Real-time image-based rigid registration of three-dimensional ultrasound

Robert J. Schneider[a,*], Douglas P. Perrin[a,b], Nikolay V. Vasilyev[b], Gerald R. Marx[c], Pedro J. del Nido[b], Robert D. Howe[a]

[a] *Harvard School of Engineering and Applied Sciences, Cambridge, MA, USA*
[b] *Department of Cardiac Surgery, Children's Hospital, Boston, MA, USA*
[c] *Department of Cardiology, Children's Hospital, Boston, MA, USA*

Abstract

Registration of three-dimensional ultrasound (3DUS) 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. Current systems that register 3DUS volumes either use external tracking systems (electromagnetic or optical), which add expense and impose limitations on acquisitions, or are image-based methods that operate offline and are incapable of providing immediate feedback to clinicians. This paper presents a real-time image-based algorithm for rigid registration of 3DUS volumes designed for acquisitions in which small probe displacements occur between frames. Described is a method for feature detection and descriptor formation that takes into account the characteristics of 3DUS imaging. Volumes are registered by determining a correspondence between these features. A global set of features is maintained and integrated into the registration, which limits the accumulation of registration error. The system operates in real-time (i.e. volumes are registered as fast or faster than they are acquired) by using an accelerated framework on a graphics processing unit. The algorithm's parameter selection and performance is analyzed and validated in studies which use both water tank and clinical images. The resulting registration accuracy is comparable to similar feature-based registration methods, but in contrast to these methods, can register 3DUS volumes in real-time.

Keywords: ultrasound, registration, real-time, image-based

1. Introduction

Three-dimensional ultrasound (3DUS) has many benefits in that it is inexpensive, portable, non-ionizing, and capable of accurately imaging fast moving structures such as heart valves. However, there are several applications in which the information contained in a single 3DUS volume is insufficient and multiple registered 3DUS volumes are necessary. Applications include creating large field-of-view (FOV) mosaics from smaller FOV 3DUS volumes (Poon and Rohling, 2006), stabilizing a temporal sequence of 3DUS volumes to account for motion of the probe or anatomy (Shekhar et al., 2004), and/or compositing images of the same anatomy acquired at various probe locations to improve image quality (Rajpoot et al., 2011). In these applications, it is clinically desirable to register volumes in real-time (i.e. as fast or faster than the volumes are acquired) so as to provide immediate feedback to clinicians.

Depending on the application, 3DUS volumes are registered in either a rigid or non-rigid fashion. Rigid registration is generally used when little or no amount of non-rigid deformation of anatomy is expected between frames, and can be modeled as a combination of rotation and translation. An example of this scenario would be when registering 3DUS volumes of the heart taken at the same time during the cardiac cycle (Rajpoot et al., 2011). Non-rigid registration is used when a considerable amount of non-rigid deformation of anatomy is expected between frames. In these instances, aligning corresponding objects from different images cannot be handled with a rigid registration model, but rather a spatially varying deformation field needs to be determined.

Various approaches have been used in non-rigid registration of ultrasound to resolve the spatially varying deformation field. Typically scale space or sub-volume approaches are used for robustness and to improve computational efficiency (Krucker et al., 2002; Xiao et al., 2002; Pratikakis et al., 2003; Zikic et al., 2006; Ledesma-Carbayo et al., 2006). Some methods have used tracking or matching of features between images, where a dense deformation field is then found from interpolation or fitting a B-spline approximation to the feature displacements (Foroughi et al., 2006; Moradi et al., 2006). It is worth noting that most of these methods, while designed for non-rigid registration, could be adapted to find a rigid registration between volumes by enforcing a globally consistent deformation field.

In this manuscript, we address the issue of real-time registration of 3DUS volumes. As real-time 3DUS volumes are acquired very quickly (typically at or above 30Hz) and

*Corresponding author (rjschn{at}seas.harvard.edu)

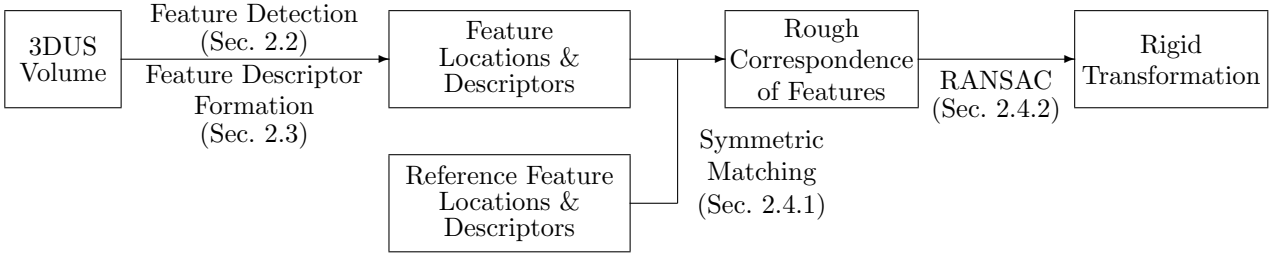


Figure 1: Summary of the feature-based 3DUS rigid registration algorithm. The specific processes can be found in the indicated sections.

typically over only a small region of interest, we assume the amount of non-rigid deformation between images is small, and therefore assume a rigid transformation model between volumes. Rigid registration of 3DUS volumes is typically done using either a tracking-based or imaged-based approach (or both). Tracking-based methods operate by tracking the transducer with an external tracking system, usually optical or electromagnetic, and computing transformations between images based on the position and orientation of the tracker (Poon and Rohling, 2006; Yao et al., 2009; Zhuang et al., 2010). These systems, along with the added expense of additional hardware, require careful calibration of the tracker to the ultrasound image (Mercier et al., 2005). Some of these systems also make use of image-based registration methods to refine the registration and account for other movements that cannot be tracked by the probe (Xiao et al., 2002; Gee et al., 2003; Poon and Rohling, 2006; Yao et al., 2009; Zhuang et al., 2010). For systems that do not use image-based refinement, to account for the displacement of anatomy due to respiration, either patients are put on breath-hold, images are respiration-gated (Makela et al., 2002), or respiration is tracked and accounted for using an additional tracker on the chest or abdomen (Wein et al., 2008). Also, careful consideration of the tracking environment needs to be considered, as electromagnetic tracking is sensitive to metal in the field and optical trackers need to maintain a clear line of sight between the sensor and markers.

Image-based 3DUS registration methods can be divided into two types: voxel-based and feature-based. Voxel-based methods compute a metric over all voxels in a volume (or overlapping voxels from volumes to be registered) and in an iterative fashion find the parameters of the transformation between volumes (Rohling et al., 1998; Shekhar et al., 2003; Francois et al., 2003; Cen et al., 2004; Neemuchwala et al., 2005; Grau et al., 2007; Wachinger et al., 2008; Rajpoot et al., 2009; Kutter et al., 2009; Wein et al., 2009). These methods have proven to be accurate, even in the presence of large translational and rotational displacements between images, but require that voxels be revisited several times. Fast implementations of this type (Kutter et al., 2009; Wein et al., 2009) used accelerated frameworks on a graphics processing unit (GPU) and simultaneously registered 3DUS volumes to each other and to CT images. The former registered larger volumes in 1-3 seconds, depending on volume size, whereas the latter registered 10 smaller intracardiac 3DUS volumes in 0.6 seconds by assuming a linear trajectory of the transducer and modifying only the first and last transformations.

The second image-based registration type consists of feature-based methods that compute a transformation between images by determining a correspondence between feature sets (i.e. segmented volume, edges, salient points, etc.) extracted from 3DUS volumes (Porter, 2004; Soler et al., 2005; Moradi et al., 2006; Wang et al., 2007; Ni et al., 2008). These methods can usually only handle small translational and rotational displacements between 3DUS volumes. The fastest of these (*not* GPU accelerated) was shown in Ni et al. (2008), which used a 3D SIFT implementation to register volumes in roughly one minute.

A limitation of current rigid 3DUS registration methods is that, because they typically take several seconds or minutes to register two volumes, they cannot operate in real-time. Using these methods, it therefore cannot be determined if sufficient or insufficient data is being acquired for a given application until well after the images are acquired. For instance, in the application of creating a large FOV mosaic, the extent of the mosaic cannot be immediately assessed, and so it is not known when sufficient coverage of the anatomy of interest has been obtained.

To address these issues, we present an image-based 3DUS rigid registration method capable of operating in real-time when using a GPU accelerated framework, thereby making the registered 3DUS images immediately available to clinicians. The method is a feature-based method, summarized in Figure 1. The presented method is designed for real-time ultrasound acquisitions where it is assumed that small probe displacements occur between images in the sequence. This does not mean that large translations or rotations cannot occur over the course of the acquisition, but merely that from frame to frame these displacements are small. As probe movement in most acquisitions is already limited, such as on the surface of the liver or within the esophagus to view the heart, these restrictions should not limit the registration algorithm's usefulness. The method, however, is not designed to perform the registration between volumes that exhibit large translational or rotational displacements, such as used in several compositing studies (Grau and Noble, 2005; Yao et al., 2009; Rajpoot et al., 2011). We also do not claim any novel contributions in the area of ultrasound compositing or improving image quality, but when using the presented

registration method for mosaicing multiple 3DUS volumes, a mean compositing method is used as it is simple, fast, and produces favorable results as found in visual inspection.

This manuscript is organized such that the details of the algorithm are outlined in Section 2. The accuracy, parameter selection, and execution times for the algorithm, using both clinical and water tank images, are then presented in Section 3.

2. Materials and Methods

2.1. Data Assumptions & Pre-Processing Strategies

This method is designed for real-time ultrasound acquisitions where data is streamed from the ultrasound machine to an external CPU. Using both the CPU and an accompanying GPU, transformations and interpolations are computed. To expedite the processing of the 3DUS volumes, we exploit several characteristics of ultrasound sequence data. First, we know the number of intermediate volumes (for instance, the Laplacian-of-Gaussian of a volume) needed to process each volume before it can be registered. Therefore, to avoid the dynamic allocation/deallocation of memory, all necessary volumes are initialized and reused as necessary. Second, there are several intermediates – for instance, eroded data masks that are used to eliminate edge effects at the boundary of the conical-shaped 3DUS image and the surrounding zero-padded region needed to fill out the rectangular volume – that are the same for every volume in a sequence. We therefore precompute and store these intermediates so that they can be quickly queried at a later time. Lastly, there are some computations that need to be performed at each voxel, however, in the zero-padded region surrounding the conical-shaped 3DUS image, these operations are meaningless. Therefore, the number and locations of the conical data voxels are stored so that efficient thread launches on the GPU can be executed to perform these computations.

