

Identification of Human Response to Virtual 3D Face Stimuli

Egidijus Vaškevičius¹, Aušra Vidugirienė², Vytautas Kaminskas³

*Department of Systems' Analysis, Vytautas Magnus University
Vileikos st. 8, LT – 44404, Kaunas Lithuania
e-mail: ¹e.vaskevicius@if.vdu.lt, ²a.vidugiriene@if.vdu.lt,
³v.kaminskas@if.vdu.lt*

crossref <http://dx.doi.org/10.5755/j01.itc.43.1.5927>

Abstract. This paper introduces identification results of human response to virtual 3D face stimuli. Observations of human reactions are done using preprocessed EEG (electroencephalogram) signals: excitement, meditation, frustration, engagement/boredom. Virtual 3D face features – distance between eyes, nose width, and chin width – are used as stimuli. Cross-correlation analysis demonstrated that dynamical relations between human reactions and stimuli exist. Input-output models describing relations between stimuli and corresponding human reactions are built. A new input-output model building method is proposed that allows building stable models with the least output prediction error. Models' validation results demonstrate relatively high prediction accuracy of human reactions.

Keywords: 3D face stimuli; human reaction; cross-correlation analysis; input-output model; model parameter estimation; model validation.

1. Introduction

Virtual environments are becoming a part of our daily life including computer games, work tasks, various mental and physical training programs, e-learning, online shopping, and specialized software. These environments affect the users in different extents and ways. The influence can be both positive and negative. Various studies introduce investigations, where human state observation is an important task. Several of them concentrates on stress evaluation and training when using virtual environments for military purposes and curing post-traumatic stress disorder [15]. Other focuses on neuro-marketing [11], [12], adaptive virtual mediator services [14], social networks [22], and learning applications [1], [4].

In order to protect people from various harmful effects to their health (e.g. stress) that virtual environment can cause [17], to create training applications, or to implement/use virtual mediator services [7], control mechanisms of virtual environment object features and the interactions have to be modeled [5], [10].

The most effective way to observe a state of a person in a real time is to monitor human bio-signals [19]. A variety of human bio-signals such as galvanic skin response, pulse, magnetoencephalography (MEG), heart rate, EEG, etc. are used for this purpose and plenty of methods and techniques help to

measure and analyze the signals [23]. We have chosen to use EEG-based signals for human affective state monitoring because of the reliability and quick response [18], [6].

In this research, a virtual 3D face was used as a stimulus for eliciting human reaction. It is known that the majority of information to another person when communicating is transferred by face features [13]. A person is used to react to the smallest face feature changes during a very short time [21]. We have created and used a realistic virtual 3D face model. Human reactions when he/she is observing visual stimuli (3D face changes) were investigated.

Our previous investigations on the dependencies between virtual stimuli and human affective responses demonstrated that in certain circumstances the dependencies exist and there is a need of further research [20].

Three types of stimuli in virtual 3D face (distance between eyes, nose width and chin width) were used for human reaction elicitation and four response signals (excitement, meditation, engagement/boredom, and frustration) were observed. The aim of the investigation is to identify the relationships between stimuli and reactions to them and to describe those using mathematical models that can predict human reactions in high accuracy and can be used for human state control systems design.

2. Observations and data

Features of virtual 3D face (distance between eyes, nose width, and chin width) were used for input and EEG-based human reaction signals (excitement, meditation, frustration, engagement/boredom) of a person were measured as output (Fig. 1). These output signals were chosen as important features in learning, and stress regulation. The output signals were recorded using Emotive Epoc device. The device records EEG inputs from 14 channels (according to international 10-20 locations): AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4 [3].

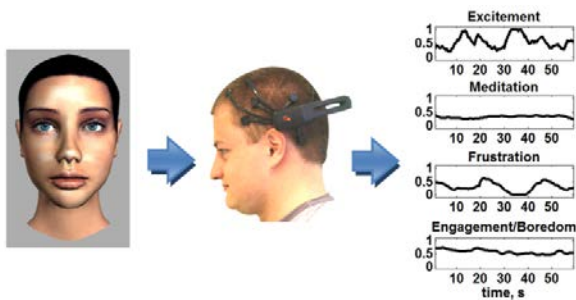


Figure 1. Input-Output scheme for the experiments

Seven 3D woman faces were created using Autodesk MAYA. One 3D woman face was used as a “neutral” one (Fig. 2, middle). The other faces were formed by changing face features in an extreme manner: large and small distance between eyes (Fig. 2, left and right, top), wide and thin nose (Fig. 2, left and right, middle), wide and thin chin (Fig. 2, left and right, bottom). After importing the 3D faces into Unity 3D engine, the transitions between the same kind of features in the face were programmed using *morph target's* method.



Figure 2. A “neutral” 3D woman face (middle) and limit cases of features: the largest (left top) and the smallest (right top) distance between eyes, the widest (left middle) and the thinnest (right middle) nose, the widest (left bottom) and the thinnest (right bottom) chin

There were two types of input signals used. TYPE 1 input signals (Fig. 3, top) were formed when neutral face was equal to 0, the largest (widest) feature was equal to 1.8 and the smallest (thinnest) feature was equal to -1.8. The values in-between were changed linearly. Time interval between

largest/smallest value and zero (normal face) was 10 s. The features were changing continuously: from normal to extremely wide, then back to normal and to extremely thin, then again to normal and to extremely wide and again back to normal. The animated tests were prepared for every of three features: distance between eyes, nose width and chin width.

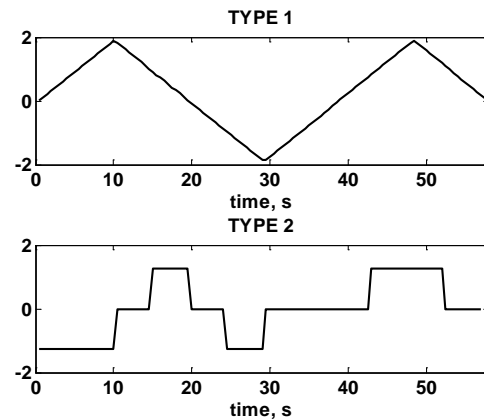


Figure 3. TYPE 1 and TYPE 2 input signals: values variation of corresponding face feature modifications

TYPE 2 input signals (Fig. 3, bottom) were formed when changing the 3D face features suddenly. “Small distance-between-eyes”, “thin nose” or “thin chin” was shown for some period of time, then the picture was suddenly changed to “normal” and after some time to “large distance-between-eyes”, “wide nose” or “wide chin” and again to “normal”. It was repeated two times with different time steps between each change. “Neutral” face was equal to 0, “largest distance-between-eyes”, “widest nose” or “widest chin” were equal to 1.25 and the smallest features – to -1.25.

