Estimation of 3D cardiac deformation using spatio-temporal elastic registration of non-scanconverted ultrasound data

An Elen\textsuperscript{a}, Dirk Loeckx\textsuperscript{a}, Hon Fai Choi\textsuperscript{b}, Hang Gao\textsuperscript{b}, Piet Claus\textsuperscript{b}, Frederik Maes\textsuperscript{a}, Paul Suetens\textsuperscript{a} and Jan D’hooge\textsuperscript{b}

\textsuperscript{a}Medical Image Computing, Department of Electrical Engineering (ESAT);  
\textsuperscript{b}Cardiovascular Imaging and Dynamics, Department of Cardiovascular Diseases;  
Katholieke Universiteit Leuven, Herestraat 49 - bus 7003, B-3000 Leuven, Belgium

ABSTRACT

Current ultrasound methods for measuring myocardial strain are often limited to measurements in one or two dimensions. Spatio-temporal elastic registration of 3D cardiac ultrasound data can however be used to estimate the 3D motion and full 3D strain tensor. In this work, the spatio-temporal elastic registration method was validated for both non-scanconverted and scanconverted images. This was done using simulated 3D pyramidal ultrasound data sets based on a thick-walled deforming ellipsoid and an adapted convolution model. A B-spline based frame-to-frame elastic registration method was applied to both the scanconverted and non-scanconverted data sets and the accuracy of the resulting deformation fields was quantified. The mean accuracy of the estimated displacement was very similar for the scanconverted and non-scanconverted data sets and thus, it was shown that 3D elastic registration to estimate the cardiac deformation from ultrasound images can be performed on non-scanconverted images, but that avoiding of the scanconversion step does not significantly improve the results of the displacement estimation.

Keywords: Non-rigid registration, 3D Ultrasound, cardiac, scanconversion

1. INTRODUCTION

Myocardial motion and strain are measured as an important indicator to quantify myocardial function.\textsuperscript{1} Within ultrasound (US) deformation imaging, Doppler methods are traditionally used for quantitative assessment of regional wall deformation.\textsuperscript{2} A major disadvantage of this technique is that it is only capable of measuring the strain along the direction of propagation of the ultrasound wave, yielding a one dimensional strain measurement.\textsuperscript{3} Therefore, non-Doppler based methods have been proposed in order to measure all myocardial strain components within the image plane. Roughly, these non-Doppler methodologies can be subdivided in methods based on speckle tracking\textsuperscript{4–7} and image registration approaches.\textsuperscript{8,9} All of these methods allow to measure the strain in the 2D image plane. Although this is an obvious improvement, it remains limited as cardiac motion and deformation are truly 3D.

With the introduction of matrix array transducer technology, 3D US imaging of the heart has become feasible.\textsuperscript{10} Each of the non-Doppler methods described above might, in theory, be generalized to estimate myocardial motion and deformation in three dimensions. However, the challenge lies in the fact that the increased field of view of the 3D ultrasound system typically comes at the expense of both spatial and temporal resolution of the data sets. Therefore, de-correlation between subsequent ultrasound volumes becomes significant. As some of the above mentioned methods assume the gray values or the gray scale patterns of each structure to be relatively constant between subsequent image frames, their application might become impractical.

In a prior study of our group an alternative method using spatio-temporal elastic registration to estimate the full 3D strain tensor from US images was shown feasible and validated using simulated 3D US data sets.\textsuperscript{11,12} In the latter study, the simulated data sets contained images of the left ventricle (LV) in cartesian coordinates, since this is the traditional reference system to interpret images. However, array transducers record the images in a spherical reference system after which a scanconversion is applied to obtain the images in cartesian coordinates.
By performing the registrations on non-scanconverted images not only one step in the post-processing chain is omitted but the introduction of interpolation errors in the image (inherent to the scanconversion) is also avoided. In addition, the anisotropic redistribution of the recorded data over the image (as scanconversion results in higher information density close to the transducer) is avoided. The major advantages of cartesian images on the other hand are the easy visual interpretability, the true reproduction of distances and the easy availability of the images. Both coordinate systems might thus have advantages towards accurate spatio-temporal image registration. The purpose of the current work was therefore to compare the accuracy of the motion estimation in both coordinate systems using simulated 3D US data sets both before and after scanconversion.

This article is organized as follows. In Section 2, the elastic or non-rigid registration method is explained, as is the acquisition of the simulated non-scanconverted and scanconverted data sets. The quality measure to compare the results obtained from the different data sets is discussed as well. The results are listed in Section 3 and discussed in Section 4. Section 5 concludes.