2.2. Feature Detection

Several methods can be used to find feature points in an image (Mikolajczyk, 2002), but the most important characteristic of the feature detection is that the feature points are stable (i.e. features are found at the same salient region) from image to image. In the 2D SIFT implementation for photos (Lowe, 2004) and 3D SIFT implementation for 3DUS (Ni et al., 2008), these features are found as the local extrema of a Difference-of-Gaussian (DoG) scale-space. While finding features across multiple scales is important for photos, as the scale of objects can change according to the orientation of the camera relative to an object, this is unnecessary for real-time 3DUS volumes that exhibit small displacements, as the scale of objects should not change.

We therefore simplify feature detection by searching for features at a single scale. As the DoG has been shown to be adequate for feature detection in 3DUS (Ni et al., 2008), and as the DoG is just an approximation to the scale-normalized Laplacian-of-Gaussian (LoG) (Lowe, 2004), we search for features by finding the local minima of the LoG of a 3DUS volume at a single user-specified scale, σ_f. We use an isotropic LoG kernel, as the resolution of most 3DUS volumes is nearly isotropic and the kernel size is typically small compared to the volume size. If this were not the case, however, an anisotropic kernel would be more appropriate. We search only for local minima of the LoG, as opposed to both local minima and maxima, because we want to avoid finding features in the middle of blood pools or shadows that appear in a 3DUS image, and the minima are found at higher intensity locations that typically correspond to tissue locations. To further enforce that features reside at tissue locations, we require that the 3DUS intensity at a feature location is above a user-specified tissue threshold, τ_{tissue}. An important trait of this feature detection method is that, while this does not find all features at all spatial scales, because features do not change scale in 3DUS, the same features at a constant scale are being found from image to image. Assuming an appropriately chosen value for the scale, σ_f, a sufficient number of features (on the order of several hundred features) should be found such that an accurate registration between images can be computed.

The intent of having the user define σ_f and τ_{tissue} is to allow the user some control over the number of features that are found in the images and the resulting registration time. However, we show later how these parameters can be automatically tuned to make the registration method fully automatic. It stands to reason that for increasingly smaller feature scales or tissue thresholds, more features will be found. Typical values for σ_f and τ_{tissue} used in studies were 1-2mm and 150-200, respectively, where 3DUS intensities were in the range of 0 to 255.

2.3. Feature Descriptors

A feature descriptor is a means of uniquely characterizing a feature location and allows for the calculation of a correspondence between features from different images. Previous efforts have been made to make these descriptors both scale and rotationally invariant in both 2D and 3D imaging for the purpose of matching features under large rotations and changes in scale (Lowe, 2004; Ni et al., 2008). However, when registering volumes in a real-time 3DUS sequence, the rotations from frame to frame are small. We therefore construct feature descriptors using a simple rotationally variant method.

We construct feature descriptors by taking a sparse sampling of the 3DUS volume at and around a feature location (Figure 2). The sampling is taken on a rectilinear $5 \times 5 \times 5$ grid of 125 points, where the grid is oriented along the image axes and centered at each feature location. The sample spacing, δ_s, is equal to $M_d \sigma_f$, where the scale factor M_d is a constant and ensures that the extent of the feature is represented. The samples are then organized as

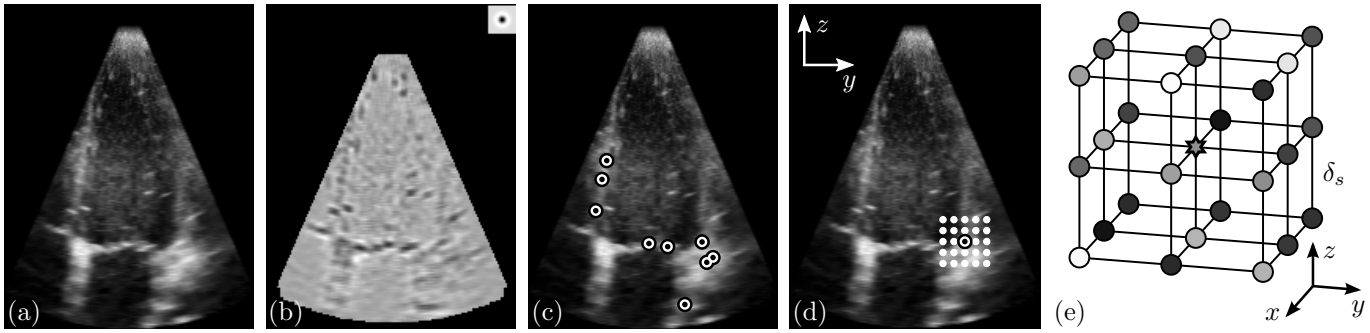


Figure 2: The feature detection and feature descriptor formation process. (a) Slice from a clinical 3DUS image of the heart. (b) Corresponding slice from the Laplacian-of-Gaussian (LoG) computed for the original volume shown in (a) ($\sigma_f = 2.0$mm). The image in the top right corner of (b) shows the center slice of the LoG kernel (to scale) corresponding to σ_f. (c) Same slice as (a) with detected feature locations found as the local minima of the volumetric LoG and with intensity greater than $\tau_{tissue} = 150$. (d) 2D depiction of sample points about a feature to construct the feature descriptor. (e) Example of points sampled in 3D relative to the feature point (starred) to construct the feature descriptor. The shaded spheres represent the sampled intensities of the 3DUS volume. The grid axis is aligned with the example image axis shown in (d). Note that while a 3×3×3 grid is shown, the actual grid used to compute feature descriptors is a 5×5×5 grid with a sample offset of δ_s.

feature vectors of dimension 125 normalized to unit norm. We show in several validation studies that these descriptors provide a sufficient level of performance with regard to registration accuracy, and additionally take little time to compute.

2.4. 3DUS Volume Registration

Volume registration is performed using the feature sets from the volumes to be registered. In this study, we refer to the reference (stationary) volume as $\mathbf{V}_{ref}$ and its features as $\mathbf{F}_{ref} = (\mathbf{X}_{ref}, \mathbf{D}_{ref})$, where $\mathbf{X}_{ref}$ and $\mathbf{D}_{ref}$ are the positions and feature descriptor vectors, respectively, of the reference feature set. The number of features in the set is n_{ref}, and the position and feature vector for feature i in the set is $\mathbf{X}_{i,ref}$ and $\mathbf{D}_{i,ref}$, respectively. Similarly, we refer to the volume to be registered as $\mathbf{V}_{new}$ and its features as $\mathbf{F}_{new} = (\mathbf{X}_{new}, \mathbf{D}_{new})$, where there are n_{new} features in the set. The source of $\mathbf{F}_{ref}$ varies depending on the registration strategy, as discussed in Section 2.7.

We compute a transformation between volumes by first finding a rough correspondence between the features in $\mathbf{F}_{ref}$ and those in $\mathbf{F}_{new}$ using the symmetric matching algorithm described in Section 2.4.1. This rough correspondence will contain true matches (inliers) and false matches (outliers), and so we subsequently employ the model fitting algorithm (RANSAC) described in Section 2.4.2 to remove outliers and to better estimate the rigid transformation between volumes.

2.4.1. Symmetric Matching

Symmetric matches between feature sets $\mathbf{F}_{ref}$ and $\mathbf{F}_{new}$ are found by first computing the pairwise distances between descriptor vectors $\mathbf{D}_{ref}$ and $\mathbf{D}_{new}$, where distances are computed as the Euclidean norm of the vector differences. If we index vectors $\mathbf{D}_{ref}$ with α, where $\alpha \in \{1, ..., n_{ref}\}$, and index vectors $\mathbf{D}_{new}$ with β, where $\beta \in \{1, ..., n_{new}\}$, then a symmetric match is said to occur when $A = \underset{\alpha}{\operatorname{argmin}} \|\mathbf{D}_{\alpha,ref} - \mathbf{D}_{B,new}\|$ and $B = \underset{\beta}{\operatorname{argmin}} \|\mathbf{D}_{A,ref} - \mathbf{D}_{\beta,new}\|$. The $\mathbf{X}_{ref}$ and $\mathbf{X}_{new}$ corresponding to the symmetric matches are stored as $\mathbf{M}_{ref}$ and $\mathbf{M}_{new}$, where the coordinate positions of the features from symmetric match i are $\mathbf{M}_{i,ref}$ and $\mathbf{M}_{i,new}$ and $i \in \{1, ..., n_{sym}\}$.

2.4.2. RANSAC

While symmetric matching generates a rough correspondence between features in two feature sets, outliers (false matches) prevent computing an accurate registration strictly from the symmetric matches. For this reason, we use a RANSAC (RANdom SAmple and Consensus) algorithm (Fischler and Bolles, 1981; Hartley and Zisserman, 2003) to remove outliers so that a more accurate registration can be computed. In each RANSAC trial, three unique symmetric matches are used with the least-squares registration algorithm in Arun et al. (1987) to estimate the transformation, $\mathbf{T}_t$, that maps $\mathbf{M}_{new}$ to $\mathbf{M}_{ref}$, where $t \in \{1, ..., n_{trials}\}$ and $n_{trials} = 10n_{sym}$. The coordinates $\mathbf{M}_{new}$ are then transformed according to $\mathbf{T}_t$, and a support for the trial, S_t, found as the number of matches where $\|\mathbf{M}_{i,ref} - \mathbf{T}_t\mathbf{M}_{i,new}\| \leq d_{ransac}$, where d_{ransac} is a distance threshold that determines the cut-off for when a transformed symmetric match is considered an inlier versus an outlier. A study to determine an appropriate value for d_{ransac} is described in Section 3.1.2. The symmetric matches that make up the support from the trial with the largest S_t are then used again with the least-squares registration algorithm in Arun et al. (1987) to estimate the final transformation, $\mathbf{T}_{final}$, that transforms the new volume to the coordinate space of the reference volume.

