Multiview RT3D Echocardiography Image Fusion

Kashif Rajpoot[1], J. Alison Noble[1], Vicente Grau[1,2],
Cezary Szmigielski[3], and Harald Becher[3]

[1] Institute of Biomedical Engineering, University of Oxford, Oxford, UK
[2] Oxford e-Research Centre, University of Oxford, Oxford, UK
[3] Department of Cardiovascular Medicine, University of Oxford, Oxford, UK
{kashif,noble}@robots.ox.ac.uk, vicente.grau@oerc.ox.ac.uk,
cezary.szmigielski@cardiov.ox.ac.uk, harald.becher@orh.nhs.uk

Abstract. Real-time three-dimensional echocardiography (RT3DE) permits the acquisition and visualization of the beating heart in 3D. However, its actual utility is limited due to missing anatomical structures and limited field-of-view (FOV). We present an automatic two-stage registration and fusion method to integrate multiple single-view RT3DE images. The registration scheme finds a rigid transformation by using a multiresolution algorithm. The fusion is based on the 3D wavelet transform, utilizing the separation of the image into low- and high-frequency wavelet subbands. The qualitative and quantitative results, from 12 subjects, demonstrate that the proposed fusion framework helps in: (i) filling-in missing anatomical information, (ii) extending the FOV, and (iii) increasing the structural information and image contrast.

Keywords: Real-time 3d, RT3DE, echocardiography, image fusion, multiview.

1 Introduction

Real-time 3D echocardiography (RT3DE) utilizes a matrix-array transducer to acquire volumetric-sequence (3D+T) images of the heart. In comparison to other 3D cardiac imaging modalities, it permits simple, near real-time, low-cost, and completely harmless image acquisition to capture the true 3D nature of the heart. The images are usually acquired by positioning the ultrasound probe at the thorax and finding the acoustic window from the apical or the parasternal views. However, due to the large-sized foot-print of the ultrasound probe, it is often a problem to find a suitable acoustic window through which the complete heart or the left-ventricle (LV) chamber can be imaged. The field-of-view (FOV) of image acquisition is, therefore, limited and cannot always capture the entire heart volume. Moreover, the images contain speckle, missing boundaries, and intensity dropout.

Medical ultrasound images are constructed by the sound reflections received at the probe from the tissue interfaces in the body. It is well known that the angle between the tissue and the ultrasound wave plays an important role in the reconstruction process. Tissue interfaces perpendicular to the ultrasound wave reflect well, while the interfaces parallel to the wave may not reflect the wave at all. Thus, the single-view apical or parasternal images are often affected by limited or missing anatomical

N. Ayache, H. Delingette, and M. Sermesant (Eds.): FIMH 2009, LNCS 5528, pp. 134–143, 2009.

information. Consequently, the use of any automated or manual quantitative analysis method on these images is prone to subjectivity and error.

We propose and investigate the use of acquiring apical view images from multiple probe positions at different angles (called multi-view) and fusing them together to: (i) fill-in missing information, (ii) extend the FOV of echocardiography images, and (iii) improve the image quality. The paper starts by providing a background about general and echocardiographic image fusion and presenting a summary of related work in Section 2. The registration and fusion algorithms are described in Section 3. The fusion results are evaluated in Section 4, with concluding remarks in Section 5.

2 Background and Related Work

In general, image fusion is the process of integrating several aligned source images to form a single fused image, preserving the salient features from all the source images. Image fusion is of great significance to combine images of varying focus, multiple sensor acquisition in remote sensing, and images of differing modalities in medical imaging [1].

In echocardiography, multiview RT3DE image fusion can be defined as the process of combining two or more single-view source images to form a single multi-view image which integrates the structural information from all the source images. Let $I_1(x,y,z),\cdots,I_N(x,y,z)$ be N aligned 3D single-view image frames, the fused image $I(x,y,z)$ is generated by applying a fusion rule θ on these images.

$$I(x,y,z) = \theta\big(I_1(x,y,z),\cdots,I_N(x,y,z)\big)$$

The fusion rule θ may take the form of minimum, maximum, average, or any other linear or non-linear function of the source intensities. The fusion rule θ decides how the values from source images are put together to form a single fused image.

