Finding Complete 3D Vertex Correspondence for Statistical Shape Modeling

Robert Ieuan Palmer, Xianghua Xie, and Gary Tam

Abstract— **A statistical shape model that accurately generalizes a family of 3D shapes requires establishing correspondences across the set of shapes. However in 3D anatomical meshes, finding a sufficient number of landmarks to accurately describe the shape can be a challenge, and often only a few points are easily identifiable due to the smooth nature of the object surface. Using a *sparse* set of landmarks, this paper finds a *dense* set of vertex correspondences across a set of 3D aortic root meshes. This is achieved by non-rigidly transforming a source mesh to a target mesh. Then, for *every* vertex on the target, a corresponding vertex on the deformed source is found, resulting in complete correspondence. A more accurate transformation results in better correspondence establishment, and our mesh registration experiments show an average Hausdorff distance of $3.65mm$, and an average point-to-mesh distance of $0.41mm$, i. e. within one voxel.**

I. INTRODUCTION

One of the most popular approaches to deformable model based segmentation is through the use of active shape model (ASM) [1] to statistically generalize shape variations and subsequently impose such generalization as a shape constraint, e.g. [2]. A common problem in deformable image segmentation is that certain vertices may deform towards the erroneous boundaries that may have strong features. Applying a statistical shape model (SSM) to the deformed model can ensure that those vertices do not deviate too far away from the training shapes, and this iterative process often leads to better segmentation. SSMs can be generated from a set of corresponding landmark points, however, finding the same number of well-defined, corresponding landmarks in all training shapes is not a trivial task. This is especially true for 3D meshes where the number of faces and vertices can vary, and the problem can be even more compounded in anatomical objects where landmark points are often sparsely separated due to their smooth surfaces.

A number of approaches to find correspondences on 3D shapes have been proposed, many of which work directly on the meshes themselves. Kim *et al.*[3] used blended intrinsic maps to find non-isometric correspondences in a range of shapes. Although this is an important piece of work, the method requires water-tight meshes which is not always possible in medical data. Meanwhile, Zhang *et al.*[4] used mesh deformation to find a set of *sparse* correspondences, however sparse correspondences are not always sufficient to build an accurate SSM of anatomical objects.

Alternatively, approaches have been proposed to find correspondences specifically on anatomical shapes. Subsol *et al.*[5], were among the first to propose using B-splines for non-rigid mesh registration, however it is thought that the line and point features used here may not be suitable for cardiac objects as they tend to have smooth surfaces [6]. Frangi *et al.*[6] proposed to find corresponding mesh landmarks by working with the corresponding image data. A non-rigid free-form-deformation (FFD) based on B-splines between each 3D mesh and an atlas was estimated using an intensity-based similarity metric. Although this method is an attractive alternative, using an image-based similarity metric is not always suitable for images that lack strong features or have varying contrast and appearance.

We propose a method to obtain point correspondence between *all* vertices in a target mesh with vertices in a source mesh, starting with an initial set of sparsely labeled corresponding landmarks. A denser set of corresponding landmarks is first determined from the initial sparse set using an interpolation method based on Dijkstra's shortest path algorithm. From these landmarks a non-rigid free-form-deformation of source to target is estimated and applied, before complete vertex correspondence is computed between the two meshes. The FFD estimation is similar to that of Frangi *et al.*[6], however our estimation is based on a similarity metric between the meshes themselves, rather than their corresponding image data. We show that this avoids inaccurate transformation estimation when meshes are generated from labelled CT data where the appearance of the root can vary significantly due to changing contrast, calcifications, or the blood flow around the root itself.

II. METHOD

This section describes how to find a dense set of correspondences across a set of training meshes. A target mesh $M_t = (V_t, E_t, F_t)$ with $|V_t| = n$ vertices is selected from the mesh set, and the remaining meshes are regarded as the source meshes. Assume a source mesh $M_s = (V_s, E_s, F_s)$ where $|V_s| = p$, and $n \neq p$, and m easily identifiable corresponding landmark points labelled on both meshes such that $P_t \subset V_t$ and $P_s \subset V_s$, where $m << n$. The challenge is to find a complete set of n vertex correspondences on the source mesh M_s. Our method obtains a set of source vertices $Q_s \subset V_s$ that are correspondent with V_t.