2.5. Interpolation

Once the transformation, $\mathbf{T}_{final}$, relating the position of $\mathbf{V}_{new}$ to $\mathbf{V}_{ref}$ is determined, we reconstruct $\mathbf{V}_{new}$ in the coordinate system of $\mathbf{V}_{ref}$. The interpolated image, $\mathbf{V}_{interp}$, is computed using a tri-linear interpolation

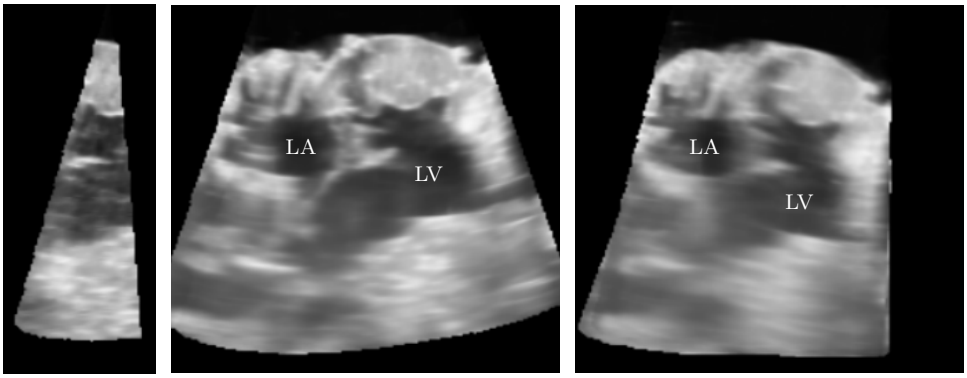


Figure 3: Mosaic of a porcine heart in a water tank created from 437 3DUS volumes. (*Left*) Example of a single 3DUS volume used in the mosaic. (*Center*) Cross-sections of the mosaic created using the RTG method and (*Right*) RTP method. Notice how the RTP mosaic lacks definition and is a different shape compared to the RTG mosaic, indicating significant registration error. (LA – left atrium; LV – left ventricle)

method. Using an eroded mask of the conical data volume, we avoid interpolating the data near the intersection (within a two voxel radius) of the conical 3DUS image and the surrounding padding of zeros, thereby avoiding the blurring effect that would otherwise be seen in the interpolated image. The interpolated volume, $\mathbf{V}_{interp}$, is made to be the same volume size as the original volume, $\mathbf{V}_{new}$, but as the conical 3DUS image does not occupy the entire rectangular volume, we use a windowing operation to limit the region within $\mathbf{V}_{interp}$ for which the interpolation is computed to the subvolume at which the transformed conical image resides.

2.6. Mosaicing

One application of the presented registration method is constructing large FOV mosaics from several smaller FOV 3DUS volumes. For this application, we composite the images using a simple averaging method, as this is simple, produces favorable results as seen through visual inspection, and can be computed quickly on a GPU.

To composite using the averaging method, we maintain two volumes, $\mathbf{\Sigma}_{data}$ and $\mathbf{\Sigma}_{count}$, where $\mathbf{\Sigma}_{data}$ is the summation of interpolated intensities from each $\mathbf{V}_{interp}$ volume added to the mosaic, and $\mathbf{\Sigma}_{count}$ the summation of the number of voxels that have contributed to the data in $\mathbf{\Sigma}_{data}$. Voxels with a value of zero (in the zero-padded region) do not factor into the compounding, and therefore do not affect the mosaic. The compounded mosaic image is then found as $\mathbf{V}_{mosaic} = \mathbf{\Sigma}_{data}/\mathbf{\Sigma}_{count}$. A similar windowing operation as was used for the interpolation procedure is used for updating the mosaic, making the process of updating the mosaic quick and efficient.

2.7. Registration Strategies

When registering volumes in a 3DUS sequence, there are two registration strategies that could be employed. One would be to register the next volume to the most recently registered volume, a method which will be referred to as the Register-to-Previous (RTP) method. While this method would be suitable for short 3DUS sequences, it would accumulate substantial error in long sequences with many volumes (Figure 3). An alternate strategy in registering new volumes is instead use a combination of the features from all volumes in the sequence previous to the current volume, a method which will be referred to as the Register-to-Global (RTG) method. The work presented in Wachinger et al. (2008) shows the benefits of this group-wise versus pair-wise registration strategy for 3DUS.

A naive approach for the RTG method might be to simply retain the features from every volume and store their positions and descriptors in a large database for use in registering subsequent volumes. However, this approach is inefficient and can lead to the storage of repeated and useless information. Rather, we record only the "good" features (i.e. features that are likely to be found in multiple volumes) and maintain a manageable database size by limiting repeated features, therefore allowing for real-time execution of the registration algorithm. To accomplish this, we start by only recording those features that make up the support from a previous 3DUS volume registration. As these corresponding features were found in at least two previous volumes, it is more likely they will be found in subsequent volumes. Also, because features are required to be local minima of the LoG, it makes sense to allow only one feature to occupy a small neighborhood. We keep track of the feature positions for features added to the database by maintaining a volume, $\mathbf{V}_{DB}$, that is the same size as $\mathbf{V}_{interp}$ (or $\mathbf{V}_{mosaic}$ when mosaicing) and that indicates the positions of the previously recorded features. A new feature is added to the database only if a feature has not already been recorded at its location. When a new, unique feature is found, its index location in the database is then written to the neighborhood of diameter σ_f centered at the feature position in $\mathbf{V}_{DB}$. Additionally, the feature position and descriptor vector are recorded to the database.

When registering new volumes, it is unnecessary to use the entire database of features to compare to the features of the new volume, as there will be a large number of features whose position will be nowhere near the volume to register. Rather, the feature set that we use as the reference feature set, $\mathbf{F}_{ref}$, when registering a new volume is a

combination of a small subset of the feature database and the features from the previously registered volume. The features that are used from the database are those that resided at the location of the conical data volume in the transformed previous volume, $\mathbf{V}_{interp}$. This is based on the assumption that there are small displacements between consecutive frames. These features can be quickly found by looking at the intersection of $\mathbf{V}_{interp}$ and $\mathbf{V}_{DB}$. Additionally, all features from the previous volume are used because neighboring volumes will have a large number of corresponding features, and these features do not necessarily already exist in the database.

2.8. Automatically Tuning Parameters

In the registration algorithm, there are four parameters that need to be defined that control how the algorithm operates and performs. These parameters are the feature descriptor sample offset scale factor (M_d), RANSAC distance threshold (d_{ransac}), feature scale (σ_f), and tissue threshold (τ_{tissue}). The sample offset scale factor determines, relative to the feature scale, the extent of the sampling grid used in the feature descriptor formation process. The feature scale is assumed to be tuned for each application depending on the image and anatomy of interest, and so it follows that this scale factor should be a constant. An optimal range and value for the scale factor is determined in the study shown in Section 3.1.1. Similarly, the RANSAC distance threshold should also not need to be tuned for each application, as it should be made as small as possible so that accurate registrations result, but not so small as to be on the order of the original 3DUS resolution. We show this to be the case in the study described in Section 3.1.2.

The feature scale and tissue threshold, however, are dependent on the imaging and anatomy being imaged, as the scale of the dominant features and the 3DUS intensity at prospective feature locations will change for different tissue and anatomic structures. These values therefore need to be tuned for each application. The effect of varying these values is studied and described in Section 3.1.3 and Section 3.1.4.

To facilitate the tuning of the feature scale and tissue threshold (i.e. to keep the user from having to tune these parameters for every application or 3DUS sequence) and make the registration algorithm fully automatic, we designed an auto-tune method for these parameters. The method starts by assuming values for these parameters, and in iteratively registering two example 3DUS volumes, the parameters are modified such that either a desired number of matching features are found or a desired registration time results.

As later shown and described in Section 3.1.3 and Section 3.1.4, suitable starting values for the feature scale and tissue threshold are 1.0mm and 170, respectively. Also shown is that, in general, registration time and support (i.e. final number of matching features) are inversely proportional to these parameters. Therefore, at each iteration, a greedy algorithm can be employed and depending on whether the registration time or support is too large or small, σ_f and τ_{tissue} are increased or decreased, respectively, in step sizes of 0.1mm for σ_f and 10 for τ_{tissue}. The parameter modification that results in the largest change to the registration time or support is adopted (i.e. only one parameter value is changed in each iteration), and the process iterated until the time or support falls within a desired range. This auto-tune method typically takes less than a few seconds to perform.

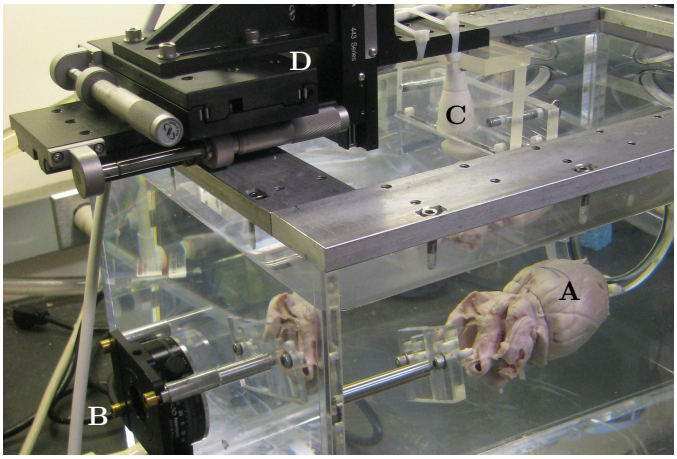


Figure 4: Water tank setup that allows for the acquisition of an ultrasound image with a known ultrasound probe position/orientation relative to the object being imaged. Shown are a porcine heart (A) attached to a rotation stage (B), and an ultrasound probe (C) attached to a translation stage (D).

3. Validation and Performance

We performed several studies to characterize the behavior of the algorithm. These studies either explored the parameter space of the algorithm to determine suitable ranges for parameter values or explored the algorithm performance for different applications. Data was acquired for these studies using a 3DUS machine (iE33 Echocardiography System with transesophageal X7-2t and transthoracic X7-2 probes, Philips Healthcare, Andover, MA, USA) with the capability to stream 3DUS images to an external PC (Dell Alienware Aurora, Intel Core i7 processor @ 2.67GHz, 6GB RAM, NVIDIA GTX260 graphics card). Several studies were done using images that were acquired at known positions (i.e. known translation and rotation offsets). This used the water tank setup shown in Figure 4, which features a translation stage (to which we attached the X7-2 ultrasound transducer) and a rotation stage (to which we attached a porcine heart for imaging). We also explore the accuracy of the registration algorithm when used for stabilizing clinical ultrasound sequences of valves in a beating heart.

