Research Article

Facial Animation of Life-Like Avatar based on Feature Point Cluster

Samuel Gandang Gunanto[1,2] *, Mochamad Hariadi[1], Eko Mulyanto Yuniarno[1] and Matahari Bhakti Nendya[2]

[1]*Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember Surabaya – Indonesia*
[2]*Department of Animation, Institut Seni Indonesia Yogyakarta – Indonesia*

___

Abstract

The problems that are often encountered in the making of facial expressions on 3D animated avatars heavily dependent on the ability of key animators to create movement and combine them into a series of facial motion animation. This research proposes an automatic preprocessing for facial animation using feature point cluster. The mechanism of a transfer system for animated facial expressions using a radial basis function technique that is effective, especially in terms of production speed. This study focuses on the improvement of animated facial expressions using the feature point cluster. The result is to provide improved motion animation more expressive. This process will be easier and faster for the formation of facial motion animation without the need for process adjustments manually.

Keywords: facial animation, avatar, feature point cluster, retargeting, radial basis function.

___

1. Introduction

One form of the use of digital characters is avatars. Avatar acts as a personal representative of a person in the digital world. Personal representation form by using avatars can be a person, robots, superheroes, cartoon characters or animals. Avatar on the genre game Real- Time Strategy (RTS) is represented in the form of the face of the characters used. The forms of the character face are very closely linked to strengthening the presence of facial animation for characters.

Traditional approaching to the formation of animated facial expressions models rely on an artist to create a movement key and then combine them into a series of movements of facial expression. A game with the rules of the interactive game requires the presence of facial animation in their communication or interaction with the game area. Due to limited resources, in-game facial animations are often denied. But lately with the advent of gameplay system of animated films interspersed therein require the presence of animated facial expressions are there to provide entertainment aspect and the story for the player. Problematic encountered for manufacturers is the use of templates and the same facial movements in different models is a very time-consuming task for the animators [1]. These problems also occur in the manufacture of facial animation motion because the adjustment of natural movement and detail needed to make a facial expression be attractive and make the characters come alive. Hence the search for a solution that is capable of handling the issue a top priority to be tackled, especially which is able to bring a design appropriate technology in the animation industry

sector.

This study proposes a method of facial animation using feature point cluster approach. Mesh deformation by considering the feature-point cluster is expected to provide the model results avatars are able to visualize the expression without the need for a conversation or dialogue.

2. State of the Arts

2.1 Computer Facial Animation

Language expression was as developed by Hjortsjo in 1970 was one of the first attempts to systematize the mapping of facial expressions by facial muscle activity was observed from various differences in facial expressions of the actors [2]. Facial Action Coding System, or commonly known as FACS as the development of methods of expression language in 1978. FACS divides the face area to 66 Action Unit (AU) were able to reconstruct the face with various expressions naturally [3]. FACS system is still often used by animators to reconstruct facial expression by utilizing the controllers in the expression of facial animation system descriptively.

Mapping technique in the movement of facial expression in animation system can be guided based on the marker. The marker-based system is a standard that has been developed commercially or fabricated. This system is used because of the lack of computing used taking into account the speed of the process is instantaneous or real-time.

2.2 Deformation-Based Approaches

Deformation directly defined on the surface of the facial mesh that often produce the quality of animation. This ignores the facial anatomy and structure of the existing muscle. So that the focus is only on making variations of

facial expression by manipulating thin-shell mesh. This category includes morphing between different models and simulated artificial muscles in shape splines [4], mesh [5], or free-form deformations [6].

2.2.1 Morphing 2D and 3D Morphing
Beier and Neely [7] shows the 2D morphing between two images that are manually selected features associated with a line. Warp function conducted based on where the affected and surrounding areas. Realistic approach of this method requires manual interaction in the form of coloring, selection of features correspondent, setting the parameters of warp and dissolve. Variations can be applied in the viewpoint of the destination image or the selection of correspondence feature. Head movement is very difficult in the synthesis process if the feature of interest is closed. To overcome this, Pighin et al. [8] combine 2D with 3D transformations morphing geometry model. They animate the facial expression animation keys with 3D geometry interpolation when the image started morphing appropriate texture. This approach looked realism, although the animation is limited to interpolate between key facial expression has been determined in advance.

