Theoretical and Applied Informatics
ISSN 1896-5334
Vol. 26 (2014), no. 3, 4
pp. 137 - 148

Realistic modelling of facial expressions using volume morphing

Przemysław KOWALSKI[a]

[a] Institute of Theoretical and Applied Informatics
Polish Academy of Science
ul. Bałtycka 5, Gliwice, Poland
przemek@iitis.gliwice.pl

Abstract: The article presents experiments in realistic modelling of facial expressions using volume morphing. The experiments use a variable number of points and face features. As the results are shown meshes deviations (between goal mesh and morphed shape). Experiments have shown that it is impossible to carry out a fully realistic morphing based on existing software. At the same time, even imperfect morphing allows you to show the expected emotional facial expression.

Keywords: 3D scanner, 3D mesh, face expression, morphing

1. Introduction

Realistic modelling of facial expressions is important for applications such as man-machine interface, video games, medical applications or virtual reality. Interest in the subject dates back to the 70s. There is a number of proposed solutions – key-frame interpolation, parameterized models, control-point models, a performance-driven approach and physically-based models [14]. Some of the solutions are strictly connected to the applications, e.g. physically-based models are used for medical applications [4], knowledge of the facial muscles location [6].

The human face depends on an underlying skull to which soft and hard tissues are attached. Because all faces are identical in basic design, we must observe rather subtle differences, mainly caused by muscle's work. Although the muscles are precisely described in textbooks of human anatomy, there are individual differences between people in the muscles themselves – some muscles are not present in some individuals, and some muscles can be asymmetric [3].

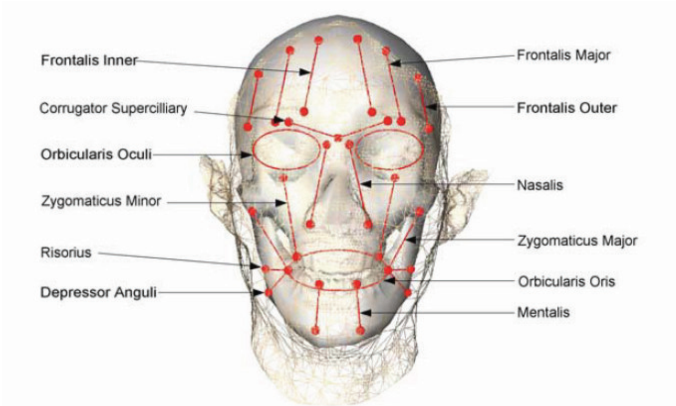

Fig. 1. Integrated face model with a hierarchical structure [13] which shows face muscles used for facial expression. For example, happiness is represented by two dominant contractions, which correspond to the left and right zygomatic majors.

Typical face features used for emotion expression are – in the eye region:

- eyebrow arch,
- eyelid,
- eyeball size, eye gaze,
- pupil size, iris size,

and features in the mouth region:

- jaw rotation,
- mouth expression, width,
- upper lip position,
- control of mouth corners [10].

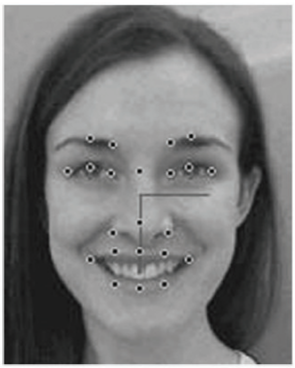

Fig. 2. The most important points tracked by NEVEN vision software [1].

People can distinguish many types of emotions – there are more than 180 kinds of smile, each of them differs by using different muscles [3]. All system developers must choose the important features of the face used by the system. Typical features are eyes, eyebrows, nose, lips corners (and outline, see fig. 2, 3, 4). For systems that track face and emotions (see fig. 5), the number of features is relatively small (e.g. system NEVEN uses 53 measurements for each frame [1]). For multimedia purposes (MPEG format) are labelled points on eyebrow arch (3 points for each eyebrow), eyes (7 points on each), nose (11 points), lips (18 points), chin (4 points), muscle hooks (2 for each chick) and some points for face outline (about 9 points) [7]. In the case of 3D graphics, the number of controlled features is relatively high (e.g. in Pixar movie „Toy Story" there are several thousands of control points [11]). The Parke's Parametric Face Model [2] consists of 300 triangles and quads for them defined ten expression parameters and about twenty conformation parameters.

