Facial expressions animation based on motion trajectories modelling

Maja Kocoń, Zbigniew Emirsajłow
West Pomeranian University of Technology

Abstract: This paper describes a way of the face characteristic points trajectory synthesis during emotion changes. The points were selected in according to human face anatomy properties and based on an available system of face movements description. The motion curve was proposed to model changes of emotions on a three-dimensional geometrical model of the human face.

Keywords: facial expression, feature points, three-dimensional model, facial animation, human-machine interactions (HMI)

1. Introduction

In the recent years the research of analysis or synthesis of emotions has become increasingly popular. Emotions play a significant role because human face is a natural medium of communication. Most of human emotions are expressed through the face and can provide intuitive communication between user and machine (HMI-Human Machine Interaction) [13]. Developments in the field of emotions synthesis are used in many applications of the HMI [16], where one of the main goals is to create believable embodied agents. Its main concern is creating a virtual representation of the human which is able to interact with the users. Emotion modelling may be used in different applications like user-friendly interfaces [18], [19], portable personal guides [15], edutainment systems [17] and social robots (human-robot interaction) [14].

For facial expressions modelling on three-dimensional face model a number of methods exist. The most popular are image-based approaches (morphing techniques) [23] or techniques that utilize captured data of face motion [24]. In paper [22] F. Erol proposed system for face animation that uses parametric interpolation techniques for modelling individualized emotions. For this purpose front and side views of the human faces with located markers were captured. Positions of markers were extracted manually and then used to deform the face model. Still popular is a key-frame method called morphing between fixed and variable elements of mesh. In [25] a group of six basic facial expressions is generated based on facial animation parameters (FAPS). The transitions between expressions are produced by interpolation between different emotional states using 2D morphing techniques and transformations of geometric model. This system can be used in applications associated with avatars or emotions recognition. Garre et al [26] have adopted a mass-spring system to implement animation of deformable objects. This method gives more realism than i.e. Free Form Deformation. In [27] motion sequences captured from video are processed by finding the optimal path between user specified key-frames for control of movement, location and timing.

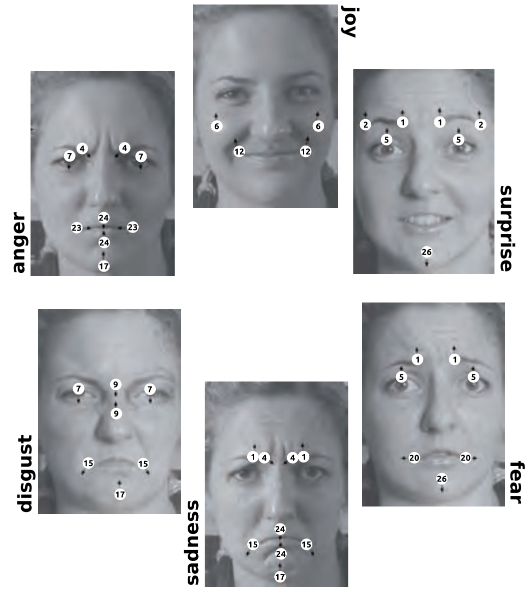
There are many approaches to the problem of animation. In this paper method of obtaining a movement model of individual landmarks, by using three-dimensional model of the face assigned with the bones model as the objects deforming the mesh object was proposed. Each bone deforms the face mesh by using vertex weights, ensuring smooth transformation and movement of points.

The paper is organised as follows. Next section is describing the proposed approach. In the section 3 experimental results and their evaluation are presented. Finally, we complete the paper with conclusions.

2. Proposed approach

The primary aim of this paper is modelling emotional states of the human face. The modelling of the object changes is done on a 3D model of the face, by using a system of the bones and a proposed trajectory function of motion. We have designed a deformable three–dimensional face model accordingly to the face anatomy [8], [2]. The knowledge of the face structure and the relationship between muscles is important because the main goal was to build a face model suitable for animation. Therefore we had to take movements aspects into account. From the movement analysis point of view, the essential step is the selection of characteristic points of the face. On the basis of the face actions classification (FACS) [3] and anatomical aspects of the face [4, 5, 9] a group of 21 landmarks was selected. Points were marked on the faces of volunteers, captured in the form of video sequences and then values of points movements were obtained from video material. Location of every point was recorded in the neutral state and in every of six chosen emotions (joy, fear, disgust, anger, surprise, sadness). Recordings were performed under the same conditions, the illumination and distance between camera and captured face were the same in each case.