Multiview RT3DE image fusion is related to two other common approaches found in the literature for integrating multiple acquisitions of ultrasound images: namely, spatial compounding and mosaicing. In the ultrasound literature, compounding is often employed to combine multiple ultrasound images having large overlap, with little or no angular difference in the acquisitions, to achieve the objective of improving the visual quality of the compounded image [2]. On the other hand, mosaicing [3] attempts to stitch together multiple images having little overlap, acquired with large angular separation in acquisitions, with the purpose of increasing the image information by extending the FOV. RT3DE image fusion can be considered as a combination of both compounding and mosaicing since its objectives include extending the FOV, increasing the image information, and improving the image quality. Moreover, in the fusion case, the multiple single-view images can have small or large overlap between them due to the arbitrary nature of probe movement during the image acquisition.

In Soler et al. [4], the authors employed a manual LV surface extraction procedure to assist in the single-view RT3DE image registration process. However, the LV surface extraction is a complex, time-consuming and subjective process due to the missing boundaries in the image. Moreover, the developed fusion methodology is restricted for fusion of only 2 source single-view RT3DE images. Grau & Noble [5] presented a local phase-based method for the fusion of apical and parasternal view

RT3DE images. Yao & Penney [6] extended the local phase-based compounding method of Grau & Noble [5] for up to 9 images, in an ad-hoc approach which is extremely computationally intensive. They used the local-phase based registration of Grau et al. [7] for aligning the images. In all these cases, the fusion technique is validated on only one or two subject cases. Szmigielski et al. [8] showed in a clinical work that the averaging based fusion technique significantly improves the quality and completeness of the RT3DE images.

3 RT3DE Image Fusion

We present a two-stage method for multiview RT3DE image fusion. During the first stage, the N images are aligned together using a standard multiresolution-based rigid registration algorithm. In the second stage, we develop a novel wavelet based image fusion technique to combine the aligned images. The proposed fusion approach is simple, fast and straightforward to fuse any number of single-view RT3DE images.

3.1 Automatic Rigid Image Registration

Registration aims to bring the reference volume I_{r} and floating volumes ${I_f}^i$ $(i = 2, \ldots N)$ into alignment with each other. We assume that a rigid transformation is sufficient to align I_{r} and ${I_f}^i$, as these images were acquired with ECG-gating, small probe movements, and over a gap of few minutes. There may be minor non-rigid misalignments due to the non-rigid motion of the heart, but finding a non-rigid transformation can introduce spurious deformations in the alignment process.

We followed a voxel-based registration process that attempts to find a geometrical transformation T such that the transformed floating image looks similar to the reference image by maximizing a similarity criterion C,

$$C = S\left(I_{\mathrm{r}}, T\left({I_f}^i\right)\right)$$

where S is a similarity function that measures the correspondence between the reference image and the floating image. In this work, the similarity function S was computed only for the image voxels region (i.e., excluding the region outside of acquisition pyramid in an echocardiography image). Since the images are acquired using the same modality and similar machine settings but they can have different intensity ranges, we employed the normalized cross-correlation (NCC) as the similarity function S. We followed a multiresolution registration approach with multiple initializations to avoid the local maxima of the similarity function. This ensures faster and robust estimation of the transformation T. The registration approach works through the following steps: (i) construct a 3-level multiresolution Gaussian pyramid, (ii) compute the similarity function S between the overlapping region of the reference and floating volumes, (iii) determine the transformation T at the coarsest level by maximizing the similarity criterion C, (iv) propagate T to the next finer level and refine the transformation T, and (v) iterate steps (i-iv) until it reaches the finest level on the multiresolution pyramid. For each resolution level in step (i), the image is blurred with a Gaussian kernel of variance $(\mathrm{shrink\_factor}/2)^2$.

