Introducing Control Cluster Method into Blend Shapes Facial Animation*

A.A. Bukatov[1]***, E.E. Gridchina***[1]***, D.A. Zastavnoy***[1,2]

baa@sfedu.ru|helen.gridchina@gmail.ru|dzast@sfedu.ru

[1]Institute of Mathematics, Mechanics and Computer Sciences, IT Department of Southern Federal University, Rostov-on-Don, Russia

[2]Russian Customs Academy, Rostov-on-Don, Russia

Rigging is a major part of facial animation process making it easier to edit facial expressions and manipulate the face itself. Blend shape models are considered to be the predominant approach of setting up the facial rig to achieve believable facial animation. Control Cluster Method proposed by the authors and previously implemented for skinning tasks could also find usage in facial animation. Control cluster method proposes to split all the vertices of the polygonal model mesh into two groups: cluster and free vertices. Free vertices position is calculated based on cluster vertices position in a non-linear fashion. Advantages of the blend shape model as well as its linear nature limitations and the ways of overcoming them with the help of Control Cluster Method are discussed in the context of solving eye blink animation task.

Keywords: *facial animation, control cluster, non-linear, blend shapes*

1. Introduction

Facial animation has always been considered a challenging area of computer graphics as face and facial expressions are fundamental to convey believable character emotions animation [1, 2]. The process known as rigging involves experienced artists introducing and manipulating controls to create a realistic facial animation. The task of creating an adequate facial rig is rather complex and labour intensive.

Among the many facial rigging algorithms proposed nowadays blend shapes are believed to be the most widespread due to providing expressiveness, interpretability and simplicity combined. This research focuses on introducing non-linear approach to blend shapes in the frame of Control Cluster Method (CCM) [3, 4].

Previously Control Cluster Method (CCM) was proposed for non- linear skinning [5]. It makes use of control clusters to facilitate the work of the animator in creating realistic animation of the character without involving much user-input data. The method provides intuitive control and it is easy to use because it allows the user to influence only a small group of vertices leaving the non-linear deformation of the rest of the vertices to the algorithm.

2. Related work

A wide range of approaches to facial animation have been proposed [1]. A group of methods focuses on using physically based models, trying to approximate such elements of the face as bones, muscles, fat tissues and others [6]. Other approaches use scans or motion capture data to acquire principal component analysis (PCA) models [7]. Some approaches are based on interpolation in an abstract pose or expression space [8]. Others create facial animation by means of dense motion capture driving low-level meshes [9]. There are also parametric and blend shape models and some hybrid approaches for facial rigging. Among the methods proposed blend shape models, also referred to as blend shape targets and morph targets, remain demanded for believable human-like characters' facial animation [10]. When more complex approaches are used, it's common to use blend shapes as a base for physically based deformations.

3. Blend shape models

3.1 Algorithm

When using a blend shape model, a facial pose is defined as a linear combination of a range of blend shape targets, i. e. facial poses with different expressions. A number of facial poses can be achieved at a low computational cost by adjusting the weights of the linear combination.

Provided n is a number of blend shape targets and bi is a long vector comprising all k three component vertices of the i-th blend shape target, the blend shape model can be formulated as follows

$$f = \sum_{i=0}^{n} w_i b_i \tag{1}$$

(1) can also be formulated in matrix notation

$$f = Bw \tag{2}$$

where f is a $3k \times 1$ vector of the resulting facial pose, B is a $3k \times n$ matrix (its column vectors b_i represent blend shape targets – $3k \times 1$ vectors) and w is a $n \times 1$ vector representing weights. b_0 is considered to be the neutral face blend shape target.

*Работа опубликована по гранту РФФИ №16-07-20482.

For scaling all the weights not to result in scaling the whole head, the weights in (2) are constrained to sum to one making linear combination (1) affine. Linear combination may be also defined as a convex one by constraining weight to the interval [0,1]. Taking into account that $\sum_{i=0}^{n} w_i = 1$ and b_0 is the neutral blend shape target, blend shape models (1) and (2) can have a local formulation as

$$f = \sum_{i=0}^{n} w_i b_i = b_0 - b_0 +$$

$$+ \sum_{i=0}^{n} w_i b_i = b_0 - \sum_{i=0}^{n} w_i b_0 + \sum_{i=0}^{n} w_i b_i =$$

$$= b_0 + \sum_{i=0}^{n} w_i (b_i - b_0) \tag{3}$$

$$f = b_0 + Bw \tag{4}$$

It's typical for the difference between a certain blend shape target bi and the neutral blend shape target b0 to be limited to a small region of the face, so the resulting parameterization provides intuitive localized control.

