

# Respiratory Motion Correction for 2D Cine Cardiac MR Images using Probabilistic Edge Maps

O Oktay, G Tarroni, W Bai, A de Marvao, D O'Regan, S Cook, D Rueckert

Imperial College London, London, UK

## Abstract

*2D short axis (SA) cine MR images are of crucial importance for the accurate assessment of cardiac anatomy and function. Since SA cine stacks are routinely acquired during multiple breath-holds, different breath-hold positions can cause a misalignment of the heart between different slices, with potential detrimental effects on a variety of clinically relevant measurements (e.g. volume or shape of the left ventricle). In this study, we propose a novel approach to spatially align motion corrupted SA slices in MR image stacks using 3D probabilistic edge maps (PEMs) generated with structured decision forests. In our technique, each 2D SA slice is associated with a 3D PEM outlining the myocardial contours in the same slice as well as in the adjacent one. In-plane spatial misalignment between adjacent slices is then corrected using a registration algorithm applied to the associated PEMs. This approach was tested against a conventional intensity-based registration method on SA cine stacks acquired from 26 healthy subjects, for whom anatomical 3D cardiac images were also available as reference. End-diastolic left ventricular volumes are estimated using a 3D multi-atlas segmentation technique and used to quantify alignment accuracy. The results show that the proposed technique successfully reduces the misalignment between slices and that the registered stacks allow a more accurate volumetric estimation than both the original and the intensity-corrected ones.*

## 1. Introduction

Cardiovascular magnetic resonance (CMR) imagery can be used for accurate volumetric reconstruction of the beating heart throughout the cardiac cycle, which is in turn necessary for a great number of clinically-relevant tasks such as volumetric estimations, wall motion analysis and inter-modality registration [1]. While fast SSFP sequences allow nowadays the direct acquisition of an anatomical 3D image (A3D) of the whole heart, they are usually limited by either relatively poor image quality or low temporal resolution, making them unsuitable for accurate functional

assessment. Consequently, the most common CMR sequence currently used in the clinical practice is still the short axis (SA) SSFP cine, consisting of 10-14 parallel slices and 20-30 frames per cardiac cycle. SA cine stacks are generated during multiple breath-holds (i.e. 1-3 slices acquired per each breath-hold). Although the subjects are instructed to hold their breath at the same breath-hold position, in practice the heart location can vary considerably. If the differences between the breath-hold positions are not accounted for with suitable inter-slice alignment procedures, the acquired image stack will not correctly represent the cardiac volume, introducing potential errors in all of the following analyses and visualizations.

Several post-processing techniques for respiratory motion correction for SA cine images have been presented in the last years. Virtually all of the proposed approaches consist in the 3D rigid registration between the acquired SA cine stack and other CMR images, such as A3D [1,2] or long axis (LA) [3,4] images. Unfortunately, in the clinical practice these additional sequences are often unavailable. Therefore, inter-slice alignment can only be performed by conventional intensity-based approaches, which are on the other hand relatively inaccurate. The reason for this mainly lies in the considerable changes in the imaged features between adjacent slices (due to the relatively high slice thickness of a typical SA cine stack), which hinders the application of registration approaches based on the comparison of information directly extracted from the two slices.

In this paper, we propose a novel approach to in-plane spatial alignment of motion corrupted SA slices in MR image stacks using 3D probabilistic edge maps (PEMs) generated with structured decision forests. PEMs are learning-based image representations outlining the contours of a specific object of interest (the myocardium, in this case). Through training, structured decision forests are able to associate to each 2D SA image for slice  $i$  a 3D PEM representing the contours at slices  $i$  and  $i+1$ . The overlap between 3D PEMs generated from adjacent slices allows to correct the in-plane spatial misalignment using a block matching registration algorithm. The proposed approach was tested on a dataset of real SA cine stacks against a conventional intensity-based registration method.