Maximization of the similarity criterion C was performed using the Powell-optimizer and tri-linear interpolation was used to approximate the transformed voxel values. The registration algorithm is applied to determine the transformation T between the end-diastolic frames of I_1 and I_i images. The spatial transformation T is then used to align the rest of the respective frames between the reference and floating sequences. A couple of example results are shown in Fig. 1, demonstrating visually the accuracy of the alignment.

3.2 Wavelets for Image Fusion

Wavelets [9] offer a simple and elegant framework of simultaneous space-frequency visualization, multiresolution analysis, and subband decomposition – permitting the separation of an image into low- and high-frequency components. The wavelet decomposition highlights interesting features in the wavelet domain which are often not visible in the spatial domain. In the 3D discrete wavelet transform (DWT), an image is divided into 8 subbands – 1 low-frequency and 7 high-frequency subbands.

We developed a wavelet based fusion scheme that is simple, efficient, fast, and extendable to fuse any number of source single-view images. Once the images are registered, the wavelet based fusion process combines the aligned single-view image volumes into a single multi-view volume that blends together the information from all the source images. The partition of an echocardiography image using wavelets into low- and high-frequency components is significant for the fusion process. In this way, we can employ suitable fusion rules for these components according to their characteristics in the image.

1. We observed in our experiments that the blood-pool speckle and the borders of the image pyramid region appear in the high-frequency wavelet components[1]. To suppress the noise[2] while still keeping the speckly appearance in the fused image similar to the original images, we average the high-frequency components.
2. On the other hand, low-frequency components are related to the more homogeneous regions of the image (e.g., myocardium). Since these regions appear brighter on ultrasound images, we choose maximum of the low-frequency components. This allows the fusion process to effectively fill-in the missing information.

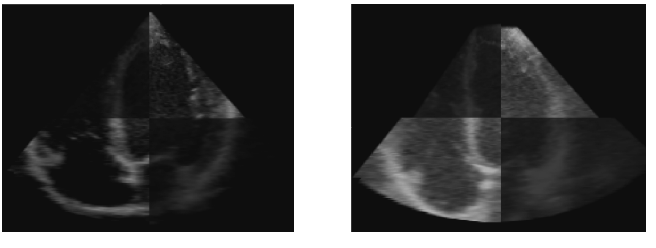


Fig. 1. Registration examples show the consistent accuracy of the registration method

[1] Note that, strictly speaking, speckle does not always have high-frequency characteristics.
[2] In this work, we define noise as highly inhomogeneous neighboring intensities.

These observations are verified by the illustration in Fig. 2, where we show the low- and high-frequency information contents of an example 2D echocardiography image slice. The 2D image, Fig. 2(a), is decomposed into low- and high-frequency wavelet components. The wavelet reconstruction from only low-frequency components is shown in Fig. 2(b), by zeroing out the high-frequency wavelet components. Fig. 2(c) depicts the wavelet reconstruction from only high-frequency wavelet components, by zeroing out the low-frequency wavelet components.

We propose the wavelet based fusion mechanism in the following 4 steps:

1. perform 1-level forward wavelet transform W^F on each single-view image I_i to separate the image into low-frequency Wav_L^i and high-frequency Wav_H^i wavelet coefficients,

$$W^F(I_i) = \left[Wav_L^i, Wav_H^i\right]$$

2. apply the maximum fusion rule θ_{max} on low-frequency coefficients Wav_L^i to obtain the fused low-frequency wavelet coefficients,