2D and 3D morphing method can produce good facial expressions and quality, but have limitations similar to interpolation approach. Selection point relating to the image of the destination is still done manually and depending on the angle of view and does not have the general form for a different face. In additional viewpoint animation also, be tied to a specific destination image only.

2.2.2 Free Form Deform
Free Form Deformation (FFD) volumetric change the object by changing the control point in three-dimensional cubic lattice [9]. In concept, a flexible object embedded in a cubic lattice-shaped 3D control grid and is composed of dots of control. Box cubic lattice of control capable of adjusting the changes done to the object so that the object is modified therein remains in the square lattice, see Fig. 1.
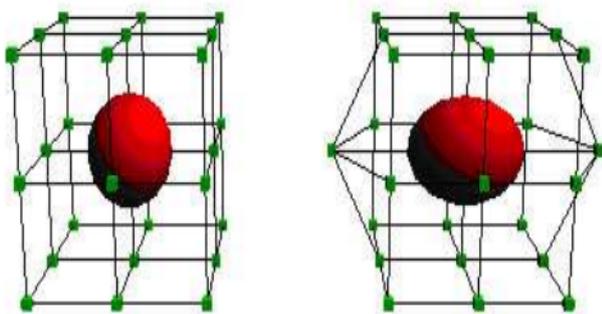

Fig 1. Free Form Deformation (FFD) [2]

2.3 Facial Motion Capture Data
Facial motion capture data used in some motion capture animated film. In motion capture obtained and cleaned by the movement of the animation. An array of good quality camera catches the data used to build a 3D marker location in the area of the face. Although the optical system is very difficult to use and expensive, provide reconstructed data timeliness and motion information. Once the data is ready, facial animation can be created by processing structure [10] muscle or blend shapes [11].
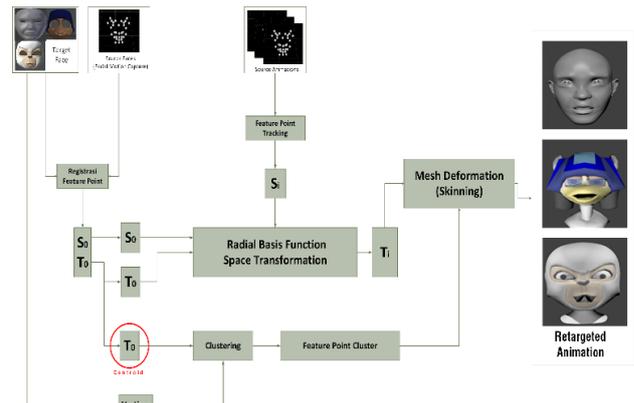
3. Experimental Design

Fig 2. Facial Animation of Life-Like Avatar based on Feature Point Cluster

3.1 Facial Rigging
Facial rigging is the process of making a control point for facial animation, this usually made by the animator. In this study, the facial rigging based on the location of facial feature points on the face mask that refers to the approach used in the FACS optitrack motion capture system.

Facial rigging process is done manually on each model of the avatar. The result of this process is a data point of the face mask features from avatar models, see Fig. 3. Prop facial feature point coordinates will be used as a centroid in the process of clustering and is used in the process of retargeting.
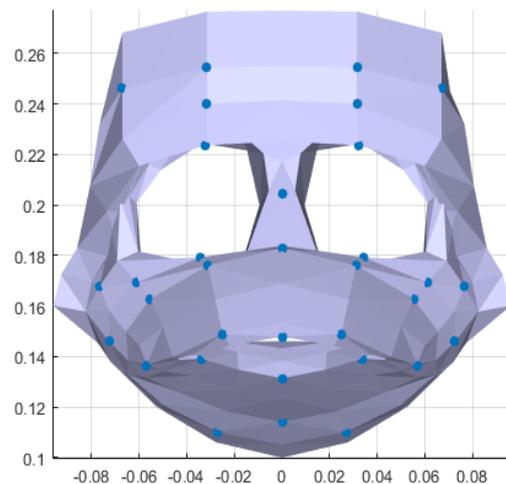

Fig. 3. *Facial Rigging* on Target Face Models

3.2 Clustering
Clustering process used in this study refers to a clustering algorithm *k*-nearest neighbor with some modification in the process of definition value of *k*. Algorithm *k*-nearest neighbor (k-NN) is a method to perform the clustering of objects based on the learning process from data that were located closest to the object. In this case, learning data is the data of vertices which are located close to the point features.

