

# A facial motion tracking and transfer method based on a key point detection

Yasuhiro AKAGI  
Kagoshima University  
Korimoto, 1-21-24  
Kagoshima  
890-8580, JAPAN  
akagi@ibe.kagoshima-u.ac.jp

Ryo FURUKAWA  
Hiroshima City University  
3-4-1, OzukaHigashi  
AsaMinami-Ku, Hiroshima  
731-3194, JAPAN  
ryo-f@hiroshima-cu.ac.jp

Ryusuke SAGAWA  
AIST  
1-1-1 Higashi  
Tsukuba, Ibaraki  
305-8561 JAPAN  
ryusuke.sagawa@aist.go.jp

Koichi OGAWARA  
Wakayama University  
Sakaedani 930  
Wakayama-city  
640-8510, JAPAN  
ogawara@sys.wakayama-u.ac.jp

Hiroshi KAWASAKI  
Kagoshima University  
Korimoto 1-21-24  
Kagoshima  
890-8580, JAPAN.  
kawasaki@ibe.kagoshima-u.ac.jp

## ABSTRACT

Facial animation is one of the most important contents in 3D CG animations. By the development of scanning and tracking methods of a facial motion, a face model which consists of more than 100,000 points can be used for the animations. To edit the facial animations, key point based approaches such as "face rigging" are still useful ways. Even if a facial tracking method gives us all point-to-point correspondences, a detection method of a suitable set of key points is needed for content creators. Then, we propose a method to detect the key points which efficiently represent motions of a face. We optimize the key points for a Radial Basis Function (RBF) based 3D deformation technique. The RBF based deformation is a common technique to represent a movement of 3D objects in CG animations. Since the key point based approaches usually deform objects by interpolating movements of the key points, these approaches cause errors between the deformed shapes and the original ones. To minimize the errors, we propose a method which automatically inserts additional key points by detecting the area where the error is larger than the surrounding area. Finally, by utilizing the suitable set of key points, the proposed method creates a motion of a face which are transferred from another motion of a face.

## Keywords

Facial Expression, Motion Transfer, Tracking, Non-rigid, Deformation

## 1 INTRODUCTION

A facial animation is one of the important topics in the area of computer vision and graphics[18, 15, 20, 21]. It is possible to obtain dense and accurate 3D points from an object with the development of 3D scanning method. In case of scanning a moving object, it is an important topic that how to detect a movement of a point from a frame to another frame. This kind of information is required for recognizing a facial expression and creating CG animations from the scanned point cloud.

Then, a number of methods have been proposed to detect the correspondences between frames. The techniques which track artificial markers on a target face are the well-known approaches to capture the motion of a face[3, 9]. In other cases, if it is difficult to place the artificial markers on a face, a method which detects a key point of a face can be a solution. There are two types of approaches to detect the key points: geometry based approaches and image based approaches. One of the solutions based on the geometry of a face is a non-rigid registration algorithm. The advantage of this approach is that it can deal with a deformable surface. If both a 3D shape and an image can be captured at the same time, the correspondence of points between frames can be found by using the image processing techniques. Moreover, by utilizing the key point of a face, methods called "facial transfer"[8, 11, 13, 7, 12] are proposed. These methods create a motion of an arbitrary face that reflects (copied from) motion of another face. To trans-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

for the facial motion, correspondences between a source face to a target one are required. Thus, the methods which can detect suitable pairs of key points between two faces are required.

## 2 RELATED WORK

Since it is difficult to create the facial animations manually, a number of methods are proposed to capture the facial motion. Two approaches have been researched to capture the facial motion: marker-based approaches and marker-less approaches. The marker-based approaches are more robust to detect motions than the marker-less approaches.

### 2.1 Marker-based approach of facial tracking

Huang *et al.* [9] proposed the method to capture a motion with high-fidelity by using one hundred markers. This method can capture dynamic wrinkles and fine-scale facial details. Bickel *et al.* [3] directly paint color point on a face to robustly detect the same points beyond frames. The marker-based approaches have a common problem that it is laborious to put the markers and it is difficult to capture the natural texture with the motion at the same time.

