

## A Study on 3D Face Similarity by Point Cloud Based Metric for Japanese Terracotta Figurines (Haniwa)\*

Xin Lu<sup>1),†</sup> Chunyuan Li<sup>2)</sup> Tsutomu Kinoshita<sup>3)</sup> Akio Kimura<sup>1),†</sup>  
Kouichi Konno<sup>1),†</sup>

1) Faculty of Science and Engineering, Iwate University

2) Faculty of Engineering, Bei Hua University

3) Faculty of Engineering, Tohoku Gakuin University

{†luxin, kimura, konno} (at) cis.iwate-u.ac.jp

### Abstract

The Haniwa are a type of terracotta figurines from ancient Japan, which have been scanned and saved as three-dimensional models composed of point cloud and triangular meshes for archaeological research. Although these three-dimensional models can be used to analyze the production technique of the Haniwa, there are several cases where archaeologists subjectively classify the Haniwa using only visual and qualitative observation approaches based on the Haniwa features, i.e., the face, clothes, etc. Here, the need for a computer-aided and quantitative evaluation method, using three-dimensionally measured point cloud, becomes evident. For a quantitative evaluation of the Haniwa, the evaluation of face similarity is considered an important issue. As the face is mainly composed of nose, mouth, and eyes, the face similarity of the Haniwa can be evaluated by comparing the corresponding facial parts between individuals. However, shapes and placements of the same facial parts for each pair of Haniwa have individual differences. Therefore, the face alignment to precisely localise the facial parts and recognize their shapes should be required. To achieve the aforementioned precise alignment, a hybrid method combining ellipsoid-fitting-based segmentation and topological hole detection is proposed, to extract the point cloud of the nose and locate the contours of the eyes and mouth. Besides the face alignment, the quantitative evaluation essentially contains a process of comparing the faces by analyzing the relationships between the facial parts. For the face comparison, when the facial parts are aligned, the point clouds of the nose for each pair of Haniwa are registered through a scaling iterative closest point algorithm to achieve a value for scale alignment. After the point clouds of all the facial parts for each pair of Haniwa are aligned on the same scale to suppress scale error generated in production, their difference is calculated using the standard iterative closest point algorithm to obtain point-based dissimilarity for each pair of Haniwa. Moreover, owing to basic visual observation, the difference in the relative positions of the eyes and mouths with respect to those of the noses was manually measured as another dissimilarity for face comparison. Our experimental results demonstrate that the proposed method is effective for the similarity evaluation of the Haniwa through the three-dimensional point cloud.

---

\*Part of the material in this article has been orally presented at conference NICOGRAPH 2019 [1]. This article has been improved based on the peer-reviewed comments at the submission of NICOGRAPH 2019, and questions of the chair at its session.

## 1 Introduction

The Haniwa [2] are a type of terracotta figurines that were made for ritual use and were buried with the dead as funerary objects during the Kofun period (3rd to 6th centuries AD) of Japan. Archaeologists mainly attempted to clarify information on their provenances with datings or evaluate their artistic values. Specifically, the archaeologists conventionally identified groups of workers and individual workers from the human Haniwa remains by the face, clothes, etc. and further evaluated the findings using detailed manufacturing techniques and tools, through a visual and qualitative observation approach [2]. However, a more quantitative evaluation of the human Haniwa revealed that clear and subjective features, such as the presence or absence of parts extracted through the visual observation approach, could not satisfy objective evaluation requirements.

Recently, as three-dimensional measurement devices have been developed, point clouds representing the surface of an object become easily accessible. Compared with conventional recording methods such as photographs and rubbed copies, the amount of data obtained through a three-dimensional measurement is significant. Additionally, compared to a visual research, the three-dimensional measurements can provide more accurate and objective data. Therefore, a quantitative evaluation method for extracting and comparing detailed three-dimensional information of human Haniwa from the target point cloud is required.

To ensure precise facial evaluation results, the objective and quantitative evaluation method should be preferable to evaluate the Haniwa face, because the face contains important manufacturing information. Moreover, the facial parts, such as nose, mouth, and eyes, need to be used preferentially, because they are the most outstanding features of the face. Facial alignment to precisely localize the facial parts and determine their shapes becomes important. As the Haniwa were not produced in the same size owing to technical limitations, for face comparison like a human being, the extracted point clouds of two individuals

need to be resized on the same scale using a simple and effective scale-aligning method. After the scale alignment, a dissimilarity metric is required for measuring the global distance between individuals directly based on point clouds. Moreover, the difference in the relative positions of eyes and mouths with respect to those of noses needs to be measured as another dissimilarity metric for face comparison because measuring relative positions is a common practice of humans for visual comparison.

To overcome the above issues, this study addresses the following: (i) extraction of the facial parts of the human Haniwa from the point cloud; (ii) adjustment of the face in a same scale between individuals; (iii) evaluation of the similarity between individuals by comparing the same facial parts based on the two dissimilarity metrics. For our proposed method, both usefulness of the face alignment, applied to the three-dimensional point cloud of human Haniwa, as well as effectiveness of the dissimilarity metrics are confirmed in this paper.

## 2 Related Studies

Thus far, studies that quantitatively evaluate the Haniwa face through three-dimensional point cloud data have rarely been introduced; however, face alignment, which is the first step of the evaluation method, has been used for the human face in the past three decades.

Facial alignment estimates the position of facial parts such as facial contour, eyes, eyebrows, and mouth from the target image. It is vital to facial recognition and can be realized using many methods, including the active appearance model (AAM) [3] and constrained local model (CLM) [4]. Since sufficient extraction accuracy could not be obtained, it was difficult to use these methods in an actual environment prior to the development of AI.

During the development of AI, Cao et al. proposed a machine learning model, referred to as explicit shape regression (ESR) [5]. ESR can be considered as a transformation from a two-dimensional facial image to a three-dimensional morphable model; it can realize facial part ex-



Figure 1: Facial part extraction using deep learning from two-dimensional facial image of the human Haniwa.

traction rapidly in an actual environment. ESR is identified by training a three-dimensional facial expression dataset [6] which contains information on 150 people with 47 different facial expressions, captured using an RGBD camera. However, in our study, the human Haniwa are too few to be collected and processed by ESR.

Meanwhile, the introduction of deep learning in 2012 impelled a rapid development in image recognition technology. Lu et al. proposed an accurate facial part extraction method using deep learning for humans [7]. However, such a deep-learning-based method for two-dimensional image requires a large number of images for learning; hence, it is difficult to apply this method to the human Haniwa without a sufficient number of samples. Their extraction method can extract facial parts from the two-dimensional facial image of the human Haniwa; however, the results are substandard with duplicate or missed extractions for facial parts, as shown in Figure 1. The extracted left eye, right eye, nose, and mouth should be marked in green, red, yellow, and blue, respectively. Unfortunately, the facial part extraction using deep learning from two-dimensional facial image of the human Haniwa did not perform well.

### 3 Our Method

Owing to the above-mentioned disadvantages, we propose a novel evaluation method involving face alignment and comparison for the Haniwa face.

For face alignment, a novel method is presented

for extracting their detailed three-dimensional information using traditional computer graphics technology from the target point cloud. The schematic diagram of this evaluation method is illustrated in Figure 2. Based on a point cloud with a triangular mesh of the Haniwa obtained using a three-dimensional laser scanner, the face surface was first estimated by fitting an ellipsoid after the front head of the Haniwa was extracted. Subsequently, a method for segmenting the prominent parts of the Haniwa using the fitted ellipsoid was proposed to mark the nose points on the face, as well as a three-dimensional hole detection method to label the contours of its eyes and mouth.