Output signals – excitement, meditation, frustration and engagement/boredom – varied from 0 to 1. If excitement, meditation, frustration and engagement were low, the value was 0 or close to 0 and if they were high, the values of parameters were 1 or close to 1. The signals were recorded with the sampling period of $T_0=0.5$ s.

Five volunteers (volunteers no. 1-3 were female, and volunteers no. 4-5 were male) were tested. A volunteer was watching three animated scenes of approximately 1 minute one after another, and EEG-based signals were measured simultaneously using Emotiv Epoc device. The recorded preprocessed signals were stored for later analysis.

3. Correlation analysis

To estimate the possible relations between virtual stimuli and human reaction signals, a cross-correlation analysis between stimuli (3D virtual face features) and response (EEG-based reaction) signals was performed.

Cross-covariation functions between input and output signals and auto-covariation functions of input and output signals are used for this purpose [16]:

$$\begin{aligned}
 R_{yx}[\tau] &= \frac{1}{N} \sum_{t=1}^{N-\tau} (y_t - \bar{y})(x_{t+\tau} - \bar{x}) \\
 R_{xx}[\tau] &= \frac{1}{N} \sum_{t=1}^{N-\tau} (x_t - \bar{x})(x_{t+\tau} - \bar{x}) \\
 R_{yy}[\tau] &= \frac{1}{N} \sum_{t=1}^{N-\tau} (y_t - \bar{y})(y_{t+\tau} - \bar{y})
 \end{aligned} \quad (1)$$

where

$$\bar{x} = \frac{1}{N} \sum_{t=1}^N x_t, \quad \bar{y} = \frac{1}{N} \sum_{t=1}^N y_t \quad (2)$$

are the averages of input and output, $\tau = 0, \pm 1, \dots$ and $N=115$.

Cross-correlation functions between both types of input signals (TYPE 1 and TYPE 2) and all output signals for volunteer no. 1 are demonstrated in Figs. 4–6.

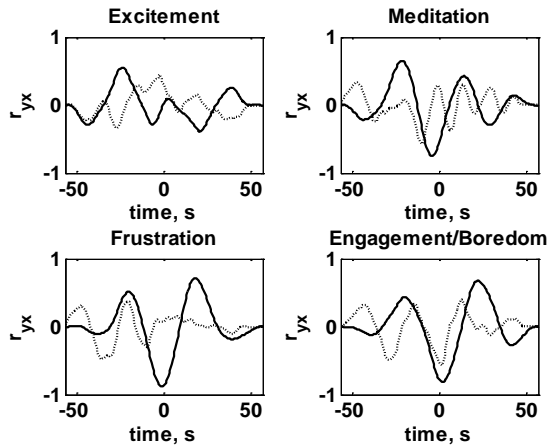


Figure 4. Cross-correlation functions between distance between eyes input feature and all output features for volunteer no.1. Solid line denotes input Type 1, dotted line denotes input Type 2

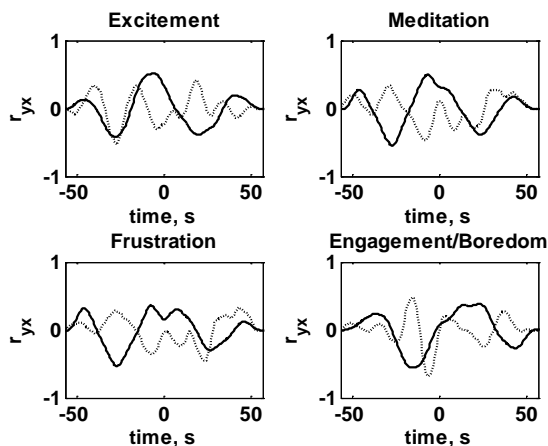


Figure 5. Cross-correlation functions between nose width input feature and all output features for volunteer no.1. Solid line denotes input Type 1, dotted line denotes input Type 2

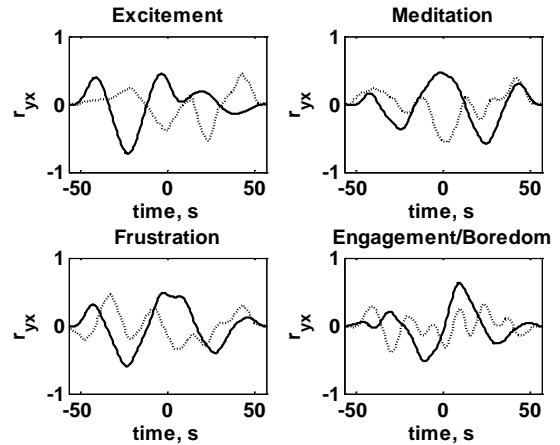


Figure 6. Cross-correlation functions between chin width input feature and all output features for volunteer no.1. Solid line denotes input Type 1, dotted line denotes input Type 2

Maximum cross-correlation function values

$$\max_{\tau} |r_{yx}[0]| = \max_{\tau} \left| \frac{R_{yx}[\tau]}{\sqrt{R_{yy}[0]R_{xx}[0]}} \right| \quad (3)$$

were estimated for each stimulus and the corresponding response signal pair. The estimates of the maximum cross-correlation function values are shown in Table 1 and Table 2.

