

# Construction of statistical shape atlases for bone structures based on two-level framework

Chenyu Wu<sup>1\*</sup>

Patricia E. Murtha, PhD

Branislav Jaramaz, PhD<sup>1</sup>

<sup>1</sup> The Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

\* Correspondence to: Chenyu Wu, The Robotics Institute, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA. Email: [chenyuwu@cmu.edu](mailto:chenyuwu@cmu.edu)

This original contribution is based in part on work previously presented at the 2<sup>nd</sup> International Conference on Computer Vision Theory (VISAPP2007) in Barcelona, Spain.

## Abstract

**Background** The statistical shape atlas is a 3D medical image analysis tool that encodes shape variations between populations. However, efficiency, accuracy, and finding the correct correspondence, are still unsolved issues during the construction of the atlas.

**Methods** We developed a two-level based framework which speeds up the registration

process while maintaining accuracy of the atlas. We also proposed a semi-automatic strategy to achieve segmentation and registration simultaneously, without knowing any prior information about the shape.

**Results** We have constructed the atlas for the femur and spine, separately. Experimental results demonstrate the efficiency and accuracy of our methods.

**Conclusions** Our two-level framework and semi-automatic strategy are able to efficiently construct the atlas for bone structures without losing accuracy. We can handle either 3D surface data or raw DICOM images.

**Keywords** statistical shape atlas; two-level; non-rigid registration; segmentation; femur; spine

## Introduction

In medical literatures, the atlas is first used as a reference to interpret CT/MR images [1], and represents the shape or appearance of human anatomical structures [2, 3]. Some atlases are based on physical properties, such as elastic models [4, 5], “snakes” [6], geometric splines [7], and finite elements based models [8]. Others are modeled from the statistical perspective [9]. In order to analyze the shape variation between patients, principal component analysis (PCA) is widely used [10, 11, 12]. The two most common statistical atlases are the shape atlas [13] and the appearance atlas [14]. The shape atlas uses only geometric information such as landmarks, surfaces (boundaries of 3D objects) or crest lines [3]. The appearance atlas uses both geometric features and intensity of pixels or voxels.

As a key to build atlases, 3D image registration has been studied for years in computer vision, but still remains a critical problem in the medical image field due to the geometrical complexity of anatomical shapes, and computational complexity caused by the enormous size of volume data. It has numerous clinical applications such as statistical atlas construction for group study and statistical parameters analysis [15, 16], mapping anatomical atlases to individual patient images for disease analysis [17, 18, 19] and segmentation [20, 21].

Depending on the type of transformation being applied, registration can be rigid or non-rigid. If the shape has no change between the two images, registration should be rigid, such as in inter-modality registration where images are captured at the same time. However, when we take into account the time factor, i.e., when two images are captured at different time, as in intra-subject registrations, most are non-rigid due to the shape variation of the anatomical structures caused by swelling, prostate poking, bone fractures, tumor growth changes, intestinal movements etc. In addition, inter-subject registrations are usually non-rigid because of the local anatomical difference between patients. So far non-rigid registration is still a challenging problem [22].

Non-rigid registration is used to find a non-rigid transformation from one 3D surface to the reference surface by minimizing the distance between two surfaces. In general, a non-rigid transformation is represented by a global rigid or affine transformation plus a local non-linear deformation, which can be described by radial basis functions (RBF) [23], octree-spline [24], thin-plate spline (TPS) [16, 25], geo-metric splines [7], finite elements [8], or free form B-spline [12] etc. To evaluate the results of the registration, different similarity measurements will be selected according to different features and imaging modalities. For example, sum of squared distances (SSD) is usually used for geometric features [26], but correlation coefficients (CC) [27], Ratio Image Uniformity (RIU) [28], and mutual information (MI) [29] are

used for intensity features. Registration problem can be simplified given known correspondences, for example using markers [30]. Nevertheless, markers are not allowed to be used or even available in many scenarios. Alternate estimation of correspondences and transformations are therefore widely used for both rigid cases [26] and non-rigid cases [15, 25, 31]. Moreover, with the increase of data size and geometrical complexity, multi-resolution strategy has been adopted into the registration framework [20, 32]. Sparse matrices are also used to handle the computational complexity [33].

Our goal in this paper is to construct shape atlases for bone structures, which can be used as references in surgical navigation systems. We developed an efficient algorithm to overcome computational complexity while maintaining accuracy. We can handle different data types including 3D surface data and raw DICOM images.

## **Materials**

Being the longest bone in the human body, the femur is often involved in hip and knee surgeries. It will be a good start with the femur due to its importance and simplicity in shape. On the other hand, spine vertebrae are much more complicated but also very important as spine surgeries are very common now. Therefore we select the femur and spine to be studied.

### **Femur data**

We have CT-scanned 87 different patients all with healthy femur in West Pennsylvania Hospital (Pittsburgh, PA) and Shadyside Hospital (Pittsburgh, PA) over several years. There are 53 males and 34 females. 43 are left femurs and 44 are right. The patients' age ranges from 39 to 78 and their femur length ranges from 400mm to 540mm. The CT volumes are

segmented to provide triangulated surface models using marching cube (MC) algorithm. All surface models are smoothed by the method proposed in [34]. Since we are more interested in the condyles given most surgeries are applied on the condyles, only the top portion including the femoral head and the bottom portion including the condyles are scanned to build surfaces.

## **Spine data**

In our experiments we use lumbar vertebrae since lumbar vertebrae are the largest segments of the movable part of the vertebral column. It is easier to segment this section from the vertebral column. We select all five lumbar vertebrae (L1-L5) and the last thoracic vertebra (T12) for our study.

We apply the CT scan to three healthy volunteers in the University of Pittsburgh Medical Center (UPMC) and collect the data in DICOM format. To enlarge the dataset, we have scanned three human-size artificial spines by the same protocol. We use the solid white plastic sawbones model (# 1352-31) made of rigid plastic. The shape of the sawbones models are cast from different natural real specimens. Fig. 1 shows the CT image of an artificial spine and a real spine.

## **Methods**

We propose a two-level non-rigid registration approach inspired by Chui and Rangarajan's non-rigid registration algorithm [25] and previous multi-resolution work [32]. After each 3D surface is registered to the reference, we use PCA to build the statistical shape atlas. We also develop a semi-automatic strategy to handle the spine surface segmentation and registration simultaneously.

## Two-level Non-rigid Registration

Since Chui and Rangarajan's algorithm [25] is not able to handle more than 2000 3D points [33], we broke down the registration procedure into a two-level process to deal with the computational and geometrical complexity. We first applied their algorithm to the simplified low-resolution surfaces, which are computed by Garland's quadric error metrics (QEM) based surface simplification algorithm [35].

To improve efficiency, instead of successively matching each resolution from coarse to fine, we directly propagate the correspondences from low resolution ( $X^L$  and  $Y^L$ ) to high resolution ( $X^H$  and  $Y^H$ ) by interpolation. The surface interpolation method is a derivative of methods known jointly as "moving least squares" [36]. Radial basis functions (RBF), finite element, multivariate spline such as thin-plate spline (2D bivariate spline) and triharmonic thin-plate spline, are popular techniques used in surface interpolation. Carr et al. [6] applied multivariate splines method into radial basis functions by using splines as kernel functions. In this work we choose Gaussian kernel due to its simple mathematical representation and less restrictions on nodes [37]. More specifically, we used a linear affine function plus a series of radial basis functions (RBFs) to construct the interpolation function:

$$y_i^L = g(x_i^L) = \mathbf{c}_1 \cdot \left[ \varphi \left( \left\| x_i^L, x_1^L \right\| \right), \dots, \varphi \left( \left\| x_i^L, x_N^L \right\| \right) \right]^T + \mathbf{c}_2 + \mathbf{c}_3 \cdot x_i^L \quad (1)$$

where  $x_i^L$  is a vertex on the low-resolution surface  $X^L$ , whose correspondence on the low-resolution surface  $Y^L$  is  $y_i^L$ ,  $i = 1, 2, \dots, N$  ( $N$  is the number of vertices on the surface  $X^L$ ).

$x_i^L$  and  $y_i^L$  are both  $3 \times 1$  vectors with three coordinates.  $\mathbf{c}_1$  is a  $3 \times N$  coefficient matrix of radial basis functions.  $\varphi \left( \left\| x_i^L, x_1^L \right\| \right)$  is a symmetric radial basis function. We have chosen a Gaussian kernel  $\varphi(u_i, u_j) = \exp(-\|u_i - u_j\| / 0.5)$ , as suggested by [38].  $\mathbf{c}_2$  and  $\mathbf{c}_3$  are

coefficients for the affine transformation.  $\mathbf{c}_2$  is a  $3 \times 1$  vector and  $\mathbf{c}_3$  is a  $3 \times 3$  matrix. Given  $N$  correspondences, we have  $N$  equations for each axis (X , Y and Z):

$$\underbrace{\begin{bmatrix} \varphi \left( \left\| x_1^L, x_1^L \right\| \right) & \cdots & \varphi \left( \left\| x_1^L, x_N^L \right\| \right) & 1 & x_1^{LT} \\ \vdots & & \vdots & \vdots & \vdots \\ \varphi \left( \left\| x_N^L, x_1^L \right\| \right) & \cdots & \varphi \left( \left\| x_N^L, x_N^L \right\| \right) & 1 & x_N^{LT} \end{bmatrix}}_{P^k} \cdot \underbrace{\begin{bmatrix} \mathbf{c}_1^{kT} \\ \mathbf{c}_2^{kT} \\ \mathbf{c}_3^{kT} \end{bmatrix}}_{\mathbf{c}^k} = \underbrace{\begin{bmatrix} y_1^{Lk} \\ \vdots \\ y_N^{Lk} \end{bmatrix}}_{y^k} \quad (2)$$

where  $\mathbf{c}_1^k$ ,  $\mathbf{c}_2^k$  and  $\mathbf{c}_3^k$  denote the  $k^{\text{th}}$  row of  $\mathbf{c}_1$ ,  $\mathbf{c}_2$  and  $\mathbf{c}_3$ , respectively.  $y_i^{Lk}$  denotes the  $k^{\text{th}}$  row of  $y_i^L$ ,  $k$  can be 1,2 or 3, corresponding to the axis X , Y and Z. Therefore we have  $3N$  equations in all:

$$P = [P^1 \ P^2 \ P^3]^T, \quad c = [c^1 \ c^2 \ c^3]^T, \quad y = [y^1 \ y^2 \ y^3]^T \quad (3)$$

$P$  is a  $3N \times (N + 4)$  matrix. In order to ensure smooth interpolation, we add the additional orthogonally constraints  $\sum_i x_i^{LT} \mathbf{c}_{1,i} = 0$  [39] to Eq. 3, where  $\mathbf{c}_{1,i}$  denotes the  $i^{\text{th}}$  column of  $\mathbf{c}_1$ :

$$\underbrace{\begin{bmatrix} P \\ x_1^L & x_2^L & \cdots & x_N^L & 0_{4 \times 4} \end{bmatrix}}_Q \cdot c = \underbrace{\begin{bmatrix} y \\ 0_{4 \times 1} \end{bmatrix}}_w \quad (4)$$

The least-squares solution for this linear system ( $Qc = w$ ) is given by  $c = (Q^T Q)^{-1} Q^T w$ . Finally, the correspondence of a vertex  $x_j^H$  on the high-resolution surface  $X^H$  can be computed by using Eq. 1:  $y_j^H = g(x_j^H)$ ,  $j = 1, \dots, M$  ( $M$  is the number of vertices on the surface  $X^H$ ).

For both low-resolution and high-resolution surfaces, we applied a refining procedure to improve accuracy. The refining procedure includes both local and global steps.

- **Local refining:** the registration can be further improved by minimizing the point-to-surface distances. This idea is illustrated by Fig. 2.  $x_i$  is a vertex on the deformed surface  $X$ , whose corresponding vertex on the surface  $Y$  is  $y_i$ . We check the neighboring triangles of  $y_i$ , which are triangles sharing the same vertex  $y_i$ , e.g.

$S_1$ ,  $S_2$  and  $S_3$  in Fig. 2. We examine the distance from  $x_i$  to each neighboring triangle (the distance computed from the vertex  $x_i$  to the plane where the triangle lies), i.e.  $d_1$ ,  $d_2$  and  $d_3$  in Fig. 2. If any of them is smaller than  $d_0 = \|x_i - y_i\|$ , we use the corresponding projected point to replace  $y_i$  to achieve a better registration. For those cases where different vertices on the surface  $X$  correspond to the same surface point on  $Y$ , we assign this corresponding surface point to the vertex on  $X$  with the smallest distance and make it unavailable to other vertices on  $X$ .

- **Global refining:** for low-resolution surfaces with sparse points, we cannot apply the local refining procedure since it may change the shape of the surfaces. It is also not applicable to the spine case when the high-resolution surface is not available. Therefore we propose a global refining procedure instead. We apply a RBF based interpolation by an iterative way to warp the high-resolution registered reference surface toward the original low-resolution points. We employ the same RBF functions used for interpolation but different sigma for the RBF kernel. Here we choose  $\sigma$  as 250. The larger  $\sigma$  represents more global deformation.

## Initial Alignment

Given the particularity of the femur data, we need to do a pre-alignment to get rid of the effect caused by the missing shape. As Fig. 3 shows, we use the bottom portion of the femur as an example. The surface  $Y$  has more femur shaft but less shaft remains on the surface  $X$ . If we align both centers as in previous work [25], experiments shows that the registration process will be very slow and may not converge in several cases. The reason is that a part of the surface  $Y$  (as bounded by blue in Fig. 3) has no counterpart on  $X$ . Therefore, in order to improve upon the process, we decided to estimate the pseudo center of  $Y$  instead of the true

center. And then the pose of two surfaces are estimated and aligned.

We use the length of  $X$  to estimate the pseudo length of  $Y$ . Assuming the axis  $Z$  is in the scan direction from the knee to hip. We select those points bounded by the pseudo length (denoted by black in Fig. 3) to estimate the pseudo center  $\kappa_{Y'}$ , also covariance matrices  $\Psi_X$  and  $\Psi_{Y'}$ :

$$\begin{aligned}\kappa_X &= \frac{1}{N_X} \sum p_X \\ \kappa_{Y'} &= \frac{1}{N_{Y'}} \sum p_{Y'}\end{aligned}\tag{5}$$

$$\begin{aligned}\Psi_X &= \frac{1}{N_X - 1} [p_X^1 - \kappa_X \quad \cdots \quad p_X^{N_X} - \kappa_X] \cdot [p_X^1 - \kappa_X \quad \cdots \quad p_X^{N_X} - \kappa_X]^T \\ \Psi_{Y'} &= \frac{1}{N_{Y'} - 1} [p_{Y'}^1 - \kappa_{Y'} \quad \cdots \quad p_{Y'}^{N_{Y'}} - \kappa_{Y'}] \cdot [p_{Y'}^1 - \kappa_{Y'} \quad \cdots \quad p_{Y'}^{N_{Y'}} - \kappa_{Y'}]^T\end{aligned}\tag{6}$$

where  $N_{Y'}$  is the number of points  $\{p_{Y'}\}$  in  $Y$  which satisfy  $(z_Y - \min z_Y) < (z_X - \min z_X)$ . We solve for the principle axes by decomposing the covariance matrix using the moment analysis:

$$\Psi_X = U_X \Lambda_X U_X^T, \quad \Psi_{Y'} = U_{Y'} \Lambda_{Y'} U_{Y'}^T\tag{7}$$

Each column of  $U_X$  represents a principle axis of points set  $\{p_X\}$ , and  $U_{Y'}$  for  $\{p_{Y'}\}$ . As Fig. 3 shows, we use three axes to describe the pose of the points set: red for  $\{p_X\}$  and green for  $\{p_{Y'}\}$ . The rotation from coordinates of  $X$  to  $Y'$  is computed by  $U_{Y'} \cdot U_X^T$ . We apply a rigid transformation  $[U_{Y'} \cdot U_X^T \mid (\kappa_{Y'} - \kappa_X)]$  to the points set  $\{p_X\}$  and two point sets are finally aligned. Experiment shows the rate of convergence has been improved from 78% to 95.2% by doing such alignment.

## Semi-automatic Strategy

Compared with the femur, spine vertebrae have much more complicated shapes which

makes it more difficult to segment the surface model from the 3D CT images using the marching cube algorithm. However it is very time consuming to manually label over ten thousand points to obtain the high-resolution surfaces. Given some shape information, Kaus et al. [40] use the learning deformable models to generate aligned 3D meshes. In our case without knowing any shape prior, we have developed a semi-automatic strategy based on the two-level registration framework to generate high-resolution aligned surface models. Through this method, we were able to combine surface segmentation and registration in the same procedure, and reduce manual work to a minimal.