As the point cloud of the nose contains more points than those of the eyes and mouth, the point clouds of the nose for each pair of Haniwa were registered through the scaling iterative closest point (sICP) algorithm [8] to achieve a scale value for scale adjustment. After the point clouds of all face parts for each pair of Haniwa were resized on the same scale, their root mean squared error (RMSE) was calculated using the standard iterative closest point (ICP) algorithm [9,10] as a dissimilarity for each pair of Haniwa. Moreover, mean angle error (MAE) of the lines from the nose centroid to the centroids of the other face parts was originally measured as another metric for facial part arrangement dissimilarity according to habitual of humans for visual comparison.

The details are described in the following sections.

## 4 Face Alignment

### 4.1 Front Head Extraction

Six three-dimensional models of the Haniwa were selected from the dataset [2]. Owing to the state of excavation and preservation, one of them had both a head and a body, whereas the others only had a head. To perform face alignment stably, each front head of these models was manually extracted and abstracted as a graph  $\mathcal{P} := (\mathcal{V}, \mathcal{E}, \mathcal{F})$ . Every  $\mathcal{P}$  includes vertices (cloud points)  $\mathbf{v} := [x \ y \ z]^T \in \mathcal{V}$ , edges  $\mathbf{e} := (\mathbf{v}^a, \mathbf{v}^b) \in \mathcal{E}$  and faces  $\mathbf{f} := (\mathbf{v}^a, \mathbf{v}^b, \mathbf{v}^c) \in \mathcal{F}$ . An example of the corresponding extracted results is shown in Figure

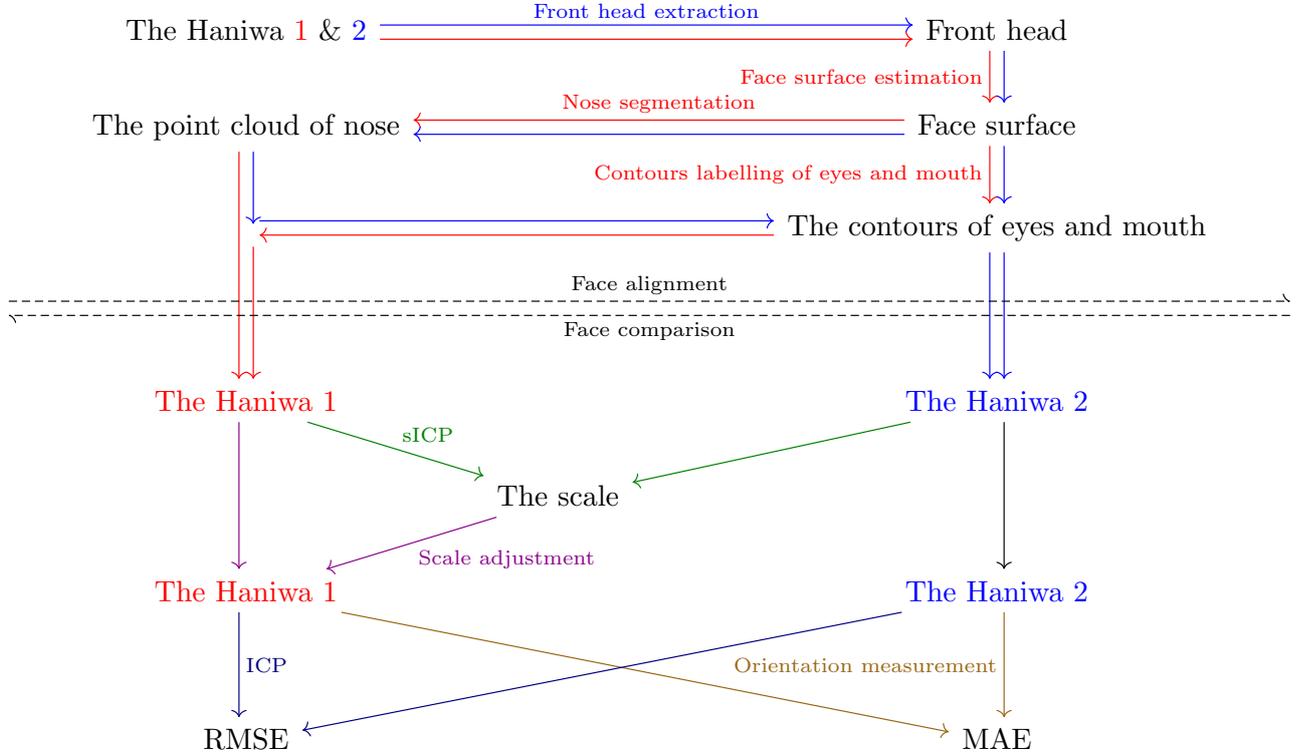


Figure 2: Schematic diagram of face alignment and comparison for three-dimensional models of the Haniwa.

3.

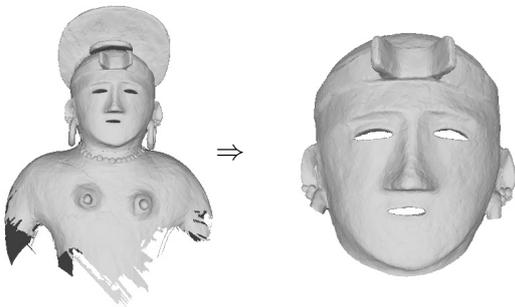


Figure 3: Front head extraction from three-dimensional models of Haniwa. Left: three-dimensional model of Haniwa; right: extracted front head.

## 4.2 Face Surface Estimation

The three-dimensional models of the Haniwa enable their heads to be fitted and approximated by ellipsoids. Since the curvature of the fitted ellipsoid is similar to that of the Haniwa face, the ellipsoid can be considered as a reference surface for determining the concave and convex portions facial parts. If a point cloud lies outside of the fitted ellipsoid, the facial part represented by this point cloud is convex. Otherwise, the facial part is concave. We suggest using the direct least-squares fitting method [11, 12] to locate the ellipsoid as follows. An ellipsoid can be defined as

$$\mathbf{v}^T \mathbf{A} \mathbf{v} + \mathbf{b}^T \mathbf{v} + c = 0, \quad (1)$$

with a symmetric and positive matrix  $\mathbf{A}$ , vector  $\mathbf{b}$ , and scalar  $c$ .  $\mathbf{A}$  and  $\mathbf{b}$  are defined as

$$\mathbf{A} := \begin{bmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{bmatrix}, \quad \mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}. \quad (2)$$

Eqn. (1) can be written as a quadric equation as follows:

$$a_1x^2 + a_2y^2 + a_3z^2 + 2a_4xy + 2a_5xz + 2a_6yz + b_1x + b_2y + b_3z + c = 0, \quad (3)$$

subject to the positive-definite constraints of  $\mathbf{A}$ , i.e.,

$$a_4^2 < a_1a_2, \quad a_5^2 < a_1a_3, \quad a_6^2 < a_2a_3, \quad (4)$$

which is referred from [13]. By rearranging the variables  $x^2, y^2, z^2, xy, xz, yz, x, y$  and  $z$  as a variable vector  $\mathbf{x}$ :

$$\mathbf{x} := [x^2 \ y^2 \ z^2 \ xy \ xz \ yz \ x \ y \ z \ 1]^T, \quad (5)$$

and the coefficient vector  $\mathbf{w}$ :

$$\mathbf{w} := [a_1 \ a_2 \ a_3 \ 2a_4 \ 2a_5 \ 2a_6 \ b_1 \ b_2 \ b_3 \ c]^T, \quad (6)$$

the ellipsoid can be written in a vector form,

$$f(\mathbf{w}, \mathbf{x}) := \mathbf{w}^T \mathbf{x} = 0. \quad (7)$$

Subsequently, ellipsoid fitting can be achieved by minimizing the sum of the squared algebraic distances of the cloud points with respect to the expected coefficient vector as follows:

