Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 381–388.

METROLOGY AND MEASUREMENT SYSTEMS

Index 330930, ISSN 0860-8229

www.metrology.pg.gda.pl

EMOTION MONITORING – VERIFICATION OF PHYSIOLOGICAL CHARACTERISTICS MEASUREMENT PROCEDURES

Agnieszka Landowska

Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Software Engineering Department, Narutowicza 11/12, 80-233 Gdańsk, Poland (✉ nailie@eti.pg.gda.pl, +48 58 347 2989)

Abstract

This paper concerns measurement procedures on an emotion monitoring stand designed for tracking human emotions in the Human-Computer Interaction with physiological characteristics. The paper addresses the key problem of physiological measurements being disturbed by a motion typical for human-computer interaction such as keyboard typing or mouse movements. An original experiment is described, that aimed at practical evaluation of measurement procedures performed at the emotion monitoring stand constructed at GUT. Different locations of sensors were considered and evaluated for suitability and measurement precision in the Human-Computer Interaction monitoring. Alternative locations (ear lobes and forearms) for skin conductance, blood volume pulse and temperature sensors were proposed and verified. Alternative locations proved correlation with traditional locations as well as lower sensitiveness to movements like typing or mouse moving, therefore they can make a better solution for monitoring the Human-Computer Interaction.

Keywords: affective computing, emotion recognition, physiology, motion artifacts, sensor location.

1. Introduction

Human emotions influence the manner people act and work, but tracking effective and counter-productive emotional states is a challenge even for qualified psychologists [1–2]. Affective computing provides a large set of methods and tools for emotion recognition, that might be used for tracking affect in different contexts. Automatic affect recognition processes are based on symptoms of emotions, which are expressed by human body, including: facial expression, body language, voice and physiological signals [3]. Physiology, comprising such parameters as blood volume pulse, galvanic skin response (skin conductance), body temperature, and muscle tension, proved to have good accuracy in affect recognition [4]. However, the accuracy of biometric-based emotion recognition algorithms was not analyzed for suitability in monitoring the human-computer interaction.

The paper describes the construction of an emotion monitoring biometric stand, which is intended to be used in monitoring human-system interaction, especially tracking emotional states of a person working with systems. The stand was constructed at Gdańsk University of Technology in 2013. The main challenge was the sensitiveness of biometric sensors to movements and their location on fingers, which is inconvenient while using a mouse and/or a keyboard in the human-computer interaction. Therefore, an experiment was held, aiming at eliminating sensors from hands and finding alternative locations that are as good as typical finger location. The following experiment hypothesis is given: *it is possible to monitor the human-computer interaction reliably with biometric sensors placed in non-disturbing locations*. The experiment's design, execution and results are presented in the paper followed by the discussion of the outcomes. The paper also provides some guidelines for measuring physiological signals in the human-systems interaction.

Article history: received on Feb. 10, 2014; accepted on Jul. 05, 2014; available online on Dec. 15, 2014; DOI: 10.2478/mms-2014-0049.

2. Related work

Works that are mostly related to this research fall into the following categories: research on physiological signals in emotion recognition, biometric sensors location and studies on motion artifacts in physiological signals.

The automatic emotion recognition can be based on a diverse characteristics analysis that can be obtained from different input channels including: face muscle activity and other muscles activity (obtained from video image or electromyography sensors) [5–6], posture analysis (obtained from a video image or pressure matte) [7], behavioural patterns (from keyboard and mouse trackers or from video) [8–9], accent and modulation (voice) [10], and physiological parameters (from sensors or video) [4, 11]. Physiological parameters are perceived as the most reliable proxy for emotion recognition, as they are not controlled by most people [12]. The highest accuracy, however, is provided by emotion recognition algorithms based on multimodal inputs, which combine different channels and parameters [3], [13]; and that approach was chosen for the emotion monitoring stand construction at Gdańsk University of Technology.

The physiological parameters most commonly used in emotion recognition cover: heart rate, skin conductance, respiration, muscle activity, peripheral temperature as well as brain activity [14–17]. The heart rate can be obtained from an electrocardiographic sensor placed on the chest or from a blood-volume pulse sensor, placed on fingertips [14]. Alternative locations of the BVP sensor are toes and ear lobes [11, 15], and the latter was chosen for the proposed emotion monitoring stand. Electromyographic sensors in emotion recognition are placed on face muscles to track microexpressions and on a trapesius muscle (shoulder) because it reflects stress [11, 16]. In emotion recognition, the typical location of skin conductance sensors is the base of a ring and an index finger or a little and a middle finger, however, three alternative locations were mentioned: a forearm, palm and foot [17]. Respiration can be measured on the thorax or abdomen, while peripheral temperature is measured on fingertips. As a change of the signal is interpreted as a change of emotional state, one restriction must be made, namely, all sensors are sensitive to movements and in most of the experiments described in the literature, a subject is asked not to move the body parts that the sensors are attached to.

There are studies on motion artifacts that influence biometric characteristics, yet they focus mainly on medical applications f.e. [18]. In this context research concentrates predominantly on the elimination of respiratory motion as well as uncontrolled body twitch artifacts. In those studies diverse methods are employed for motion artifacts rejection, including adaptive filters, Bayes filters, Canonical Component Analysis, Blind Source Separation as well as methods and procedures based on hardware such as a differential analysis based on multiple electrodes [19, 20]. Motion artifacts are quantified in those studies using diverse time, frequency and time-frequency domains metrics [21].

