Efficient Post-processing of Diffusion Tensor Cardiac Magnetic Imaging Using Texture-conserving Deformable Registration

Fanwen Wang^{a,b,c}, Pedro F.Ferreira^{b,c}, Yinzhe Wu ^{a,b,c}, Andrew D. Scott ^{b,c}, Camila Munoz^{b,c}, Ke Wen^{b,c}, Yaqing Luo^{b,c}, Jiahao Huang^{a,b,c}, Sonia Nielles-Vallespin ^{b,c}, Dudley J.Pennell ^{b,c} and Guang Yang*a,b,c,d

aBioengineering Department and Imperial-X, Imperial College London, London W12 7SL, UK;
bNational Heart and Lung Institute, Imperial College London, London SW7 2AZ, UK;
Cardiovascular Magnetic Resonance Unit, Royal Brompton Hospital, Guy's and St Thomas' NHS Foundation Trust, London SW3 6NP, UK;
dSchool of Biomedical Engineering & Imaging Sciences, King's College London, London WC2R 2LS, UK;

ABSTRACT

Diffusion tensor based cardiac magnetic resonance (DT-CMR) is a method capable of providing non-invasive measurements of myocardial microstructure. Image registration is essential to correct image shifts due to intra and inter breath-hold motion. Registration is challenging in DT-CMR due to the low signal-to-noise and various contrasts induced by the diffusion encoding in the myocardial and surrounding organs. Traditional deformable registration destroys the texture information while rigid registration inefficiently discards frames with local deformation. In this study, we explored the possibility of deep learning-based deformable registration on DT-CMR. Based on the noise suppression using low-rank features and diffusion encoding suppression using variational auto encoder-decoder, a B-spline based registration network extracted the displacement fields and maintained the texture features of DT-CMR. In this way, our method improved the efficiency of frame utilization, manual cropping, and computational speed.

Keywords: Image Registration, Deep Learning, Disentanglement, Diffusion Tensor Cardiac Magnetic Resonance.

1. DESCRIPTION OF PURPOSE

Diffusion tensor based cardiac magnetic resonance (DT-CMR) is a non-invasive imaging modality that can probe the microstructure of the heart. After acquiring data with different diffusion encoding directions, a tensor can be calculated based on the self-diffusion of water molecules in the myocardium. Cardiac triggering is combined with breath-holding to minimize the effects of motion of the images.

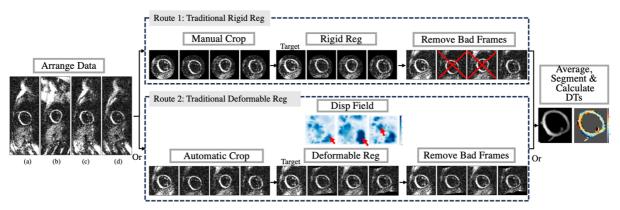


Figure 1. The post-processing pipeline of DT-CMR with details on registration. Reg denotes registration, Dis denotes the displacement field. The severe changes in the displacement field, pointed out by red arrows, indicate the destruction of texture information.

Image registration is a process that finds the mapping from one image to another and aligns them in one coordinate system. Image registration in DT-CMR poses a significant challenge due to the complex respiratory and cardiac motion, the variability of heart rates, as well as the heterogeneity of signal intensities and contrasts present in both the myocardial and surrounding chest wall and stomach¹. Any misalignment induces bias in the final parameter estimation. Most DT-CMR pipelines include manual cropping to exclude the signal from neighboring organs,

rigid registration of different frames, and manual removal of bad frames (Fig 1, Route 1 panel). Varying contrast and intrinsically low SNR induced by diffusion encoding in the myocardium makes traditional and deep-learning based deformable registration incompatible. The red arrows in the displacement field show the undesirable warping of the texture information (Fig 1, Route 2 panel). The mapping of intensity within the myocardial performed by the majority of registration methods makes the frames anatomical alike but destroys the texture information embedded. However, simple rigid registration followed by bad frames removal (as is typical) results in longer scanning time.

In this study, we explored the possibility of using deformable registration on DT-CMR post-processing to improve efficiency. According to the underlying physics, DT-CMR can be modeled as a combination of anatomical tissue property, diffusion encoding information, and acquisition noise. By suppressing the noise using low-rank features and the diverse contrasts induced by diffusion encoding using a beta variational autoencoder, we disentangled the anatomy and contrast of DT-CMR. Then a diffeomorphic B-spline deep learning network on the different anatomies but similar contrast images were used to derive the smooth displacement field. The proposed method improved the efficiency in 3 ways: more efficient use of frames, less need for manual cropping, and faster computational speed.

