Non-rigid Registration of 3D Ultrasound Images Using Model-based Segmentation

Babak Matinfar, Lyubomir Zagrochev
Philips Research North America
345 Scarborough Road, Briarcliff Manor, NY 10510, U.S.A.
babak.matinfar@philips.com

Abstract

Non-rigid image registration of large 3D volumes depends heavily on image size. Voxel-based approaches for registration of large data sets are computationally intensive and take a lot of processing time, usually exponentially proportional to their resolution. In this paper, we present a method for non-rigid image registration that is independent of the image size. It can be applied to large data sets without sacrificing performance. The method is based on computing a thin plate spline (TPS) transformation between corresponding control points derived from a rapid and highly accurate model-based segmentation. The segmentation process does not depend on the image size and defines accurate point-based correspondence. Those properties were used to estimate an approximating TPS between a source and a target volumes and to deform the target to align with the source. The method was applied to image sequences of pediatric 3D ultrasound data. Registration of 3D ultrasound images is challenging because of poor-defined image gradients, artifacts, low contrast, noise and speckle. Quantitative and qualitative results indicate that the method is capable of registering large volumes acquired throughout the cardiac cycle from both intra- and inter-patient ultrasound sequences. More importantly, the actual registration does not depend on image intensity and can be applied to multi-modal applications.

1. Introduction

Image registration is a computational method for establishing point-by-point correspondence between two data sets [5]. The registration process typically relies on corresponding image points known as control points or landmarks. These points are used to estimate a transformation function that deforms one image to align with the other. Non-rigid registration uses non-rigid transformation functions such as polynomials, splines, or multiquadrics [6], [2], [3]. These could be formulated as either approximating or interpolating functions that require the solution of a system of equations. The non-rigid registration is usually achieved in an iterative fashion by minimizing a cost or energy function that represents the geometric and/or intensity differences between the two data sets. The performance of the registration algorithm depends on the quality of the corresponding landmarks and the image resolution in terms of accuracy and speed. In voxel-based non-rigid image registration, the cost function is estimated by optimizing an intensity-based similarity metric, such as mutual information or normalized cross correlation. These techniques could be accurate but are time consuming and impractical for many interventional procedures. Furthermore, the optimization process could be trapped in false local maxima, resulting in large registration errors.

Model-based segmentation is a fully-automated method for organ segmentation [11], [1]. It can account for large variations of the segmented shape. It also defines accurate point-based correspondence between segmented volumes. Furthermore, the segmentation speed depends on the resolution of the mesh representing the model rather than the image resolution. Those properties make the approach extremely attractive for registration of large volumes. Having a set of accurate corresponding control-points enables the use of point registration techniques to derive a transformation function which does not depend on intensity-based similarity measures. More importantly, the registration is independent of the image size and the registration speed remains constant with increased image resolution. This is of the key importance for registration of large volume data, where the number of image voxels could be really high and the registration could take a significant amount of time.

2. Method

We propose the use of rapid and accurate shape-constrained segmentation [12] to establish point-based correspondence between two volumes. A shape-constrained deformable model can be adapted to source and target volumes. The geometry of the model consists of a triangular mesh with a fixed topology and a constant number of vertices and triangles. When the model is adapted to different volumes, the triangles of the mesh enlarge or shrink but their number stays the same. They are also attracted to the same part in different images, a property that provides accurate point-based correspondence. A non-rigid transformation function could then be defined in various ways [13]. In this work, we used an approximating thin plate spline to deform the target image to align with the source. The details of each step are provided below.