Although there are multiple studies on emotion recognition based on biomeasurements, according to the author’s knowledge, the sensitiveness of biosensor readings to movements was not evaluated in monitoring the human-computer interaction before.

3. Emotion monitoring stand construction at GUT

The objective of the emotion monitoring stand is to conduct experiments on computer users affective states retrieval and analysis. The stand is equipped with computers, cameras and a set of biosensors, which allow to monitor user activities and record multiple user observation channels at the same time. In 2013, an emotion monitoring stand was constructed at Gdańsk University of Technology as a dedicated stand in a research laboratory room. The

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 381–388.

equipment of the stand was chosen on the basis of its capabilities and availability. The FlexComp Infiniti coder by Thought Technology, Canada was chosen as a biosensor analytical device. The coder is a ten-channel multimodal device dedicated for real-time biofeedback, psychophysiology training and monitoring. However, it was not dedicated for emotion recognition and not checked against suitability in the human-computer interaction. The coder is connected with a computer database via TT-USB device and allows to record simultaneously up to ten channels with the sampling rate of 2048 samples per second. The coder removes noise from all input channels and performs signal amplification and preliminary filtration that is adjusted to the sensor type.

The biometric sensor set for the emotion monitoring stand was configured to measure physiological parameters that are commonly used in affect recognition and the choice was justified by the literature review. The sensors are compatible with the FlexComp Infiniti coder and other coders produced by Thought Technology. Some of the sensors in the predefined set were doubled in order to try out multiple locations at the same time. A detailed list of the sensor types, values measured and accuracy as provided by the producer [15] is shown in Table 1. All of the sensors were produced by Thought Technology, Canada.

Table 1. Parameters of biometric sensors for emotion monitoring stand.

Sensor	Input values scope	Accuracy
skin conductance SC-Flex/Pro	0 – 30.0µS	±0.2µS ±5%
electromyography MyoScan Pro EMG	0 – 2000µVRMS	±0.3µVRMS ±5%
respiration RESP	relative measure 30–65%	not available
EEG-Z sensor	0 – 200µVRMS	±0.3µVRMS ±5% for 10-40°C
blood-volume pulse HR/BVP Flex/Pro	relative measure 0% – 100%	±5%
temperature TEMP	10-45°C	± 1.0°C

Thought Technology BioGraph Infiniti application was installed for gathering biometric data from the coder. The system was chosen according to the compatibility with the coder, but also basing on signal quality optimization features including verifying signal quality and adjusting the sensor location, integrated electrode impedance measurement as well as artifact rejection feature, both automatic and manual. Additionally, the optional Physiology Suite was installed, specifically designed for monitoring and assessing physiological functions: recording biomeasurement sessions, reviewing recorded data for the purpose of artifact rejection, generating session reports and demonstrating the results.

4. Experiment design

To verify the research hypothesis that *it is possible to monitor the human-computer interaction reliably with biometric sensors placed in non-disturbing locations,* an experiment was designed and conducted. In the study of the emotion monitoring stand hardware layer, two research methods were employed, namely a controlled experiment and a questionnaire.

The controlled experiment was carried out with a set of predefined independent variables including sensor locations that were manipulated to record differences in dependent variables (measured values from the biometric sensors). Full randomization of subject selection was not possible in a laboratory setting located in one place only; therefore, while assigning subjects to groups, randomization type II was performed. Moreover, the group of convenience was constructed in such a way that it represented some range of possible confounding variables such as the age and gender. The experiment was followed by a questionnaire to determine subjective assessment of disturbing monitoring.

4.1. Operationalization of variables and detailed thesis

To verify the overall paper research thesis on non-disturbing and reliable sensor locations, the analysis was decomposed into the evaluation of three detailed characteristics:

– subjective notion of disturbing and non-disturbing sensor locations (in HCI),
– sensitiveness of biomeasurements' values to body parts movements,
– correlation of measurements in on-finger and off-finger locations.

In other words, the first step of the hypothesis verification is to spot the locations that are subjectively the most disturbing during interaction with computers and that are the most interfered with typical movements like typing or mouse movement. The next step is to verify whether the alternative locations are good enough to be a reliable replacement for typical finger locations. That step required checking whether the outcomes of alternative sensor locations are similar or correlated with the typical finger location.

Sensors such as SC, BVP and TEMP are sensitive to movements and at the same time the human-computer interaction is based on intensive hand and finger mobility. Standard locations of BVP, TEMP and SC sensors include fingers; BVP and TEMP are placed on fingertips, while two SC sensors are placed at the finger base of a ring and an index finger or a little and a middle finger. These locations are not convenient for monitoring the human-computer interaction, as they disturb keyboard and mouse usage, therefore, alternative sensor locations had to be identified and tested. Based on a literature analysis, alternative locations were chosen, and the following ones were tested: a forearm for the skin conductance sensor and ear lobes for the temperature and pulse sensors. The detailed information on dependent and independent variables and the experiment thesis is provided in Table 2.

Table 2. Operationalization of variables and detailed thesis for the experiment.