A large number of points necessary to reproduce realistic emotions. On the other hand, the use of too many points is expensive (in terms of memory and time), and may transfer to model the details of the individual, not representative of all people.

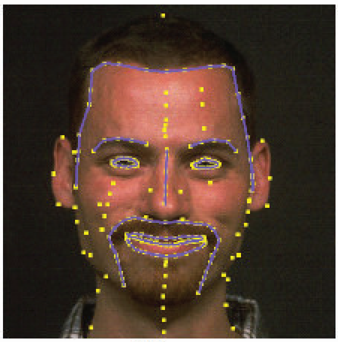

Fig. 3. Facial features annotated using a set of curves [11].

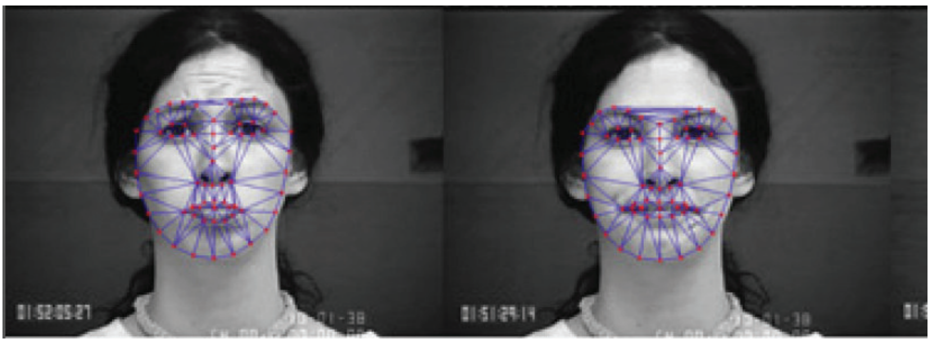

Fig. 4. Mesh generated for AAM algorithm [Zhanga]

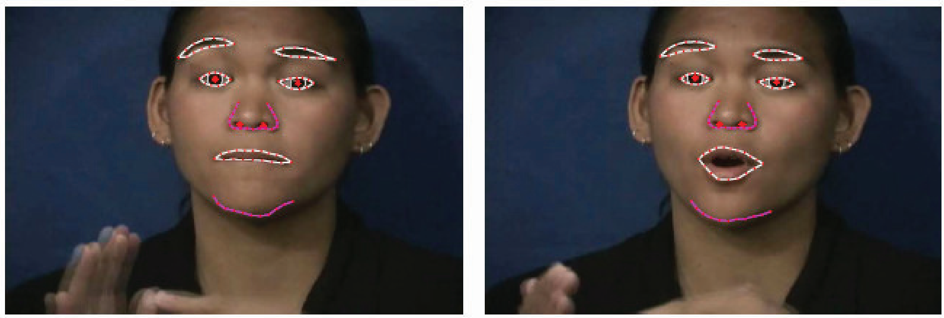

Fig. 5. The example of precise detection of facial features [8].

Most of the features in presented applications are tracked in 2D that does not take into account the variability of 3D face surface. We can see that people do not need 3D information to perceive human emotions.

2. Morphing

Morphing technique transforms one shape into another shape, with intermediate stages shapes among them. The morphing technique was popularized for 2D images, but can be used in 3D. Typical 3D realization morphs one 3D volume into another 3D volume while the set of point defines transformation. It means the mesh between defined points is not morphed into a given mesh – only the defined points on the mesh are morphed on the points defined by the destination mesh [5]. We can calculate the marker's influence on the mesh vertices [9]. Morphing between natural objects provides information and tools for the study of the morphological and physiogenetic relationships between them [5].

In our case, we transform 3D mesh, but the mesh is transformed using volume morphing – the procedure morphs general geometry of the face, but the transformation of the local characteristics requires additional characteristic points.

Examples were prepared using RapidForm 2006 package and morphing procedure. The procedure morphs one 3D mesh to another. The procedure creates a new 'goal' mesh (made from the 'start' mesh and position of features on the given 'goal' mesh). Function allows to generate several (up to five) intermediate meshes. Before morphing, we should transform the 'start' mesh to the proper position.