Finding correspondences between shapes that vary in appearance can be challenging, however the problem becomes much easier when the shapes are similar to one another. By deforming M_s to match M_t, vertex correspondences can be obtained using a nearest-neighbor metric. First, M_s and M_t are transformed to a manually defined *natural coordinate*
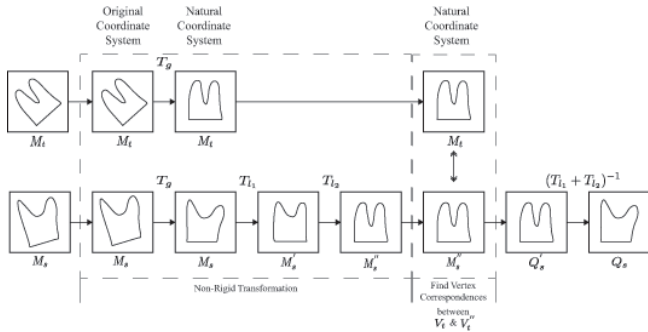
Department of Computer Science, Swansea University, Swansea, UK, http://csvision.swan.ac.uk/. Correspondence: x.xie@swansea.ac.uk.

Fig. 1. The source mesh undergoes a transformation $M_s \rightarrow M_s''$. For *every* vertex V_t in M_t, a corresponding vertex is found in V_s'' to get a re-ordered set of vertices Q_s'. Finally, $(T_{l_1} + T_{l_2})^{-1}$, is applied to get the correspondences on the original source mesh Q_s in *natural coordinates*.

system, so that they are globally aligned. This is followed by a non-rigid transformation estimation between M_s and M_t using P_s and P_t, which is applied to M_s. A nearest neighbor calculation is performed between the vertices of the deformed source mesh and V_t to obtain a new set of correspondences Q_s'. Finally the inverse transform is applied leaving the desired Q_s. This is repeated for all source meshes remaining in the set. An overview is shown in Figure 1.

A. Transformation Estimation

Let $T(x, y, z)$ be the transformation that deforms mesh M_s so that $T : M_s \mapsto M_t$. Applying a global transformation T_g alone accommodates affine differences between the source and target, but does not consider the local differences in shape. To allow a more flexible matching between the pair, a local transformation T_l is also estimated based on B-spline FFD. Our approach estimates T_l in two separate stages; a coarser transformation T_{l_1}, followed by a finer transformation T_{l_2}. For both cases, the FFDs themselves are also estimated using a multi-resolutional procedure T_l^H, where H is the number of mesh resolutions [6], [7]. As a result the full transformation consists of a global, followed by two local transformations.

$$T(x, y, z) = T_g(x, y, z) + [T_{l_1}^{H_1}(x, y, z) + T_{l_2}^{H_2}(x, y, z)]. \quad (1)$$

B. Global Transformation

To initialize the full transformation, M_t and M_s must be globally aligned. This is done by individually transforming each mesh to an external coordinate system, using the initial landmark points P_s and P_t. Alignment to this coordinate system is estimated through an affine transformation: $T_g(x, y, z) = Ax + b$, where A is a matrix containing rotation and scale parameters, and b is a translation vector.

The translation vector is determined by shifting each mesh so that a selected anchor point is at coordinate $[0, 0, 0]$. The x, y, z axes of the natural coordinate system are manually defined, and Euler angles can be determined by finding the angles between the new and original axes. Finally, scaling parameters are estimated by using the sizes of the mesh bounding box. A practical example is given later.

C. Local Transformations

The local transformation terms, T_{l1} and T_{l2}, are estimated in a multi-resolution procedure [6], [7]. Both are expressed as a summation of FFDs at multiple resolutions H:

$$T_l^H(x, y, z) = \sum_{h=1}^{H} T_l^h(x, y, z) \quad (2)$$

At each mesh resolution h, the voxel lattice is warped by moving a set of voxel lattice control points $\phi_{i,j,k}^h$ of size $[n_x \times n_y \times n_z]$, and an FFD is estimated [8]. These control points act as parameters of the B-spline FFD. If δ_0 is the original control point spacing, then at each resolution h, the spacing is defined as $\delta_h = \delta_0 / 2^h$. Decreasing δ_0 decreases the flexibility of the spline, whereas increasing δ_0 allows a more local deformation. This FFD deformation is defined as:

$$T_l^h(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) \phi_{i+l,j+m,k+n}^h \quad (3)$$

where B_l represents the lth basis function of the B-spline [8], and $i = \lfloor x/n_x \rfloor - 1$, $j = \lfloor y/n_y \rfloor - 1$, $k = \lfloor z/n_z \rfloor - 1$, $u = x/n_x - \lfloor x/n_x \rfloor$, $v = y/n_y - \lfloor y/n_y \rfloor$, and $w = z/n_z - \lfloor z/n_z \rfloor$ [6], [7].

