

A Multiresolutional Approach for Facial Motion Retargetting Using Subdivision Wavelets

Kyungha Min and Moon-Ryul Jung*

Dept. of Media Technology,
Graduate School of Media Communications, Sogang Univ., Seoul, Korea
{minkh, moon}@sogang.ac.kr

Abstract. We propose a new facial motion retargetting method using wavelet-based multiresolutional analysis of triangular meshes. We define the displacement of the source model as the difference between the neutral and the current expressions of the source model, and decompose it to the displacement of the base mesh and the “displacements” of the wavelet coefficients. We compute the wavelet coefficients of a mesh by using the wavelet basis functions derived from a mesh subdivision scheme. To get the motion of the target model, we add the displacement of the source model to the target model level by level starting from the base mesh. The wavelet-based representation of mesh facilitates the comparison between two expressions of the same model and the comparison between the expressions of the source and target models.

1 Introduction

Digital facial animation is one of the most popular research areas in computer graphics. The development of entertainment markets such as computer games, animations and movies accelerates the progress of digital facial animation techniques to build an efficient framework of realistic facial animation. In human body animation, motion retargetting, which creates new motions of the target model from the motion data of the source model, has introduced an innovative tool for creating realistic character animation [3]. The major difficulty of applying motion retargetting to facial models is that faces do not have the obvious control parameters to represent and control facial motions such as joint angles for body motion [7].

The problem of motion retargetting can be formulated as the problem of obtaining the current expression of the target model by adding the scaled version of the displacement of the source model to the neutral expression of the target model (See Fig. 1). To handle the problem of facial motion retargetting, we propose a multiresolution approach. To materialize this approach, we present two key ideas.

* This research has been supported by the fund (R01-2002-000-00311-01) from KOSEF (Korea Science and Engineering Foundation).

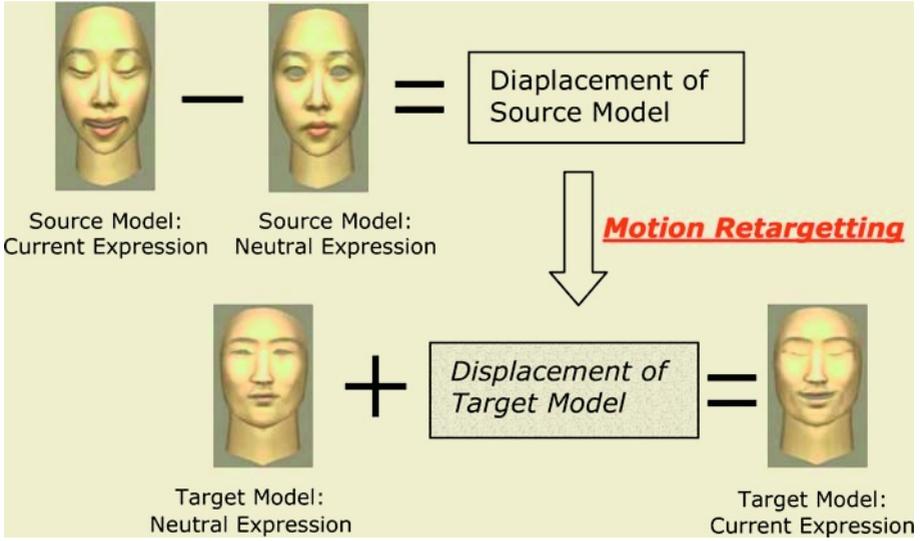


Fig. 1. The process of motion retargetting: Add the scaled version of the displacement of the source model to the neutral expression of the target model to get the current expression of the target model

The first key idea is that we retarget the displacement of the base mesh of the source model by using **the displacement parameters with respect to the local coordinate system defined by the normal vector and the principal directions on the base mesh**. Since the principal directions are good at representing the local geometry and orthogonal to each other [1], we use the normal vector and the two principal directions to define the local coordinate systems. It contrasts to the approach of Noh and Nuemann [6], where the source model is deformed to the target model to define the local coordinate systems for the target model, which correspond to those on the source model.

The second key idea is to represent the facial mesh as **a multi-resolution mesh and use the wavelet coefficients as the “control parameters” of the facial motion**. To represent the source facial model faithfully, we decompose the model into the base model and wavelet coefficients that store the details of the model using a multiresolution scheme called “subdivision wavelets” [4]. We decompose the displacement of the source model into the displacement of the base model and those of the wavelet coefficients. We then retarget the multi-level displacements to the target model.