3. The face data

We use the faces scanned by 3dMD in The Institute of Theoretical and Applied Informatics of the PAS. The 3dMD is a structured light scanner and produces 3D mesh with texture. The mesh can be processed to reproduce the face better [12]. The set of faces presented a group of young people, some of them in different emotional states.

4. Examples of 3D face mesh morphing

In the first experiment, we use a minimal number of points. The list of points consists of two points for lips, one for a chin, two for forehead, two for a nose, two additional for ears positions.

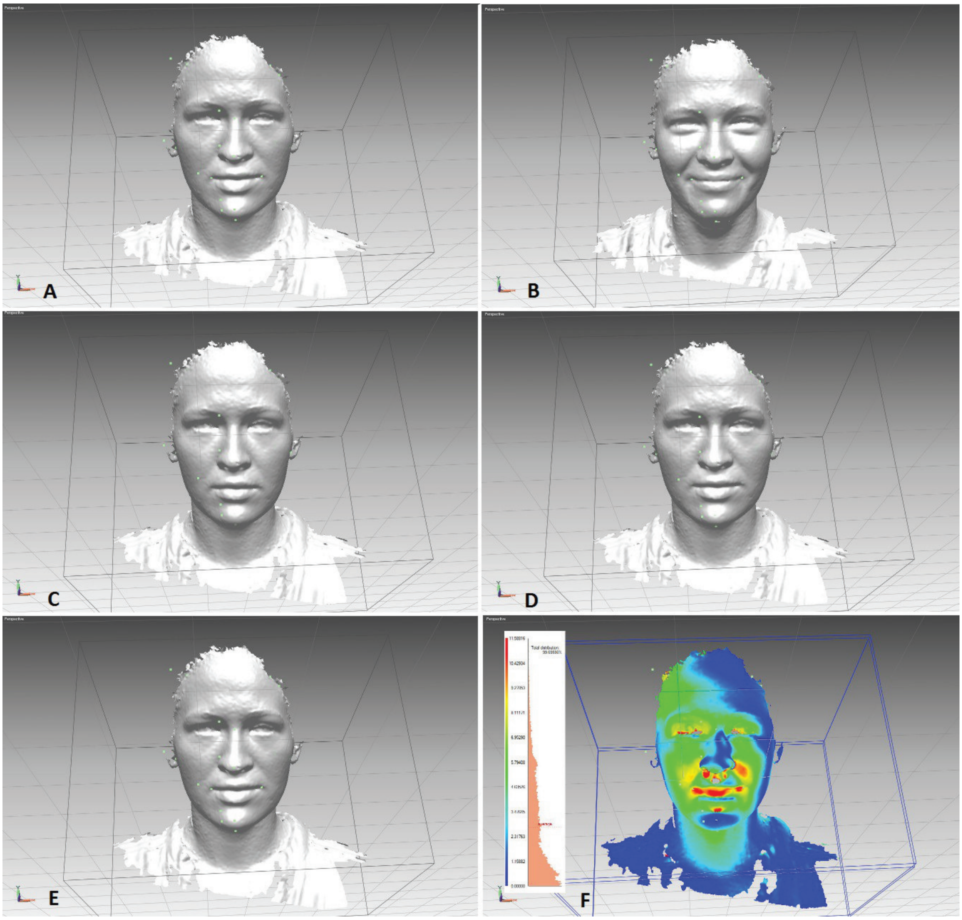

Fig. 6. The experiment with a small number of points. A: the source face, B: the target face, C-E: intermediate stages, F: difference between target face and morphed face.

The morphed face looks 'nice', but does not show expression of the given target face (see fig. 6).

The experiments were inspired by points defined in known systems, especially NEVEN [1], because of limitations of RapidForm (maximal number of feature points is set to 50). In the first experiment, we use only twenty-nine feature points.