$$Wav_L^f = \theta_{max}(Wav_L^1, \cdots, Wav_L^N)$$

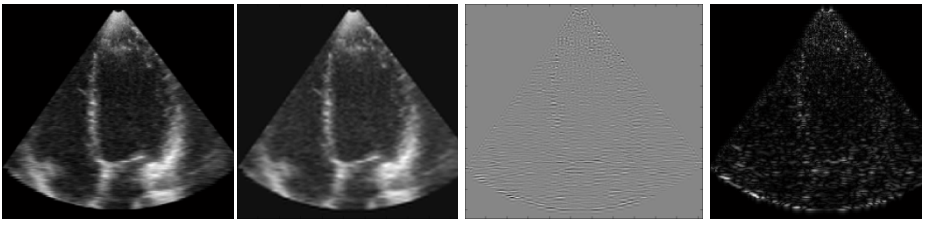
3. apply the mean fusion rule θ_{mean} on high-frequency coefficients Wav_H^i to obtain the fused high-frequency wavelet coefficients,

$$Wav_H^f = \theta_{mean}(Wav_H^1, \cdots, Wav_H^N)$$

4. perform inverse wavelet transform W^I from fused wavelet coefficients (Wav_L^f and Wav_H^f) to re-construct the fused image I.

$$I(x, y, z) = W^I\left[Wav_L^f, Wav_H^f\right]$$

This complete process of wavelet based fusion is graphically demonstrated in Fig. 3. We highlight the fact that the above proposed fusion strategy for the low- and high-frequency wavelet coefficients is contrary to the conventionally used wavelet based fusion of ordinary (usually non-medical) images [10]. In the conventional approach, the maximum of the high-frequency wavelet coefficients magnitude is selected to preserve the edges with the assumption that the edges lie in the high-frequency



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

Fig. 2. Low- and high-frequency characteristics of an echocardiography image. (a) An example 2D echocardiography image, (b) low-frequency wavelet reconstruction, (c) high-frequency wavelet reconstruction, and (d) absolute values of (c). Due to poor print quality on paper, the information of these images will be better visible in the electronic version.

components. However, this assumption is invalid for the echocardiography images which, unlike natural images, do not contain any sharp edges in the high-frequency wavelet coefficients (see Fig. 2(c) and Fig. 2(d)). Thus, the conventional approach does not yield the desired results in our application.

4 Experimental Validation

4.1 Data

Volumetric images were obtained using the Philips iE33 machine (Philips Medical Systems, Andover, USA) with a matrix-array transducer, acquiring a full volume sequence using ECG-triggering over 4 heart-beats. The usual spatial dimensions are 224x208x208 with nearly isotropic resolution of 0.82x0.84x0.76mm^3. To assess the fusion performance, the proposed method was applied on images from 12 healthy young subjects, with the number of source single-view images for each subject in the range of 3 to 6. The single-view images were recorded from different transducer positions using the following protocol. The first full-volume sequence was acquired by placing the transducer probe at the LV apex. This image was used as the reference volume for the image registration stage in 3.1. Two more full-volume sequences were acquired by translating the probe from the apex towards the lateral wall of the LV by approximately 1cm and 2cm, respectively. Another full-volume sequence was acquired by translating the probe from the apex towards the interventricular septum by approximately 1cm. Finally, two more volumes were captured by moving the probe one inter-costal space above and below the optimal apical position. We used 3 to 6 single-view images for some subjects as it was not possible to acquire all the 6 single-view images using the above protocol due to poor acoustic window.

4.2 Validation Measures

In order to evaluate the wavelet based fusion method, we computed four quantitative validation measures as defined below.

1. The % improvement in global image contrast, $\Delta(contrast)$, due to fusion is,

$$\Delta(contrast) = \left[\frac{\overline{\rho_f^{MY}} - \overline{\rho_f^{BP}}}{\frac{1}{N}\sum_{i=1}^{N}\left(\overline{\rho_i^{MY}} - \overline{\rho_i^{BP}}\right)} - 1 \right] * 100$$

where $\overline{\rho^{MY}}$ and $\overline{\rho^{BP}}$ denote the mean value in the myocardium and the blood-pool, respectively; N represents the total number of source single-view images; and i or f represent the i^{th} or fused image, respectively.

