A facial tracking and transfer method with a key point refinement

Yasuhiro Akagi* Kagoshima Univ.
Ryo Furukawa Hiroshima City Univ.
Ryusuke Sagawa AIST
Koichi Ogawara Wakayama Univ.
Hiroshi Kawasaki† Kagoshima Univ.

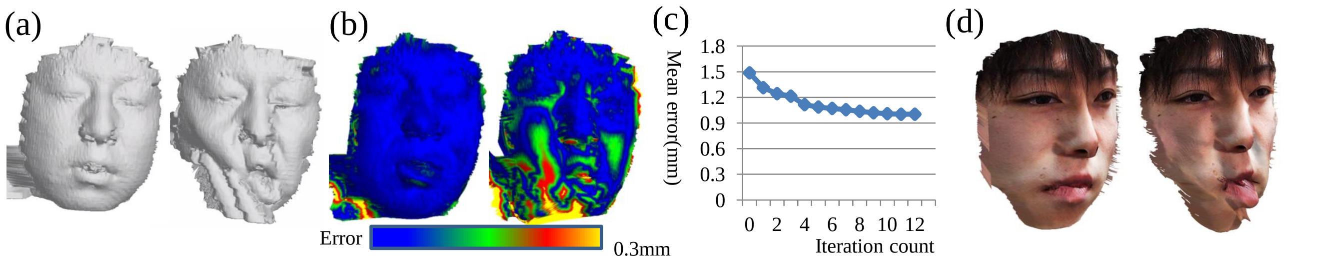

Figure 1: *(a)Input facial motion (slapping). (b)Visualization of the errors between the original face and the tracked face. (c)Transition of the mean errors in the minimization process. (d)Initial (left) and transferred face (right) from the slapping motion.*

Abstract

We propose a key point detection and refinement method for a facial motion tracking and transfer. Since key-point-based approaches for representing facial expressions usually deform faces by interpolating movements of the key points, the approaches cause errors between the deformed face and the original one. To solve this problem, our method tracks non-rigid deformations of surfaces to detects additional key points for minimizing the errors.

Keywords: Facial animation, Motion tracking, Facial transfer, Key point detection

1 Introduction

By the development of scanning and tracking methods of a facial motion, a face model consists of more than 100,000 points. On the other hand, key point based approaches such as "face rigging" are still useful ways to create and edit facial animations. Even if a facial tracking method gives us all point-to-point correspondences, an suitable set of key points is needed for such kind of applications. Therefore, the set of key points which efficiently represents the motion of a face is needed. Moreover, to create a motion of "an artificial face" such as animal characters, virtual humans and etc. in CG animations, facial transfer(cloning) methods are needed. We think a detection method of the suitable set of key points between a face to another face can be a solution for the facial transfer.

2 Facial tracking

The key idea of our facial tracking method is that it detects the deformation $D^t(F^t) \simeq F^0$ and the inverse deformation $D^{t-1}(F^0) \simeq F^t$ (F^t is a shape of tth frame in a facial motion). We utilize a deformation method based on a Radial Basis Function [Sumner et al. 2007]. If the difference between $D^t(F^t)$ and F^0 can be minimized, the deformation D^{t-1} represents the deformations from the F^0 to the closest shape of F^t. The detection process of the key points is as follows: (1)Calculating a temporal D^{t-1} by using temporal key points given by a landmark detection method[Cao et al. 2012]. (2)Sampling the candidates of key points randomly form F^0. (3)To evaluate the importance of the candidates, our method calculates a correlation coefficient between a key point and a motion of a neighboring area in $D^{t-1}(F^0)$ as a degree of representation of the motion. If the value of the correlation coefficient of a point is below the average value of all sampled points, the point is rejected from the candidates. (4)Even though the deformation D^{t-1} contains errors, the each position of the key points in each frame is close to its true position. Then, our method detects the true positions of the key points by using the Non-Rigid Registration method proposed by Jian et. al.[Jian and Vemuri 2011]. (5)Returning to the process (2) till the ratio of error decreases to less than 1%. By using this recursive process, it can minimize the error between $D^{t-1}(F^0)$ and F^t.

3 Facial transfer

The key idea of a facial transfer method is how to interpolate the difference between a source face and a target one. Our facial tracking method deforms 3D space around a face based on the transitions of key points. This method can be a solution to represent a transfer between two faces. The process of a proposed facial transfer method is as follows: (1)Creating deformations $D^t_{fit}(F^t) \simeq T^0$ from a source motion of a face F^t to a target one T^0. (2)Using the key point sampler described in section 2 to define correspondences between $D^t_{fit}(F^t)$ and T^0. (3)By using the distance from the source face to the deformed source face $Dif^t = ||D^t_{fit}(F^t) - F^t||$, calculating a deformation for facial transfer $D^t_{trns}(T^0) \simeq D^t_{fit}(F^t) + Dif^t$.

4 Results

Figure (b) and (c) shows the error between the input motion and tracking results. In these results, the mean error between the input faces and tracked faces is less than $1.0mm$ in the facial motion. However, the result (b) shows that the errors are still remaining around the area where the hand touches the face. Figure 1 (d) and (e) show the transfered shapes of the face. These results show that the movements of the mouth and cheeks can be transferred.

References

CAO, X., WEI, Y., WEN, F., AND SUN, J. 2012. Face alignment by explicit shape regression. In *CVPR*, IEEE, 2887–2894.

JIAN, B., AND VEMURI, B. C. 2011. Robust point set registration using gaussian mixture models. *IEEE Trans. Pattern Anal. Mach. Intell. 33*, 8 (aug), 1633–1645.

SUMNER, R. W., SCHMID, J., AND PAULY, M. 2007. Embedded deformation for shape manipulation. In *ACM SIGGRAPH 2007 papers*, ACM, SIGGRAPH '07.

*e-mail:akagi@ibe.kagoshima-u.ac.jp
†e-mail:kawasaki@ibe.kagoshima-u.ac.jp

SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.

ACM 978-1-4503-2261-4/13/07