Table 1. Maximum cross-correlation function values when Type 1 input signal is used

Volunteer no.	Input (stimulus)	Output (reaction)			
		Excitement	Meditation	Frustration	Engagement/ Boredom
1	Distance between eyes	0.55	0.75	0.88	0.81
	Nose width	0.51	0.55	0.53	0.55
	Chin width	0.73	0.57	0.59	0.63
2	Distance between eyes	0.47	0.30	0.44	0.38
	Nose width	0.73	0.51	0.47	0.66
	Chin width	0.35	0.75	0.43	0.37
3	Distance between eyes	0.54	0.52	0.58	0.46
	Nose width	0.30	0.67	0.80	0.62
	Chin width	0.38	0.74	0.43	0.38
4	Distance between eyes	0.76	0.69	0.64	0.44
	Nose width	0.47	0.65	0.59	0.66
	Chin width	0.40	0.64	0.43	0.55
5	Distance between eyes	0.30	0.57	0.54	0.69
	Nose width	0.71	0.68	0.75	0.82
	Chin width	0.70	0.66	0.54	0.62

The shift of the maximal values of cross-correlation functions in relation to $r_{yx}[0]$ allow stating that there exist dynamic relations between virtual

stimuli and human reactions to them. High cross-correlation values justify a possibility to use linear dynamic models.

Table 2. Maximum cross-correlation function values when Type 2 input signal is used

Volunteer no.	Input (stimulus)	Output (reaction)			
		Excitement	Meditation	Frustration	Engagement/ Boredom
1	Distance between eyes	0.45	0.57	0.49	0.56
	Nose width	0.54	0.47	0.45	0.68
	Chin width	0.53	0.56	0.47	0.38
2	Distance between eyes	0.46	0.54	0.56	0.43
	Nose width	0.49	0.41	0.43	0.61
	Chin width	0.63	0.56	0.57	0.38
3	Distance between eyes	0.48	0.45	0.33	0.63
	Nose width	0.47	0.49	0.47	0.68
	Chin width	0.55	0.40	0.42	0.57
4	Distance between eyes	0.61	0.59	0.43	0.44
	Nose width	0.50	0.44	0.46	0.41
	Chin width	0.53	0.63	0.64	0.54
5	Distance between eyes	0.36	0.37	0.45	0.60
	Nose width	0.43	0.50	0.45	0.52
	Chin width	0.46	0.56	0.36	0.48

4. Mathematical model building

All dependencies between virtual object stimuli and human reactions are described by input-output structure models [8]:

$$A(z^{-1})y_t = B(z^{-1})x_t + \varepsilon_t \quad (4)$$

$$B(z^{-1}) = \sum_{j=0}^m b_j z^{-j}, \quad A(z^{-1}) = 1 + \sum_{i=1}^n a_i z^{-i} \quad (5)$$

where y_t is an output (excitement, engagement/boredom meditation, frustration or) and x_t is an input (distance between eyes, nose width, chin width) signals respectively calculated as

$$y_t = y(tT_0), \quad x_t = x(tT_0) \quad (6)$$

with sampling period T_0 , ε_t corresponds to noise signal and z^{-1} is the backward-shift operator ($z^{-1}x_t = x_{t-1}$).

Eq. (4) can be expressed in the following form:

$$y_t = \sum_{j=0}^m b_j x_{t-j} - \sum_{i=1}^n a_i y_{t-i} + \varepsilon_t \quad (7)$$

Model parameters (coefficients of the polynomials (5)) and model order (degrees m and n of polynomials (5)) are unknown. They have to be estimated

according to the observations obtained during the experiments with the volunteers.

It is not difficult to show that in the case of using the model, the following relationship exists between covariation functions:

$$R_{yx}[\tau] = \sum_{j=0}^m b_j R_{xx}[\tau - j] - \sum_{i=1}^n a_i R_{yx}[\tau - i] \quad (8)$$

Equation (8) can be expressed as a linear regression equation:

$$R_{yx}[\tau] = \beta_\tau^T \mathbf{c} \quad (9)$$

where

$$\beta_\tau^T = [R_{xx}[\tau], R_{xx}[\tau - 1], \dots, R_{xx}[\tau - m], \\ -R_{yx}[\tau - 1], \dots, -R_{yx}[\tau - n]] \quad (10)$$

$$\mathbf{c}^T = [b_0, b_1, \dots, b_m, a_1, a_2, \dots, a_n] \quad (11)$$

and T is a vector transpose sign.

For the estimation of unknown parameter vector \mathbf{c} we use a method of least squares [8]:

$$\hat{\mathbf{c}} = \mathbf{Q}^{-1} \mathbf{q} \quad (12)$$

$$\mathbf{Q} = \sum_{\tau=-M}^M \beta_\tau \beta_\tau^T, \quad \mathbf{q} = \sum_{\tau=-M}^M R_{yx}[\tau] \beta_\tau \quad (13)$$

where M is the number of covariation function values used.

Model parameter estimates' vector $\hat{\mathbf{c}}$ is calculated when $M=0, 1, \dots$ and model stability condition is verified [8]. It means that the following polynomial

$$\hat{A}_M(z) = z^n \hat{A}_M(z^{-1}) = z^n + \sum_{i=1}^n \hat{a}_i z^{n-i} \quad (14)$$

roots

$$z_i^A: \hat{A}_M(z) = 0, i = 1, 2, \dots, n \quad (15)$$

have to be in the unit disk

$$|z_i^A| \leq 1 \quad (16)$$

In that way, we get a subset Ω_M with M values where the models are stable. From this subset we choose

$$\hat{M}: \min\{\sigma_\varepsilon[M|m, n], M \in \Omega_M\} \quad (17)$$

where

$$\sigma_\varepsilon[M|m, n] = \sqrt{\frac{1}{N} \sum_{t=1}^N \hat{\varepsilon}_t^2[M|m, n]} \quad (18)$$

is one step output prediction error standard deviation,

$$\hat{\varepsilon}_t[M|m, n] = y_t - \hat{y}_{t|t-1}[M|m, n] \quad (19)$$

is one step output prediction error,

$$\begin{aligned} \hat{y}_{t|t-1}[M|m, n] &= \\ &= z[1 - \hat{A}_M(z^{-1})]y_{t-1} + \hat{B}_M(z^{-1})x_t \end{aligned} \quad (20)$$

is one step output prediction [9] and z is the forward-shift operator ($zy_t = y_{t+1}$).

Figs. 7-8 demonstrate examples of prediction error standard deviations with every M value for each output signal (when the same input signal is used).

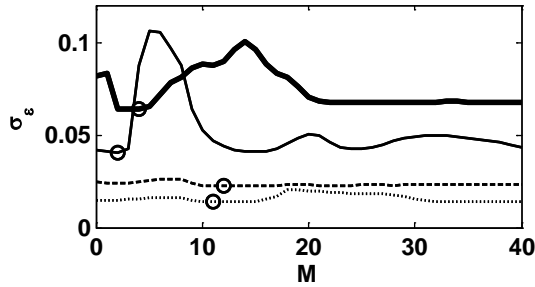


Figure 7. Prediction error standard deviation, nose width input (Type 2), volunteer no. 4, model order $m=1$, $n=1$.