### 2.2 Marker-less approach of facial tracking

On the other hand, the marker-less approaches are proposed. Valgaerts *et al.* [18] proposed dense tracking method of movements of a face. This method uses a stereo-camera system and tracks the movement of each pixel of the camera images by using an optical-flow detection method. This method deforms a 3D face based on the optical-flow. Bradley *et al.* [4] proposed a facial tracking method which tracks both the movement of a texture and that of a geometry at the same time under the constant light condition. They also detect a shape of a mouth by using a facial parts recognition method. Sibbing *et al.* [15] uses the feature tracker like the KLT tracker to detect the motion of a face. Weise *et al.* [20] constructs the facial performance database of a person to detect the motion of the face from 2D image and 3D point set. Then, some approaches use the image processing methods to find the key points of a face such as a pupil of the eye, the outline of a lip, the tip of a nose and etc. The accuracy of these kind of methods is over 95% in some recent researches[5, 2].

### 2.3 3D shape based approach of facial tracking

The non-rigid registration algorithm is one of the useful methods to detect the motion of the point sets[17].

Jian *et al.* [10] proposed this kind of approach by using the L2 distance between Gaussian mixtures representing two point sets. However, since the features of the motion of a face vary in each facial part, it complicates the registration problem. If we can find a fine initial guess of the motion, the non-rigid registration algorithm will be a useful way to solve the problem.

### 2.4 Facial transfer

To create a motion of "an artificial face" such as animal characters, virtual humans and etc. in an animation, facial transfer(cloning) methods which copy the motion from a person to another virtual face are proposed[8, 11, 13, 7, 12, 14]. Huang *et al.* [8]utilizes a key point based deformation for a facial transfer. To represent motions of a face, a set of key points called Active Appearance Models(AAMs) are used. This method minimizes the deformations of AAMs to fit to a target face. Vieira *et al.* [19] proposed the facial transfer method which defines a zone of influence and a weight map for interpolating the movement of key points. Cosker *et al.* [6] generate a map for representing facial expressions (movements of the key points) based on Downhill Simplex Minimisation tracker. The map is calculate by analyzing training sets of the facial expressions with PCA. To perform the facial transfer, this method normalizes the movements of the key points and creates the weight vector between a source face and a target one. In this kind of the approaches, weighting values for key points are an important factor to represent the movement of a face. Moreover, since the targets of these methods are sparse polygon models or 2D images, a fitness of the shape in the area where there are no key points is not evaluated. This point can be a problem if these methods apply to dense face models.

## 3 OVERVIEW OF THE PROPOSED METHOD

We describe the overview of our study. The proposed method in this study consists of the following three steps.

- 1) **Initialization of key points tracking** In the paper, we assume the scanning method which can capture both 3D point clouds and 2D images at the same time. Therefore, we utilize the method proposed by Cao *et al.* [5] for finding facial landmarks from a 2D image of the face to detect key points in each frame as the initial tracking result for the following process.
- 2) **Non-rigid shape deformation for dense motion** In our method, a dense motion of a facial expression from an initial frame to other frames is represented by 3D non-rigid shape deformation. We utilize a Radial Basis Function based deformation

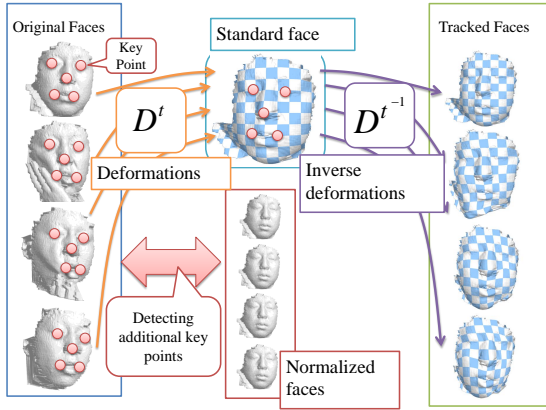


Figure 1: Overview of our approach for tracking a facial motion.

method [16]. The method generates deformed shapes from the shape of the initial frame to the other frames by using a pair of key points: the key points of the initial frame and those of the other frames (Figure1).

**3) Extra key points addition** Since the number of key points detected as the initial process is relatively small compared to all the 3D points, there is naturally a considerable amount of errors existing between scanned shapes and the deformed ones. To solve this problem, our method automatically detects additional key points to decrease the errors. Our method searches an area where errors between the scanned shapes and the deformed ones are larger than other areas to find the additional key points. By repeating the process, errors are minimized.