2.1. Segmentation of the Aortic Root

3D echocardiography is a common procedure for detailed anatomical assessment of cardiac pathology, particularly valvular defects [4]. Real-time 3D (RT3D) echocardiography is capable of acquiring volumetric data at frame rates sufficient to depict cardiac motion. By using transesophageal (TEE) or transthoracic (TTE) multiplane probes, it is possible to obtain rotational images at defined interval angles around a fixed axis. To minimize reconstruction artifacts, sequential images are gated to both electrocardiography (ECG) and respiration. RT3D echocardiography systems use matrix-array transducer technology typically containing several thousand imaging elements, to acquire large volumes of data within a single heart beat. Rapid data acquisition and online reconstruction of 3D data has made 4D and multidimensional echocardiography possible. However, processing of TTE ultrasound images is challenging due to low image quality, limited field of view and apical haze. Registration of 3D ultrasound volumes is necessary in several applications, such as when stitching volumes to expand the field of view or when stabilizing a temporal sequence of volumes to cancel out motion of the probe or anatomy [9].

Shape constrained deformable models (SCDM) with supervised learning have been used for developing heart models from CT [12]. The entire segmentation framework presented in [12], is not tailored for a specific modality and can easily be adapted or extended to other modalities if suitable boundary detection functions can be defined and trained. To enable use of these features for ultrasound imaging modalities that do have uncalibrated gray values, we use a simple histogram-based calibration scheme. In particular, we determine the gray-value histogram of the image and map the gray-value interval between the L and $(100 - L)$-percentile linearly to a standardized interval. This calibration makes the appearance of images acquired with the same protocol from different individuals more similar. We do, however, not transform the entire image, but include this calibration step in the boundary detection functions. Fully automatic segmentation is achieved by a segmentation chain comprising a localization step based on the Generalized Hough Transformation (GHT) and subsequent model adaptation with increasing degrees of freedom. The automatic initialization of the model is performed by application of GHT for localizing anatomical structures in 3D images and the fact that patients are scanned in standardized poses and that the orientation of organs shows little variability. In many cases, it is therefore sufficient to consider translations and possibly scaling. After initial model placement, parametric model adaptation with one global similarity transformation was used to refine pose and size of the aortic root model. Then, using a multi-affine transformation, the locations of suborgans were adjusted to further adapt to the 3D structure of image. Finally SCDM adaptation is performed by iterating boundary detection and mesh deformation (see below).

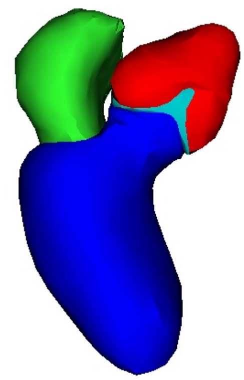

Figure 1. The mesh model used for training of segmentation of 3D ultrasound pediatric data. The model only includes the aorta (red), aortic valve (light blue), left atrium (green) and left ventricle (dark blue).

We followed the same methodology and developed a new aortic root model that could adapt to 3D ultrasound data from pediatric hearts. We created a heart model which only included aorta, aortic valve, left atrium and left ventricle, Figure 1. Each vertex and each triangle in the model is described with an anatomical label and a unique identifier, and therefore can be used as a landmark. For model training, we segmented volumes from pediatric 3D TTE. The model population consisted of 21 different patients who prospectively underwent 3D TTE for reasons other than aortic valve (AoV) or aortic root diseases and were found to have morphologically and functionally normal AoV and aortic root on TTE [1]. Since consistency of data is important for robust segmentation of arbitrary volumes, the model was only

[1]The Institutional Review Board of the Boston Children's Hospital approved the study.

