Registration of multi-view 3D echocardiography images

Harriët W. Mulder[1]
harriet@isi.uu.nl
Marijn van Stralen[1]
marijn@isi.uu.nl
Heleen B. van der Zwaan[2]
h.vanderzwaan@erasmusmc.nl
K.Y. Esther Leung[3]
e.leung@asz.nl
Johan G. Bosch[3]
j.bosch@erasmusmc.nl
Josien P.W. Pluim[1]
josien@isi.uu.nl

[1] Image Sciences Institute
University Medical Center Utrecht
Utrecht, the Netherlands
[2] Department of Cardiology
Thoraxcenter
Erasmus Medical Center
Rotterdam, the Netherlands
[3] Department of Biomedical Engineering
Thoraxcenter
Erasmus Medical Center
Rotterdam, the Netherlands

Abstract

Real-time 3D echocardiography (RT3DE) suffers from non-uniform image quality and a limited field of view. This can be improved by fusion of multiple images. Since accurate registration is essential for fusion of good quality, this study examines the performance of different methods for intrasubject registration of multi-view apical RT3DE images.

RT3DE images focused on the left ventricle were rigidly registered to images focused on the right ventricle. The performance of single-frame and multi-frame registration, that optimizes the metric for several time frames simultaneously, was examined. Furthermore, the suitability of mutual information (MI) as similarity measure was compared to normalized cross-correlation (NCC). Evaluation of the results was based on annotations made in the data by two observers.

It was found that multi-frame registration can improve registration results with respect to single-frame registration. In addition, NCC outperformed MI as metric. If NCC was optimized in a multi-frame registration strategy including end-diastolic and end-systolic time frames, the automatic method performed as good as manual registration.

1 Introduction

Real-time three-dimensional echocardiography (RT3DE) visualizes the heart in a non-invasive manner. To overcome its limitations, like non-uniform image quality and a limited field of view, registration and fusion of multiple echocardiography images can be used [7]. Accurate registration is required for this and therefore knowledge of the performance of different methods for echocardiographic registration is important.

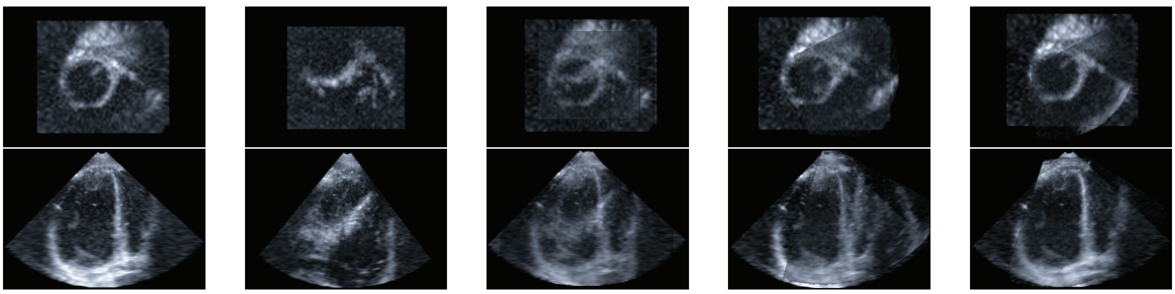
Fusion of echocardiography images is addressed by Rajpoot *et al*. [7] and Szmigielski *et al*. [9]. They used normalized cross-correlation (NCC) to register the images. To register prestress to poststress echocardiography images Leung *et al*. [5] showed the superiority of NCC over normalized MI. Mutual information (MI) was used by Shekhar *et al*. [8] for pre- to poststress RT3DE registration. In the work previously cited only end-diastolic (ED) or end-systolic (ES) time frames were registered. However, the multi-frame registration approach of Grau *et al*. [1] incorperates information available in ED and ES time frames to register RT3DE images. Nevertheless, any differences in performance between multi-frame and single-frame registration were not addressed.

We assess the performance of several methods to register multi-view RT3DE images, as presented at SPIE Medical Imaging 2011 [6]. The metrics NCC and MI, both commonly used to register echocardiography images, were examined. The performance of single-frame and multi-frame registration was compared.

2 Methods

2.1 Data

In 15 healthy volunteers two RT3DE scans were acquired using an iE33 ultrasound system (Philips Medical Systems, Best, the Netherlands) and an X3-1 matrix array transducer. The scans were acquired in harmonic mode during a single end-expiratory breath-hold. The left ventricular (LV) data set (mean volume rate: 31.6 frames per heart cycle; range: 27 to 38 frames) was taken from the standard apical view, while the right ventricular (RV) images (mean volume rate: 28.8 frames per heart cycle; range: 24 to 35) were acquired using a modified apical view. ED frames were automatically detected by the *4D RV-Function* program (TomTec Imaging Systems, Unterschleissheim, Germany), ES frames were identified by visual inspection. One of the data sets was excluded from analysis due to lack of visibility of the LV walls in the RV data. A set of LV and RV images is supplied in Figure 1.