We have a high-resolution 3D surface of vertebra generated by triangulating 40682 points which were carefully labeled by hand. This step takes several hours but only need to be done once. We trim each spine CT images into small files with individual vertebra. We manually label about 250 surface points (of the same resolution as the simplified reference surface) for each DICOM file. It takes 10-15 minutes to label each vertebra. Basically we label around 8-9 points on each image slice (12 slices from the coronal plane, 12 slides from the sagittal plane and 6 slices from the transverse plane). Examples are showed in Figs. 4-6. The labeling is assumed to be consistent since it was conducted by the same person based on the study of the structure of the spine vertebrae and DICOM images.

We apply the two-level registration algorithm on a high-resolution reference surface and a low-resolution 250 points data set. The only difference between the femur and the spine is in the last step. Since the high-resolution spine surfaces are not available, the local refine procedure cannot be applied. Only global refine procedure will be utilized to generate the final aligned high-resolution surfaces.

## **Atlas Construction**

Rigid pose alignments are applied to eliminate the effect of imaging poses [41] prior to atlas construction. Suppose we have  $K$  aligned 3D surfaces and each surface is represented by a  $3M \times 1$  vector  $v_i$  ( $i=1, \dots, K$ ), where  $M$  is the number of points in each surface and 3 denotes three coordinates X, Y and Z. We compute the mean vector  $\kappa$  and covariance matrix  $\Psi$  and then apply PCA on the data set:

$$\kappa = \frac{1}{K} \sum v_i, \quad \Psi = \frac{1}{M-1} [v_1 - \kappa \quad \dots \quad v_K - \kappa] \cdot [v_1 - \kappa \quad \dots \quad v_K - \kappa]^T \quad (8)$$

$$\Psi = U \Lambda U^T \quad (9)$$

Therefore, any surface vector in the data set can be represented by a mean vector plus a linear combination of principal components (modes):

$$v_i = \kappa + U \eta_i \quad (10)$$

where  $\eta_i$  is a  $K \times 1$  coefficient vector obtained by projecting  $v_i$  onto each principal axis. New models, not included in the data set, can be generated by manipulating the coefficient  $\eta_i$ .

## Selection of Simplification Parameters

In the two-level framework, the resolution of the simplified surfaces affects the final result. The fewer the number of points, the faster low-resolution registration is achieved but results in high-resolution registration are less accurate. To maintain both accuracy and efficiency, an appropriate number is selected based on a series of leave-one-out experiments.

Let  $N_{ref}^{low}$  denote the number of vertices in the low-resolution surfaces. As Fig. 7 shows, for each surface  $v_i$  ( $i=1, \dots, K$ ), we use other  $K-1$  surfaces to construct the atlas using Eqs. 8-9. Let  $U_i^S$  denote the first  $S$  columns of the principal component matrix  $U_i$ , which consists of 95% of shape variations. Then  $v_i$  ( $i=1, \dots, K$ ) can be reconstructed by this atlas:

$$\tilde{v}_i = \kappa + U_i^S U_i^{S^T} (v_i - \kappa) \quad (11)$$

We compare the surface distance between the original surface  $v_i$  and the reconstructed surface  $\tilde{v}_i$  by computing the mean error and root mean square error. We repeat this procedure for each surface and compute the average mean error and RMS error:

$$\overline{d_{Mean}} = \frac{1}{K} \sum_i d_{Mean}(v_i, \tilde{v}_i), \quad \overline{d_{RMS}} = \frac{1}{K} \sum_i d_{RMS}(v_i, \tilde{v}_i) \quad (12)$$

By tuning the number  $N_{ref}^{low}$ , we compare the  $\overline{d_{Mean}}$ ,  $\overline{d_{RMS}}$ , and the processing time  $\bar{T}$  for the two-level registration. Fig. 8 shows when  $N_{ref}^{low} \geq 0.2\% N_{ref}^{high}$ ,  $\overline{d_{Mean}}$  will be less than 1mm, which is a practical number in the clinical applications. Fig. 9 shows when  $N_{ref}^{low} \geq 0.3\% N_{ref}^{high}$ , the average processing time of the two-level registration will exceed 5 minutes (2.4GHz Pentium PC with 1GB RAM). So  $N_{ref}^{low} = 0.2\% N_{ref}^{high}$  is finally used in our framework to build the atlas and estimate the error distribution.

## Results

### Evaluation of Two-level Registration

We use the bottom portion of the femur as an example given its important relationship with the knee. Fig. 10 shows an example of how the surfaces distance is decreased in each step of the two-level frame-work. Figs. 11-16 show the result in each step. In this case we have two high-resolution surfaces  $X^H$  (21130 vertices, 42256 triangles, 65.84mm in z-axis) and  $Y^H$  (26652 vertices, 53300 triangles, 105.89mm in z-axis) (Patient  $X$  is a 79 years old female, with 472.55mm femur length; Patient  $Y$  is a 53 years old female, with 477.59mm femur length). We first compute point-to-surface distance from  $X^H$  to  $Y^H$  [42]:

$$d(p, Y^H) = \min_{p \in Y^H} \|p - p^*\|_2, p \in X^H \quad (13)$$

where  $\|\cdot\|_2$  is Euclidean norm. Then we compute the mean error  $d_{Mean}(X^H, Y^H)$  and root mean square (RMS) error  $d_{RMS}(X^H, Y^H)$  between surfaces  $X^H$  and  $Y^H$  from Eq. 13.

$$\begin{aligned} d_{Mean}(X^H, Y^H) &= \frac{1}{|X^H|} \sum_{p \in X^H} d(p, Y^H) dX^H \\ d_{RMS}(X^H, Y^H) &= \sqrt{\frac{1}{|X^H|} \sum_{p \in X^H} d(p, Y^H)^2 dX^H} \end{aligned} \quad (14)$$

1. High-res (see Fig. 11): with respect to the bounding box diagonal of the surface  $Y^H$  (158.48mm), the mean error is 6.49% and root mean square error is 7.70%.
2. Low-res (see Fig. 12): for the low-resolution surfaces  $X^L$  (169 vertices, 334 triangles) and  $Y^L$  (213 vertices, 422 triangles), with respect to the bounding box diagonal of the surface  $Y^L$  (158.28mm), the mean error is 6.53% and root mean square error is 7.74%.
3. Low-res after TPS (see Fig. 13): with respect to the bounding box diagonal of the surface  $Y^L$ , the mean error is 1.68% and root mean square error is 2.13%. Surface distance has been significantly decreased by performing Rangarajan's non-rigid registration method [25].
4. Low-res after refining (see Fig. 14): with respect to the bounding box diagonal of the surface  $Y^L$ , the mean error is 0.68% and root mean square error is 1.42%, which shows a refining step (global RBF interpolation) can decrease the surface distance further once two surfaces are similar enough.
5. High-res after interpolation (see Fig. 15): with respect to the bounding box diagonal of the surface  $Y^H$ , the mean error is 1.65% and root mean square error is 2.10%. The reason why the surface distance slightly increases after interpolation is: only 0.80% of vertices on the surface  $X^{H(1)}$  have correspondences computed by low-level registration,

others obtain correspondences by interpolation.

6. High-res after refine (see Fig. 16): with respect to the bounding box diagonal of the surface  $Y^H$ , the mean error is 0.28% and root mean square error is 1.26%, which once again demonstrates that the refining steps (local and global deformation) are helpful for decreasing high-resolution surfaces distance.

## Compare to Other Registration Methods based on Femur Data

- **Compare to rigid registration:** We compared our method with the classical iterative closet point (ICP) algorithm since it is widely used in many medical image registration problems, even though ICP is a rigid method. The distributions of the RMS errors and mean errors of the two approaches are showed in Fig. 17. For the RMS errors, they are centered between 0.61-1.00mm (the maximum is 1.35mm) in our method, but in ICP the range is between 3.10-6.50mm (the maximum is 8.50mm). For the mean errors, they are centered between 0.51-0.71mm (the maximum is 0.85mm) in our method, but in ICP the range is 2.50-6.50mm (the maximum is 7.50mm).
- **Compare to non-rigid registration:** Since Chui and Rangarajan's thin-plate spine (TPS) based method [25] cannot deal with the complexity of our data, we cannot directly compare two methods. Alternatively, we noticed from Fig. 9 that our method only needs 5 minutes or less to match any size of surfaces with less than 200,000 vertices. However, their method costs 5 minutes for 350 vertices, 10 minutes for 460 vertices, 20 minutes for 610 vertices, etc. It is obvious that our algorithm significantly improves efficiency.

## Evaluation of Atlases

Styner et al. [43] has proposed some measurements for evaluating the atlas in terms of compactness, generalization, specificity, and distance to manual landmarks. For the statistical atlases which encode the shape variation among different subjects, the comparison to the landmarks is not applicable in this case. The generalization is measured by a series of leave-one-out experiments. By tuning the simplification parameter, we select the model which has the best generalization (less reconstruction errors). The compactness is analyzed below for the femur atlas (Fig. 27) and spine atlas (Fig. 29) separately. The specificity is not studied in this paper since we focus on shape analysis without generating new objects.

**Femur Atlas Statistics:** Fig. 18 shows the shape variations in the first three modes of the atlas for the top portion of the femur. We use color (pink and blue) to highlight the local variations. E.g. the size of the femoral head (denoted by the pink circle) gradually reduces along the first mode, but remains almost the same along the second mode. The greater trochanter (denoted by the blue area) becomes longer but narrower along the first mode. It remains the same length but becomes wider along the third mode.

Fig. 19 shows the shape variations in the first three modes of the atlas for the bottom portion of the femur. We notice that the top of the condyles (denoted by the blue area) changes a lot along each mode. This is because each scan is cut differently. The top area can only be estimated from the reference surface without any refining from the original surfaces. The lateral epicondyle (denoted by the blue point on the right) becomes larger along the first mode, compared to the medial epicondyle (denoted by the blue point on the left). It is similar to the third mode but the change is smaller. The condyles become longer but narrower along the second mode.

Figs. 20-21 show the shape variations in the first two modes of the entire femur. For the first mode, the femoral head and greater trochanter follow the similar change as Fig. 18

shows, and the condyles become longer and narrower as Fig. 19 shows. The atlas of the entire femur behaves in a similar way as the individual portion does for the first mode. For the second mode, the greater trochanter becomes shorter but it remains the same as in Fig. 18. The condyles become much longer but they keep the same length as in Fig. 19. It means that the second dominant mode in the individual atlas may not be the second dominant mode in the conjunction case.

Fig. 22 shows the shape variations encoded in the modes of each atlas. To keep 95% of shape variations we only need 26 modes for the top portion of the femur, 49 modes for the bottom portion of the femur, and 20 modes for the entire femur atlas. The reason why the bottom portion of the femur needs more modes to keep most of the shape variations is because the training surfaces of the bottom portion of the femur change a lot in shape (i.e. The model contains different length of the femur shaft).

**Spine Atlas Statistics:** Fig. 23 shows the shape variations in the first three modes of the spine vertebra atlas. We can see that the vertebra becomes narrower but remains the same height among the first mode. The vertebral body (centrum) becomes larger along both the second and third modes. The spinous process becomes longer along the second mode but shorter along the third mode. It is twisted in the same way from the left to the right. We noticed that the sampled surfaces given the bigger standard deviation (e.g.  $1\sigma$ ) have odd shapes. This may be due to insufficient training data. Fig. 24 shows the shape variations encoded in modes of each atlas. To keep 95% of shape variations we only need 10 modes.

## Discussions

The semi-automatic procedure requires some manual input, e.g. to label 40682 points for a

reference model. But this labeling is a one-time task and can be conducted offline. It also requires 250 points to be labeled for each image studied since we don't have any prior information about the shape. It normally takes 10-15 minutes, which is much less than several hours used for labeling over ten thousand points. In future work we will incorporate the atlas as a prior to enhance performance and reduce manual labeling.

The salient points may be removed during the surface simplification and mis-matching may occur. To handle this problem, we need to find a proper resolution for the simplified surface. The lower the faster but the more mis-matching could be established. That's why we apply the leave-one-out experiments to find a good parameter for the resolution. Besides, we propose a refining process by using the original data (either the high resolution surface or labeled point set) as a reference to reduce the mis-matching caused by the simplification.

## Conclusions

In this paper we developed a two-level framework to efficiently tackle non-rigid registration challenges caused by geometrical and computational complexity. Our work is mainly inspired by Chui and Rangarajan's method [25] but performs much faster and requires less memory. Experiments demonstrate that our method significantly improves efficiency of registration without decreasing accuracy of atlases. Our framework can handle both 3D surface data and raw CT DICOM images. For those anatomical structures with complex shapes which are hard to be segmented directly using the standard marching cube algorithm, we proposed a semi-automatic strategy to successfully generate high-resolution and align 3D surfaces simultaneously.

The shape variations learned from both femur and spine atlases can be used for clinical studies and diagnosis. Besides, the shape atlases can be used as a statistical prior for shape reconstruction from other imaging modalities such as the endoscope, and also for the

segmentation of 3D surfaces from DICOM images, which will require less manual input and perform much faster.

## References

1. McInerney T, Terzopoulos D. Deformable models in medical image analysis: a survey. *Med Image Anal* 1996; **1**: 91–10.
2. Subsol G, Thirion JP, Ayache N. First steps towards automatic building of anatomical atlas. INRIA, France 2216; 1994 March.
3. Subsol G, Thirion JP, Ayache N, Epidaure P. A general scheme for automatically building of anatomical atlas: application to a skull atlas. *Med Image Anal* 1995; **2**: 37-60.
4. Broit C. Optimal registration of deformed images. PhD Thesis, Computer and Information Science Department, University of Pennsylvania, Philadelphia, PA, USA; 1981.
5. Bajcsy R, Kovacic S. Multiresolution elastic matching. *Comput Vis Graph, Image Process* 1989; **46**: 1–21.
6. Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. *Int J Comput Vis* 1988; **1**: 321–331.
7. Farin G. Curves and surfaces for cagd. Academic Press: New York, NY, 1993.
8. Park J, Mataxas D, Young A, Axel L. Deformable models with parameter functions for cardiac motion analysis from tagged mri data. *IEEE Trans Med Imaging* 1996; **15**: 278–289.
9. Szeliski R. Bayesian modeling of uncertainty in low-level vision. *Int J Comput Vis* 1990; **5**: 271–301.
10. Cootes T, Beeston C, Edwards G, Taylor C. A unified framework for atlas matching using active appearance models. Proceedings of the XVIth International Conference on Information Processing in Medical Imaging; 1999 Jun 28-Jul 2; Visegrad, Hungary; 322–333.

11. Meller S, Kalender W. Building a statistical shape model of the penis. In CARS 2004 - Computer Assisted Radiology and Surgery: Proceedings of the 18th International Congress and Exhibition; 2004 Jun 23 – 26; Chicago, USA; **1268**: 561–566.
12. Rueckert D, Frangi A, Schnabel J. Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. *IEEE Trans Med Imaging* 2003; **22**: 1014–1025.
13. Cootes T, Taylor C, Cooper D, Graham J. Active shape models: their training and application. *Comput Vis Image Und* 1995; **61**: 38–59.
14. Cootes T, Taylor C. Statistical models of appearance for medical image analysis and computer vision. *SPIE Med Imaging* 2001; **4322**: 238–248.
15. Chui H, Rangarajan A, Zhang J, Leonard CM. Unsupervised learning of an atlas from unlabeled point-sets. *IEEE Trans Pattern Anal Mach Intell* 2004; **2**: 160–172.
16. Hill D, Batchelor P, Holden M, Hawkes D. Medical image registration. *J Phys Med Bio* 2001; **46**: 1–45.
17. M. Chen. 3-d deformable registration using a statistical atlas with applications in medicine. PhD Thesis, The Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA; 1999.
18. Fleute M, Lavallée S, Desbat L. Integrated approach for matching statistical shape models with intraoperative 2d and 3d data. Proceedings of the 5<sup>th</sup> Medical Image Computing and Computer-Assisted Intervention (MICCAI02); 2002 Sept 25-28; Tokyo, Japan; **2**: 364–372.
19. Jaume S, Macq B, Warfield SK. Labeling the brain surface using a deformable multiresolution mesh. Proceedings of the 5<sup>th</sup> Medical Image Computing and Computer-Assisted Intervention (MICCAI02); 2002 Sept 25-28; Tokyo, Japan; 451–458.
20. Rohlfing T, Russakoff D, Maurer C. Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. *IEEE Trans Med Imaging* 2004; **23**: 983–994.
21. Straka M, Cruz A, Dimitrov L, Sramek M, Fleischmann D, Groller E. Bone segmentation in ctangiography data using a probabilistic atlas. Proceedings of Vision, Modeling and Visualization ( VMV03); 2003 Nov 19-21; Munich, Germany; 505–512.