3.2 Advantages and limitations

But simplicity, blend shape models have two key advantages. First of all, the approach offers semantic parameterization as the weights have intuitive meaning of different facial shapes influences (Figure 1), with other linear models such as PCA failing to provide it. The second major advantage of the approach is keeping the face model consistent as arbitrary deformations are not allowed by the model.

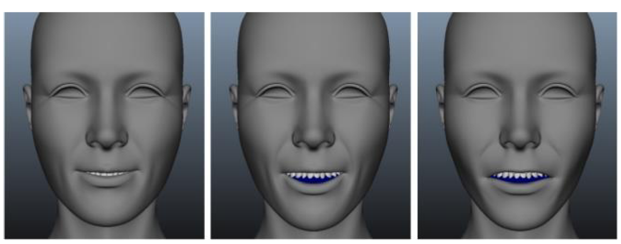

Рис. 1: Blend shape model as a semantic parameterization of facial expression. From left to right: a half-smile, a smile, open mouth. (courtesy of [10]).

Despite being conceptually simple blend shape models are hard to develop. To have a complete assortment of realistic facial expressions one has to create a wide range of blend shape targets which is time-consuming. Moreover, the blend shape model is sensitive to changes in the shape topology.

In comparison with PCA blend shapes targets are not orthogonal. While offering the benefit of interpretability, this feature of the approach can be seen as a shortcoming. The parameters being dependent, adjusting an individual parameter can damage previous edits.

The linear nature of the method can also have an impact on the resulting animation of the face. The transformation of each vertex involved has to be along a line as two weight are altering in a convex combination.

Control Cluster Method could defuse the impact of some limitations.

4. Control cluster method application to facial animation usage example control cluster method

4.1 Control Cluster Method

The main idea of Control Cluster Method consists in all the vertices of the model being split into two groups: cluster vertices and so-called free vertices. The position of the cluster vertices is to be defined by some input data. Cluster vertices control the deformation of the model as the key areas of the mesh defined by cluster vertices form the shape of the model. The position of free vertices is calculated automatically by a non-linear algorithm.

Cluster vertices can be transformed in a great variety of ways, either by direct mesh editing or by applying some deformers including lattices, cage or skeleton. Cluster vertices transformation could be also defined by data ranges from some type of measuring equipment, for example water level measuring device or motion capture data.

If a vertex does not belong to the cluster, it is considered a free one and its position is recalculated based on its position in the model topology and its nearest cluster vertices position. Free vertices position can be defined in several non-linear ways so that it corresponds to the position of the cluster vertices. This feature of CCM can help reduce shortcomings of linear nature of the blend shape model.

In previous works on Control Cluster Method [4, 5] Catmull-Rom splines [11], cubic cardinal splines generalizing cubic Catmull- Rom splines and Bezier patch meshes were used as a non-linear item of the method.

4.2 Blend shape models of an eye blink

The animation of an eye blink is a common task in facial animation. It's natural for the artist to create two blend shape targets to fulfill the task: the neutral blend shape target with an open eye and a blend shape target with a closed eye (Figure 2). But because of the linear nature of the blend shape model, the weight corresponding to the closed eye blend shape target in formula (4) alters from zero to one that causes the vertices of the eyelid to move the shortest way possible, i. e. along a line. And it results in eyelid vertices crossing the eye itself.

To achieve a better result in the frame of the blend shape model, it is possible to create an intermediate blend shape target of a half- closed eye. Production implementations of blend shape models, e. g. Autodesk Maya [12], allow that blend shape targets are created at intermediate weight values, thus providing piecewise linear interpolation. But it's necessary to create a large amount of intermediate blend shape targets for the eye blink animation to look believable.