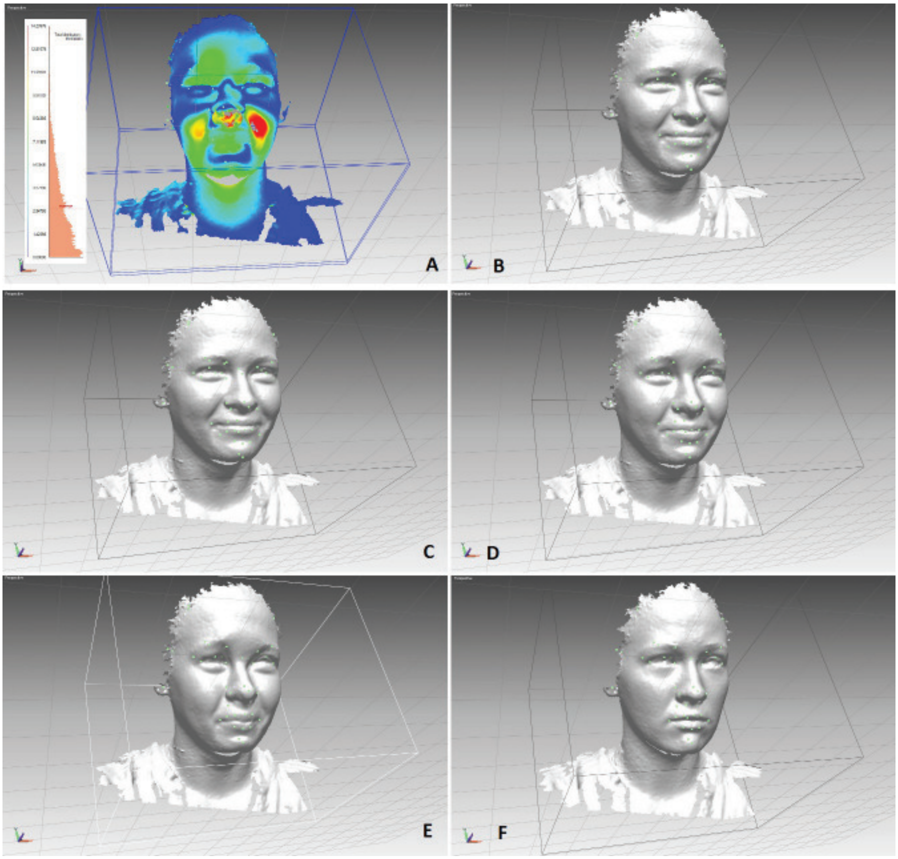

Fig. 7. The experiment with reduced number of points (inspired by NEVEN system).

The result of the experiment is shown on the fig. 7. The girl with a smile (B) was morphed into a sad girl (E) while the given mesh of the sad girl (F) differs (A), especially on left chick. Parts of the morphed mesh with the biggest errors are situated near lips, especially its nasolabial folds.

In the presented example (morphing from 'smile' to 'sadness', for the same person), we use thirty-nine feature points: three for a forehead (left, central and right), four for a nose (root of the nose, tip of the nose, both wings of the nose), four for an each eye (the left and right corners of the eye, the bottom and upper part of the eye), three for eyebrows (left, right, and central), nasolabial folds (both sides, each given by two points: upper and bottom), three additional points for each cheek (near check bone, near ear, and behind the mandible), five points for lips (left and right corners, bottom and upper part of the lips, central point), two points for chin (left and right), and one below nose (see fig. 8).

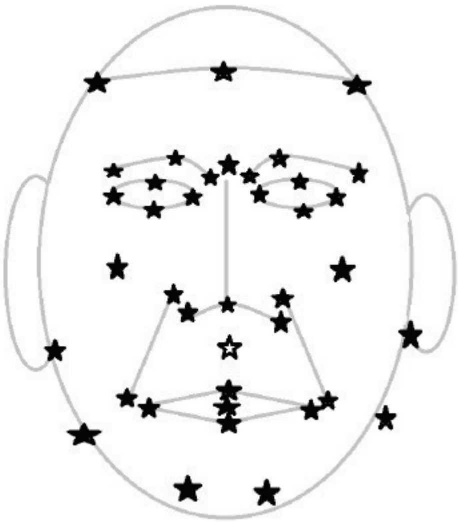

Fig. 8. Feature points used in the test (see fig. 9, 10.).