The process of modified clustering using *k*-NN to find a feature point cluster can be seen in the following flowchart, see Fig 4.
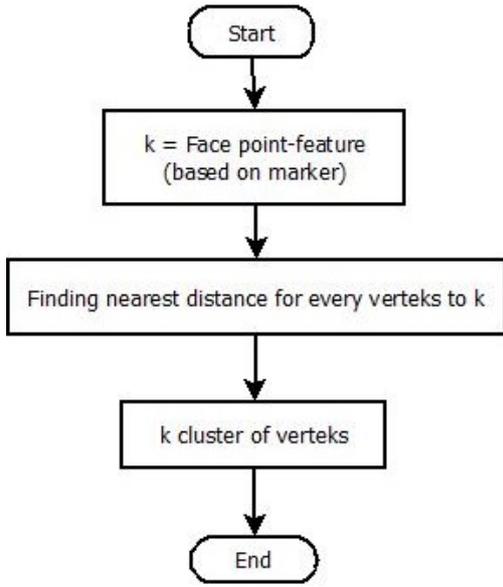
Fig. 4. Flowchart Proposed Method of Modified k-NN Based on Feature Point Cluster

A major component in the clustering algorithm is measuring the distance between the vertex and point features and the decision to make a cluster. Data in the form of the same vectors are able to use a euclidean distance matrix approach to measuring the distance between vertex and point features, see Eq. 1. The decision for clustering process based on the similarity factor between vertex to their centroid.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (1)$$

3.3 Radial Basis Function

Radial Basis Functions (RBF) is often used in computer graphics applications for estimation process and surface interpolation. In this study, RBF is used as a transformation of space. To do this first defined two rooms with two feature point set. Suppose S_0 as the set point of the source feature, T_0 as the target feature point set, and N is the size of the set. Each controls point $\vec{s_i} \in S_0$ had a relationship with $\vec{t_i} \in T_0$. After training with two pieces of the feature point set, RBF can transform the position of the source space to the target space with RBF formulation in Eq. 2.

$$F(\vec{s_j}) = \sum_{i=1}^{N} \vec{w_i} \cdot h(\|\vec{s_j} - \vec{s_i}\|) \quad (2)$$

with $h(\|\vec{s_j} - \vec{s_i}\|) = \sqrt{(\|\vec{s_j} - \vec{s_i}\|)^2 + sc_j^2}$ (multi-quadric function) and $sc_j = min_{j \neq i}(\|\vec{s_j} - \vec{s_i}\|)$.

The training network consists of the completion of 3 linear system of N (in the case of 3 dimensional) like Eq. 3.

$$\vec{t_j} = F(\vec{s_j}) \quad (3)$$

Suppose H is a matrix like $H_{ij} = h(\|\vec{s_j} - \vec{s_i}\|)$ and $T_x = (t_1^x t_2^x \dots t_N^x)^t$. So by using Eq. 3, the system can be defined as:

$$T_x = H \cdot W_x \quad (4)$$

with weights $W_x = (w_1^x w_2^x \dots w_N^x)^t$. So as to solve the system we calculate the value $W_x = H^{-1}T_x$. RBF network once trained for each axis, position in space $\vec{t}$ targets for each point $\vec{s}$ of source space is obtained by applying the transform $F(\vec{s})$ is generated.

3.4 Mesh Deformation (Skinning)

Linear blend skinning (LBS) purposed for changing the mesh according to the local transformation of the skeleton. Deformation skinning techniques adapt to the changing location of point features, the feature point position is also considered as a point of joint motion. Weights for each point defined $\sum w_i = 1$ weight to bone- i $w_i \in [0,1]$. Weighting with a value of 0 means that the feature does not affect the point vertex in the mesh point, whereas if the weight value of 1 means that the point vertex in the mesh is affected only by the feature point only. Vertex position $\vec{v_f}$ the frame f is defined by Eq. 5.

$$\vec{v_f} = \vec{v_o} + \sum_{i=1}^{N} w_i \cdot \vec{d_{if}} \quad (5)$$

Where N is the number of feature points and $\vec{d_{if}}$ is a change in the location of the i-th feature point on the frame f.

