

Phase-Based Registration of Multi-view Real-Time Three-Dimensional Echocardiographic Sequences

Vicente Grau¹, Harald Becher², and J. Alison Noble¹

¹ Wolfson Medical Vision Laboratory, Department of Engineering Science, University of Oxford, Parks Road, OX1 3PJ Oxford, United Kingdom,
{vicente, noble}@robots.ox.ac.uk
<http://www.robots.ox.ac.uk>

² Department of Cardiology, John Radcliffe Hospital, Headley Way
Headington, OX3 9DU Oxford, United Kingdom,
harald.becher@orh.nhs.uk

Abstract. Real time three-dimensional echocardiography opens the possibility of interactive, fast three-dimensional analysis of cardiac anatomy and function. However, at the present time these capabilities cannot be fully exploited due to the low image quality associated to this modality. We propose to increase image quality and information content by combining images acquired from different echocardiographic windows. In this paper, we present an algorithm to register these datasets. Phase-based measures have been proposed as a suitable alternative to intensity-based ones for ultrasound image analysis. The proposed algorithm uses a new cost function, based on local orientation and phase differences, to align the datasets. Visual observation of results, and preliminary numerical analysis, show the robustness and accuracy of this method.

1 Introduction

Among the imaging modalities that can be used for diagnosis of cardiovascular patients, echocardiography possesses a number of advantages that make it particularly useful: it is fast, low cost, portable and compatible with implanted cardiac devices and life-support equipment. However, the low quality of echocardiographic images when compared to cardiac MR is a serious drawback when calculating quantitative measures of anatomy and function. Furthermore, intensities depend on acquisition factors such as the incidence angle of the ultrasound beam and the depth within the tissue.

Recently, real-time three-dimensional echocardiography (RT3DE) has become available commercially. There is an agreement in the cardiology field that RT3DE is superior to two-dimensional echocardiography for the quantification of left ventricular volumes and function, but in spite of these advantages RT3DE has not yet become routine in the clinical practice due to difficult image acquisition, limited image quality and laborious manual data analysis (see e.g. [1]).

Echocardiographic acquisition can only be performed from certain probe locations or windows. The availability of three-dimensional echocardiographic data allows to

combine datasets acquired from different windows. In [3], we showed the enhancement in image quality and information content that could be obtained by combining these datasets, assuming that the datasets were previously aligned. In Fig. 1, images acquired from the apical and parasternal windows, corresponding to the same heart plane, are shown. Big differences between the two images make the registration process extremely difficult if intensity-based approaches are used.

As mentioned above, the characteristics of ultrasound acquisition reduce the usefulness of intensity to identify different tissues. Previous research groups, including our own [2-4] have used phase-based algorithms with considerable success. The main advantage of phase-based algorithms is their invariance to image brightness and contrast, which makes them particularly suitable for ultrasound imaging.

In this paper, we demonstrate the feasibility of a phase-based algorithm to perform registration of RT3DE datasets acquired from the apical and parasternal windows. The user input required is limited to the selection of three landmarks on each of the datasets. The algorithm makes use of different image scales and is applied in a multiresolution framework to improve speed and robustness. Results obtained on 20 datasets from 3 subjects show a consistent, accurate behavior. Our contributions include, on the algorithmic side, the definition of a registration cost function based on local phase and orientation and calculated using the monogenic signal, and on the applications side, the proposal of a working compounding method to improve the information of RT3DE datasets by combining acquisitions from several windows.

2 Methods

2.1 Phase Calculation Using the Monogenic Signal

Several methods have been proposed for calculating phase in an N-dimensional signal. Among them, the monogenic signal [5], an isotropic extension of the analytic signal to N dimensions, provides a convenient and accurate calculation method. The monogenic signal of an N-dimensional signal is based on the Riesz transform, calculated by applying the set of N filters given, in the frequency domain, by

$$H_i(\mathbf{u}) = \frac{u_i}{|\mathbf{u}|}, \quad i=1\dots N \quad (1)$$

where $\mathbf{u}=[u_1\dots u_N]^T$, with u_i representing the i th coordinate unit vector (in 3D, $u_0=x$, $u_1=y$, $u_2=z$). To achieve frequency as well as spatial localization, a band-pass filter is applied to the signal prior to calculating the monogenic signal. In our case, we used log-Gabor filters.