Comprehensive emotions classification system called FACS describes how muscle actions are related to the facial appearance [1]. In this system each muscle movement is described as minimal units called facial Action Units (AU's) [7]. Through a combination of individual units it is possible to create almost any facial expression. On the basis of this system, a set of facial characteristic points and 14 units of the movements were selected, which were used to generate animation in this work. In most cases 14 units



AU No.	Descriptor
1	Inner Brow Raiser
2	Outer Brow Raiser
4	Brow Lowerer
5	Upper Lid Raiser
6	Cheek Raiser
7	Lid Tightner
9	Nose Wrinkler
12	Lip Corner Puller
15	Lip Corner Depressor
17	Chin Raiser
20	Lip Stretcher
23	Lip Tightner
24	Lip Pressor
25	Lips Part
26	Jaw Drop

Rys. 1. Ekspresje 6 podstawowych emocji według systemu FACS oraz wybrane jednostki ruchu
Fig. 1. Expression of 6 basic emotions according to FACS movement rules and selected Action Units

are enough to create 6 basic emotional states [3]. Figure 1 shows a number of Action Units which were selected for data acquisition and animation. These units were chosen in according to [3] and based on analysis presented in [11] and [12].

For the purpose of mapping movement trajectories on the 3D model, bone-based deformation method was chosen. The bone model has been assigned to the 3D face model using characteristic points. Therefore manipulating the face geometry to change the appearance is possible. The number of the bones corresponds to the number of characteristic points - hence the 21 bones in our study were created. The bones are rigid non-deformable objects [10] and they are assigned to characteristic points of the model. Proposed bones system is depicted in figure 2. An additional related issue is bone modification. The area closest to the bone is the one most strongly modified, while the areas located further are less affected. The strength of the influence of the bone is fixed using weights and based on the methods described in [20] and [21].

Having mapped points and the system of the bones on 3D model each bone has been translated in according to the values recorded from the video sequences. The trajectory of the movement of individual bones has been altered in order to get the change in the appearance of the face. In animation one of the Hermit polynomials [6] has been used. There is a limitation that one of three bone coordinates will reach its maximum value quicker than others. Therefore, in order to obtain desired movement shape, additional phases denoted by parameters t_d, t_s have been added (figure 3). The factor t_d is initial delay, t_s is the saturation time at desired level, t_c denotes movement time and u is normalised difference of bone positions in neutral and target state.

In case when all parameters of trajectories are defined, the complete animation is then performed by translation of bones positions as expressed by equation 1.

$$\begin{bmatrix} x_i'(t) \\ y_i'(t) \\ z_i'(t) \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} + \begin{bmatrix} \alpha \cdot \hat{q}\left(t, t_d^{(i_x)}, t_s^{(i_x)}, u^{(i_x)}\right) \\ \alpha \cdot \hat{q}\left(t, t_d^{(i_y)}, t_s^{(i_y)}, u^{(i_y)}\right) \\ \alpha \cdot \hat{q}\left(t, t_d^{(i_z)}, t_s^{(i_z)}, u^{(i_z)}\right) \end{bmatrix}. \quad (1)$$

The new bone position $[x_i'(t), y_i'(t), z_i'(t)]^T$ is determined by initial value of bone $[x_i, y_i, z_i]^T$ and translation function $\hat{q}$ depicted in the equation 2 (i_x, i_y, i_z are i-th indices for each coordinate).

$$\hat{q}(t, t_d, t_s, u) = \begin{cases} 0 & \text{if} \quad t_d \geq t > 0 \\ u \cdot p(t, t_d, t_s) & \text{if} \quad t_s > t > t_d \\ u & \text{if} \quad 1 > t \geq t_s \end{cases}, \quad (2)$$

where:

$$p(t, t_d, t_s) = -2 \cdot \left(\frac{t - t_d}{1 - t_s - t_d}\right)^3 + 3 \cdot \left(\frac{t - t_d}{1 - t_s - t_d}\right)^2,$$

