

Groupwise Pose Normalization for Craniofacial Applications

Jiun-Hung Chen and Linda G. Shapiro
University of Washington
Seattle, WA 98195

{jhchen, shapiro}@cs.washington.edu

Abstract

A general framework is proposed for solving groupwise pose normalization problems and is analyzed in detail under different feature spaces. The analysis shows that using principal component analysis for pose normalization is a special case of using the proposed framework under a special feature space. The experimental results on two craniofacial datasets show the proposed method achieved promising results for solving groupwise pose normalization problems for craniofacial applications.

1. Introduction

3D shape analysis is an important tool for both recognition and retrieval of 3D shapes [19]. If two 3D shapes are members of the same class of objects, their comparison is simplified if they are both canonical views of their class. For example, if the shapes being compared are human faces, it is common and natural to have them both facing forward. For other shape classes, such as kitchenware or furniture, there may be no natural canonical pose, but it is possible and desirable to define such a pose for purpose of analysis. Given a 3D shape of a particular class, *pose normalization* is the process of applying a 3D transformation to the shape to transform it to a canonical pose for its shape class.

Most algorithms for pose normalization operate on a single 3D shape or try to bring two shapes into alignment [19][22][9][13]. Single-object pose normalization is used in 3D object retrieval, in which all models are considered separately since the group of an unseen object is unknown. Two-object pose normalization is used for alignment tasks. In some analysis tasks, particularly in the medical domain, a set of 3D models of a single class is provided and the analysis requires that all of those models be aligned in a single canonical pose. The methodology proposed here is for the last case and is motivated by our work on 3D craniofacial image analysis. Our data are sets of 3D meshes of children’s heads from a 3dMD®12-camera stereo imaging system. The children in the study have one or more cranio-

facial abnormalities such as midface flattening, cleft lip, or cleft palate. Because many of them are infants or toddlers, they cannot be expected to sit still for long or to be able to keep their heads in a specified pose. In order to analyze the abnormalities and compare individuals, we must first pose normalize the meshes so that all the heads face in the same direction.

Thus, the problem of interest in this paper is *groupwise pose normalization*: how to *jointly* pose normalize a group of 3D shape models of the same class. In this paper, the groupwise pose normalization problem is formulated as an optimization problem, and a gradient descent approach is proposed solving the optimization problem.

There are three main contributions to our work. First, the groupwise pose normalization problem is formally defined. Second, a general framework for solving this problem is proposed and is analyzed to show that using principal component analysis (PCA) for pose normalization [22] is a special case of using the proposed framework under a particular feature space. Third, the framework is applied to the problem of pose normalization of 3D meshes of children’s heads and experimental results are given.

2. Related Work

We will discuss related work in pose normalization as well as groupwise techniques for 3D data analysis.

2.1. Pose normalization

Let $X = (V, E)$ denote a triangular mesh model of a 3D shape, where V and E are the sets of vertices and edges in the mesh model, respectively. Vranic and Saube [22] proposed to use principal components analysis (PCA), a common approach to modeling shape, for pose normalization. Given the mesh X , PCA computes its scatter matrix S and finds a projection axis b that maximizes $b^t S b$. Intuitively, the total scatter of the projected samples is maximized after the projection of the samples onto b . The optimal Q projection axes b_q , $q = 1, \dots, Q$ that maximize the above criterion are the eigenvectors of S corresponding to the largest

Q eigenvalues, $\{\lambda_q | q = 1, \dots, Q\}$. The three eigenvectors are then chosen as the resultant x,y,z axes, respectively.

While PCA is a standard, easy-to-use approach, it is not effective in general shape recognition and retrieval problems [19]. For this reason, several different PCA variants [23][15][14][11] have been proposed to deal with the problems faced by basic PCA for pose normalization. Other non-PCA approaches include a rectilinearity measure [10] and a symmetry-based measure [1].

While most pose normalization algorithms operate on a single 3D object, some methods have been proposed for aligning two 3D models. Kazhdan [9] used a spherical function to represent a 3D shape model and minimized the L2-distance between pairs shapes. Martinek and Grosso [13] used an image-like representation to store depth information from all perspectives for a 3D shape and a weighted ratio of the intersection to the union of two images as a similarity measure [13]. Specialized pose normalization approaches exist for specific 3D shape classes such as faces and human bodies [20][7]. To the best of our knowledge, groupwise pose normalization problems have neither been formally defined nor solved.