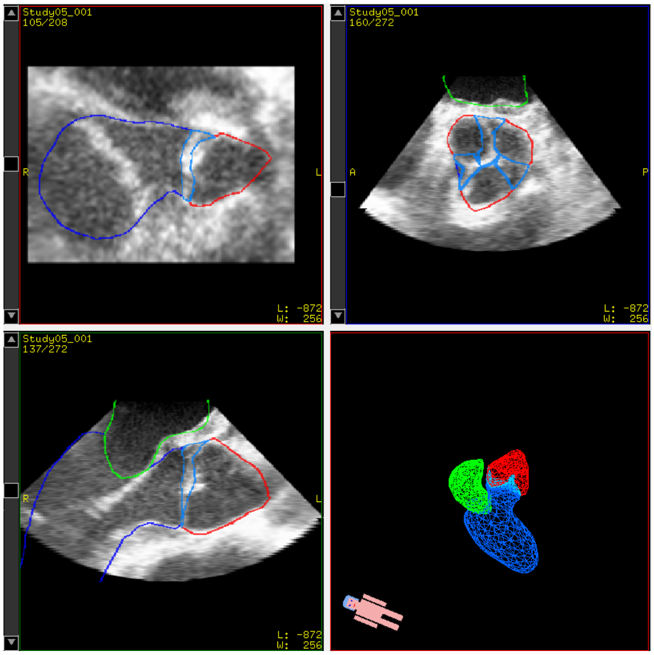

Figure 2. Segmentation of the aortic root model

trained on the normal aortic valves data. Different models could be trained for normal versus abnormal cases for geometrical measurements and comparison.

The model is represented with a mesh that consists of vertices and triangles. Various properties describing the local image appearance and shape variation are extracted for model creation. These properties are learned from a set of pediatric 3D TTE training volumes and their corresponding ground truth segmentation. Each triangle in the model is associated with a feature function. The goal of the feature function is to pull its corresponding triangle closer to the object boundary during adaptation. The feature functions are selected and assigned to each triangle in a supervised feature learning step. A set of possible feature function candidates is evaluated in terms of accuracy with which target points are detected in a simulated search. Object boundaries in general can be detected by projecting the image gradient at location $\nabla I(X_i)$ onto the corresponding triangle normal n_i. This suppresses the effect of edges (high gradients), which deviate from the expected surface orientation. Furthermore, gradients exceeding a certain threshold, g_{max}, can be damped, using:

$$G_{proj} = \pm n_i \bigtriangledown I(X) \frac{g_{max}(g_{max} + \| \bigtriangledown I(X_i) \|)}{g_{max}^2 + \| \bigtriangledown I(X_i) \|^2} \quad (1)$$

Next, the adaptation of the model to a volume is guided by the minimization of an energy term defined as the weighted sum of external and internal energies:

$$E = E_{int} + \alpha E_{ext} \quad (2)$$

The external energy, E_{ext}, moves the mesh towards detected image boundaries while the internal energy, E_{int}, penalizes deviation from a mean geometric (prior) model. The parameter α acts as a weighting term that balances the effect of the two energies. The adaptation is an iterative process that combines a boundary detection step followed by a deformation step.The aortic root model used in this study is a modification of the four chamber heart model containing only the left ventricle, left atrium, aorta, and aortic valve. We found this model is sufficient for aortic leaflet segmentation in general and application to pediatrics in particular. A segmentation of such a model is shown in Figure 2.

Expert tracers manually annotated the set of training images and the data was used to build a mean mesh and to train the boundary feature functions following the simulated search approach [8]. To generate a well behaved patient-specific mesh model that can adapt to the youngest of patients, image enhancement techniques have been applied. In young children, due to their fast heart rates and small anatomical structures, it is difficult to get high resolution imaging. Two ways to improve imaging are to increase the effective frame rate, compensating thusly for their fast heart rate [7], and de-noising the images to improve the signal to noise ratio. This is done through combining a number of overlapping transthoracic images of the aortic valve to de-noise the image [9], [10].

2.2. Thin Plate Spline (TPS)

Given a set of 3D points as defined by

$$\{(x_i, y_i, z_i), (X_i, Y_i, Z_i) : i = 1, ..., N\} \quad (3)$$

the TPS interpolating the points is defined by [13]

$$f(x, y, z) = A_1 + A_2 x + A_3 y + A_4 z + \sum_{i=1}^{N} F_i r_i^2 \ln r_i^2 \quad (4)$$

where $r_i^2 = (x_i - X_i)^2 + (y_i - Y_i)^2 + (z_i - Z_i)^2 + d^2$. This is the equation of a plate of infinite extent deforming under loads centered at $\{(x_i, y_i, z_i) : i = 1, ..., N\}$. The plate deflects under the imposition of loads to take values $\{f_i, i = 1, ..., N\}$. The value d^2 acts like stiffness, and as d^2 approaches zero, the load approaches point loads. As d^2 increases, the loads become more widely distributed producing a smoother surface. Four more constraints are as follows:

$$\sum_{i=1}^{N} F_i = 0 \quad (5)$$

$$\sum_{i=1}^{N} x_i F_i = 0, \quad \sum_{i=1}^{N} y_i F_i = 0, \quad \sum_{i=1}^{N} z_i F_i = 0. \quad (6)$$

