasel, J. Chenu, T. Kanda, H. Ishiguro, and J. R. Movellan, Towards social robots: Automatic evaluation of human-robot interaction by face detection and expression classification, *Adv. Neural Inform. Process. Syst.* **16** (2004) 1563–1570.
18. F. Nasoz, K. Alvarex, C. L. Lisetti, and N. Finkelstein, Emotion recognition from physiological signals using wireless sensors for presence technologies, *Cognit. Tech. Work* **6**(1) (2004) 4–14.
19. T. Nwe, S. Foo, and L. De Silva, Speech emotion recognition using hidden Markov models, *Speech Commun.* **41** (2003) 603–623.
20. J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, San Mateo, 1988).
21. R. W. Picard, E. Vyzas, and J. Healy, Toward machine emotional intelligence: Analysis of affective psychological states, *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(10) (2001) 1175–1191.
22. J. Quinlan, *C4.5: Programs for Machine Learning* (Morgan Kaufmann, San Mateo, 1993).
23. P. Rani, C. Liu, N. Sarkar, and E. Vanman, An empirical study of machine learning techniques for affect recognition in human-robot interaction, *Pattern Anal. Appl.* **9** (2006) 58–69.
24. V. N. Vapnik, *Statistical Learning Theory* (New York: Wiley, 1998).
25. Y. Yacoob and L. S. Davis, Recognizing human facial expression from long image sequences using optical flow, *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(6) (1996) 636–642.
26. C. Zhou and X. Lin, Facial expressional image synthesis controlled by emotional parameters, *Patt. Recogn. Lett.* **26**(16) (2005) 2611–2627.

Tung-Hung Chueh received his Ph.D. from the Institute of Statistics at the National Chiao Tung University, Hsinchu, Taiwan in 2008. He has been working as postdoctorate researcher in Sinica and National Chiao Tung University in Taiwan. Since then, he has worked in the Industrial Technology Research Institute. His research interests include data mining, machine learning, bioinformatics, statistical learning and energy economics.

Tai-Been Chen is an Associate Professor in the Department of Medical Imaging and Radiological Science, I-Shou University, Taiwan, ROC (e-mail ctb@isu.edu.tw and personal webpage http://140.133.62.66). He obtained his master degree and Ph.D., in 1999 and 2007 respectively, at the Institute of Statistics, National Chiao Tung University, Taiwan. His research interests include biomedical image information, medical image processing, and statistical computing.

Henry Horng-Shing Lu received his Ph.D. and M.S. degree in Statistics from Cornell University, USA in 1994 and 1990, respectively, and his B.S. degree in Electric Engineering from the National Taiwan University, Taiwan, in 1986. He is a professor at the Institute of Statistics, National Chiao Tung University, Taiwan. He currently serves as the Dean of College of Science at the National Chiao Tung University. He is an elected member of the International Statistical Institute. His research interests include statistics, image science, and bioinformatics.

Shan-Shan Ju acquired her Ph.D. in Chemistry from the University of Pennsylvania, USA in 1994. She worked as a postdoctoral fellow in IAMS, Sinica, Taiwan, conducting a study of surface vibrational spectroscopy using non-linear optics for five years, before she joined ITRI in 2002. Her emotion/BCI recognition work via physiological and EEG signals in ITRI was conducted from 2004−2011. Her current position is as a patent examiner in the field of physical measurement at the Taiwan Intellectual Property Office.

Teh-Ho Tao received his M.S. degree in Electrical Engineering from the University of Arizona, Tucson, Arizona in 1981 and his Ph.D. in Biomedical Engineering from Case Western University, Cleveland, Ohio in 1988. Currently he is with Holux Technology, Inc. in the Science-Based Industrial Park, Hsinchu, Taiwan. He established the health care product business unit at Holux Technology in 2010 and has been responsible for the commercialization of ultra wideband radar technology for vital signs monitoring applications. From 1989 to 2010, he was with the Center for Measurement Standards, Industrial Technology Research Institute in Hsinchu, Taiwan where he worked on the development and transfer of key technologies in sensor and measurement instrument to the medical as well as industrial instrumentation industries in Taiwan. He received the Ministry of Economic Affairs Excellent Project Award for the Medical Device Technology Five Year Development Project in 2000, the 2006 IEEE/AESS Harry Rowe Mimno Award for the paper "Short-Distance Ultra Wideband Radars" and the 2009 IEEE/AESS Harry Rowe Mimno Award for the paper "Ultra Wideband Radar for Patient Monitoring".

Jiunn-Haur Shaw received his Ph.D. in Mechanical Engineering from the Pennsylvania State University in 1989. Since then, he joined the Center for Measurement Standards/ITRI, and was responsible for the construction, facilitating, commissioning of national flow measurement standards in Taiwan. These include water, hydrocarbon, low pressure, high pressure air and air speed flow facilities. In the past 22 years of his career with CMS, Dr. Shaw has led the flow metrology group in disseminating flow measurement standards to local industries, participating in international key comparisons to safeguard the measurement capabilities of the national standards in equivalence with other NMIs, and providing technical consulting services to domestic and international companies and regional NMIs. In 2005, he received the Algie Lance Best Paper Award from the Measurement Science Conference. He also served as APMP technical committee chair of Fluid Flow from 2009−2010. During 2004, Dr. Shaw joined a research project in developing sensing technology for child safety and pattern recognition of mental states.