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} \sum_{i=1}^N f(\mathbf{w}, \mathbf{x}_i)^2, \quad (8)$$

where the variable vector of every cloud point  $\mathbf{x}_i$  is defined by

$$\mathbf{x}_i := [x_i^2 \ y_i^2 \ z_i^2 \ x_i y_i \ x_i z_i \ y_i z_i \ x_i \ y_i \ z_i \ 1]^T, \quad (9)$$

with equality constraints:

$$a_1a_2 - a_4^2 = 1, \quad a_1a_3 - a_5^2 = 1, \quad a_2a_3 - a_6^2 = 1, \quad (10)$$

changed from Eqn. (4). Therefore, the ellipsoid-fitting problem is reformulated as

$$\min_{\mathbf{w}} \|\mathbf{D}\mathbf{w}\|^2, \quad \text{subject to } \mathbf{w}^T \mathbf{C}\mathbf{w} = 1, \quad (11)$$

where  $\mathbf{D}$  is an  $N \times 10$  design matrix

$$\mathbf{D} := \begin{bmatrix} x_1^2 & y_1^2 & z_1^2 & x_1y_1 & x_1z_1 & y_1z_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ x_N^2 & y_N^2 & z_N^2 & x_Ny_N & x_Nz_N & y_Nz_N & \dots \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ x_1 & y_1 & z_1 & 1 & & & \\ \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ x_N & y_N & z_N & 1 & & & \end{bmatrix}, \quad (12)$$

for  $N$  points in the cloud.  $\mathbf{C}$  is a  $10 \times 10$  constraint matrix

$$\mathbf{C} := \begin{bmatrix} \mathbf{C}_1 & \mathbf{0}_{[6 \times 4]} \\ \mathbf{0}_{[4 \times 6]} & \mathbf{0}_{[4 \times 4]} \end{bmatrix},$$

$$\mathbf{C}_1 := \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix}, \quad (13)$$

representing the ellipsoid constraint as shown in Eqn. (10). Applying the Lagrange multiplier  $\lambda$  in Eqn. (11), we obtain the following Lagrange function

$$\mathcal{L}(\mathbf{w}, \lambda) := \|\mathbf{D}\mathbf{w}\|^2 + \lambda(\mathbf{w}^T \mathbf{C}\mathbf{w} - 1). \quad (14)$$

Its optimal form for  $\frac{\partial \mathcal{L}}{\partial \mathbf{w}} = 0$  and  $\frac{\partial \mathcal{L}}{\partial \lambda} = 0$  is written as

$$\mathbf{S}\mathbf{w} = \lambda \mathbf{C}\mathbf{w}, \quad \text{subject to } \mathbf{w}^T \mathbf{C}\mathbf{w} = 1, \quad (15)$$

where  $\mathbf{S}$  is a  $10 \times 10$  scatter matrix, i.e.,

$$\mathbf{S} := \mathbf{D}^T \mathbf{D}. \quad (16)$$

Ten real solutions  $(\lambda_k, \mathbf{w}_k)$ ,  $k \leq 10$  can be calculated from

$$\mathbf{C}^{-1} \mathbf{S}\mathbf{w} = \lambda \mathbf{w} \quad (17)$$

through eigen-decomposition. Because

$$\|\mathbf{D}\mathbf{w}\|^2 = \mathbf{w}^T \mathbf{D}^T \mathbf{D}\mathbf{w} = \mathbf{w}^T \mathbf{S}\mathbf{w} = \lambda \mathbf{w}^T \mathbf{C}\mathbf{w} = \lambda \quad (18)$$

the eigenvector  $\mathbf{w}_k$  corresponding to the minimal positive eigenvalue  $\lambda_k$  must to be obtained to satisfy Eqn. (11). Subsequently, the center of the

fitted ellipsoid  $\mathbf{c}$  with three axes  $\mathbf{u}_1, \mathbf{u}_2$ , and  $\mathbf{u}_3$  should be calculated as unit eigenvectors  $\mathbf{y}_1, \mathbf{y}_2$ , and  $\mathbf{y}_3$  multiplied by the squared-root of the corresponding eigenvalues  $\lambda_1, \lambda_2$ , and  $\lambda_3$  as

$$\mathbf{u}_i = \sqrt{\lambda_i} \mathbf{y}_i, \quad i = 1, 2, 3. \quad (19)$$

Note that  $\|\mathbf{u}_1\| \geq \|\mathbf{u}_2\| \geq \|\mathbf{u}_3\|$  according to  $\lambda_1 \geq \lambda_2 \geq \lambda_3$ . The fitting result is shown in Figure 4.

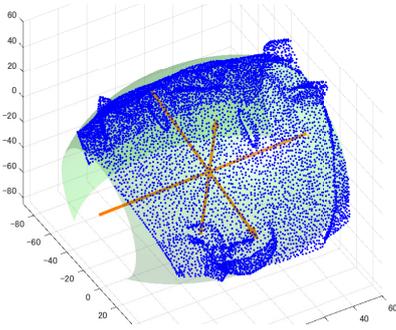


Figure 4: Face surface estimation by ellipsoid fitting.

### 4.3 Contour Labelling of Eyes and Mouth

It is important to label the eyes, nose, and mouth in facial features because they contain the most visual information of the face. For an extracted point cloud with the corresponding meshes of the front head, it is clear that the eyes and mouth comprise of topological holes of the technological level at that time. Each topological hole can be detected using Algorithm 1. Additionally, the each edge has left and right faces if the mesh is represented by oriented polygon model. In other words, an edge has one face if the edge is boundary of a hole. According to this principle, this algorithm selects the edges that only appear once in the face list, and then separates them into closed rings based on their overlapping vertices<sup>1</sup>.

<sup>1</sup>Since the edges of the gaps and silhouettes of the faces are not included in the boundary set  $\mathcal{B}$ , they do not affect the hole detection of the eyes and mouth.