Lines: solid thick– excitement, dotted– meditation, solid thin– frustration, dashed– engagement/boredom

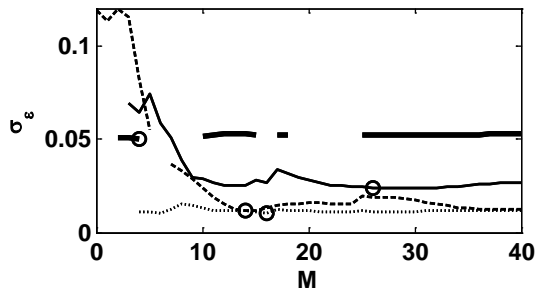


Figure 8. Prediction error standard deviation, distance-between-eyes input (Type 2), volunteer no. 1, model order $m=0$, $n=2$.

Lines: solid thick– excitement, dotted– meditation, solid thin– frustration, dashed– engagement/boredom

The circles indicate M values where the error standard deviations were the smallest in each output case. If models are not stable with certain M values, the error standard deviations are not estimated and there are discontinuities in the error standard deviations signal.

Estimates of the model orders – \hat{m} and \hat{n} – are defined from the following conditions:

$$\begin{aligned} \hat{n}: \quad & \left| \frac{\sigma_\varepsilon[\hat{M}|m, n+1] - \sigma_\varepsilon[\hat{M}|m, n]}{\sigma_\varepsilon[\hat{M}|m, n]} \right| \leq \delta, \\ & n = 1, 2, \dots \\ \hat{m}: \quad & \left| \frac{\sigma_\varepsilon[\hat{M}|m+1, n] - \sigma_\varepsilon[\hat{M}|m, n]}{\sigma_\varepsilon[\hat{M}|m, n]} \right| \leq \delta, \\ & m = 0, 1, \dots, n \end{aligned} \quad (21)$$

where $\delta > 0$ is a chosen constant value. Usually in the practice of identification $\delta \in [0,001 \div 0,01]$ what

corresponds to a relative variation of prediction error standard deviation from 0,1% to 1%.

This way a stable input-output model is built that ensures the best one step output signal prediction.

Figs. 9-10 demonstrate prediction error standard deviations for an input-output pair when $n=1, 2$ and $m=0, 1, 2$ for each model. Fig. 9. shows error standard deviation when distance-between-eyes input (Type 1) and engagement/boredom output is used for prediction for every of five models. Fig. 10. shows the same with chin width input (Type 2) and excitement output signal.

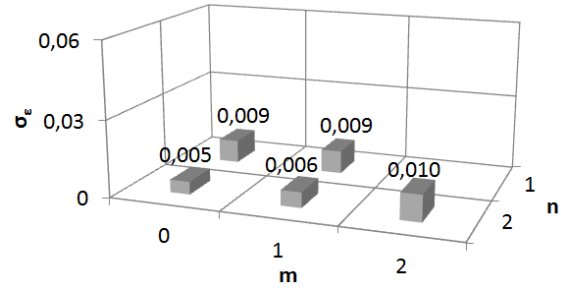


Figure 9. Prediction error standard deviation, volunteer no. 2, distance-between-eyes input signal (Type 1), engagement/boredom output signal

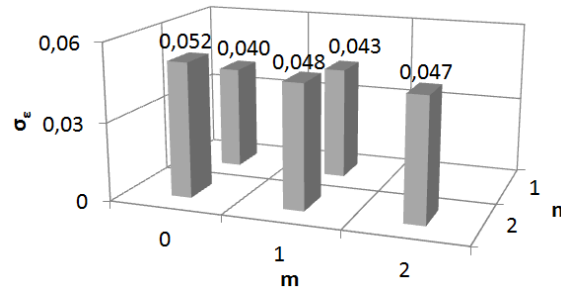


Figure 10. Prediction error standard deviation, volunteer no. 3, chin width input signal (Type 2), excitement output signal

5. Models validation

Every stimuli type requires 12 input-output models for each volunteer. Each model is selected from five possible models (when $n=1, 2$; $m=0, \dots, n$) using the rule (21). Two types of stimuli for each of five volunteers were used, so 120 models were analyzed. The analysis of five volunteers' data showed that relations between each of three stimuli (distance between eyes, nose width, and chin width) and excitement output signal can be modeled when model order is $\hat{m} = 0$ and $\hat{n} = 1$. Relations between each of three stimuli and meditation, frustration and engagement/boredom output signals are best modeled when model order is $\hat{m} = 0$ and $\hat{n} = 2$.

Every model was validated through the output signal (excitement, meditation, engagement/boredom, frustration) prediction analysis. The predicted output signals have the following expression [8]:

$$\hat{y}_{t+1|t} = z[1 - \hat{A}(z^{-1})]y_t + \hat{B}(z^{-1})x_{t+1} = -\sum_{i=1}^{\hat{n}} \hat{a}_i y_{t+1-i} + \sum_{j=0}^{\hat{m}} \hat{b}_j x_{t+1-j} \quad (22)$$

Prediction accuracies were evaluated using the following measures:

- prediction error standard deviation

$$\sigma_\varepsilon = \sqrt{\frac{1}{N} \sum_{t=0}^{N-1} (y_{t+1} - \hat{y}_{t+1|t})^2} \quad (23)$$

- relative prediction error standard deviation

$$\tilde{\sigma}_\varepsilon = \sqrt{\frac{1}{N} \sum_{t=0}^{N-1} \left(\frac{y_{t+1} - \hat{y}_{t+1|t}}{y_{t+1}}\right)^2} * 100\% \quad (24)$$

- and average absolute relative prediction error

$$|\bar{\varepsilon}| = \frac{1}{N} \sum_{t=0}^{N-1} \left|\frac{y_{t+1} - \hat{y}_{t+1|t}}{y_{t+1}}\right| * 100\% \quad (25)$$

Measures of prediction accuracy (23) - (25) are given in Tables 3-7.