Research objective 1. Subjective evaluation of sensor interference with HCI	
Independent variable:	sensor location
Dependent variable:	subjective feeling of comfort (measured with questionnaire)
Hypothesis	Off-finger locations of sensors are less disturbing when compared to finger locations
Null hypothesis	Off-finger locations of sensors are more or the same disturbing when compared to finger locations
Research objective 2. Sensitiveness to motion	
Independent variable:	movement (keyboard typing, mouse movements, head movements, body movements) and lack-of-movement (as baseline)
Dependent variable:	biometric values provided by the sensor (values depending on sensor types, sensor placed in on-hand and off-hand locations)
Hypothesis	Measurements at off-finger location of sensors is less disturbed by body part movements than typical finger location
Null hypothesis	Both on-finger and off-finger locations of sensors are significantly disturbed by body part movements
Research objective 3. On-fingers and off-fingers sensor locations	
Independent variable:	sensor location with two values assigned: finger location (SC: finger base, BVP, TEMP: fingertips) and off-finger location (SC: forearm, TEMP and BVP: ear lobe)
Dependent variable:	biometric values provided by the sensor
Hypothesis	Readings from an alternative sensor location are good proxy for measuring change in bioparameter value as provided with a typical (finger) location
Null hypothesis	Readings from an alternative sensor location do not correlate with bioparameter values from a typical location

In this experiment, four movements typical in the human-computer interaction were considered, including typing with a keyboard, moving a mouse, moving the head (as looking

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 381–388.

at multiple monitors or moving the head forward to see a small detail on the screen) and changing the body position (slight chair movements).

A decision was made to combine the research hypothesis verification in one experiment, due to the following reasons. The verification of research objective 1 (subjective evaluation of the sensor locations) must follow the measurement procedure, as otherwise disturbance would be evaluated only hypothetically. For the rest of the hypothesis (2 and 3), the thesis verification requires the same subject group and the same dependent variable is measured. Both typical and alternative locations must be tested against sensitiveness to motion (hypothesis 2).

4.2. Experiment plan

The experiment was held using the emotion monitoring stand described in point 3, however, the results may apply to other biometric stands as well, especially those equipped with similar or equivalent biometric equipment. There were two test groups differing in skin conductance sensor location, as two sensors working at the same time interfere with each other and falsify the measurements. The temperature and BVP sensors were doubled in each group. There was no control group, as the treatment and measurement are the same activity. A random assignment to test groups was performed (even subject numbers were assigned to group A, while odd numbers to group B, all numbers being assigned in a sequence of attendance, which was not controlled by the investigator).

The experiment process consisted of the following steps: information on the investigation and signing a consent, number assignment and group assignment, attachment of sensors, measurement sequence 1, re-attachment of sensors, measurement sequence 2 and filling in a personal characteristics form (age, gender) and a questionnaire on disturbing monitoring.

The first measurement sequence consists of attachment of the sensors and baseline measurement (rest measurement). In the second measurement sequence the baseline is recorded first and then four different movement patterns are measured, intermediated with a rest. In the movement measurement the subject is guided with an instruction displayed on the additional monitor. The detailed sensor location is presented in Fig. 1.

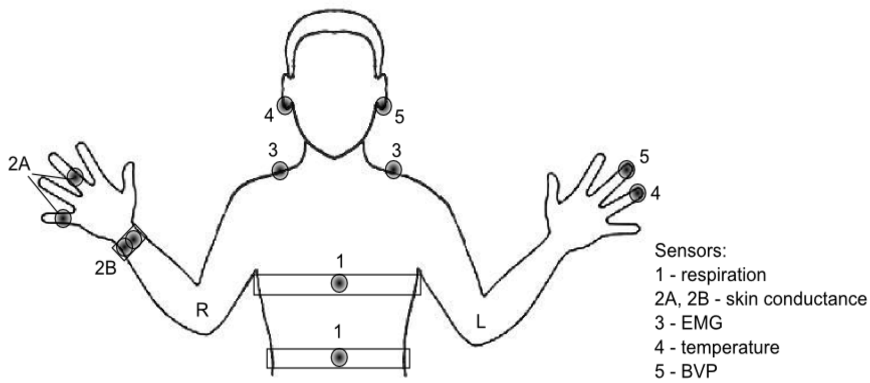

Fig. 1. Sensor location in emotion monitoring verification experiment.

Sensors assigned only with numbers were placed in these locations for both test groups. Sensor 2 (skin conductance) was located at fingers (2A) for test group A and at a forearm (2B) for test group B.

The questionnaire on disturbing monitoring was passed to the subjects after both measurement phases and it was intended to survey individual subjective evaluation of comfort in monitoring environment, as well as to identify the most disturbing locations, both in the experimental setting and in the human-computer interaction. The questionnaire started with a sensor location picture (similar to the one provided in Fig. 3, however differentiated among the groups), in which all sensor locations were assigned with single letter captions. Questions were asked on subjective evaluation of sensor location disturbance during the experiment and the human-computer interaction.

4.3. Emotion monitoring study execution

A group of 31 subjects was involved in the experiment, including 13 females and 18 males, with age ranging from 21 to 59 years. During the experiment, subjects were instructed not to move or to perform specific movements and that instruction was provided on the screen in front of them. An exemplary session recording in Thought Technology application is provided in Fig. 2. Different measurement phases are separated with vertical lines and their description is presented at the bottom of the window. As it can be seen, there is a leap in most of the parameters between baseline 1 and baseline 2, where sensors were detached and re-attached, illustrating sensitiveness to sensor relocation. For most of the parameters some disturbances, caused by different movement phases (keyboard usage, mouse movement, head movement and body movement on a chair, respectively), can be observed starting from minute 2.

Fig. 2. Exemplary session track visualization of selected biometrics.

The sampling rate in data acquisition was 2048 Hz for all sensors, although such a value is required only for the BVP sensor. Data set from the experiment for 31 subjects exported with 256 Hz and timestamp is available at the Emotion in HCI Research Group website [22].