2. METHODS

During elastic or non-rigid image registration, the optimal deformation field to transform an image (the floating image) is searched for, in such a way that the deformed floating image matches a second image (the reference image) as good as possible. If the reference and floating image are taken equal to two time frames of the same image sequence, the deformation of the left ventricle can be derived from the found deformation field. In this work, a B-spline transformation model was chosen to regularize the deformation field but other possibilities exist. Mutual information between the reference and deformed floating image was used to quantify the match between both. Penalties were added to this similarity measure for extra regularization. The parameters of the deformation field were optimized by maximization of the similarity of the similarity measure. All of these steps are discussed in more detail in the next paragraphs.

2.1 Non-rigid image registration

2.1.1 Transformation model

The B-spline transformation field \( g(r_R; \mu) \) maps each position \( r_R = (r^x_R, r^y_R, r^z_R) \in \mathbb{R}^3 \) in the reference image onto a position \( g(r_R; \mu) \in \mathbb{R}^3 \) in the floating image, depending on some set of transformation parameters \( \mu \).

This transformation field is defined by overlaying a regular mesh upon the reference image, with mesh spacing \( \Delta x \), \( \Delta y \) and \( \Delta z \) in the \( x \), \( y \) and \( z \) direction respectively. Each of the \( n \) knots \( ijk \) of the mesh is considered as a control point, with a 3D parameter vector \( \mu_{ijk} \) at position \( k_{ijk} = (k^x_{ijk}, k^y_{ijk}, k^z_{ijk}) \) expressing the displacement of that knot before regularization, thus \( \mu \in \mathbb{R}^{3n} \). The transformation field in each voxel is then given by the B-spline average of these knots’ displacement vectors:

\[
g(r_R; \mu_{ijk}) = \sum_{ijk} \mu_{ijk} \cdot \beta^n \left( \frac{r^x_R - k^x_{ijk}}{\Delta x} \right) \cdot \beta^n \left( \frac{r^y_R - k^y_{ijk}}{\Delta y} \right) \cdot \beta^n \left( \frac{r^z_R - k^z_{ijk}}{\Delta z} \right) \cdot \frac{1}{\Delta x \cdot \Delta y \cdot \Delta z},
\]

with \( \beta^n \) the B-spline of degree \( n \).

2.1.2 Similarity measure

The similarity measure expresses the quality of the match between the transformed floating image and the reference image, as a function of the transformation parameters \( \mu \). Maximization of mutual information (MMI) of the image intensity values of corresponding voxel pairs is a generally accepted similarity measure for image registration. Mutual information (MI) quantifies the mutual dependence of both images on each other by
measuring the distance between the current joint probability distribution of the intensities \( p(r, f; \mu) \) and the distribution associated to the case of complete independence \( p(r; \mu) \cdot p(f; \mu) \):

\[
I(R, F; \mu) = \sum_{r \in R} \sum_{f \in F} p(r, f; \mu) \cdot \log \frac{p(r, f; \mu)}{p(r; \mu) \cdot p(f; \mu)}
\]

with \( I(R, F; \mu) \) the MI between \( R \) and \( F \) given \( \mu \). \( p(x; \mu) \) the probability distribution of \( x \), given \( \mu \). The joint and marginal probability distributions can be obtained by normalization of the joint and marginal histograms of the overlapping parts of both images. The joint histogram expresses the frequency of the joint appearance of intensities \( r \) in the reference image and intensities \( f \) in the deformed floating image. It is divided in bins \( R \) for the reference image and \( F \) for the floating image. The strength of mutual information lies in the fact that it does not assume conservation of intensities between both images.

### 2.1.3 Penalties

The B-spline interpolation of the displacement vectors \( \mu \) ensures a certain degree of smoothness of the deformation field. This inherent smoothness is however not controllable and does not always suffice. For this reason, three penalties which allow to calculate properties of the deformation, independent of the image, were introduced. When a cartesian coordinate system is used and voxel size is taken into account, these penalties have a physical interpretation. The introduced penalties are:

- A smoothness penalty based upon the 3D equivalent of the 2D bending energy of a thin sheet metal:

\[
E_{Sm}(\mu) = \frac{1}{V_R} \int_R \left[ \frac{\partial^2 g(r; \mu)}{\partial x^2} \right]^2 + \left[ \frac{\partial^2 g(r; \mu)}{\partial y^2} \right]^2 + \left[ \frac{\partial^2 g(r; \mu)}{\partial z^2} \right]^2 + 2 \left[ \frac{\partial^2 g(r; \mu)}{\partial x \partial y} \right]^2 + 2 \left[ \frac{\partial^2 g(r; \mu)}{\partial y \partial z} \right]^2 + 2 \left[ \frac{\partial^2 g(r; \mu)}{\partial z \partial x} \right]^2 \, dr,
\]

with \( V_R \) the reference image volume.