2.2. Groupwise Techniques for 3D Data

Groupwise approaches have been applied to several other 3D data analysis tasks including point registration[4][24], point correspondence [5], image registration [12] and Procrustes analysis [16][6]. To avoid solving point correspondence problems, one class of groupwise point-set registration approach [4][24] models point sets as probability distributions and uses information theoretic measures [4][24] to determine how similar a group of probability distributions are after transformation.

There are several important contrasts between groupwise *point-set registration* and groupwise *pose normalization*. First, groupwise point-set registration focuses on point sets (2D/3D point clouds), while groupwise pose normalization focuses on shape models (2D/3D meshes). Richer information including oriented position, descriptors such as the light field descriptor [2], and geometric properties such as surface-normal angles and curvatures, can be taken into consideration when working with 3D shape models.

Davies *et al.* [5] proposed an information theoretic MDL-based objective function to quantize the quality of the point correspondences. A simplified version G proposed by Thodberg [21] as defined below is commonly-used.

$$G = \sum_{k=1}^N L_k \text{ with } L_k = \begin{cases} 1 + \log(\lambda_k/\lambda_{cut}), & \text{if } \lambda_k \geq \lambda_{cut} \\ \lambda_k/\lambda_{cut}, & \text{otherwise} \end{cases} \quad (1)$$

In [5], given a set of computed point correspondences among a set of shapes, PCA is computed on the set of point correspondences, and the computed eigenvalues, $\{\lambda_k | k =$

$1, \dots, N\}$, are used to calculate G in (1). λ_{cut} is a parameter that determines the point at which to effectively switch between the determinant-type term (the if-part in (1)) and the trace-type term (the otherwise-part in (1)). The determinant-type terms jointly measure the volume of the training set after correspondence in shape space, which favors compactness. The trace-type terms jointly measure similarity of each pair of the training shapes after correspondence via Euclidean distance. The point correspondences of the i th shape are assumed to be controlled by some parameter vector β_i , for which the individual parameters are given by $\{\beta_{i,a} | a = 1, \dots, A\}$. The gradient descent approach is used to minimize G with respect to a parameter vector β_i . The Jacobian matrix for the gradient of the objective function is defined in [8]:

$$\frac{\partial G}{\partial \beta_{i,a}} = \sum_{k=1}^N \frac{\partial L_k}{\partial \lambda_k} \frac{\partial \lambda_k}{\partial \beta_{i,a}} \quad (2)$$

It is easy to compute $\frac{\partial L_k}{\partial \lambda_k}$ (see (1)) and so we focus on $\frac{\partial \lambda_k}{\partial \beta_{i,a}}$ in the following discussions. $\frac{\partial \lambda_k}{\partial \beta_{i,a}}$ can be computed by using the following chain rule for derivatives.

$$\frac{\partial \lambda_k}{\partial \beta_{i,a}} = \frac{\partial \lambda_k}{\partial p_i} \frac{\partial p_i}{\partial \beta_{i,a}} \quad (3)$$

where p_i is a vector which contains the positions of corresponding points on the i -th mesh.

While $\frac{\partial p_i}{\partial \beta_{i,a}}$ is typically computed by using finite differences, the following analytic form for $\frac{\partial \lambda_k}{\partial p_i}$ exists:

$$\frac{\partial \lambda_k}{\partial p_i} = 2(1 - \frac{1}{N})c_{i,k}b_k. \quad (4)$$

where $c_{i,k}$ is the projection coefficient of the i -th position vector p_i onto the k -th eigenvector b_k .

Chen, Zheng and Shapiro [3] further generalized the MDL framework [5] to feature spaces and proved that the Davies framework is a special case of their proposed framework. Although the MDL framework [5] and its generalization [3] have been used in point correspondence and image registration [12], to the best of our knowledge, this class of framework has not been applied for solving groupwise pose normalization problems.

