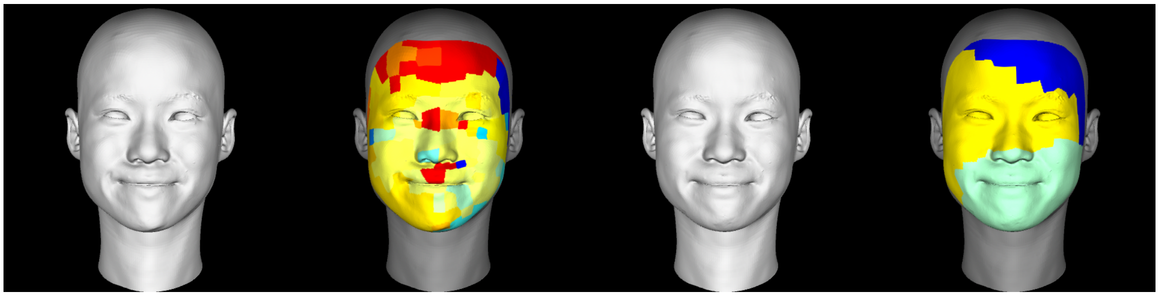
Optimized Local Blendshape Mapping for Facial Motion Retargeting

Wan-Chun Ma Graham Fyffe Paul Debevec

USC Institute for Creative Technologies



(a) (b) (c) (d)

Figure 1: ***Optimized Local Blendshape Mapping.*** *(a) Retargeted result from the previous technique. (b) A visualization of the pose sets that are used in (a). Each color represents a different pose set. (c) Retargeted result from the proposed technique. (d) A visualization of the pose sets that are used in (c).*

1 Concept

One of the popular methods for facial motion retargeting is local blendshape mapping [Pighin and Lewis 2006], where each local facial region is controlled by a tracked feature (for example, a vertex in motion capture data). To map a target motion input onto blendshapes, a pose set is chosen for each facial region with minimal retargeting error. However, since the best pose set for each region is chosen independently, the solution likely has unorganized pose sets across the face regions, as shown in Figure 1(b). Therefore, even though every pose set matches the local features, the retargeting result is not guaranteed to be spatially smooth. In addition, previous methods ignored temporal coherence which is key for jitter-free results.

In order to deal with these problems, we consider the facial motion retargeting algorithm as an optimization problem which takes the following criteria into consideration:

1. The number of poses that are (fully or partially) used to represent the current shape should be minimal.
2. The pose sets of all tracked features should vary smoothly across both spatial and temporal domains.
3. The retargeting error should be as small as possible.

2 Optimization

We formulate the criteria as the following cost function, which we solve using belief propagation [Yedidia et al. 2003]:

$$\min_{T} \mathcal{F}(T) = \sum_{i \in V} \mathcal{R}_i(t_i) + k_t S(t_i, t_i') + k_s \sum_{\{i,j\} \in E} \mathcal{S}(t_i, t_j).$$

V and E are the set of vertices and edges respectively. $T = \{t_i | 1 \leq i \leq n_v\}$ is a configuration of facial deformation; t_i describes a pose set that is associated with a vertex v_i. k_t and k_s are weighting factor of temporal and spatial smoothness. We assume that local deformation around a vertex can be represented by a small number of blendshape poses. Therefore, we define each pose set as a vector of three pose indices $t_i = (p_0^i, p_1^i, p_2^i)$. Given a pose set, each local patch can be approximated as a convex linear combination of corresponding local patches $\hat{\rho}_i(t_i)$ from the three poses in the set. We set p_0^i be always neutral pose, and therefore there are $C_2^{n_p-1}$ possible pose sets where n_p is the number of blendshape poses. We then regularize the pose sets using the following terms.

The retargeting error term: $\mathcal{R}_i$ is the sum of distances between all local patch vertices of the target shape ρ_i around v_i and the reconstructed shape $\hat{\rho}_i(t_i)$ based on a pose set t_i:

$$\mathcal{R}_i(t_i) = \| \rho_i - \hat{\rho}_i(t_i) \|^2$$

We compute the convex linear combination weights which minimizes $\mathcal{R}_i$ by quadratic programming. We use the resulting weights later on to blend the final result.

The smoothness terms: $\mathcal{S}(t_i, t_i')$ describes the temporal cost of assigning a pose set t_i to v_i. It is the Hamming distance between the current pose set t_i and the pose set t_i' computed in the previous frame. Similarly, the spatial smoothness term $\mathcal{S}(t_i, t_j)$ is the Hamming distance between the two pose sets t_i and t_j of adjacent features v_i and v_j.

3 Conclusion

Our technique is able to deliver better visual quality over traditional local blendshape mapping methods. It computes facial motion retargeting for blendshapes without any prior knowledge of facial segmentation, which is required for most of the blendshape retargeting methods. The number of blendshape poses that are used for rendering a frame is reduced hence computing resources is saved.

References

Pighin, F., and Lewis, J. P. 2006. Facial motion retargeting. In *SIGGRAPH Courses*.

Yedidia, J., Freeman, W., and Weiss, Y. 2003. *Understanding Belief Propagation and Its Generalizations*. ch. 8, 239–236.

SIGGRAPH 2011, Vancouver, British Columbia, Canada, August 7 – 11, 2011.
ISBN 978-1-4503-0921-9/11/0008