Facial expression reconstruction on the basis of selected vertices of triangle mesh

Damian Pęszor* and Marzena Wojciechowska†

*Institute of Informatics, The Silesian University of Technology, Akademicka 16, 44-100 Gliwice, Poland
†Polish-Japanese Academy of Information Technology, Koszykowa 86, 02-008 Warsaw, Poland

Abstract.
Facial expression reconstruction is an important issue in the field of computer graphics. While it is relatively easily to create an animation based on meshes constructed through video recordings, this kind of high-quality data is not often transferred to another model because of lack of intermediary, anthropometry-based way to do so. However, if high-quality mesh is sampled with enough density, it is possible to use obtained feature points to encode the shape of surrounding vertices in a way that can be easily transferred to another mesh with corresponding feature points. In this paper we present method used for obtaining information used to reconstruct changes in facial surface on the basis of selected feature points.

Keywords: facial expressions, reconstruction, vertices selection, facial animation
PACS: 07.05.Tp

INTRODUCTION

Video-based markerless performance capture is a set of techniques which are able to reconstruct the shape of facial surface on the basis of video recordings. While the issue of finding correspondences between cameras and tracking specific features throughout recording is much more complex than in case of marker-based techniques, using video data has it's significant advantages. Most importantly, it is not limited by the number of markers that can be applied to human face, instead working with the quality that depends mostly on the resolution of recordings, which is reflected by high polygon counts of resulting meshes. The data obtained this way can be easily used in the form of animation of face that is reconstructed in the process of performance capture. However, animating another character requires a way to reconstruct facial expression on the basis of movement of subset of vertices composing obtained mesh, which have known correspondences in the mesh that needs to be animated. Proposed algorithm is used to reconstruct facial expression on each segment of facial mesh constrained by given feature points.

FACIAL MESH

The method presented in this paper is designed to reconstruct facial expression on the basis of small subset of vertices of original mesh obtained from video recordings. While the method was inspired by issues associated with facial animation and tested on meshes representing human facial surface, there is no reason to contain it to this field only, as it can be used with any soft surfaces characterized by feature points. Meshes tested with presented algorithm were obtained using structured-light based scanner and stereophotogrammetry. Depending on the technique used to obtain facial mesh, various errors might be present in the model, such as; discontinous surfaces, surfaces representing different parts of human body or clothes, topological errors like non-manifold edges and vertices, noise on edges, gaps and holes. Approach used for preprocessing of this data was described in [1].

SEGMENTATION

Subset of vertices of the mesh which will be a basis of surface deformation is topologically connected to other vertices. However, no explicit data containing relation between feature point and mesh' vertices is available as a standard for mesh construction algorithms. This information is needed to decide which feature points affect given vertex and how much each of them do so. In fact, feature points and their correspondence on animated model is not a standard output of construction algorithms, therefore there is a need to either find them manually, or automatically on the basis of colour, curvatures, anthropometric data and position of one feature point with relation to other. An attempt to analyse three-dimensional mesh to find basic fiducial points was addressed as part of previous research in [2].

With feature points selected, the mesh can be divided into almost triangular segments. Such segments can be defined as topologically correct mesh, in order to maintain properties of facial structure, however with enough feature points this is not really necessary and can be done automatically through tessellation algorithms. Few guidelines should be followed when creating segmentation mesh. First, every segment should represent possibly flat area, high density of feature points will result in this being fulfilled automatically. Second, to distribute influence of each feature point evenly, segments should be possibly equilateral, again, with high density of feature points standard tessellation algorithms produce appropriate results. Third, segments should be possibly small - again, a property easily obtainable through high density and proper tessellation. Due to the fact, that there is no clear association between vertices of analysed mesh and segmentation mesh, such an association has to be derived from geometrical and topological properties of analysed mesh.

As part of proposed approach, two methods for determining association between segment and vertices are used. First, graph-based method is used to cut out non-expressive facial areas (or areas that are not covered by feature points). With densely distributed feature points represented by vertices of triangular mesh, the shortest topological connection between them will correspond to smallest distance. A* [3] algorithm is therefore used to obtain paths between feature points connected in segmentation mesh. Finding each topological path corresponding to edges of segmentation triangle gives constraints that can be used to associate vertices enclosed by those paths to the segment by traversing through them with e.g. breadth first search algorithm. This method, however, emphasises topological nature of the mesh rather than surface properties.

