PROCEEDINGS OF SPIE

Synthetic velocity mapping cardiac MRI coupled with automated left ventricle segmentation

Xiaodan Xing, Yinzhe Wu, David Firmin, Peter Gatehouse, Guang Yang

Synthetic Velocity Mapping Cardiac MRI Coupled with Automated Left Ventricle Segmentation

Xiaodan Xing[a,*], Yinzhe Wu[a, b], David Firmin[a, c], Peter Gatehouse[a, c], and Guang Yang[a,*]

[a]National Heart and Lung Institute, Imperial College London, London, UK
[b]Department of Bioengineering, Imperial College London, London, UK
[c]Cardiovascular Research Centre, Royal Brompton Hospital, London, UK

ABSTRACT

Temporal patterns of cardiac motion provide important information for cardiac disease diagnosis. This pattern could be obtained by three-directional CINE multi-slice left ventricular myocardial velocity mapping (3Dir MVM), which is a cardiac MR technique providing magnitude and phase information of the myocardial motion simultaneously. However, long acquisition time limits the usage of this technique by causing breathing artifacts, while shortening the time causes low temporal resolution and may provide an inaccurate assessment of cardiac motion. In this study, we proposed a frame synthesis algorithm to increase the temporal resolution of 3Dir MVM data. Our algorithm is featured by 1) three attention-based encoders which accept magnitude images, phase images, and myocardium segmentation masks respectively as inputs; 2) three decoders that output the interpolated frames and corresponding myocardium segmentation results; and 3) loss functions highlighting myocardium pixels. Our algorithm can not only increase the temporal resolution 3Dir MVMs, but can also generates the myocardium segmentation results at the same time.

Keywords: Non-linear Frame Synthesis, Cardiac MR, Myocardial Velocity Mapping

1. DESCRIPTION OF PURPOSE

Three-directional cine multi-slice left ventricular myocardial velocity mapping (3Dir MVM) is a unique CMR technique that provides both spatial and temporal patterns of cardiac motion by acquiring magnitude images and velocity-encoded phase images correspondingly. Magnitude images, as shown in Figure 1(a), provide the spatial location and shape information of left ventricles. Velocity encoded phase images in Figure 1 (c), encodes the velocities of cardiac movement in three orthogonal directions.[1] However, the application of this technique is limited due to the long acquisition time. Reducing the acquisition time while maintaining the temporal resolution becomes an interesting problem. To increase the temporal resolution of velocity mappings with a short acquisition time, generating the intermediate velocity mapping frames from existing frames is a potential a solution.

In this study, we propose a frame synthesis algorithm to increase the temporal resolution of 3Dir MVM data. Our aim is to synthesize unknown target images from their existing neighbors using a deep neural network. By improving the temporal resolution of 3Dir MVM data, we could reduce the acquisition time while not corrupting temporal patterns of the cardiac motion.

Problem description: We choose $t = \tau$ as a a start time point and $t = \tau + 4$ as an end time point. The goal of our project is to synthesize images at $t = \tau + k$, $k \in \{1, 2, 3\}$. Since cardiac cycles for all healthy controls are similar, the relative time position of each image can provide decisive information of the shape of myocardium and surrounding tissues. Thus, our algorithm aims to achieve a conditional synthesis problem with τ and k as conditions.