5. Experiment results

The experiment results were analyzed in the following order: first, a subjective evaluation of the sensor location was summarized to identify the most disturbing sensor locations in the human-computer interaction, then the locations were tested against sensitivity to the movements. For selected sensors (with two alternative locations) correlation analysis of measurements was performed.

5.1. Subjective evaluation of sensor locations

After the measurement was completed, the subjects were asked to assign the sensor locations with numbers from 0 to 3 (0– non-disturbing, 3– very disturbing). As the scale in the questionnaire was ordinal, in analysis position metrics were used instead of means.

People found EMG sensor location on the trapesius muscle and forearm location of the skin conductance sensor the least disturbing both during the experiment and in HCI (30 out of 31 subjects rated EMG location with 0). Both locations of respiration sensors were perceived as non-disturbing (with mode 0), however the location on thorax was a little more disturbing (with median 1) than the location on abdomen (median 0).

Finger locations of the BVP and TEMP sensors were found most disturbing (with both mode and median equal to 3), while alternative locations of those sensors on ear lobes were significantly less disturbing (with both mode and median equal to 0). Skin conductance location on fingers was less disturbing than the one of the BVP and TEMP sensor (median 2), although it was on the right hand used for the mouse movement. The skin conductance forearm location was considered as one of the least disturbing during the experiment and in HCI (median and mode equal to 0). Histograms of the notes for typical and alternative locations of blood-volume pulse, temperature and skin conductance sensors are provided in Fig. 3.

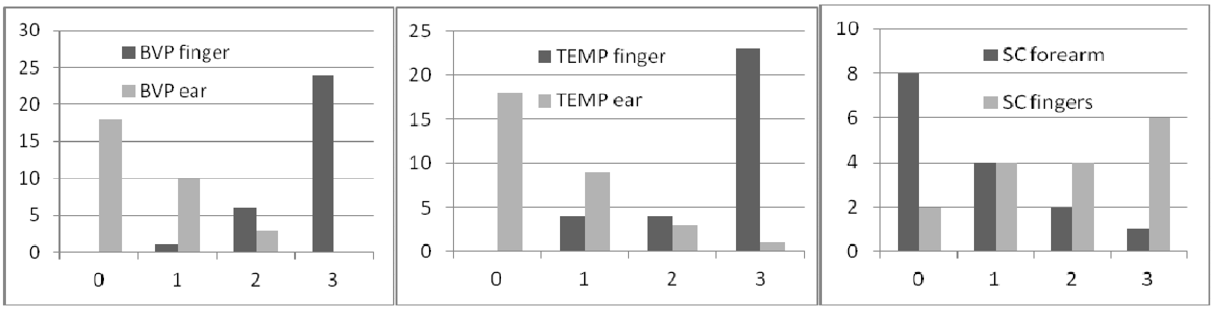

Fig. 3. Subjective evaluation of sensor locations – histograms for BVP, temperature and SC sensors.

Verification of the hypothesis on alternative and typical locations of the temperature, blood-volume pulse and skin conductance sensors, requires checking significance of the observed differences and therefore Mann-Whitney U test for BVP and temperature and Wilcoxon test for skin conductance were calculated, as they can be used for ordinal data comparison.

The Mann-Whitney independent observations test was calculated for the blood volume pulse and temperature with results of U metric 17 and 55.5 respectively, which allows to accept the hypothesis on less disturbing alternative locations with $p<0.0001$. The Wilcoxon independent observations test was calculated for skin conductance providing W value 2.5, which allows to accept the hypothesis on less disturbing alternative locations with $p<0.02$.

Based on this analysis, a hypothesis can be accepted, suggesting that the chosen off-finger location of sensors (ear lobes for temperature and blood volume pulse sensors and forearm for skin conductance sensor) are less disturbing in comparison to finger locations.

5.2. Sensitiveness to body part movements

As biometric sensors readings are sensitive to human body movements, it is important to measure the effect in different sensor locations. In this experiment four movements were considered, including typing with a keyboard, moving a mouse, moving the head and changing the body position. All of the movements were performed separately. Descriptive statistics such as mean, standard deviation and maximum for the baseline (no movement) and all four movement types are provided in Table 3. Values are averaged for all participants. EMG is provided in μVRMS, skin conductance in μS, temperature in Celsius scale, BVP signal was recalculated to retrieve heart rate, while respiration is provided in a relative % measure.

Table 3. Descriptive statistics of baseline and movements for all participants.