## 4 DEFORMATION METHOD OF A FACE

We describe the details of a deformation method of a facial motion. In the paper,  $F^t$  denote a vector of 3D points  $(p_0^t, \dots, p_n^t)$  of the face in the  $t$ -th frame. Then, our purpose of the facial motion tracking is to find deformations from  $F^0$  to  $F^t$ . In the following section, we describe a basic idea of the deformation method.

### 4.1 Shape deformation based on the Radial Basis Function

We introduce a deformation method of a 3D object based on a key point interpolation. We utilize the deformation method based on RBF [16]. The 3D object is created by adding all the shapes which are calculated by multiplying each key point by RBF (Figure 2). Therefore, let the pairs of key points to be  $K^0$  and  $K^t$  where  $K^t$  is a vectors of 3D points  $(k_0^t, \dots, k_M^t)^T$  of the face  $F^t$  of  $t$ -th frame, then,  $K^t$  can be calculated by a sum of the

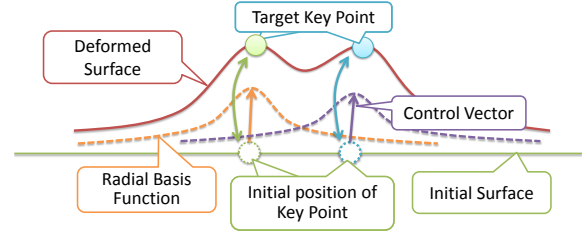


Figure 2: RBF based deformation

multiplication of all  $k_i^0 (i \in M)$  by RBF with the control vector  $C^t = (c_0^t, \dots, c_M^t)^T$  as follows (Figure 2):

$$K^t = RBF(W(K^0, K^0))C^t \quad (1)$$

where

$$W(K^0, K^0) = \begin{bmatrix} \|k_0^0 - k_0^0\| & \dots & \|k_0^0 - k_m^0\| \\ \vdots & \ddots & \vdots \\ \|k_m^0 - k_0^0\| & \dots & \|k_m^0 - k_m^0\| \end{bmatrix}. \quad (2)$$

The values of the matrix  $W$  are the parameters of the RBF according with the distance from each point in  $K^0$  to other points. If the inverse matrix of  $RBF(W(K^0, K^0))$  exists, we can solve the equation (1) to find the value of  $C^t$ . By using the inverse matrix of  $RBF(W(K^0, K^0))$  and the vectors  $C^t$ , the key points  $K^0$  are correctly transformed to  $K^t$ . Finally, by using the set of key points  $K^0$  and the control vectors  $C^t$ , an arbitrary point  $x$  is deformed by the following equation.

$$x^t = RBF(\hat{W}(x, K^0))C^t \quad (3)$$

$x^t$  is a deformed point to the  $t$ -th frame. By using this deformation method, we can deform all the points of frame 0 to any frame. In the following sections, we represent this process of the deformation from the face  $F^0$  to  $F^t$  as the following equation.

$$\hat{F}^t = D(F^0 \rightarrow F^t) \quad (4)$$

$\hat{F}^t$  is a face which is deformed from  $F^0$  to fit to  $F^t$ . And by using the set of key points  $K^t$ , we can also make deformations form arbitrary frame number to another frame.

#### 4.1.1 Kernel function

The feature of the kernel function of the RBF is important to deform a 3D object smoothly. In our study, we define the kernel as the Gaussian function.

$$RBF(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x}{\sigma}} \quad (5)$$

The Gaussian function is positive and symmetric. Thus, if the value of  $\sigma$  is the same in each key point, the weighting matrix  $RBF(W)$  has an inverse matrix. By the reason, we use the fixed value for the  $\sigma$ .

## 5 ADDITION OF EXTRA KEY POINTS

We detect and add extra key points to reduce the difference between  $F^0$  and  $D^t(F^t \rightarrow F^0)$ . To calculate the difference between  $F^0$  and  $D^t(F^t \rightarrow F^0)$ , we use  $e_i^t$  which is a distance between the point  $p_i^0$  in  $F^0$  and its nearest point in  $D^t(F^t \rightarrow F^0)$ . Since our deformation method transforms the key points to target position and multiply with RBF which decrease the weight according to the distances from the key points, an error becomes large in the area where the key points are sparse (Figure 2 and Equation (3)). To overcome this problem, the area where the error is large is detected and a new key point is added.