2 Previous Work

Retargetting motions to new bodies were first developed by Gleicher[3] to make captured motions reusable. This idea was also applied to facial motion, and

there are some work[5, 6, 8, 10] in facial motion retargetting. Noh and Neumann [6] compute the displacement of each vertex on the source mesh at each frame relative to the neutral frame, and add the scaled version of the displacements to the neutral frame of the target mesh. If there are frames whose displacements from the neutral frame are large, the retargetted version of the displacements tend to intersect each other. Noh and Neumann’s method has to apply filtering to the retargetted displacements to remove the intersections.

Pyun and Shin [8] proposed an example-based retargetting scheme. In this approach, the system uses a set of correspondences between example poses of the source and target facial models provided by the animator. This method produces more accurate results than that of Noh and Neumann, because of sufficient mapping data between source and target expressions. However, the more detailed the facial model is, the larger number of example expressions are needed.

Na and Jung[5] represent each frame of the source mesh by the base mesh and a sequence of normal offsets from the base mesh. The displacement of the source base mesh is retargetted by using an example-based approach of [8]. The fine motion of the target model is obtained by adding the normal offset displacements of the source mesh to the normal offsets of the target mesh at the neutral frame, level by level.

Wang et al. [10] learn a decomposable generative model from a set of facial motions of different people. A given motion of the source face is first decomposed into the content (e.g. smile) and the style parameters (a particular style of smile). Then the content motion of the source is transformed to the target motion by applying the scale factor between the source and target faces and the style parameters to the decomposable generative model.

3 Multiresolution Representation of Meshes

3.1 Construction of the Facial Models

The facial models used in this paper are triangular meshes with subdivision connectivity created by an experienced designer. Given a mesh without subdivision connectivity, we can remesh it so that it has subdivision connectivity by using the method proposed by Eck et al. [2].

3.2 Mesh Representation Using Subdivision Wavelets

For our experimentation we apply multi-resolution analysis to a given mesh three times to obtain the wavelet coefficients for three levels of the mesh. More levels of wavelet coefficients can be used to better approximate given meshes.

To represent a function using wavelet coefficients, we need to have appropriate basis functions, that is, scaling functions and wavelet functions. As Lounsbery and his colleagues [4] have shown that appropriate scaling and wavelet functions can be constructed for a mesh created by means of a given subdivision scheme, we construct appropriate scaling and wavelet functions to represent the mesh.

The wavelet functions constructed this way are called "subdivision wavelets". In this paper, we use the Butterfly subdivision scheme, an interpolation subdivision scheme, since it facilitates the computation of the analysis and synthesis filters. The analysis filters A^j, B^j and synthesis filters P^j, Q^j are determined by the subdivision scheme employed and the scaling functions derived from the subdivision scheme.

As shown in Figure 3.2, analysis filter A^j filters a given mesh to the next lower resolution. Analysis filter B^j computes \mathbf{v}^j , the wavelet coefficients, which is the difference between the consecutive resolutions.

The filters A^j, B^j are represented as matrices, as follows:

$$\mathbf{V}^j = A^j \mathbf{V}^{j+1} \qquad \mathbf{v}^j = B^j \mathbf{V}^{j+1}.$$

Given \mathbf{V}^j and \mathbf{v}^j , \mathbf{V}^{j+1} is reconstructed by applying the synthesis filters P^j and Q^j as follows:

$$\mathbf{V}^{j+1} = P^j \mathbf{V}^j + Q^j \mathbf{v}^j$$

In our subdivision scheme, P^j and Q^j are computed as follows:

$$(P^j \quad Q^j) = \begin{pmatrix} A^j \\ B^j \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{I}^j & -\alpha^j \\ \mathbf{N}^j & \mathbf{I} - \mathbf{N}^j \alpha^j \end{pmatrix} \quad (1)$$

Here \mathbf{N}^j represents the portion of the subdivision matrix which weight the "new vertices" introduced by each application of subdivision. \mathbf{I}^j is the portion of the subdivision matrix which weight the "old vertices", which is the identity matrix, because the Butterfly subdivision is an interpolation scheme. The matrix α^j are used to define the wavelet functions in terms of the scaling functions, and are computed by solving a linear system involving the inner products of the scaling functions. See Lounsbery and et al. [4] for details.

