2011 International Conference on Telecommunication Technology and Applications
Proc .of CSIT vol.5 (2011) © (2011) IACSIT Press, Singapore

Determination of Human Physiological State by Multimodal Biometric Device

Gennady Shishkin [1], Alexey Shishkin [2], Igor Ageev [1], Vladimir Litvinov [1], Yury Rybin [1], Andrey Plokhikh [1 +] , and Galina Soganova [1]

[1] Moscow Aviation Institute (State Technical University), Moscow, Russia
[2] Moscow State University, Moscow, Russia

Abstract. The important problem of human psychophysiological state recognition is studied. An integrated hardware–software setup has been developed to achieve automatic assessment of the affective status of a human. Several signal processing techniques are applied to the collected signals to extract the set of the most relevant in the physiological responses features. The support vector machine was used as a classifier. As a result 93.4% functional state prediction accuracy was obtained.

Keywords: physiological state recognition, tremor, galvanic skin response, speech, suprathermal radiance

1. Introduction

Physiological data can provide useful information about human emotional or cognitive states and can help in the recognition of the level of cognitive load or of the presence of stress. In many cases, people must work in extreme conditions that lead to strong emotional strain, e.g. air traffic control officers, operators, drivers, pilots etc. There is a great interest to human functional state evaluation in modern biomedical practice. In spite of considerable achievements in this area during the past several decades, there are still a lot of problems, and many researchers are trying to solve them. Most of them use two modalities: facial and vocal expressions [1] and relatively little attention has been paid to combine several other physiological signals [2]. In contrast to speech and facial expression other physiological signals are not under voluntary control, and thus cannot be masked up to the same extent. One of the works with several modalities used is the paper of R. Picard et al [3] where authors analysed five physiological signals: facial muscle tension along the masseter, blood volume pressure, skin conductivity, a Hall effect respiration and a heart rate and achieved 81% recognition accuracy on eight classes of emotion including neutral.

In this paper an integrated hardware – software setup has been developed to achieve automatic assessment of the functional status of a human [4]. The device uses several modalities such as suprathermal radiance of biological object (SRBO), galvanic skin response (GSR), muscle microtremor and speech analysis. We utilized a support vector machine in pattern classification. Based on the results of signal processing the most relevant features were extracted by rating them by their importance. The values of these features depend neither on subject nor operator opinion about human functional state.

2. Experimental Setup and Operating Principles

The developed hardware-software setup allows non-invasive measurement of several biometric parameters, their processing, data storage and documentation. If special procedure is designed it could be used for recognition of human physiological state, for the estimation of human state change under the

+ Corresponding author. Tel.: + 7 (499) 158-0367; fax: +7 (499) 158-0367.
E-mail address: plokhikh2001@mail.ru

influence of various external factors, for person identification, for the determination of body physiological reserve, for medical diagnostics, as a lie detector etc.

For our research we have chosen the following techniques to measure physiological signals: suprathermal radiance of biological object (SRBO); hand muscle microtremor; galvanic skin response (GSR); human voice. The complete instrumental setup developed for this research consists of a water sensor of hand suprathermal radiance, a galvanic skin response sensor, a microtremor sensor, a microphone, a power source of primary sensors, 16 channels ISA Multifunction Card Advantech PCL-818HG, a computer sound card.

The analog signals (water electrical conductivity and temperature, skin resistance, skin oscillations and microtremor) are received by data acquisition card. After analog-digital conversion these signals along with voice signal from the sound card are processed and analysed by specially developed software. It sets operating mode of the card with simultaneous acquisition of SRBO, GSR and microtremor data, calculates the human psychophysiological parameters and provides visual demonstration of results and graphical user interface. The sampling rate of SRBO, GSR and microtremor collecting is 50 Hz (see [4] for more details).

2.1. Suprathermal Radiance of Biological Objects

Fig. 1 shows specific conductivity of water as a function of temperature. Measurements were made in two identical conductometric cells simultaneously. During the heating of one of the cells (curve 1) the heater was replaced by the experimenter's hand at the time moment corresponding to the water temperature of about 24.8°C. Heating of the second cell by the heater continued with no conditions changed. Substantial variation in the dependence of water specific conductivity on temperature was observed [5].

More intense growth of the water conductivity at its heating by hand than at heating by common electric heater testifies to the presence of additional energy factor associated with the influence of a human being. We called this factor as suprathermal radiation.