3.1. Parameter Selection

As previously described, there are four parameters (σ_f, τ_{tissue}, M_d, d_{ransac}) that control algorithm operation and

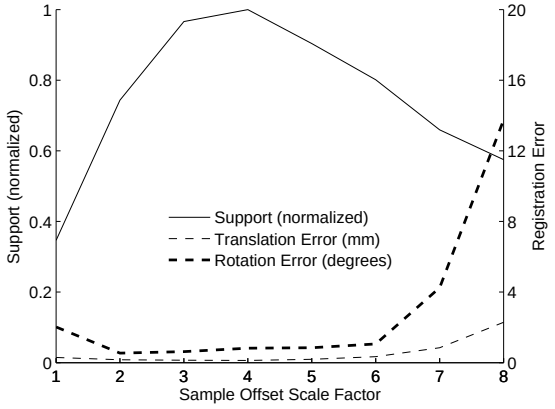


Figure 5: Effect of the feature descriptor sample offset scale factor (M_d) on registration performance.

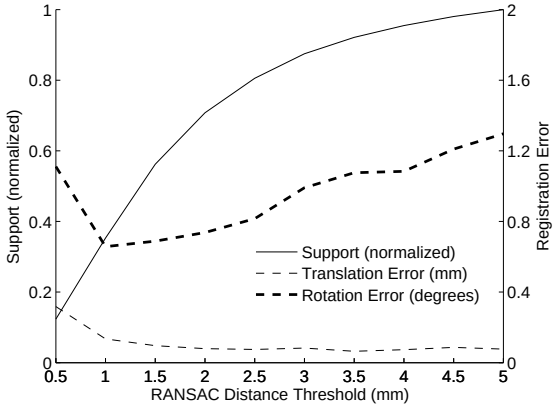


Figure 6: Effect of the RANSAC distance threshold (d_{ransac}) on the registration performance.

performance. The studies contained in this section show how varying these parameters affects algorithm performance. As explained earlier, M_d and d_{ransac} should not need to be tuned, and so optimal ranges and values for these parameters are determined. The effect of varying σ_f and τ_{tissue} on registration time and accuracy are also described, and suitable ranges for these parameters are explored.

The studies described in this section were performed using images acquired in a water tank (Figure 4) using combinations of known translations (0, 3, 6, and 9 mm) and rotations (0, 4, and 8 degrees). The objective of the registration algorithm is to register images at a frame rate of roughly 30Hz, and so this range translates to probe motions upwards to 0.27 meters per second and 242 degrees per second, which is expected to be an upper bound on expected probe motion. The ultrasound volume size was $144 \times 80 \times 112$ voxels along the lateral, elevational, and axial directions, respectively, with a resolution of $0.75 \times 0.65 \times 0.83$ mm/voxel. For the following studies, the images were registered to an image at 0 mm of translation and 0 degrees of rotation using the described parameters, and registration errors computed according to the offset of each image. The measurement of "support" shown in the accompanying plots comes from the RANSAC algorithm and is a measure of the number of final matches used to compute the transformation between images.

3.1.1. Sample Offset Scale Factor, M_d

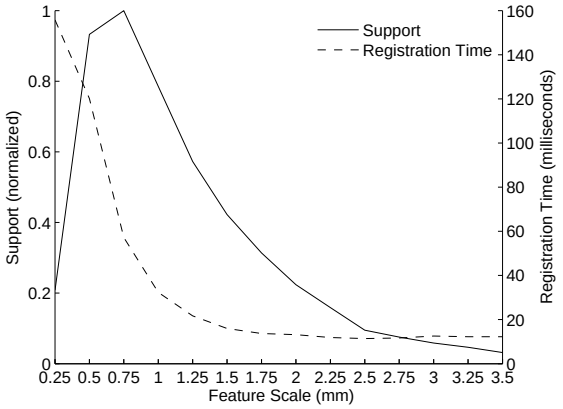
The size of a feature is related to the designated feature scale at which the LoG is computed. When computing the LoG, for a detection error of less than 0.1%, the width of the LoG kernel is suggested to be roughly 8.5 times the feature scale (Gunn, 1999). Therefore, the extent of the grid used to compute the feature descriptor should be *at least* as large, meaning the corresponding sample offset scale factor would have a value of at least 2.125. If M_d was made to be too large, sample points would be so spread out that the descriptor would include distant image data unrelated to the feature.

We therefore explored how varying M_d would affect registration performance to determine an optimal range for the value of M_d. We did this by registering the images of known position using varying M_d from 1 to 8. The registrations were computed using $d_{ransac} = 1.5$mm, $\tau_{tissue} = 150$, and three different feature scales ($\sigma_f = 1$, 1.5, and 2mm). The three spatial scales were determined qualitatively to be a reasonable range for the feature scale for the given images. As the RANSAC algorithm is a random process, the registration results can vary slightly from one execution to the next, and so we performed the registrations 10 times for each feature scale, resulting in 30 trials for each scale factor. The results of this analysis are averaged over all feature scales and trials, a plot of which can be seen in Figure 5.

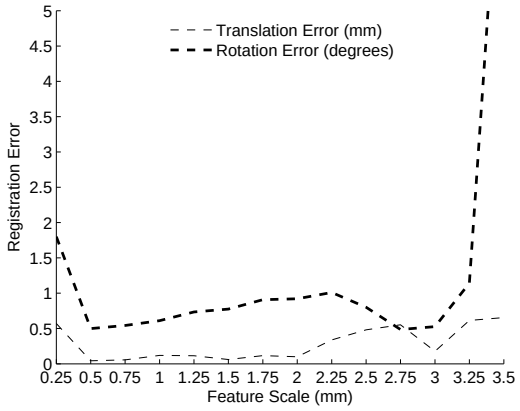
The plot indicates that when M_d is between 3 and 4, the support is at its maximum. Meanwhile, the registration error (both translation and rotation) is roughly constant for M_d between 2 and 6. Therefore, in the interests of generating the maximum support and limiting error, we use $M_d = 3.5$ for the remaining studies.

3.1.2. RANSAC Distance Threshold, d_{ransac}

The RANSAC distance threshold should be made as small as possible to limit registration error, but the lower bound on the threshold value is near the spatial resolution of the 3DUS volume. To show this, we analyzed the effect of the RANSAC distance threshold, d_{ransac}, on the algorithm performance in much the same way as was done for M_d, with the difference being that for this study we fixed M_d at 3.5 while d_{ransac} varied between 0.5 and 5mm. The results were averaged over all feature scales and trials (Figure 6). As anticipated, for d_{ransac} below the spatial resolution of the volume (about 0.75mm/voxel) and for large values, the registration performance diminished. Appropriate values for d_{ransac} would then be those between 1 and 2mm, as registration error is comparatively small and constant in this region. Accordingly, we use a d_{ransac} value of 1.5mm for the remaining studies.

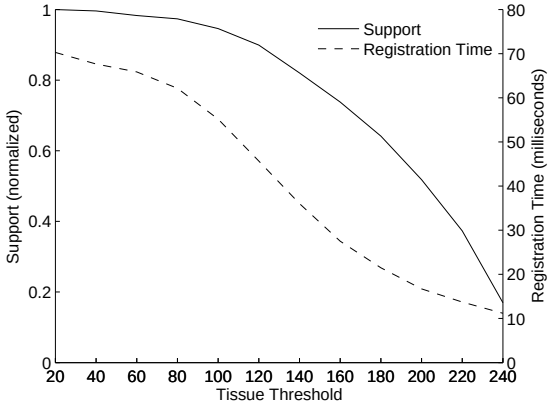


(a)

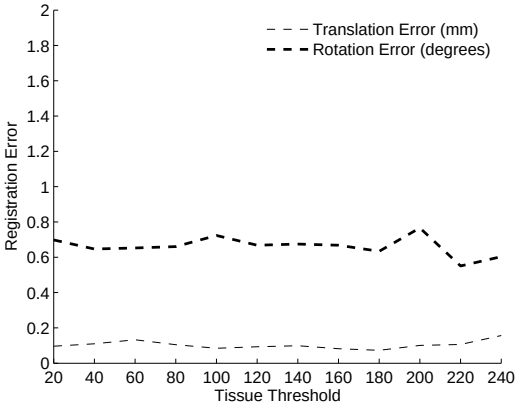


(b)

Figure 7: Effect of feature scale (σ_f) on (a) the registration support and execution time and (b) the registration accuracy.



(a)



(b)

Figure 8: Effect of the tissue threshold (τ_{tissue}) on (a) the registration support and execution time and (b) the registration accuracy.

3.1.3. Feature Scale, σ_f

As previously mentioned, an appropriate feature scale at which to detect features will depend on the imaged tissue and anatomy. To elucidate the effect of varying feature scale on registration performance, we computed the registrations for the same images as used in the previous two studies, using the optimal values of $M_d = 3.5$ and $d_{ransac} = 1.5$mm and making $\tau_{tissue} = 150$. The value of the feature scale was varied between 0.25mm and 3.5mm, and similar to the previous studies, several trials (25 trials) were performed for each feature scale. The performance values were then averaged over all trials. The effect of the feature scale on the number of matches found and the registration time can be seen in Figure 7(a), while Figure 7(b) shows the effect of the scale on the registration accuracy.

With regards to the accuracy, it can be seen in Figure 7(b) that in an effort to keep the translation error as low as possible, a value of σ_f between 0.5mm and 2mm should be used. In this range, the rotation error is smallest for smaller σ_f, and so it might be assumed that a σ_f of 0.5mm should be used. However, as seen in Figure 7(a), the registration time for small feature scales becomes large due to the increased number of features being detected. It would therefore be prudent to choose a small enough feature scale to maintain registration accuracy, but large enough that the registration time will allow for real-time operation. For example, if the frame rate is 30Hz, the registration time should be $\leq$ 30ms, therefore an appropriate feature scale would be 1-1.5mm. It is important to note that the registration time is dependent on the GPU hardware and architecture. While this study was performed using an NVIDIA GTX260 graphics card, more recent hardware releases would prove to substantially reduce the reported registration times. A more detailed analysis of registration time versus GPU hardware can be found in Section 3.6.