---

#### Algorithm 1 Topological Holes Detection

---

- 1: List the edges from the faces  $\mathcal{F}$  to build an edge set  $\tilde{\mathcal{E}}$ .
  - 2: Select the edges without replicates from  $\tilde{\mathcal{E}}$  to build a boundary set  $\mathcal{B}$ .
  - 3:  $i \leftarrow 1$ .
  - 4: **while**  $\mathcal{B} \neq \emptyset$  **do**
  - 5:   Select  $\mathbf{e}_1 := (\mathbf{v}_1^a, \mathbf{v}_1^b)$  from  $\mathcal{B}$
  - 6:   Define a topological hole  $\mathcal{H}^{(i)} := (\mathcal{V}^{(i)}, \mathcal{E}^{(i)})$ , where  $\mathcal{V}^{(i)} \leftarrow \{\mathbf{v}_1^a, \mathbf{v}_1^b\}$  and  $\mathcal{E}^{(i)} \leftarrow \{\mathbf{e}_1\}$ .
  - 7:   Initialize a trace  $\mathcal{T}^{(i)} := \{\mathbf{v}_1, \mathbf{v}_2\}$ , where  $\mathbf{v}_1 \leftarrow \mathbf{v}_1^a$  and  $\mathbf{v}_2 \leftarrow \mathbf{v}_1^b$ .
  - 8:    $k \leftarrow 2$ .
  - 9:   **for**  $j = 2$  **to** the number of element of  $\mathcal{B}$  **do**
  - 10:     **while**  $\mathbf{v}_1 \neq \mathbf{v}_k$  **do**
  - 11:       Select  $\mathbf{e}_j := (\mathbf{v}_j^a, \mathbf{v}_j^b)$  from  $\mathcal{B}$ .
  - 12:       **if**  $\mathbf{v}_j^a = \mathbf{v}_k$  or  $\mathbf{v}_j^b = \mathbf{v}_k$  **then**
  - 13:         **if**  $\mathbf{v}_j^a = \mathbf{v}_k$  **then**
  - 14:          $\mathbf{v}_{k+1} \leftarrow \mathbf{v}_j^b$ .
  - 15:         **end if**
  - 16:         **if**  $\mathbf{v}_j^b = \mathbf{v}_k$  **then**
  - 17:          $\mathbf{v}_{k+1} \leftarrow \mathbf{v}_j^a$ .
  - 18:         **end if**
  - 19:          $\mathcal{V}^{(i)} \leftarrow \mathcal{V}^{(i)} \cup \{\mathbf{v}_{k+1}\}$ .
  - 20:          $\mathcal{E}^{(i)} \leftarrow \mathcal{E}^{(i)} \cup \{\mathbf{e}_j\}$ .
  - 21:          $\mathcal{T}^{(i)} \leftarrow \mathcal{T}^{(i)} \cup \{\mathbf{v}_{k+1}\}$ .
  - 22:          $k \leftarrow k + 1$ .
  - 23:         **end if**
  - 24:       **end while**
  - 25:     **end for**
  - 26:      $\mathcal{B} \leftarrow \mathcal{B} / \mathcal{E}^{(i)}$ .
  - 27:      $i \leftarrow i + 1$ .
  - 28: **end while**
  - 29:  $N \leftarrow i$ .
  - 30: Output  $\mathcal{H}^{(i)} := (\mathcal{V}^{(i)}, \mathcal{E}^{(i)})$ ,  $i = 1, \dots, N$ .
- 

After detecting all the topological holes on the front head, we selected the three largest holes with their vertices and edges as

$$\mathcal{H}^{(t)} := (\mathcal{V}^{(t)}, \mathcal{E}^{(t)}), \quad t = 1, 2, 3. \quad (20)$$

and marked them in red, as shown in Figure 5. Their perimeters  $L^{(t)}$  and centroids  $\mathbf{c}^{(t)}$  were cal-

culated as

$$L^{(t)} := \sum_{\mathbf{e}^{(t)} \in \mathcal{E}^{(t)}} \|\mathbf{e}^{(t)}\|, \quad \mathbf{c}^{(t)} := \frac{1}{N^{(t)}} \sum_{\mathbf{v}^{(t)} \in \mathcal{V}^{(t)}} \mathbf{v}^{(t)} \quad (21)$$

and constrained by  $L^{(1)} \geq L^{(2)} \geq L^{(3)}$ , where  $N^{(t)}$  is the number of  $\mathbf{v}^{(t)}$  in  $\mathcal{V}^{(t)}$ .

Next, the three holes were assigned to their corresponding facial parts. Random indices  $t_1$ ,  $t_2$ , and  $t_3$  are selected from  $\{1, 2, 3\}$ , where  $t_1 \neq t_2 \neq t_3$ . By confirming whether the line  $(\mathbf{c}^{(t_1)}, \mathbf{c}^{(t_2, t_3)})$  from  $\mathbf{c}^{(t_1)}$  to the midpoint of every two centroids  $\mathbf{c}^{(t_2, t_3)} := (\mathbf{c}^{(t_2)} + \mathbf{c}^{(t_3)})/2$  is parallel with the major axis  $\mathbf{e}_1$  of the fitted ellipsoid by

$$\hat{t}_1, \hat{t}_2, \hat{t}_3 = \underset{t_1, t_2, t_3 \in \{1, 2, 3\}, t_1 \neq t_2 \neq t_3}{\operatorname{argmin}} \theta^{(t_1 - \overline{t_2, t_3})}, \quad (22)$$

the left and right eyes with the mouth can be distinguished from each other, where

$$\begin{aligned} \theta^{(t_1 - \overline{t_2, t_3})} &:= \angle(\mathbf{l}^{(t_1 - \overline{t_2, t_3})}, \mathbf{u}_1) \\ &= \arccos \frac{\mathbf{l}^{(t_1 - \overline{t_2, t_3})} \top \mathbf{u}_1}{\|\mathbf{l}^{(t_1 - \overline{t_2, t_3})}\| \|\mathbf{u}_1\|}, \\ \mathbf{l}^{(t_1 - \overline{t_2, t_3})} &:= \mathbf{c}^{(\overline{t_2, t_3})} - \mathbf{c}^{(t_1)}. \end{aligned} \quad (23)$$

Therefore, topological holes  $\mathcal{H}^{(\hat{t}_1)}$ ,  $\mathcal{H}^{(\hat{t}_2)}$  and  $\mathcal{H}^{(\hat{t}_3)}$  can be used to express  $\mathcal{H}^{le}$ ,  $\mathcal{H}^{re}$  for the eyes and  $\mathcal{H}^m$  for the mouth, respectively, which are marked in red in Figure 6. Moreover, the orbital thickness of the Haniwa resulted in many scanning depth errors. To suppress these errors, the point clouds of the mouth and eyes must be projected onto the fitted ellipsoid, as shown in Figure 7.

#### 4.4 Nose Segmentation

As illustrated in Figure 8, the convex parts in the face are marked in orange, where the nose expresses a convex part at the center of face; hence, the concave-convex portion of the facial parts based on the fitted ellipsoid can be applied to quickly segment the nose from the front head. However, in the actual discrimination, other convex parts (noise) exist that interfered with the nose segmentation. Therefore, some constraints must be added for noise suppression. Here, the most obvious constraint is that the nose is located within the triangle of the centroids of the

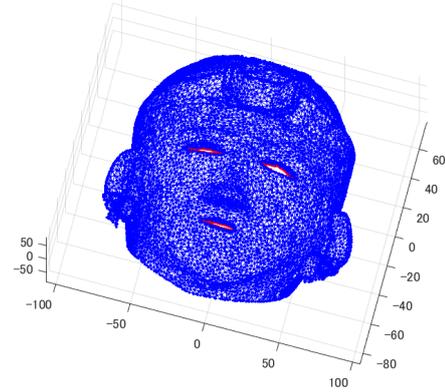


Figure 5: Contour labelling of the three largest topological holes.

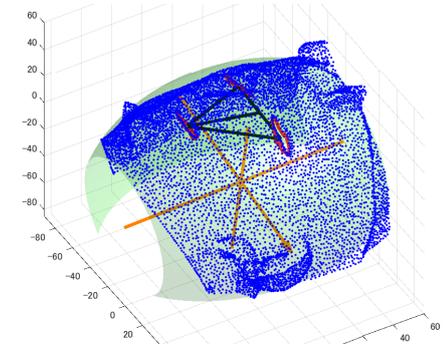


Figure 6: Confirmation of left and right eyes with mouth.

eyes and mouth considered as the vertices. This constraint can be achieved by calculating the angular distances as follows.