2. METHOD

2.1 Data Acquisition and Preparation

20 healthy volunteers were scanned at 3T and 1.5T using SIEMENS MRI scanners with a Stimulated Echo Acquisition Mode (STEAM) based Echo Planar Imaging (EPI) sequence with a spatial resolution of $2.8 \times 2.8 \times 8$ mm³. Nine diffusion directions with b = 150 (b150) and 600 s/mm² (b600) and one reference b = 0 s/mm² in short-axis midventricular slice at both end-systole (ES) and end-diastole (ED). We got 20 (subjects) \times 2 (1.5 T/3T) \times 2 (ES/ED) = 80 cases. Since every scan had different orientations and volunteer positions, we treated each case as independent. For a fair comparison, all the frames of 7 averages of b600 acquisition and 2 averages of b150 were used. The training, validation, and testing dataset included b150 for the images was automatically cropped. The images from the same subject were normalized using the max intensity of the brightest frame as the input of diffusion encoding suppression and registration.

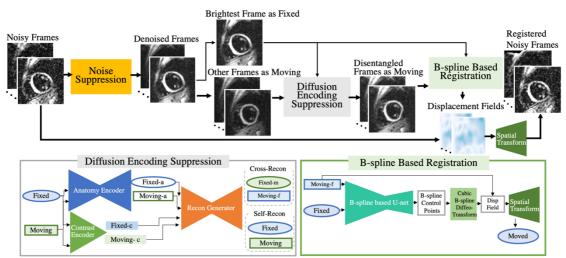


Figure 2. The architecture of the texture-conserving network. -c and -a denote the disentangled contrast and anatomical information. -f and -m denote the cross-reconstruction using the contrast from fixed or moving frames respectively.

2.2 Network Architecture

Noise Suppression: We first developed a groupwise Principal Component (PC) analysis denoising² method to suppress noise and retain the diffusion encoding information (Fig 2). The PCs with 97% of the signal information were included with an autocorrelation function for genuine diffusion effects. The rejected PCs were set to zero and multiplied by corresponding PC scores to generate denoised frames.

Diffusion Encoding Suppression: We implemented a deep-learning based disentanglement network³ for the denoised frames in a pairwise manner (Fig 2). We chose the frame with the highest intensity as the fixed d^F and all the others as moving d^M . Variational autoencoders extracted anatomy information (a^F, a^M) and the mean and

contrast information (c^F, c^M) . Then we synthesized the frames using self-reconstruction and cross-reconstruction and took the cross-reconstructed image d^{FM} with the initial anatomy a^M but similar c^F as the input of the registration network. Loss for self-reconstruction $L_{self-recon}$, cross-reconstruction $L_{cross-recon}$, perceptual L_{perce} , anatomy similarity L_{ana} and contrast bottleneck $L_{contrast}$ were utilized to supervise the network:

$$L = \lambda_1 (L_{self-recon} + L_{cross-recon}) + \lambda_2 L_{perce} + \lambda_3 L_{ana} + \lambda_4 L_{contrast}$$

 $L = \lambda_1 \big(L_{self-recon} + L_{cross-recon} \big) + \lambda_2 L_{perce} + \lambda_3 L_{ana} + \lambda_4 L_{contrast}$ **B-spline Based Registration:** To suppress undesirable drastic displacement fields, we adopted a diffeomorphic b-spline based registration network⁴. Based on a variant of U-net, we generated the control points of the image (Fig 2). With a differentiable mutual information metric, the registration network succeeded in generating a smooth displacement field for registration. The loss function was defined as:

$$L(d^F, d^{FM} \circ \psi) = L_{NMI}(d^F, d^{FM} \circ \psi) + \lambda L_{Reg}$$

 $L(d^F, d^{FM} \circ \psi) = L_{NMI}(d^F, d^{FM} \circ \psi) + \lambda L_{Reg}$ A Parzen window⁴ was applied on the joint histogram to make the loss differentiable. After extracting the desired displacement field ψ , we set the brightest frame as fixed and registered all the other frames accordingly.

2.3 Experiments Details

Our proposed texture-conserving network was compared with traditional rigid (Rigid) and deformable registration (Deformable) using Elastix⁵. Ablation studies used Deep Learning-based (DL) registration on noisy frames (Noisy+DL), and DL on denoised frames (Denoised+DL). For the disentangled network, we set $\lambda_1 = 1$, $\lambda_2 = 0.03$, $\lambda_3 = 0.02, \lambda_4 = 10^{-8}$. For the registration network, we set $\lambda = 0.1$, control point spacing = 4. Both networks were trained using epoch = 100 and chose the model with the best performance in the validation set.

Our evaluation criteria were 1) The number of negative eigenvalues which increases due to the effects of motion, noise, or misregistration artifacts⁶. 2) Image stack visualization which shows a stacked intensity profile from all images in the horizontal and vertical direction from all frames. 3) Helix Angle Gradient (HAG) a transmural line profile in HA on healthy volunteers should be linear from endocardium to epicardium⁶. Line profiles with negative slopes and fitting R² larger than 0.3 were included. R² and root-mean-square error (RMSE) of the linear regression were compared.

RESULTS

The output of images using denoising and disentanglement and their displacement fields were shown (Fig 3).