The first constraint shows that the sum of the loads applied to the plate should be zero. This is needed to ensure that

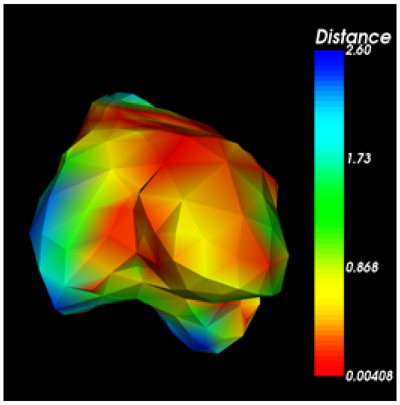

Figure 3. Surface-to-surface distance error between a ground truth mesh (expert tracer) and a mesh obtained from the automated segmentation using a shape-constrained deformable model.

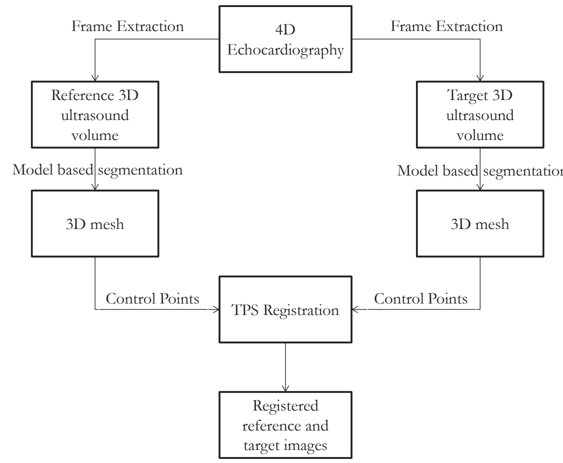

Figure 4. The registration pipeline for 3D ultrasound volumes.

plate does not move under the imposition of the loads and remains stationary. The other constraints require that the moments with respect to each axis be zero, ensuring that the plate would not rotate under the imposition of the loads.

2.3. Registration of 3D Ultrasound Data

3D ultrasound volumes were extracted from 4D echocardiography at different points in the cardiac cycle. Registration of these 3D images was performed using the 3D meshes resulting from the segmentation of the volumes with the shape-constrained deformable aortic root model. The model was adapted automatically to each image and the results were saved as a series of control points. As mentioned earlier, the model-based segmentation preserves mesh correspondence from one image to another.

To evaluate the quantitative accuracy of the segmentation, we used a leave-one-out cross validation. Our training set consisted of 21 volumes. Their corresponding ground truth meshes were provided by expert tracers. The model was trained on 20 volumes, leaving one out. The volume that was left out was then segmented with the trained model and the mesh resulting from the segmentation was compared with the ground truth mesh for that patient. The surface-to-surface distance error for each vertex and the average error were computed. The process was repeated 21 times, leaving a different volume out at each iteration. Figure 3 shows the surface-to-surface error for one of the data sets. The mean surface to surface error for the aortic root was $1.21 \pm 0.6mm$, which is well within the voxel size (3.0 mm) of the ultrasound volume. The accuracy of the segmentation is important because the accuracy of the registration ultimately depends on the accuracy of the segmentation.

Using the thin plate splines described in Section 2.2, we used the segmented meshes to estimate the transformation function. The method was applied in two different ways: first to inter-patient images at different time points in the cardiac cycle, and then to inter-patient images at the same time point in the cardiac cycle. After segmentation, the TPS function was estimated from corresponding meshes and used to deform the target volume. The registration pipeline is depicted in Figure 4.