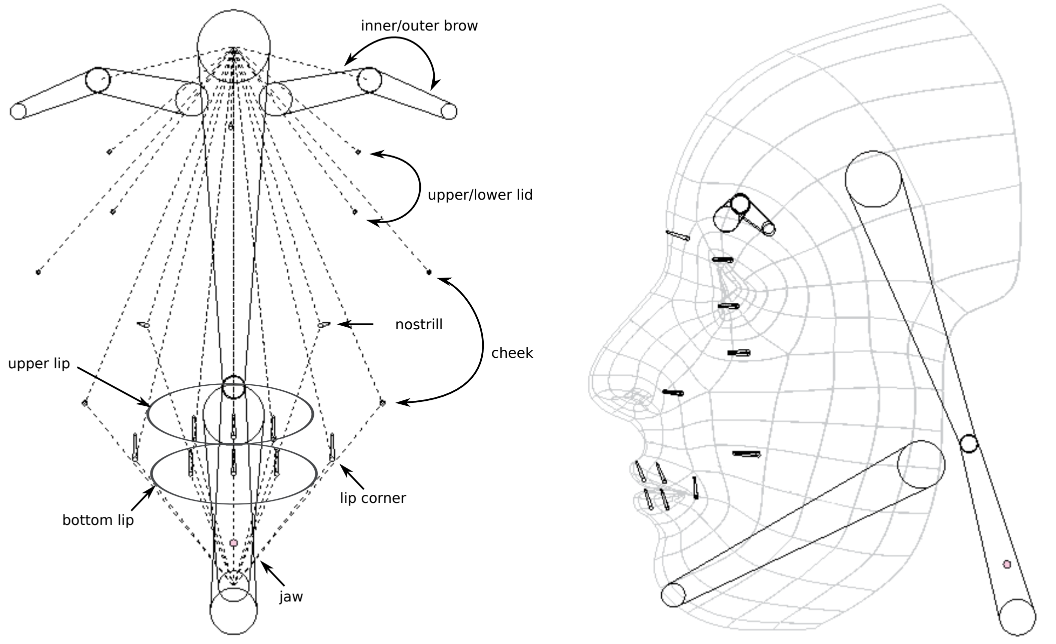
$t_d + t_c + t_s = 1$, α is the scale factor selected depending on movement strength, i denotes the bone index, $i = 1, \ldots, N$ and N is the total number of bones used in model.

In figure 4 the examples of movement trajectories obtained by modification of the values t_d, t_s (for three bones in joy state) are shown and emotions modeled by the change this values are depicted in figure 8.

3. Experimental evaluation

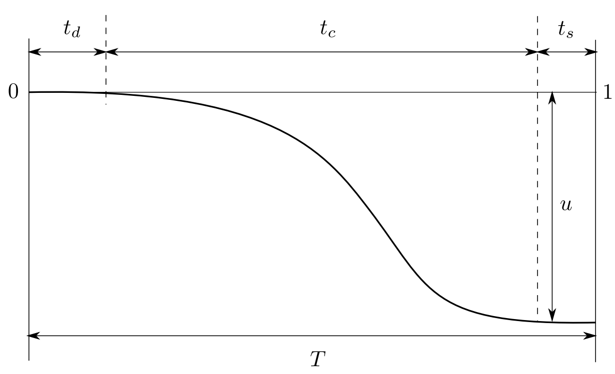
The main medium for transferring information about emotions are the regions of the eyes and mouth, therefore as it can be seen in figure 2, distribution of the bones is concentrated in these regions.

Obtained directions and values of bone movements as well as selected t_d, t_s can be used to determine the groups of bones with similar properties. This approach is facilitating the translation of a group of points, because we can move a number of points with one function. Creating groups allows



(d) (e)

Rys. 2. Proponowany system kości z oznaczeniami (a), widok z boku systemu kości przypisanego do modelu twarzy (b)
Fig. 2. The proposed bones' system with bones' names (a), bones assigned to the face model in the side view (b)



Rys. 3. Proponowana funkcja trajektorii ruchu z opóźnieniem początkowym (t_d) oraz czasem nasycenia (t_s)
Fig. 3. Proposed trajectory function with delay (t_d) and saturation (t_s) intervals

us to simplify the controlled model, which leads to faster calculations. Taking into account bone system described in previous section and excluding the aspect of symmetry (left and right bones are treated as different areas) of the face it's possible to create 8 groups (figure 5).

Knowing the values of u, we can determine which point has the highest activity. The most active points for every emotion are located in mouth area, in the lips corners to be exact. For most points, the coordinate of depth (in our case Y) has very small value and reaches its target value before others. As an example in figure 6 bones' coordinates assigned for the mouth are shown.