22. Dawant B. Non-rigid registration of medical images: purpose and methods, a short survey. Proceedings of IEEE International Symposium on Biomedical Imaging; 2002 Jul 7-10; Washington, DC, USA; 465–468.
23. Yu H. Automatic rigid and deformable medical image registration. PhD Thesis, Mechanical Engineering Department, Worcester Polytechnic Institute, Worcester, MA, USA; 2005.
24. Szeliski R, Lavalée S. Matching 3-d anatomical surfaces with non-rigid deformations using octreesplines. *Int J Comput Vis* 1996; **18**: 171–186.
25. Chui H, Rangarajan A. A new point matching algorithm for non-rigid registration. *Comput Vis Image Und* 2003; **89**: 114–141.
26. Besl P, McKay N. A method for registration of 3-d shapes. *IEEE Trans Pattern Anal Mach Intell* 1992; **14**: 239–256.
27. Kim J and Fessler JA. Intensity-based image registration using robust correlation coefficients. *IEEE Trans. Pattern Anal Mach Intell* 2004; **23**: 1430–1444.
28. Woods RP, Mazziotta JC, Cherry SR. Mri-pet registration with automated algorithm. *J Comput Assist Tomogr* 1994; **17**: 536–546.
29. Wells WM, Viola P, Atsumi H, Nakajima S, Kikinis R. Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1996; **1**: 35–51.
30. Maurer CR, Fitzpatrick JM, Wang MY, Galloway RL, Maciunas RJ, Allen GG. Registration of head volume images using implantable fiducial markers. *IEEE Trans. Pattern Anal Mach Intell* 1997; **16**: 447–462.
31. Glaunes J, Trounev A, Younes L. Diffeomorphic matching of distributions: a new approach for unlabelled point-sets and sub-manifolds matching. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2004 Jun 27-Jul 2; Washington, DC, USA; **2**: 712–718.
32. Shen C, Davatzikos D. Hammer: hierarchical attribute matching mechanism for elastic registration. *IEEE Trans Med Imaging* 2002; **21**: 1421–1439.
33. Papademetris X, Jackowski A, Schultz R, Staib L, Duncan J. Non-rigid brain registration using extended robust point matching for composite multisubject fmri analysis. Proceedings of 6<sup>th</sup>

- Medical Image Computing and Computer-Assisted Intervention (MICCAI03); 2003 Nov 15-18; Montreal, Canada; **2879**: 788–795.
34. Desbrun M, Meyer M, Schroder P, Barr A. Implicit fairing of irregular meshes using diffusion and curvature flow. Proceedings of the 26<sup>th</sup> ACM Annual Conference on Computer Graphics (SIGGRAPH99); 1999 Aug 8-13; Los Angeles, CA, USA; 317–324.
  35. Garland M, Heckbert P. Surface simplification using quadric error metrics. Proceedings of the 24<sup>th</sup> ACM Annual Conference on Computer Graphics (SIGGRAPH97); 1997 Aug 3-8; Los Angeles, CA, USA; 209–216.
  36. Yngve G, Turk G. Robust creation of implicit surfaces from polygonal meshes. *IEEE Trans Visual Comput Graph* 2002; **8**: 346–359.
  37. Carr JC, Fright WR, Beatson RK. Surface interpolation with radial basis functions for medical imaging. *IEEE Trans Med Imaging* 1997; **16**: 96–107.
  38. Pighin F, Hecker J, Lischanski D, Szeliski R, Salesin D. Synthesizing realistic facial expressions from photographs. . Proceedings of the 25<sup>th</sup> ACM Annual Conference on Computer Graphics (SIGGRAPH98); 1998 Jul 19-24; Orlando, FL, USA; 75–84.
  39. Carr JC, Beatson RK, Cherrie JB, Mitchell TJ, Fright WR, Mccallum BC, Evans TR. Reconstruction and representation of 3d objects with radial basis functions. Proceedings of the 28<sup>th</sup> ACM Annual Conference on Computer Graphics (SIGGRAPH01); 2001 Aug 12-17; Los Angeles, CA, USA; 67–76.
  40. Kaus MR, Pekar V, Lorenz C, Truyen R, Lobregt S, Weese J. Automated 3-D PDM construction from segmented images using deformable models. . *IEEE Trans Pattern Anal Mach Intell* 2003 Aug; **22(8)**: 1005-1013.
  41. Goryn D, Hein S. On the estimation of rigid body rotation from noisy data. *IEEE Trans Pattern Anal Mach Intell* 1995; **17**: 1219–1220.
  42. Aspert N, Santa-Cruz D, Ebrahimi T. Mesh: measuring errors between surfaces using the hausdorff distance. Proceedings of IEEE International Conference on Multimedia and Expo; 2002 Aug 26-29; Lausanne, Switzerland; **1**: 705–708.

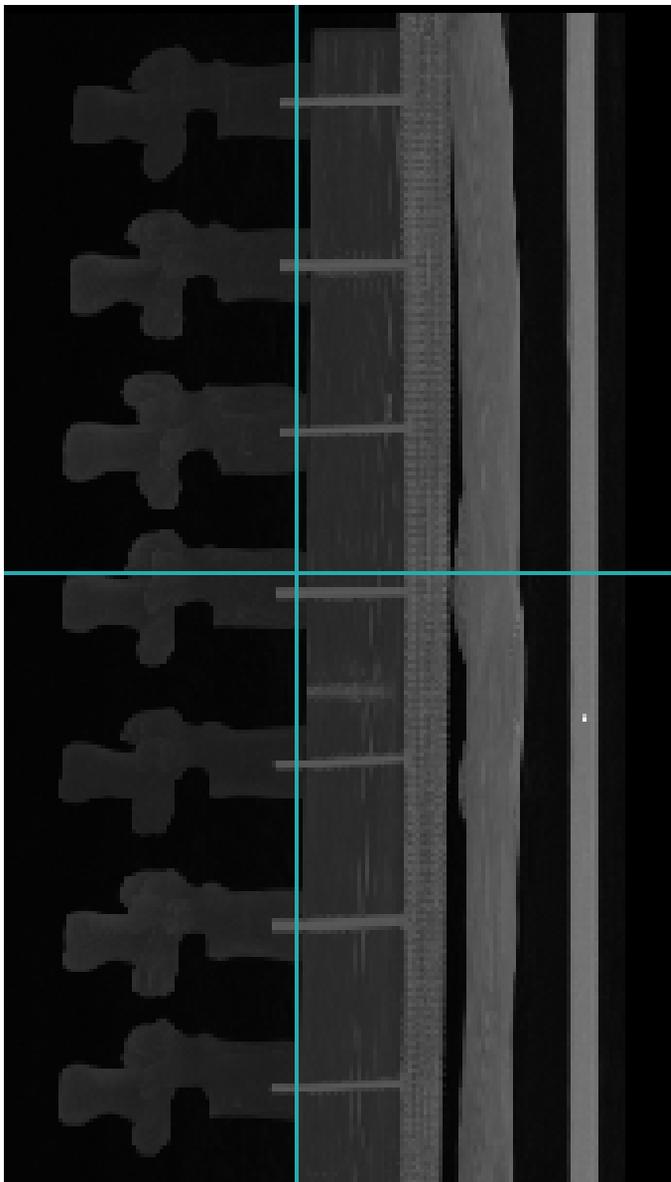
43. Styner MA, Rajamani KT, Nolte LP, Zsemlye G, Szekely G, Taylor CJ, Davies RH. Evaluation of 3D correspondence methods for model building. *Inf Process Med Imaging*; 2003 Jul; **18**: 63-75.

## Figures

1. Spine CT images for the artificial case (a) and real patient (b)
2. Illustration of local non-rigid registration between the vertex and surface.  $x_i$  is a vertex on the deformed surface  $X$ , whose correspondence on the surface  $Y$  is  $y_i$ . The projection of  $x_i$  to each triangle  $S_t$  sharing  $y_i$  is denoted by  $y_i^t$ .  $d_t$  is the distance between  $x_i$  and  $y_i^t$  ( $t = 1, 2, 3, \dots$ ).
3. Illustration of the rigid transformation from the surface  $X$  to the surface  $Y$ . The first row compares the translated surfaces  $X$  with  $Y$ . Black points on the surface  $Y$  are used to compute the pseudo center. The second row compares the translated and rotated surfaces  $X$  with  $Y$ . Red axes represent the pose of the surface  $X$ , green axes for  $Y$ .
4. Illustration of labeling spine surface points on the coronal plane.
5. Illustration of labeling spine surface points on the sagittal plane.
6. Illustration of labeling spine surface points on the transverse plane.
7. Illustration of the leave-one-out experiments.
8. Illustration of the reconstruction error  $\overline{d_{Mean}}$  and  $\overline{d_{RMS}}$  in terms of  $N_{ref}^{low}$ , from the leave-one-out experiments.
9. Illustration of the processing time  $\overline{T}$  in terms of  $N_{ref}^{low}$ .
10. Illustration of how the surface distance is decreased in each step of the two-level framework.

11. Illustration of the high resolution surface  $X^H$  and  $Y^H$ . Point-to-surface distances from  $X^H$  to  $Y^H$  are illustrated by a HSV color map: the color of each vertex on  $X^H$  represents the distance  $d(p, Y^H), p \in X^H$ . (Both surface (left) and mesh (right) are showed in the bottom row.)
12. Illustration of the low resolution surface  $X^L$  and  $Y^L$  after applying the mesh simplification to both  $X^H$  and  $Y^H$ . Point-to-surface distances from  $X^L$  to  $Y^L$  are illustrated by a HSV color map: the color of each vertex on  $X^L$  represents the distance  $d(p, Y^L), p \in X^L$ . (Both surface (left) and mesh (right) are showed in the bottom row.)
13. Illustration of the deformed low resolution surfaces  $X^{L(1)}$  and  $Y^L$  after applying Chui and Rangarajan's non-rigid registration to  $X^L$ . Point-to-surface distances from  $X^{L(1)}$  to  $Y^L$  are illustrated by a HSV color map: the color of each vertex on  $X^{L(1)}$  represents the distance  $d(p, Y^L), p \in X^{L(1)}$ . (Both surface (left) and mesh (right) are showed in the bottom row.)
14. Illustration of the deformed low resolution surfaces  $X^{L(2)}$  and  $Y^L$  after applying a refining procedure to  $X^{L(1)}$ . Point-to-surface distances from  $X^{L(2)}$  to  $Y^L$  are illustrated by a HSV color map: the color of each vertex on  $X^{L(2)}$  represents the distance  $d(p, Y^L), p \in X^{L(2)}$ . (Both surface (left) and mesh (right) are showed in the bottom row.)
15. Illustration of the deformed high resolution surfaces  $X^{H(1)}$  and  $Y^H$  after applying the interpolation to  $X^H$ . Point-to-surface distances from  $X^{H(1)}$  to  $Y^H$  are illustrated by a HSV color map: the color of each vertex on  $X^{H(1)}$  represents the distance  $d(p, Y^H), p \in X^{H(1)}$ . (Both surface (left) and mesh (right) are showed in the bottom row.)

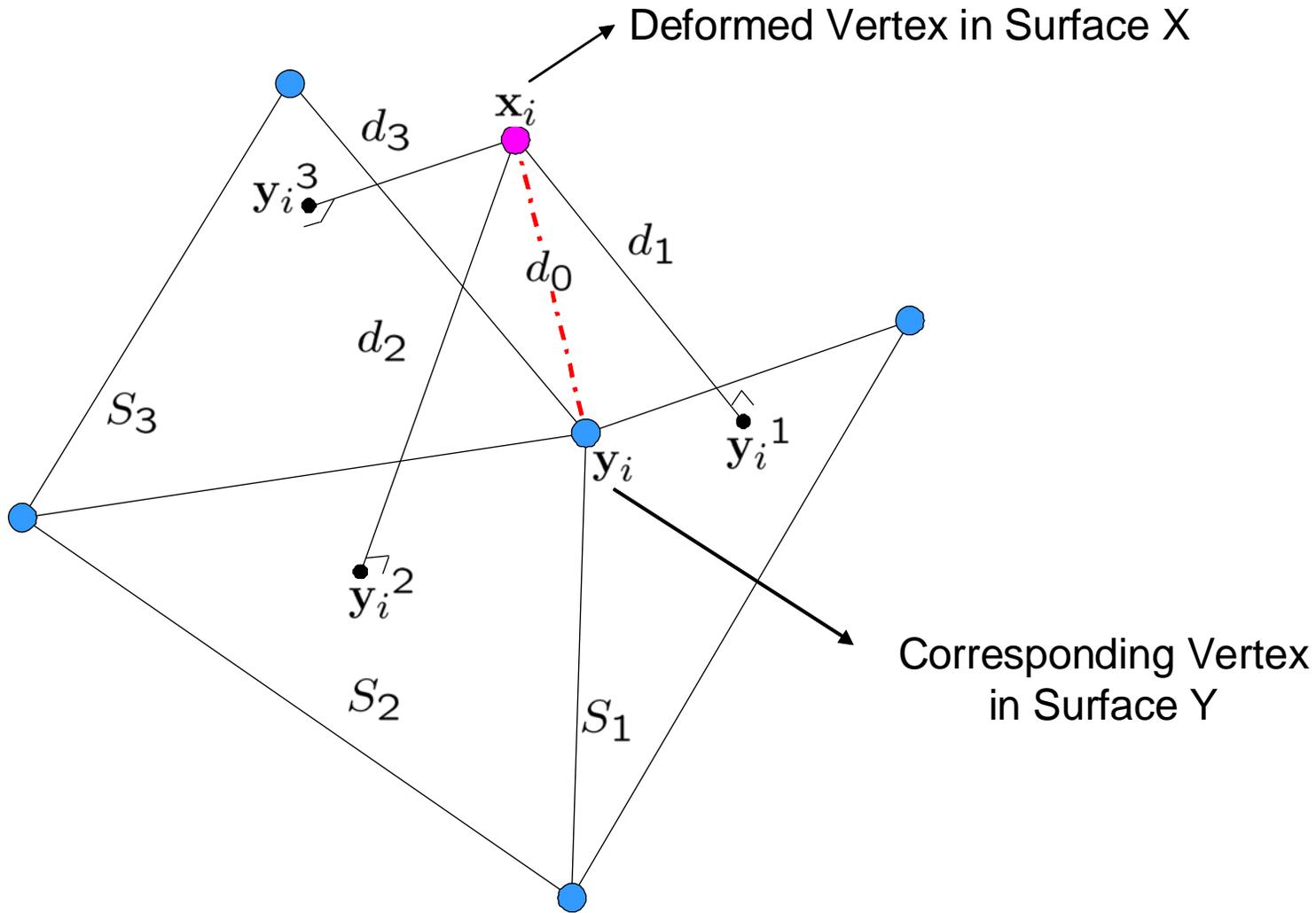
16. Illustration of the deformed high resolution surfaces  $X^{H(2)}$  and  $Y^H$  after applying a refining procedure to  $X^{H(1)}$ . Point-to-surface distances from  $X^{H(2)}$  to  $Y^H$  are illustrated by a HSV color map: the color of each vertex on  $X^{H(2)}$  represents the distance  $d(p, Y^H), p \in X^{H(2)}$ . (Both surface (left) and mesh (right) are showed in the bottom row.)
17. Illustration of the RMS and Mean error distributions for our method and ICP.
18. Femoral head atlas: illustration of the shape variations in the first, second and third mode. The first mode encodes 41.50% of shape variations, the second and third modes encode 19.78% and 7.64% of shape variations. The color highlights the local shape variations.
19. Condyles atlas: illustration of the shape variations in the first, second and third mode. The first mode encodes 25.70% of shape variations, the second and third modes encode 9.75% and 8.21% of shape variations. The color highlights the local shape variations.
20. Entire femur atlas: 27.67% of shape variations is encoded in the first mode. The color highlights the local shape variations.
21. Entire femur atlas: 18.63% of shape variations is encoded in the second mode. The color highlights the local shape variations.
22. Illustration of the shape variations encoded in modes of atlases of the femoral head (blue), condyles (pink) and entire femur (green).
23. Spine vertebra atlas: 39.0% of shape variations is encoded in the first mode. The second and third modes encode 20.4% and 12.8% of shape variations.
24. Illustration of the shape variations encoded in modes of the spine vertebra atlas.