3. Groupwise Pose Normalization

In this section we formulate the groupwise pose normalization problem and both describe and analyze our approach.

3.1. Problem Formulation

Assume that Ψ is the space of all triangular mesh models of 3D shapes. A group of 3D shape models of the same

class will be jointly pose normalized, so that the shapes can be more effectively compared.

Let Γ be the space of all transformations $T : \Psi \rightarrow \Psi$. Let Ω be a feature space and $\phi : \Psi \rightarrow \Omega$ be a mapping from a mesh to a feature representation. Let $S = \{X_i \in \Psi | i = 1, \dots, N\}$ be a set of N 3D shape models of a particular class. Let $F : \Omega^N \rightarrow R$ be a groupwise shape dissimilarity function that measures the quality of pose normalization for a set of models. The groupwise pose normalization problem is to find the set of transformations T^* that minimize the value of function F as defined below.

$$T^* = \arg \min_{\{T_i \in \Gamma | i=1, \dots, N\}} F(\{\phi(T_i(X_i)) | i = 1, \dots, N\}) \quad (5)$$

3.2. Proposed Method

To solve (5), we will adapt the general MDL framework of [3] to groupwise pose normalization problems. We will restrict Γ to be the space of rigid transformations in 3D space. While the framework of [3] allowed an arbitrary reproducing kernel Hilbert space (RKHS) Ω [18], in this work we will consider Ω to be a space of sets of local features, in which PCA and the framework of [3] cannot be applied directly. In other words, the function $\phi : \Psi \rightarrow \Omega$ will map a mesh to a set of local features that belongs to Υ , the space of local features vectors.

For the dissimilarity function F , we will adapt (1), which was also used in [3], from the Davies *et al.* work [5]. Because PCA cannot be performed on Ω , the space of sets of local features, it is instead performed on Υ , the space of local features, so that the resultant eigenvalues can be used for computing the value of F . To elaborate, PCA is performed on $\Delta = \{f \in \phi(T_i(X_i)) | i = 1, \dots, N\}$, the union of the sets, $\{\phi(T_i(X_i)) | i = 1, \dots, N\}$ that is the input to (5).

The intuition is that the better the local features are represented by PCA on Υ , the better the resultant meshes are pose normalized.

The proposed objective function is defined below.

$$F = \sum_{k=1}^d L_k \text{ with } L_k = \begin{cases} 1 + \log(\lambda_k/\lambda_{cut}), & \text{if } \lambda_k \geq \lambda_{cut} \\ \lambda_k/\lambda_{cut}, & \text{otherwise} \end{cases} \quad (6)$$

where d is the minimum of N and the dimension of a local feature vector in Υ and $\{\lambda_k | k = 1, \dots, d\}$ are the eigenvalues obtained from PCA performed on Δ .

The set of local feature vectors of the i th model is assumed to be controlled by some transformation parameter vector α_i , for which the individual parameters are given by $\{\alpha_{i,a} | a = 1, \dots, A\}$. The gradient descent approach is proposed to minimize F with respect to a parameter vector α_i . The Jacobian matrix for the gradient of the objective function is defined as

$$\frac{\partial F}{\partial \alpha_{i,a}} = \sum_{k=1}^d \frac{\partial L_k}{\partial \lambda_k} \frac{\partial \lambda_k}{\partial \alpha_{i,a}} \quad (7)$$

In contrast with (3), $\frac{\partial \lambda_k}{\partial \alpha_{i,a}}$ can be computed by using the following chain rule for derivatives

$$\frac{\partial \lambda_k}{\partial \alpha_{i,a}} = \sum_{f \in \phi(T_i(X_i))} \frac{\partial \lambda_k}{\partial f} \frac{\partial f}{\partial \alpha_{i,a}} \quad (8)$$

While $\frac{\partial f}{\partial \alpha_{i,a}}$ is typically computed by using finite differences, the following analytic form for $\frac{\partial \lambda_k}{\partial f}$ exists:

$$\frac{\partial \lambda_k}{\partial f} = 2(1 - \frac{1}{M})cb_k. \quad (9)$$

where c is the projection coefficient of f onto the k -th eigenvector b_k and M is the cardinality of Δ .