2. The improvement in SNR, $\Delta(SNR)$, in the fused image is defined as,

$$\Delta(SNR) = 20 * log\left(\frac{\frac{1}{N}\sum_{i=1}^{N}\mu_i^{BP}}{\mu_f^{BP}}\right) + 20 * log\left(\frac{\frac{1}{N}\sum_{i=1}^{N}\mu_i^{MY}}{\mu_f^{MY}}\right)$$

where μ^{MY} and μ^{BP} denote the variance in the myocardium and the blood-pool, respectively.

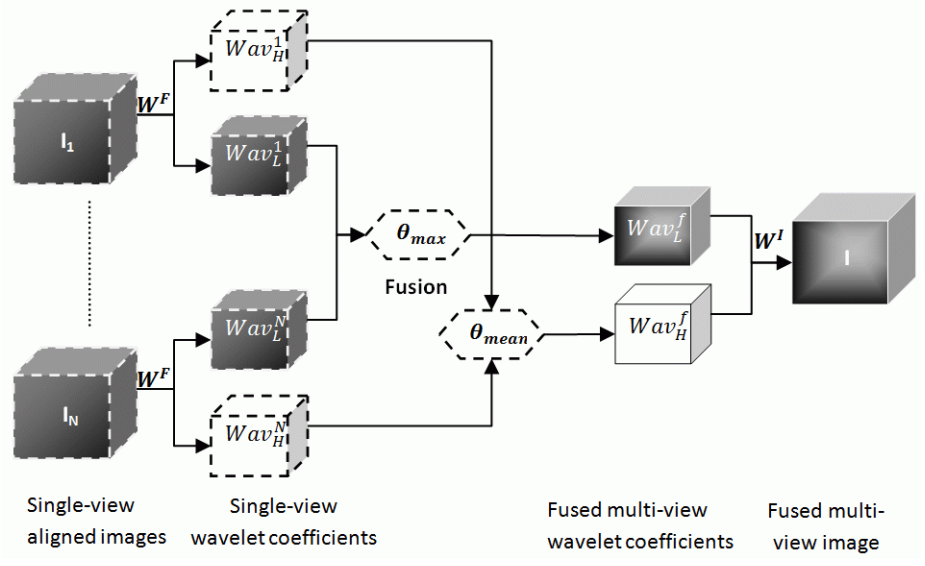


Fig. 3. Wavelet based multiview RT3DE image fusion scheme. For simplicity, only 1 high-frequency subband (Wav_H) is shown; in practice, there are 7 high-frequency subbands.

3. The quantitative change in the image's structural information is defined as the increase in the amount of total edge or feature points. The % change in feature information, $\Delta(feature)$, is computed as following,

$$\Delta(feature) = \left[\frac{nF_f}{\frac{1}{N}\sum_{i=1}^{N} nF_i} - 1\right] * 100$$

where nF represents the total number of feature points in the image. The feature map is computed using the feature detector in [11], derived from local-phase based feature asymmetry measure to detect step-like low contrast endocardial or epicardial features. These features are vital due to their potential utility in left ventricle segmentation using an active contour based algorithm, as in [12].

4. The % change in FOV, $\Delta(fov)$, due to the fusion is computed as,

$$\Delta(fov) = \left[\frac{FOV_f}{\frac{1}{N}\sum_{i=1}^{N} FOV_i} - 1\right] * 100$$

$$FOV = \sum_{i=1}^{n} vox_i$$

where vox_i denote the image voxels (1 for foreground, and 0 for background region outside the image pyramid), and n indicates the total number of voxels. The $\Delta(fov)$ is independent of the fusion technique and is a factor of image acquisition geometry.