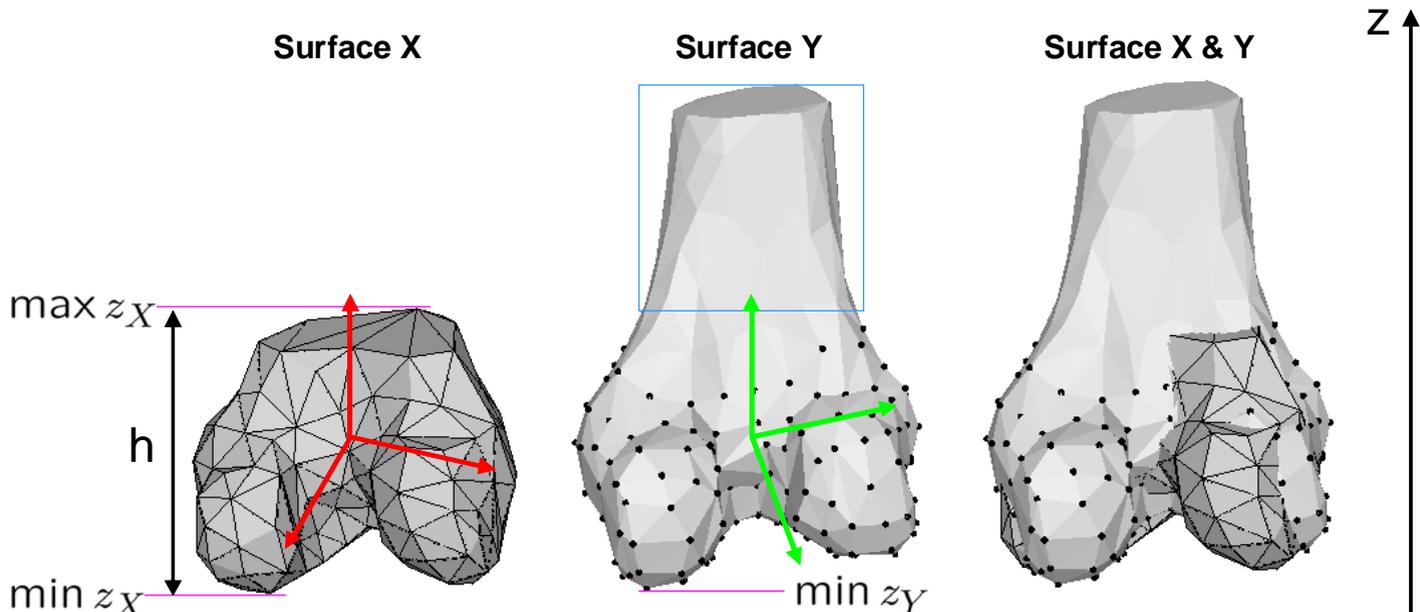


(a)



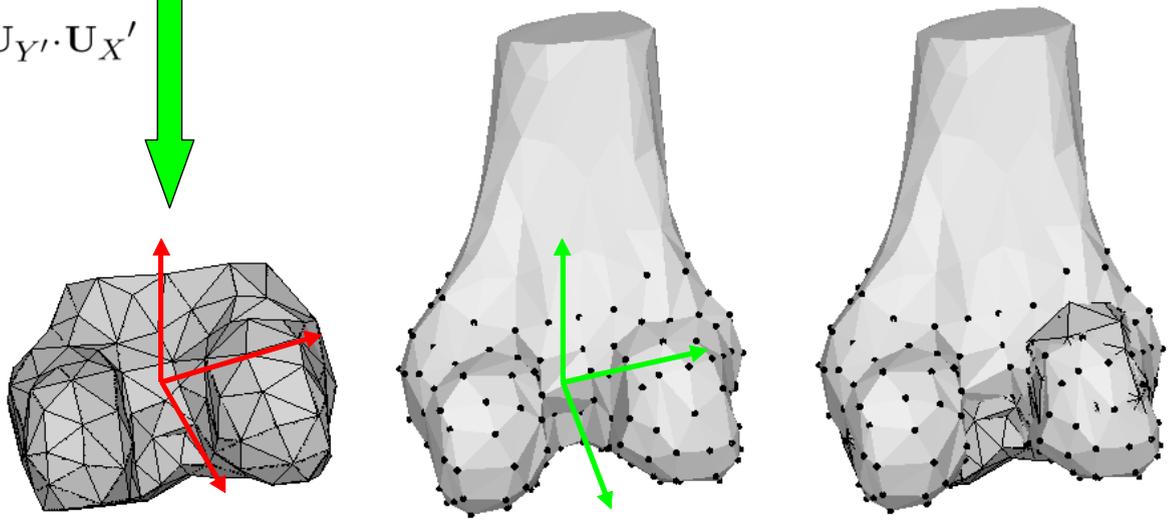
(b)



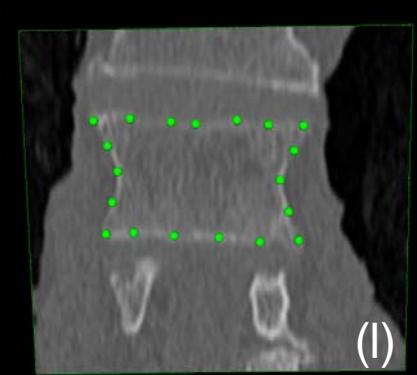
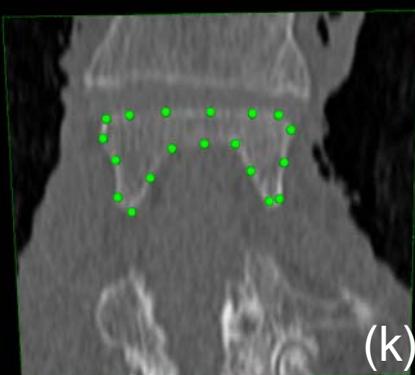
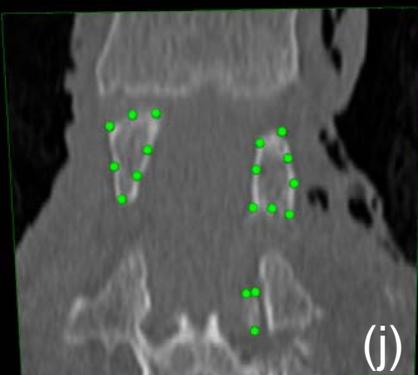
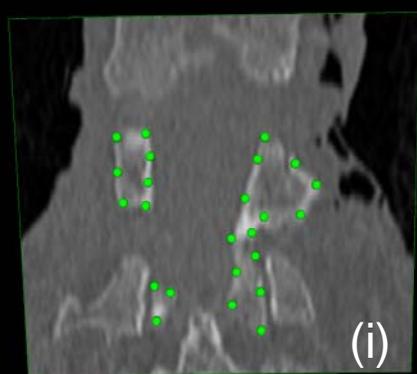
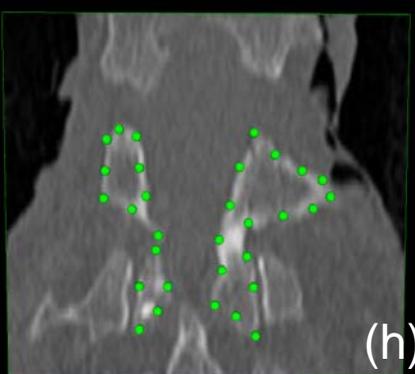
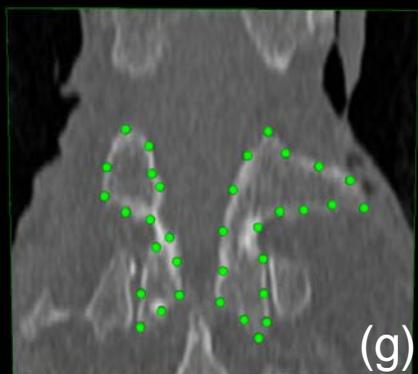
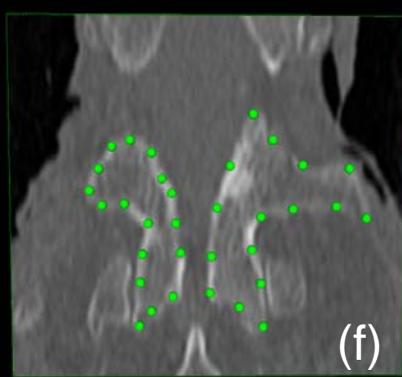
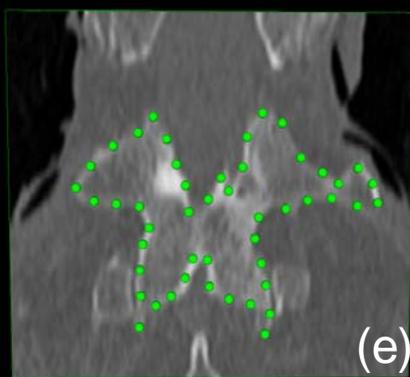
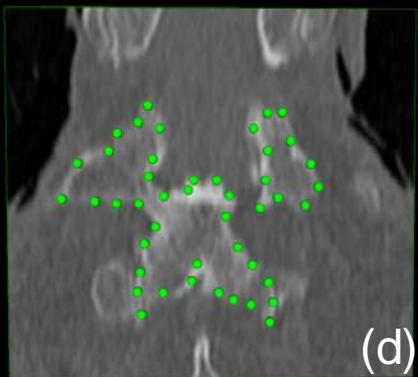
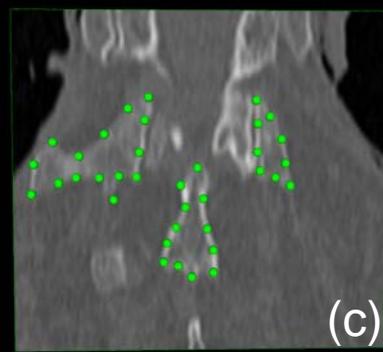
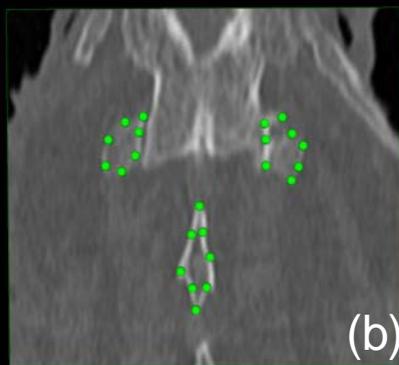
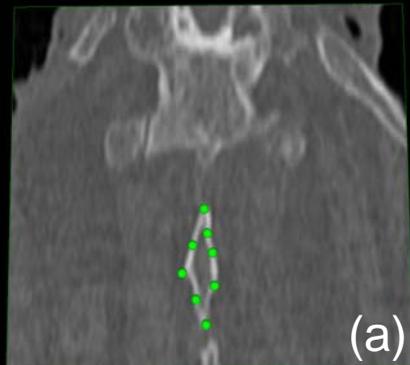


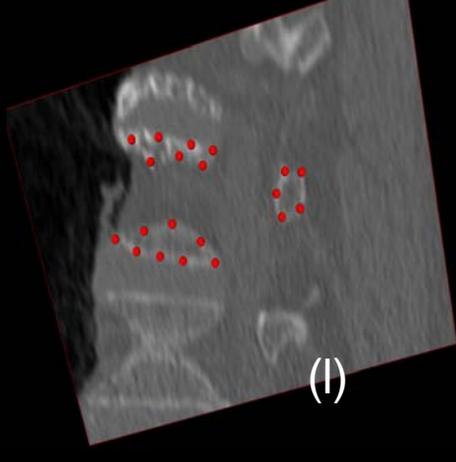
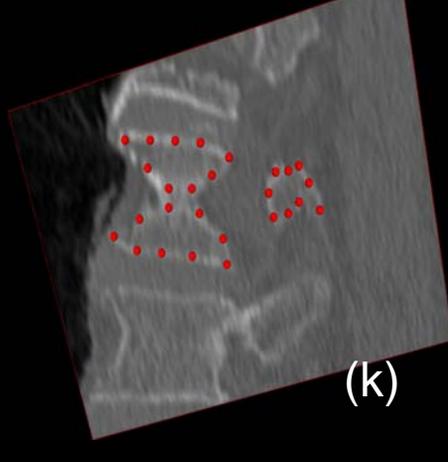
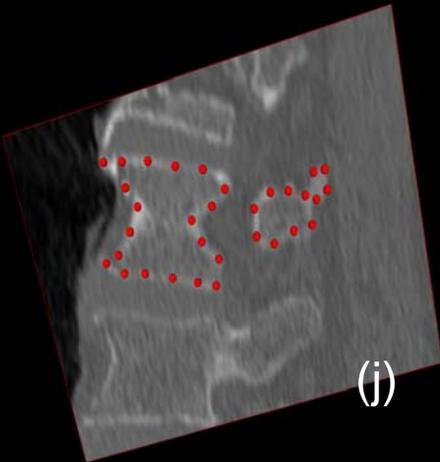
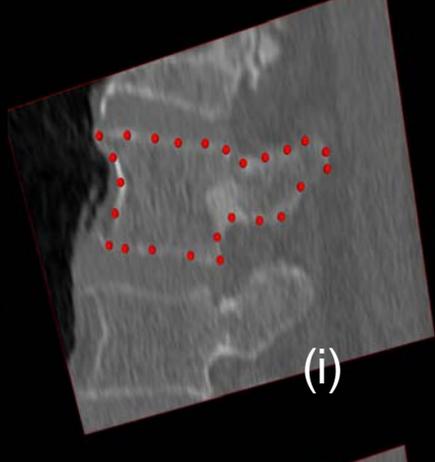
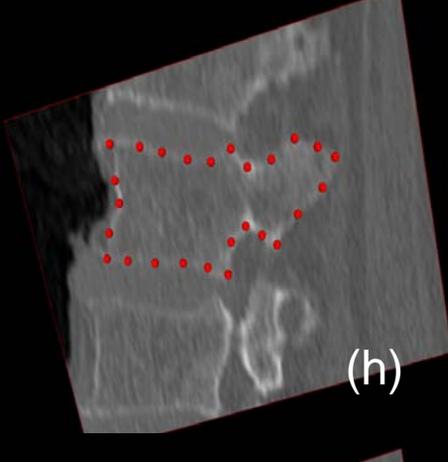
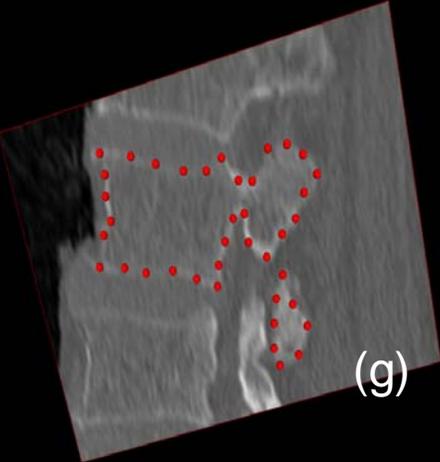
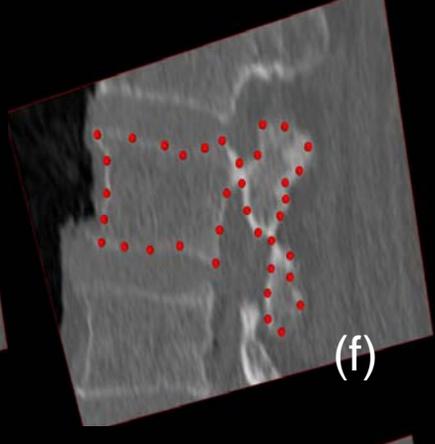
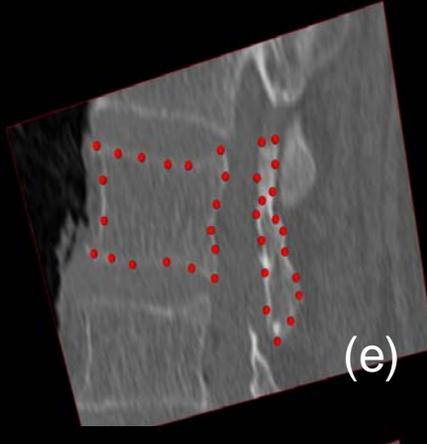
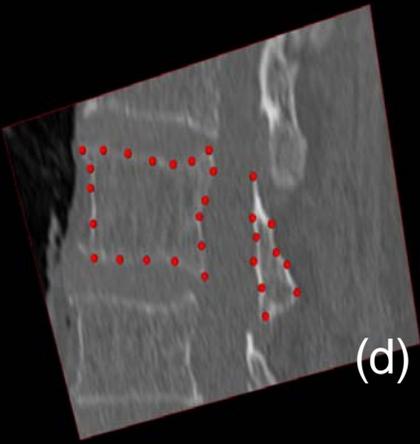
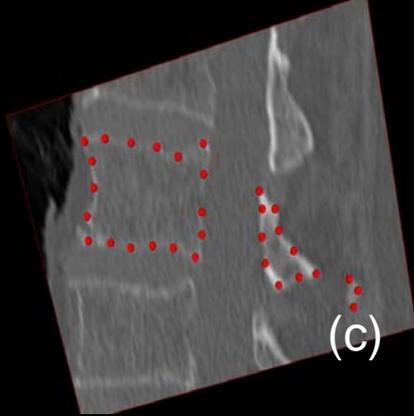
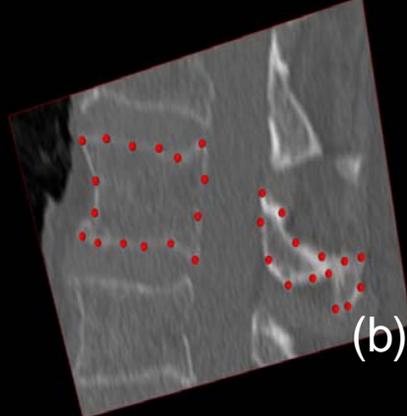
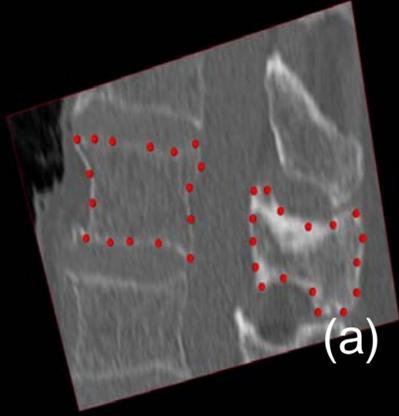
Translate Surface X to the pseudo center of Surface Y