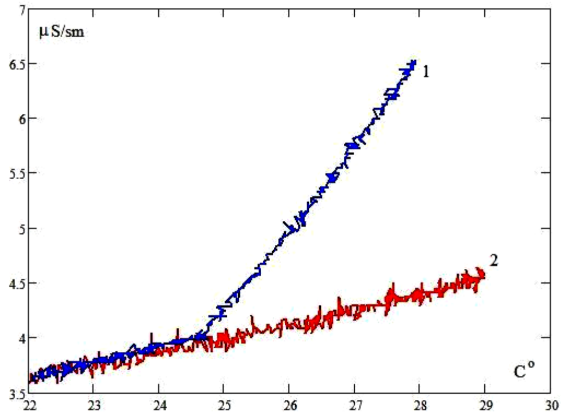

Fig. 1: Electroconductivity dependence on water temperature for usual heating and bioaction.

To measure quantitatively the observed deviation the following two parameters were used:

$$a = \frac{1}{\sigma}\frac{\partial \sigma}{\partial T}, \quad B = \frac{\Delta\sigma_B - \Delta\sigma_T}{\sigma_0} \cdot \frac{1}{\Delta t} \tag{1}$$

where σ is an electric conductivity, $\Delta\sigma_B$ is an electrical conductivity increment during the action of a biological object (bioaction), $\Delta\sigma_T$ is an electrical conductivity increment during the usual heating, Δt is a period of bioaction/heating, σ_0 is an initial value of electric conductivity and T is a temperature. The parameter a is a relative temperature coefficient of electric conductivity and describes the water properties.

Nominally parameter B can be viewed as a bioaction power (a relative excess of an electrical conductivity increment during bioaction over an electrical conductivity increment during heating by a physical heat source per time unit).

In practice we used slightly different expressions for B leaving unchanged its meaning but utilizing linear dependence of conductivity on temperature during the usual heating, the constant heating time, the difference of initial temperature (conductivity) on two measurement steps etc.

2.2. Galvanic Skin Response

To measure GSR parameters the electrodes made from biologically inert materials were used. The following parameters were analysed: the value of an absolute skin resistance (ASR), its time dynamics and the GSR spectrum.

There are two methods of GSR measurement: the passive one with registration of skin electrical potentials (Tarkhanov's method, PSR) and the active one when the external source of electrical field is used for measuring the resistance (Ferre's method, ASR).

In our device we used both methods: PSR and ASR were measured simultaneously. The most interesting information that can be obtained from PSR is an energetic of skin resistive oscillations directly connected with local emotional state of humans, the spectrum of these oscillations and ASR dynamics. For that purpose we defined the following parameters:

- integral coefficient I proportional to PSR oscillation power : $I = \sum_{i=1}^{n} SP_i$ where SP_i are harmonics of absolute value of GSR spectrum from 0 to maximum frequency, n is a number of harmonics;
- the hidden frequency of oscillations F_e: $F_e = \frac{\sum dSP_i}{k \cdot I}$ where dSP_i are harmonics of spectrum of PSR time derivative, k is a coefficient that depends on the duration of signal;
- the absolute skin resistance R_k;
- coefficient D_k that describes the skin resistance changes during measurement time.

2.3. Hand Muscle Microtremor

To measure the frequency and intensity of human hand microtremor the standard accelerometer was used as a sensor. Two main parameters connected with signal spectrum were calculated. The first one is an integral coefficient A proportional to signal oscillations power $A = \sum A_i$ where A_i are components of absolute value of tremor spectrum density up to frequency of 20 Hz except zero frequency A_0. The second parameter is an average oscillation frequency f_{av}:

$$\sum A_i(f_i < f_{av}) = \sum A_i(f_i > f_{av}) \quad (2)$$

where A_i and f_i are the amplitude and the frequency of corresponding spectrum component at the range $0 < f_i < 20$ Hz.

2.4. Voice Analysis

Despite the complexity of voice stress analysis the speech parameters used along with SRBO, GSR and microtremor parameters allow to draw quite reliable conclusions about human functional state. In our experiments about 10 speech signals were recorded for each human in different physiological states.