For five emotions a symmetry in the movement has been observed, however for one emotion called 'disgust' evident asymmetry has appeared. The right side of the face has a larger values of t_s, so has saturated faster than the left side. For comparison in figures 6 and 7 trajectories of the movement for mouth and left eye are shown. As can be seen they are determined at different times, despite the apparent symmetry. For the remaining emotions the saturation time was similar, but keep in mind that the movement largely depends on the test items, i.e. biological conditions of the face and emotions intensity.

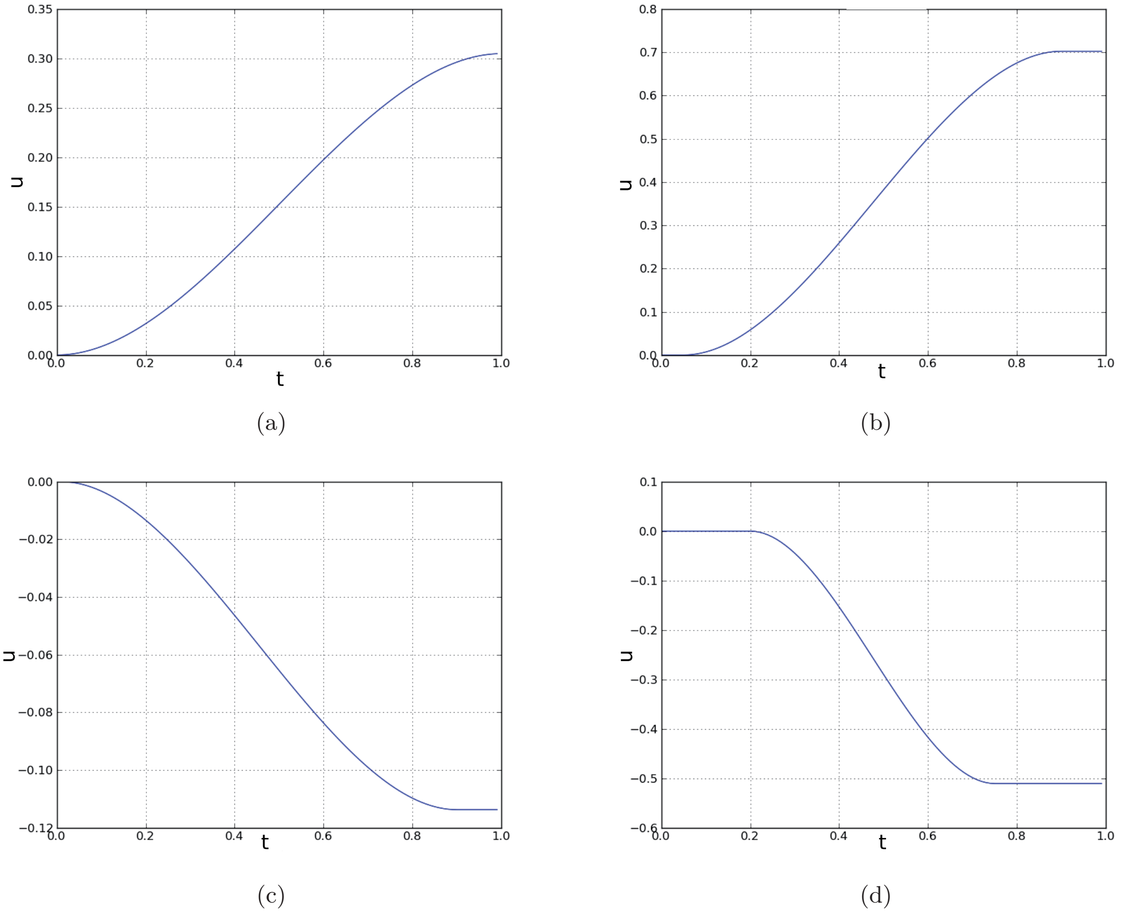
Comparing the t_d values, it is possible to notice which part of the face initiates emotions. The area of eyes is responsible for fear, surprise, disgust emotions and joy, sadness, anger corresponds to the mouth area. This information can also be useful in case of initial forming of the bone groups.