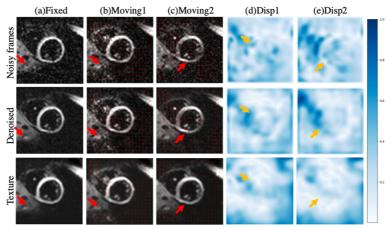


Figure 3. The displacement fields in ablation studies. Red arrows point out the diverse contrast in the surrounding organs and myocardium. Yellow arrows show the corresponding drastic deformation.

Table 1. Comparison and Ablation methods. Negative X denotes the number of pixels with X negative eigenvalues. All the data are shown with mean and standard deviation.

Methods	R-squared (†)	RMSE(↓)	Negative 1(↓)	Negative 2(↓)	Negative 3(↓)
Rigid	0.876 ± 0.155	6.509 ± 5.265	59.90±48.62	7.80 ± 10.70	1.30±3.77
Deformable	0.850 ± 0.164	6.557 ± 5.006	56.80 ± 39.28	7.30 ± 9.83	1.00 ± 3.16
Noisy-DL	0.844 ± 0.164	6.900 ± 5.494	57.30±35.31	9.70 ± 11.15	0.40 ± 0.70
Denoised-DL	0.857 ± 0.161	6.769 ± 5.645	56.20±31.72	7.60 ± 7.78	0.40 ± 0.97
Proposed	0.881 ± 0.139	5.716±4.560	37.00±27.74	4.70 ± 9.12	0.30 ± 0.95

Compared with the original frames, denoised ones were less noisy with smoother displacement fields. The contrast within the myocardial in the moving frames (c) differed from the fixed frame (a). Moreover, the contrast in the bottom right of the myocardial, pointed by a red arrow, was suppressed using the disentanglement with smaller deformation in the displacement field. Compared with the registration networks applied on the noisy frames or denoised frames, the texture-conserving network showed less undesirable deformation within the myocardial, pointed by yellow arrows.

The tensor parameter estimations of different methods were evaluated (Fig 4 and Table 1). The proposed method had the lowest RMSE, and lowest R^2 with the least number of negative eigenvalues.

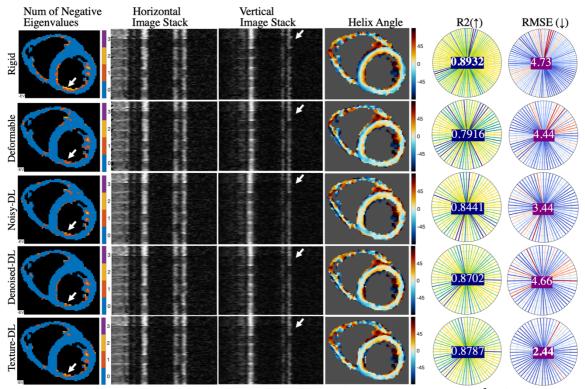


Figure 4. HA is measured in degrees. The visualization of the diffusion tensor parameters with R^2 and RMSE of the fitting of helix angle line profiles.

4. CONCLUSIONS

We investigated the potential of performing deformable registration on DT-CMR by mitigating the effects of noise and removing the diffusion encoding, ultimately enabling more precise registration. Using PCA-based denoising to suppress noise and disentangling network to suppress contrast changes due to diffusion encoding, we applied a diffeomorphic B-spline registration network with differentiable mutual information on DT-CMR. Results showed that compared with traditional rigid and deformable registration methods, the proposed methods achieved the highest R², minimum RMSE, and lowest number of negative eigenvalues on healthy subjects.

REFERENCES

- [1] P. F. Ferreira, R. R. Martin, A. D. Scott *et al.*, "Automating in vivo cardiac diffusion tensor postprocessing with deep learning–based segmentation," Magnetic Resonance in Medicine, 84(5), 2801-2814 (2020).
- [2] O. J. Gurney-Champion, D. J. Collins, A. Wetscherek *et al.*, "Principal component analysis for fast and model-free denoising of multi b-value diffusion-weighted MR images," Phys Med Biol, 64(10), 105015 (2019).
- [3] C. Yang, Y. Zhao, L. Huang *et al.*, "DisQ: Disentangling Quantitative MRI Mapping of the Heart," Lecture Notes in Computer Science. 291-300.
- [4] H. Qiu, C. Qin, A. Schuh et al., "Learning Diffeomorphic and Modality-invariant Registration using B-splines."
- [5] S. Klein, M. Staring, K. Murphy *et al.*, "elastix: A Toolbox for Intensity-Based Medical Image Registration," IEEE Transactions on Medical Imaging, 29(1), 196-205 (2010).
- [6] P. F. Ferreira, P. J. Kilner, L.-A. McGill *et al.*, "In vivo cardiovascular magnetic resonance diffusion tensor imaging shows evidence of abnormal myocardial laminar orientations and mobility in hypertrophic cardiomyopathy," Journal of Cardiovascular Magnetic Resonance, 16(1), 87 (2014).