Points were chosen using two main criteria: the ease of location and suitability to describe the 3D surface of the face. On the fig 10 (d), we can see that the biggest problem for the proper realistic morphing of the mouth. The mouth region is highly variable, due to the great importance of soft tissue; and the difficulty in pointing invariant features points in such region.

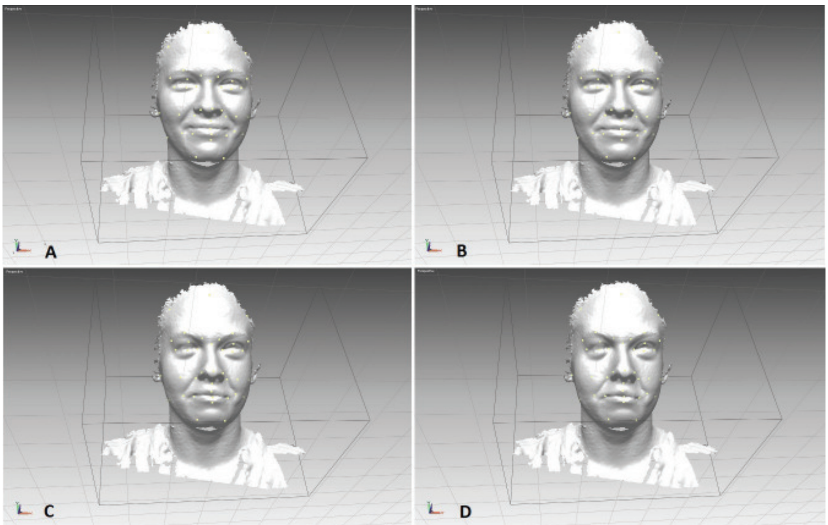

Fig. 9. Intermediate stages of mesh morphing (from „A" – smile, to „D" – sadness). For a 'start' and 'goal' mesh, see fig. G.

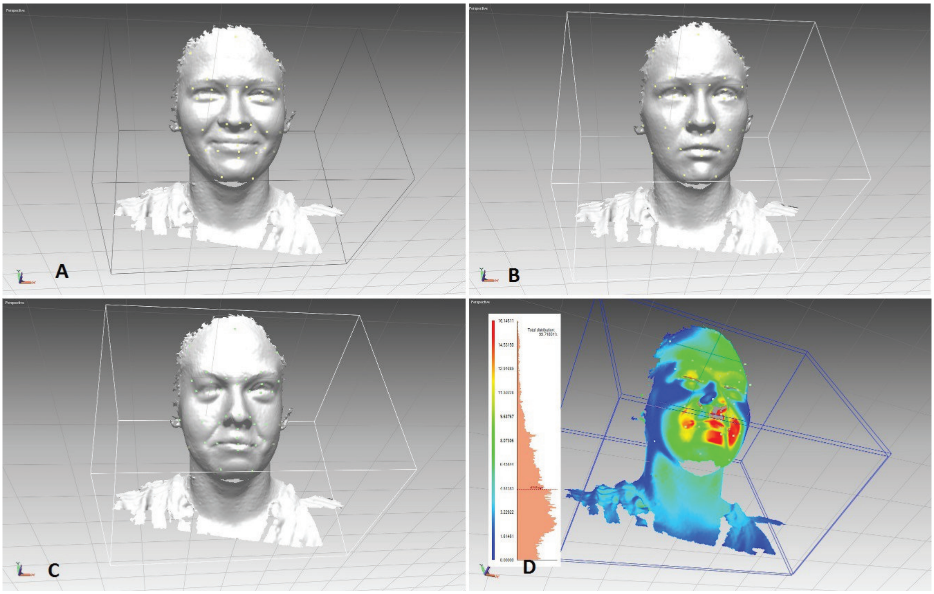

Fig. 10. The same example (see fig. 9). A – the 'start' mesh. B – the given 'goal' mesh. C – the generated (morphed) 'goal mesh. D – the difference between the given and the generated goal.