4. Conclusion

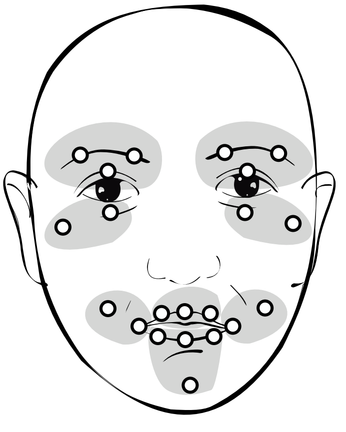
In this paper an attempt to obtain the trajectory of individual characteristic point of the face was made and in the result animation of emotional states' changes were modelled. At the stage of modification, non-deformable bone objects and functions of movement changes were used. Method of generating the trajectories of the movement, as well as a proper selection of the t_d, t_s movement values can be developed later on. Currently we are working on markerless trajectories using natural facial features only. As mentioned earlier, each face has different properties and each person expresses emotions in different way. In this work only one emotion synthesis animation approach was presented. In future works there is a plan to increase a set of analysed video sequences with larger number of different faces and emotional states.

References

1. Takeo Kanade, Yingli Tian and Jeffrey F. Cohn: *Comprehensive Database for Facial Expression Analysis.* IEEE Computer Society, 2000.



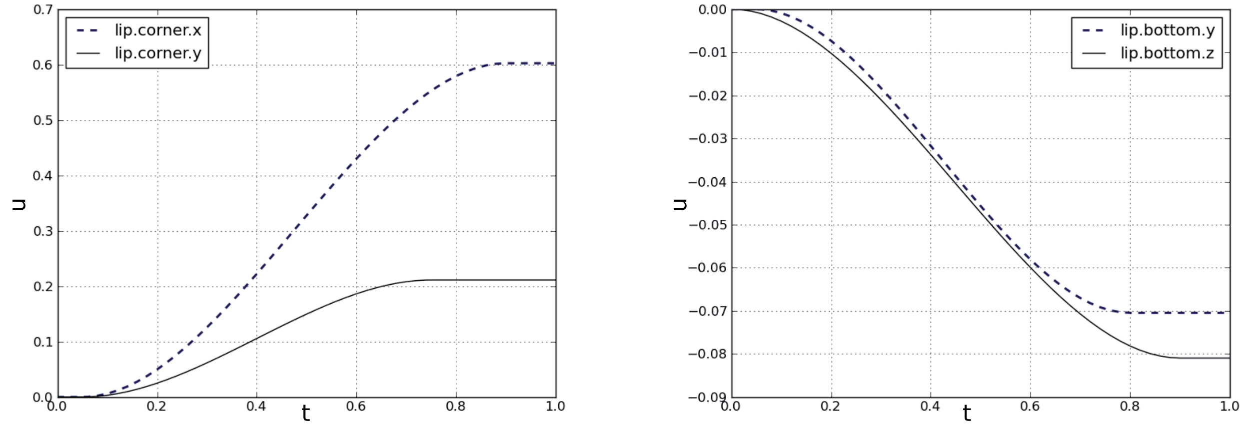
Rys. 4. Przykładowe trajektorie ruchu dla różnych parametrów u, t_d oraz t_s w przypadku współrzędnych kości: cheek2.l-x (a), lip.bottom.loc.l-y (b), lip.upper.loc.r-z (c) i cheek1.l-z (d).
Fig. 4. The example of movement trajectories for different parameters u, t_d and t_s in the case of bones' coordinates: cheek2.l-x (a), lip.bottom.loc.l-y (b), lip.upper.loc.r-z (c) and cheek1.l-z (d).



Group		Group	
1	brow.inner.l brow.outer.l upperlid.l	5	brow.inner.r brow.outer.r upperlid.r
2	lowerlid.l cheek1.l	6	lowerlid.r cheek1.r
3	lip.corner.l cheek2.l	7	lip.corner.r cheek2.r
4	lip.upper.l lip.upper lip.upper.r	8	lip.bottom.l lip.bottom lip.bottom.r jaw