Table 3. Prediction accuracy measures, volunteer no. 1

Input stimulus	Output reaction	Input TYPE 1			Input TYPE 2		
		σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$	σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$
Distance between eyes	Excitement	0.044	11.5	7.79	0.054	13.8	8.6
	Meditation	0.006	1.9	1.5	0.010	3.0	2.6
	Frustration	0.011	2.1	1.4	0.023	4.0	2.8
	Eng/Bor	0.010	1.6	1.5	0.010	1.7	1.7
Nose width	Excitement	0.038	16.0	10.5	0.031	16.1	9.4
	Meditation	0.005	2.0	1.4	0.007	2.2	1.7
	Frustration	0.017	4.9	3.3	0.019	7.4	5.0
	Eng/Bor	0.003	0.6	0.5	0.013	2.4	2.2
Chin width	Excitement	0.026	10.9	7.0	0.017	8.2	5.2
	Meditation	0.007	2.7	1.7	0.005	2.6	1.7
	Frustration	0.017	4.1	2.9	0.016	4.8	3.6
	Eng/Bor	0.004	0.8	0.5	0.014	2.9	2.8

Table 4. Prediction accuracy measures, volunteer no. 2

Input stimulus	Output reaction	Input TYPE 1			Input TYPE 2		
		σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$	σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$
Distance between eyes	Excitement	0.057	10.8	8.3	0.041	20.3	15.4
	Meditation	0.037	1.1	1.0	0.005	1.4	1.1
	Frustration	0.016	3.4	2.5	0.023	7.4	5.2
	Eng/Bor	0.005	0.7	0.5	0.008	1.4	1.1
Nose width	Excitement	0.062	18.3	12.3	0.055	12.2	8.0
	Meditation	0.002	0.8	0.7	0.008	2.5	2.4
	Frustration	0.016	4.7	3.4	0.024	3.8	2.6
	Eng/Bor	0.005	0.8	0.6	0.007	1.8	1.2
Chin width	Excitement	0.040	18.7	11.8	0.077	20.6	14.4
	Meditation	0.002	0.8	0.6	0.005	1.6	1.2
	Frustration	0.015	4.5	3.3	0.020	3.9	2.9
	Eng/Bor	0.006	1.0	0.8	0.006	1.3	0.9

Table 5. Prediction accuracy measures, volunteer no. 3

Input stimulus	Output reaction	Input TYPE 1			Input TYPE 2		
		σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$	σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$
Distance between eyes	Excitement	0.027	10.6	8.7	0.025	9.2	7.4
	Meditation	0.007	4.1	2.9	0.010	5.2	4.2
	Frustration	0.009	2.0	1.4	0.014	3.3	2.8
	Eng/Bor	0.013	1.5	1.2	0.015	1.7	1.0
Nose width	Excitement	0.026	9.1	7.3	0.033	11.0	8.2
	Meditation	0.006	2.8	1.7	0.006	3.2	2.2
	Frustration	0.010	1.9	1.5	0.021	3.3	2.4
	Eng/Bor	0.010	1.2	0.9	0.011	1.5	1.2
Chin width	Excitement	0.040	18.7	11.8	0.040	11.8	9.7
	Meditation	0.002	0.8	0.6	0.006	3.3	2.3
	Frustration	0.015	4.5	3.3	0.025	4.4	3.6
	Eng/Bor	0.006	1.0	0.8	0.010	1.4	1.0

Table 6. Prediction accuracy measures, volunteer no. 4

Input stimulus	Output reaction	Input TYPE 1			Input TYPE 2		
		σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$	σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$
Distance between eyes	Excitement	0.052	11.5	8.7	0.060	17.1	12.7
	Meditation	0.002	0.7	0.5	0.011	3.5	3.3
	Frustration	0.021	4.2	2.8	0.026	7.3	4.8
	Eng/Bor	0.008	1.2	0.9	0.011	2.0	1.3
Nose width	Excitement	0.057	17.0	12.4	0.064	12.4	10.1
	Meditation	0.003	0.9	0.7	0.006	1.6	0.9
	Frustration	0.022	5.8	4.1	0.033	5.7	3.8
	Eng/Bor	0.010	1.7	1.1	0.012	2.9	1.8
Chin width	Excitement	0.066	13.4	10.6	0.063	16.6	11.6
	Meditation	0.004	1.2	0.9	0.005	1.8	1.1
	Frustration	0.021	5.7	3.5	0.033	7.9	5.6
	Eng/Bor	0.008	1.4	1.0	0.010	2.0	1.6

Table 7. Prediction accuracy measures, volunteer no. 5

Input stimulus	Output reaction	Input TYPE 1			Input TYPE 2		
		σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$	σ_ε	$\tilde{\sigma}_\varepsilon, \%$	$ \bar{\varepsilon} , \%$
Distance between eyes	Excitement	0.053	16.9	12.4	0.068	10.2	7.9
	Meditation	0.002	0.9	0.6	0.005	1.4	1.2
	Frustration	0.015	5.2	3.4	0.012	3.1	2.2
	Eng/Bor	0.010	2.8	1.6	0.012	2.7	2.3
Nose width	Excitement	0.062	15.6	10.2	0.048	14.6	9.9
	Meditation	0.003	1.0	0.8	0.006	1.9	1.6
	Frustration	0.015	4.4	2.7	0.014	3.0	2.2
	Eng/Bor	0.012	3.0	2.0	0.014	3.6	2.6
Chin width	Excitement	0.059	18.3	10.8	0.054	18.8	13.5
	Meditation	0.005	1.6	1.2	0.005	1.5	1.0
	Frustration	0.009	2.7	1.9	0.018	4.1	2.7
	Eng/Bor	0.008	1.5	1.1	0.012	3.0	2.1

Figs. 11-22 demonstrate observed output signals, their predictions and prediction errors.

Figs. 11-16 demonstrate observed output signals, their predictions and prediction errors of female volunteer when Type 1 and Type 2 inputs are used. Figs. 17-22 demonstrate the same signals of a male volunteer when Type 1 and Type 2 inputs are used.