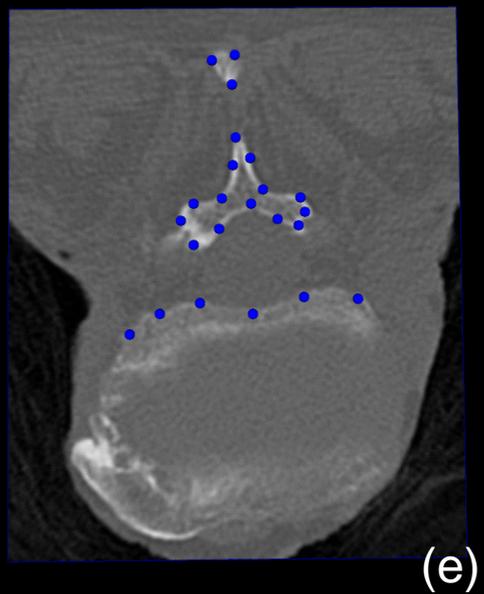
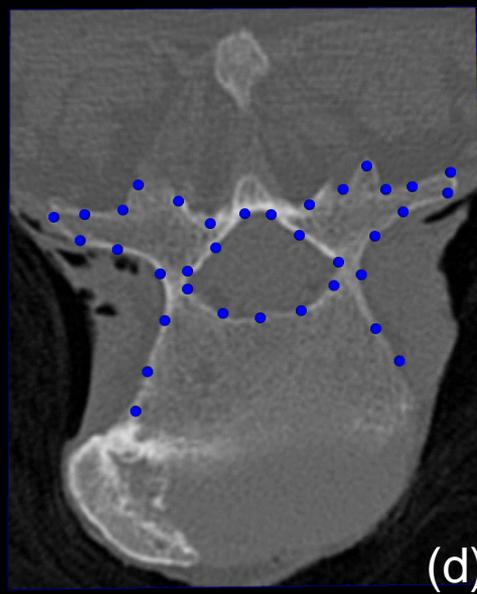
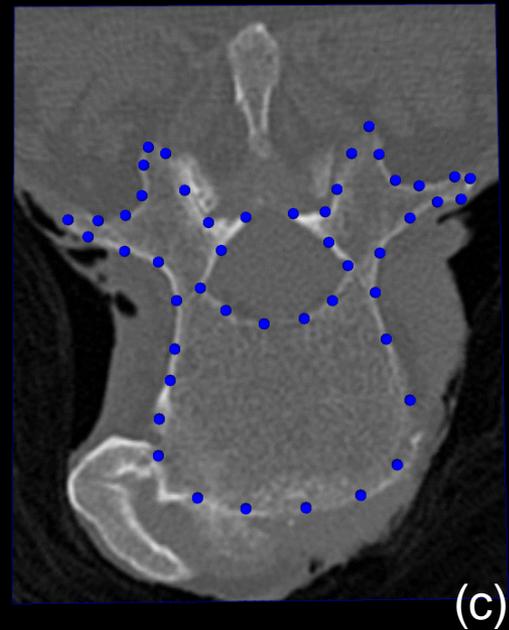
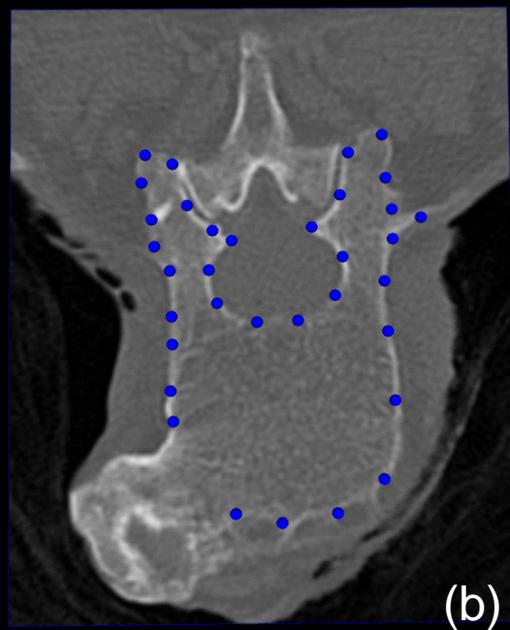
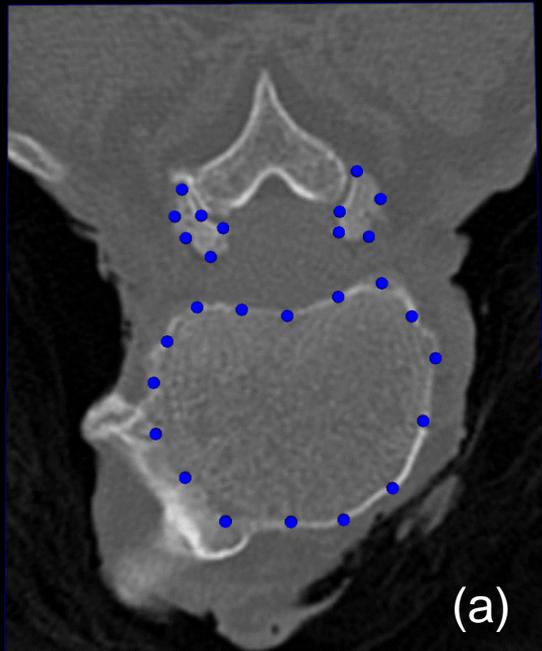
$U_{Y'} \cdot U_{X'}$

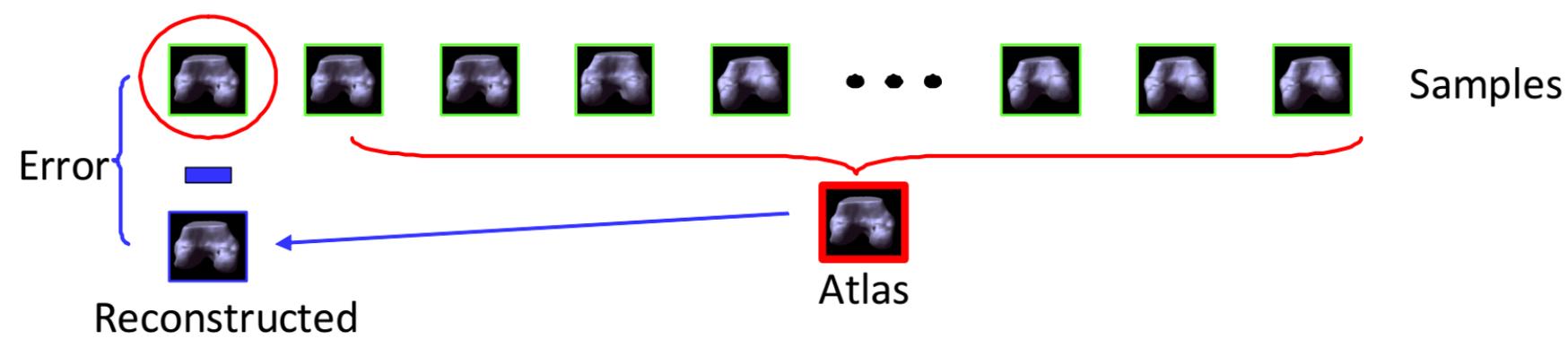


Rotate Surface X to the same pose as Surface Y

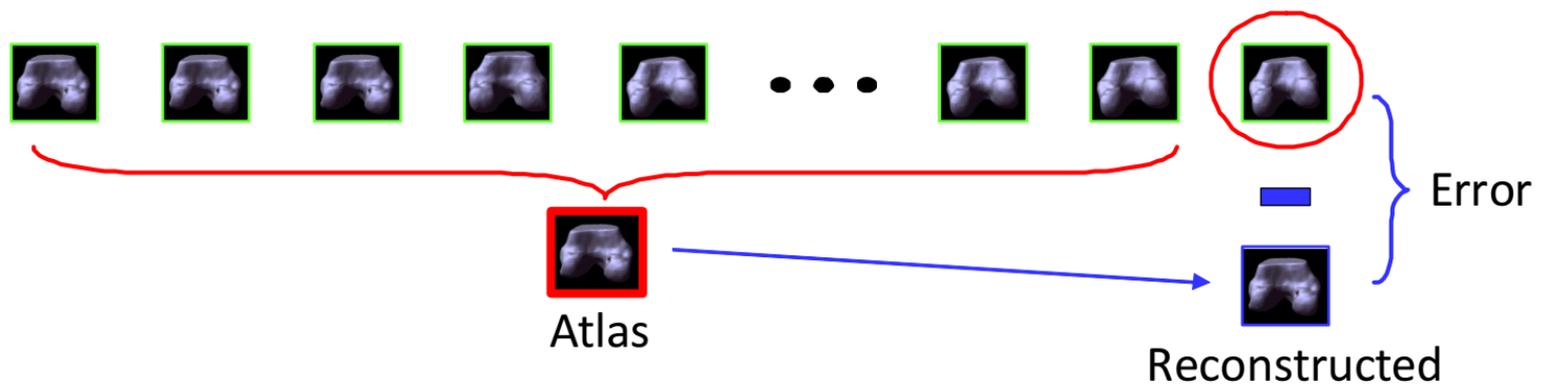






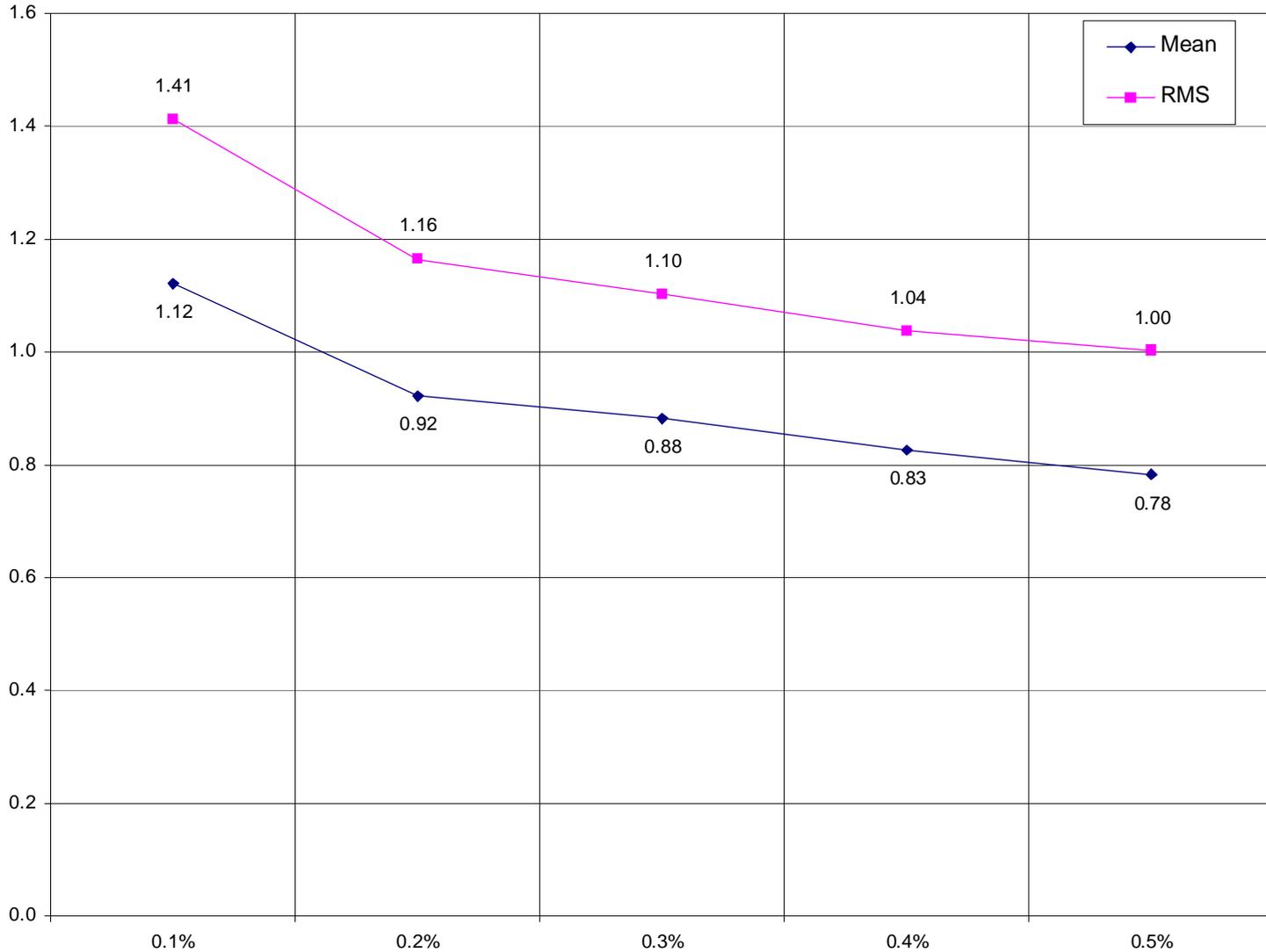


⋮



# Average Surface Errors (Based on 87\*5 leave-one-out experiments)

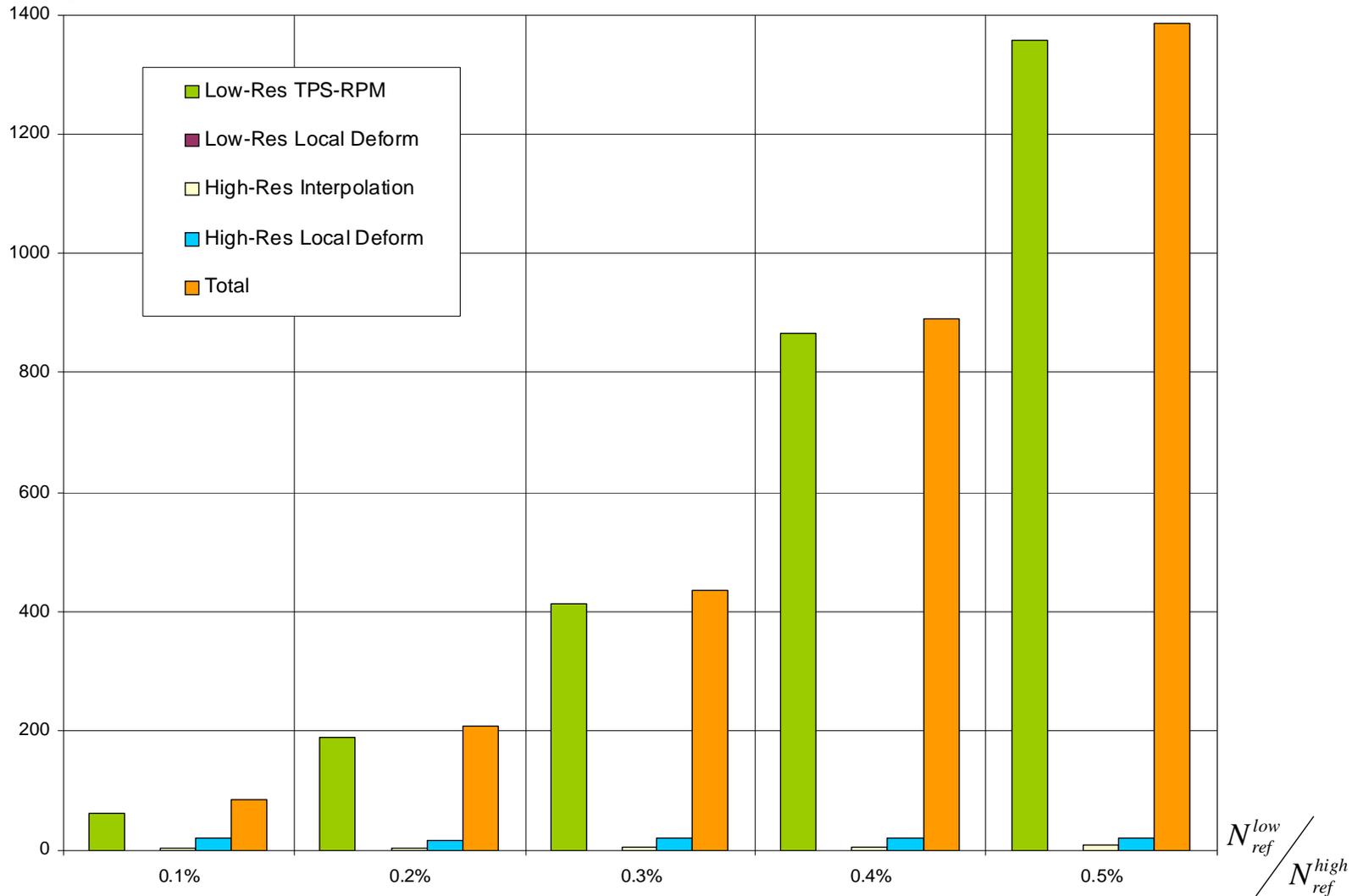
Error (mm)