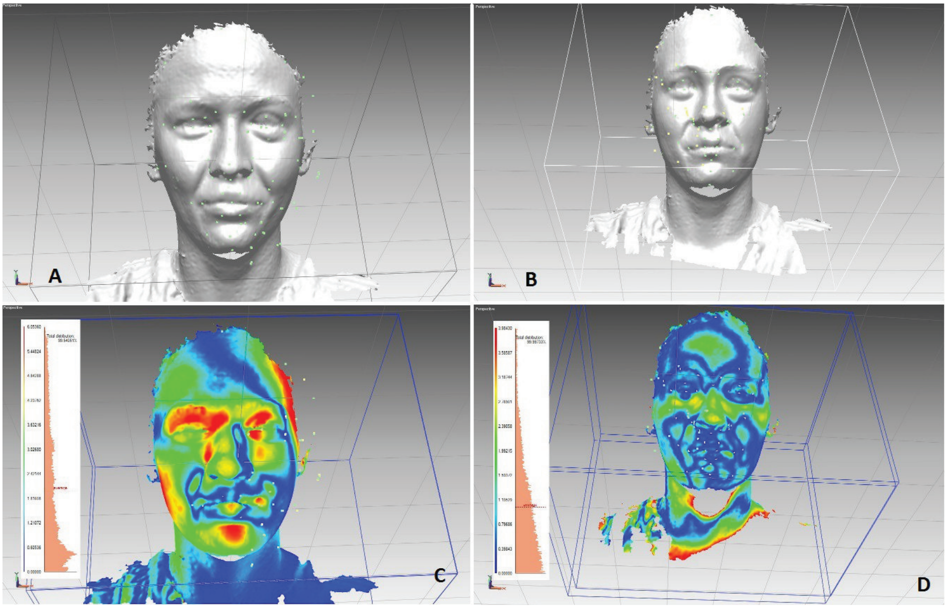

Fig. 11. Two morphings between the same pair of faces, with different number of feature points. (A-B: final morphed faces; C-D – meshes deviation.)

The fig. 11 presents small impact of the additional feature points. For the example A-C about 38 points were used, for the B-D example the number of used points is 48. Additional feature points slightly reduced errors, but the overall appearance of the face is still unsatisfactory.

On the fig. 11, we can also observe that errors in overall appearance of the face (observed by humans) does not respond to the mesh deviation image. The biggest differences between meshes are on cheeks, nose and forehead, the parts of the face that changes are not important for emotion expression.