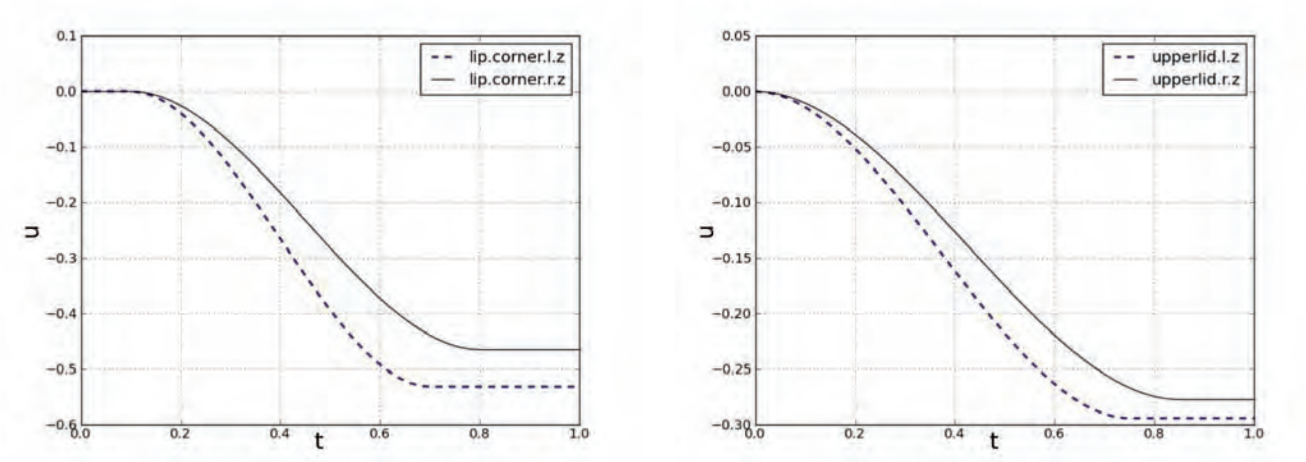
Rys. 5. Grupy kości zgodne z kierunkiem ruchu
Fig. 5. Groups of bones consistent with the direction of movement

2. Leslie G. Farkas and Ian R. Munro: *Antropometric facial proportions in medicine*. Charles C. Thomas, 1987.
3. Paul Ekman, Wallace V. Friesen and Joseph C. Hager: *Facial Action Coding System. The Manual*. Research Nexus division of Network Information Research Corporation, 2002.
4. Reyes Enciso, Alex Shawa, Ulrich Neumannb, and JamesMaha: *3D head anthropometric analysis*. In proceedings of the international society for optical engineering, SPIE Medical Imaging, vol. 5029, pp. 590-597, 2003.
5. Marco Fratarcangeli: *Physically Based Synthesis of Animatable Face Models*. Workshop On Virtual Reality Interaction and Physical Simulation, 2005.
6. Kenny Erleben, Jon Sporring, Knud Henriksen, and Henrik Dohlmann: *Physics Based Animation*. Charles River Media, Inc., 2005.



(a) (b)