4. Result and Discussion

4.1 Feature Point Cluster

Feature point cluster formation performed by the clustering model involving all the target face vertex points that exist within the target model's face mask. Determined 33 points vertex associated with facial feature points as the center of the cluster or centroid. Calculations the similarity done by using the Euclidean distance to get the membership of each cluster. The cluster results obtained then used in the process of determining the affected area of movement.
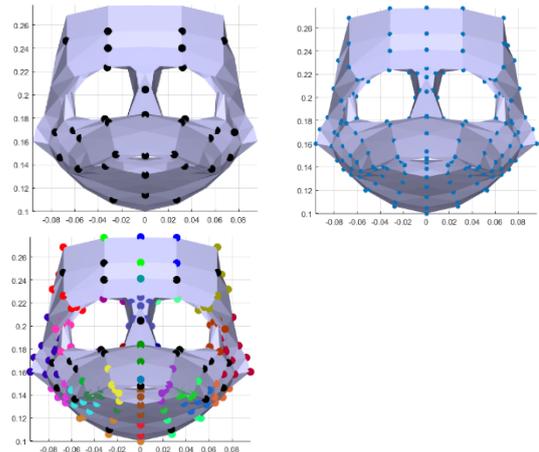

Fig. 5. Feature Point Cluster Processing.

In this study, the process of using the clustering algorithm *k*-nearest neighbor (*k*-NN), which is a method to classify the object based on the learning data that was located closest to the object, are modified using feature point, see Fig.4. The value of *k* in the *k*-NN is determined by the position of the vertex associated with facial feature points that have been determined on the facial rigging.

4.2 Radial Basis Function Space Transform

Transformation of space using radial basis function is used as a way to transform the space between the facial expressions, based on facial motion capture data, and the target face model, the avatar. Morphological forms of human faces, in general, are very varied and have differences with the morphological form of 3D characters such as cartoon characters, monsters or animals. This resulted in the movement of facial feature points that use humans as a source of animation data can't be directly used. There are several things to consider in the use of animation data such as scale and orientation.
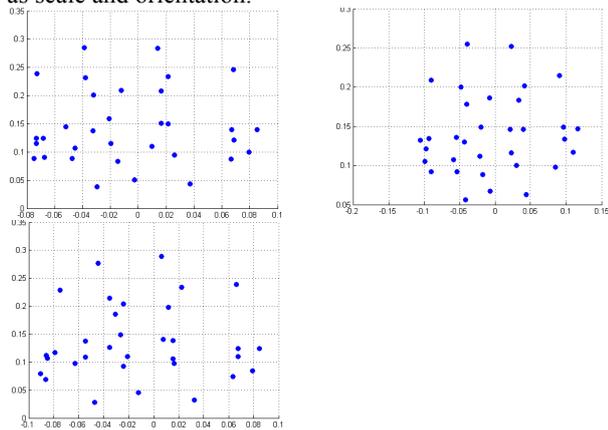

Fig. 6. Radial Basis Function Result to Avatar Face Model

To determine the feature points displacement on the face of the avatar, the feature point displacement calculation from the RBF conducted. As the discussion and observation, taken 12 points from the character models considered to represent the avatar character formation of facial expressions. Fig. 7 presents 12 feature points are used as a discussion.
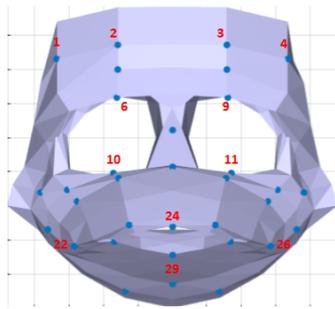

Fig. 7. Feature Point Used For Analyzed On Avatar Face Model

Fig. 7 is a position of the feature point of the avatar's face. The process of space transformation RBF provides results for each point feature that has undergone a process of transformation.