### 5.1 Initial key point in 2D space

Our method requires several numbers (10 - 30) of initial key points on a 2D image. Finally we detect a suitable set of key points automatically. However, since our method evaluates the difference between a standard face and normalized ones to detect the area lacking of key points, the method requires an initial guess of a deformation.

There is a lot of method to detect features of 3D object is proposed[1]. These methods are useful to track rigid movements of the objects. However these methods cannot detect the same positions on a face between two faces. And it is difficult to track the same position of a face with topological transformations (e.g. open a mouth or blink eyes) by using these methods. Therefore we utilize the method proposed by Cao *et al.* [5] for finding facial landmarks from a 2D image of the face. This method can find the facial landmarks with the accuracy over 95%. This accuracy is enough for the initial set of key points for our method. Before the detection process of extra key points, we apply this method to initialize the set of key points  $K^t$ .

### 5.2 Key point candidate selection

Since the shape deformation in our method is calculated by a key point and its interpolation, the error near the key point is naturally small. Therefore, it is a straightforward to find a new key point from the large error area. In our method, we select the candidates for the new key point which satisfies the following conditions:

1. The sum of the errors of neighboring points is larger than the threshold.
2. An error of the key point is larger than neighboring points.

Since the number of 3D points of a face is usually large, comparing the criteria for all the point is impractical, we randomly sample the candidates of key points  $Kc_k^0$  from the  $F^0$ . In our implementation, we sample 4,000 points.

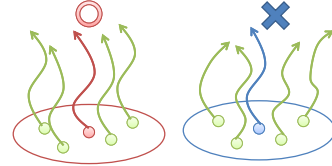


Figure 3: Filtering of candidates of key points. (a)The area where the correlation is high. (b)The area where the correlation is low because of there are outliers.

### 5.3 Filtering the candidate using temporal consistency

Since selected candidates usually include many outliers because of noise, occlusion and other reasons, filtering is necessarily required. In our method, we check the temporal consistency of neighboring points (Figure 3). The detailed calculation process is as follows:

1. In terms of the candidates of key points  $Kc_k^0$ , find the neighboring points  $np_j$  inside the sphere of radius  $R$ .
2. The error values of  $e_k^t$  and  $e_j^t$  ( $j \in R$ ) are calculated.  $e_k^t$  is the minimum distance form  $Kc_k^0$  to  $D^t(F^t \rightarrow F^0)$  and  $e_j^t$  is the distances is the minimum distance form  $np_j$  to  $D^t(F^t \rightarrow F^0)$ .
3. Time series of  $e_k^t$  and  $e_j^t$  is calculated for all the frames of the face  $F^t$  to make a sequence of errors  $e_k$  and  $e_j$ .
4. The correlation between  $e_k$  and  $e_j$  are calculated. The higher value of the correlation means that this candidate of a key point  $Kc_k^0$  is placed on the center of a large movement of a face.
5. If the correlation is below the average of all the candidates, the point  $Kc_k^0$  is rejected.

Through this process, we reduce the candidates  $Kc_k^0$  under 40.

### 5.4 Refinement of the new key points

Since the candidates of key points are selected from the area where the errors are larger than other area, the nearest point in  $D^t(F^t \rightarrow F^0)$  form  $Kc_k^0$  is usually the wrong point for deciding positions of the key points in other frames  $Kc_k^t$ . Therefore, we utilize the Non-Rigid Registration method which uses the Gaussian Mixture model for representing 3D points for registrations [10] for a local search for the positions of the key points  $Kc_k^t$  (Figure 4). The detailed process is as follows:

1. By using  $D^t(F^0 \rightarrow F^t)$  (with the current set of the key points), we deform the  $F^0$  for creating "currently tracked" faces in each frame. Then, the  $Kc_k$  is also deformed to  $Kc_k^t$  by the deformation.