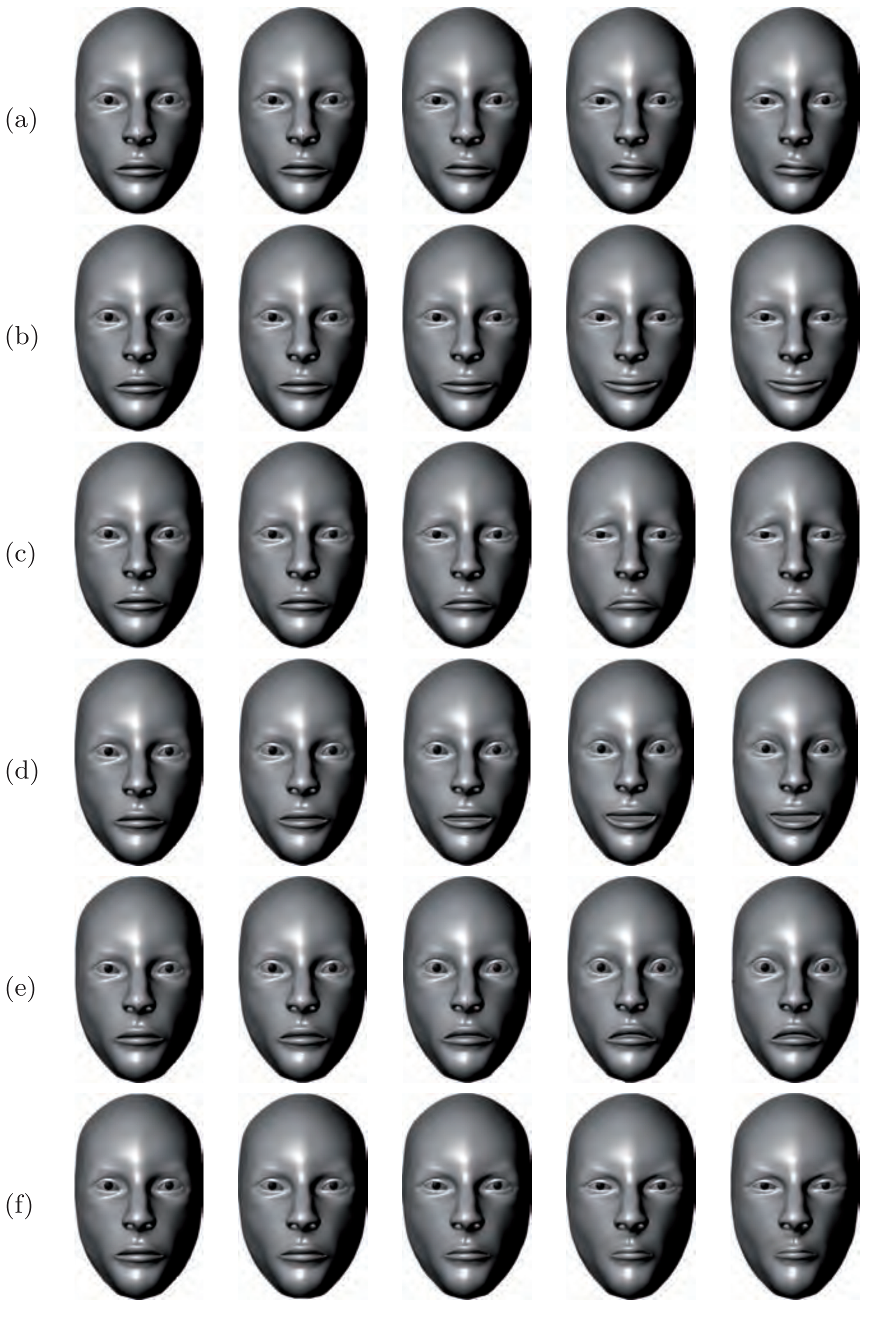
Rys. 6. Trajektorie ruchu dla współrzędnych kości z obszaru ust: lip.corner w stanie zdziwienia (a) oraz lip.bottom w stanie radości (b)
Fig. 6. Movement trajectory for bones' coordinates of lips area: lip.corner in surprise state (a) and lip.bottom in joy state (b)



(a) (b)

Rys. 7. Asymetria ruchu na przykładzie stanu niesmak dla współrzędnych kości z obszaru ust: lip.corner.z strona prawa i strona lewa twarzy (a), dla współrzędnych kości z oszaru oczu: upperlid.z strona prawa i lewa (b)
Fig. 7. Movement asymmetry on the example of disgust state for the bones' coordinates from the mouth area: lip.corner.z right and left side of face (a), for the bones' coordinates of eyes area: upperlid.z right and left side (b)

7. Terence Benja Kuderle: *Muscle-Based Facial Animation*. University Of Applied Science Wedel (FH) Diploma Thesis, 2005.
8. Margaret J. Fehrenbach and Susan W. Herring: *Illustrated Anatomy of the Head and Neck*. Saunders; 3 edition, 2006.
9. V. Costa Orvalho, Ernesto Zacur and Antonio Susin: *Transferring Facial Expressions to Different Face Models*. Ibero-American Symposium on Computer Graphics, 2006.
10. Rick Parent: *Computer Animation. Algorithms and Techniques*. Elsevier, 2008.
11. Yan Tong, Wenhui Liao and Qiang Ji: *Inferring Facial Action Units with Causal Relations*. Computer Vision and Pattern Recognition, 2008.
12. Lei Zhang, Yan Tong, and Qiang Ji: *Interactive Labeling of Facial Action Units*. 19th International Conference on Pattern Recognition, 2008.
13. Nadia Mana and Fabio Pianesi: *Modelling of Emotional Facial Expressions during Speech in Synthetic Talking Heads using a Hybrid Approach*. Auditory-Visual Speech Processing (AVSP2007), 2007.
14. Breazeal, C.,: *Designing Sociable Robots*, MIT Press, 2002.
15. Vidrascu, L. and Devillers, L.: *Real-Life Emotion Representation and Detection in Call Centers Data*. Lecture Notes in Computer Science, Vol. 3784, pp. 739-746, 2005.
16. Fragopanagos, N. and Taylor, J. G.: *Emotion recognition in human-computer interaction*. Neural networks : the offcial journal of the International Neural Network Society, vol. 18, no. 4, pp. 389-405, 2005.
17. Preben Wik and Anna Hjalmarsson: *Embodied conversational agents in Computer Assisted Language Learning*. Speech Communication, vol. 51, no. 10, 2009.
18. Catherine Pelachaud: *Modelling Multimodal Expression*