Sensor	Location	Baseline		Movement							
				Keyboard typing		Mouse movement		Head movement		Body movement	
		Mean (SD)	Max	Mean (SD)	Max	Mean (SD)	Max	Mean (SD)	Max	Mean (SD)	Max
EMG	Trapesius muscle right	6.5 (3.1)	47.9	16.4 (13.9)	160.1	42.9 (31.9)	281.5	8.7 (6.3)	68.3	35.7 (32.7)	287.3
EMG	Trapesius muscle left	5.1 (3.9)	32.8	12.5 (10.2)	105.1	27.7 (18.5)	243.7	6.9 (5.09)	56.0	34.5 (21.8)	294.9
RESP	Thorax	16.7 (0.7)	20.1	15.6 (0.6)	17.5	15.6 (0.9)	18.5	16.4 (0.53)	17.9	16.8 (1.9)	23.4
RESP	Abdomen	25.7 (0.9)	28.6	25.8 (0.7)	28.1	25.9 (0.9)	28.4	25.9 (0.6)	27.7	25.2 (1.8)	30.5
TEMP	Finger	34.3 (0.07)	34.6	34.4 (0.07)	34.6	34.2 (0.09)	34.4	34.2 (0.04)	34.3	34.4 (0.08)	34.5
TEMP	Ear lobe	32.7 (0.04)	32.8	32.8 (0.02)	32.8	32.9 (0.02)	32.9	33.0 (0.01)	33.0	33.1 (0.01)	33.1
BVP (Heart Rate)	Finger	69.9 (10.8)	85.8	77.2 (24.6)	140.2	73.5 (18.6)	108.6	67.2 (15.1)	85.5	83.4 (24.5)	138.7
BVP (Heart Rate)	Ear lobe	69.7 (11.0)	81.8	71.5 (16.1)	83.6	71.9 (16.4)	85.1	67.1 (14.7)	82.6	80.9 (18.9)	102.6
SC	Fingers	1.10 (0.07)	1.30	1.31 (0.11)	1.52	1.79 (0.14)	2.01	1.59 (0.09)	1.80	1.86 (0.22)	2.22
SC	Forearm	0.74 (0.04)	1.36	0.85 (0.05)	0.96	0.89 (0.05)	1.00	0.88 (0.04)	0.99	0.98 (0.11)	1.25

In the movement analysis there is no point comparing the means, as the mean of a biometric signal may not express peaks caused by movements (two different functions e.g. linear and periodical can have the same mean). Instead, an analysis of variability should be performed to express the possible disturbance of the signal by movements. The following metrics in time domain can be used for quantification of the movement artifacts: variance, range, maximum value, power, absolute difference, relative difference, impulse count, autocorrelation, moving average. The choice of the metric should be justified by correlation with observable disturbance as well as computational complexity. As variance and maximum values showed accurate correlation with the observations as well as low complexity, they were chosen for the comparison of sensor locations. Also an analysis of the EMG outputs provided some confirmation to the correctness of the chosen approach. As trapesius muscles

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 381–388.

are involved in the movements performed by the subjects, EMG readings are severely affected and that is reflected in the variance metric. There is also a noticeable difference between the two locations on right and left shoulders during the mouse movement, as the right hand was involved (there were no left-handed representatives within the group) and that difference is also reflected in the variance metric. An additional benefit of the variance metrics is that there are statistical tests directly designed for variance comparison. Moreover, research is being conducted by the author, aiming at analyzing the movement artifacts from this experiment with the use of metrics from frequency or time-frequency domain.

A detailed analysis of each sensor readings is performed using F-Snedecor statistical test that allows to estimate significance of difference in variance. Detailed results of F-test for all sensor locations and movements are provided in Table 4 (where *F* is calculated F-test value, *df* – degrees of freedom, thresholds are obtained form F-Snedecor test tables, and the final test result is denoted as S- significant or IS- insignificant).

Table 4. F-Snedecor test results on significance of differences between baselines and movements.

Sensor	**Location**	**F-test parameters ($p<0.01$)**		**F values**							
				Keyboard typing		**Mouse movement**		**Head movement**		**Body movement**	
		df	**threshold**	**F**	**S/IS**	**F**	**S/IS**	**F**	**S/IS**	**F**	**S/IS**
EMG	Trapesius muscle right	30	2.386	20.10	**S**	106.40	**S**	4.14	**S**	111.36	**S**
EMG	Trapesius muscle left	30	2.386	12.29	**S**	80.83	**S**	2.99	**S**	145.47	**S**
RESP	Thorax	30	2.386	0.63	IS	1.53	IS	0.50	IS	6.93	**S**
RESP	Abdomen	30	2.386	0.70	IS	1.13	IS	0.55	IS	4.51	**S**
TEMP	Finger	30	2.386	1.00	IS	1.57	IS	0.37	IS	1.56	IS
TEMP	Ear lobe	30	2.386	0.21	IS	0.18	IS	0.12	IS	0.08	IS
BVP (HR)	Finger	30	2.386	5.19	**S**	2.98	**S**	1.97	IS	5.16	**S**
BVP (HR)	Ear lobe	30	2.386	2.14	IS	2.23	IS	1.79	IS	2.96	**S**
SC	Fingers	15	3.485	2.24	IS	4.00	**S**	1.39	IS	9.50	**S**
SC	Forearm	14	3.698	1.69	IS	1.44	IS	1.03	IS	7.80	**S**

Differences of respiration signal variability in the baseline and movement measurements in two locations (thorax and abdomen) also exist, though, they were found insignificant for the keyboard typing, mouse movements and head movements, while significant only for the body movements (detailed *F*-values are provided in Table 4).

Differences of temperature signal variability for the baseline and movement measurements on both the ear lobe and fingers were found insignificant. This result is natural as the overall variability of temperature signals is very low (standard deviations range is 0.01-0.09). Nevertheless, F-values calculated for the ear lobe are considerably lower than for finger location of the sensor (e.g. five times lower for the keyboard typing).

Differences of the heart rate (derived from the blood-volume pulse signal) variability for the baseline and movement measurements on the fingers are significantly disturbed by the keyboard typing, mouse movements and body movements, but not by head movements. The location of the BVP sensor on the ear lobe is less disturbed – the difference in variability was found significant only for the body movements.

Differences of skin conductance signal variability for the baseline and movement measurements on fingers were found insignificant for the keyboard typing and head movement, yet significant for mouse movements and body movements. The location of the skin conductance sensor on a forearm is less disturbed – the difference in variability was found significant only for the body movements. The body movements cause larger variability for almost all sensors, and therefore such large movements must be rejected as artifacts.