Note that if Υ is a reproducing kernel Hilbert space (RKHS), (6)-(9) can be generalized with Mercer kernels [18] by using the techniques developed in [3]. In the following, two types of local features, positions and rotation variant local features, are considered. Unlike using finite differences to approximate the gradient $\frac{\partial f}{\partial \alpha_{i,a}}$, the analytic formulas for computing the gradient $\frac{\partial f}{\partial \alpha_{i,a}}$ can be derived for these two types of local features.

3.2.1 Positions as Local Features:

Consider the 3D transformation $T = \{t_x, t_y, t_z, q_1, q_2, q_3\}$ where t_x, t_y, t_z are the three translation parameters to represent a translation vector, $t = [t_x \ t_y \ t_z]^t$, and q_1, q_2, q_3 are the classical Rodrigues parameters to represent a rotation matrix, R (see (10)) [17]. Let x be a position vector before application of the transformation T and $x' = R(x - t)$ be the corresponding position vector after the transformation. The analytic formulas for computing derivatives with respect to translation and rotation are given in (11) to (15).

3.2.2 Rotation-Variant Local Features

Consider a feature map f_m that maps each vertex in a mesh to a feature vector. For example, f_m might map each vertex v to a vector n_v , the normal to the tangent plane at v . Assume that feature vectors are invariant with respect to translation and scale and change via a rotation matrix with the following rule, $l' = Rl$, where l is a feature vector before the transformation and l' is the corresponding feature vector after the transformation. Then the analytic formulas for computing gradients with respect to rotations are:

$$R = \frac{1}{1 + q_1^2 + q_2^2 + q_3^2} \begin{bmatrix} 1 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_3) & 2(q_1q_3 - q_2) \\ 2(q_2q_1 - q_3) & 1 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_1) \\ 2(q_3q_1 + q_2) & 2(q_3q_2 - q_1) & 1 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (10)$$

$$\frac{\partial R}{\partial q_1} = \frac{1}{1 + q_1^2 + q_2^2 + q_3^2} \begin{bmatrix} 2q_1 & 2q_2 & 2q_3 \\ 2q_2 & -2q_1 & 2 \\ 2q_3 & -2 & -2q_1 \end{bmatrix} - \frac{2q_1}{1 + q_1^2 + q_2^2 + q_3^2} R \quad (11)$$

$$\frac{\partial R}{\partial q_2} = \frac{1}{1 + q_1^2 + q_2^2 + q_3^2} \begin{bmatrix} -2q_2 & 2q_1 & -2 \\ 2q_1 & 2q_2 & 2q_3 \\ 2 & 2q_3 & -2q_2 \end{bmatrix} - \frac{2q_2}{1 + q_1^2 + q_2^2 + q_3^2} R \quad (12)$$

$$\frac{\partial R}{\partial q_3} = \frac{1}{1 + q_1^2 + q_2^2 + q_3^2} \begin{bmatrix} -2q_3 & 2 & 2q_1 \\ -2 & -2q_3 & 2q_2 \\ 2q_1 & 2q_2 & 2q_3 \end{bmatrix} - \frac{2q_3}{1 + q_1^2 + q_2^2 + q_3^2} R \quad (13)$$

$$\begin{aligned} \frac{\partial x'}{\partial q_1} &= \frac{\partial R}{\partial q_1}(x - t) \\ \frac{\partial x'}{\partial q_2} &= \frac{\partial R}{\partial q_2}(x - t) \\ \frac{\partial x'}{\partial q_3} &= \frac{\partial R}{\partial q_3}(x - t) \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\partial x'}{\partial t_x} &= R[-1 \ 0 \ 0]^t \\ \frac{\partial x'}{\partial t_y} &= R[0 \ -1 \ 0]^t \\ \frac{\partial x'}{\partial t_z} &= R[0 \ 0 \ -1]^t \end{aligned} \quad (15)$$

$$\begin{aligned} \frac{\partial l'}{\partial \theta_x} &= \frac{\partial R}{\partial \theta_x} l \\ \frac{\partial l'}{\partial \theta_y} &= \frac{\partial R}{\partial \theta_y} l \\ \frac{\partial l'}{\partial \theta_z} &= \frac{\partial R}{\partial \theta_z} l \end{aligned} \quad (16)$$

3.2.3 Relation to PCA

If $N = 1$, $\phi(X) = V$, $\Upsilon = R^3$, eqn. (6) is used as the objective function, and λ_{cut} is a number that is larger than the largest eigenvalue from the PCA analysis on V , then eqn. (5) degenerates to the original PCA problem. This relation

shows that the proposed objective function is a general objective function to minimize and (5) is a general framework.