## 2. Methods

### 2.1. Image Analysis

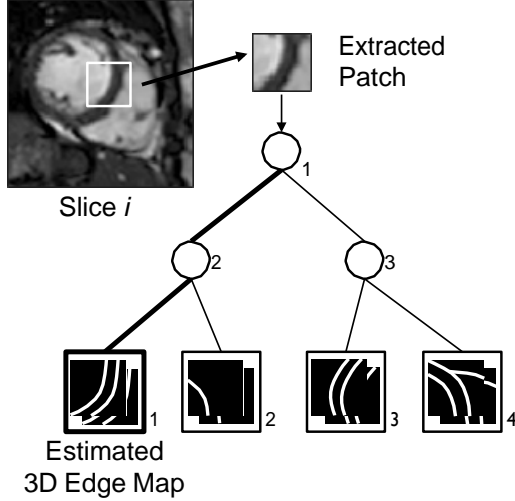


Figure 1. Simplified graph of the structured decision tree. In the proposed approach, each image patch extracted from slice  $i$  is associated with a 3D edge map outlining the myocardial contours on slices  $i$  and  $i+1$ .

Structured decision forests can be considered as extensions of more traditional classification forests [5]. In standard classification forests, each decision tree classifies a sample  $\mathbf{x} \in X$  by recursively branching left or right at each split node  $j$  (circles in Fig. 1) until a leaf node  $k$  (squares in Fig. 1) is reached. Each leaf node is associated with a label  $y \in Y \subset Z$  or a distribution over the labels. Each split node  $j$  is associated with a binary split function  $h(\mathbf{x}, \theta_j) \in \{0, 1\}$ , defined by the set of parameters  $\theta_j$ : if the outcome is 0 the node sends  $x$  to the left, otherwise to the right. In most cases,  $h$  is a decision *stump*, i.e. a single feature dimension  $n$  of  $\mathbf{x}$  is compared with a threshold  $\tau$ :  $\theta = (n, \tau)$  and  $h(\mathbf{x}, \theta) = [\mathbf{x}(n) < \tau]$ . A classification forest is an ensemble of  $T$  independent decision trees: during testing, given a sample  $\mathbf{x}$ , the predictions of the different trees are combined into a single output using an ensemble model (e.g. based on majority voting). During training, the goal at each node is to find the set of parameters  $\theta_j$  which maximize a previously defined *information gain*  $I_j$ . A standard definition for  $I_j$  is

$$I_j = H(S_j) - \sum_{i \in \{0,1\}} \frac{|S_j^i|}{|S_j|} H(S_j^i), \quad (1)$$

where  $S_j$ ,  $S_j^0$  and  $S_j^1$  are respectively the training set (comprising of samples  $\mathbf{x}$  and associated labels  $y$ ) arriving at node  $j$ , leaving the node to the left and leaving the

node to the right, and where  $H(S)$  can be defined as the Shannon entropy [5].

Structured decision forests extend the concept of classification forests by using structured output spaces for  $Y$  in lieu of integer labels. For the particular task at hand, the samples  $\mathbf{x} \in X$  represent image patches, while the labels  $y \in Y$  represent edge maps (see. Fig. 1). In order to train a structured decision tree, it is necessary to find a way to cluster structured labels at each split node into two subgroups depending on some similarity measure between them. This can be done in two steps [5]. First,  $Y$  is mapped to an intermediate space  $Z$  in which a distance between maps can be computed: more specifically, each edge map  $y$  is associated with a binary vector  $\mathbf{z}$  encoding whether every pair of pixels belong to a simply-connected region in the map or not. Then, at each split node the edge maps  $y$  are mapped into a binary set of labels  $c \in C = \{0, 1\}$ : this is achieved performing PCA on the associated  $\mathbf{z}$  vectors followed by binary quantization applied to the principal component. This way, at each split node similar edge maps will be assigned to the same binary label  $c$ . Importantly, thanks to this discretization process the information gain defined in (1) can be directly applied. On a different note, the required input features are computed in a multi-scale fashion and consist of intensity values, gradient magnitudes and HoG-like channels [5]. During testing, the edge maps predicted by each decision tree are averaged to provide the estimate for the input image patch  $\mathbf{x}$ . Edge maps estimated for overlapping image patches are also averaged, yielding to the final probabilistic representation of the edge map (PEM) for the input image.