Based on this analysis, a hypothesis stating that measurements at off-finger location of sensors is less disturbed by the body part movements than the typical finger location, can be accepted for skin conductance and BVP (heart rate). However, it must be rejected for temperature (sensitiveness to movements for both locations was found insignificant).

5.3. Comparison of typical and alternative location of sensors

Measurements in alternative sensor locations were evaluated for correlation with typical sensor locations. Typical locations on fingers are used for affect recognition from skin conductance, temperature and blood volume pulse, however, due to great variability of baseline values among people, in recognition algorithms a change vector is used instead of absolute values [14]. This allows to assume that if readings in both locations are correlated (signal variability is concurrent in time and sign), then readings (change vectors) from typical locations can be replaced with the proposed alternatives on ear lobes. For the purpose of this analysis, only baselines were used (two recordings per subject). Both the temperature and BVP were concurrently recorded in both locations at the same time. Due to interference of two concurrent skin conductance sensors, this input was excluded from the correlation analysis. In the correlation analysis, Pearson linear correlation coefficients are calculated for individual baseline recordings. As Person coefficient is not an additive metric, therefore a determination coefficient (Pearson squared) is calculated to estimate average values for all recordings. For Pearson coefficient the t-test was calculated to confirm significance of the result. The histogram of Pearson coefficients for the temperature and BVP was provided in Fig. 4.

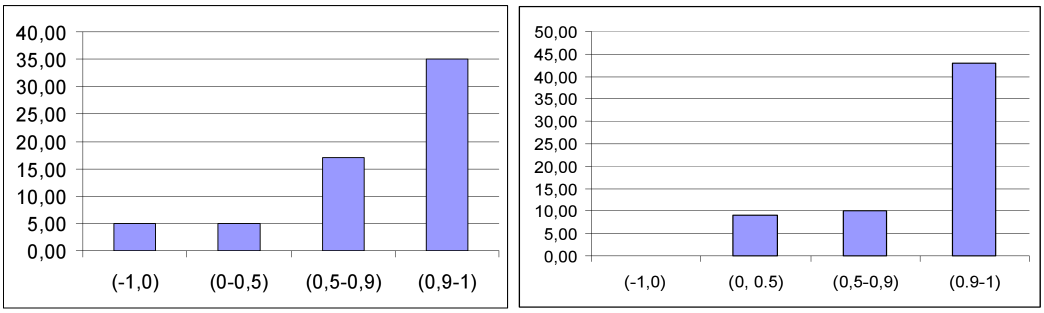

Fig. 4. Histogram of individual Pearson coefficients for finger and ear lobe location of temperature sensor (left) and heart rate (BVP sensors) (right).

There is observable great interpersonal variability of the finger and ear lobe temperature, i.e. for 24 subjects average temperature on fingers was higher than on ear lobes and for the other 7 the opposite was the case. For selected subjects the highest temperature measured on fingers (i.e. 28.87 degrees Celsius) was lower than the minimum value for other subjects (i.e. 34.73). Thus a detailed analysis was performed and Pearson linear correlation coefficients were calculated for each of the experiment participants. The detailed analysis of individual correlation for each subject showed different results with average determination factor of 78%. For 52 out of 62 measurements (over 84%) Pearson coefficient was higher than 0.5 (strong positive correlation), whereas for 35 measurements (56%) the coefficient was higher than 0.9 (very strong positive correlation).

Metrol. Meas. Syst., Vol. XXI (2014), No. 4, pp. 381–388.

A detailed analysis was performed for the cases in which the correlation coefficient was lower than 0.5 (and sometimes even negative). For most of them the range of measurements was lower than the sensor accuracy. The average determination coefficient of measurements in two sensor locations was 72.7%, and after eliminating measurements with the range (max-min per baseline) lower than 0.2 degrees Celsius, it amounted to 77.7%. The T-test (t =10.09) confirmed the significance of calculated Pearson correlation coefficients with p<0.0005. This allows to accept a hypothesis that readings from the alternative location of the temperature sensor (ear lobe) are good proxy for temperature change measurements provided with the typical (finger) location.

An analogous procedure was performed for the analysis of blood-volume pulse locations on fingers and ear lobes. For correlation calculation, the heart rate derived from power spectrum was considered, as the blood volume pulse is a periodic signal. In the emotion recognition, change in the heart rate is used rather than absolute values or raw blood-volume pulse signal [14]. For 53 out of 62 measurements (over 85%) Pearson coefficient was higher than 0.5 (strong positive correlation), while for 43 measurements (69%) the coefficient was higher than 0.9 (very strong positive correlation). No cases of negative correlation were observed.

The average determination coefficient was 75.9% for all of the baseline recordings, whereas after eliminating recordings with the lowest range, it rises up to 81.1%, which means that 81.1% of one location measurement variability is explained with the variability of the alternative sensor location readings. The T-test (t =13.61) confirmed the significance of the calculated Pearson correlation coefficients with p<0.0005. This allows to accept a hypothesis stating that readings from the alternative location (ear lobe) of the BVP sensor is a good proxy for measuring change in the heart rate as provided with the typical (finger) location.

6. Discussion

In the evaluation of all sensor locations, multiple criteria were taken into account: sensitiveness to movements, correlation with the typical location (for alternative ones) and a subjective evaluation of disturbance in the human-computer interaction.

The summary of the experiment results is provided in Table 5 (S-significant result, IS – insignificant results).

Table 5. Summary of sensor location evaluation (S-significant. IS- insignificant).