Each record was analysed to detect signal endpoints by calculating signal's energy and zero crossing rate. Then we extracted fundamental frequency, the first three formants and vocal energy features such as maximum, minimum, mean, median values, maximum, minimum and mean duration of rising and falling slopes etc. and Teager Energy Operator (TEO) based parameters [6]. To find fundamental frequency a Robust Algorithm for Pitch Tracking (RAPT) was applied [7]. Then we smoothed the fundamental frequency contour using cubic splines. It enabled us to measure its features: the derivatives, slopes and the behaviour of their minima and maxima over time. The formant frequencies were estimated by finding the poles of the autoregressive transfer function obtained from the linear predictive coding (LPC) coefficients. To find TEO based features we first filtered speech signal through four band-pass filters with passbands 0-1 kHz, 1-2 kHz, 2-3 kHz and 3-4 kHz. Then we applied TEO to each output signal $x(n)$:

$$\psi(x(n)) = x^2(n) - x(n-1)x(n+1) \quad (3)$$

and filtered the result through a Gabor bandpass filter centered at the median fundamental frequency with a bandwidth of half of fundamental frequency. We segmented the signal into frames and computed normalized autocorrelation function and finally an area under its envelope for each frame. Thus we obtained four parameters TEO_1, TEO_2, TEO_3, TEO_4. We calculated AM and FM components of TEO profile as well:

$$|a(n)| = \sqrt{\frac{\psi(x(n))}{1-\left(\frac{\psi(y(n))+\psi(y(n+1))}{4\psi(x(n))}\right)^2}}, \quad f(n) = \frac{1}{2\pi T}\arccos\left(1-\frac{\psi(y(n))+\psi(y(n+1))}{4\psi(x(n))}\right) \tag{4}$$

where $y(n) = x(n) - x(n-1)$. In addition 14 Mel-Frequency Cepstral Coefficients (MFCC) were calculated for each frame. As a result we obtained a set of 211 speech features.

3. Device Testing

The developed setup allows to estimate the human physiological state by the following biometric features: the rise of B coefficient of water conductivity under the influence of suprathermal radiance of human hand (SRBO); average oscillation frequency for hand and finger muscles (microtremor); integral intensity of oscillation for hand and finger muscles (microtremor); integral intensity of skin oscillations (GSR); average frequency of skin oscillations (GSR); absolute skin resistance (GSR); skin resistance change during measurements (GSR); fundamental frequency features (voice); formant intensities (voice); word duration (voice); vocal energy features (voice).

Preliminary studies showed both the presence of individual combination of above parameters and close connection of these parameters with human physiological state. The setup software allows changing the duration and sampling frequency of measurements and some other input parameters.

Collected features from all modalities formed high-dimensional vectors that we gave to the support vector machine (SVM) classifier. Since different features with complex interrelationships are extracted from human physiological data we choose a Gaussian kernel.

To test the developed hardware-software setup several series of experiments showing that it registers the changes in human functional state were performed. In particular we collected physiological signals from 132 students of our department and divided them into 436 data entries since each participant generated data under certain number of normal and stressful conditions. The device determined the human state changes under the influence of the following stress factors: exams; mental workload; alcohol consumption and smoking.

First the human biometric parameters were measured. Then we exposed the subject to one of the abovementioned factors changing his functional state. Finally the biometric parameters were measured once more. The comparison between two parameter groups allowed testing our device.

One of examples of registration in functional state changes is a growth of student average muscle microtremor intensity before exam relative to the values in a normal state. These changes are also reflected in microtremor spectrum and in PSR and ASR signals. The typical histogram of water heating parameter B, microtremor parameter A and galvanic skin response parameters R_k and D_k are shown in Fig. 2a. The measurements were made just before examinations and 10-15 minutes after the examination. The reaction of each students is strongly individual but the changes were always observed. There are differences in parameters between male and female students.

The changes in human psychophysiological state during alcohol consumption and smoking were registered as well. It was revealed that smoking led to substantial human energy losses and hence to the rise of parameter B (it is proportional to palm radiance energy and reflects integral intensity of metabolic processes). The larger the value of B before smoking, the lesser its changes after smoking (see Fig. 2b with y-axis as a ratio Δ of values of B after and before smoking and x-axis as a value of B before smoking).

In Fig. 3a statistically significant results for SRBO parameter B for both human hands are presented. Solid curves refer to emotional persons and dashed curves refer to rational persons. The experimental data are smoothed by cubic splines to clearly show a trend. In several experiments the relative position of curves can be reversed in certain time moments.