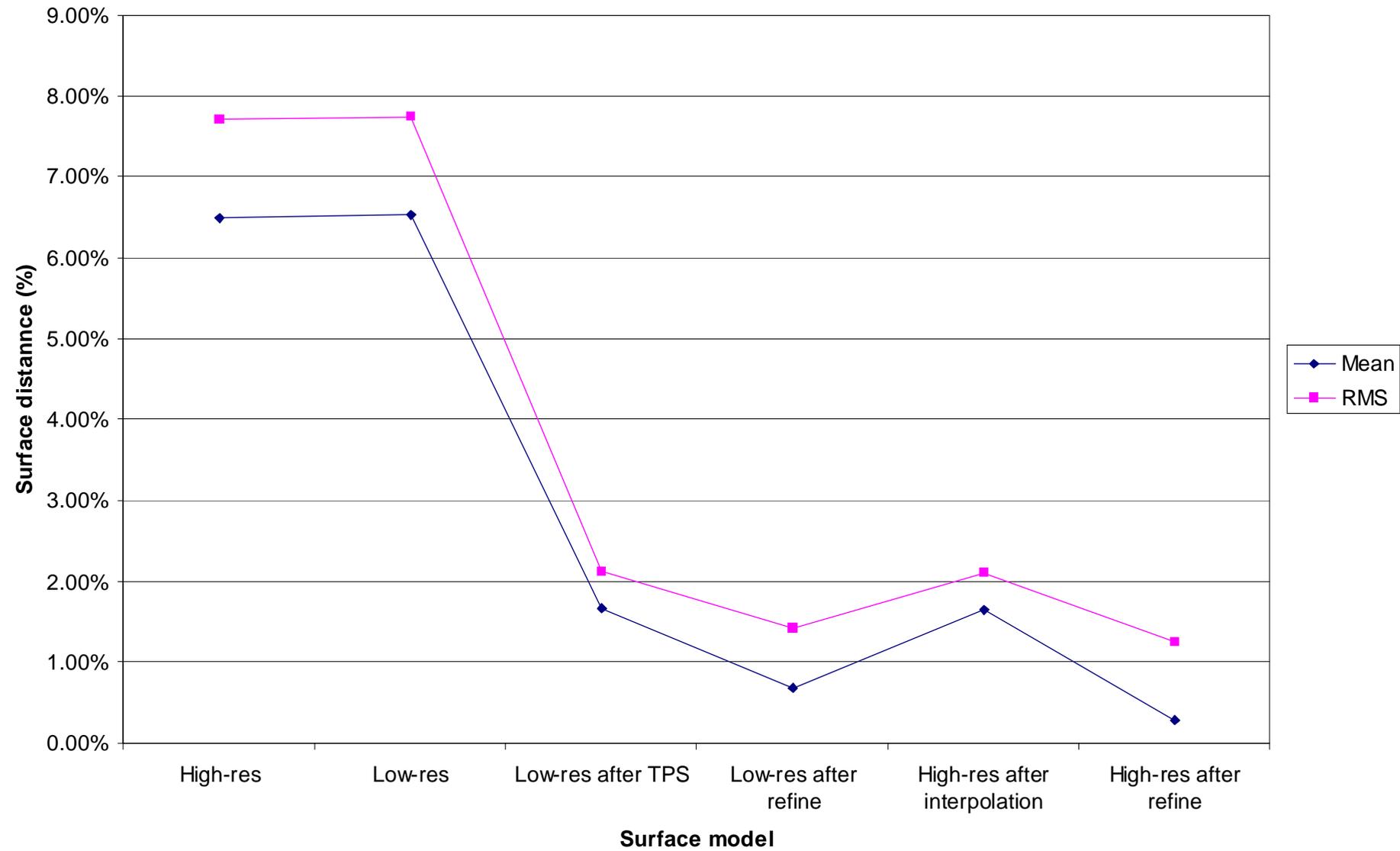
$N_{ref}^{low} / N_{ref}^{high}$

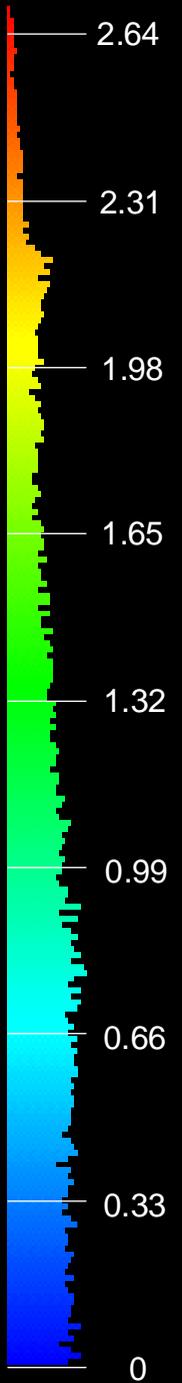
# Average Processing Time (Based on 86\*5 registration results)

Processing time (second)



# Registration error decreased by two-level framework

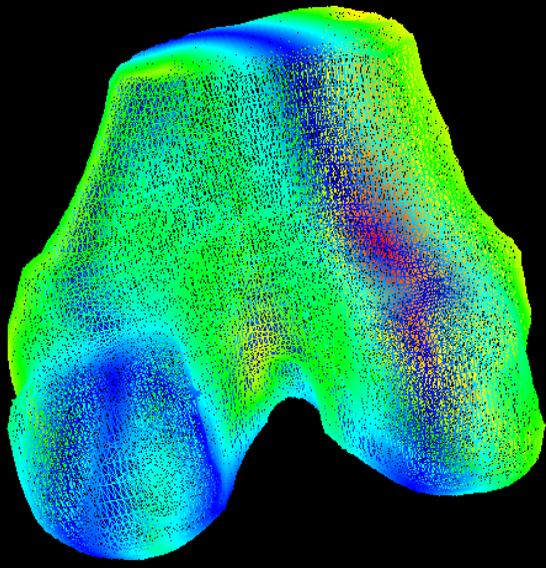
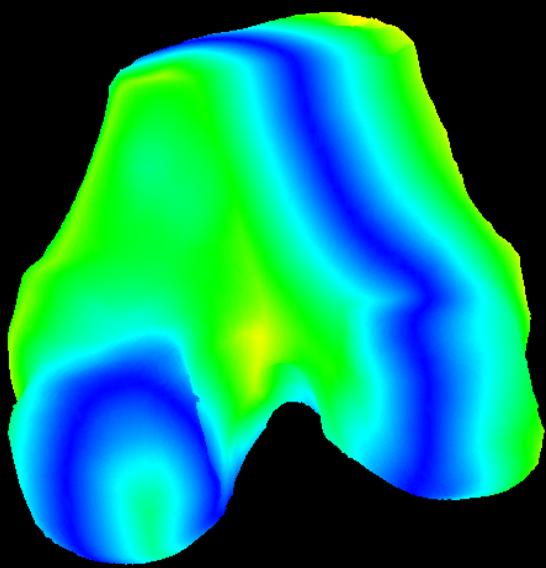


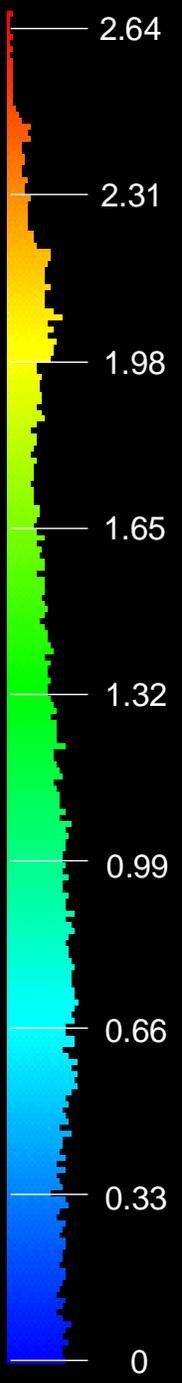


**High-Resolution**

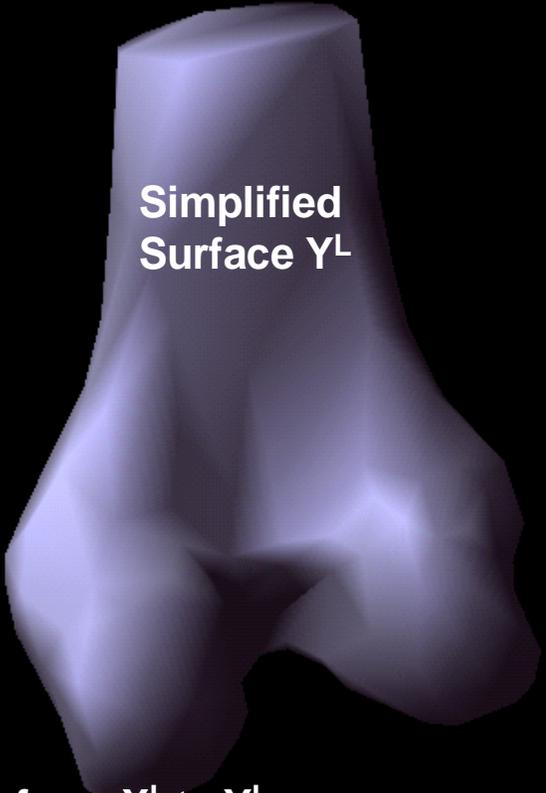
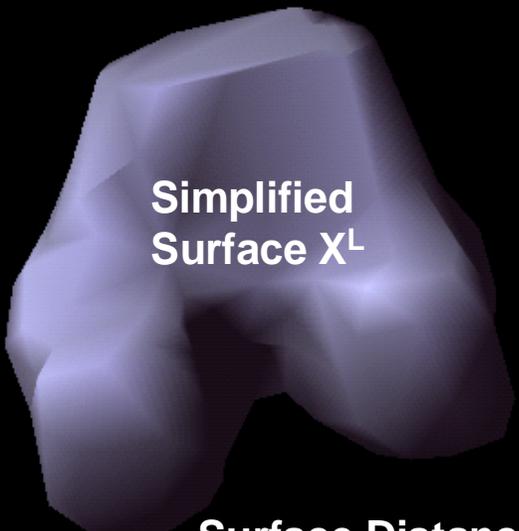


**Surface Distance from  $X^H$  to  $Y^H$**

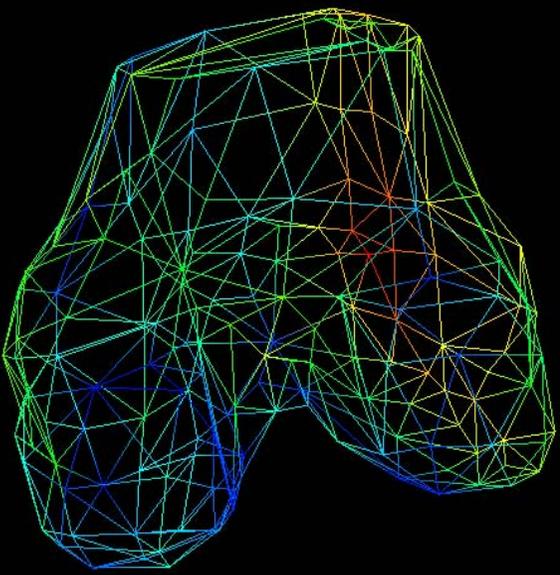
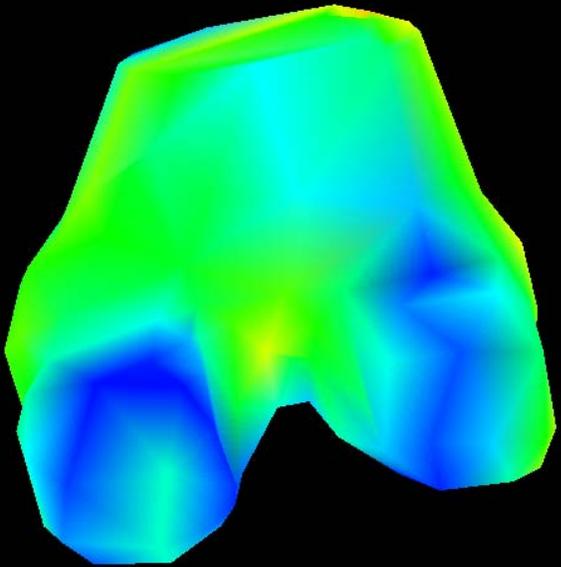