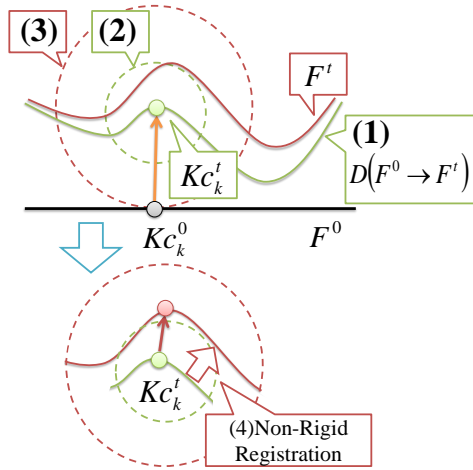


Figure 4: Refinement of key points

2. A surface of  $D^t(F^0 \rightarrow F^t)$  inside the sphere of radius  $R$  and the center of which is the key points  $Kc_k^t$  is selected. The selected surface represents a shape of the current face which contains the errors.
3. Similarly, another surface of  $F^t$  inside the sphere of radius  $2R$  and the center of which is  $Kc_k^t$  is selected. The second surface represents a shape of the original face  $F^t$  around the key point.
4. By using the Non-Rigid Registration method, the first surface fits into the second one. Then, the positions of the key points  $Kc_k^t$  are also fitted to the  $F^t$ .

Through the process, pairs of additional key points  $Kc_k^0$  and  $Kc_k^t$  are detected and added to the list of the key points  $K^t$ .

### 5.5 Iterative step for adding key points

We iterate the process of adding key points until the ratio of error is decreased to less than 1% from the previous process. During the iteration process, we also remove the existing key points which have larger error than newly added key points. Initial key points are also removed in the actual process in our experiments.

## 6 FACIAL TRANSFER METHOD BASED ON THE FACE DEFORMATION

Facial expressions and motions on a face are important presentation in animations. To utilize the captured motions efficiently, the methods transfer from the facial motion of a person to other persons are proposed. The transfer methods are used in the following examples: (a) re-product facial animations of a person by using a motion data of another person. (b) "Stuntman" of the

facial motion. (c) to create animations of virtual characters.

These methods use pairs of key points to define the correspondence between a transfer source shape and a target one. One of the important points of this kind of methods is how to interpolate the difference of both shapes. One of the proposed methods[19] defines a zone of influence and a weight map for interpolating the movement of key points. It is a time consuming process to set the parameters of key points. To solve this problem, we utilize the face space deformation method to transfer facial motions. This method deforms 3D space around a face based on positions of key points. Therefore, the method can be a solution to represent a conversion from a movement of a face to that of other faces. This method consists of following four steps.

- (1) Creating deformations from a source motion of a face to a target one.
- (2) Key point sampling to define correspondences between the source and target faces.
- (3) Detection of weighting values of key points for deformations.
- (4) Creating deformations of the target face.

The details of these are explained in the following sections.

### 6.1 Deformation from a source motion of a face to a target face

First, we deform a sequence of a facial motion  $F^t$  (source face) to fit into a shape of another face  $T^0$  (target face). To define the deformation based on the face space deformation, we use a set of key points described in the section 5.1. Then, the deformations are given by the equation 6.

$$\hat{F}T^t = D^t(F^t \rightarrow T^0) \quad (6)$$

By using the deformation  $D^t(F^t \rightarrow T^0)$ , each source face  $F^t$  is deformed to fit to the target face  $T^0$ . Although the structures of  $\hat{F}T^t$  differ to  $T^0$ , the shapes of  $\hat{F}T^t$  close to  $T^0$ . Since the  $T^0$  is a static, the remaining difference between  $T^0$  and  $\hat{F}T^t$  means the movement of the face in the space of the target face. Then, we define a correspondences between  $\hat{F}T^t$  and  $T^0$ .

### 6.2 Key point detection for a face transfer

Since the number of the points in the  $T^0$  and  $\hat{F}T^t$  is too large for defining deformations, our method defines pairs of key points as follows:

- (1) We define a set of key points  $Kt_i$  by random sampling from the target face  $T^0$ .

- (2) Another set of key points  $Kf_i^t$  which are the nearest points in  $F\hat{T}^t$  from each point of  $Kt_i$  are detected.

The relationships between  $Kt_i$  and  $Kf_i^t$  are the pairs of key points.