In our technique, the described approach to edge representation was applied to associate to each 2D SA image for slice  $i$  a 3D PEM outlining the myocardial contours at slices  $i$  and  $i+1$  (see Fig. 2). (The PEM at slice  $i-1$  could be also predicted, but this would be at the expense of the prediction accuracy as the dimensionality of output space would be increased.) To build a model with motion-free data, the training dataset consisted of slices extracted from the A3D acquisitions (downsampled in the  $z$  direction to match the slice thickness of the 2D cine data later used for testing) together with myocardial contours (the edge maps used as labels). During testing, the structured random forest predicts for each 2D SA image at slice  $i$  not only the myocardial contours location for the same slice but also for the following one ( $i+1$ ) in a motion-free scenario. The overlap between 3D PEMs estimated for adjacent slices of the stack allows to apply a block matching registration algorithm [6] to perform rigid in-plane spatial re-alignment. The rigid transformation estimated for the PEMs is then applied to the actual image at slice  $i+1$ . Each slice (from second to last) undergoes the same process, resulting in the complete re-alignment of the stack from base to apex.

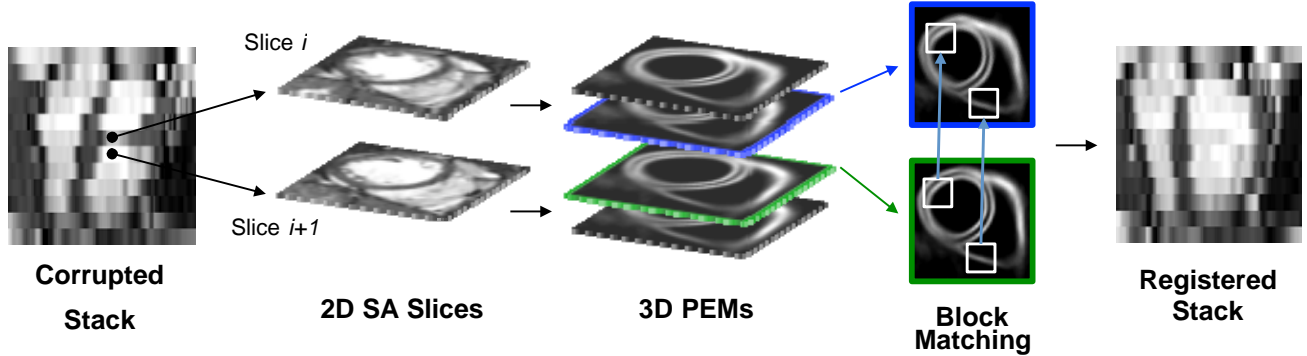


Figure 2. Pipeline of the proposed motion correction technique: 3D PEMs are estimated from each pair of adjacent SA slices and then aligned using a block matching algorithm. The estimated transformation is then applied to the actual slice.

## 2.2. Image Acquisition

CMR imaging was performed on 435 healthy subjects using a 1.5T Philips Achieva system (Best, Netherlands) equipped with a 32 element cardiac phased-array coil. The maximum gradient strength was 33 mT/m and the maximum slew rate 160 mT/m/ms. 2D cine balanced steady-state free precession (b-SSFP) images were acquired in the left ventricular short axis (LVSA) plane from base to apex using the following parameters: repetition time/echo time 3.0/1.5 ms, flip angle  $60^\circ$ , acquired pixel size  $2.0 \times 2.2$  mm, slice thickness 8 mm (2 mm gap); reconstructed voxel size  $1.2 \times 1.2 \times 8$  mm; number of slices 10-12, cardiac phases 30. A single breath-hold A3D LVSA b-SSFP sequence was also acquired in the same orientation using the following parameters: 3.0/1.5 ms, flip angle  $50^\circ$ , pixel size  $2.0 \times 2.0$  mm, slice thickness 2 mm overlapping, reconstructed voxel size  $1.2 \times 1.2 \times 2$  mm, number of slices 50-60, cardiac phases 20, sensitivity encoding (SENSE) factor 2.0 in both anterior-posterior and right-left directions.