3.3. Groupwise Pose Normalization by Example

If additional prior information about the shape class is known, it is necessary to add this information to the groupwise pose-normalization problems. For example, some pose-normalized examples can be manually provided. In that case, PCA can be replaced by weighted PCA and the pose-normalized examples can be given more weight than the remaining 3D models.

Consider $F_w : \Omega^N \times R^N \rightarrow R$ as an objective function for groupwise pose normalization problems.

$$T^* = \arg \min F_w(\{\phi(T_i(X_i)), \kappa_i | i = 1, \dots, N\}) \quad (17)$$

If (6) is used as the objective function to minimize, in

contrast with (8), the derivatives have the following form.

$$\frac{\partial \lambda_k}{\partial \alpha_{i,a}} = \kappa_i^2 \sum_{f \in \phi(T_i(X_i))} \frac{\partial \lambda_k}{\partial f} \frac{\partial f}{\partial \alpha_{i,a}} \quad (18)$$

In other words, the derivatives for the weighted version are the corresponding ones in (8) multiplied by κ_i^2 .

4. Experimental Results and Discussions

Two separate sets of 3D craniofacial data for studying craniofacial abnormalities were used in our pose normalization tests. One set of 40 heads is for studying midface flattening (M set) and the other set of 19 heads is for studying cleft lip or cleft palate (C set). The heads in these two sets have been interactively pose normalized and thus serve as ground truth for our experiments. Some examples from the two sets can be found in the first two rows in Figures 1 and 2, respectively¹.

Given a set of heads, a random test set is created by randomly rotating each head with respect to its mass center. For each random test set, the proposed method is used to solve the pose normalization problem and a result set is generated. The errors between the ground truth set and the result set are measured in terms of degrees of rotations. In addition, the mean of the standard deviations along the three resultant main axes is used for measuring how good the pose normalization result is. A total of 4 random test sets is generated with different degrees of rotation and the means and standard deviations of the errors are recorded for comparisons using a 5-fold cross validation-like method (i.e., the examples in a fold are used as reference examples and the remaining folds are used as a test set for pose normalization).

The first experiment is to study how different features and reference examples affect the performances of the proposed method. Two different features, 3D positions and normals, are compared and reference examples are selected as pose-normalized heads that can be used to constrain the groupwise pose normalization to force it to choose a standard pose for craniofacial analysis². A standard PCA approach is also implemented for comparisons.

The last five rows in Figure 1 and 2 show examples of the pose normalized results for the two sets with different features. Table 1 shows the detailed error comparisons. The figures and table show that the performances with normals as features are better than those with positions as features, and that the PCA performance is the worst. One reason for the weak performance of the positions as features is that

¹To follow the IRB protocol to "deidentify" the images, the eyes were blackened out.

²Total 4 combinations, positions with references (P/w), positions without references (P/w.o), normals with references (N/w) and normals without references (N/w.o) are tested.

the heads are not scaled, since scaling is not generally used in craniofacial analysis. In addition, the performances with the reference examples are generally better than those without the reference examples, and the improvements are larger with normal features than those with position features.

The second experiment is to investigate how the value of λ_{cut} affects the pose normalization performance. A fold selected from the random datasets (M1 and C1) is selected as a reference set, and the remaining folds are used for pose normalization. Different values of λ_{cut} are tested for both features, and the experimental results are shown in Figure 3. The performance generally decreases with the values of λ_{cut} for normal features, and the ranges that achieve the best performance is $[2^{-5} \ 2^{-10}]$. In contrast, no clear trends between the performance and the values of λ_{cut} are observed for position features.