Given the base mesh \mathbf{V}^0 , a series of wavelet coefficients \mathbf{v}^j , and the synthesis filters P^j, Q^j , the original mesh \mathbf{V}^n is reconstructed by the following formula:

$$\mathbf{V}^n = \left(\prod_{i=0}^{n-1} P^i \right) \mathbf{V}^0 + \sum_{j=0}^{n-1} \left(\prod_{k=j+1}^{n-1} P^k \right) Q^j \mathbf{v}^j \quad (2)$$

Let the neutral expression and the current expression of the target model be \mathbf{T}_0 and \mathbf{T}_j . They have the following representations:

$$\mathbf{T}_0 = (\mathbf{T}_0^0, \mathbf{t}_0^0, \mathbf{t}_0^1, \mathbf{t}_0^2) \qquad \mathbf{T}_j = (\mathbf{T}_j^0, \mathbf{t}_j^0, \mathbf{t}_j^1, \mathbf{t}_j^2)$$

The motion retargetting is achieved by adding the displacement between \mathbf{S}_0 and \mathbf{S}_j to \mathbf{T}_0 in order to obtain \mathbf{T}_j .

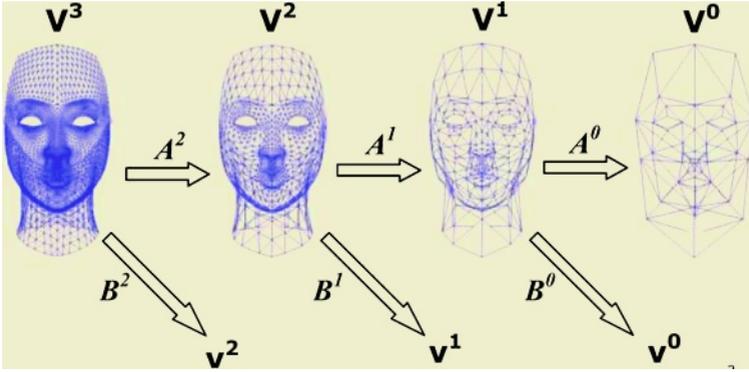


Fig. 2. Multiresolutional analysis of a facial model: The original model V^3 is filtered down to the base mesh V^0 , and the wavelet coefficients, v^j , are computed in the filtering process

3.3 The Motion Retargetting Procedure

The multiresolutional motion retargetting is achieved in the following steps.

Step 1. Compute the displacement M_j of the source model between the neutral expression and the current expression, as follows:

$$M_j = (M_j^0, m_j^0, m_j^1, m_j^2), \text{ where}$$

$$M_j^0 \equiv S_j^0 - S_0^0, \text{ and } m_j^i \equiv s_j^i - s_0^i, \text{ for } 0 \leq i \leq 2. \tag{3}$$

The displacement of the base mesh M_j^0 is defined as the difference of the current expression S_j^0 from the neutral expression S_0^0 , as follows:

$$M_j^0 (\equiv S_j^0 - S_0^0) = \{ \mathbf{d}_i \mid \mathbf{d}_i = \mathbf{p}_j^i - \mathbf{p}_0^i, \text{ where } \mathbf{p}_j^i \in S_j^0 \text{ and } \mathbf{p}_0^i \in S_0^0 \},$$

where \mathbf{p}_0^i and \mathbf{p}_j^i are the corresponding vertices between the base mesh of the neutral expression and the base mesh of the current expression.

As shown in Figure 3, the displacement vector \mathbf{d}_i is represented as follows:

$$\mathbf{d}_i = \alpha_i \mathbf{N}_i + \beta_i \mathbf{J}_i + \gamma_i \mathbf{K}_i \tag{4}$$

Here \mathbf{N}_i is the normal vector, and \mathbf{J}_i and \mathbf{K}_i are the principal directions at the location \mathbf{p}_0^i . The principal directions of a triangular mesh are computed by the algorithm in [9].

As shown in Eqn 3, the displacement of upper levels of the mesh, m_j^i , is defined to be the difference of the wavelet coefficients between the neutral expression and the target expression. The wavelet coefficients can be compared and thus subtracted because the wavelet basis functions on the neutral expression and the target expression correspond to each other.

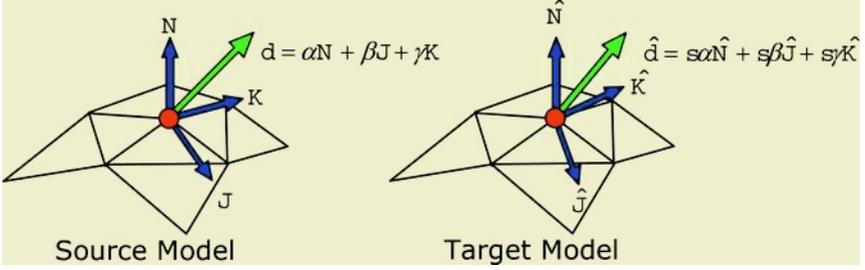


Fig. 3. Retargetting the displacement \mathbf{d}_i on the source base mesh to the displacement $\hat{\mathbf{d}}_i$ on the target base mesh

Step 2. Transform the displacement of the source model, $(M_j^0, m_j^0, m_j^1, m_j^2)$ to the displacement of the target model, $(\hat{M}_j^0, \hat{m}_j^0, \hat{m}_j^1, \hat{m}_j^2)$.