4.3 Results and Evaluation

To assess the performance of the proposed fusion method, we implemented two other common fusion methods that use either averaging or maximum as their decision rule, respectively. The averaging fusion rule is popular in ultrasound compounding literature because of its ability to suppress noise and improve the signal-to-noise (SNR) ratio. On the other hand, the maximum fusion rule is particularly good for ultrasound images, since ultrasound images have high intensity values corresponding to important anatomical structures (e.g., myocardium). The three methods are denoted as: wavelet fusion – WAV, maximum fusion – MAX, and averaging fusion – AVG. For the computation of $\Delta(feature)$ and $\Delta(fov)$, the whole 3D volumes were used. However, in the case of computation of $\Delta(contrast)$ and $\Delta(SNR)$, representative 2D slices (apical 4-chamber plane) were used for the interactive selection of region-of-interests (ROIs) in the myocardium tissue and the blood-pool cavity (three arbitrary sized rectangle ROIs each in the myocardium and the blood-pool cavity). It was assumed that these ROIs represent the whole volume image for the computation of $\Delta(contrast)$ and $\Delta(SNR)$ measures.

Table 1 shows the mean values, over 12 images, for $\Delta(feature)$, $\Delta(contrast)$, and $\Delta(SNR)$. It shows significant contrast improvement gain with WAV (24.07%) in comparison to the popular averaging based compounding method AVG (– 5.32%). The SNR improvement with AVG ($8.01dB$) is slightly higher than WAV ($7.21dB$), because AVG involves averaging operation which achieves significantly more smoothing resulting in relatively higher SNR improvement. However, the SNR improvement with WAV is significant and comparable to AVG. The mean $\Delta(fov)$ value for 12 images is 34.15% and is independent of the image fusion technique. The amount of FOV enlargement is large, considering that an individual single-view image can already capture a big part of the heart in 3D images. The feature improvement with WAV (15.85%) and MAX (16.25%) is considerably higher than AVG (8.99%), due to the better preservation of important structures like myocardium tissue. We attribute only part of this feature improvement to be due to FOV enlargement, while the remaining is due to filling-in of missing anatomical information in the multiview combination. The above comparison demonstrates that the wavelet based fusion outperforms both of the other methods when analyzed collectively over all the quantitative evaluation measures.

Fig. 4 shows the visual results of RT3DE image fusion for a representative 2D slice. The visual results indicate that the fusion process significantly increases the structural information in the image and that every additional source single-view image, even if acquired from small probe movements, brings in some complementary feature information. The visual results show that the WAV method blends the source single-view images very well and has no stitching artifacts around the overlapping image borders as are visible on AVG fusion results. Moreover, the fusion increases the FOV, contrast and SNR of the image. Another notable gain of fusion is that it increases the density of the sampling resolution due to multiview combination. This is clearly visible on myocardium muscle tissue, papillary muscles and valves which appear thicker and have more uniform appearance after fusion.

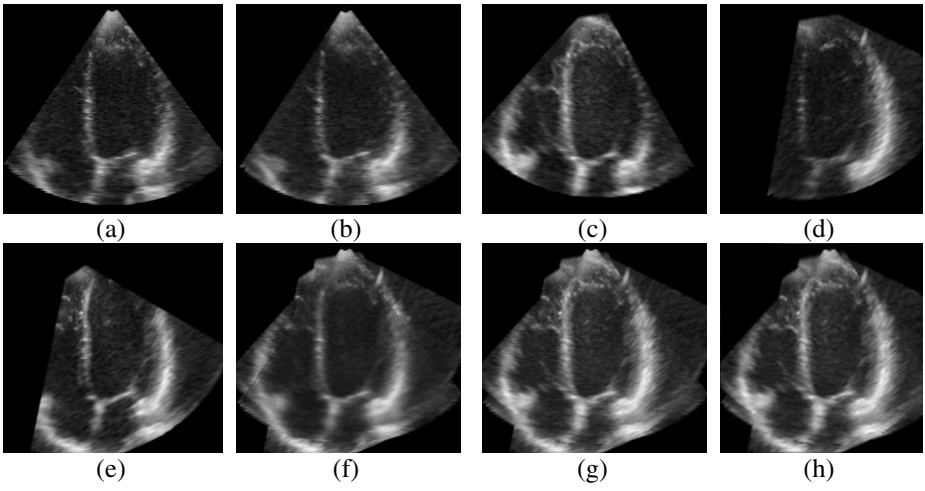


Fig. 4. Image fusion example, shows representative 2D slices. (a) – (e) 5 single-view 4-chamber images; (f) AVG fusion, (g) MAX fusion, and (h) WAV fusion.