The B-spline parameters $\phi_{i,j,k}^h$, are optimized using gradient descent, formulated as $E(\phi) = E_s(target, source) + \lambda E_r(T_l)$, where E_r is a smoothness cost and λ is a regularization term. E_s, is a similarity metric based on a distance measure between sets of mesh vertices.

The transformations are applied sequentially. First T_{l_1} is estimated and applied to M_s to get M_s'. T_{l_2} is then estimated and applied to M_s', to get M_s''. Two factors contribute to making T_{l_1} a coarser transformation than T_{l_2}. The first is the similarity metric calculation. For T_{l_1} this is estimated from the sum-of-squared-distance (SSD) between the sparse landmarks P_s and P_t. The similarity metric for T_{l_2} however is based on the SSD between *all* vertices V_s' and V_t. The other factor is the control point spacing δ_0, during FFD estimation. T_{l_1} has spacing of δ_{0_1}, and T_{l_2} has spacing of δ_{0_2}, where $\delta_{0_1} > \delta_{0_2}$.

D. Finding Complete Correspondence

Finding landmark correspondences between two similar shapes is now a much simpler task, and once the source has been deformed to M_s'', a nearest neighbor algorithm can be used to find complete vertex correspondence. For *every* vertex in V_t, the nearest neighbor based on Euclidean distance is found in V_s''. The nearest neighbor in V_s'' is considered the corresponding vertex, and this leads to a new set of re-ordered deformed source vertices Q_s'. Finally, the inverse local transformation $(T_{l_1} + T_{l_2})^{-1}$ is applied to Q_s' to get the complete source correspondences Q_s in the natural coordinate system.

III. APPLICATION TO AORTIC ROOT MESHES

Here, we demonstrate the proposed method by applying it to computing complete mesh correspondence in aorta geometries obtained from CT transcatheter aortic valve implantation
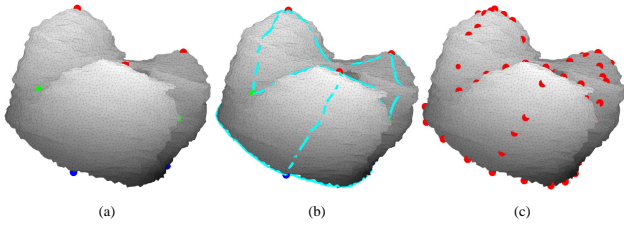
Fig. 2. Initial landmark labelling. (a) Initial sparse landmark points showing the *hinges* (red), *commissures* (green), and *sinotubular junction points* (blue); (b) Interpolation paths between pairs of initial landmarks; (c) New initial landmarks after path interpolation.

Fig. 3. Deformation comparison. (a) column of source meshes (M_s); (b) column of target meshes (M_t); (c) source deformation (M_s'') using proposed method; (d) source deformation (M_s'') based on image data [6], [7].

(TAVI) images. When labelling the aortic root, only a few landmarks can be easily identified. This can often lead to an SSM that inadequately describes the shape. For example when it comes to image segmentation on CT TAVI images where the aortic root boundary is often diffused or close to the blood flow, there is a very high chance that a number of vertices can end up on the wrong boundary. If these vertices are far away from the landmarks used to build the SSM, it is possible that applying the SSM has little or no effect on their positions, potentially leaving a shape that does not resemble the root at all. A denser set of corresponding landmarks is therefore desirable in order to adequately describe the aortic root.

SSMs of the aortic root have previously been implemented using sparsely separated, easily identified landmarks [9], [10]. We propose that obtaining a denser set of correspondences not only describes the shape better, but also makes it possible to exclude or include landmarks in order to build an SSM of varying degrees of freedom. To the best of our knowledge, we are the first to propose finding a complete set of correspondences between a set of aortic root meshes.

A. Finding Initial Landmarks

Each aortic root mesh was labelled with 10 corresponding landmark points. Three of these were aortic valve *hinge points*; the first of which was the nearest to the *aortic arch*, with the remaining two labelled in a clockwise fashion. Three *commissure points* were labelled between the hinges, with the first commissure point between hinges 1 and 2, and the remaining labelled in a clockwise fashion. Three points were labelled on the *sinotubular junction*, directly below the three hinge points. The first of which was below the first hinge point, with the remaining two once again labelled in a clockwise fashion. Finally, a *centre point* was labelled on the surface at the centre of the root.