Each motion vector $\hat{\mathbf{d}}_i$ in \hat{M}_j^0 is computed as follows:

$$\hat{\mathbf{d}}_i = s\alpha_i\hat{\mathbf{N}}_i + s\beta_i\hat{\mathbf{J}}_i + s\gamma_i\hat{\mathbf{K}}_i \quad (5)$$

as shown in the right of Figure 3. Here $(\hat{\mathbf{N}}_i, \hat{\mathbf{J}}_i, \hat{\mathbf{K}}_i)$ is the target local coordinate system defined at the corresponding target vertex. Here s is the scale factor, which is defined to the ratio between the surface area of the source base mesh and that of the target mesh. We use $\hat{m}_j^i = m_j^i$. In other words, we use the displacements of the wavelet coefficients on the source model as the displacements of the wavelet coefficients on the target model, for the following reasons. The wavelet functions for each level of the target model have the same forms as those of the source model. Though the supports of the wavelet functions are not the same between the source model and the target model, these supports correspond to each other between the source and target models, starting from the base mesh.

Step 3. Obtain the current expression of the target model $T_j = (T_j^0, t_j^0, t_j^1, t_j^2)$ by adding the displacement of the target model computed in **Step 2** to the neutral expression $T_0 = (T_0^0, t_0^0, t_0^1, t_0^2)$. That is, compute:

$$T_j = (T_j^0, t_j^0, t_j^1, t_j^2), \text{ where}$$

$$T_j^0 = T_0^0 + \hat{M}_j^0, \text{ and } t_j^i = t_0^i + \hat{m}_j^i, \text{ for } 0 \leq i \leq 2.$$

Step 4. Finally, reconstruct the mesh corresponding to T_j by using Eqn 2:

$$\text{mesh}(T_j) = P^2P^1P^0T_j^0 + P^2P^1Q^0t_j^0 + P^2Q^1t_j^1 + Q^2t_j^2.$$

4 Implementation and Results

We implemented the proposed algorithm in a personal computer with 2.0 GHz Pentium-4 CPU and 512 MByte Main Memory. Each of the facial models has