3.1.4. Tissue Threshold, τ_{tissue}

The tissue threshold is primarily used to limit how many features are found in an image, and so it is important to understand its effect on the registration performance. We examine the effects of the tissue threshold on the registration performance using the same water tank images from the previous three studies. For this study, $M_d = 3.5$, $d_{ransac} = 1.5$mm, and $\sigma_f = 1.0$mm, while we made τ_{tissue} to vary between 20 and 240 (assuming intensities are in the range of 0-255). For each value of τ_{tissue}, 25 trials were

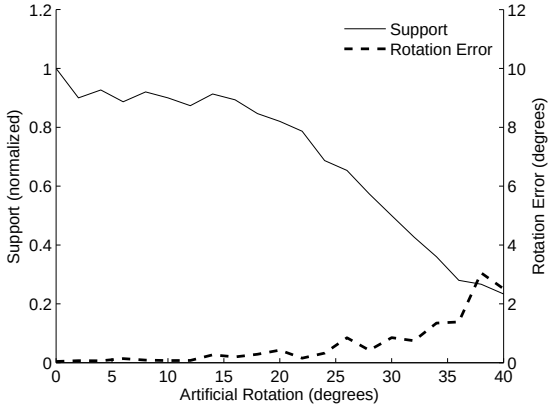


Figure 9: Registration accuracy and normalized support for an artificially rotated clinical 3DUS image.

performed. The results of the registration performance, shown in Figure 8, were then found by averaging across these trials.

For larger values of τ_{tissue}, fewer features are being found in each image and therefore a lower support results. While this has the effect of lowering the registration time, this has little effect on the registration accuracy when τ_{tissue} is below roughly 200. As a result, τ_{tissue} is a suitable parameter to tune to adjust the registration time to below acceptable limits without affecting accuracy. For instance, it can be seen in Figure 8 that changing τ_{tissue} from 120 to 170 has almost no effect on the registration accuracy, yet the registration time is cut in half.

3.2. Registration Accuracy Under Artificial Rotation

Given that we have chosen to construct rotationally variant feature descriptors, it is helpful to know the extent to which these features can accurately register two images that are rotated relative to each other. In an ideal scenario, where the imaging is independent of the ultrasound probe orientation to the anatomy, the feature descriptors are affected by rotation but not translation. To analyze the effect of image rotation on the registration accuracy, we perform a study similar to that done in Ni et al. (2008), where the stability of feature matching under rotation was analyzed for a registration method that used the rotationally invariant 3D SIFT feature descriptor.

For this study, we generated a set of artificially rotated images by applying a known rotational transformation to a baseline image. We then registered the transformed image to the baseline image ($M_d = 3.5$, $d_{ransac} = 1.5$mm, $\sigma_f = 1.5$mm, and $\tau_{tissue} = 150$) and computed the registration error. This was done for rotation angles between 0 and 40 degrees. The resulting support (normalized by the support from the 0 degree image registration) and the registration accuracy for a clinical image generated using the X7-2 transducer can be seen in Figure 9. The results suggest that rotations up to roughly 20° can be resolved using the presented registration method.

Table 1: Registration strategy comparison – position and orientation error between the first and last images registered in an ultrasound sequence acquired using a loop trajectory of the transducer

		Error	
Trajectory	Registration Strategy	Translation (mm)	Rotation (degrees)
1	RTG	0.36	0.39
	RTP	0.56	1.06
2	RTG	0.28	0.36
	RTP	1.73	2.81
3	RTG	0.27	0.31
	RTP	4.70	4.15
Avg.± Std. Dev.	RTG	0.30 ± 0.05	0.35 ± 0.04
	RTP	2.33 ± 2.13	2.67 ± 1.55

3.3. Drift Analysis for Different Registration Strategies

As we are proposing to use the presented registration algorithm to generate arbitrarily large 3DUS volumes, it is important that the algorithm accumulate minimal error as more and more volumes are added to the mosaic, especially if the same region is imaged several times throughout an acquisition. To analyze the accumulation of error for the RTP and RTG registration strategies discussed in Section 2.7, we conducted a study which analyzed the accumulation of registration error by computing the difference between the first and last frame in a loop trajectory of the ultrasound probe. For this study, we again used the water tank setup shown in Figure 4, allowing us to accurately acquire the first and last frames of the loop in the same position. The parameter values used were $M_d = 3.5$, $d_{ransac} = 1.5$mm, $\sigma_f = 1.5$mm, and $\tau_{tissue} = 150$. The trajectories were created using a series of translations and rotations along and about the y-axis of the image, where translations and rotations are traversed in 3mm and 4° increments, respectively. If we designate a translation as T and rotation as R, then the studied trajectories were as follows:

1. T(48mm), T(-48mm)
2. T(48mm), R(20°), T(-48mm), R(−20°)
3. T(48mm), R(60°), T(-48mm), R(−60°)

The accumulated error for the three trajectories can be seen in Table 1. While the RTG method exhibited roughly constant error for the three different trajectories, the RTP method exhibited a larger error than the RTG method for every trajectory, and also accumulated an increasing amount of error as the number of frames and total rotation throughout the trajectory increased. This suggests that the RTG method is better suited for registering several volumes to create a 3DUS mosaic. Qualitative verification of this conclusion can be seen by comparing the mosaics generated using the RTP and RTG methods shown in Figure 3 ($M_d = 3.5$, $d_{ransac} = 1.5mm$, $\sigma_f = 1.5$mm, $\tau_{tissue} = 150$), which were created from a left ventricular axis sweep (437 3DUS volumes of size $160 \times 64 \times 208$ voxels and resolution

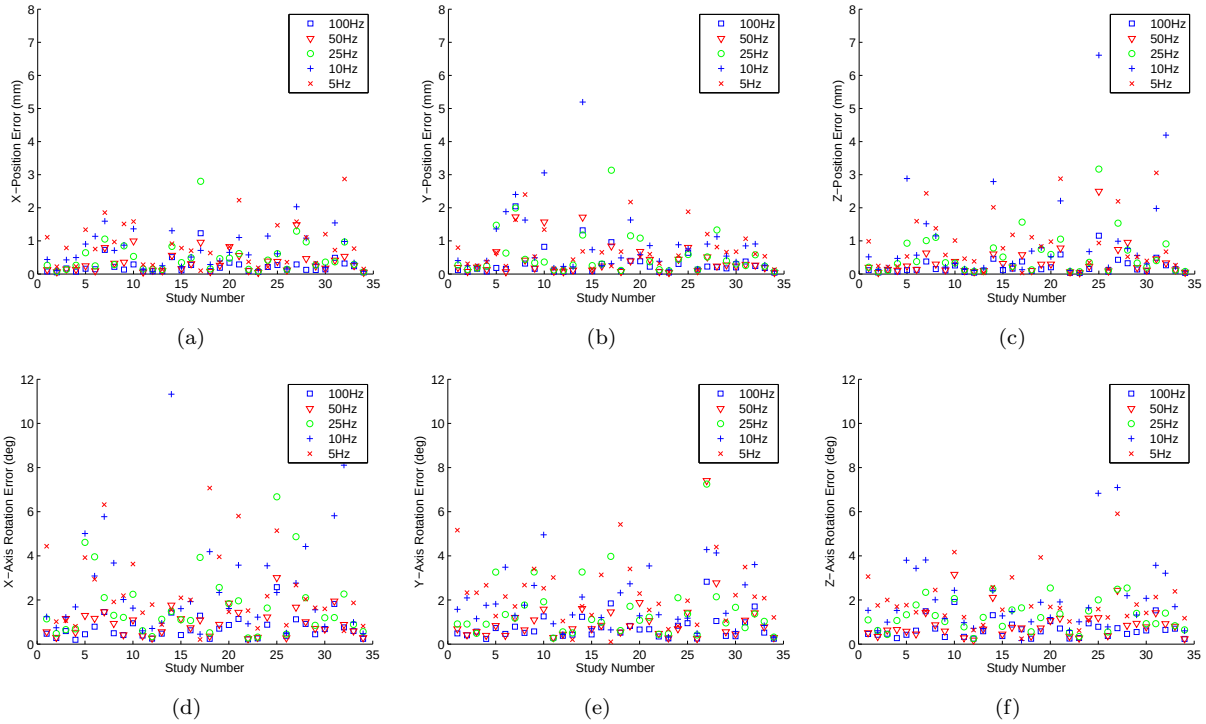


Figure 10: Registration position (a-c) and orientation (d-f) accuracy for the stabilized ECG-gated 3DUS sequences. Gating was done at varying rates over the cardiac cycle. Images were of mitral and aortic valves acquired in 34 clinical studies using either a transesophageal or transthoracic approach. Respiration was not stopped nor was respiration-gating used for the acquisitions.

$0.56 \times 0.70 \times 0.58$mm/voxel) of a porcine heart in a water tank. The mosaic created using the RTG method shows clear boundaries and texture within the tissue, while the mosaic created using the RTP method has much less definition and a different overall shape. This again supports the conclusion that the RTG method is better suited for creating 3DUS mosaics from many images.

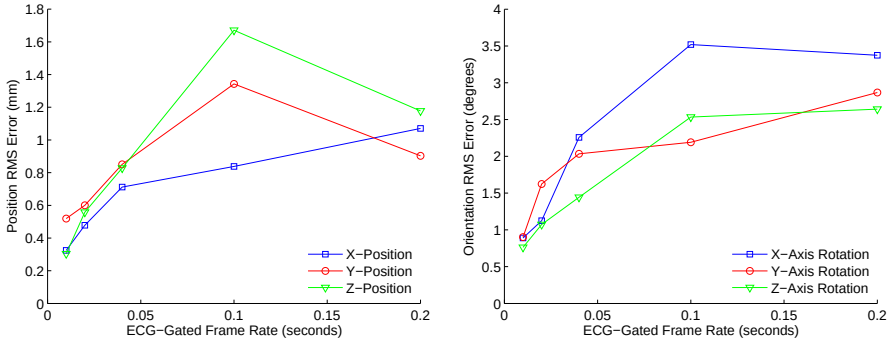
3.4. Registration Accuracy in Stabilizing 3DUS Volumes of Heart Valves Over a Cardiac Cycle

To study fast moving heart valves using 3DUS, it is beneficial to view the valves at a higher frame rate than what can be originally acquired on a 3DUS machine. This can be done using ECG-gating and acquiring images at specific times during the cardiac cycle over several heartbeats. However, depending on the desired frame rate, the acquisition process can take a long time, upwards to several minutes. During this time, it is likely that the ultrasound probe will shift, and it is typically not possible to cease respiration during this time. Therefore, in the gated sequence there will be small shifts of the valve relative to the probe which appear in the images. These shifts can be accounted for using a rigid transformation model.