## 2.3. Performance Testing

The structured decision forest model was trained on 410 A3D images at end diastole - together with the associated myocardial contours - following the present implementation details: image patch size  $36 \times 36$  px, edge map labels size  $12 \times 12 \times 2$  px, number of trees  $T = 16$ . The proposed registration technique was tested on 25 2D cine stacks at the end diastolic phase. For comparison, we also tested a conventional intensity-based registration technique, in which the block matching algorithm [6] was directly applied to the image intensity values between adjacent slices. The block matching algorithm was selected as competitor because in our experiments it consistently outperformed the more common methods based on normalized mutual information. A 3D multi-atlas segmentation technique [7] was applied to estimate end-diastolic left ventricular vol-

Table 1. Results for volume differences (mean  $\pm$  std).

	$\Delta V$ (ml)
$ EDV_{pre} - EDV_{ref} $	$10.34 \pm 9.41$
$ EDV_I - EDV_{ref} $	$10.76 \pm 9.64$
$ EDV_{PEMs} - EDV_{ref} $	$7.73 \pm 6.55$

umes respectively from the A3D images ( $EDV_{ref}$ ), the 2D SA image stacks before ( $EDV_{pre}$ ) and after slice-alignment using the proposed technique ( $EDV_{PEMs}$ ) and using the conventional intensity-based registration method ( $EDV_I$ ). Since the motion-free A3D images and 2D SA image stacks were acquired from the same subjects,  $EDV_{ref}$  could be used as reference, and volume differences between  $EDV_{ref}$  and  $EDV_{pre}$ ,  $EDV_{PEMs}$  and  $EDV_I$ , respectively, could be used as proxy to registration accuracy.

## 3. Results

Fig. 3 shows examples of obtained registration results: qualitatively the proposed registration technique (PEMs-based) does a better job in slice realignment compared to the conventional (intensity-based) one. Statistics obtained from the computed volume differences are reported in Table 1: volume estimation performed on stacks registered using the proposed technique is more reliable in comparison to the conventional technique (statistically significant improvement,  $p = 0.016$  using t-test).

## 4. Discussion and Conclusion

In this study we presented a registration technique to spatially align motion corrupted 2D SA MR image stacks using 3D PEMs. The proposed approach has been successfully applied to 3D multi-modal registration between MR and either ultrasound or CT images [8]. Differently from that work, in which 3D PEMs were associated with 3D im-