### 6.3 Weighting values of key points

We define direction vectors  $V_i$  from the target face to the source motion based on the pairs of key points  $Kt_i$  and  $Kf_i^t$  as follows:

$$V_i^t = \frac{Kf_i^t - Kt_i}{\|Kf_i^t - Kt_i\|} \quad (7)$$

Since the difference between  $T^0$  and  $F\hat{T}^t$  are caused by the movement of the face  $F^t$ , these vectors  $V_i^t$  represent directions of movement of the face in each key point. However, since the positions of key points  $Kf_i^t$  are fitted to the target face ( $F^t$  is deformed in the space of another face  $T^0$ ), the features of the movement of the source face are decreased by the deformations. Then, we introduce weighting values of key points to improve the feature of the movement of  $F^t$ . We consider the initial position of the key point  $Kf_i^0$  to calculate the weighting values. Then, the weighting values  $W_i$  of key points  $Kf_i^t$  are given by following equation 8.

$$W_i = \frac{\alpha}{N_k} \|Kf_i^0 - Kf_i^t\| \quad (8)$$

$\alpha$ : a parameter for controlling the strength of transfer effect.

$N_k$ : a number of the key point  $Kf_i^0$

### 6.4 Deformations of the facial transfer

Finally, we construct deformations of the facial transfer  $D_{trms}^t$ . By using the direction vectors of the key points  $V_i$  and the weighting values  $W_i$ , the deformations  $D_{trms}^t$  are given by following equation.

$$D_{trms}^t(Kt_i \rightarrow (W_i^t V_i^t + Kt_i)) = \hat{T}^t \quad (9)$$

$\hat{T}^t$  represents the transferred face from  $T^t$  with the pair of key points  $Kt_i$  and  $W_i^t V_i^t + Kt_i$ . Though these four steps mentioned in the beginning of this chapter, the facial transfer method generates a motion of an arbitrary face that reflects motion of another face.

## 7 RESULTS AND DISCUSSIONS

In this section, we show the results of a facial motion transfer and the efficiency of a key point refinement technique. First, we evaluate the accuracy of deformations by calculating errors between the original faces to the deformed faces. Next, we discuss about the results of our facial transfer method. In our experiment, we use three types of facial motions: "Slap", "Smile" and "Stretch" (Fig. 5). These motions consist of about 100,000 to 200,000 points and 47 to 60 frames. We implement the proposed method on PC with Xeon X5650 processor. It takes about 10 minutes for processing one of the iteration described in section 5.5.

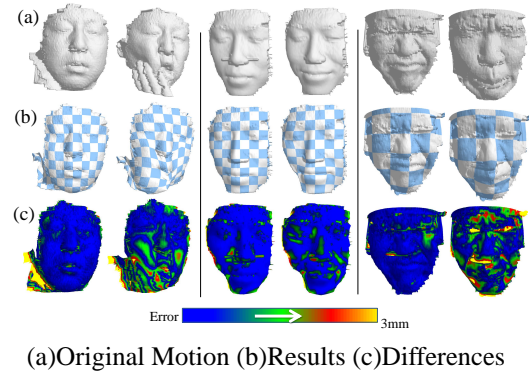


Figure 5: The comparison of the original and captured face.

Table 1: The errors between the original motion and captured motion.

Name	mean error (mm)		RMSE (mm)	
	Initial	Final	Initial	Final
Slap	1.49	1.00	1.89	1.41
Smile	0.64	0.44	0.87	0.66
Stretch	1.23	0.84	1.62	1.22

### 7.1 Accuracy of the motion capture

In this section we discuss about the error between the original motion and a motion which is deformed from the standard face by using the key points. Table 1 shows the mean errors and RMSE tracked by using the proposed method. The initial error is the mean distance between the face which is deformed with the initial set of key points and the original face. The final error is the mean distance when the recursive process reaches to the convergence condition. The proposed method enables us to minimize the mean error less than 1.0mm on each motion.

Table 2 shows the errors without the key point selection mentioned in the section 5.3 and 5.4 for comparisons. The table shows that the errors are larger than two times those of the proposed method. This result shows that the proposed key point selection method can select the suitable set of key points.

Figure 5(b) and (c) show the deformed motions and the distributions of the errors. This result (b) shows that the proposed method smoothly keeps the continuity of the texture in each frame. The distribution of the error shows that the errors are still remaining around the area where the hand touches the face. Since our method evaluates the motion of key points based on the standard face, it is difficult to detect the motion of an additional object such as a hand. Although there are some errors, our method succeeded in detecting the correspondence between the standard face and the other motions.

Table 2: The errors based on randomly adding the key points.