In our smoking experiments the main microtremor parameter is an integral spectral parameter A representing a muscle tremor energy connected with human metabolic processes. Note that about 26% of total metabolic energy is transformed into energy in muscles. After smoking the tremor energy grows substantially that testifies to an increase of stress in human body. Besides energy changes the microtremor is modified as well. In Fig. 3b the spectra before (solid curve) and after (dashed curve) smoking are given. The typical spectrum feature after smoking is an appearance of more regular oscillations and their harmonics.

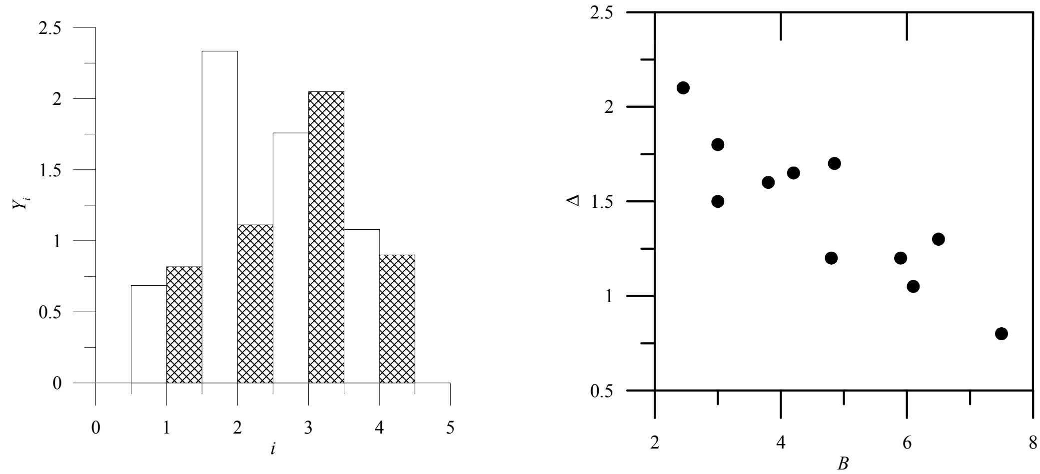

Fig. 2: a) The typical histogram of values Y_i of different modalities i before and after (cross hatched) exams. i=1 – bioaction (parameter B), i=2 – microtremor (parameter A), i=3 and i=4 - skin response (parameters R_k and D_k respectively). b) The ratio of B after and before smoking vs value of B before smoking

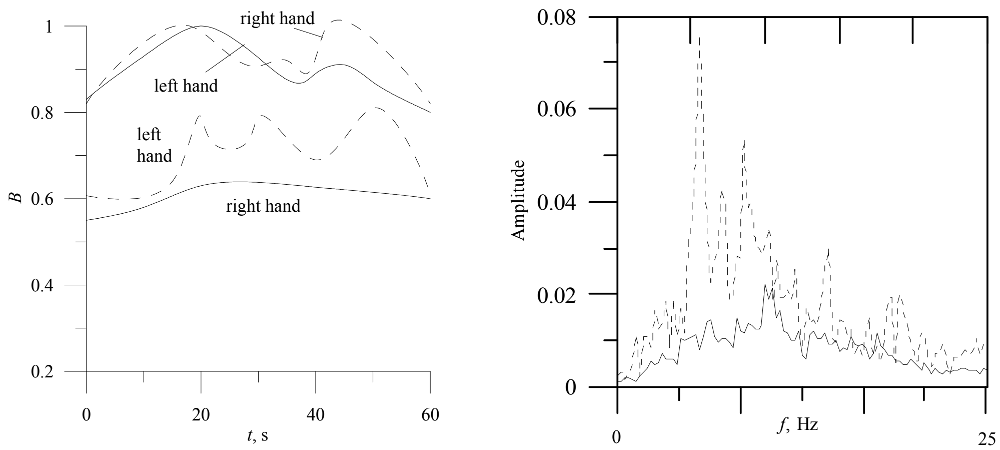

Fig. 3: a) The time dependence of normalized water heating parameter B for both hands for emotional (solid curves) and rational (dashed curves) persons; b) The spectrum of microtremor signal before (solid curve) and after (dashed curve) smoking.