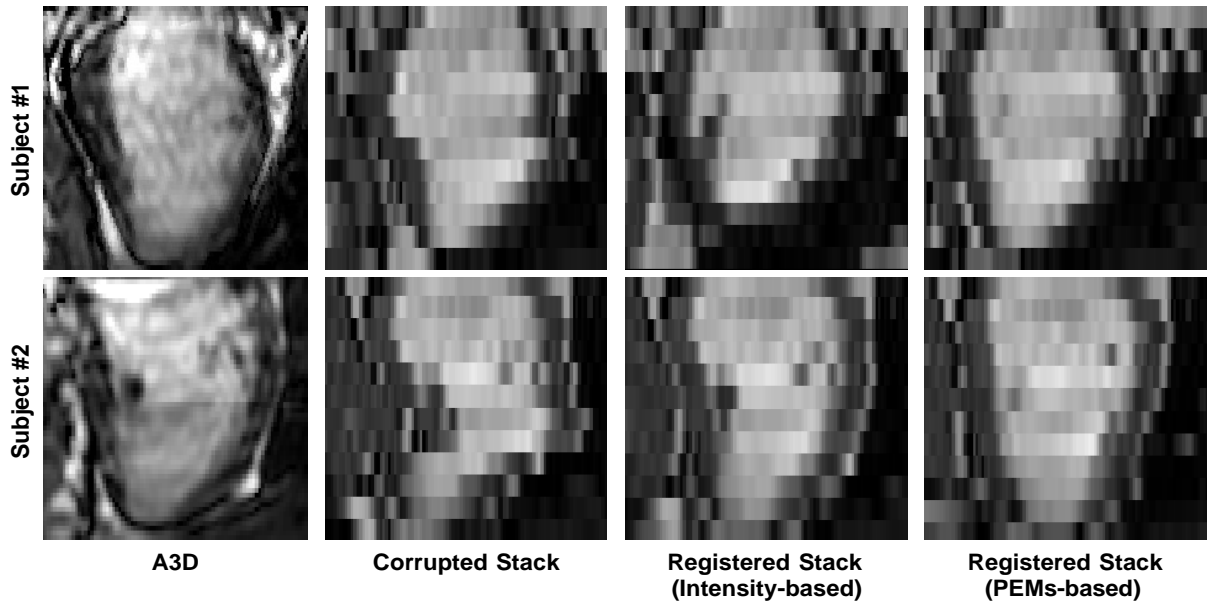


Figure 3. Registration results obtained in two different subjects using the proposed technique (PEMs-based) and a conventional (intensity-based) approach. A3D images and acquired 2D SA stacks are also displayed for comparison.

age patches, the structured decision forest presented in this paper associates 3D PEMs to 2D image patches extracted from a single image slice. The obtained results show that the proposed approach was able to perform slice realignment with higher accuracy than the conventional intensity-based technique (as far as volume estimation was concerned). Importantly, the proposed approach is currently limited to in-plane spatial alignment. Out-of-plane motion correction from a single image stack is in fact an ill-posed problem, and while the availability of other sequences (e.g. other SA or LA stacks) would allow to perform registration also in the z direction, these data are often unavailable.

In conclusion, the proposed technique provides a higher registration accuracy for the spatial alignment of 2D SA images in comparison to a conventional intensity-based method, and therefore - when no additional sequences are available - should be adopted as a pre-processing step to volumetric estimation from cine CMR image stacks.

## References

[1] Zhuang X, Shi W, Duckett S, Wang H, Razavi R, Hawkes D, Rueckert D, Ourselin S. A Framework Combining Multi-sequence MRI for Fully Automated Quantitative Analysis of Cardiac Global And Regional Functions. *Functional Imaging and Modeling of the Heart 2011*;LNCS 6666:367–374.

[2] Chandler AG, Pinder RJ, Netsch T, Schnabel JA, Hawkes DJ, Hill DLG, Razavi R. Correction of misaligned slices in multi-slice cardiovascular magnetic resonance using slice-to-volume registration. *Journal of cardiovascular magnetic resonance 2008*;10:13.

[3] Carminati MC, Maffessanti F, Caiani EG. Automated motion artifacts removal between cardiac long- and short-axis magnetic resonance images. *Computing in Cardiology 2012*; 689–692.

[4] Betancur J, Simon A, Schnell F, Donal E, Hernández A, Garreau M. Evaluation of a motion artifacts removal approach on breath-hold cine-magnetic resonance images of hypertrophic cardiomyopathy subjects. *IX International Seminar on Medical Information Processing and Analysis 2013*;892205.

[5] Dollar P, Zitnick CL. Fast Edge Detection Using Structured Forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence 2015*;37(8):1558–1570.

[6] Ourselin S, Roche A, Subsol G, Pennec X, Ayache N. Reconstructing a 3D structure from serial histological sections. *Image and Vision Computing 2001*;19(1-2):25–31.

[7] Bai W, Shi W, O’Regan DP, Tong T, Wang H, Jamil-Copley S, Peters NS, Rueckert D. A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images. *IEEE Transactions on Medical Imaging 2013*;32(7):1302–1315.

[8] Oktay O, Schuh A, Rajchl M, Keraudren K, Gomez A, Heinrich MP, Penney G, Rueckert D. Structured Decision Forests for Multi-modal Ultrasound Image Registration. *Medical Image Computing and Computer Assisted Intervention 2015*; 9350:363–371.

Address for correspondence:

Giacomo Tarroni  
Exhibition Road, London SW7 2AZ, UK  
g.tarroni@imperial.ac.uk