Using the 10 landmarks, a denser set of corresponding landmarks were found using an interpolation approach to reduce manual work. Pairs of landmark points were defined and the surface paths between them were determined using Dijkstra's shortest path algorithm. Fifteen paths were determined in all; six between the hinge points and their nearest commissure points, three between the hinge points and the centre point, three between the hinge points and their
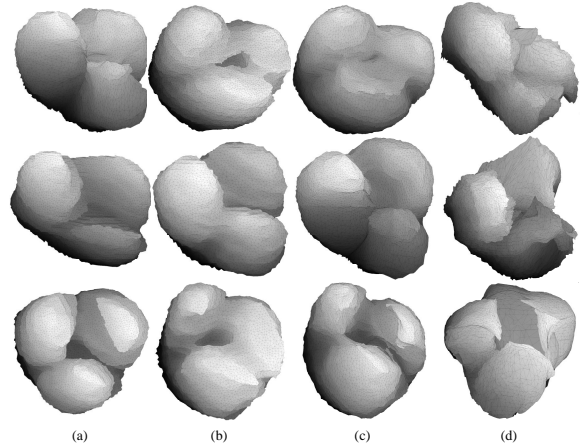
corresponding sinotubular junction points, and three between the sinotubular junction points. 5 evenly spaced points were then interpolated along each path, leaving a total of 65 landmark points. Figure 2 shows the initial root landmarks, the interpolation paths and the new landmark points.

B. Complete Vertex Correspondence

One mesh was selected as the target and the remaining meshes were used as the source meshes. The *natural co-ordinate system* was defined as follows. The anchor point was defined as the labelled *centre point*, and the translation parameters were determined. The z-axis was perpendicular to the plane on which the three sinotubular junction points lay, and the x-axis was parallel to the line between the first sinotubular junction point and the second. Scaling parameters for the source were estimated using the ratio between the local bounding boxes of M_s and M_t.

For our estimation of T_{l_1} we decided to use $H_1 = 3$ mesh resolutions, and an initial control point spacing of $\delta_0 = 15mm$. A relatively large δ_0 was selected here as to suppress the amount of local deformation as the FFD was estimated using the sparse set of corresponding landmarks. This avoided any sharp peaks or troughs in the surface of the mesh, while also reducing the chance of edge overlapping. $H_2 = 3$ was also used for the estimation of T_{l_2}, however here we decided to use $\delta_0 = 5mm$ in order to capture a more local deformation by allowing the FFD more degrees of freedom.

C. Results

An experiment was conducted using 37 aortic root meshes, where each mesh was used as the target in a leave-one-out fashion. The aortic root, including the ascending aorta and aortic arch were labelled in CT TAVI images, of size $512 \times 512 \times (500 - 800)$, and voxel size was $0.48mm \times 0.48mm \times 0.62mm$. In order to ensure that the hinges were clearly seen, multi-planar-rotation software was used
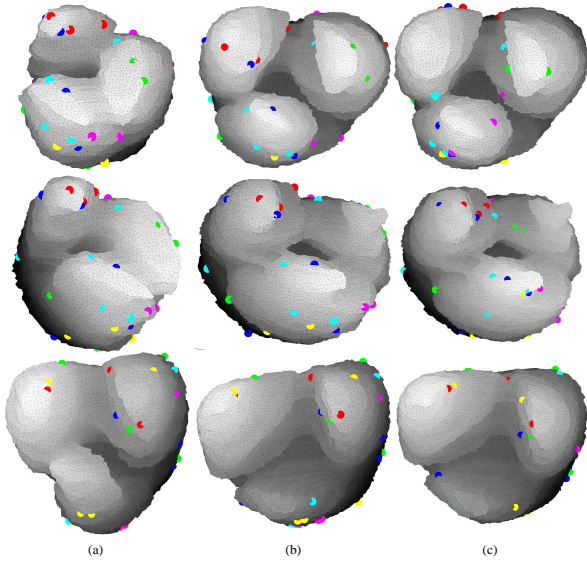
Fig. 4. Vertex correspondence comparison. (a) target mesh; (b) corresponding vertices using proposed method; (c) corresponding vertices using image-based transformation estimation [6], [7].
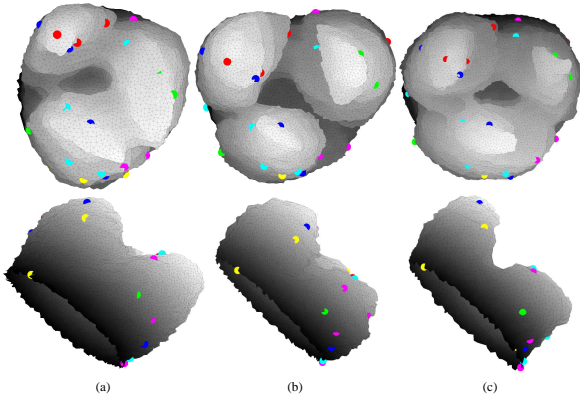