As the shifts are generally small relative to the size of the volume, which is typically made large enough to view the entire valve, we studied the effectiveness of the presented rigid registration algorithm to stabilize the images across the gated sequence. As the gated 3DUS sequence is of a beating heart, there is inevitably some amount of non-rigid deformation of the tissue between frames. We therefore acquired gated sequences at multiple sample rates across the cardiac cycle to determine at what rates the rigid registration algorithm could accurately register the gated sequence in the presence of this non-rigid tissue deformation.

For this study, we acquired gated 3DUS sequences across a cardiac cycle in 34 clinical cases. Images were taken of either the mitral or aortic valve, and were acquired using either a transesophageal or transthoracic approach. Gated sequences were acquired at 5Hz, 10Hz, 25Hz, 50Hz, and 100Hz across the cardiac cycle. We did not cease respiration nor did we perform respiration gating for the acquisitions. In the context of this study, "gated sequence" refers to the resulting collection of frames acquired over several cardiac cycles and represented relative to the single cardiac cycle. Any frame in this sequence is referred to as a "gated frame" and the frequency of the gated frames in the gated sequence is referred to as the "gated frequency." For each patient, we constructed several gated sequences of varying gated frequencies.

To quantify registration error, we first registered the gated frames in a gated sequence in a successive fashion using the RTG method. We did this in several trials for each gated sequence such that the start frame was varied among all gated frames in a gated sequence. For instance, if there were three gated frames in the gated sequence, $\{F_1,F_2,F_3\}$, we would run three different trials, where the registration order would be either $\{F_1,F_2,F_3\}$, $\{F_2,F_3,F_1\}$, or $\{F_3,F_1,F_2\}$. In the first trial, for instance, F_2 would be registered to F_1, and then F_3 registered to the already



(a) (b)

Figure 11: Position and orientation RMS registration error for corresponding frame rates across the 34 clinical studies. Acquisitions were ECG-gated over a cardiac cycle. Images were of mitral and aortic valves using either a transesophageal or transthoracic approach. Respiration was not stopped nor was respiration-gating used for the acquisitions.

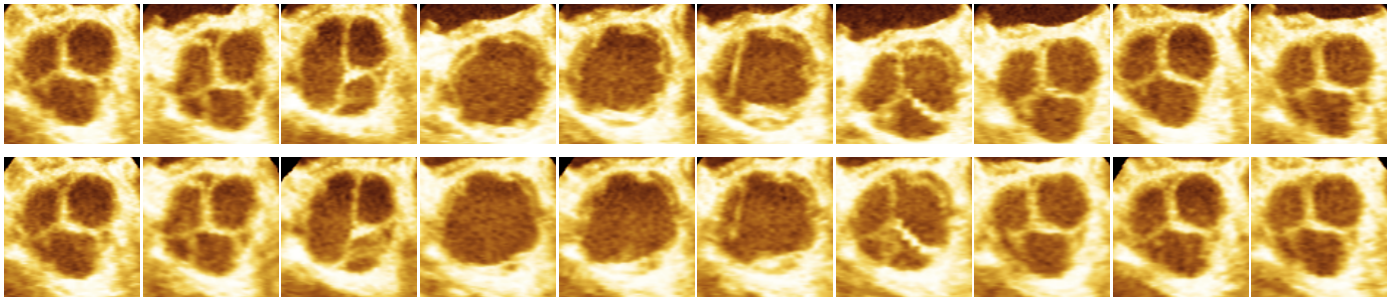


Figure 12: Short axis images of the aortic valve in a gated 3DUS sequence across the cardiac cycle acquired using a transesophageal approach. (*Top*) Originally acquired gated 3DUS sequence. (*Bottom*) Stabilized 3DUS sequence. Notice the position of the valve in the original sequence varies within the image, whereas the position of the valve in the stabilized images is relatively constant throughout the sequence. The displacements in the original images were due to respiration and movement of the probe during the gated acquisition.

registered F_2. We then registered the first frame in the trial to the already registered last frame. For instance, in the first example trial, we would register F_1 to the already registered F_3. The registration error was then the position and orientation offset of the registered first frame from the original first frame in the trial, where the position offset is measured at the center of the conical 3DUS volume. Ideally, if there was no registration error, if we were to register the first frame in the gated sequence to the last frame after all frames were stabilized, the first frame and the registered first frame would be in the same location and orientation as the heart follows a cyclic motion. The position and orientation displacements were averaged across all trials performed on the gated sequence and represent the average registration error that accumulated across the entire gated sequence. The feature scale and tissue threshold were tuned only once for each gated sequence using the auto-tuning method and two consecutive gated frames, and the determined parameter values then used for all trials. For the auto-tuning method, parameters were modified until the support was in the range of 100-150. This resulted in average registration times of roughly 100ms. The longer registration time was a result of the volume being quite large so that the entire valve could be imaged. The average time to acquire each volume was roughly 125ms, and the average cardiac cycle time was roughly 670ms (i.e. heart rate of roughly 90 beats per minute).

The average registration errors in position and orientation across all trials for the different gated frames rates in the 34 clinical cases are shown in Figure 10. It can be seen that larger registration errors were typically found for lower gated frame rates. In these cases, the time between gated frames was larger and therefore a larger amount of non-rigid tissue deformation occurred. The RMS registration errors in position and orientation across the 34 clinical cases are summarized in Figure 11. If we assume suitable registrations to be those whose error was less than the limits of visual inspection, which were shown in Bankman (2000) to be 2mm in position and 2^o in orientation in the registration of MR and CT brain images (Shekhar et al., 2004), then the lowest allowable gated frame rate at which the presented registration algorithm can accurately stabilize the gated sequence is roughly 25Hz (i.e. maximum allowable time between gated frames is roughly 40ms). Images from a gated sequence of the aortic valve before and after stabilization can be seen in Figure 3.4.

3.5. Registration Accuracy in a Water Tank Mosaic

To assess the accuracy of the registration algorithm for the purpose of generating a 3DUS mosaic, we manually se-

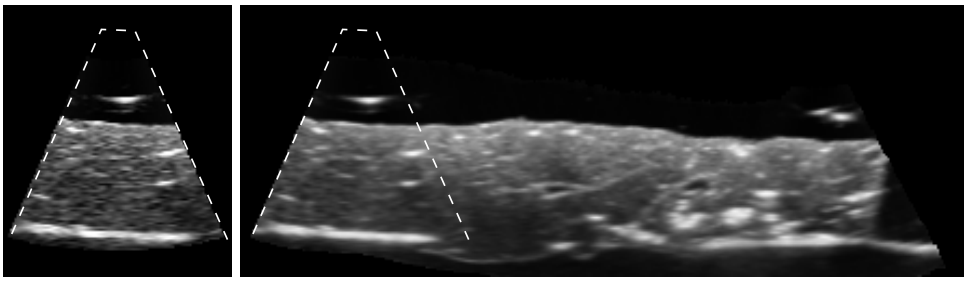


Figure 14: Cross-sections from a single 3DUS volume (*left*) and corresponding mosaic (*right*) of a porcine liver in a water tank. The dotted lines in the image on the right correspond to the size and location of the volume on the left within the mosaic.

Table 2: Summary of the mosaic accuracy study which measured inter-fiducial distances AD and BC. Variability is reported as $avg. \pm std.dev.$

	AD	BC	Both
n	15	15	30
Volumes in Mosaic	68.9 ± 7.9	65.2 ± 9.2	–
Actual Length (mm)	121	118	–
Length in Mosaic (mm)	120.5 ± 0.1	116.7 ± 0.9	–
Error (mm)	0.9 ± 0.8	1.4 ± 0.7	1.1 ± 0.8
Error (%)	0.72 ± 0.64	1.20 ± 0.59	0.96 ± 0.65

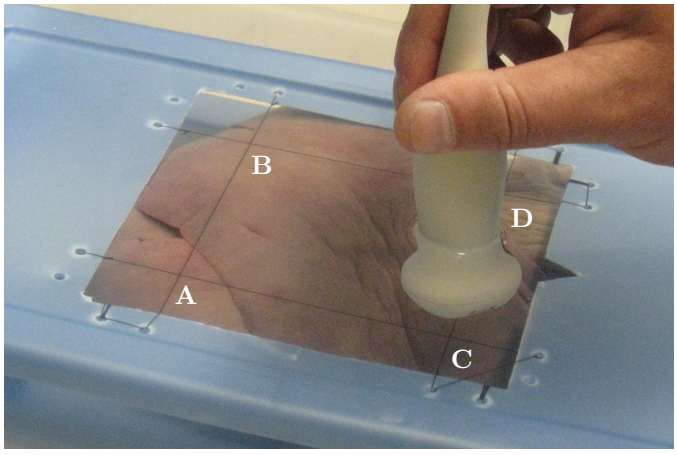


Figure 13: Water tank setup showing a porcine liver with surrounding fiducials - suture intersections A, B, C, and D (occluded).

lected fiducials in a mosaic and compared the image-based inter-fiducial distances versus the actual inter-fiducial distances. The object being imaged was a porcine liver in a water tank (Figure 13), and the fiducials (A, B, C, and D in Figure 13) were the intersections of sutures placed in a grid pattern directly above the liver. We acquired sweeping ultrasound data (using the X7-2 ultrasound transducer) by starting at fiducial A or B and ending at D or C, respectively. Typically 50-80 volumes were acquired in each sweeping acquisition, where each volume had dimensions of $144 \times 112 \times 112$ voxels and a resolution of $0.55 \times 0.54 \times 0.63$mm/voxel. For each fiducial pair (AD or BC), 15 mosaics were created using the RTG strategy and the fiducials in those mosaics manually selected. A cross-section from one of the mosaics is shown in Figure 14. The parameters used for the registration were $M_d = 3.5$, $d_{ransac} = 1.5$mm, $\sigma_f = 1.5$mm, and $\tau_{tissue} = 150$. A summary of the results from this study are shown in Table 2, where it is shown that an average error of less than 1% was found.