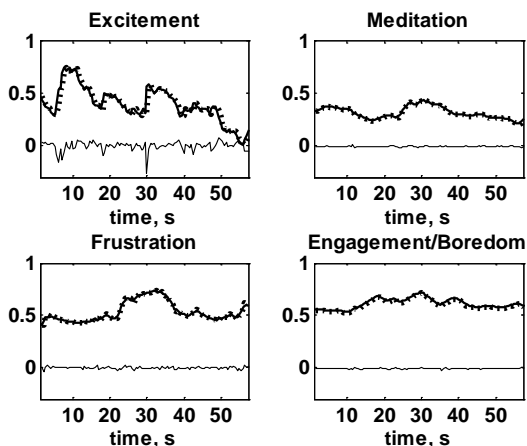


Figure 11. Volunteer no. 1 (female), distance between eyes TYPE 1 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

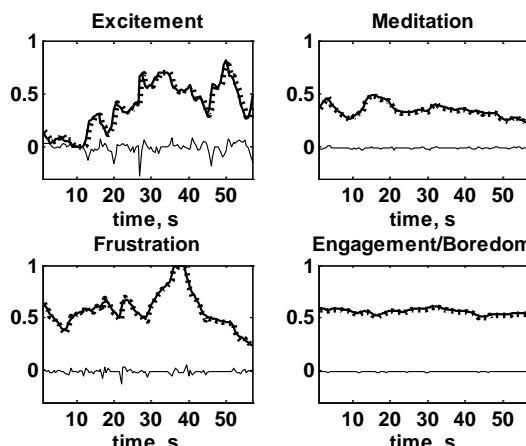


Figure 14. Volunteer no. 1 (female), distance between eyes TYPE 2 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

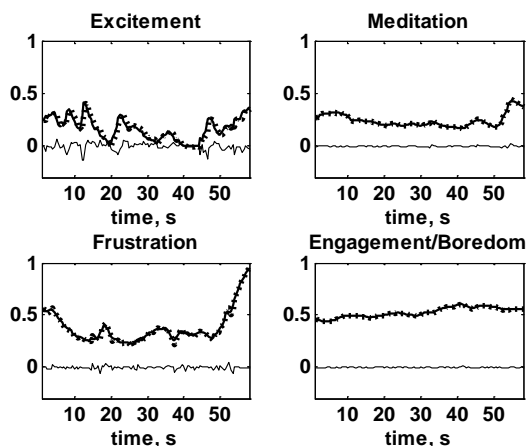


Figure 12. Volunteer no. 1 (female), nose width TYPE 1 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

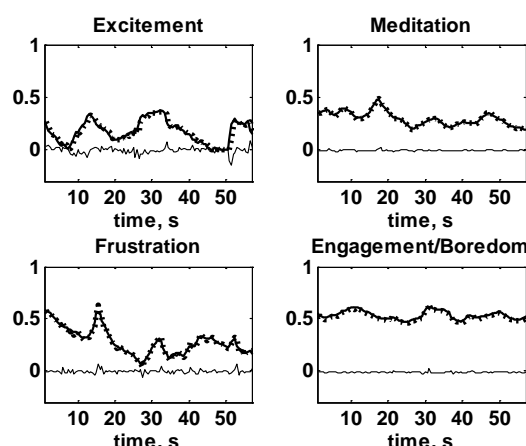


Figure 15. Volunteer no. 1 (female), nose width TYPE 2 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

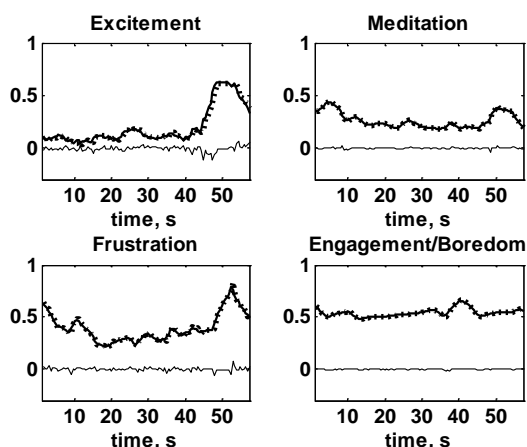


Figure 13. Volunteer no. 1 (female), chin width TYPE 1 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

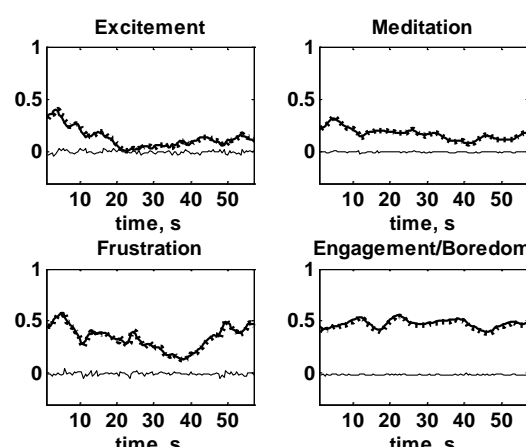


Figure 16. Volunteer no. 1 (female), chin width TYPE 2 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

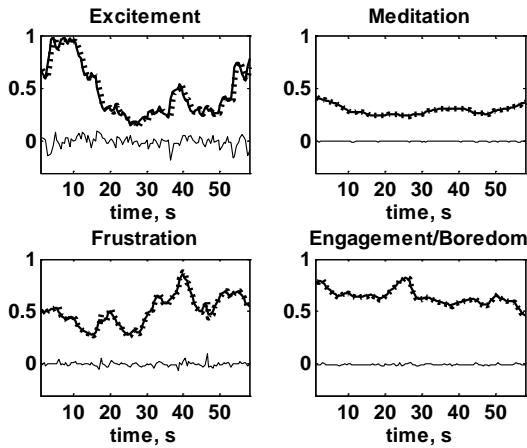


Figure 17. Volunteer no. 4 (male), distance between eyes TYPE 1 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

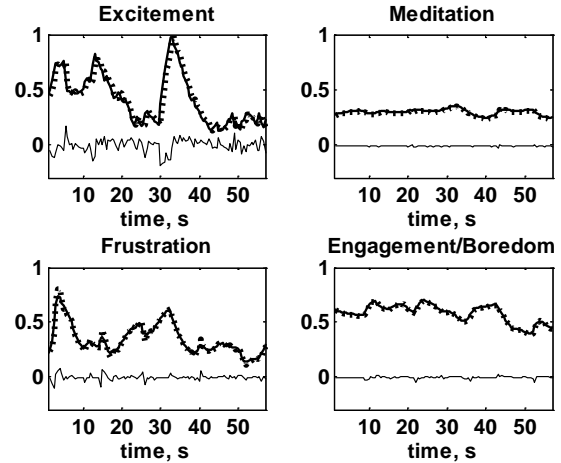


Figure 20. Volunteer no. 4 (male), distance between eyes TYPE 2 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