- A local volume conservation penalty penalising deviations of the determinant of the Jacobian \( J_g(r; \mu) \) of the displacement function \( g(r; \mu) \) from unity:

\[
E_{Vo}(\mu) = \frac{1}{V_R} \int_R (\det(J_g(r; \mu)) - 1)^2 \, dr.
\]

- A local rigidity penalty favouring an orthogonal Jacobian matrix:

\[
E_{Ri}(\mu) = \frac{1}{V_R} \int_R \left\| J_g(r; \mu) J_g(r; \mu)^T - I \right\|_F^2 \, dr,
\]

with \( \|A\|_F^2 = \sum_{ij} A_{ij}^2 \) the Frobenius norm.

The similarity measure and penalties were combined into the overall cost function \( E(R, F; \mu) \) as a weighted sum:

\[
E(R, F; \mu) = w_I I(R, F; \mu) - w_S E_{Sm}(\mu) - w_V E_{Vo}(\mu) - w_R E_{Ri}(\mu),
\]

with weights \( w \) expressing the relative influence of each factor.
2.1.4 Optimizer

The transformation parameters $\mu$ optimizing the match between the deformed floating image and the reference image are searched for by maximizing equation (6). This was done using a multiresolution approach starting with a coarse mesh and refining it at each iteration step and using the full resolution data set at each iteration. In this work, a limited memory quasi-Newton optimizer\cite{17} was used to find the optimum $\mu$ in each stage. All derivatives with respect to $\mu$, both of the MI and of the regularization penalties, were calculated analytically.\cite{18}

2.1.5 Frame-to-frame registration

The speckles observed in two consecutive frames of a US data set are better mutually correlated than the ones in two frames separated by a larger time interval. For this reason, each frame was registered to the former one (frame-to-frame registration) and the transformation field to transform one frame into another was obtained as the cumulation of all the intermediate frame-to-frame transformation fields.

2.2 Simulated data

2.2.1 Left ventricular model

The left ventricle was geometrically approximated as a thick-walled ellipsoid with physiologically relevant end-diastolic dimensions (8.8 cm long axis, 5 cm short axis and 1 cm equatorial wall thickness) which was filled with randomly positioned point scatterers with a density of 1875 per square centimeter i.e. sufficient to generate Rayleigh scattering conditions.\cite{19} To simulate the left ventricular mechanics and hence the motion of these scatterers, a kinematical model was implemented.\cite{20,21} In this model, the deformations were determined by experimentally measured kinematical parameters. For the purpose of simplicity, only transformations based on the measured changes of the cavity volume (ejection fraction) and the torsion over a cardiac cycle were considered and used as input to the model.

2.2.2 US image data

3D pyramidal US data sets were simulated from the LV models using an adapted convolution model assuming an US system sampling at 10MHz and equipped with an array transducer centered at 2.5MHz, according to a previously described methodology.\cite{22} As realistic parameters, the simulated images consisted of 65x65 lines in azimuth and elevation direction over an angle of 60 degrees resulting in a frame rate of 20Hz when gating over four cardiac cycles. A 3D demodulated data set of 1690x65x65 voxels per time frame in spherical coordinates was thus obtained. The simulated heart rate was 60bpm. The thus obtained non-scanconverted data set was used for the validation of the non-rigid registration method to estimate 3D motion. A second data set was obtained by scanconversion of this first data set. After scanconversion it was exported in a Cartesian coordinate system resulting in a data set of 282x282x282 voxels per time frame. The voxels were isotropic with a voxel size of 0.46 mm. The outer border (containing no anatomy) of both the non-scanconverted and the scanconverted images was omitted to reduce calculation time, thus obtaining a data set of 1340x65x65x20 and one of 182x232x182x20 respectively.

2.2.3 Noise model

To simulate noise in- and outside the ventricle, a new model was created having all regions not covered by the LV-model filled with uniformly distributed random noise. US images before and after scan-conversion were obtained from this model in the same way as was done for the LV-model. These noise-images could be added to the images of the LV-model after scaling their maximal amplitude to a certain percentage of the maximum amplitude of the LV-signal. A slice through the middle of the resulting 3D end-diastolic (ED) and end-systolic (ES) data set with the maximum amplitude of the noise equal to 30% of the maximum amplitude of the LV-signal is shown in Figure 1 (a) and (b) respectively. Figures 1 (c) and (d) show the ED and ES images with 60% of added noise, before scan-conversion. In all images, the outer border has already been omitted.
Figure 1. Slice through the 3D end-diastolic (a) and end-systolic (b) frame of the simulated scanconverted data set with 30% of added noise and arbitrarily scaled slice through the end-diastolic (c) and end-systolic (d) frame of the non-scanconverted data set with 60% of added noise.