Sensor	**Location**	**Level of disturbance in HCI**	**Sensitiveness to body part movements**				**Correlation with typical locations (determination)**
			Keyboard typing	**Mouse move**	**Head move**	**Body move**	
EMG	Shoulder (left)	Almost None	**S**	**S**	**S**	**S**	-
EMG	Shoulder (right)	Almost none	**S**	**S**	**S**	**S**	-
RESP	Thorax	Small	IS	IS	IS	**S**	-
RESP	Abdomen	Small	IS	IS	IS	**S**	-
TEMP	Finger	**Large**	IS	IS	IS	IS	-
TEMP	Ear lobe	Small	IS	IS	IS	IS	78% (S)
BVP (HR)	Finger	**Large**	**S**	**S**	IS	**S**	-
BVP (HR)	Ear lobe	Small	IS	IS	IS	**S**	81% (S)
SC	Fingers	**Moderate**	IS	**S**	IS	**S**	-
SC	Forearm	Small	IS	IS	IS	**S**	not available

The experiment allowed to draw some conclusions on human-computer interaction monitoring based on bio-measurements of muscle electric activity, respiration, temperature, pulse or skin conductance:

1. The emotion recognition in the human-computer interaction should not use EMG measurements placed on a trapesius muscle. Although this muscle is the one that reflects stress, the signal is significantly disturbed by the mouse movement and keyboard typing.
2. Both locations of the respiration sensor are perceived as low disturbing and insensitive to movements, except from the body movements on a chair. For the signal recorded by those sensors, artifacts connected with large movements (body movements) should be removed.
3. The temperature sensor placed on finger is one of the most disturbing and therefore should be moved to an ear lobe.
4. For the temperature sensor, an alternative location provides acceptable estimate of the value obtained from the typical finger location, nevertheless, the change may be too small to be interpreted as a symptom of emotions (the temperature ranges were small for both locations).
5. The BVP sensor placed on fingers also belongs to the most disturbing ones and should be moved from fingers to the alternative location on ear lobes. For the signal recorded in the latter location, the artifacts connected with large movements should be removed. For the BVP sensor, the alternative location on an ear lobe provides acceptable estimate of the value obtained from the typical finger location.
6. The skin conductance sensor location on a forearm is perceived as less disturbing than the location on fingers. The location on a forearm is also less sensitive to the mouse movements.
7. For the signal recorded by all sensors, artifacts connected with large movements (body movements) should be removed regardless of their location on the body.

In emotion recognition algorithms, which are based on physiological signals, relative values should be provided rather than absolute values, as there are significant differences between individuals. A normalization or standardisation procedure should be performed with a personal average, instead of an overall average calculated for all subjects. This requires baseline recordings before tracking stimuli response. Baseline recording is important for experimental settings as well as natural environments.

The application of emotion monitoring based on biometrics in real-time application run in a non-experimental setting is also a challenge, since it requires some algorithms for baseline extraction and algorithms for the automatic rejection of signal artifacts resulting from the movements.

Experiment validity

There are two main threats to external validity of this research that are typical for laboratory experiments: conducting biomeasurements on one stand in one location only and using a group of convenience instead of full randomization from the whole population. However, the results may apply to stands with similar or equivalent equipment as operation methods of sensors from other vendors are similar, so it might be expected that they would show the same characteristics. Moreover, such an experiment and analysis of sensitiveness to the movement and individual variability should be performed for new stands and equipment that are to be used for measuring biometric parameters in any atypical contexts. A group of convenience in this experiment was constructed in such a way that it represents some range of confounding variables such as age and gender. Moreover, randomization type II was performed while assigning subjects to test groups. The size of an experiment group (31 subjects) was convenient for verification of the hypothesis for most parameters. One restriction should be made for testing a hypothesis on relocation of the skin conductance sensor. Although the difference between baseline averages was large, the test found that difference insignificant. which was caused by a too small testing group.

Internal validity of this experiment is threatened by the testing effect (testing causes change of parameters) and Hawthorne effect (people behave differently when they know that they are being observed). Both effects are very hard to eliminate in measurements which require placement of sensors on the skin. There are works performed on measuring pulse form video [23] and measuring skin conductance with a mouse [24]; however, both solutions are in an experimental phase and are not yet available for broad applications. In this experiment a subjective evaluation of the sensor location disturbance level was performed post-test, after the sensors had been detached but immediately after the measurement phase, in order to avoid a history effect in the verification of the hypothesis on disturbing monitoring.

7. Conclusions

The problem of measurements sensitiveness to motion must be taken into account in both experimental and application contexts; yet, it had not been analyzed this way before in monitoring the human-computer interaction. The main contribution of the experiment lies in identifying the sensor locations which were neither significantly influenced by the keyboard typing and mouse movements, nor perceived as disturbing. This allows to accept the hypothesis that *it is possible to monitor the human-computer interaction reliably with biometric sensors placed in non-disturbing locations.*

There are wo main challenges that must be addressed in real-life applications, namely natural baseline recording and automatic rejection algorithms for movement artifacts. In experimental settings it is easy to spot a missing, noisy or disturbed signal as it is monitored, whereas in practical applications the algorithms must be prepared to deal with missing or noisy input channels.

Future works on emotion monitoring stand include building the software layer – especially an algorithm for combining emotion recognition from diverse recognition algorithms, which would be based on different input channel sets and emotion visualization tool. The emotion monitoring stand is planned to be used in quantification of many observations from the human-computer interaction [25–26].