The fundamental frequency, formant features and word duration are human specific. Parameters TEO_2, TEO_3, TEO_4 are robust features for human state recognition. For example, an average value of TEO_2 after smoking was in the range 0.85-0.94 whereas before smoking it changed from 0.65 to 0.73. Similarly average values of TEO_3 and TEO_4 after smoking fell within the limits of 0.11-0.13 and 0.18-0.22 respectively whereas before smoking they ranged within 0.06-0.09 and 0.11-0.16 respectively. Hence smoking leads to the increase of TEO_2, TEO_3, TEO_4.

On the whole the parameters changes are strictly human specific. Overall performance was quantified as the percentage of correctly classified data vectors. As a result we obtained 93.4% functional state prediction accuracies.

In order to study the classification ability of each voice feature, a rating method has been implemented. Each feature was evaluated by its importance

$$p_k = \left(\nu_k \Big/ \sum_{k=1}^{n} \nu_k \right) \times 100\%\,,\quad \nu_k = \frac{1}{N} \sum_{i=1}^{N} d_i^k \tag{5}$$

where d_i^k is a distance to separating hyperplane along k-feature for feature vector i. Our experiments showed that the importance of more than a half of features is less than 0.5 %. We kept the features with importance higher than 0.5%, 1%, 1.5% and 2%. The dependence of prediction accuracy on the number of features is shown in Table 1.

Table 1. Dependence of Prediction Accuracy on Number of Voice Features.

Feature importance, %	**Number of features**	**Prediction accuracy, %**
>0	211	93.4
>0.5	57	92.2
>1	34	91.7
>1.5	29	85.6
>2	15	82.1

It can be seen from Table 1 that the reduction in number of features from 211 to 57 leads to small decrease in prediction accuracy (from 93.4% to 92.2%). Moreover the accuracy is still quite acceptable for only 34 features (91.7%). Such analysis has to be performed for all modalities. It will be the subject of future works.

4. Conclusions

A hardware-software setup for human physiological state recognition is developed and tested. The techniques used are multimodal, non-invasive and diagnostic time is less than 2 minutes. So this real-time assessment of affective states can be successfully applied in various fields. The device measures both well-known physiological parameters (GSR, microtremor) and an unique experimentally discovered phenomenon (SRBO). To analyse human voice many features are extracted. The acceptable level of human state recognition accuracy is achieved.

The possible potential applications of the device include human functional state control in various areas; patient's state diagnostics in clinics and hospitals; credibility assessment; competency analysis; lie detector; evaluation of environmental and other external factors influence on human state.

5. Acknowledgements

This work was partly funded by Russian Foundation for Basic Research grant №10-07-00207.

6. References

[1] L.C. De Silva, and P.C. Ng. Bimodal Emotional Recognition. In: *Proc. 4th IEEE International Conf. on Automatic Face and Gesture Recognition*. IEEE Press. 2000, pp. 332-335.

[2] J. Kim. Bimodal emotion recognition using speech and physiological changes. In: M. Grimm, and K. Kroschel (eds) *Robust Speech Recognition and Understanding*. Vienna: Eds. I-Tech. 2007.

[3] R. Picard, E. Vyzas, and J. Healy. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *Trans. Pattern Anal. Mach. Intell*. 2001, **23**: 1175-1191.

[4] G.G. Shishkin, A.G. Shishkin, I.M. Ageev, et al. A Novel Device for Recognition of Human Physiological State. In: *Proc. 6th WSEAS International Conf. on Applied Electromagnetics, Wireless and Optical Communications*. 2008, pp. 15-20.

[5] I.M. Ageev, and G.G. Shishkin. Water Conductivity Changes During Its Heating by Different Heat Sources Including Biological Objects. *Biophysics*. 2001, **46**: 829-832.

[6] H.M. Teager, and S.M. Teager. Evidence for Nonlinear Sound Production Mechanisms in the Vocal Tract. In: W.J. Hardcastle, and A. Marchal (eds) *Speech Production and Speech Modelling*. NATO Advanced Study Institute, Series D, vol. 55, 1989.

[7] D. Talkin. A Robust Algorithm for Pitch Tracking (RAPT. In: K.K. Paliwal, and W.B. Kleign (eds) *Speech coding and synthesis*. Elsevier Science, Amsterdam, 1995.