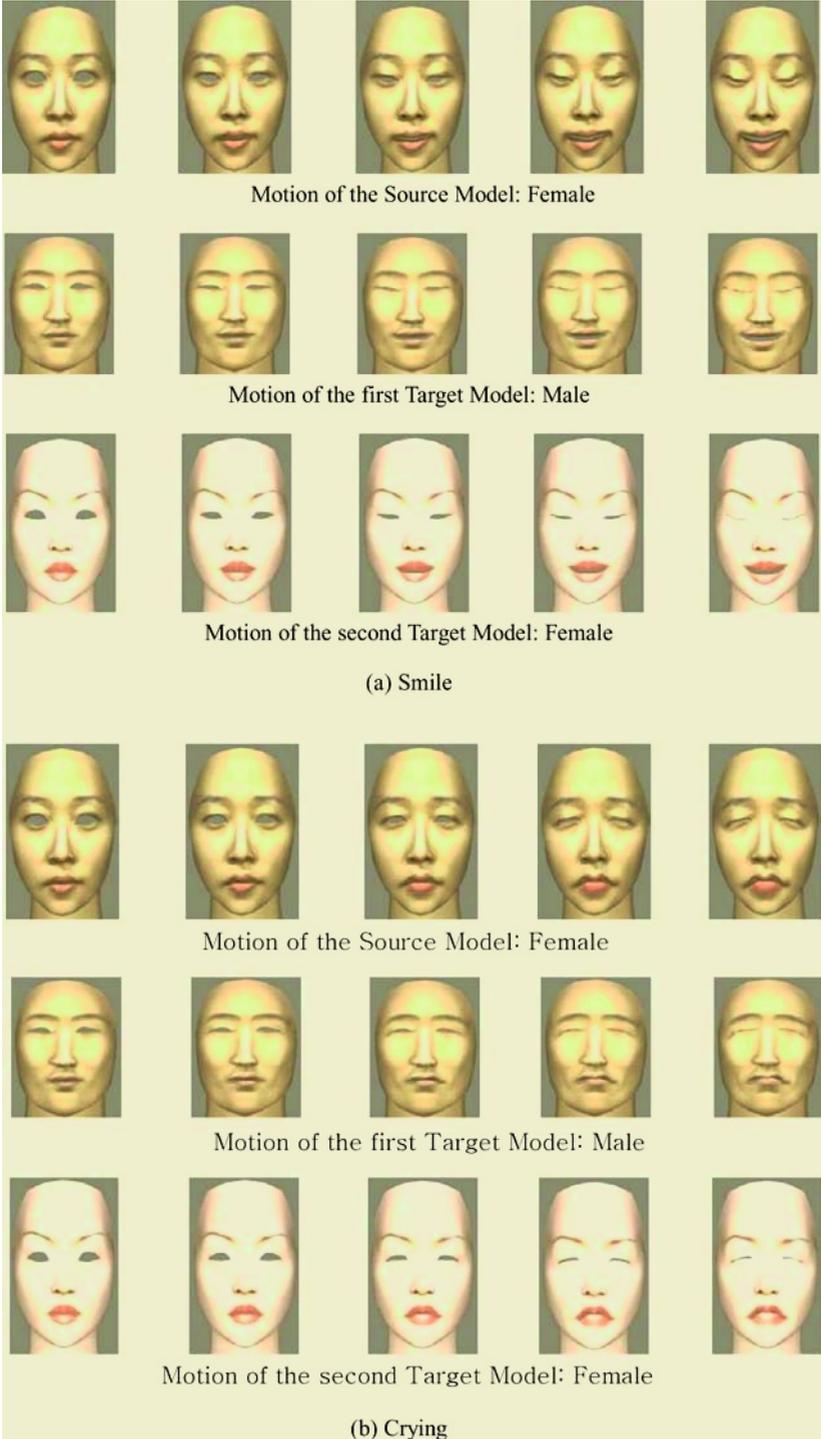


Fig. 4. Retargetting of the facial motions

2964 vertices and 5760 triangles. We tested the proposed algorithm for one source model and two target models. The source model is created from one female volunteer, and two target models are created from another female volunteer and one male volunteer. The motions we tested in this paper are smiling (See Fig. 4 (a)) and crying (See Fig. 4 (b)).

5 Conclusion

In this paper, we presented a new facial motion retargetting scheme based on the multiresolutional analysis of polygonal meshes. The displacement between the neutral and the current expressions are extracted from the base mesh of the source model and as well as from the wavelet coefficients. The algorithm creates the motion of the target model by adding the multiresolution displacements of the source model to the target base mesh and the wavelet coefficients of the target model at the neutral frame, level by level. In this process, the wavelet coefficients play the role of control parameters to represent and control facial motion. Hence we solved one of the inherent difficulties in facial motion retargetting, which is the lack of obvious control parameters to control facial motion.

References

1. do Carmo, M. P., *Differential Geometry of Curves and Surfaces*, Pearson Education, 1976.
2. Eck, M., DeRose, T., Duchamp, T., Hoppe, H., Lounsbery, M., and Stuetzle, W., "Multiresolution Analysis of Arbitrary Meshes", *Proceedings of SIGGRAPH 1998 Annual Conference Series*, pp. 173–182, 1995.
3. Gleicher, M., "Retargetting motion to new characters," *Proceedings of SIGGRAPH 1998 Annual Conference Series*, pp. 33-42, 1998.
4. Lounsbery, M., DeRose, T. D., and Warren, J., "Multiresolution analysis for surfaces of arbitrary topological type," *ACM Transactions on Graphics*, Vol. 16, No. 1, pp. 34-73, 1997.
5. Na, K. and Jung, M., "Hierarchical Retargetting of Fine Facial Motions," *Proceedings of Eurographics 2004*, pp. 687-695, 2004.
6. Noh, J., Neumann, U., "Expression cloning," *Proceedings of SIGGRAPH 2001 Annual Conference Series*, pp. 277-288, 2001.
7. Parke, F. I., Waters, W., *Computer Facial Animation*, AK Peters, 1996.
8. Pyun, H. and Shin, S. Y., "An example-based approach for facial expression cloning," In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (2003)*, PP. 167-176.
9. Taubin, G., "Estimating the Tensor of Curvature of a Surface from a Polyhedral Approximation," *Proceedings of IEEE ICCV 1995*, pp. 852 - 857, 1995.
10. Wang, Y., Huang, X., Lee, C., Zhang, S., and Li, Z., "High Resolution Acquisition, Learning and Transfer of Dynamic 3D Facial Expressions," *Proceedings of Eurographics 2004*, pp. 677-686, 2004.