Data observed on seven frames from the data of facial motion capture data and the results of the transformation process RBF space. Facial motion capture data is used as a reference to determine the transfer point features. When the feature points of facial motion capture data are changed, then all the feature points on the avatar's face moving in accordance with the data source. The Changes of feature points from the result of the transformation of space in the avatar's face are presented in the Tab. 1.

Table 1 Results of The Displacement Feature Points on The Avatar Face

Feature Point	Frame						
	0	1	2	3	4	5	6
1	0	0.0328	0.0200	0.0221	0.0592	0.0412	0.0423
2	0	0.0427	0.0598	0.0576	0.0829	0.0458	0.0477
3	0	0.0405	0.0653	0.0627	0.0917	0.0675	0.0596
4	0	0.0043	0.0304	0.0178	0.0543	0.0651	0.0609
6	0	0.0398	0.0102	0.0108	0.0735	0.0386	0.0347
9	0	0.0276	0.0186	0.0207	0.0706	0.0546	0.0465
10	0	0.0502	0.0568	0.0500	0.0839	0.0249	0.0260
11	0	0.0333	0.0533	0.0546	0.0803	0.0536	0.0417
22	0	0.0601	0.0807	0.0676	0.0665	0.0372	0.0516
24	0	0.0716	0.0873	0.0640	0.0502	0.0257	0.0140
26	0	0.0498	0.0794	0.0751	0.0691	0.0459	0.0370
29	0	0.0799	0.0904	0.0801	0.0599	0.0353	0.0236

At frame 0, the data displacement is 0, this is because the frame 0 is the origin frame for feature point. In the next frame (1-7) the displacement of the frame, which indicates the movement of feature points describe in Fig. 8.
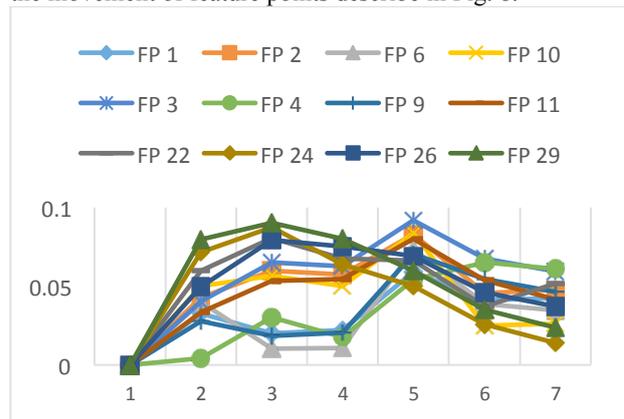

Fig. 8. Feature Points Displacement Graph of The Avatar's Face Model.

From the graph in Fig 8 is known that the displacement of facial feature points on the avatar's face is well formed. The process of moving facial feature points occurs linearly and depend on the number of mesh faces and how the morphological form of the model used. The morphological form is similar to a human and visually good recognized clearly, see Fig 9.

The level of conformity result of transformation processes in the model space RBF human face character can be calculated using the standardized approach like standard deviation of the facial motion capture data with the transformed data RBF. The result of the calculation of standard deviation for all the feature points can be seen in Tab. 2. The results of transformation of space in the face of the character models have a standard deviation of 0.0034.

Table 2 The Standard Deviation of Feature Points Displacement.

Feature Point	Standard Deviation
1	0.0177
2	0.0233
3	0.0265
4	0.0251
6	0.0231
9	0.0224
10	0.0251
11	0.0228
22	0.0247
24	0.0299
26	0.0255
29	0.0314
Average Standard Deviation	**0.0034**

4.3 Mesh Deformation

Mesh deformation process begins by determining the weight of each vertex are interconnected in a cluster member. The weighting of point refers to the displacement of the vertex when the process of movement occurs. The process of determining the point weighting effect on the formation of an animation on a mesh face is calculating using LBS. Tab. 3 are an example of the weight calculation results for each vertex.