For any cloud point  $\mathbf{v}^\exists$  on the outside of the fitted ellipsoid, the vectors from it to the ellipsoid center  $\mathbf{c}$ , centroid of left eye  $\mathbf{c}^{le}$ , centroid of right eye  $\mathbf{c}^{re}$ , and centroid of mouth  $\mathbf{c}^m$  are defined by

$$\begin{aligned} \mathbf{l}^\exists &:= \mathbf{v}^\exists - \mathbf{c}, & \mathbf{l}^{le} &:= \mathbf{c}^{le} - \mathbf{c}, \\ \mathbf{l}^m &:= \mathbf{c}^m - \mathbf{c}, & \mathbf{l}^{re} &:= \mathbf{c}^{re} - \mathbf{c}. \end{aligned} \quad (24)$$

By calculating the angular distances of these vectors, i.e.,

$$\begin{aligned} \theta^{(le, m)} &:= \angle(\mathbf{l}^{le}, \mathbf{l}^m), & \theta^{(re, m)} &:= \angle(\mathbf{l}^{re}, \mathbf{l}^m), \\ \theta^{(\exists, le)} &:= \angle(\mathbf{l}^\exists, \mathbf{l}^{le}), & \theta^{(\exists, re)} &:= \angle(\mathbf{l}^\exists, \mathbf{l}^{re}), \\ \theta^{(\exists, m)} &:= \angle(\mathbf{l}^\exists, \mathbf{l}^m), \end{aligned} \quad (25)$$

the expected constraint can be obtained as fol-

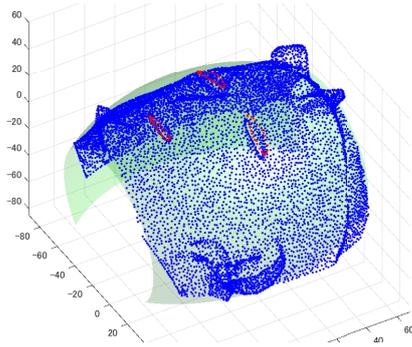


Figure 7: Point-cloud projection of eyes and mouth on face surface.

lows:

$$\begin{aligned} \theta(\exists, le) &\leq \theta^{(le, m)}, \quad \theta(\exists, re) \leq \theta^{(re, m)}, \\ \theta(\exists, m) &\leq \min(\theta^{(le, m)}, \theta^{(re, m)}), \end{aligned} \quad (26)$$

which enable us to locate and segment the nose by extracting its cloud points.

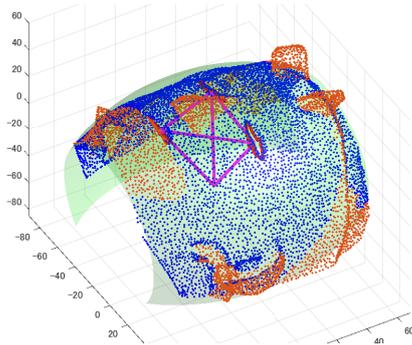


Figure 8: Nose segmentation by angle constraints.

## 5 Face Comparison

The ICP algorithm [9, 10] is employed to minimize the difference between two point clouds on the same scale, but it does not register two point clouds with different scales. Du et al. [8] introduced a novel approach known as the sICP algorithm, which integrates a scale matrix into the original ICP algorithm for scaling registration.

The point clouds of the Haniwa faces must be adjusted on the same scale before comparing the dissimilarity; this is because creating the Haniwa faces on the same scale is challenging even when

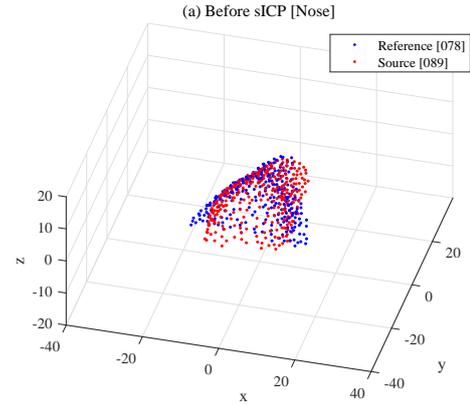


Figure 9: Point clouds of nose for reference and source before sICP.

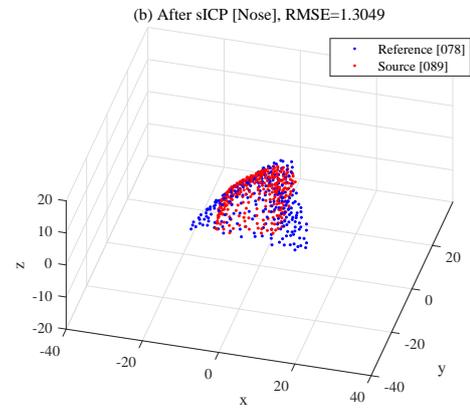


Figure 10: Point clouds of nose for reference and source after sICP.

the same maker is used when handmaking them. Meanwhile, the point cloud of the nose contains more points than those of the other parts. These points are not modified and are uniformly distributed even in depth; hence, they include three-dimensional information with higher aboriginality and precision<sup>2</sup>. Therefore, as shown in Figures 9 and 10, the sICP algorithm was first used for the point clouds of the nose to obtain their scale value. Subsequently, based on the scale value, two point clouds of all parts were adjusted on the same scale, as shown in Figure 11. Finally, the two scaled point clouds of all parts were regi-

<sup>2</sup>If the point clouds of all face parts such as those in [14] are used in sICP to obtain the scale, more errors will be generated in the following ICP compared with the scaling method based on only the nose.

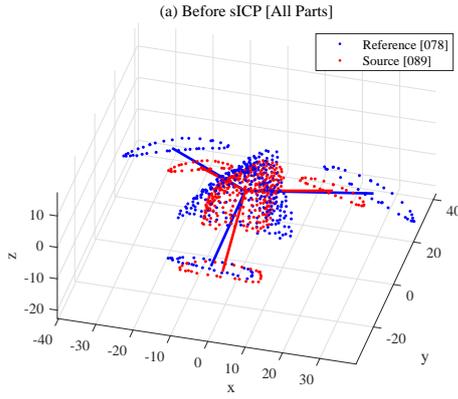


Figure 11: Point clouds of all parts for reference and source before ICP.

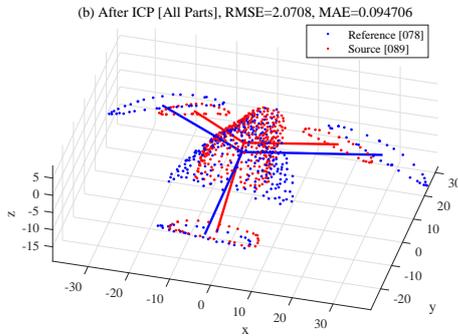


Figure 12: Point clouds of all parts for reference and source after ICP.

tered by the ICP algorithm, as shown in Figure 12.

Here, two metrics of face dissimilarity are proposed for the face comparison. Naturally, the RMSE of the ICP algorithm can be considered as one metric of face dissimilarity, although it is usually used for error evaluation. Moreover, the spatial relationship of all parts is also important to be compared, and thus the MAE of lines from nose centroid to the centroids of other face parts is regarded as another metric for facial parts arrangement dissimilarity. Here these lines are indicated in Figures 11 and 12.