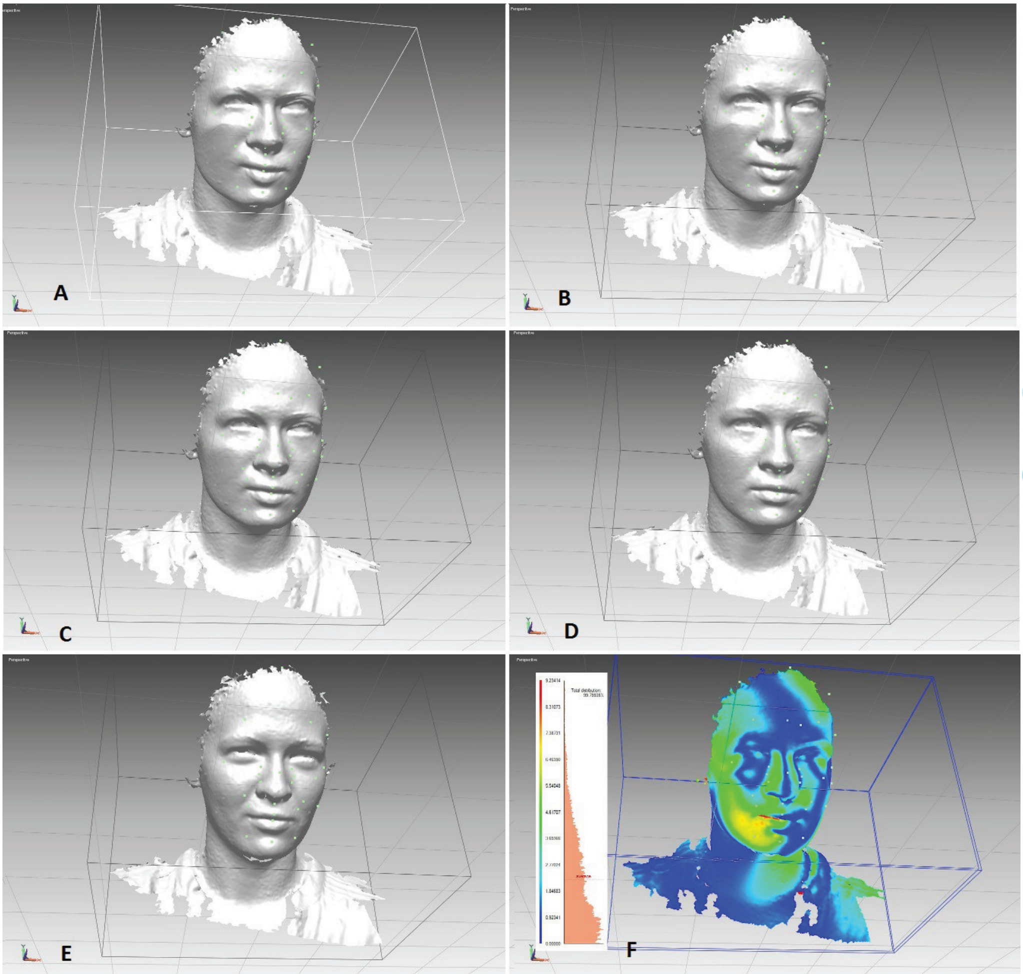

Fig. 12. Morphed face mesh with a wrong position of feature point on the right chick.

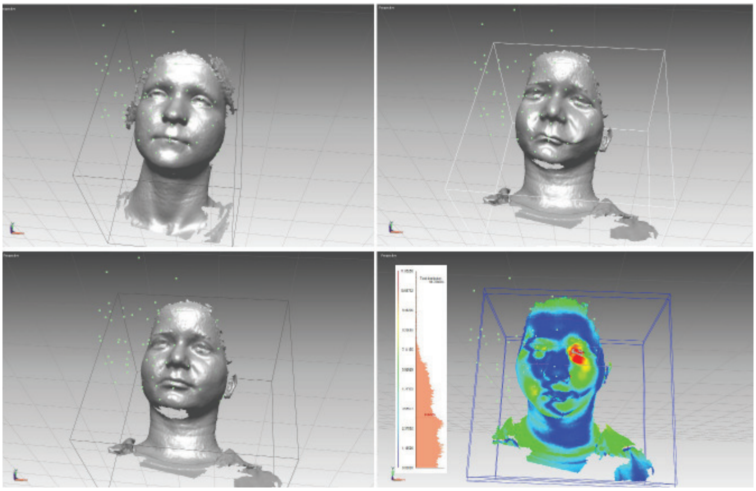

Fig. 13. Extremely bad morphing caused by problem with positioning of the feature points.

The worst matching points on the face shown on the fig. 13. The problem was caused by positioning of some feature points, needed to model part of the face. Some points cannot be pointed precisely – in this example, the problems occur with chicks, one eye (wrong position), and nasolabial folds (not visible in this example). We may note that the visually visible error is much more important than absolute measurements of error, e.g. mean error.

5. Summary

The requirements and capabilities of selecting points to morphing 3D grids depicting faces, give a small range of choice. On the one hand, we are limited by the requirements reflect the facial expressions of the face (indication of mouth, eyebrows, eyes), on the other hand, we encounter two restrictions. The first comes with the software requirements (the maximum number of points). The second is the number of potential landmarks that you can point to the human face.

A small range of choice of answering a slight range obtained quality morphing. The exception here is the experiment with too few points (nine points, fig. K). A positive aspect is the fact that even a small realism in reconstituted soft tissue enough to determine the emotional facial expression.

Acknowledgments

This work has been supported by the National Centre for Research and Development, Poland (project: INNOTECH In_Tech ID 182645 „Nowe technologie wysokorozdzielczej akwizycji i animacji mimiki twarzy" – „New technologies of high resolution acquisition and animation of facial mimicry").