The monogenic signal assigns an N+1-dimensional vector (the combination of the original signal value and the Riesz transform) to each spatial location. The N angles that define the vector can be used to estimate the N-D local signal orientation (θ) and phase (φ). Specifically, orientation and phase can be calculated as

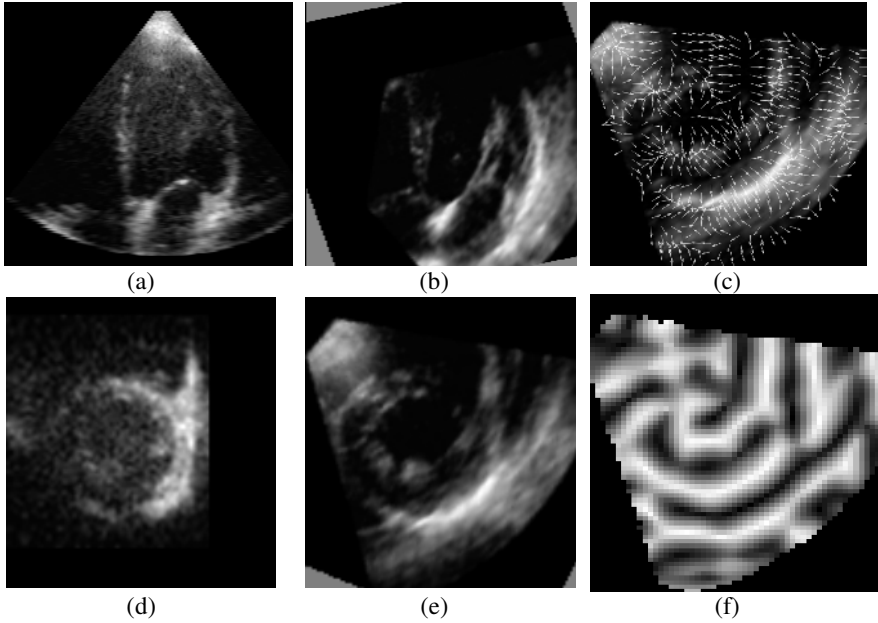


Fig. 1. Comparison of apical (a,d) and parasternal (b,e) datasets after registration. (a) and (b) correspond to long axis slices, while (d) and (e) are short axis. In (c) and (f), the orientation and phase of the short axis parasternal slice in (e) are shown. A mask corresponding to the ultrasound acquisition volume was used to avoid border effects.

$$o_i(I(x),s) = \frac{g_s(x) * h_i(x) * I(x)}{\sqrt{\sum_{i=1}^N (g_s(x) * h_i(x) * I(x))^2}} \quad , \quad i = 1 \dots N \tag{2}$$

$$\varphi(I(x),s) = \text{atan2} \left(\sqrt{\sum_{i=1}^N (g_s(x) * h_i(x) * I(x))^2}, \quad g_s(x) * I(x) \right)$$

where o_i is the i th component of the normalized orientation vector, $g_s(x)$ is the band-pass filter with central wavelength s and $h_i(x)$ are the filters from (1), in the spatial domain. Confidence in the estimation of orientation angles depends on the value of the phase. We can quantify this confidence as $\sin^2(\varphi)$ (more details are available in [5]).

Orientation and phase of the image in Fig. 1 (e), calculated after anisotropic filtering as explained in Section 2.3, are shown in the right column of Fig. 1. Even in low-contrast areas of the image, the algorithm returns a reasonable value for orientation. Notice that, at the central line of ridges, artefacts in orientation may appear. These are locations where the phase is close to zero, and thus have a low orientation confidence.

Our working hypothesis is that we can register multi-window RT3DE datasets by aligning local phase and orientation vectors calculated in (2). Sections 2.2 to 2.4 provide

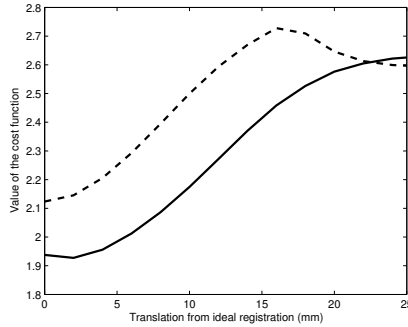


Fig. 2. Cost function values with respect to translation from the global minimum, for filters centered on wavelengths equal to 4 mm (dotted line) and 6 mm (solid line). Note that the image resolution is of the order of 1 mm. The capture range is wider in the case of the larger scale, while the lower scale achieves a higher accuracy. To obtain this figure, for each distance value d (horizontal axis), translations with norm d in different directions were calculated. The value shown in the figure corresponds to the minimum value of the cost function for each distance d .

more details about the complete registration algorithm. Section 3 shows the ability of the method to produce consistent, accurate results with limited user interaction.