Table 3. Example of The Weight Calculation Results for Each Vertex

No	Vertex Coordinates			Weight
	x	**y**	**z**	
1	-0.0559	0.216127	0.057465	0.638748
2	-0.05564	0.224933	0.064467	0.740125
3	-0.05564	0.224933	0.064467	0.740125
4	-0.05564	0.224933	0.064467	0.740125
5	-0.0559	0.216127	0.057465	0.638748
6	-0.0559	0.216127	0.057465	0.638748
7	-0.05564	0.224933	0.064467	0.740125
8	-0.06709	0.226293	0.078495	0.796049
9	-0.06709	0.226293	0.078495	0.796049
10	-0.06002	0.215447	0.075391	0.684178

Vertex weight calculation results in Tab. 3 was used as a reference in determining the new vertex position in accordance with the process of the formation of an expression. Fig. 9 is presented the process of mesh deformation on the avatar's face model, especially at the point features that affect the formation of animated facial expressions.
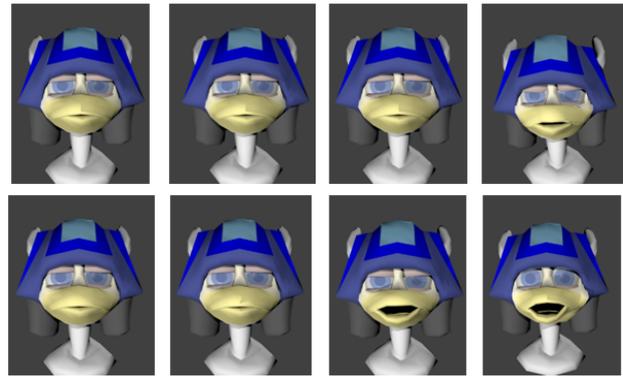

Fig. 9. Results of Facial Animation on The Avatar's Face Models

5. Conclusion And Future Works

Reuse of animation techniques (retargeting) using the transformation of the space-based Radial Basis Function (RBF) produces character expressions that follow the movements of the actors and the expression of the same approach with the actors. RBF space transformation is a technique with a linear approach that does not require a complicated computation. Mechanical merging with mesh models features point on the stages of deformation mesh provides a visual form of the transformation result. Transformation of space RBF provides solutions in the formation of facial animation in accordance with the movement of the actors and relies on the need to control the bone system, blend shape weight, and other parameters in the formation of specific facial animation.

The complexity of this research can be improved by building a transfer system for animated facial expressions in real time. The transformation process of facial expressions can also be tested with other approaches, such as using geodesic distance where considering the measurements of the facial surface.

References

1. Orvalho, V. C., Zacur, E., & Susin, A. Transferring Facial Expressions to Different Face Models. Ibero-American Symposium on Computer Graphics - SIACG. The Eurographics Association. (2006).
2. Hjortsjo, C. Man's Face and Mimic Language. Malm: Nordens Boktryckeri. (1970).
3. Ekman, P., & Friesen, W. V. Facial Action Coding System: a technique for the measurement of facial movement. Palo Alto: Consulting Psychologists Press. (1978).
4. Viad, M., & Yahia, H. Facial animation with wrinkles. Proceedings of the Third Eurographics Workshop on Animation and Simulation. (1992).
5. Singh, K., & Fiume, E. Wires: A geometric deformation technique. SIGGRAPH Proceedings. pp. 405–414. (1998).
6. Kalra, P., Mangili, A., Thalmann, N., & Thalmann, D. Simulation of facial muscle actions based on rational free from deformations. Eurographics, vol **11**. pp. 59–69. (1992).
7. Beier, T., & Neely, S. Feature-based image metamorphosis. SIGGRAPH Proceedings. ACM Press. pp 35-42. (1992).
8. Pighin, F., Auslander, J., Lischinski, D., Salesin, D., & Szeliski, R. Realistic facial animation using image-based 3d morphing. (1997).
9. Sederberg, T., & Parry, S. Free-form deformation of solid geometry models. Computer Graphics, SIGGRAPH, volume **20**. pp 151–160. (1996).
10. Choe, B., Lee, H., & Ko, H. (2001). Performance driven muscle-based facial animation. The Journal of Visualization and Computer Animation, **12(2)**. pp 67-79. (2001).
11. Deng, Z., Chiang, P., Fox, P., & Neumann, U. Animating blend shape faces by cross-mapping motion capture data. Proceedings of the 2006 symposium on Interactive 3D graphics and games. pp 43-48. (2006)