---

*Send correspondence to X. Xing (xxing@imperial.ac.uk) and G. Yang (g.yang@imperial.ac.uk).
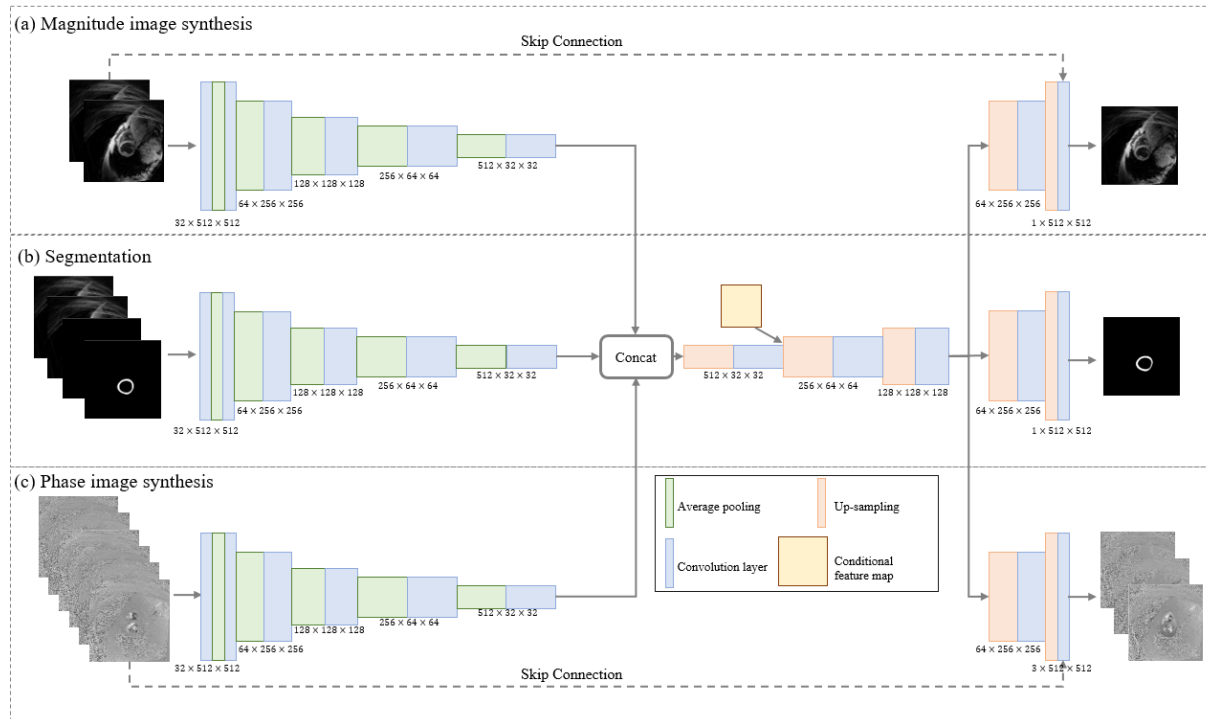
Figure 1. Multi-task attention UNet for different tasks. The resolution specified under each block is the resolution of the output feature map.

2. METHOD

We introduce below the proposed multi-head multi-tail architecture for solving the synthesis problem. We develop an attention-based UNet structure[2] and then split it into three threads with different encoders and decoders. Three encoders receive different inputs and the feature maps encoded are concatenated with a condition map. This condition map is a two-channel feature map that contains the normalized time point information of input images and the target image, which is aforementioned τ and k. To speed up the loss convergence, we incorporate skip connections at image synthesis stages.

Data acquisition We trained our algorithm on our in-house dataset,[3] which was obtained from Royal Brompton Hospital and contained 30 healthy subjects. The myocardium of all frames were manually segmented. Temporally, 50 frames were reconstructed per cardiac cycle. For each frame, one 2D magnitude image and three 2D phase images encoding velocities at three orthogonal directions were obtained. The resolution of each frame was $1.7 \times 1.7 mm^2$ and was then re-sampled into $0.85 \times 0.85 mm^2$ with a matrix size of 512×512. We will address this series of frames as *a series* in the following statement. Since 3Dir MVM had multiple slices, each subject had 3-5 slices acquired from base to apex of the left ventricle, which indicated that there were 3-5 series for each subject. We randomly selected 20 subjects (90 series) for training, 5 subjects (28 series) for validating and 5 subjects (27 series) for an independent testing.

Network parameters Our model was trained on 512×512 frames, but the network could be adaptable to all input resolutions because of the pooling layers. The two channel condition map had a size of $2 \times 32 \times 32$. The learning rate was 0.001, and batch size was set to 32. To avoid information leakage between frames, we replaced batch normalization by instance normalization.