Rys. 8. Przykłady syntezy ekspresji twarzy uzyskane w procesie translacji współrzędnych kości: niesmak (a), radość (b), smutek (c), zdziwienie (d), strach (e) oraz złość (f)

Fig. 8. Examples of synthesis of facial expressions obtained in the translation of bones' coordinates process: disgust (a), joy (b), sadness (c), surprise (d), fear (e) and anger (f)

of Emotion in a Virtual Agent. 2009.

19. Clavel, C., Plessier, J., Martin,J.-C., Ach, L., Morel, B.: *Combining Facial and Postural Expressions of Emotions in a Virtual Character*. Published in: Proceeding IVA '09 Proceedings of the 9th International Conference on Intelligent Virtual Agents, 2009.
20. E. Chaudhry L. H. You Jian J. Zhang: *Character skin deformation: A survey*. Seventh International Conference on Computer Graphics, Imaging and Visualization, 2010.
21. Jason Weber: *Run-Time Skin Deformation*. In Proc of Game Developers Conference, 2000.
22. F. Erol: *Modeling and Animating Personalized Faces* M.Sc. Thesis, Bilkent university, January 2002.
23. V. Blanz and T. Vetter: *A Morphable Model for the Synthesis of 3D Faces*. Proc. SIGGRAPH, vol. 33, pp. 187-194, 1999.
24. Yu Zhang, Edmond C. Prakash and Eric Sung: *A New Physical Model with Multilayer Architecture for Facial Expression Animation Using Dynamic Adaptive Mesh*. IEEE Transactions on Visualization and Computer Graphics, vol. 10, no. 3, 2004.

25. Narendra Patel, Mukesh Zaveri: *3D Facial Model Construction and Expressions Synthesis using a Single Frontal Face Image.* International Journal of Graphics Vol. 1, No. 1, 2010.
26. Carlos Garre, Alvaro Pérez: *A simple Mass-Spring system for Character Animation.* Computer Animation Final Project, 2008.
27. Peng Huang, Adrian Hilton, Jonathan Starck: *Human Motion Synthesis from 3D Video.* Computer Vision and Pattern Recognition, pp. 1478-1485, 2009.

Animacja ekspresji twarzy na podstawie modelowania trajektorii ruchu

Streszczenie: Niniejszy dokument opisuje sposób syntezy trajektorii punktów charakterystycznych twarzy podczas zmian emocji. Punkty zostały wybrane zgodnie z właściwościami anatomicznymi twarzy człowieka oraz na podstawie systemu FACS opisującego aktywność twarzy. Do modelowania zmian emocjonalnych na trójwymiarowym geometrycznym modelu twarzy człowieka zaproponowano uogólnioną postać krzywej ruchu.

Słowa kluczowe: ekspresje twarzy, punkty charakterystyczne, trójwymiarowy model, animacja twarzy, interakcja człowiek-maszyna (HMI)

mgr inż. Maja Kocoń

Maja Kocoń received her M.Sc. (2008) in Automatics and Robotics from Szczecin University of Technology. She is currently pursuing the Ph.D. degree at the West Pomeranian University of Technology. Her primary research interests include human-robot interaction, sociable robots and emotional context analysis of robots and avatars.
e-mail: kocon@zut.edu.pl

prof. dr hab. inż. Zbigniew Emirsajłow

Zbigniew Emirsajłow graduated in 1978 from Faculty of Electrical Engineering at Szczecin University of Technology, where in the years 1993-2009 he was heading the Department of Control Theory and Simulation Techniques. He is currently heading the Department of Control and Measurements. He is an author of over 90 papers, dedicated mainly to linear distributed parameter systems, infinite-dimensional Sylvester and Lyapunov equations and nonlinear systems.
e-mail: emirsaj@zut.edu.pl