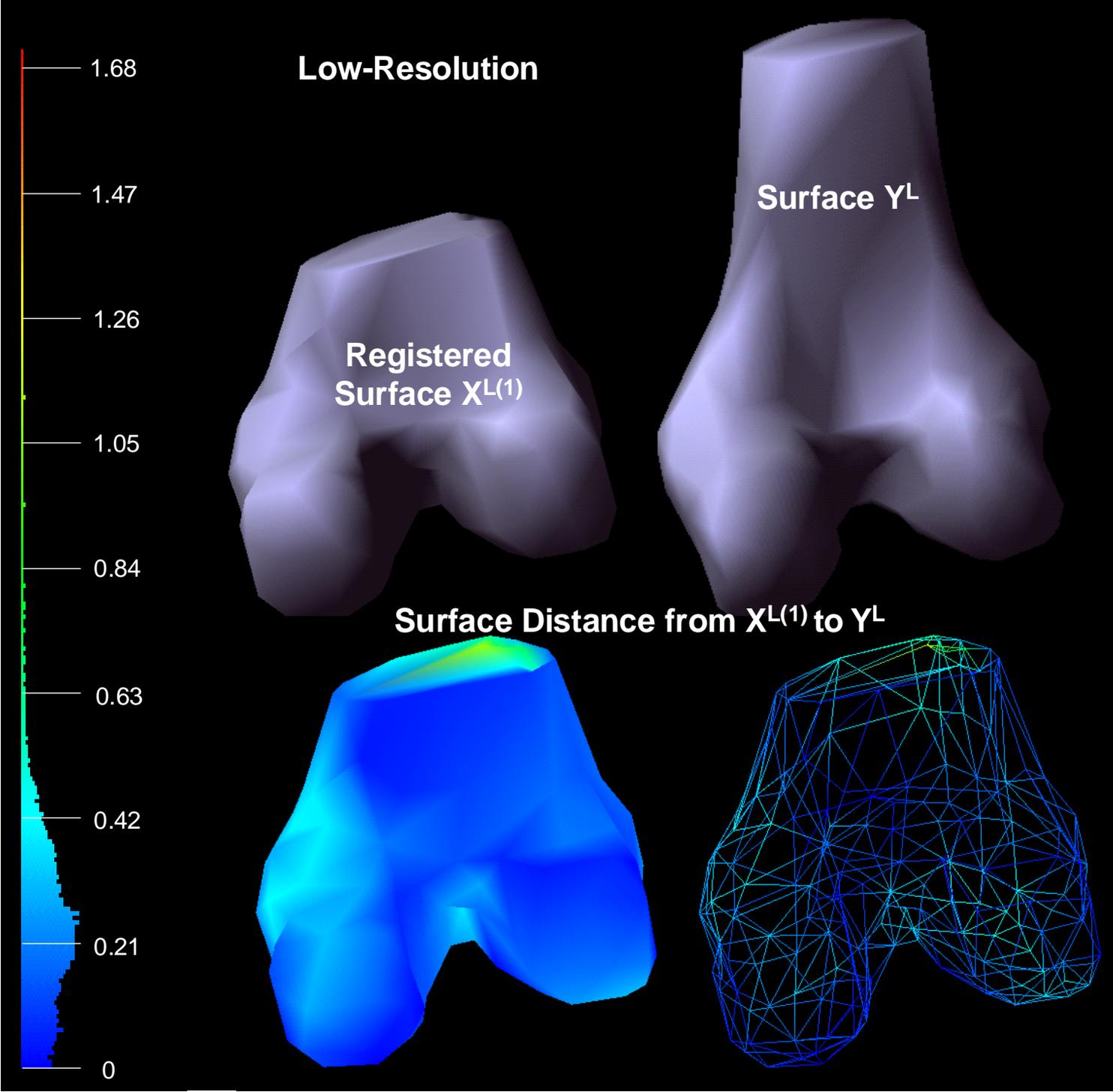


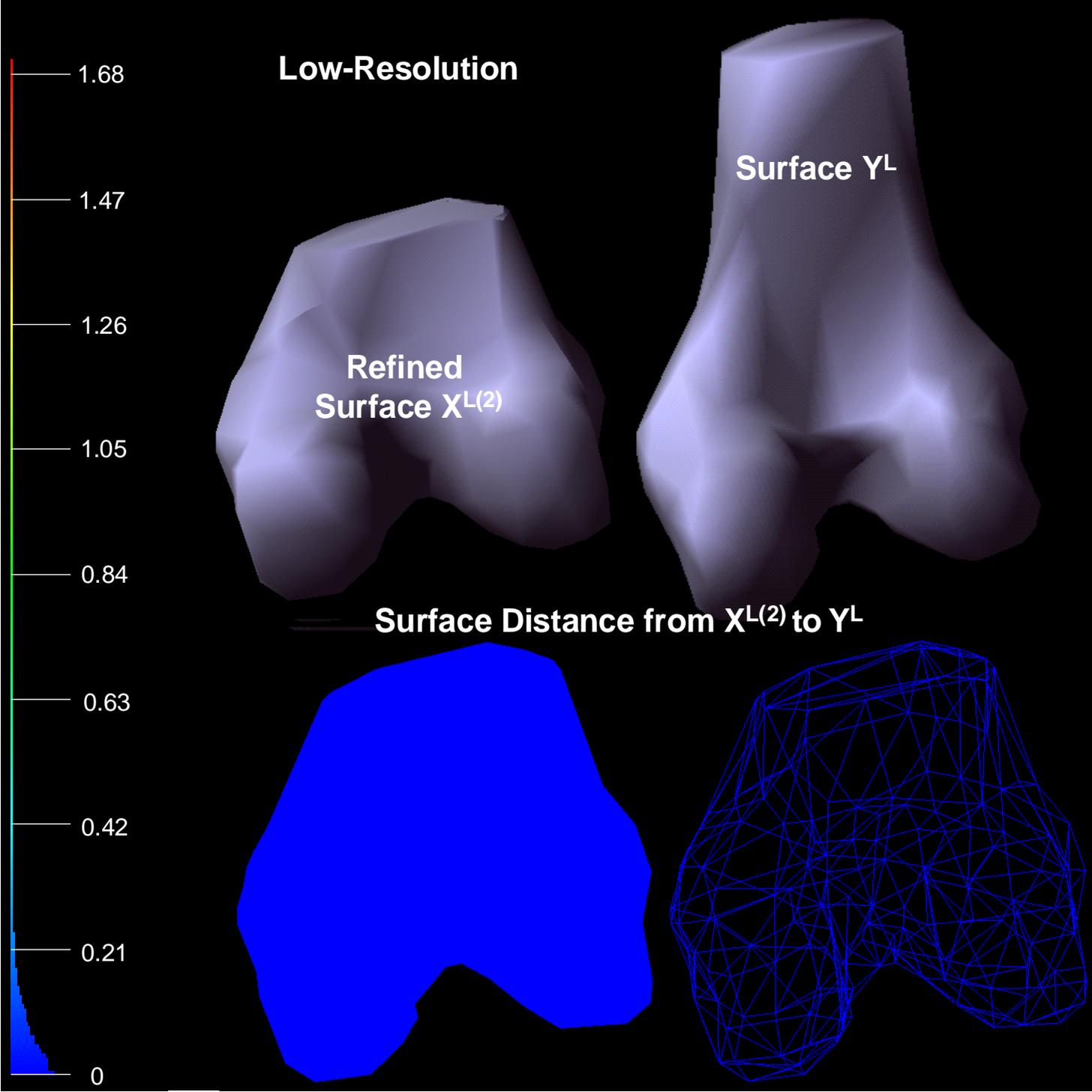
**Low-Resolution**

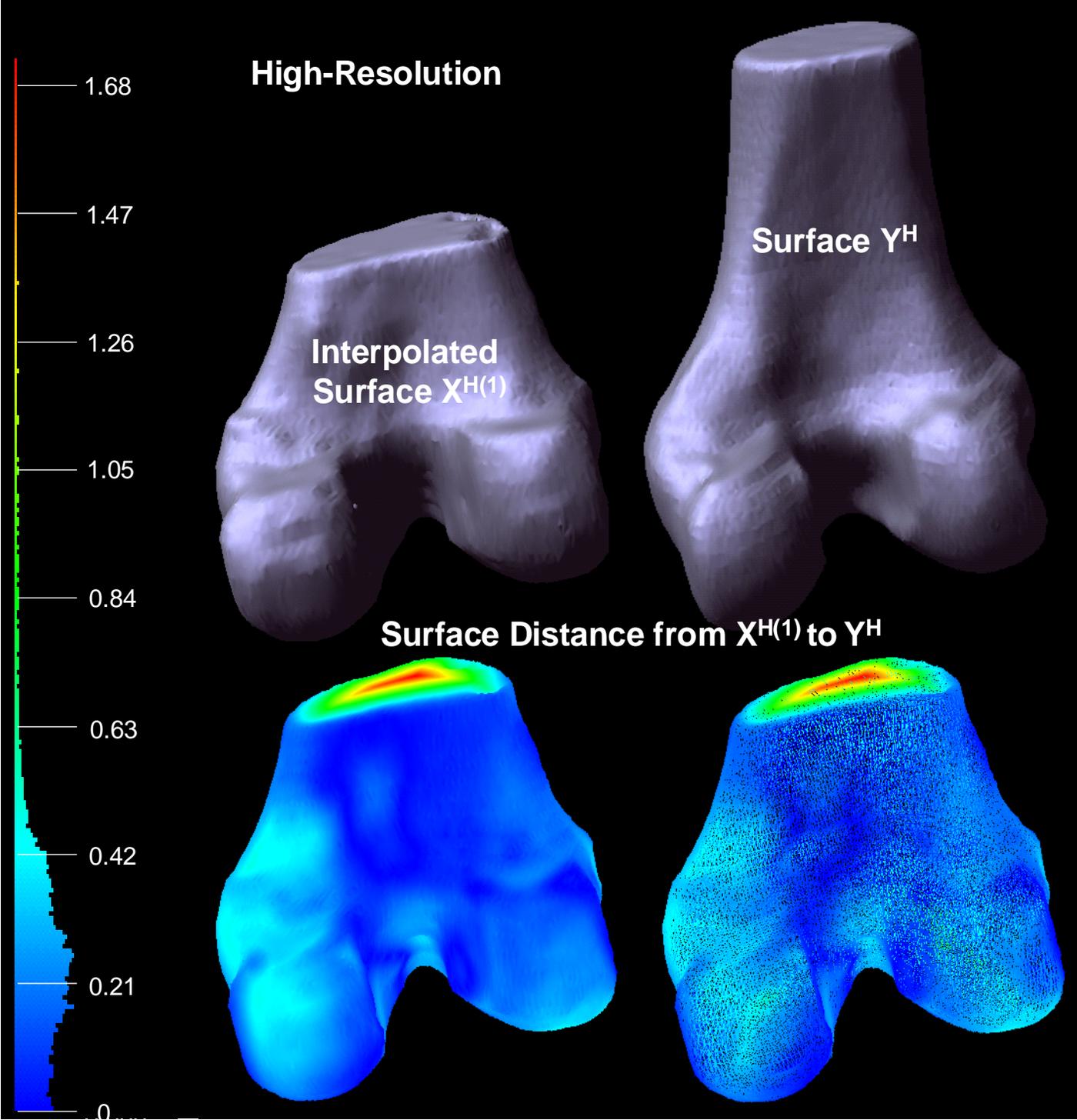


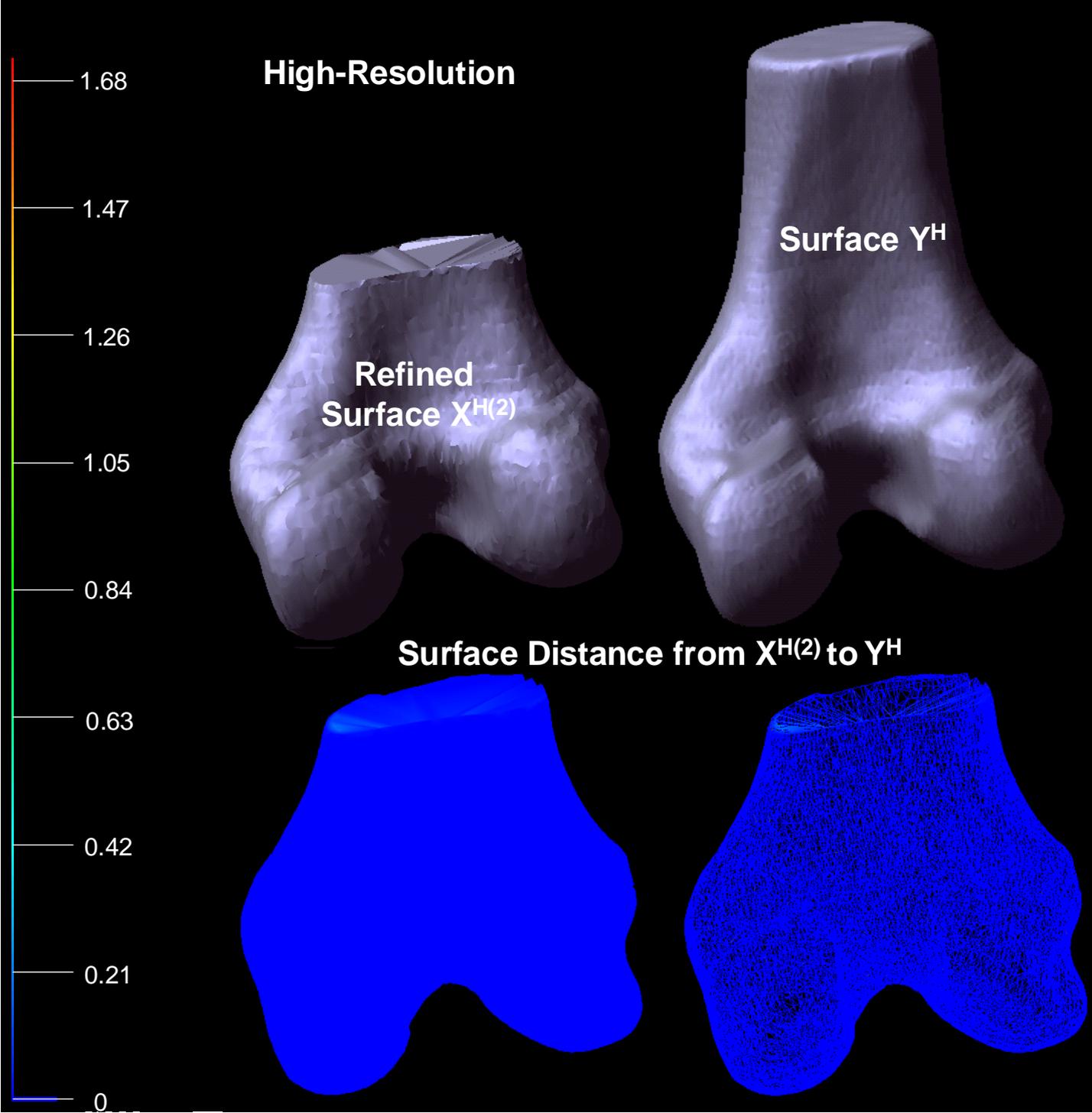
**Surface Distance from  $X^L$  to  $Y^L$**



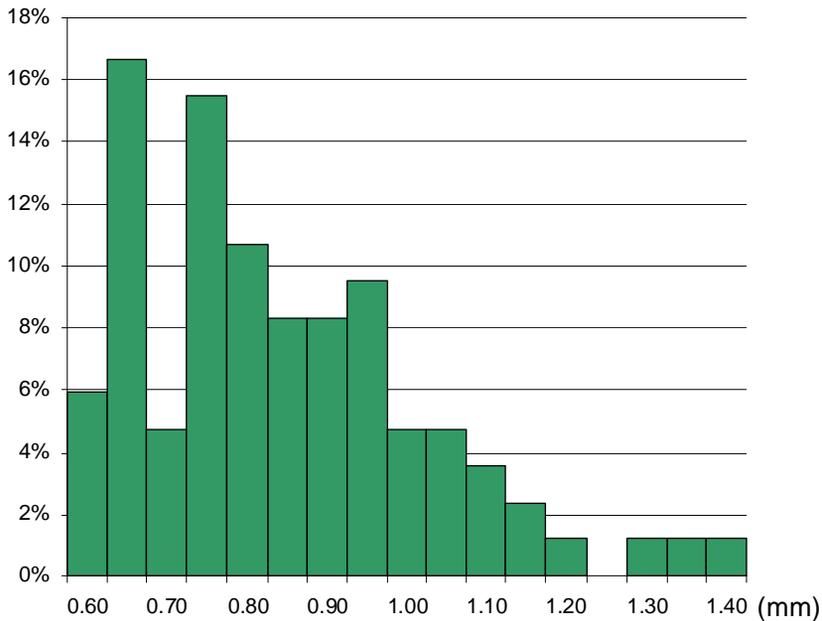




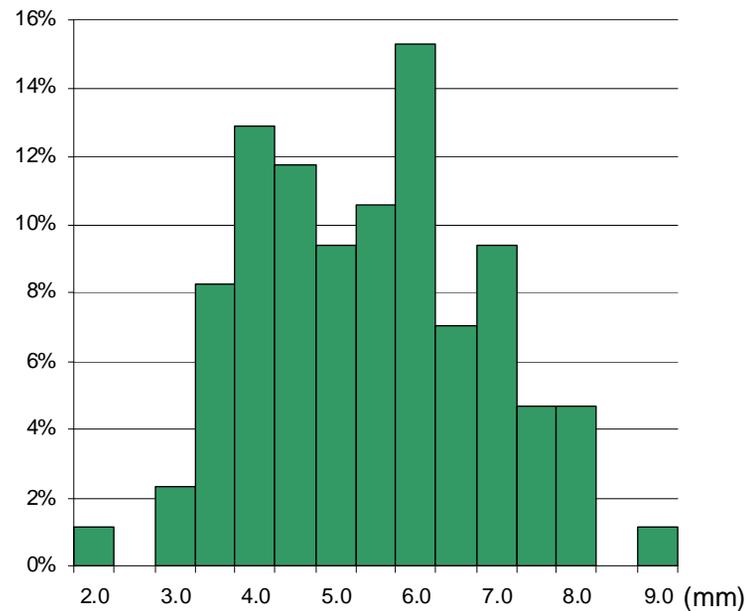




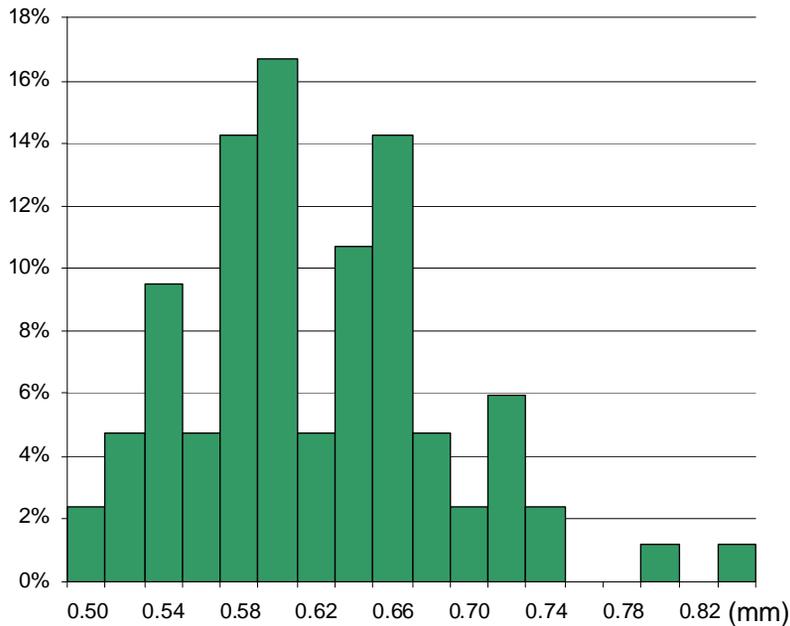
**RMS Error Distribution for Our Method**  
(Based on 87 patients)



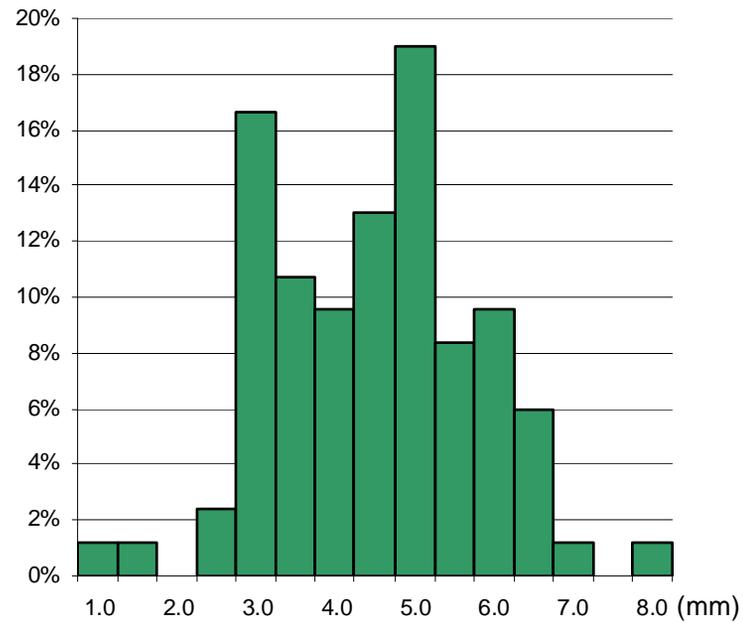
**RMS Error Distribution for ICP**  
(Based on 87 patients)