Although the two sets of experimental results show the promising performances of the proposed method, there is still room for improvement. One weakness is that the number of normal vectors of different heads may not be equal, which results in a bias toward a head with a higher number of normal vectors. How to select part of the normal vectors of a head and weight them differently will be part of our future work. From the experimental results, it is observed that different sets of reference examples can affect the groupwise pose normalization performance since different sets of examples can capture different prior information. Hence, how to select (or construct) representative examples for pose normalization will be an important future direction.

5. Conclusions and Future Work

Groupwise pose normalization problems are formally formulated in this paper. A general framework is proposed for solving groupwise pose normalization problems. It is shown that using PCA for pose normalization is a special case of using the proposed framework under a special feature space. From the experimental results, the proposed method achieved better performances than the PCA approach to which it was compared.

In addition to how to select normal vectors differently and how to select the representative examples, another important future direction is to customize the proposed method for specific objects and to take different geometric and shape information such as curvature and torsion into consideration. In contrast with using the MDL-based objective function, finding a better objective function also merits future research.

Acknowledgment

This work was supported by NIH/NIDCR Grant No. 1U01DE020050.

Table 1: Performance comparisons. The three numbers in each slot in order are the means and standard deviations of the errors with respect to the ground truth set and the mean of the standard deviations along three resultant main axes.

Set	P/w.o	P/w	N/w.o	N/w	PCA
M1	9.99(5.29),5.39	8.25(5.09),5.47	9.91(4.06),4.09	5.86 (4.04),5.02	42.70(26.20),35.04
M2	11.12(7.40),9.13	10.14(7.39),9.42	10.64(5.99),7.02	6.52 (5.73),7.60	43.61(29.33),47.53
C1	8.56(5.85),6.55	9.05(6.66),7.94	8.00(3.81),3.93	5.39 (3.25),4.52	50.24(25.06),51.72
C2	15.64(8.53),9.91	15.30(8.83),10.78	14.67(6.40),7.01	9.50 (6.05),8.26	50.46(22.54),49.73

References

- [1] M. Chaouch and A. Verroust-Blondet. Alignment of 3D models. *Graphical Models*, 71(2):63–76, 2009. 2
- [2] D. Chen, X. Tian, and M. Yu-Te Shen. On visual similarity based 3D model retrieval. In *Computer graphics forum*, volume 22, pages 223–232, 2003. 2
- [3] J.-H. Chen, K. C. Zheng, and L. G. Shapiro. 3d point correspondence by minimum description length in feature space. In *ECCV*, 2010. 2, 3
- [4] T. Chen, B. C. Vemuri, A. Rangarajan, and S. J. Eisenschenk. Group-wise point-set registration using a novel cdf-based havrda-charvát divergence. *IJCV*, 86(1):111–124, 2010. 2
- [5] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor. A minimum description length approach to statistical shape modeling. *IEEE TMI*, 21(5):525–537, 2002. 2, 3
- [6] C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):285–339, 1991. 2
- [7] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, and H. Seidel. A statistical model of human pose and body shape. In *Computer Graphics Forum*, volume 28, pages 337–346. Blackwell Publishing, 2009. 2
- [8] T. Heimann, I. Wolf, T. G. Williams, and H.-P. Meinzer. 3d active shape models using gradient descent optimization of description length. In *IPMI*, pages 566–577, 2005. 2
- [9] M. Kazhdan. An approximate and efficient method for optimal rotation alignment of 3d models. *PAMI*, pages 1221–1229, 2007. 1, 2
- [10] Z. Lian, P. Rosin, and X. Sun. Rectilinearity of 3D meshes. *IJCV*, 89(2):130–151, 2010. 2
- [11] H. Liu, J. Yan, and D. Zhang. What is wrong with mesh PCA in coordinate direction normalization. *Pattern Recognition*, 39(11):2244–2247, 2006. 2
- [12] S. Marsland, C. Twining, and C. Taylor. A minimum description length objective function for groupwise non-rigid image registration. *IVC*, 26(3):333–346, 2008. 2
- [13] M. Martinek and R. Grosso. Optimal rotation alignment of 3D objects using a GPU-based similarity function. *Computers & Graphics*, 33(3):291–298, 2009. 1, 2
- [14] M. Novotni and R. Klein. A geometric approach to 3D object comparison. In *SMI*, pages 167–175, 2001. 2
- [15] E. Paquet, M. Rioux, A. Murching, and T. Naveen. Description of shape information for 2-D and 3-D objects. *SPIC*, 16(1-2):103–122, 2000. 2
- [16] F. Rohlf and D. Slice. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology*, 39(1):40, 1990. 2
- [17] H. Schaub and J. Junkins. *Analytical mechanics of space systems*. Aiaa, 2003. 3
- [18] B. Scholkopf and A. Smola. *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. the MIT Press, 2002. 3
- [19] J. Tangelder and R. Veltkamp. A survey of content based 3D shape retrieval methods. *Multimedia Tools and Applications*, 39(3):441–471, 2008. 1, 2
- [20] F. B. ter Haar and R. C. Veltkamp. A 3D Face Matching Framework. In *Shape Modeling and Applications (SMI)*, pages 103–110, 2008. 2
- [21] H. H. Thodberg. Minimum description length shape and appearance models. In *IPMI*, pages 51–62, 2003. 2
- [22] D. Vranic and D. Saupe. 3D model retrieval. In *Proc. Spring Conference on Computer Graphics and its Applications, May*, pages 3–6, 2000. 1
- [23] D. Vranic, D. Saupe, and J. Richter. Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics. In *MMSP*, pages 293–298, 2001. 2
- [24] F. Wang, B. C. Vemuri, and T. Syeda-Mahmood. Generalized l2-divergence and its application to shape alignment. In *IPMI*. Springer, 2009. 2