(a) (b) (c) (d) (e)

Figure 1: RT3DE data acquired at ED. Short-axis view (top), four-chamber view (bottom). (a) LV image; (b) RV image; (c) The fused LV and RV image (alignment of image centers); (d) The fused LV and RV image after the initial transform was applied; (e) The fused LV and RV image after multi-frame registration (ED and ES time frames, using NCC as metric).

2.2 Registration

Two observers independently annotated the LV and RV data by indicating the junction of the mitral valve leaflets and mitral valve ring, as well as the position of the LV apex. Annotations were performed in ED and ES images. The method was adapted from Leung *et al.* [5].

LV and RV images acquired at the same cardiac phase in the same subject were rigidly registered in a multi-resolution strategy. To initialize the registration, a similar difference in field of view was assumed for all subjects. This was approximated by the average of manual transformations of five data sets. MI and NCC as implemented in *elastix* (Image Sciences Institute, UMC Utrecht, Utrecht, the Netherlands) [4] were examined as similarity measures. The adaptive stochastic gradient descent method [3] was used to optimize the metric. By means of single-frame registration ED or ES frames were independently registered. In contrast, multi-frame registration was performed to register different time frames simultaneously. It optimizes a cost function that is the average of the metric of all separate image pairs included in the registration process. The effect of the number of image pairs was examined by including ED and ES time frames or all available time frames. Inclusion of all time frames required linear interpolation between ED and ES as well as between ES and the last time frame of the data set to obtain the same number of LV and RV frames. Mean intensity fusion was used to fuse the registered images. All methods were implemented in *MeVisLab* (MeVis Medical Solutions, Bremen, Germany).

2.3 Evaluation

The difference between two sets of annotations, $d(\mathbf{M},\mathbf{N})$, was expressed as the average of the Euclidean distances between corresponding points, $\frac{1}{P}\sum_{j=1}^{P} d(M_j,N_j)$. Where $\mathbf{M}$ and $\mathbf{N}$ are sets of P points.

Rigid registration (in a least squares sense [2]) of the annotations in the LV and RV image resulted in the transformations $\mathbf{T}_{\mu_A}$ and $\mathbf{T}_{\mu_B}$, the subscripts A and B denote the observers and the vector μ contains the transformation parameters. To compare the manual transformations of the observers to each other, the *interobserver transformation variability* was calculated as

$$\frac{1}{S}\sum_{q=1}^{S} d(\mathbf{T}_{\mu_A^q}(\bar{\mathbf{L}}^q),\mathbf{T}_{\mu_B^q}(\bar{\mathbf{L}}^q)) \tag{1}$$

S is the number of datasets and $\bar{\mathbf{L}}$ is the vector of mean positions of the points annotated in the LV image, it equals $\frac{\mathbf{L}_A+\mathbf{L}_B}{2}$. $\mathbf{L}_A$ and $\mathbf{L}_B$ are the points annotated in the LV image by observer A or B.

The misalignment between corresponding points when the image centers are aligned was calculated as $\frac{1}{S}\sum_{q=1}^{S} d(\bar{\mathbf{L}}^q,\bar{\mathbf{R}}^q)$. After applying an initial transformation $\mathbf{T}_{\mu_{init}}$, it was defined as $\frac{1}{S}\sum_{q=1}^{S} d(\mathbf{T}_{\mu_{init}}(\bar{\mathbf{L}}^q),\bar{\mathbf{R}}^q)$. $\bar{\mathbf{R}}$ contains the RV annotations averaged over the observers.

The *registration error*, defined as the difference between an automatic transformation, $\mathbf{T}_{\mu_{auto}}$, and the transformation $\mathbf{T}_{\mu_{manual}}$ that resulted from rigidly registering the mean annotations $\bar{\mathbf{L}}$ and $\bar{\mathbf{R}}$, is given by Equation 2.

$$\frac{1}{S}\sum_{q=1}^{S} d(\mathbf{T}_{\mu_{auto}^q}(\bar{\mathbf{L}}^q),\mathbf{T}_{\mu_{manual}^q}(\bar{\mathbf{L}}^q)) \tag{2}$$

The performance of automatic registration with respect to manually registering the data was evaluated by comparing the registration error to the interobserver transformation variability. Statistical significance was assessed using a Wilcoxon signed ranks test.

3 Results

The interobserver transformation variability was 3.8 ± 1.6 mm for ED time frames and 3.3 ± 1.5 mm for ES time frames (mean ± standard deviation).

The amount of misalignment when the image centers were aligned was 37.9 ± 9.5 mm (ED) and 32.0 ± 8.4 mm (ES) and was reduced to 15.6 ± 8.1 mm (ED) and 14.7 ± 7.9 mm (ES) after transforming the images with the initial transformation.