3. Results

The registration of the same patient at two different time points in the cardiac cycle, the end of diastole (ED) and the end of systole (ES), are shown in Figures 5 and 6, respectively. Each row represents three orthogonal cross-sections from the same ultrasound volume. The third and fourth rows illustrate image overlays of the source and target volumes.

The volumes shown in Figure 5 were sequentially extracted from a DICOM heart beat sequence and downsampled to the same grid size: $56 \times 74 \times 34$ with a $3.0\,mm$ isotropic voxel resolution. Each volume was segmented with the model and the TPS-based registration was performed. The result of the registration is shown in Figure 6. The second row illustrates the deformed target volume. It can be compared to the target volume before registration shown in Figure 5. The overlaid images in Figure 6 (third and fourth rows) show that the described method was able to account for the nonrigid deformation of the heart from ED to ES.

The same procedure was followed for the inter-patient registration. Here the volumes were extracted at the same time point in the cardiac cycle, which was ED. The source and target images are shown in Figure 7 (first and second rows). It can be seen from Figure 7 (third row) that there is a significant misalignment between the two volumes. Like with the intra-patient case, the two volumes were segmented

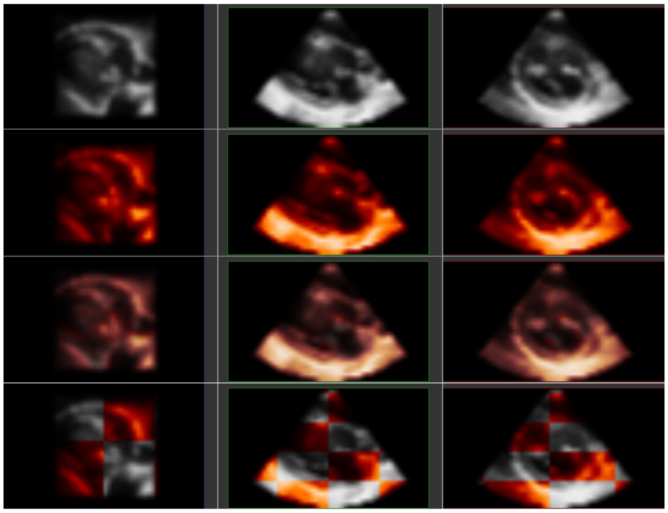

Figure 5. Image overlay of non-registered ED and ES data of the same patient TTE data.

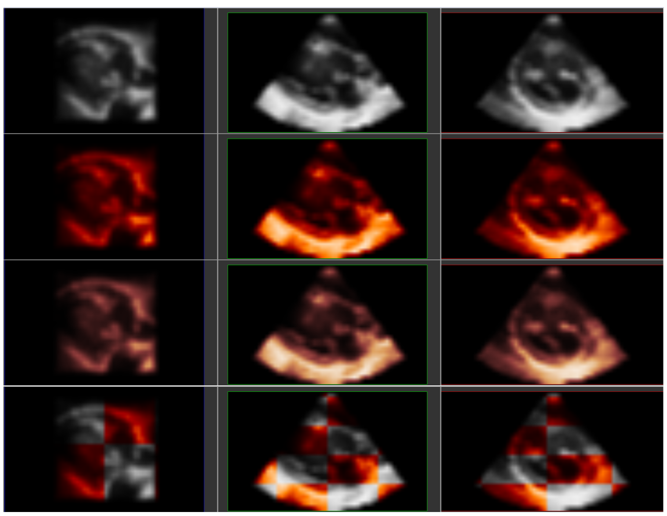

Figure 6. Image overlay of registered ED and ES of the same patient TTE data.

and registered with the thin plate splines. The result is shown in Figure 8, with the target image after registration shown in second row. The overlaid images shown in the third and fourth rows display the performance of registration. To further illustrates the performance of the registration we used Canny edge detector to extract boundaries in the image sets. Figures 9 and 10 show the overlay of image boundaries in the registered and non registered datasets. It is clear that registration performed well in aligning the edges and boundaries in images. Therefore, the registration was accurate and computationally efficient to be applied in interventional procedures.