3.6. Registration Running Times

The presented feature-based registration algorithm can operate in real-time, but it is helpful to know how the registration time is distributed over the different components of the algorithm. For this, we have included a detailed time analysis (Figure 15) of the registrations performed to compute the mosaics shown in Figures 3 and 14. For these mosaics, the same M_d and d_{ransac} values were used as found in the studies described in Sections 3.1.1 and 3.1.2, while τ_{tissue} and σ_f were tuned for the different types of images. The time analysis does not include the time to render the mosaic as this can be done in a separate thread.

To determine the effect of the graphics card on the registration time, we also performed the same analysis as was done above but instead of using an NVIDIA GTX 260 (which has 192 cores), we used an NVIDIA GeForce 8600M GT (which has 32 cores). We found nearly identical results with regard to the percentage of time that each component took, however, the average total registration time was roughly six times longer (about 170ms). This indicates a linear relationship between the number of cores and the registration time. This also suggests that if for instance the same registrations were performed on the latest NVIDIA hardware (i.e. NVIDIA GTX 580 with 512 cores), a 2.6× speed increase could be achieved (i.e. registration time decrease from 30ms to 12ms).

It is worth noting that with the CUDA and device architecture used in all of the presented studies (NVIDIA

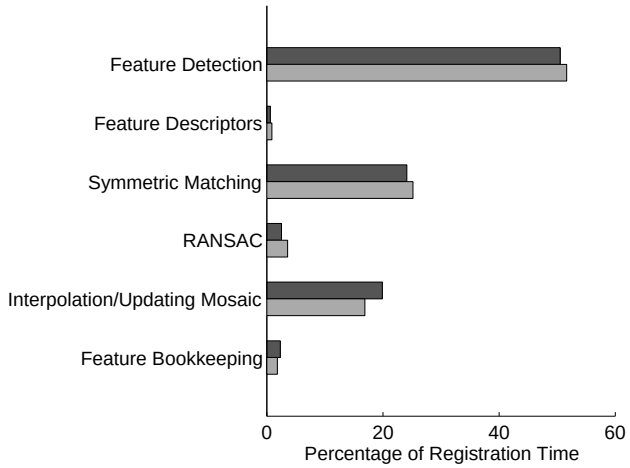


Figure 15: Average running times for the components of the registration algorithm relative to the average total registration time. Results are for the RTG mosaic shown in Figure 3 (*dark gray*) and the mosaic shown in Figure 14 (*light gray*), where the average registration times were 28.7ms and 24.4ms, respectively.

GTX 260 Graphics Processing Unit, Driver/Runtime Version 3.0, Compute Capability 1.1, NVIDIA Corporation, Santa Clara, CA, USA), the time to load a 3DUS volume ($\sim 128^3$) onto the GPU as a texture took about 20 times longer than loading the volume onto the GPU as a linear memory block. Therefore, linear memory blocks in global memory were used for all data loaded onto the GPU in this study.

4. Discussion

4.1. Registration Algorithm Performance

In this paper we have presented a feature-based 3DUS rigid registration algorithm capable of registering volumes in a 3DUS sequence in real-time. The registration algorithm is most closely related to the method in Ni et al. (2008), which is a SIFT-based method that ensures that features are detected at multiple spatial scales and also that feature descriptors are scale and rotationally invariant. This is a consequence of the original design of SIFT, which was designed for detecting and matching features in 2D photos (Lowe, 2004). However, while features in 2D photos may appear at different scales simply by changing the position and orientation of the camera relative to the imaged objects, the scale of features in 3DUS should not change, especially in the case of real-time 3DUS where it is assumed that the probe undergoes small displacements between frames. Also, in assuming the probe undergoes small displacements, the need for a feature descriptor that can handle large rotational and translational displacements becomes unnecessary. It has also been shown that ultrasound is not a rotationally invariant imaging modality (Grau et al., 2007; Schneider et al., 2010), as the appearance of anatomy is dependent on the direction of acoustic propagation relative to the imaged structures, making rotationally invariant feature descriptors unnecessary.

We therefore only search for features at a single spatial scale. We also devised a simple rotationally variant feature descriptor formation method. The simplified feature detection and descriptor formation methods, as well as using a GPU accelerated framework, results in at least a 1000× speed increase (from 1 minute down to under 100ms for single volume registration) compared to the method in Ni et al. (2008). While a substantial speed increased was realized, we showed in several validation studies that this was not at the expense of registration accuracy. Comparable performance was in fact found in those studies (Sec. 3.2 and Sec. 3.5) where a similar study was performed for the 3D SIFT method.

Comparable registration accuracy and fast execution was made possible through the use of the RTG registration strategy. In maintaining a minimal but representative set of global features and storing and indexing these features efficiently, we were able to limit the accumulation of registration error, even when mosaicing several hundred 3DUS volumes. The most promising results for the RTG method were those described in Sec. 3.3, where for loop trajectories in a water tank, the RTG method accumulated less than 0.5mm and 0.5^o of translation and rotation error, respectively.

To determine the clinical usefulness of the presented rigid registration algorithm, and to also determine the ability of the algorithm to handle varying degrees of non-rigid tissue deformation, we studied the accuracy of the registration method in stabilizing an ECG-gated 3DUS sequence acquired over the cardiac cycle at varying gated frame rates. The 3DUS images in the clinical studies were of the mitral or aortic valve in a beating heart, acquired using either a transesophageal or transthoracic approach without the use of respiration gating or ceasing respiration. Due to small probe movements and respiration during the long acquisition times, the location of the valves would shift relative to the ultrasound probe causing the gated images to appear to shift (Figure 3.4). The rigid registration algorithm was therefore employed to cancel out this undesired movement as seen in the images. We found that the registrations were accurate (within the limits of visual inspection (Bankman, 2000)) for gated frame rates of 25Hz or less (i.e. time between frames is less than 40ms). This suggests that the rigid registration algorithm is accurate even in the presence of small degrees of non-rigid deformation in 3DUS volumes. A similar study of rigidly aligning misaligned or shifted 3DUS images of the beating heart was shown in Shekhar et al. (2004) for stress echocardiography studies, which suggests that the presented method might also be suitable for this application.

As there was no gold standard to use for directly measuring accuracy in the clinical validation study, the registration error was found as the drift across the sequence by finding the position and orientation difference between the frame at the start of the sequence and the same frame

registered to the end of the stabilized sequence. Since the heart follows a cyclic motion, these frames should ideally be in the same location were there no registration error. A limitation of this measurement is that it relies on the assumption of perfect cyclic motion of the heart, or that the heart is always in the same configuration at corresponding times relative to the R-peak in the ECG signal (i.e. the event by which frames were gated). This assumption breaks down in the presence of increasing beat-to-beat variation or arrhythmias, however, we did not observe any significant heart rate variations in the examined cases.

In the clinical validation study, we showed that registration accuracy declined for low frame rates, which was due to larger amounts of non-rigid deformation between frames that could not be resolved by the rigid transformation model. For this reason, we can assume that the proposed registration method would not be suitable for motorized 3D probes when imaging fast moving cardiac structures, as these probes generally have lower frame rates and a large amount of non-rigid tissue deformation would be expected. However, the proposed method should prove sufficient when imaging relatively static structures, such as the liver, using a motorized 3D probe and assuming small displacements between frames. The same could also be assumed for freehand 3DUS volumes.

While the registration method is accurate even in the presence of small degrees of non-rigid deformation, this type of deformation will be most noticeable in the application of stitching together (mosaicing) several 3DUS volumes. For larger degrees of non-rigid deformation, stitching artifacts would be expected. However, the presented registration algorithm is not designed for these situations, but rather a non-rigid registration method should be employed. In the case of mosaicing, the presented method is best suited for images of static tissue, or for images of moving tissue taken at the same time during its motion (for instance, images of the heart taken at the same time during the cardiac cycle). Also, while we used an averaging method for compositing mosaics, this was done because this method is simple, operates quickly on the GPU, and produces favorable results as seen in visual inspection. However, this is not a novel contribution in the area of improving image quality. There are more sophisticated compositing methods which better integrate the information from multiple 3DUS images (Grau and Noble, 2005; Yao et al., 2009; Rajpoot et al., 2011).

The presented registration algorithm operates under a small-displacement assumption to compute accurate transformations in real-time. While it is tempting to define "small" as a percentage of volume size, in the case of the presented feature-based registration method, it is rather a function of image content. The method requires a coincidence of features from frame to frame, and so more important than volume overlap is overlap of salient regions. It is important to take this into consideration when registering anatomy with "dead zones" that have little or no saliency, as these regions will not contribute to the registration. These regions have the potential to cause larger than expected registration errors if 3DUS volumes are roughly the same size (i.e. if the "dead zone" occupies a large percentage of the 3DUS image). For instance, this likely will not be an issue when imaging the liver, as it is a dense organ with unique and varying internal structure and texture. However, regions such as the pools of amniotic fluid as seen in fetal ultrasound, or the blood pool of the left ventricle as seen in echocardiography, have the potential to be more problematic.

While the algorithm requires that the feature scale and tissue threshold parameters be tuned for a given application, we have shown that these parameters can be tuned automatically using a very simple greedy algorithm that attempts to force the resulting registration support or time to fall within a desired range. In particular, this auto-tuning method was used extensively in the clinical validation study to stabilize the gated 3DUS volumes. The auto-tuning method exploits information that we gathered from the sensitivity studies, specifically, that the registration support and time are inversely proportional to the feature scale and tissue threshold. While not shown, it can be reasoned that an online auto-tuner (i.e. one that adjusts the parameter values during the registration process) could also prove useful if, for instance, it was important to regulate registration time. While the feature scale could not be changed (as the detected features need to be at a single scale) the tissue threshold could be adjusted accordingly from one registration to the next to keep the registration time within acceptable limits.