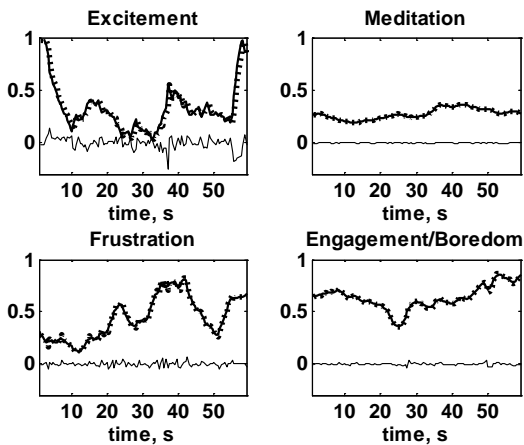


Figure 18. Volunteer no. 4, (male) nose width TYPE 1 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

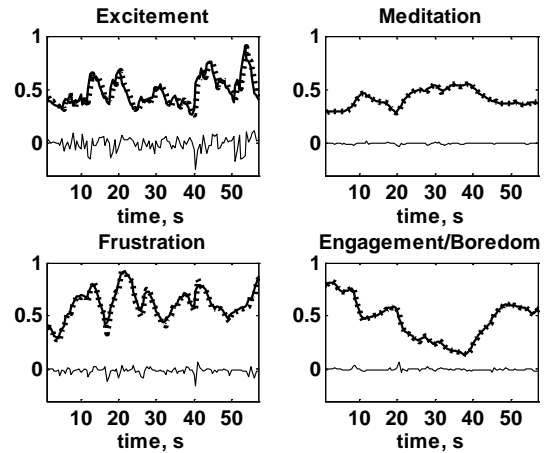


Figure 21. Volunteer no. 4 (male), nose width TYPE 2 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

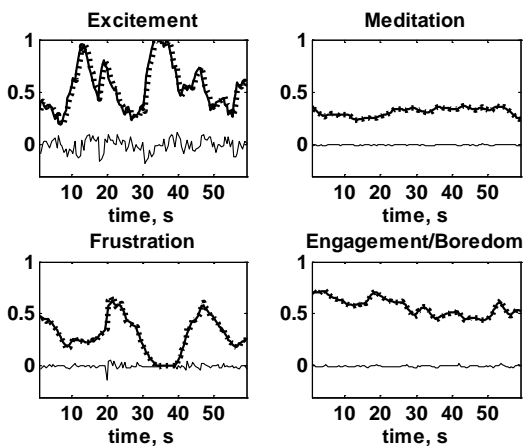


Figure 19. Volunteer no. 4 (male), chin width TYPE 1 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

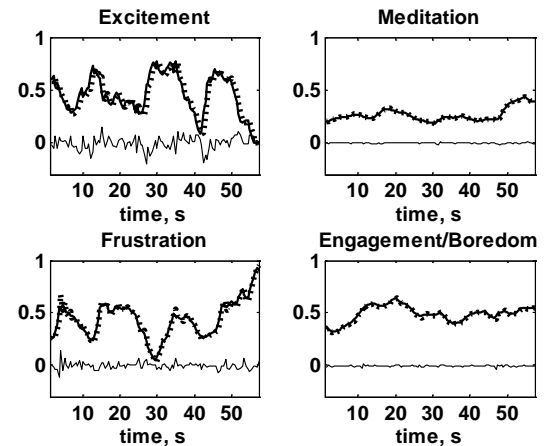


Figure 22. Volunteer no. 4 (male), chin width TYPE 2 input. Solid thick line – observed signal, dotted line – predicted signal, solid thin line – prediction error between observed and predicted signal

Results of model validation demonstrate that average absolute relative prediction errors of meditation and engagement/boredom signals in all five volunteer cases were less than 1.5 %, frustration – less than 3.5 %, and excitement – less than 10 %. The corresponding relative error standard deviation averages are less than 2 % for meditation and engagement/boredom, less than 4.5 % for frustration and less than 15 % for excitement output signals.

Predictive models are necessary in the design of predictor-based self-tuning control systems [2], [9]. High prediction accuracies of the built models allow expecting that they can be a basis for this type control systems of human reactions to virtual stimuli. Such systems could serve as a mean for regulation of human attention level in learning applications as well as in computerized workplaces for dispatchers.

6. Conclusions

Cross-correlation analysis demonstrated that there is a relatively high correlation between stimuli (input) and human response (output) and that evidences the existence of linear dynamic relations between them.

When using cross- and auto-correlation functions, a new method for building an input-output model was proposed. It allows building a stable model for output signal prediction with the least prediction error. The models allow describing the dynamic dependencies between virtual 3D face that changes during time and the reactions to it.

Validation results of the models demonstrate that every volunteer reacted to the stimuli individually and the reactions are described by input-output models with different estimates of polynomial coefficient.

The models describe the dependencies between every input signal (distance between eyes, nose width and chin width) and every corresponding output signal (excitement, meditation, engagement/boredom, and frustration) in high accuracy: averages of absolute relative prediction errors of engagement/boredom and meditation signals in five volunteer cases were less than 1.5 %, frustration – less than 3.5 %, and excitement – less than 10 %.

Acknowledgments

Postdoctoral fellowship of Ausra Vidugiriene is funded by European Union Structural Funds project "Postdoctoral Fellowship Implementation in Lithuania" within the framework of the Measure for Enhancing Mobility of Scholars and Other Researchers and the Promotion of Student Research (VP1-3.1-ŠMM-01) of the Program of Human Resources Development Action Plan.

References

[1] **R. Ciceri, S. Balzarotti.** From Signals to Emotions: Applying Emotion Models to HM Affective

Interactions. In: *Affective Computing: Emotion Modeling, Synthesis and Recognition*, ITech, Education and Publishing, 2008, pp. 271-296. Available at: <http://dx.doi.org/10.5772/6176>.