2.2 Definition of a Registration Cost Function Based on Orientation Differences

The novel part of the applied algorithm is the definition of the cost function $M(T)$. We propose to use an alignment of local orientation and phase, taking into account the confidence in orientation estimation, over the cardiac cycle:

$$\begin{aligned}
 M(T) = & \sum_t \frac{x}{\sum_x W(x,t)} \cos\left(\left|\theta(I_1(x,t),s) - \theta(I_2(T(x,t)),s)\right|\right) + \dots \\
 & \dots + \sum_t \frac{x}{N_x} \frac{\sum_x \varphi(I_1(x,t),s) - \varphi(I_2(T(x,t)),s)}{N_x} \\
 W(x,t) = & \min\left(\sin^2(\varphi(I_1(x,t),s)), \sin^2(\varphi(I_2(x,t),s))\right)
 \end{aligned} \tag{3}$$

with T being the geometrical transform mapping image I_2 to image I_1 , φ and θ being phase and orientation, respectively, as introduced in Section 2.1, and $W(x,t)$ being the relative weights calculated from the orientation confidence. In (3), the cost function is estimated by accumulating the contributions of all spatial locations x and all frames t in the cardiac cycle. In general, calculation time can be reduced by using only a small set of frames. In our experiments, we used only the end-systolic and end-diastolic frames.

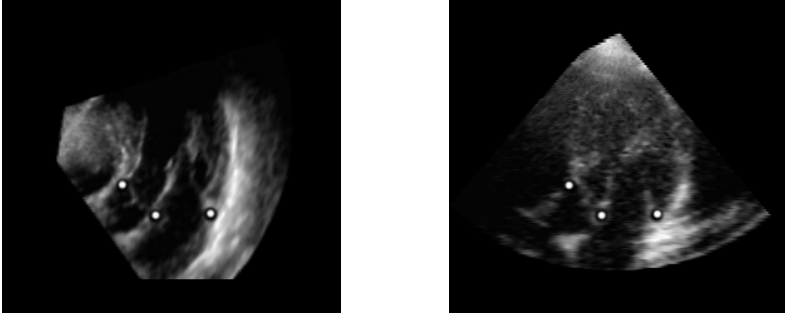


Fig. 3. Slices of the parasternal (left) and apical (right) datasets of a sample subject, with the points corresponding to the valve hinges highlighted

2.3 Registration Method

As we apply our method on multi-window datasets of the same patient, acquired with a time lapse of minutes, we do not expect a significant variation of shape and heart motion to appear between sequences. We thus use a rigid (translation + rotation) spatial transform. Regarding temporal registration, heart rate can change by a few frames between acquisitions, particularly when the subject moves to facilitate access to the scanning window. We perform this alignment by identifying the end-systolic and end-diastolic frames and performing piecewise linear temporal registration between them.

Minimization of the metric function (3) was done using Powell optimization, using only the end-systolic and end-diastolic frames in the calculation.

The influence of scale s is illustrated in Fig. 2: a larger scale enlarges the capture range of the metric and thus increases the probability of reaching the approximate location of the global minimum, while reduced scales provide a narrower minimum increasing the reproducibility of the final result when initialization landmarks vary. To combine these advantages, we embed the process in a multiscale scheme, by initially applying a large scale and using the result to initialize the process at the next, smaller scale. Furthermore, to increase robustness and decrease computation time, the whole process is carried out in a multiresolution framework. As a preprocessing step, spatial coherence-enhancing anisotropic diffusion [6] was applied.

The registration method was implemented using a combination of Matlab programs, to calculate phase and orientation, and ITK functions to perform the registration. New ITK classes were written to calculate the new cost function in (3).

2.4 Initialization

A reasonably good starting point is required to avoid the optimization method getting trapped in a local minimum. Though the cost function in (3) does not have a particularly narrow capture range (see Fig. 2), in the case of multi-window acquisition there is too much difference between the original datasets for the algorithm to converge.

Though one might calculate an approximate probe location, based on an estimation of approximate apical and parasternal positions for the general case, one would still have to estimate the approximate rotation of the probe about its axis, which in the 3D acquisitions can be modified at will by the clinician. We have thus chosen to use a landmark-based initialization. We identify three landmarks corresponding to the hinges of the mitral and aortic valves, selected at the plane crossing approximately through the centers of the mitral and aortic valve rings and parallel to the long axis of the left ventricle. In Fig. 3, the landmarks detected in a sample case are shown. Initialization is done by minimizing the sum of squared distances between landmarks.