Name	mean error (mm)	RMSE (mm)
Slap	2.25	2.76
Smile	1.07	1.52
Pinch	1.67	2.09

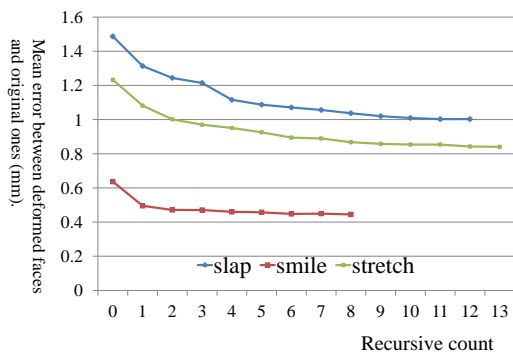


Figure 6: The convergence of the mean error.

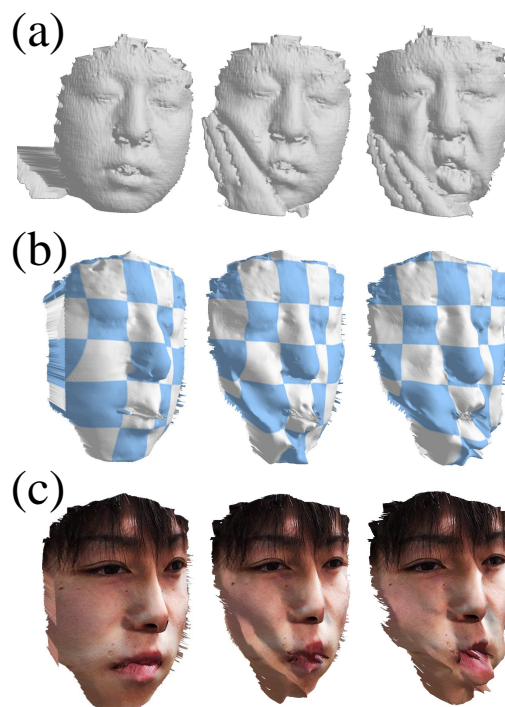
## 7.2 Results of the facial transfer

Figure 7 shows the results of the proposed facial transfer method. (a) is the original motion (Slap). The left image of (b) is the target shape of a face for applying the facial transfer. The results in (b) and (c) show that the movement of the left cheek and the lip can be transferred to the target face. In the right images of (b) and (c), there are noises around the lip and lower jaw. Since our transfer method randomly samples candidates of key points, key points are placed on the noisy area of the source motion in sometimes. Although the clustering method mentioned in section 5.4 chooses the center of the candidates as a key point, this problem has still happened in this case. To solve this problem, we should apply a constraint function such as the Tevs's facial tracking method[17] for a deformation of a face.

## 8 CONCLUSIONS

We proposed the key point detection and refinement method for a facial motion which is represented by dense point cloud. We also proposed the facial transfer method base on the key points. The contributions of this paper are as follows:

- (1) We propose the face space deformation which can represent the movement of a face as the RBF based deformations.
- (2) By utilizing the face space deformation, the proposed method can define the suitable pairs of key points to minimize the errors.



(a)Original Motion (b)Transfer target(left) and transferred shape of the face. (c)Transferred motion with a texture.

Figure 7: The comparison of the original motion and the transferred motion.

- (3) In our results, we show that the mean errors between the original motions and tracked motions can reduce less than 1.0mm in all three types of facial motions.
- (4) We also show our facial transfer method can transfer the complex motion on a face to another face.

In the future work, a refinement technique of a surface of a face is needed to generate smoother animations.

## 9 ACKNOWLEDGMENT

This work was supported in part by the Funding Program for Next Generation World-Leading Researchers(NEXT Program) No.LR030 and Grant-in-Aid for Young Scientists(B) No.25870570 in Japan.

## 10 REFERENCES

- [1] Bariya, P., Novatnack, J., Schwartz, G., Nishino, K.: 3d geometric scale variability in range images: Features and descriptors. *Int. J. Comput. Vision* **99**(2), 232–255 (2012)
- [2] Belhumeur, P., Jacobs, D., Kriegman, D., Kumar, N.: Localizing parts of faces using a consensus