An example of a fused ED data set is shown in Figure 1. Registration errors are given in Table 1. The error obtained by multi-frame registration (including only ED and ES time frames) using NCC or MI as metric was comparable to the interobserver transformation variability at both ED and ES. Multi-frame registration outperformed single-frame registration when only ED and ES time frames were included. In contrast, inclusion of all time frames lowered registration accuracy. Better results were obtained using NCC as metric than by use of MI. No significant differences were found between the accuracy of the different registration methods.

Table 1: Registration errors of automatic registration (mean ± standard deviation in mm). The interobserver transformation variability was 3.8 ± 1.6 mm (ED) and 3.3 ± 1.5 mm (ES).

	ED		**ES**	
	MI	**NCC**	**MI**	**NCC**
Single-frame	6.4 ± 7.2	5.1 ± 4.0	7.6 ± 11.2	2.9 ± 1.5
Multi-frame (ED & ES)	4.3 ± 2.2	3.8 ± 0.8	3.8 ± 3.7	3.0 ± 1.4
Multi-frame (all)	7.5 ± 10.1	5.9 ± 7.0	6.9 ± 11.5	5.4 ± 8.1

4 Discussion

Multi-frame registration performs better than single-frame registration when only ED and ES time frames are used to optimize the metric. Presumably, the amount of useful information is increased by involving ED as well as ES time frames in the registration process. However, inclusion of all time-frames lowers registration accuracy. Several factors might be of influence on this. First, since ED and ES time frames contain the most diverse information, little additional information will be added by inclusion of more time frames as has been remarked in literature [1]. Second, interpolation of the RV data can introduce artifacts. Third, calculation of the registration error is solely based on ED and ES time frames. Inclusion of all time frames will result in a transformation that is optimized for the whole heart cycle. Therefore, the chosen error measure might not reflect the real error well in case of a multi-frame registration strategy including all time frames.

5 Conclusion

Knowledge about the performance of different methods to register RT3DE images is important since accurate registration is essential for image fusion, a tool to improve the quality of

RT3DE data. We compared the performance of different methods to register apical multi-view RT3DE images and showed that the performance of automatic registration of these images is as good as that of manual registration. This is found for both ED and ES frames. An initial transformation that approximates the protocol-specific systematic misalignment is applied to reach the starting point for registration. Best results are obtained using NCC as similarity measure in a multi-frame registration strategy including ED and ES time frames.

6 Acknowledgements

The authors like to thank G. van Burken (Erasmus University Medical Center, Rotterdam, the Netherlands) for transferring data and software. This research is supported by the Dutch Technology Foundation STW.

References

[1] V. Grau, H. Becher, and J.A. Noble. Registration of multiview real-time 3-d echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 26(9):1154–1165, 2007.

[2] B.K.P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America a-Optics Image Science and Vision*, 4(4):629–642, 1987.

[3] S. Klein, J.P.W. Pluim, M. Staring, and M. Viergever. Adaptive stochastic gradient descent optimisation for image registration. *International Journal of Computer Vision*, 81(3):227–239, 2009.

[4] S. Klein, M. Staring, K. Murphy, M.A. Viergever, and J.P.W. Pluim. elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.

[5] K.Y.E. Leung, M. van Stralen, A. Nemes, M. M. Voormolen, G. van Burken, M.L. Geleijnse, F.J. ten Cate, J.H.C. Reiber, N. de Jong, A.F.W. van der Steen, and J.G. Bosch. Sparse registration for three-dimensional stress echocardiography. *IEEE Transactions on Medical Imaging*, 27(11):1568–1579, 2008.

[6] H.W. Mulder, M. van Stralen, H.B. van der Zwaan, K.Y.E. Leung, J.G. Bosch, and J.P.W. Pluim. Registration of multi-view apical 3d echocardiography images. In *Proceedings of SPIE Medical Imaging*, 2011.

[7] K. Rajpoot, J.A. Noble, V. Grau, C. Szmigielski, and H. Becher. Multiview rt3d echocardiography image fusion. In *Proceedings of the Functional Imaging and Modelling of the Heart*, 2011.

[8] R. Shekhar, V. Zagrodsky, M.J. Garcia, and J.D. Thomas. Registration of real-time 3-d ultrasound images of the heart for novel 3-d stress echocardiography. *IEEE Transactions on Medical Imaging*, 23(9):1141–1149, 2004.

[9] C. Szmigielski, K. Rajpoot, V. Grau, S.G. Myerson, C. Holloway, J.A. Noble, R. Kerber, and H. Becher. Real-time 3d fusion echocardiography. *JACC Cardiovasc Imaging*, 3(7): 682–90, 2010.