2.3 Accuracy of the motion estimation

Once each frame is registered with the former one, the displacement of each point of the myocardium from ED to ES can be obtained from the cumulative transformation field. For a dense set of myocardial points \( r_R \) in ED, the estimated displacement \( g(r_R; \mu) \) from ED to ES was compared to the true/theoretical deformation \( G(r_R) \) and the percentage error on the estimated displacement was averaged to yield a quality-measure:

\[
D[\%] = \frac{1}{n} \sum_{r_R: G(r_R) \geq 2v} \left| \frac{g(r_R; \mu) - G(r_R)}{|G(r_R)|} \right| \cdot 100, \tag{7}
\]

with \( D \) the mean percentage error on the estimated displacement, \( n \) the number of ED myocardial points for which \( G(r_R) \geq 2v \) is true and \( v \) the voxel spacing. Because the division by \( |G(r_R)| \) could cause this error measure to become very large for points with a small displacement, the points which move less than two voxel spacings were excluded from the error measure. The error measure \( D \) was always calculated in cartesian coordinates, so for the non-scanconverted data sets, the deformation vectors were calculated in spherical coordinates and then transformed to cartesian coordinates to calculate the mean percentage error on the estimated displacement.

To quantify the error in the radial, longitudinal and circumferential directions (the directions in which the strain is traditionally measured), the transformation vectors in cartesian coordinates were projected onto this cardiac coordinate system and the mean percentage error on the estimated displacement was recalculated in each of the directions separately.

3. RESULTS

The non-rigid registration method described in section 2.1 was applied to the simulated scanconverted and non-scanconverted data sets with 0, 30 and 60% of added noise. For every registration, the transformation field was defined using equation (1) with second-degree B-splines in all three dimensions. Thus, the influence of \( \mu_{ijk} \) was restricted to 1.5 inter knot distances in each direction. 32 discrete bins were used for the intensities of both the reference and floating image and the mutual information was calculated using partial volume distribution interpolation.\(^{14,18}\) The remaining registration parameters were tuned depending on the data set.

3.1 Motion estimation

3.1.1 Scanconverted data

For the scanconverted data sets, the initial grid size in each direction was taken equal to 64 voxels (or 29.5 mm) in the first iteration and reduced to 32 and 16 voxels in the next two iterations. For the weights of the penalties, several combinations were tried. The optimal parameter setting was found to be 1 for the weighing factor of the
volume penalty, 0.01 for the rigidity penalty and 100 for the smoothness penalty for the data set without added noise. The weight of the volume penalty should be decreased to 0.1 and 0.01 when background noise of 30% and 60% is added.

The resulting mean percentage error on the estimated displacement from ED to ES was calculated for the part of the LV between the apex and the equator and was found to be 7.45 ± 3.25% for the data set with 0% noise, 10.65 ± 3.26% for the one with 30% noise and 14.86 ± 5.15% for the one with 60% noise. Figure 2 shows the results for the motion estimation from the data set with 0% noise. An ED long axis slice through the middle of the data set, a short axis slice near the apex and one near the base are shown. The white arrows indicate the true motion from ED to ES for a selection of points, while the gray arrows represent the estimated motion. The white arrows are almost invisible due to the good match between both sets of arrows.

The first three columns of Table 1 list the resulting mean percentage error, as well as the mean percentage error in the radial, longitudinal and circumferential direction.

![Figure 2](image)

Figure 2. Slices from the ED-frame of the scanconverted data set without added noise. The horizontal lines indicate the relative position of the slices. The white arrows indicate the true motion from ED to ES for a selection of points, the gray arrows the estimated motion.

### 3.1.2 Non-scanconverted data

The non-scanconverted data sets are very small in the theta- and phi-direction, compared to the axial-direction. For this reason, the voxel size was adapted so that the non-scanconverted images were considered as being almost square. The initial grid size was taken equal to 16 voxels in the theta- and phi-direction and 512 voxels in the axial-direction. As such, the total number of grid points was similar to the number for the scanconverted data sets. The grid size in every direction was divided by two in both following iterations. The optimal weighing factors for the penalties were found to be 1 for the volume penalty, 0.1 for the rigidity penalty and 100 for the smoothness penalty. When noise was added to the images, the weight of the volume penalty had to be decreased to 0.1 and 0.01 for 30 and 60% of added noise respectively. The mean percentage error on the estimated displacement between the apex and equator for all non-scanconverted data sets is listed in the last three columns of Table 1.