In our experiments, the clinician could always locate the landmarks with enough accuracy. However, the complexity of the images makes this task difficult for a non-expert. Automatic initialization would be desirable and is a topic for further research.

3 Results and Conclusion

To evaluate the performance of the algorithm, 3 subjects were scanned at the John Radcliffe Hospital, Oxford, using a Philips Sonos 7500 scanner. Between 1 and 3 parasternal datasets and between 3 and 6 apical ones were obtained for each subject, changing the angle of the probe within each window to vary the information contained on each dataset. A total of 20 datasets were acquired in this way. For each dataset, the frames closest to end-systole and end-diastole were identified. Three landmarks were then selected for each one of the datasets, as described in Section 2.3.

Sample results for two of the subjects are shown in Fig. 4. It is clear that the algorithm succeeds in keeping continuity of the main structures. The full sequences are available at <http://www.robots.ox.ac.uk/~mvl/htdocs/research/gallery.html>.

Quantitative evaluation of image analysis techniques in RT3DE datasets is particularly complicated due to the difficulty in obtaining the ground truth. Using a 3D localizer during image acquisition is prone to errors caused by patient movement and displacement of the heart in the thoracic cavity due to patient breathing. Our own tests attaching the ultrasound probe to a mechanical arm (Faro Arm) showed alignment errors that were obviously larger than the ones in the results of the presented algorithm, thus making the procedure not suitable for ground truth estimation. Manual alignment of the datasets using 3D visualization software also proved impractical.

Manual segmentation of the datasets cannot be used to establish ground truth, as the quality of the original images is not sufficient to perform an exact manual delineation. As an alternative, an experienced cardiologist performed a visual assessment of the local accuracy of registration using the 16-segment model recommended by the American Society of Echocardiography to assess ventricular wall function. A diagram showing the names of the segments is shown in Fig. 5(a). Seven pairs of apical-parasternal datasets from 3 subjects, aligned using our method, were presented to the expert, who qualitatively scored the registration accuracy at each segment between 0 (worst) and 2 (best), with 2 meaning registration errors smaller than approximately 2 mm, 1 meaning errors between approximately 2 and 5 mm, and 0 bigger than approximately 5 mm (note that the resolution of the images is in the range of 1 mm).

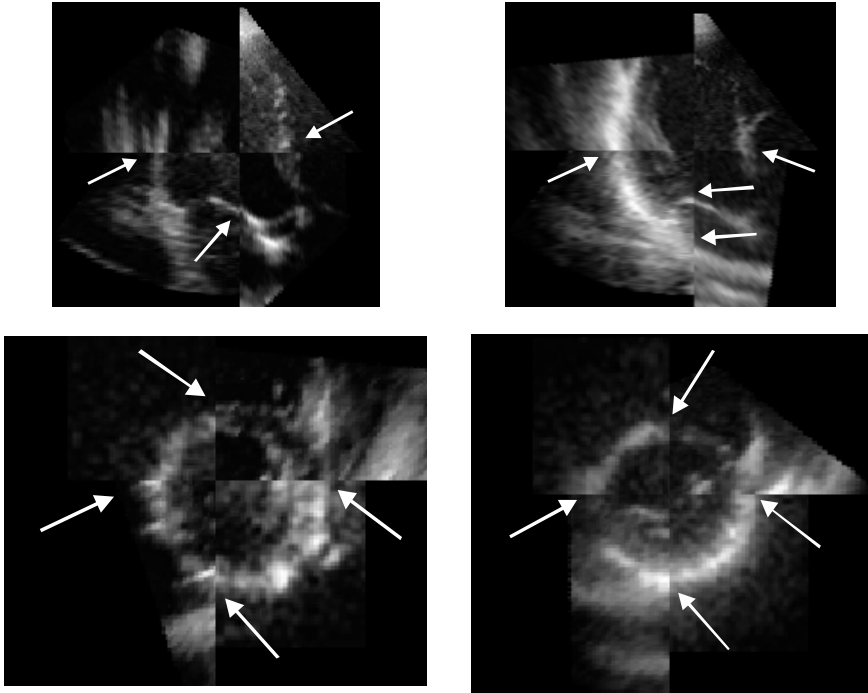


Fig. 4. Sample slices of two different subjects after alignment. Top and bottom rows show long and short axis cuts, respectively. Apical and parasternal datasets have been superimposed in a 2x2 checkerboard pattern. Notice continuity of the ventricular walls and valves at locations highlighted by the arrows.