Table 1. Comparison of AVG, MAX, and WAV fusion methods

	$\Delta(feature)$ (%)	$\Delta(contrast)$ (%)	$\Delta(SNR)$ (dB)
AVG	8.99	– 5.32	8.01
MAX	16.25	15.06	4.31
WAV	15.85	24.07	7.21

5 Conclusions

This work presented a completely automatic two stage registration and fusion method to integrate multiple single-view RT3DE apical-view images. The images were aligned using automatic multiresolution-based rigid registration algorithm, followed by wavelet based fusion method that allows the decomposition of the aligned images into low- and high-frequency components, thus permitting the use of relevant fusion rules. The quantitative validation of wavelet based fusion, against averaging and maximum based fusion methods, showed that it performs well to increase features, contrast, and SNR in the fused image.

Acknowledgement

Kashif Rajpoot is funded by a Pakistan Higher Education Commission PhD scholarship.

References

1. Goshtasby, A., Nikolov, S.: Image Fusion: Advances in the state of the art. Information Fusion, 114–118 (2007)
2. Rohling, R.N., Gee, A., Berman, L.: Three-Dimensional Spatial Compounding of Ultrasound Images. Medical Image Analysis 1, 177–193 (1996)

3. Wachinger, C., Wein, W., Navab, N.: Three-Dimensional Ultrasound Mosaicing. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 327–335. Springer, Heidelberg (2007)
4. Soler, P., Gerard, O., Allain, P., Saloux, E., Angelini, E., Bloch, I.: Comparison of Fusion Techniques for 3D+T Echocardiography Acquisitions from Different Acoustic Windows. Computers in Cardiology, 141–144 (2005)
5. Grau, V., Noble, A.: Adaptive Multiscale Ultrasound Compounding using Phase Information. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3749, pp. 589–596. Springer, Heidelberg (2005)
6. Yao, C., Penney, G.: Spatial Compounding of 3D Echocardiography: Novel Methodologies for Large Sets of Images. In: Medical Image Understanding & Analsysis (MIUA), Dundee, Scotland (2008)
7. Grau, V., Becher, H., Noble, J.A.: Registration of Multiview Real-Time 3-D Echocardiographic Sequences. IEEE Transactions on Medical Imaging 26, 1154–1165 (2007)
8. Szmigielski, C., Grau, V., Hoeschen, M., Noble, J.A., Kerber, R., Becher, H.: New Fusion Technique Significantly Improves Image Quality and Completeness of Datasets in RT3D Echocardiography. In: American Heart Association Scientific Sessions, II_401, Orlando, USA (2007)
9. Mallat, S.: A Wavelet Tour of Signal Processing. Academic Press, USA(1998)
10. Nikolov, S., Hill, P., Bull, D., Canagarajah, N.: Wavelets for Image Fusion. In: Petrosian, A.A., Meyer, F.G. (eds.) Wavelets in Signal & Image Analysis. Kluwer Academic Publishers, Dordrecht (2001)
11. Rajpoot, K., Noble, A., Grau, V., Rajpoot, N.: Feature Detection from Echocardiographic Images using Local Phase Information. In: Medical Image Understanding and Analysis (MIUA), Dundee, UK (2008), http://www.robots.ox.ac.uk/~kashif/pubs/miua2008kmr.pdf
12. Zagrodsky, V., Walimbe, V., Castro-Pareja, C., Qin, J.X., Song, J.-M., Shekhar, R.: Registration-Assisted Segmentation of Real-Time 3-D Echocardiographic Data Using Deformable Models. IEEE Transactions on Medical Imaging 24, 1089–1099 (2005)