References

[1] Bailenson J.N., Pontikakis E.D., Mauss I.B., Gross J.J., Jabon M.E., Hutcherson C.A.C., Nass C., John O., *Real-time classification of evoked emotions using facial feature tracking and physiological responses*, International Journal of Human-Computer Studies, vol. 66, 2008, pp. 303-317

[2] Blanz V., *Overview Facial Animation Techniques*, SIGGRAPH 2004

[3] Bruce V., Young A., *Face Perception*, Psychology Press, 2012

[4] Delingette H., Subsol G., Cotin S., Pignon J., *A Craniofacial Surgery Simulation Testbed*, Proceedings of the Visualization for Biomedical Computing, Rochester, 1994

[5] Fang S., Srinivasan R., Raghavan R., Richtsmeier J.T., *Volume morphing and rendering—An integrated approach*, Computer Aided Geometric Design, vol. 17, 2000, pp. 59-81

[6] Hui K.C., Leung H.C., *Expression modeling – a boundary element approach*, Computer & Graphics, vol. 30, 2006, pp. 981-993

[7] Kshirsagar S., Garchery S., Magnenat-Thalmann N., *Feature Point Based Mesh Deformation Applied To MPEG-4 Facial Animation*, DEFORM '00/AVATARS '00 Proceedings of the IFIP TC5/WG5.10 DEFORM'2000 Workshop and AVATARS'2000 Workshop on Deformable Avatars 2001

[8] Martinez A., Du S., *A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives*, Journal of Machine Learning Research, vol. 13, 2012, pp. 1589-1608

[9] Pęszor D., Wojciechowski K., Wojciechowska M., *Automatic markers' influence calculation for facial animation based on performance capture*, ACIIDS 2015 (paper accepted)

[10]. Parke F., *Parametrized Face Models*, SIGGRAPH2004

[11] Pighin F., Hecker J., Lischinski D., Szeliski R., Salesin D., *Synthesizing Realistic Facial Expressions from Photographs*, Proceedings of ACM SIGGRAPH'98, New York, NY, 1998, pp. 75-84

[12] Skabek K., Łapczyński D., *Reconstruction of Head Surface Model from Single Scan*, CORES 2013, pp. 483-492

[13] Zhang Y., Prakash E.C., Sung E., *Hierarchical Facial Data Modeling for Visual Expression Synthesis*, Journal of Visualization, Vol. 6., No. 3, 2003, pp. 313-320

[14] Zhang Y., Wei W., *A realistic dynamic facial expression transfer method*, Neurocomputng, vol. 89, 2012, pp. 21-29

Realistyczne modelowanie wyrazu twarzy z wykorzystaniem morfingu

Streszczenie

Artykuł prezentuje wyniki eksperymentów dla przestrzennej siatki twarzy z jednego jej wyrazu (ekspresji) do drugiego. Dla eksperymentów wykorzystano dane pozyskane przy pomocy skanera 3dMD, obejmujące kilka twarzy, z których część przedstawia różne stany emocjonalne. Ograniczono się przy tym, do wyraźnie rozróżnialnych stanów (śmiech, skupienie, smutek). Do przekształcania siatek wykorzystano procedurę 'polygon morphing' pakietu RapidForm2006.

W trakcie prac skupiono się na zagadnieniu wyboru zestawu znaczników (punktów charakterystycznych), niezbędnego dla realistycznego modelowania wyrazu emocjonalnego twarzy przy pomocy morfingu. Uzyskane wyniki poddano porównaniu z docelową siatką. Rozkład błędów dopasowania wskazuje na te części twarzy, które nie zostały odpowiednio dopasowane. Rozkład ten pozwala wnioskować na temat istotnych obszarów twarzy, które wymagają zdefiniowania dodatkowych punktów charakterystycznych.

Definiowanie punktów charakterystycznych napotyka dwa ograniczenia:

- ograniczenie programowe, związane z wykorzystywanym oprogramowaniem (maksymalnie pięćdziesiąt par punktów charakterystycznych);
- ograniczenia rozpoznawalnych cech charakterystycznych, które mogą zostać wskazane przez operatora.

W trakcie prac osiągnięto pewien kompromis, ograniczając się do mniej niż 50 punktów charakterystycznych, które odtwarzają najważniejsze elementy twarzy istotne dla przedstawienia emocji. Uzyskany wynik nie jest jednak w pełni zadowalający, gdyż część elementów twarzy, zwłaszcza okolice ust, pozostaje bardzo plastycznych, nie zapewniając jednocześnie potencjalnych punktów charakterystycznych.