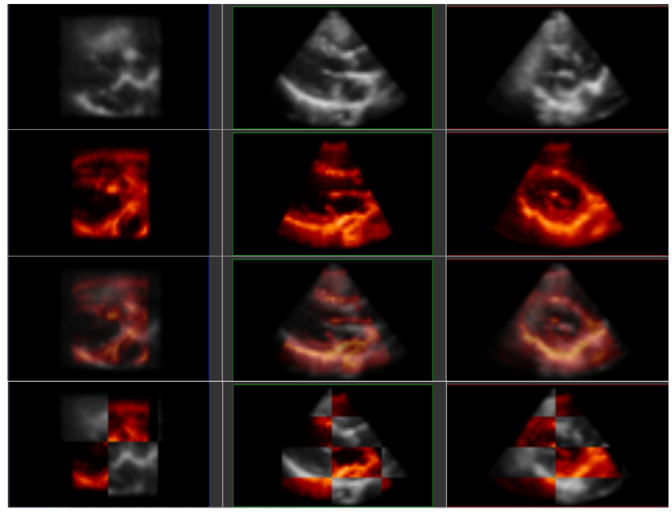

Figure 7. Image overlay of non-registered ED data of different patient TTE data.

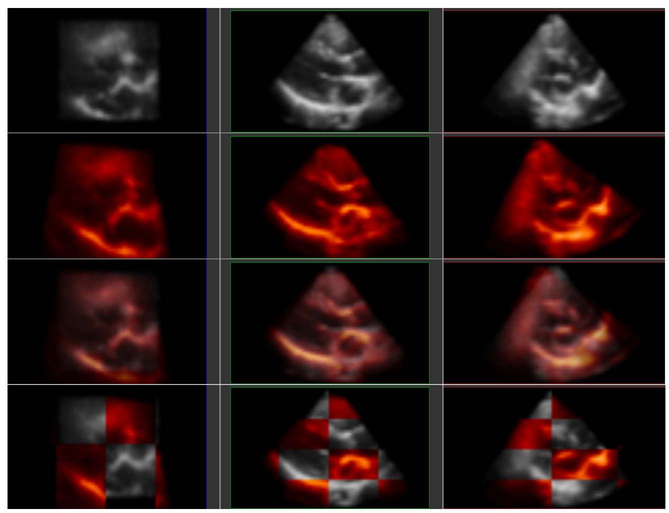

Figure 8. Image overlay of registered ED data of different patient TTE data.

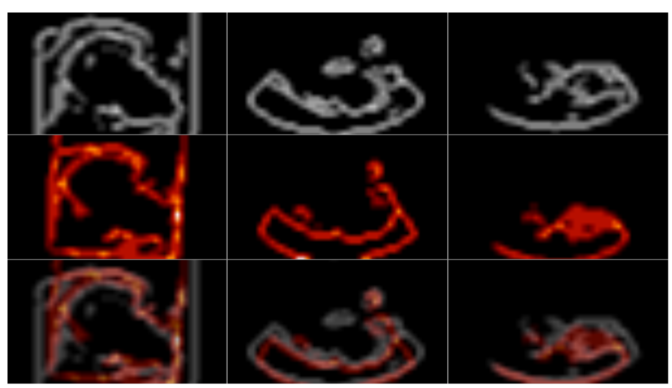

Figure 9. Overlay of image boundaries in non-registered ED and ES of the same patient TTE data.

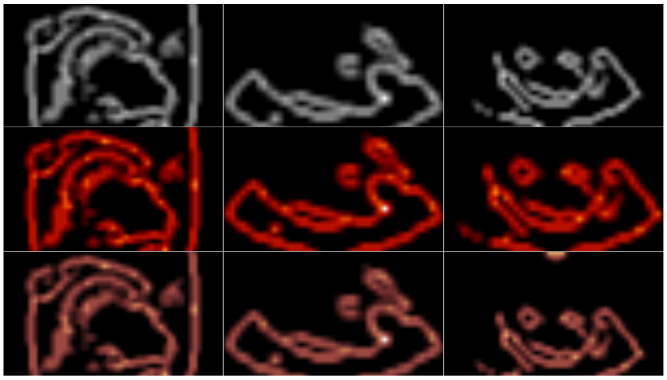

Figure 10. Overlay of image boundaries in registered ED and ES of the same patient TTE data.