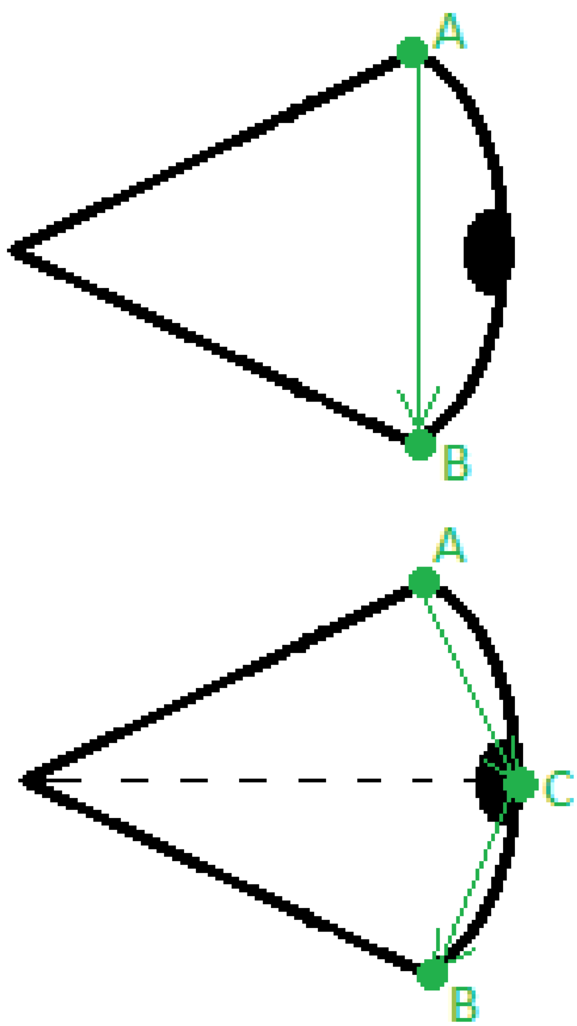

Рис. 2: Blend shape linearity shortcomings (shown schematically). A – opened eye blend shape target, B – closed eye blend shape target, C – half-closed eye blend shape target. From top to bottom: animation of an eye blink with two blend shape targets A and B; animation of an eye blink with an additional intermediate blend shape target C. The eyelid goes through the eye in both cases.

4.3 Introducing CCM to blend shape models

According to the Control Cluster Method all the vertices of the eyelid are to be divided into cluster and free vertices. The margin of the upper eyelid vertices are considered cluster vertices, the rest of the eyelid vertices are free (Figure 3).

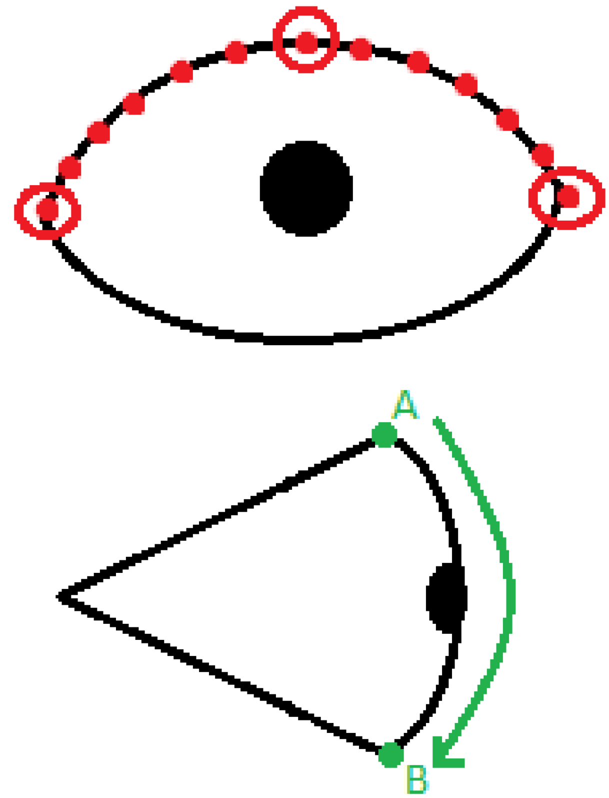

Рис. 3: Up: splitting vertices into groups: all the vertices of the upper eyelid margin are cluster vertices; the rest of the eyelid vertices are free. Bottom: the CCM transformation of the blend shape targets does not go through the eye, but follows its shape..