The registration algorithm makes it possible to register 3DUS volumes online without the need for external tracking systems (such as electromagnetic or optical tracking systems), which means that when acquiring data to use with the presented method, any existing 3DUS system should suffice. A limitation of the presented method, however, compared to systems that use a tracking system, is that because a continuous stream of 3DUS volumes is required, any interruption in that stream, such as lifting the probe from the object being imaged, will likely cause subsequent volume registrations to fail under the RTG and RTP registration strategies. For this reason, tracking systems can still be a valuable registration tool, especially if used in combination with an image-based registration method.

4.2. Future Work

The presented work has obvious applications other than those shown, such as in hepatic tumor location and intervention, echocardiography and intracardiac echocardiography, and fetal ultrasound. However, to be more robust and useful for these applications, there are several aspects of the algorithm that could be improved. The development of a real-time non-rigid registration algorithm would be especially helpful for applications where soft tissue, such as the breast, is imaged, as soft-tissue deforms in a largely

non-rigid fashion, especially due to ultrasound probe contact. Other areas of possible improvement are to integrate a real-time temporal registration algorithm based on the ECG signal so that, for instance, a large FOV 4D image of the heart could be constructed in real-time. Lastly, in order to composite images that were acquired in acquisitions where the probe underwent large displacements, it would be best to use instead of averaging a more elegant compositing method (for example, Rajpoot et al. (2011)) that takes into account the appearance of anatomy under largely varying probe orientations. The challenges with the proposed future work are not the methods themselves, but integrating the methods into the current system and maintaining a real-time architecture.

5. Conclusion

The presented 3DUS rigid registration algorithm takes into account the nature of real-time 3DUS volumes to simplify the process of feature detection and extraction, and in doing so, along with using a GPU accelerated framework, is able to operate in real-time (i.e. register volumes as fast as they are acquired). The method also uses techniques to maintain and integrate a set of global features into the registration process, which helps to limit the accumulation of registration error. Several studies revealed appropriate ranges and values for the algorithm parameters, and provided insight into the effects of varying parameter values on the registration performance. An auto-tune method was also described which automatically tunes necessary parameters, making the registration method fully automatic. Accuracy studies showed that the algorithm's performance is comparable to similar existing methods, and a clinical validation study showed that the registration method could accurately stabilize gated sequences of the beating heart for gated frame rates at and below 25Hz, indicating the registration method is able to operate in the presence of small degrees of non-rigid deformation.

6. Acknowledgments

This work was supported in part by the U.S. National Institutes of Health under grants RO1 HL073647-01 and 5 F32 HL084965-03, and by the U.S. National Science Foundation Graduate Research Fellowship Program.

7. References

Arun, K., Huang, T., Blostein, S., 1987. Least-squares fitting of two 3-D point sets. IEEE Transactions on Pattern Analysis and Machine Intelligence 9, 698–700.

Bankman, I., 2000. Handbook of medical imaging: processing and analysis. Academic Press, San Diego, CA.

Cen, F., Jiang, Y., Zhang, Z., Tsui, H., Lau, T., Xie, H., 2004. Robust registration of 3-D ultrasound images based on gabor filter and mean-shift method. Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis LNCS 3117, 304–316.

Fischler, M., Bolles, R., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24, 381–395.

Foroughi, P., Abolmaesumi, P., Hashtrudi-Zaad, K., 2006. Intra-subject elastic registration of 3d ultrasound images. Medical image analysis 10, 713–725.

Francois, R., Fablet, R., Barillot, C., 2003. Robust statistical registration of 3D ultrasound images using texture information, in: IEEE Int. Conf. on Image Processing, Citeseer. pp. I–581–4.

Gee, A., Treece, G., Prager, R., Cash, C., Berman, L., Street, T., 2003. Rapid registration for wide field-of-view freehand 3d ultrasound. IEEE Transactions on Medical Imaging 22, 1344–1357.

Grau, V., Becher, H., Noble, J., 2007. Registration of multiview real-time 3-D echocardiographic sequences. IEEE Transactions on Medical Imaging 26, 1154–1165.

Grau, V., Noble, J., 2005. Adaptive multiscale ultrasound compounding using phase information. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005 , 589–596.

Gunn, S., 1999. On the discrete representation of the Laplacian of Gaussian. Pattern Recognition 32, 1463–1472.

Hartley, R., Zisserman, A., 2003. Multiple view geometry in computer vision. Cambridge University Press New York, NY, USA. 2nd edition.

Krucker, J., LeCarpentier, G., Fowlkes, J., Carson, P., 2002. Rapid elastic image registration for 3-d ultrasound. Medical Imaging, IEEE Transactions on 21, 1384–1394.

Kutter, O., Wein, W., Navab, N., 2009. Multi-modal Registration Based Ultrasound Mosaicing. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009 , 763–770.

Ledesma-Carbayo, M., Mahia-Casado, P., Santos, A., Perez-David, E., Garcia-Fernandez, M., Desco, M., 2006. Cardiac motion analysis from ultrasound sequences using nonrigid registration: validation against doppler tissue velocity. Ultrasound in Medicine & Biology 32, 483–490.

Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60, 91–110.

Makela, T., Clarysse, P., Sipila, O., Pauna, N., Pham, Q., Katila, T., Magnin, I., 2002. A review of cardiac image registration methods. IEEE Transactions on Medical Imaging 21, 1011–1021.

Mercier, L., Langu, T., Lindseth, F., Collins, L., 2005. A review of calibration techniques for freehand 3-D ultrasound systems. Ultrasound in Medicine & Biology 31, 143–165.

Mikolajczyk, K., 2002. Detection of local features invariant to affine transformations. PhD Thesis, Institut National Polytechnique de Grenoble, France .

Moradi, M., Abolmaesoumi, P., Mousavi, P., 2006. Deformable registration using scale space keypoints, in: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pp. 791–798.

Neemuchwala, H., Hero, A., Carson, P., 2005. Image matching using alpha-entropy measures and entropic graphs. Signal Processing 85, 277–296.

Ni, D., Qu, Y., Yang, X., Chui, Y., Wong, T., Ho, S., Heng, P., 2008. Volumetric ultrasound panorama based on 3D SIFT. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008 , 52–60.

Poon, T., Rohling, R., 2006. Three-dimensional extended field-of-view ultrasound. Ultrasound in Medicine & Biology 32, 357–369.

Porter, B., 2004. Three-dimensional medical ultrasound acquisition and data registration and fusion. Ph.D. thesis. University of Rochester.

Pratikakis, I., Barillot, C., Hellier, P., Memin, E., et al., 2003. Robust multiscale deformable registration of 3d ultrasound images. International Journal of Image and Graphics 3, 547–565.

Rajpoot, K., Grau, V., Noble, J.A., Becher, H., Szmigielski, C., 2011. The evaluation of single-view and multi-view fusion 3D echocardiography using image-driven segmentation and tracking. Medical Image Analysis 15, 514–528.

Rajpoot, K., Noble, J., Grau, V., Szmigielski, C., Becher, H., 2009. Multiview RT3D Echocardiography Image Fusion, in: Proceedings of the 5th International Conference on Functional Imaging and Modeling of the Heart, Springer-Verlag. pp. 134–143.

Rohling, R., Gee, A., Berman, L., 1998. Automatic registration of 3-D ultrasound images. Ultrasound in Medicine & Biology 24, 841–854.

Schneider, R., Perrin, D., Vasilyev, N., Marx, G., del Nido, P., Howe, R., 2010. Mitral Annulus Segmentation from 3D Ultrasound Using Graph Cuts. IEEE Transactions on Medical Imaging 29, 1676–87.

Shekhar, R., Zagrodsky, V., Castro-Pareja, C., Walimbe, V., Jagadeesh, J., 2003. High-Speed Registration of Three-and Four-dimensional Medical Images by Using Voxel Similarity. Radiographics 23, 1673.

Shekhar, R., Zagrodsky, V., Garcia, M., Thomas, J., 2004. Registration of real-time 3-d ultrasound images of the heart for novel 3-d stress echocardiography. IEEE Transactions on Medical Imaging 23, 1141–1149.

Soler, P., Gerard, O., Allain, P., Saloux, E., Angelini, E., Bloch, I., 2005. Comparison of fusion techniques for 3D+T echocardiography acquisitions from different acoustic windows. Computers in Cardiology, 2005 32, 141–144.

Wachinger, C., Wein, W., Navab, N., 2008. Registration Strategies and Similarity Measures for Three-dimensional Ultrasound Mosaicing. Academic Radiology 15, 1404–1415.

Wang, Z., Slabaugh, G., Unal, G., Fang, T., 2007. Registration of Ultrasound Images using an Information-Theoretic Feature Detector, in: Proc. IEEE International Symposium on Biomedical Imaging, IEEE, Washington, D.C.. pp. 736–739.

Wein, W., Camus, E., John, M., Diallo, M., Duong, C., Al-Ahmad, A., Fahrig, R., Khamene, A., Xu, C., 2009. Towards Guidance of Electrophysiological Procedures with Real-Time 3D Intracardiac Echocardiography Fusion to C-arm CT. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009 , 9–16.

Wein, W., Cheng, J., Khamene, A., 2008. Ultrasound based respiratory motion compensation in the abdomen. MICCAI 2008 Worshop on Image Guidance and Computer Assistance for Softissue Intervetions 32, 294.

Xiao, G., Brady, J., Noble, J., Burcher, M., English, R., 2002. Non-rigid registration of 3-D free-hand ultrasound images of the breast. IEEE Transactions on Medical Imaging 21, 405.

Yao, C., Simpson, J., Jansen, C., King, A., Penney, G., 2009. Spatial compounding of large sets of 3D echocardiography images, in: Proc. SPIE - Medical Imaging, pp. 726515–1–8.

Zhuang, X., Yao, C., Ma, Y., Hawkes, D., Penney, G., Ourselin, S., 2010. Registration-based propagation for whole heart segmentation from compounded 3d echocardiography, in: Proc. IEEE International Symposium on Biomedical Imaging, IEEE, Rotterdam, The Netherlands. pp. 1093–96.

Zikic, D., Wein, W., Khamene, A., Clevert, D., Navab, N., 2006. Fast deformable registration of 3d-ultrasound data using a variational approach. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006 9, 915–923.