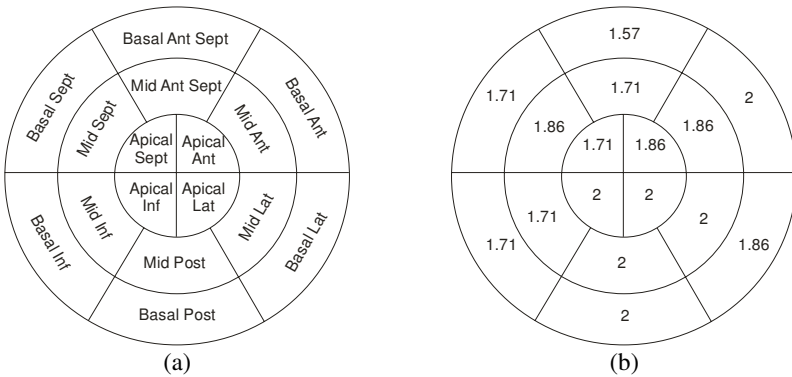


Fig. 5. Results of visual scoring of registration accuracy. (a) Names of the 16 segments in the model; (b) average score for each segment.

Average results for each segment in Fig. 5(b) show the accuracy of the method. Globally, 95 out of 112 individual segments received the maximum (2) score, 17 received a 1, and none received a 0 score. The global average score for all segments was 1.85.

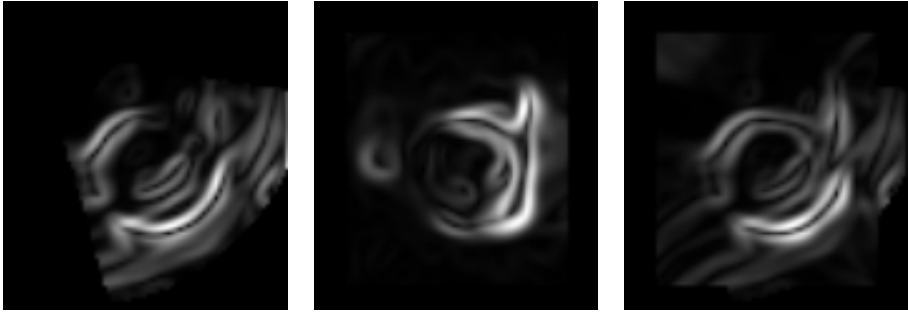


Fig. 6. Results of the Sobel filter on short axis slices from the parasternal (left) and apical windows (center), and on the combination of both (right). The completion of the circular pattern defining the left ventricle can be clearly appreciated.

The presented results, though obtained on a small sample, demonstrate the ability of the algorithm to produce accurate results for the apical-parasternal RT3DE registration problem. Fused images can be obtained using methods such as the one in [3]. In the end, the accuracy of our results will be validated by the improvement provided in real clinical tasks such as motion evaluation or volume measurements. In Fig. 6, we provide an illustration of the potential benefits of the newly generated images. The increased amount of edge information, corresponding to the endo- and epicardial surfaces, can be appreciated just by showing the result of a simple gradient estimation filter. We expect similar advantages for state of the art algorithms for motion estimation and ventricle segmentation.

Acknowledgment. The presented research was funded by the EPSRC grant GR/S72801.

References

1. Kühl, H.P. et al, "High-resolution transthoracic real-time three-dimensional echocardiography", *Journal of the American College of Cardiology* 43(11), 2083-2090, 2004.
2. Mulet-Parada, M., Noble, J.A., "2D+T Acoustic Boundary Detection in Echocardiography", *Medical Image Analysis*, 4(1):21-30 (2000)
3. Grau, V., Noble, J. A., "Adaptive Multiscale Ultrasound Compounding Using Phase Information", *MICCAI 2005*: 589-596
4. Hemmendorff, M. et al, "Phase-based multidimensional volume registration", *IEEE TMI* 21(12), 1536-1543, (2003)
5. Felsberg, Sommer, "The monogenic signal", *IEEE TSP*, 49(12):3136-3144, (2001)
6. Weickert, J., "Coherence-enhancing diffusion filtering," *International Journal of Computer Vision*, vol. 31, pp. 111-27, 1999.