Acknowledgements

This study was partially supported by the Foundation for Polish Science within the Grant no 173/UD/SKILLS/2012 and by Polish-Norwegian Research Programme operated by the National Centre for Research and Development under the Norwegian Financial Mechanism 2009-2014 within Project Contract no Pol-Nor/210629/51/2013 as well as by DS Programs of the ETI Faculty, Gdańsk University of Technology.

Literature

[1] Boehner, K., dePaula, R., Dourish, P., Sengers, P. (2007). *How emotion is made and measured*. Int. J. Human-Computer Studies. 65, 275–291.

[2] Landowska, A. (2013). *Affect-awareness Framework for Intelligent Tutoring Systems*. In *Proc of HSI*. Gdańsk. Poland. 540–547.

[3] Gunes, H., Schuller, H. (2013). *Categorical and dimensional affect analysis in continuous input: Current trends and future directions*. Image and Vision Computing. 31. 120–136.

[4] Szwoch, W. (2013). *Using physiological signals for emotion recognition*. In *Proc of HSI*. Gdańsk. Poland. 556–561.

[5] Wang, J., Yin, L., Wei, X., Sun, Y., (2006). *3D facial expression recognition based on primitive surface feature distribution*. Comp. Vision and Pattern Recogn.. 2. 1399–1406.

[6] Zhang, J., Lipp, O., Oei, T., Zhou, R. (2011). *The effects of arousal and valence on facial electromyograpjic asymmetry during blocked picture viewing*. International Journal of Psychophysiology. 79. 378–384.

[7] Gunes, H. M., Piccardi, M. (2005). *Affect Recognition from Face and Body: Early Fusion vs. Late Fusion*. IEEE International Conference on Systems. Man and Cybernetics. 4. 3437–3443.

[8] Vizer, L. M., Zhou, L., Sears, A. (2009). *Automated stress detection using keystroke and linguistic features*. Int. Journal of Human-Computer Studies. 67. 870–886.

[9] Kołakowska, A. (2013). *A review of emotion recognition methods based on keystroke dynamics and mouse movements*. In *Proc of HSI*. Gdańsk. Poland. 548–555.

[10] Zeng, Z., Pantic, M., Roisman, G., Huang, T. (2009). *A survey of affect recognition methods: Audio. visual. and spontaneous expressions*. IEEE Transactions on Pattern Analysis and Machine Intelligence. 31(1). 39–58.

[11] Healey, J., Picard, R. (2005). *Detecting stress during real-world driving tasks using physiological sensors*. IEEE Trans. on Intelligent Transportation Systems. 6 (2). 156–166.

[12] Lisetti, C., Nasoz, F. (2004). *Using noninvasive wearable computers to recognize human emotions from physiological signals*. EURASIP Journal on Applied Signal Processing. 11. 1672–1687.

[13] Mauss, I., Robinson, M. (2009). *Measures of emotion: A review*. Cogn. Emot.. 23(2). 209–237.

[14] Jang, E., Rak, B., Kim, S., Sohn, J. (2012). *Emotion classification by Machine Learning Algorithm using Physiological Signals*. In. Proc. of Computer Science and Information Technology. Singapore. 25. 1–5.

[15] *BioGraph Infiniti and FlexComp Infiniti User Manual*. Thought Technology. Canada. *www.thoughttechnology.com* (VII 2013).

[16] Boxtel, A. *(2010). Facial EMG as a Tool for Inferring Affective States. In Proc. of Measuring Behavior. Eindhoven. The Netherlands.* 104–108.

[17] Bach, D., Flandin, G., Friston, K., Dolan, R. (2010). *Modelling event-related skin conductance responses*. Int J Psychophysiol. 75(3). 349–356.

[18] Yousefi, R., Nourani, M., Ostadabbas, S., Panashi, I. (2013). *A motion-tolerant adaptive algorithm for wearable photoplethymographic biosensors*. IEEE J. Biomed. Health Inform.. 18. 670–681.

[19] Sweeney, K. T., Ward, T. E., McLoone, S. F. (2012). *Artifact Removal in Physiological Signals – Practices and Possibilities*. IEEE Transactions on Information Technology in Biomedicine. 16 (3). 488–500.

[20] Sweeney, K. T., Kelly, D., Ward, T. E., McLoone, S. F. (2011). *A Review of the State of the Art in Artifact Removal Technologies as used in an Assisted Living Domain*. IET Conference on Assisted Living.

[21] Sweeney, K., McLoone, S., Ward, T. (2010). *A simple bio-signals quality measure for in-home monitoring*. 7th IASTED International Conference.

[22] Emotions in Human-Computer Interaction Research Group. www.emorg.eu (I 2014).

[23] Balakrishnan, G., Durand, F., Guttag, J. (2013). *Detecting Pulse from Head Motions in Video. IEEE Conf. on* Computer Vision and Pattern Recognition (CVPR). 3430–3437.

[24] Nimwegen, C., Leuven, K. (2009). *Unobtrusive physiological measures to adapt system behavior: The GSR mouse*. Presented at Key issues in sensory augmentation. Institute for Media Studies. KU Leuven. Nederlands.

[25] Landowska, A. (2013). *Affective computing and affective learning – methods. tools and prospects*. EduAction. Electronic education magazine. 1(5). 16–31.

[26] Kołakowska, A., Landowska, A., Szwoch, M., Szwoch, W., Wrobel, M. R. (2013). *Emotion Recognition and its Application in Software Engineering*. In HSI *Proc.*. Poland. 532–539.