- of exemplars. In: *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 545–552 (2011)
- [3] Bickel, B., Lang, M., Botsch, M., Otaduy, M.A., Gross, M.: Pose-space animation and transfer of facial details. In: *Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '08*, pp. 57–66. Eurographics Association (2008)
- [4] Bradley, D., Heidrich, W., Popa, T., Sheffer, A.: High resolution passive facial performance capture. *ACM Trans. Graph.* **29**(4), 41:1–41:10 (2010)
- [5] Cao, X., Wei, Y., Wen, F., Sun, J.: Face alignment by explicit shape regression. In: *CVPR*, pp. 2887–2894. IEEE (2012)
- [6] Cosker, D., Roy, S., Rosin, P.L., Marshall, D.: Re-mapping animation parameters between multiple types of facial model. In: *Proceedings of the 3rd international conference on Computer vision/computer graphics collaboration techniques, MIRAGE'07*, pp. 365–376. Springer-Verlag, Berlin, Heidelberg (2007)
- [7] Fratarcangeli, M., Schaerf, M., Forchheimer, R.: Facial motion cloning with radial basis functions in mpeg-4 fba. *Graph. Models* **69**(2), 106–118 (2007)
- [8] Huang, D., De La Torre, F.: Facial action transfer with personalized bilinear regression. In: *Proceedings of the 12th European conference on Computer Vision - Volume Part II, ECCV'12*, pp. 144–158. Springer-Verlag, Berlin, Heidelberg (2012)
- [9] Huang, H., Chai, J., Tong, X., Wu, H.T.: Leveraging motion capture and 3d scanning for high-fidelity facial performance acquisition. *ACM Trans. Graph.* **30**(4), 74:1–74:10 (2011)
- [10] Jian, B., Vemuri, B.C.: Robust point set registration using gaussian mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(8), 1633–1645 (2011)
- [11] Moro, A., Mumolo, E., Nolich, M.: Automatic 3d virtual cloning of a speaking human face. In: *Proceedings of the 2010 ACM workshop on Surreal media and virtual cloning, SMVC '10*, pp. 45–50. ACM, New York, NY, USA (2010)
- [12] Pyun, H., Kim, Y., Chae, W., Kang, H.W., Shin, S.Y.: An example-based approach for facial expression cloning. In: *ACM SIGGRAPH 2006 Courses, SIGGRAPH '06*. ACM, New York, NY, USA (2006)
- [13] Sattar, A., Stoiber, N., Seguier, R., Breton, G.: Gamer's facial cloning for online interactive games. *Int. J. Comput. Games Technol.* **2009**, 9:1–9:16 (2009)
- [14] Seol, Y., Lewis, J., Seo, J., Choi, B., Anjyo, K., Noh, J.: Spacetime expression cloning for blendshapes. *ACM Trans. Graph.* **31**(2), 14:1–14:12 (2012)
- [15] Sibbing, D., Habbecke, M., Kobbelt, L.: Markerless reconstruction and synthesis of dynamic facial expressions. *Comput. Vis. Image Underst.* **115**(5), 668–680 (2011)
- [16] Sumner, R.W., Schmid, J., Pauly, M.: Embedded deformation for shape manipulation. In: *ACM SIGGRAPH 2007 papers, SIGGRAPH '07*. ACM (2007)
- [17] Tevs, A., Berner, A., Wand, M., Ihrke, I., Bokeloh, M., Kerber, J., Seidel, H.P.: Animation cartography –intrinsic reconstruction of shape and motion. *ACM Trans. Graph.* **31**(2), 12:1–12:15 (2012)
- [18] Valgaerts, L., Wu, C., Bruhn, A., Seidel, H.P., Theobalt, C.: Lightweight binocular facial performance capture under uncontrolled lighting. *ACM Trans. Graph.* **31**(6), 187:1–187:11 (2012)
- [19] Vieira, R.C.C., Vidal, C., Cavalcante-Neto, J.B.: Expression cloning based on anthropometric proportions and deformations by motion of spherical influence zones. In: *SBC Journal on 3D Interactive Systems*, pp. 14–22 (2011)
- [20] Weise, T., Bouaziz, S., Li, H., Pauly, M.: Realtime performance-based facial animation. In: *ACM SIGGRAPH 2011 papers, SIGGRAPH '11*, pp. 77:1–77:10. ACM, New York, NY, USA (2011)
- [21] Zhang, W., Wang, Q., Tang, X.: Real time feature based 3-d deformable face tracking. In: *Proceedings of the 10th European Conference on Computer Vision: Part II, ECCV '08*, pp. 720–732. Springer-Verlag, Berlin, Heidelberg (2008)