[2] **D. W. Clarke.** Advances in Model Predictive Control. *Oxford Science Publications*, UK, 1994.

[3] **Emotiv Epoc specifications.** Brain-computer interface technology. Available at: <http://www.emotiv.com/upload/manual/sdk/EPOCSpecifications.pdf>.

[4] **S. Fatahi, N. Ghasem-Aghae.** An Effective Intelligent Educational Model Using Agent with Personality and Emotional Filters. In: *Proceedings of the World Congress on Engineering*, Vol. 1, 2010, pp. 142-147. Available at: http://www.iaeng.org/publication/WCE2010/WCE2010_pp142-147.pdf.

[5] **J. Hoey, T. Schroeder, A. Althothali.** Bayesian Affect Control Theory. In: *IEEE 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Geneva, Switzerland, 2013, pp. 166-172. Available at: <http://dx.doi.org/10.1109/ACII.2013.34>.

[6] **C. Hondrou, G. Caridakis.** Affective, Natural Interaction Using EEG: Sensors, Application and Future Directions. In: *Artificial Intelligence: Theories and Applications*. LNCS Vol. 7297, Springer-Verlag Berlin Heidelberg, 2012, pp. 331-338. Available at: http://dx.doi.org/10.1007/978-3-642-30448-4_42.

[7] **M. E. Hoque, M. Courgeon, J.-C. Martin, B. Mutlu, R. W. Picard.** MACH: My Automated Conversation coach. In: *15th International Conference on Ubiquitous Computing (Ubicomp)*, Zurich, Switzerland, September, 2013, pp. 697-706. Available at: <http://dx.doi.org/10.1145/2493432.2493502>.

[8] **V. Kaminskas.** Dynamic Systems Identification via Discrete-Time Observations, *Part I – Statistical Method Foundations. Estimation in Linear Systems*, Vilnius: Mokslas, 1982, (in Russian).

[9] **V. Kaminskas.** Predictor-based self tuning control with constraints. *Model and Algorithms for Global Optimization, Optimization and Its Applications*, Vol. 4, Springer, 2007, pp. 333-341. Available at: http://dx.doi.org/10.1007/978-0-387-36721-7_20.

[10] **R. N. Khushaba, L. Greenacre, S. Kodagoda, J. Louviere, S. Burke, G. Dissanayake.** Choice modeling and the brain: A study on the Electroencephalogram (EEG) of preferences. *Expert Systems with Applications*, Vol. 39, Is. 16, 2012, pp. 12378–12388. Available at: <http://dx.doi.org/10.1016/j.eswa.2012.04.084>.

[11] **R. N. Khushaba, Ch. Wise, S. Kodagoda, J. Louviere, B. E. Kahn, C. Townsend.** Consumer neuroscience: Assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking. *Expert Systems with Applications*, Vol. 40, Is. 9, pp. 3803-3812. Available at <http://dx.doi.org/10.1016/j.eswa.2012.12.095>.

[12] **Y. Liu, O. Sourina, M. Rizqi Hafiyandi.** EEG-based Emotion-adaptive Advertising. In: *IEEE 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Geneva, Switzerland, 2013, pp. 843-848. Available at: <http://dx.doi.org/10.1109/ACII.2013.158>.

[13] **A. Mehrabian.** Silent messages: Implicit communication of emotions and attitudes. Belmont, CA: *Wadsworth Publishing Company*, 1981.

[14] **J. Renny Octavia, K. Luyten, J. Vermeulen, B. Mommen, K. Coninx.** Exploring Psycho-

- physiological Measures for the Design and Behavior of Socially-Aware Avatars in Ubicomp Environments. In: *CHI 2010*, April, Atlanta, GA, USA pp. 1-4.
- [15] **A. Rizzo, J. G. Buckwalter, B. John, B. Newman, T. Parsons, P. Kenny, J. Williams.** STRIVE: Stress Resilience In Virtual Environments: A Pre-Deployment VR System for Training Emotional Coping Skills and Assessing Chronic and Acute Stress Responses. In: *Medicine Meets Virtual Reality 19*, IOS Press, 2012, pp. 379-385. Available at: <http://dx.doi.org/10.3233/978-1-61499-022-2-379>.
- [16] **J. G. Proakis, D. G. Manolakis.** Digital signal Processing: Principles, Algorithms, and Application, 4th edition, *Pearson Prentice Hall*, 2007.
- [17] **A. Sano, R. W. Picard.** Stress recognition using wearable sensors and mobile phones. In: *IEEE 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, Geneva, Switzerland, 2013, pp. 671-676. Available at: <http://dx.doi.org/10.1109/ACII.2013.117>.
- [18] **O. Sourina, Y. Liu.** A Fractal-based Algorithm of Emotion Recognition from EEG using Arousal-valence model. In: *Proc. Biosignals*, Rome, 2011, pp. 209-214.
- [19] **Suprijanto, L. Sari, V. Nadhira, IG.N. Merthayasa, I. M. Farida.** Development System for Emotion Detection Based on Brain Signals and Facial Images. In: *World Academy of Science, Engineering and Technology* 26, 2009, pp. 931-938. Available at: <http://waset.org/publications/8769>.
- [20] **E. Vaskevičius, A. Vidugiriene, V. Kaminskas.** Investigation of dependencies between virtual 3d face stimuli and emotion-based human responses. In: *Proc. of the 8th International Conference on Electrical and Control Technologies*, Kaunas, 2013, pp. 64-69.
- [21] **J. Willis, A. Todorov.** First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. In: *Psychological science*, Vol. 17, No. 7, 2006, pp. 592–598. Available at: <http://dx.doi.org/10.1111/j.1467-9280.2006.01750.x>.
- [22] **Y. Zhang, J. Tang, J. Sun, Y. Chen, J. Rao.** MoodCast: Emotion Prediction via Dynamic Continuous Factor Graph Model. In: *IEEE 10th International Conference on Data Mining*, Sydney, 2010, pp. 1193-1198. Available at: <http://dx.doi.org/10.1109/ICDM.2010.105>.
- [23] **M. Zisook, J. Hernandez, M. S. Goodwin, R. W. Picard.** Enabling Visual Exploration of Long-term Physiological Data. In: *Proceedings of the 2013 IEEE Conference on Visual Analytics Science and Technology*, Atlanta, Georgia, USA, October 2013.

Received December 2013.