Geometric method, on the other hand, is based on spatial properties of vertices. Each segment is represent by a plane restricted by edges connecting feature points. Between two segments, spherical linear interpolation [4] of normals can be used to obtain a plane passing through edge which will split the surface of the mesh. With three edges of each segment and therefore three intersecting planes, every vertex can be associated with the segment which is basis for pyramid that contains given vertex. Due to the nature of facial structure, parts of mesh that are not connected (for example, opposite parts of cheeks) could be identified as associated with same segment. With information gathered by graph based method, those ambiguities can be easily resolved.

FEATURE POINTS AS A BASIS FOR RECONSTRUCTION

While presented approach was first tested on marker-based performance capture, it proved to suffer from low density of markers and insufficient separation of facial features (see [5]). Data obtained from video-based performance capture, however, can be sampled by feature points with enough density to disregard previously encountered issues.

The apex p_a of pyramid constructed during geometric segmentation is calculated from three feature points $\{A,B,C\}$ and normals n_i to planes composing sides of pyramid:

$$p_a = \begin{vmatrix} n_{00} n_{10} n_{20} \\ n_{01} n_{11} n_{21} \\ n_{02} n_{12} n_{22} \end{vmatrix}^{-1} \quad [(A \cdot n_0)(n_1 \times n_2) + (B \cdot n_1)(n_2 \times n_0) + (C \cdot n_2)(n_0 \times n_1)] \tag{1}$$

Then, the analysed vertex v_i is projected (v_i') from p_a onto the base plane of the segment using it's normal n and one of feature points (A).

$$v_i' = (v_i - p_a)\frac{(A - p_a) \cdot n}{(v_i - p_a) \cdot n} + p_a \tag{2}$$

Finally, the barycentric coordinates of projection are calculated:

$$\begin{aligned} a &= \frac{(y_B - y_C)(x_{v'} - x_C) + (x_C - x_B)(y_{v'} - y_C)}{(y_B - y_C)(x_A - X_C) + (x_C - x_B)(y_A - y_C)} \\ b &= \frac{(y_C - y_A)(x_{v'} - x_C) + (x_A - x_C)(y_{v'} - y_C)}{(y_B - y_C)(x_A - X_C) + (x_C - x_B)(y_A - y_C)} \\ c &= 1 - a - b \end{aligned} \tag{3}$$

In order to include data from neighbouring segments so that there will not be any visible transitions between them, the surface is fitted to quadratic (or higher) Bézier triangle (see [6]). Surface of quadratic Bézier triangle composed of n points $p_i | i \in \mathbb{N} \wedge i \in [0, n-1]$ each having barycentric coordinates of a, b, and c related to feature points α^2, β^2 and γ^2 respectively and influenced by control points $\alpha\beta$, $\alpha\gamma$ and $\beta\gamma$ is described with following equation:

$$p_i(a,b,c) = (\alpha a + \beta b + \gamma c)^2 = \alpha^2 a^2 + \beta^2 b^2 + \gamma^2 c^2 + 2\alpha\beta ab + 2\alpha\gamma ac + 2\beta\gamma bc \tag{4}$$