## 6 Experiments

The proposed method was used to perform the face alignment and comparison for the six Haniwa models that were selected from the dataset [2] in

Figure 13.

Figures 14 and 15 show the experimental results of face alignment from two different viewpoints, separately. In these figures, the edges of the eyes and mouth are indicated in red, and the cloud points of nose are marked in orange. All of them indicate whether headwear exists for Models 1, 3, and 4, and whether the hair existed on the head top of Models 5 and 6. Our proposed face alignment was executed precisely without interference at all times. Therefore, high precision with good anti-interference ability can be considered as characteristic of our proposed method.

Table 1 shows the RMSE of the point clouds of all face parts and the MAE of lines from the nose centroid to the centroids of other face parts between every two models in Models 1–6 after the ICP. Figure 16 shows the registration results between Model 1 and the other models to express the face similarities measured based on the RMSE and MAE in three viewpoints.

It is obvious that Models 1 and 4 are more similar than others according to the first lowest MAE (0.0537 rad) and second lowest RMSE (1.7464), because their corresponding face parts basically coincide with each other. Meanwhile, the MAE of Models 1 and 3 (0.0947 rad) and that of Models 3 and 4 (0.0940 rad) are definitely close to that of Models 1 and 4. It denotes that there might exist a making rule to locate nose, mouth, and eyes in the Wazumi technique.

Because Models 2 and 5 have a low RMSE (1.5462), they can be considered as similar objects. In fact, Models 2 and 5 can be classified together with Models 1, 3, and 4 from the human perspective. However, Model 2 exhibited a larger nose, whereas Model 5 exhibited a broken one; therefore, the RMSE and MAE between them compared with those of other models were significant. In addition, Model 6 exhibited the largest nose with the highest bridge; therefore, its RMSE and MAE compared with those of the other models remained high.

Our proposed method was implemented in MATLAB using an Intel Core i3-4350 CPU and 8 GB RAM. The total processing time of the face alignment was approximately 1.1108 seconds on an organized cloud with 20,000 points with the

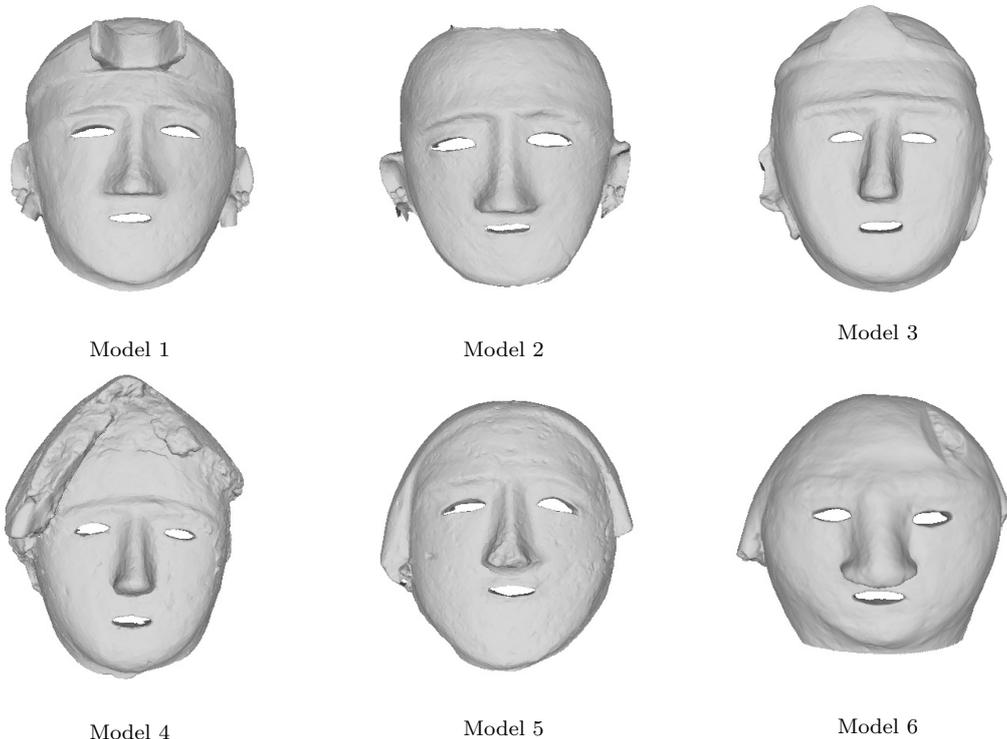


Figure 13: Three-dimensional models of the Haniwa.

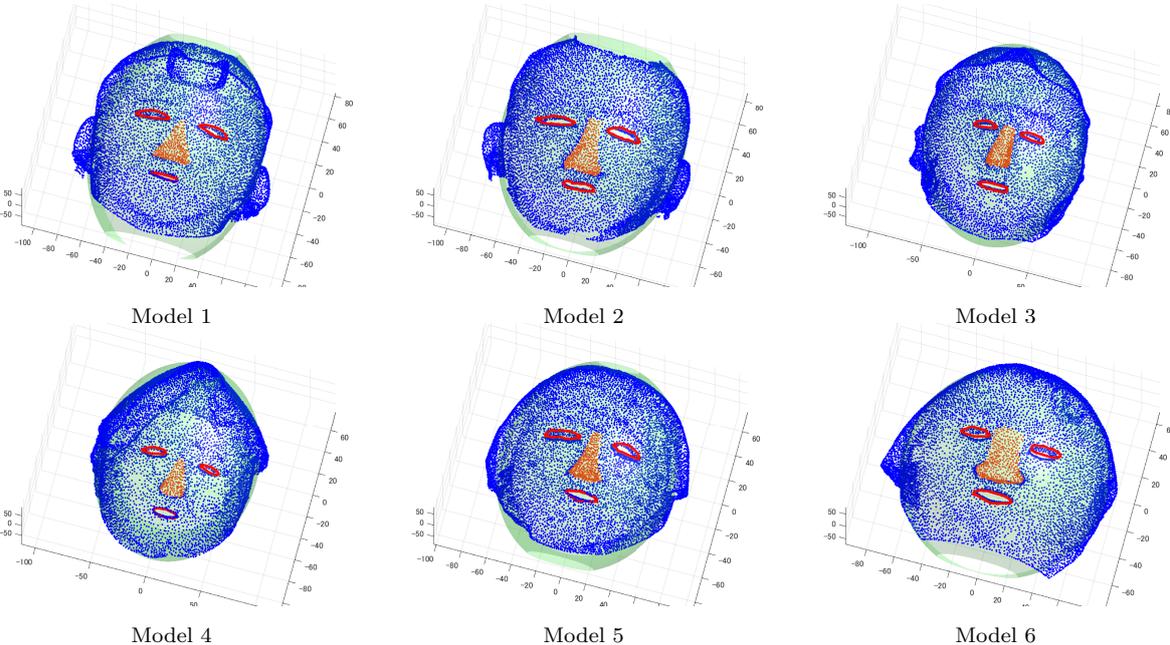


Figure 14: Face alignment results from first viewpoint.