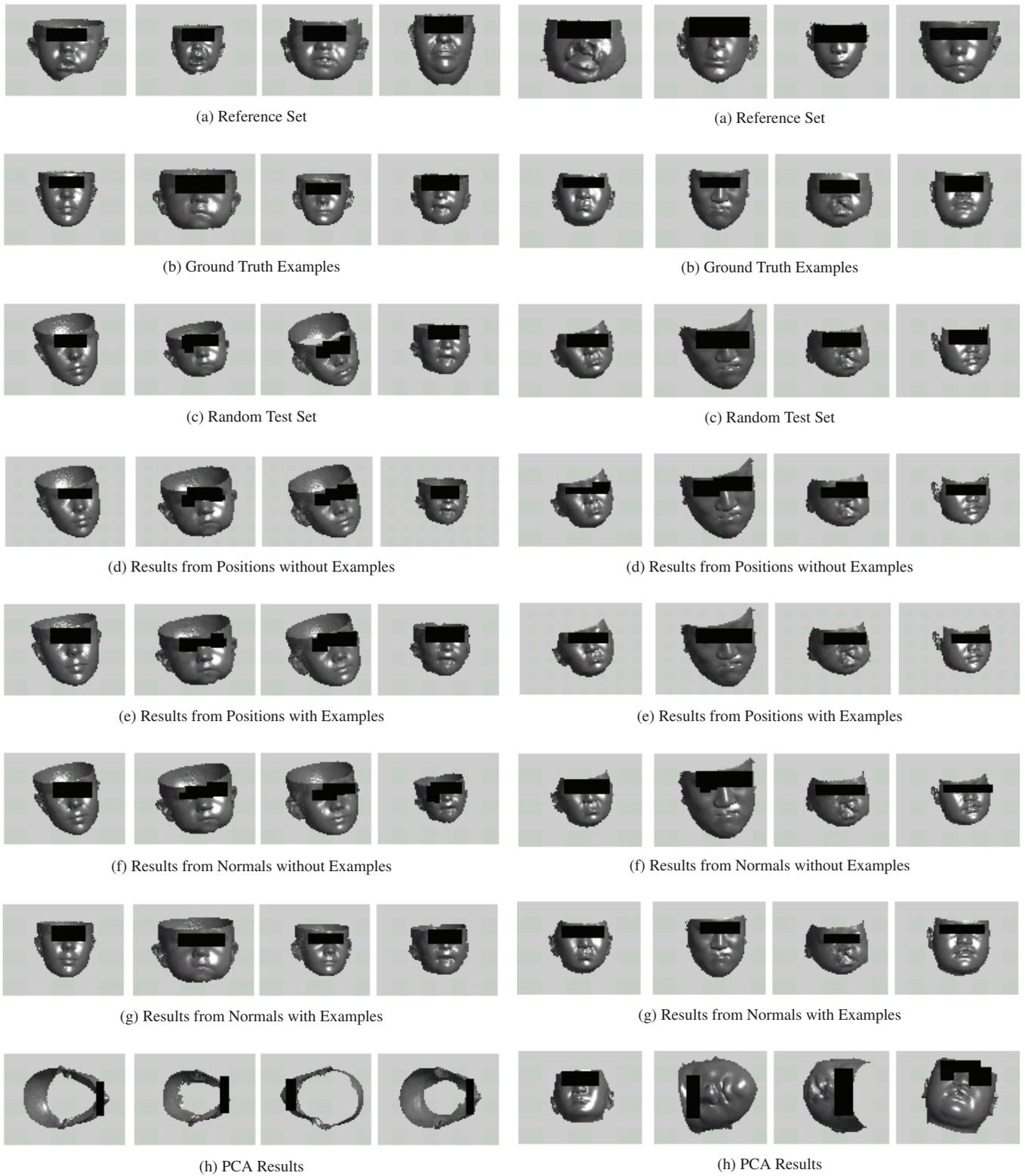
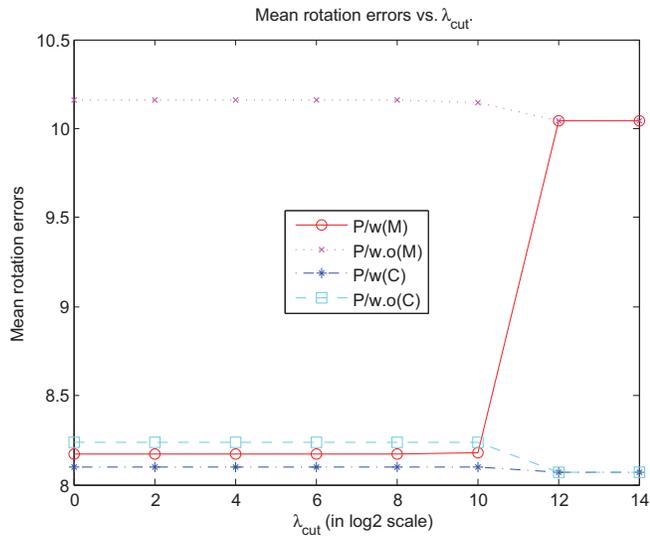
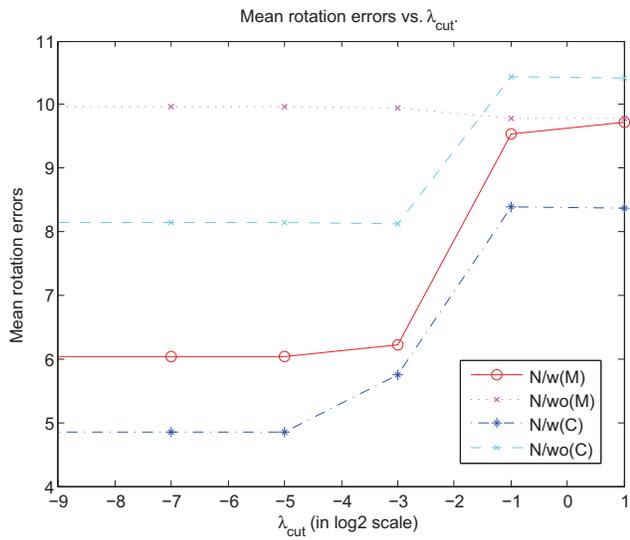


Figure 1: Some examples of the pose normalized 3D heads for the M set.

Figure 2: Some examples of the pose normalized 3D heads for the C set.



(a) Position



(b) Normal

Figure 3: The impact of λ_{cut} on the pose normalization results for position (a) and surface normal (b).