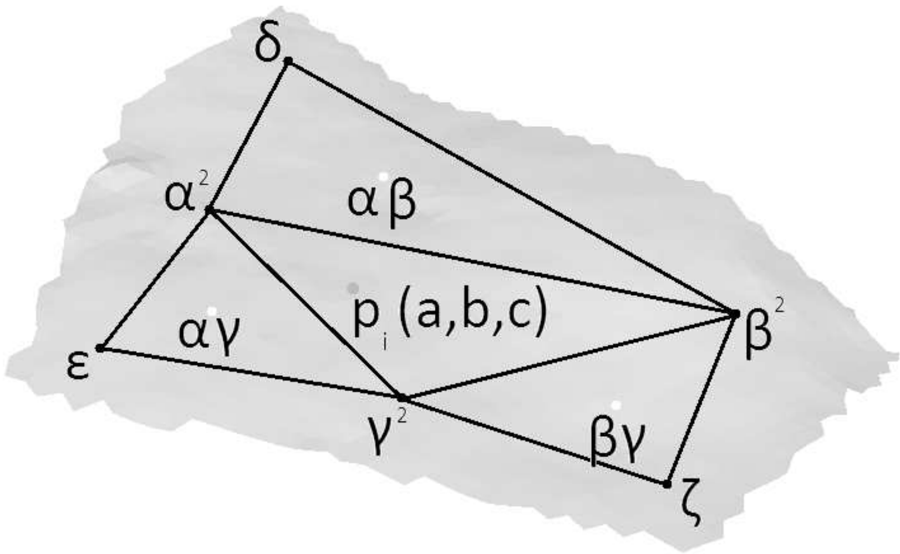

FIGURE 1. Example of surface. Black dots represent feature points, white dots represent additional control points of Bézier triangle and grey dot represents example point

Control points are estimated using nonlinear least squares fitting [7] in order to find a surface approximating actual points' positions. Once obtained, each control point weighted by barycentric coordinates is replaced with position of third vertex of neighboring segment $(\delta, \varepsilon, \zeta)$ with new weight. Similarly, influence of each considered feature point is calculated:

$$\begin{aligned} A &= a^2 \\ B &= b^2 \\ C &= c^2 \\ D &= \frac{2\alpha\beta ab}{\delta} \\ E &= \frac{2\alpha\gamma ac}{\varepsilon} \\ F &= \frac{2\beta\gamma bc}{\zeta} \end{aligned} \tag{5}$$

This can be used to reconstruct given vertex on the basis of feature points' positions' changes from neutral frame to current frame f:

$$\begin{aligned} v_{i,f}(A,B,C,D,E,F) = v_{i,0}(A,B,C,D,E,F) + A(\alpha_i - \alpha_0) \\ + B(\beta_i - \beta_0) + C(\gamma_i - \gamma_0) + D(\delta_i - \delta_0) + E(\varepsilon_i - \varepsilon_0) + F(\zeta_i - \zeta_0) \end{aligned} \tag{6}$$

CONCLUSION AND FURTHER WORKS

The approach described in this paper can be used to reconstruct facial expression data using subset of vertices of original mesh constructed from video recordings in order to minimize the data required to animate facial mesh. With proper selection of feature points on the basis of anthropometric properties, the reconstruction presented in this paper allows to retarget facial expression data from actor's face recorded as part of performance capture session to different model of facial structure, either constructed manually or obtained from different person.

While quadratic Bézier triangles correspond to the segmentation using triangular mesh, and higher order of Bézier triangles improve the results, it seems that substituting control points with feature points might not produce perfect results in case of low-density feature points and extreme expressions. It is therefore planned to test different parametric surfaces in order to find one better suitable to the task at hand.

ACKNOWLEDGMENTS

This project 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. D. Pęszor, A. Polański and K. Wojciechowski, "Preprocessing of 3D scanned images for facial animation on the basis of realistic acquisition," *AIP Conference Proceedings*, 2015, Vol. 1648 Issue 1
2. D. Pęszor, A. Polański and K. Wojciechowski, "Estimation of marker placement based on fiducial points for automatic facial animation," *AIP Conference Proceedings*, 2015, Vol. 1648 Issue 1
3. P. E. Hart, N. J. Nilsson, B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics*, IEEE, Vol. 4, Issue 2, July 1968, pp. 100-107
4. K. Shoemake, "Animating rotation with quaternion curves" *SIGGRAPH '85 Proceedings of the 12th annual conference on Computer graphics and interactive techniques*, pp. 245-254, July 1985
5. D. Pęszor, K. Wojciechowski, M. Wojciechowska, "Automatic Markers' Influence Calculation for Facial Animation Based on Performance Capture," *Lecture Notes in Computer Science*, 2015, Vol. 9012, pp. 287-296
6. G. Farin, *"Curves and Surfaces for CAGD"*, 5th ed., Academic Press. ISBN 1-55860-737-4
7. D. M. Bates, D. G. Watts, *"Nonlinear Regression Analysis and Its Applications"*, New York, Wiley, 1988, ISBN 978-0-471-81643-0