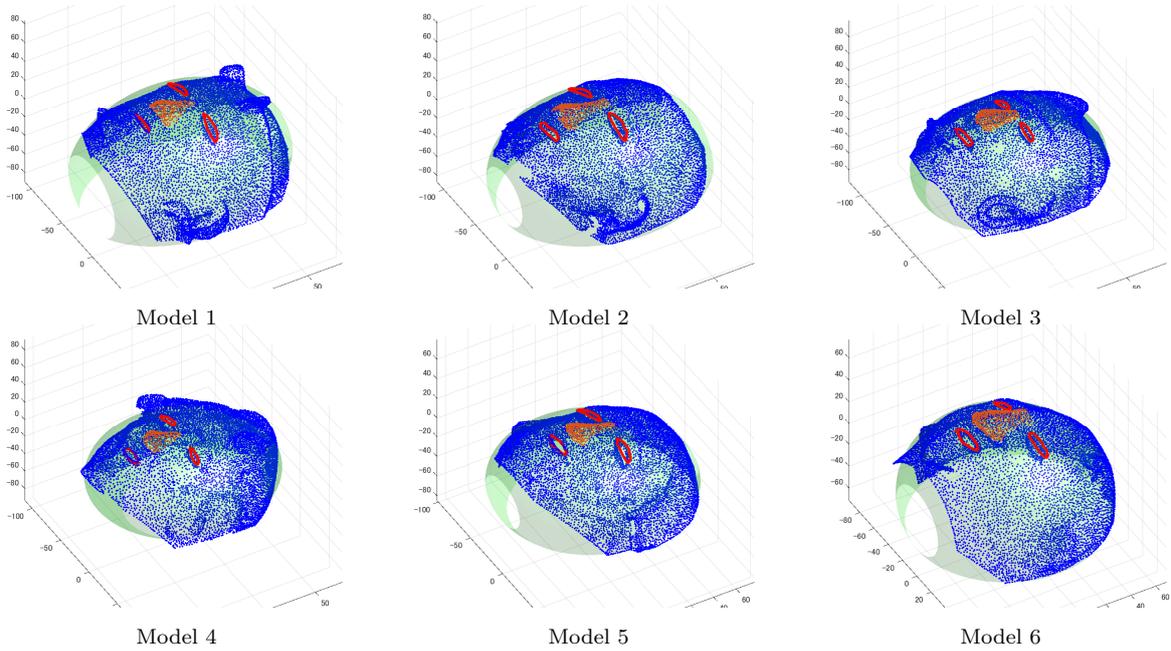


Figure 15: Face alignment results from second viewpoint.

Table 1: Root mean squared error (RMSE) of point clouds of all face parts, and mean angle error (MAE) of lines from nose centroid to centroids of other face parts between every two models in Models 1-6 after ICP. Lower-left: MAE in rad; upper-right: RMSE in millimeters.

		RMSE and MAE					
		Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
MAE	RMSE						
	Model 1			2.2535	2.0908	<b>1.7464</b>	2.2398
Model 2	0.1538			2.1311	2.6485	<b>1.5462</b>	2.8546
Model 3	0.0947	0.1280			2.0588	2.5971	2.3985
Model 4	<b>0.0537</b>	0.1295	0.0940			2.4905	3.1309
Model 5	0.1503	0.1260	0.0965	0.1278			2.7455
Model 6	0.1267	0.1346	0.1223	0.0829	0.1109		

corresponding meshes. The processing time for the face surface estimation, contour labeling of the eyes and mouth, and nose segmentation were 0.7185, 0.3510, and 0.0413 seconds, respectively. Moreover, the total processing time taken to compare the two faces was approximately 0.1338 sec-

onds. The processing times of the sICP and ICP, and the computation of the MAE were 0.1007, 0.0132, and 0.0199 seconds, respectively.

## 7 Conclusion

A hybrid method combining ellipsoid-fitting-based segmentation and topological hole detection, has been proposed to perform the three-dimensional face alignment of the Haniwa with high precision. Furthermore, the similarity between Haniwa faces was evaluated digitally for the first time in this study based on the point-cloud-based metrics. Our experimental results proved the archaeological assumption that the maker located the nose, mouth, and eyes of the Haniwa at fixed angles. Archaeologists have confirmed the proposed method is valuable based on the experimental results. In our future studies, we will perform face alignment and comparison using more Haniwa figurines, including other types, e.g., those wearing a helmet or with a vertically long nose. Based on the comparison results, we will analyze their artistic values and provenances with datings.

## 8 Acknowledgements

The authors would like to thank the Shibayama Town Board of Education and Kannon Temple for providing the data. The authors would like to take this opportunity to thank Professor Jokura Masayoshi for his professional advice and valuable support. He is an East Asian archaeologist in Faculty of Letters, Arts and Sciences, Waseda University. The authors would also like to thank the anonymous reviewers for their helpful and constructive comments that contributed significantly in improving the manuscript, as well as the editors for their generous comments and support during the review process. Part of this study was supported by JSPS KAKENHI (Grant Number JP18H00734).

## References

- [1] Xin Lu, Chunyuan Li, Tsutomu Kinoshita, and Konno Kouichi. A Study on 3D Face Similarity by Point Cloud Based Metric for Japanese Terracotta Figurines (Haniwa). Society for Art and Science (NICOGRAPH) Conference Series, 2019.
- [2] Masayoshi Jokura. *A study of Tonozuka-Himetsuka ancient tomb*. Rokuichi-Syobou, Tokyo, JP, 1st edition, 2017 (in Japanese).
- [3] Iain Matthews and Simon Baker. Active appearance models revisited. *International Journal of Computer Vision*, 60(2):135–164, 2004.
- [4] David Cristinacce and Tim Cootes. Automatic feature localisation with constrained local models. *Pattern Recognition*, 41(10):3054–3067, 2008.
- [5] Xudong Cao, Yichen Wei, Fang Wen, and Jian Sun. Face alignment by explicit shape regression. *International Journal of Computer Vision*, 107(2):177–190, 2014.
- [6] Chen Cao, Yanlin Weng, Shun Zhou, Yiyong Tong, and Kun Zhou. FaceWarehouse: A 3d facial expression database for visual computing. *IEEE Transactions on Visualization and Computer Graphics*, 20(3):413–425, 2014.
- [7] Naoki Degawa, Xin Lu, and Akio Kimura. A performance improvement of Mask R-CNN using region proposal expansion. volume 11049 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 1104929, 2019.
- [8] Shaoyi Du, Nanning Zheng, Lei Xiong, Shihui Ying, and Jianru Xue. Scaling iterative closest point algorithm for registration of m-d point sets. *Journal of Visual Communication and Image Representation*, 21(5-6):442–452, 2010.
- [9] Paul J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [10] Yang Chen and Gérard Medioni. Object modelling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

- [11] A. Fitzgibbon, M. Pilu, and R.B. Fisher. Direct least square fitting of ellipse. *The IEEE transactions on Pattern Analysis and Machine Intelligence*, 21(5):476–480, 1999.
- [12] R. Halir and J. Flusser. Numerically stable direct least squares fitting of ellipses. In *Proceedings of International Conference in Central Europe on Computer Graphics and Visualization Visualization*, pages 125–132, 1998.
- [13] R.C. Gunning. *An Introduction to Analysis*. Princeton University Press, 2018.
- [14] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *Proceedings of International Conference on Computer Vision (ICCV)*, December 2015.

#### Xin Lu



Xin Lu is an assistant professor of Faculty of Science and Engineering at Iwate University. He received the B.E. degree in computer science and information technology from Hohai University, China, in 2000. He earned the M.E. and Dr.Eng. degrees in information science from the University of Tokushima in 2003 and 2006, respectively. In 2006, he joined Iwate University. His current research interests include statistics, machine learning, and signal processing. He is a member of IEEE, IEICE, and SICE.