**Mean Error Distribution for Our Method**  
(Based on 87 patients)



**Mean Error Distribution for ICP**  
(Based on 87 patients)



$\bar{y} - 2\sigma_i$

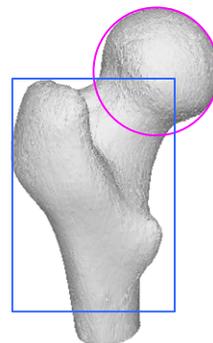
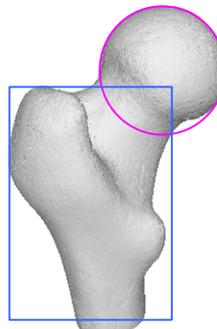
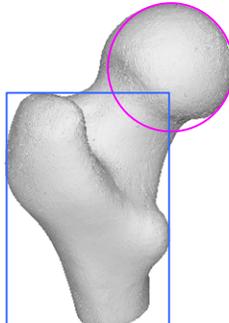
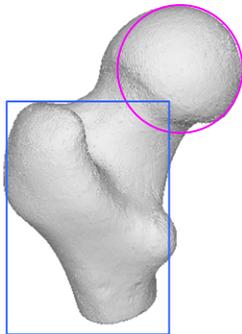
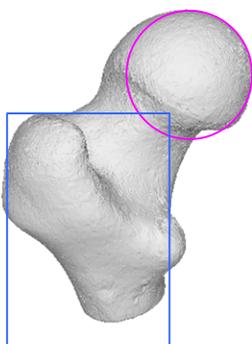
$\bar{y} - \sigma_i$

$\bar{y}$

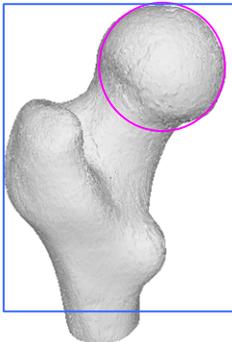
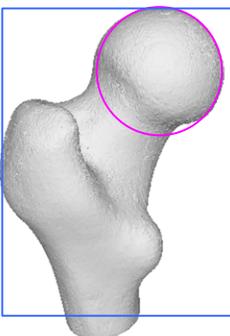
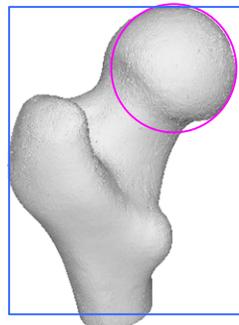
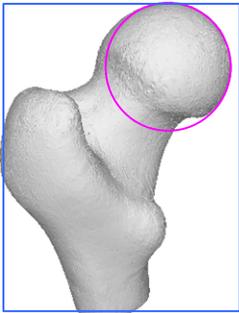
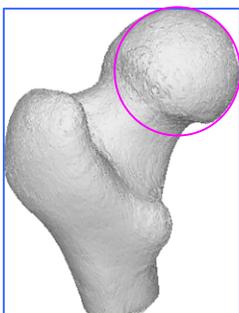
$\bar{y} + \sigma_i$

$\bar{y} + 2\sigma_i$

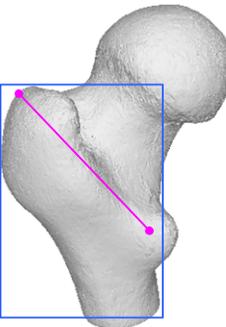
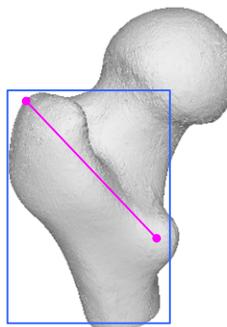
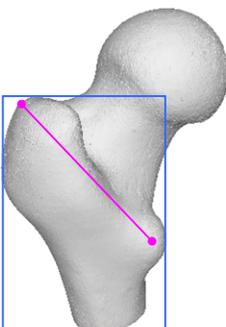
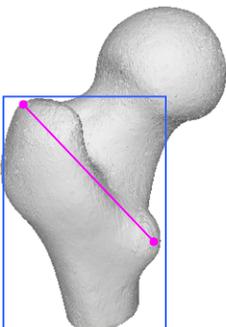
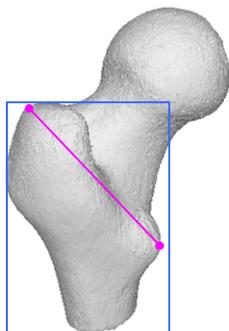
**mode 1  
(41.50%)**

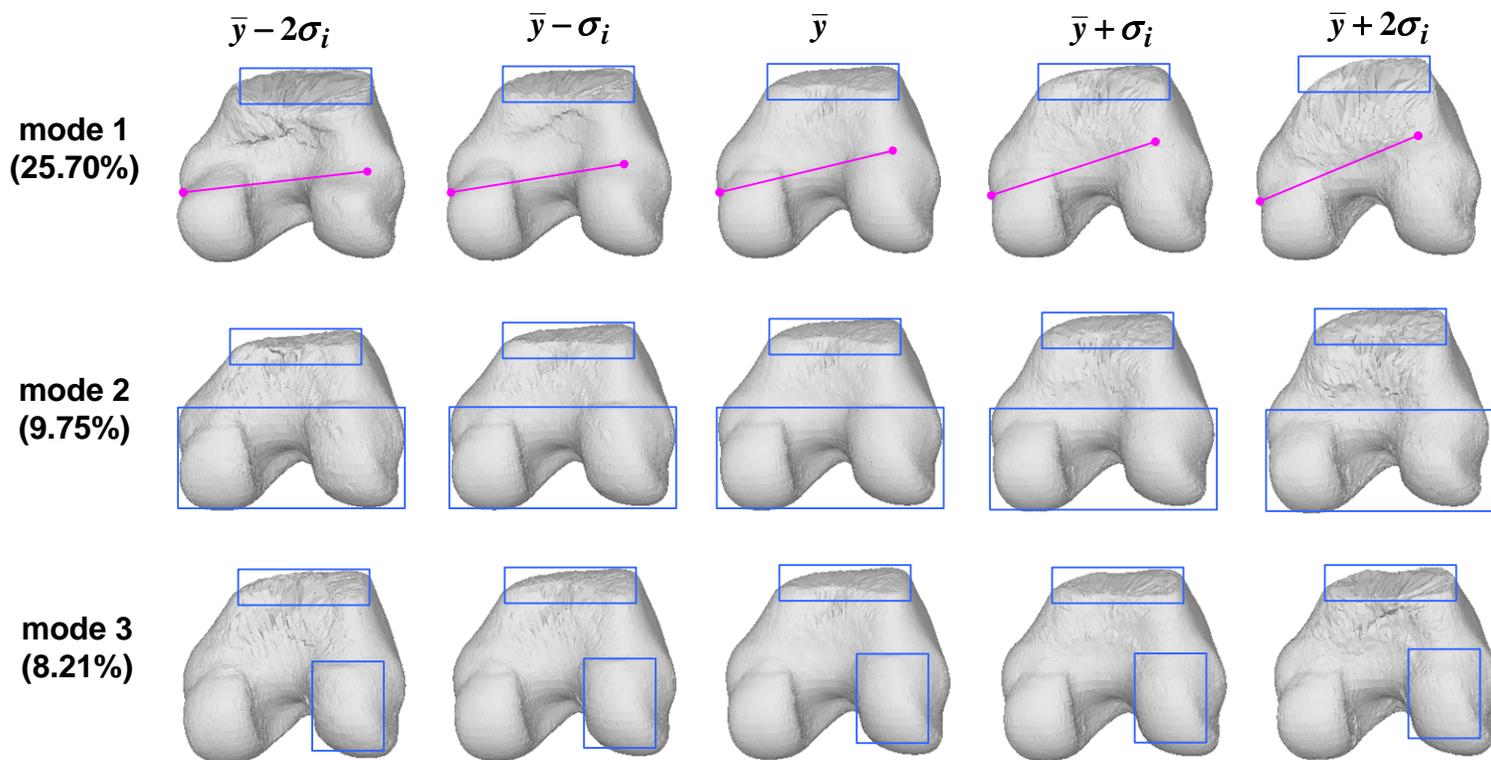


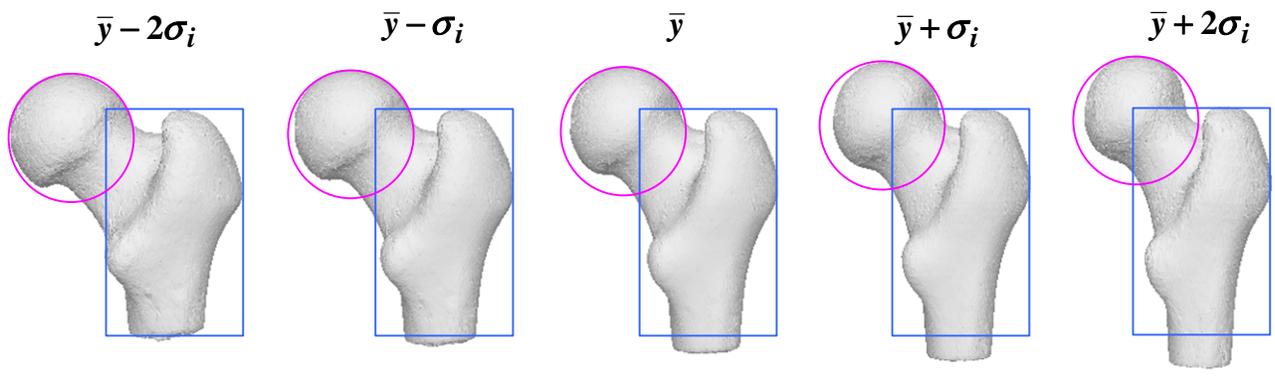
**mode 2  
(19.78%)**



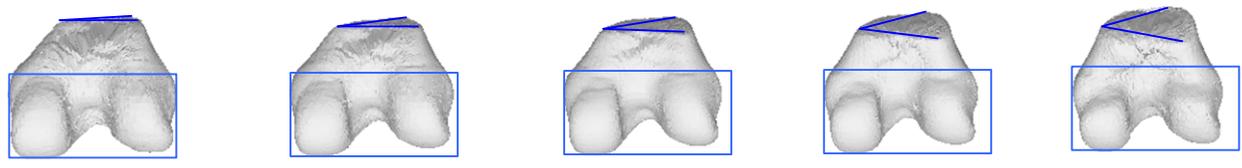
**mode 3  
(7.64%)**







**mode 1**  
**(27.67%)**



$\bar{y} - 2\sigma_i$



$\bar{y} - \sigma_i$



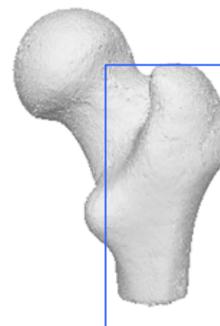
$\bar{y}$



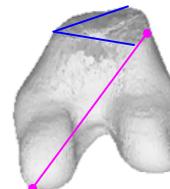
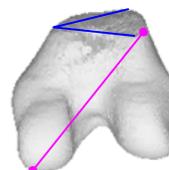
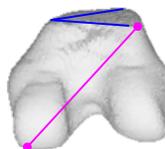
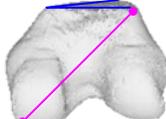
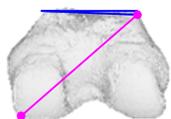
$\bar{y} + \sigma_i$

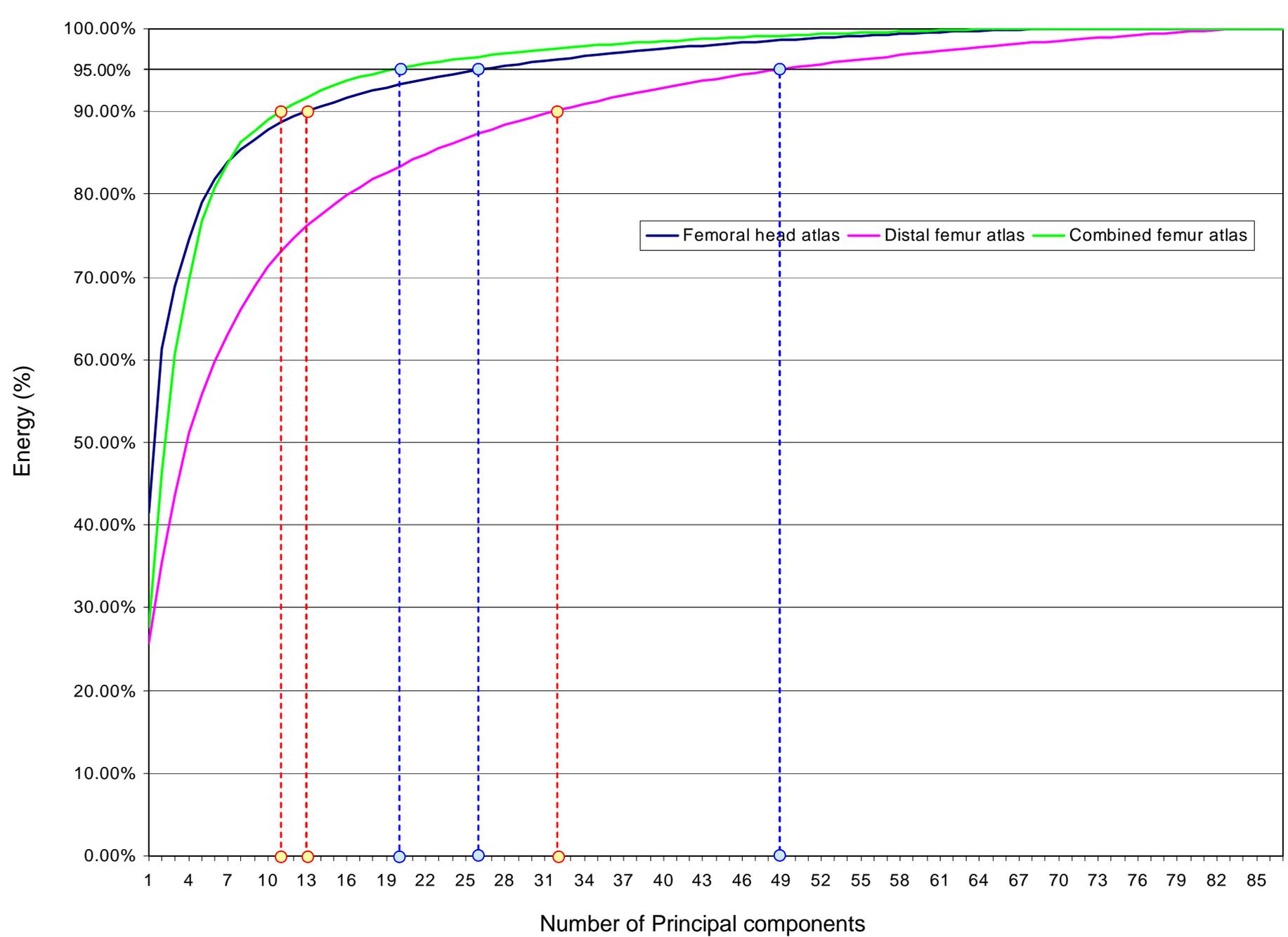


$\bar{y} + 2\sigma_i$



**mode 2**  
**(18.63%)**





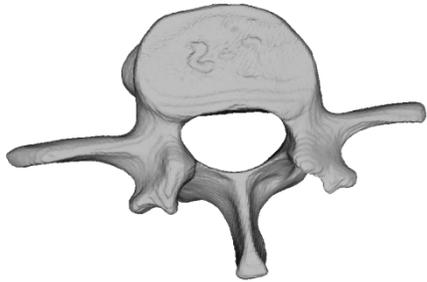
$\bar{y} - \sigma$

$\bar{y} - 0.5\sigma$

$\bar{y}$

$\bar{y} + 0.5\sigma$

$\bar{y} + \sigma$



mode 1 (39.0%)

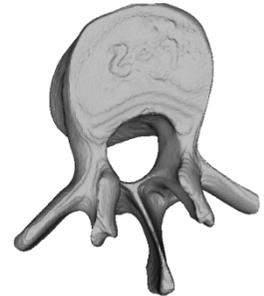
$\bar{y} - \sigma$

$\bar{y} - 0.5\sigma$

$\bar{y}$

$\bar{y} + 0.5\sigma$

$\bar{y} + \sigma$



mode 2 (20.4%)

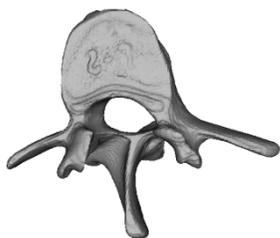
$\bar{y} - \sigma$

$\bar{y} - 0.5\sigma$

$\bar{y}$

$\bar{y} + 0.5\sigma$

$\bar{y} + \sigma$



mode 3 (12.8%)

# Spine Vertebra Atlas