The cluster vertices could be transformed from an open eye position to a closed eye position with the help of manually added curve (the curve can be dependent on the eye shape). The margin eyelid cluster itself is defined as non-deformable. The central vertex and vertices at the corner of the eye are defined the main ones. The corner vertices position is fixed, they are allowed only to rotate. The central vertex travels along the curve lugging the reset of the cluster vertices.
The free vertices could be transformed by the blend shape model with the use of the two blend shape targets and they should also correspond to the new position of the cluster vertices.

5. Conclusion

The process of rigging a face is considered to be a strenuous task for an artist to fulfill as there is no standard definition for the rig and rigging itself is very labourious. The number of approaches to facial rigging is abundant, while blend shape models method proves to be the most prevalent but not without some limitations. Control Cluster Method is an approach for geometric non-linear deformation of the model in a labour-saving way and is capable of tackling the shortcomings of the blend shape model as well as of taking advantage of its benefits. The variety of non-linear ways to transform free vertices of the facial model is in the scope of the future research.

Литература

[1] Verynica Orvalho, Pedro Bastos, Frederic I. Parke, Bruno Oliveira, Xenxo Alvarez: "A Facial Rigging Survey", Eurographics (STARs) 2012, pp. 183-204.

[2] "Context Aware Human-Robot and Human-Agent Interaction". Edited by Nadia Magnenat-Thalmann, Junsong Yuan, Daniel Thalmann and Bum-Jae You. London: Springer, 2015.

[3] Alexander Bukatov, Elena Gridchina, and Dmitry Zastavnoy, "A spline-based approach to control cluster deformation," World Applied Sciences Journal, vol. 26, no. 6, pp. 724–727, 2013.

[4] Alexander Bukatov, Elena Gridchina, and Dmitry Zastavnoy, "A control cluster approach to non-linear deformation," in Poster Paper Proceedings of WSCG, Vaclav Skala, Ed., Plzen, Czech Republic, 2014, pp. 97-100.

[5] Alexander Bukatov, Elena Gridchina, and Dmitry Zastavnoy, "Using Dual Quaternions for Control Cluster Method," in Proceedings of the 24th International Conference on Computer Graphics and Vision GraphiCon'2014, GraphiCon, pp. 34–37, Rostov-on-Don, Russia, 2014.

[6] Eftychios Sifakis, Igor Neverov, and Ronald Fedkiw. 2005. "Automatic determination of facial muscle activations from sparse motion capture marker data". ACM Trans. Graph. 24, 3 (July 2005), pp. 417-425.

[7] Li Zhang, Noah Snavely, Brian Curless, and Steven M. Seitz. 2004. Spacetime faces: high resolution capture for modeling and animation. In ACM SIGGRAPH 2004 Papers (SIGGRAPH '04), Joe Marks (Ed.). ACM, New York, NY, USA, 548-558.

[8] Taehyun Rhee, Youngkyoo Hwang, James Dokyoon Kim, and Changyeong Kim. 2011. "Real-time facial animation from live video tracking". In Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '11), Stephen N. Spencer (Ed.). ACM, New York, NY, USA, 215-224.

[9] Derek Bradley, Wolfgang Heidrich, Tiberiu Popa, and Alla Sheffer. 2010. High resolution passive facial performance capture. In ACM SIGGRAPH 2010 papers (SIGGRAPH '10), Hugues Hoppe (Ed.). ACM, New York, NY, USA, pp. 41:1-41:10.

[10] J.P. Lewis, Ken Anjyo, Taehyun Rhee, Mengjie Zhang, Fred Pighin, Zhigang Deng, "Practice and theory of blendshape facial models", EUROGRAPHICS State of the Art Reports, pp. 199-218.

[11] E. Catmull and R. Rom, "A class of local interpolating splines," in Computer Aided Geometric Design, R. Barnhill and R. Riesenfeld, Eds., pp. 317–326. Academic Press, 1974.

[12] Tickoo S.: Autodesk Maya 2015: A Comprehensive Guide. CADCIM Technologies, USA, 2014.