### 3.2 Error in time and space

The percentage error averaged over each short axis slice going from apex to base is shown in Figure 3(a) for the scanconverted data sets and in Figure 3(b) for the non-scanconverted data sets. Figures 4(a) and 4(b) show the error averaged over the entire LV for the motion from ED to every time frame up to ES, for the scanconverted and non-scanconverted data sets respectively.
Table 1. Total mean percentage error on the estimated displacement ($D$) and mean percentage error in the radial ($D_R$), longitudinal ($D_L$) and circumferential ($D_C$) direction for the data sets before scan-conversion (Sfer.) and after (Cart.) with 0%, 30% and 60% of added noise. Only points below the equator with a movement of more than two voxel spacings were taken into account.

<table>
<thead>
<tr>
<th></th>
<th>Cart. 0%</th>
<th>Cart. 30%</th>
<th>Cart. 60%</th>
<th>Spher. 0%</th>
<th>Spher. 30%</th>
<th>Spher 60%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_L$</td>
<td>5.38 ± 7.55</td>
<td>5.60 ± 7.01</td>
<td>5.46 ± 4.76</td>
<td>3.49 ± 5.06</td>
<td>3.89 ± 4.62</td>
<td>4.55 ± 4.40</td>
</tr>
<tr>
<td>$D_C$</td>
<td>3.68 ± 15.85</td>
<td>2.06 ± 17.59</td>
<td>13.74 ± 22.33</td>
<td>1.47 ± 11.14</td>
<td>13.02 ± 11.70</td>
<td>7.85 ± 25.02</td>
</tr>
</tbody>
</table>

Figure 3. Percentage error on the estimated displacement from ED to ES averaged over each slice, for the scanconverted data sets (a) and the non-scanconverted data sets (b).

Figure 4. Average percentage error on the estimated displacement from ED to every time frame up to ES, for the scanconverted data sets (a) and the non-scanconverted data sets (b). Time frame 0 is equal to ED and time frame 7 to ES.

4. DISCUSSION

In this work the accuracy of spatio-temporal elastic registration to estimate 3D cardiac motion was investigated using simulated data sets before and after scan-conversion and with three levels of added noise. For each data set, the total mean percentage error on the estimated displacement was calculated, as well as the mean percentage error in the radial, longitudinal and circumferential directions. All errors were calculated including only the short axis slices between the apex and the equator, as the circumferential error increases tremendously above the equator where the direction of the torsion reverses. This is probably caused by the very low spatial resolution here at the base, as the US beams diverge. It is noted that $D$, $D_R$ and $D_L$ change only slightly when the slices at the base are included, but $D_C$ can increase up to 20%. These effects are the same for the scanconverted and non-scanconverted data sets.

Table 1 shows no significant differences between the mean percentage error on the displacement estimated from the data sets before and after scanconversion. Figures 3 and 4 show that there also are no major differences in the spatial or temporal error distribution between the scanconverted and non-scanconverted data sets. As
such, the interpolation errors introduced by scanconversion seem to have little effect on the results of the non-rigid registration method to estimate the 3D motion. The higher resolution in the axial-direction (which approximates the longitudinal direction of the left ventricle) causes the error on the estimated displacement in this direction to decrease slightly, but since a fairly isotropic deformation is searched, the higher resolution in only one direction seems irrelevant for the total displacement estimation.

To obtain the scanconverted images, one extra step in the post-processing chain is needed, but standard ultrasound systems more easily allow exportation of images after scanconversion.

5. CONCLUSION

In this study, simulated data sets representing 3D US image sequences, before and after scanconversion were used to investigate the accuracy of spatio-temporal non-rigid registration to estimate the 3D motion and deformation of the left ventricle. Although scanconversion inherently introduces interpolation errors in the images and redistributes the recorded data anisotropically, it was shown here that the 3D deformation can be estimated with a clinically acceptable and very similar accuracy for both scanconverted and non-scanconverted data sets. As such, it was shown that the scanconversion step does not necessarily have a negative influence on the non-rigid image registration method and that the easily available scanconverted images may be used for 3D motion and strain estimation.

ACKNOWLEDGMENTS

A. Elen and H.F. Choi are Research Assistants of the Research Foundation - Flanders (Belgium, FWO-Vlaanderen).

REFERENCES