4. Conclusion

We presented a method for non-rigid registration of large 3D data volumes that does not depend on the image size and any intensity-based similarity metrics. Voxel-based registration methods for large volumes are computationally intensive and depend on the image size. That renders them impractical for a large number of interventional procedures. Using the properties of the shape-constrained deformable models, we created a mesh segmentation scheme that adapts well to 3D TTE data. The adaptation of the model to different volumes provides accurate point-based correspondence that can be used to estimate a non-rigid thin plate spline transformation. The speed of the registration depends on the mesh resolution and stays constant with increased image size/resolution. More importantly, our method is capable of registering multi-modal data. As long as the 3D mesh used for segmentation has the same topology in different modalities, the control points can be used to calculate the TPS transformation. This is another key feature of our method that makes it suitable for multi-modal clinical applications.

References

[1] O. Ecabert, J. Peter, H. Schramm, C. Lorenz, J. von Berg, M. J. Walker, M. Vembar, M. E. Olszewski, K. Subramanyan, G. Lavi, and J. Weese. Automatic model-based segmentation of the heart in CT images. *IEEE Transaction on Medical Imaging*, 27(9):1189–1201, Sept. 2008. 1

[2] R. L. Harder and R. N. Desmarais. Interpolation using surface splines. *Journal of Aircarft*, 9(2):189–191, 1972. 1

[3] R. L. Hardy. Multiquadric equations of topography and other irregular surfaces. *Journal of Geophysics Research*, 86(7):1905–1915, 1971. 1

[4] J. Hung and *et al*. 3D echocardiography: A review of the current status and future directions. *Journal of American Society of Echocardiography*, 20:213–233, 2007. 2

[5] H. Lester and S. R. Arridge. A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, 32(1):129–149, January 1999. 1

[6] J. Meinguet. An intrinsic approach to multivariate spline interpolation at arbitrary points. In *Polynomial and Spline Approximation*, pages 163–190. Reidel, 1979. 1

[7] D. P. Perrin, N. V. Vasilyev, G. R. Marx, and P. J. del Nido. Temporal enhancement of 3d echocardiography by frame reordering. *Journal of the American College of Cardiology*, 5(3):300–304, 2012. 3

[8] J. Peters, O. Ecabert, C. Meyer, R. Kenser, and J. Weese. Optimizaing boundary detection via simulated search with applications to multi-modal heart segementation. *Medical Image Analyses*, 14(1):70–84, 2010. 3

[9] R. Schneider, D. P. Perrin, N. V. Vasilyev, G. R. Marx, P. J. D. Nido, and R. D. Howe. Real-time image-based rigid registration of three-dimensional ultrasound. *Journal of Medical Image Analysis*, 16(2):404–414, Feb 2012. 2, 3

[10] K. Thiele, J. Jago, R. Entrekin, and R. Peterson. Exploring nsight imaging a totally new architecture for premium ultrasound. *www.philips.com/epiq*, 2013. 3

[11] J. Weese, M. Kaus, C. Lorenz, S. Lobregt, R. Truyen, and V. Pekar. Shape constrained deformable models for 3D medical image segmentation. *Information Processing in Medical Imaging, Lecture Notes in Computer Science*, 2082:380–387, 2001. 1

[12] J. Weese, I. Wächter-Stehle, L. Zagorchev, and J. Peters. Shape-constrained deformable models and applications in medical imaging. In *Shape Analysis in Medical Image Analysis*, pages 151–184. Springer, 2014. 1, 2

[13] L. Zagorchev and A. Goshtasby. A comparative study of transformation functions for nonrigid image registration. *IEEE Transactions in Image Processing*, 15(3):99–108, March 2006. 2, 3