Fig. 5. Further examples of vertex correspondence using proposed method. (a) Target mesh with randomly sampled vertices; (b) Randomly selected source mesh with corresponding vertices; (c) Another randomly selected source mesh with corresponding vertices.

during labelling. This was followed by the marching-cubes algorithm for mesh generation, and the 10 landmarks were manually labelled. The ascending aorta and aortic arch were then discarded from the meshes below the plane on which the 3 sinotubular junction points lay. The average local mesh size was $35mm \times 34mm \times 24mm$.

Registration of source to target was evaluated using the mean symmetrical Hausdorff distance H_{dist}, and mean point-to-mesh distance Epm_{dist} between M_t and M_s''. The results were $H_{dist} = 3.65 \pm 1.19mm$ and $Epm_{dist} = 0.41 \pm 0.25mm$. This represents a mean distance error of less than one voxel, and a deformed mesh with no distant outlying vertices from the surface. Figure 3 compares the deformation of source to target using an image-based metric [6], [7], and our method using the mesh-based approach. This shows that our method deforms the source meshes so that they have a greater

visual resemblance to the target than the deformed meshes using image-based transformation estimation. In addition, the resulting meshes using our method have no compressed or stretched mesh faces, and no tangled mesh edges, leaving a smooth, regularized mesh. This is in contrast to the deformed meshes using an image-based similarity metric, where the meshes appear stretched and compressed in numerous areas.

This close similarity in appearance between M_t and M_s'' allowed more accurate corresponding vertices to be found using the nearest-neighbor calculation. Figure 4 compares the vertex correspondences using our method, and using the image-based transformation estimation.

Additional results of vertex correspondences between target and source meshes using our method are shown qualitatively in Figure 5. Twenty randomly sampled vertices have been plotted on the target, and the corresponding vertices on two source meshes are shown using a colour coded scheme.

IV. CONCLUSION

We have presented a mesh based registration method for finding complete vertex correspondence from a set of sparsely spaced corresponding landmarks. In particular, we have demonstrated this method using complex aortic root meshes, which have corresponding images with varying appearance, and showed the advantages of the proposed method compared against image based approaches, such as [6].

REFERENCES

[1] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[2] R. Yang, M. Mirmehdi, X. Xie, and D. Hall, "Shape and appearance priors for level set-based lv segmentation," *IET Journal of Computer Vision*, vol. 7, no. 3, pp. 170–183, 2013.

[3] V. G. Kim, Y. Lipman, and T. Funkhouser, "Blended intrinsic maps," *SIGGRAPH*, vol. 79, no. 12, 2011.

[4] H. Zhang, A. Sheffer, D. Cohen-Or, Q. Zhou, O. van Kaick, and A. Tagliasacchi, "Deformation-driven shape correspondence," *Eurographics Symposium on Geometry Processing*, vol. 27, no. 5, 2008.

[5] G. Subsol, J. P. Thirion, and N. Ayache, "A scheme for automatically building three-dimensional morphometric anatomical atlases: Application to a skull," *Medical Image Analysis*, vol. 2, no. 1, pp. 37–60, 1998.

[6] A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *Transactions On Medical Imaging*, vol. 21, no. 9, pp. 1151–1166, 2002.

[7] D. Rueckert, L. I. Sonoda, C. Hayes, D. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast mr images," *Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[8] S. Lee, G. Wolberg, and S Y. Shin, "Scattered data interpolation with multilevel b-splines," *Tansactions on Visualization and Computer Graphics*, vol. 3, no. 3, 1997.

[9] R. I. Ionasec, I. Voigt, B. Georgescu, Y. Wang, H. Houle, F. Vega-Higuera, N. Navab, and D. Comaniciu, "Patient-specific modeling and quantification of the aortic and mitral valves from 4d cardiac ct and tee," *Transactions on Medical Imaging*, vol. 29, no. 9, pp. 1636–1651, 2010.

[10] S. Grbic, R. I. Ionasec, Y. Zheng, D. Zauner, B. Georgescu, and D. Comaniciu, "Aortic valve and ascending aortic root modeling from 3d and 3d+t ct," *SPIE Medical Imaging*, 2010.