#### Chunyuan Li



Chunyuan Li is a lecturers at Beihua University in China. She received M.E. and Dr.Eng. degrees in computer science from Iwate University in 2012 and 2017, respectively. She worked on a researcher at Iwate University from 2017 to 2018. Her current research interests include computer graphics and interactive systems.

#### Tsutomu Kinoshita



Tsutomu Kinoshita is an associate professor of Faculty of Engineering at Tohoku Gakuin University. He received a BS in Mathematics in 1993 from the Tokyo University of Science. He earned a Dr.Eng. in computer science from Iwate University in 2013. He joined the 3D geometry handling and data translation project at TOYOTA from 1993 to 1998, and the XVL project at Lattice Technology from 2001 to 2014. He worked on an associate professor of Faculty of Environmental and Information Sciences at Fukui University of Technology from 2015 to 2017. His research interests include geometric modeling and computer graphics. He is a member of the Society for Art and Science.

### Akio Kimura



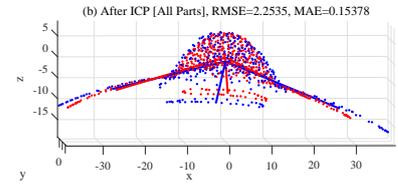
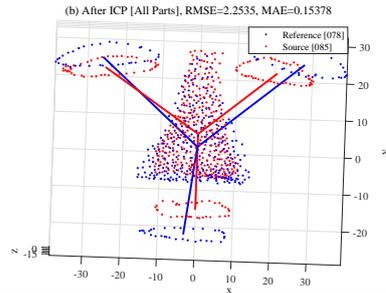
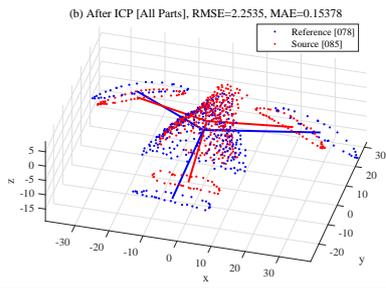
Akio Kimura completed master's program of Computer and Information Sciences, Graduate School of Engineering, Iwate University in 1993, and joined Sony Corporation. While at Sony, he was engaged in research and development of magnetic recording. In 1995, he joined Iwate university as an assistant professor, and is now an associate professor of Faculty of Science and Engineering. He is engaged in research related to image processing, computer vision, and machine learning. He holds Dr.Eng. degree, and is a member of IEICE, IPSJ, and IIEEJ.

### Kouichi Konno

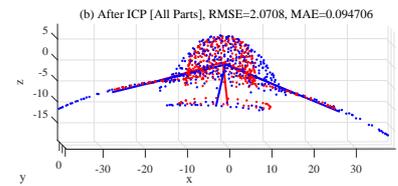
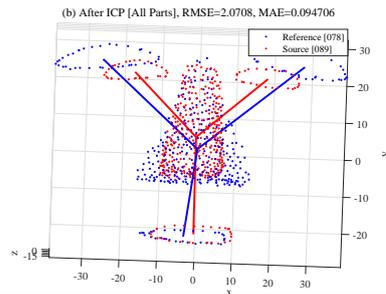
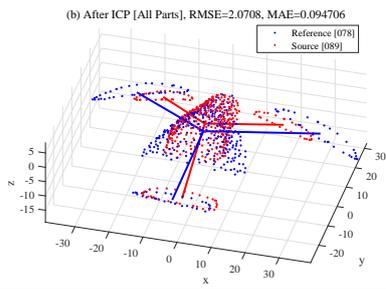


Kouichi Konno is a professor of Faculty of Engineering at Iwate University. He received a BS in Information Science in 1985 from the University of Tsukuba. He earned his Dr.Eng. in precision machinery engineering from the University of Tokyo in 1996. He joined the solid modeling project at RICOH from 1985 to 1999, and the XVL project at Lattice Technology in 2000. He worked on an associate professor of Faculty of Engineering at Iwate University from 2001 to 2009. His research interests include virtual reality, geometric modeling, 3D measurement systems, and computer graphics. He is a member of EuroGraphics.

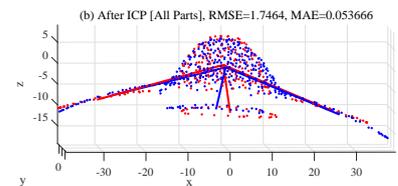
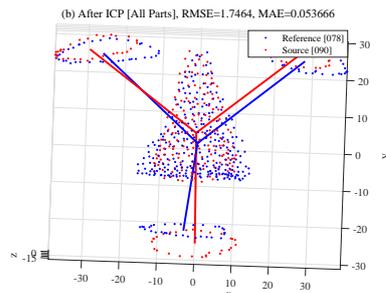
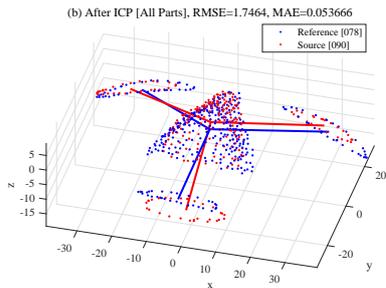
Models 1 & 2



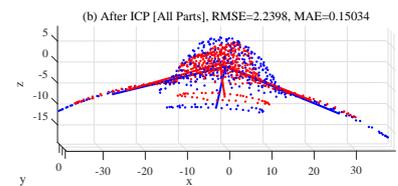
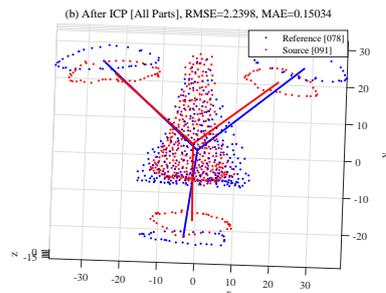
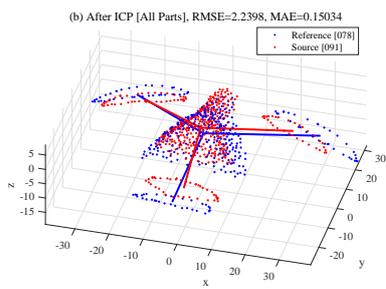
Models 1 & 3



Models 1 & 4



Models 1 & 5



Models 1 & 6

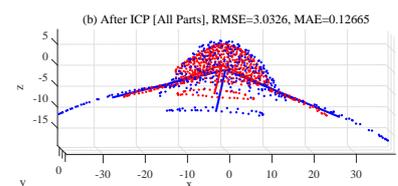
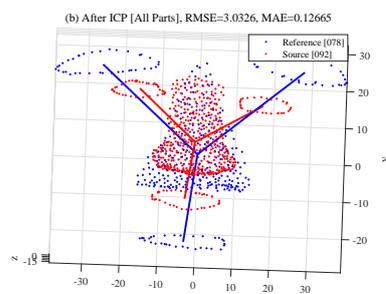
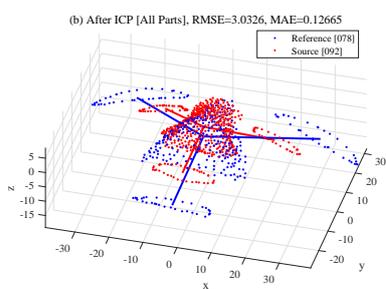


Figure 16: Results of ICP among different models from three viewpoints.