Loss functions and evaluation metrics For synthesis, we observed that vanilla distance error over all pixels, such as Mean Absolute Error (MAE), might bring network to a local minimum because most pixels in magnitude images and phase images are noise. In this case, we used a weighted MAE l_{MAE} to highlight the pixels within the region of interest (ROI). We used two weight maps: ω_1 is the de-noising weight map, which is computed by original magnitude images and only non-background (non-black) pixels are highlighted; ω_2 is the myocardium
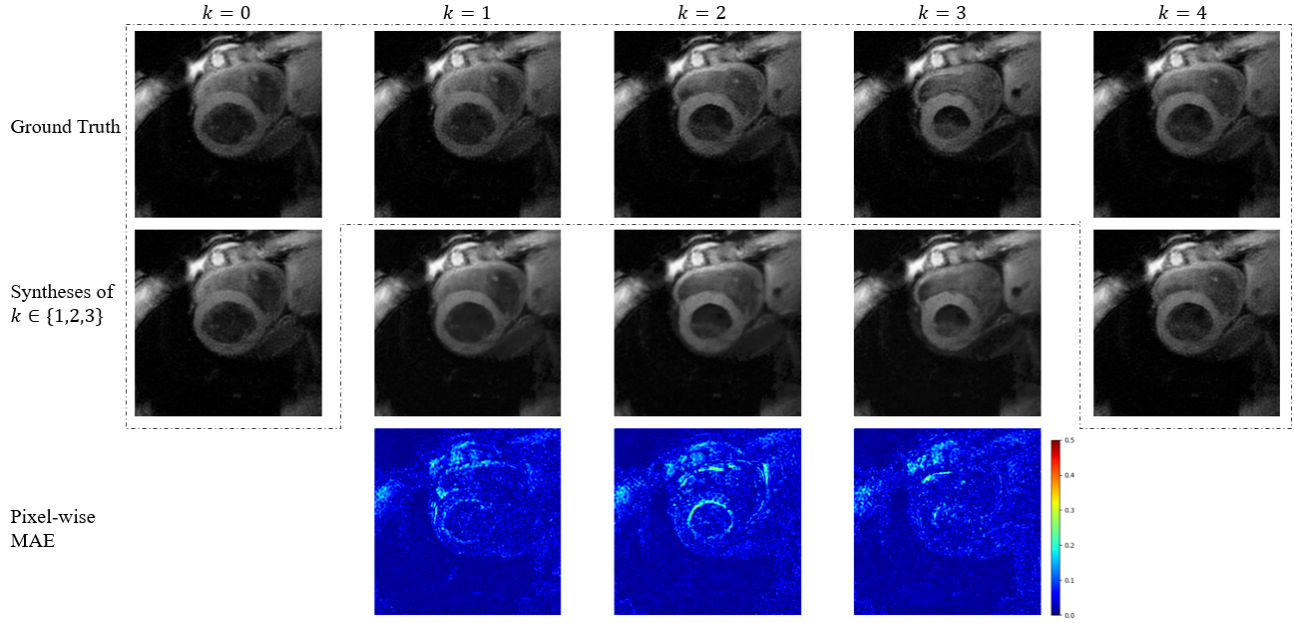
Figure 2. Example magnitude synthesis result. This example has a $\tau = 0$. Magnitude images in dash line boxes are ground truth magnitude images, while others are synthesized results.

weight map, which is obtained by highlighting all pixels inside the epicardiums of all frames with a dilation of 2 pixels.

For segmentation, since the shape variation mainly occurs at the boundary of myocardium, we used the summation of the Dice loss l_{DC} and the boundary loss l_{BD}[4] as the loss function. The overall loss function is shown below:

$$l = w_{syn}(\omega_1 + \omega_2)l_{MAE} + w_{seg}(l_{DC} + l_{BD}) \tag{1}$$

Here w_{syn} and w_{seg} represent the weight parameters for the synthesis and the segmentation respectively. To evaluate the performance of our synthesis model, we used three quantitive metrics: 1) MAE; 2) Peak signal-to-noise ratio (PSNR); 3) Structural similarity (SSIM).

3. RESULTS

Quantitative results of our proposed synthesis model are shown in Table 1. We first compared our results with linear interpolation method and optical flow-based method, which was achieved by Horn-Schunck optical flow. Our algorithm could conditionally synthesize the intermediate 3Dir MVM frames with a promising accuracy on magnitude image. Example results of magnitude synthesis and corresponding segmentation and phase synthesis can be visualized in Figure 2, 3 and 4.

We then validated the performance of the proposed modules in our model via ablation studies. For ablation studies, we added two metrics to evaluate the performance of predicted segmentation. To evaluate our segmentation, Dice scores were used, which can also be considered as an additional validation for our synthesis tasks. The velocity coefficient is calculated based on the velocity curves. Using segmentation masks and phase images, we could calculate the longitudinal, radial and circumferential velocity of the myocardium with a self-developed Python toolkit. The velocity coefficient is calculated by the averaged Pearson's correlation between predicted and true velocities.
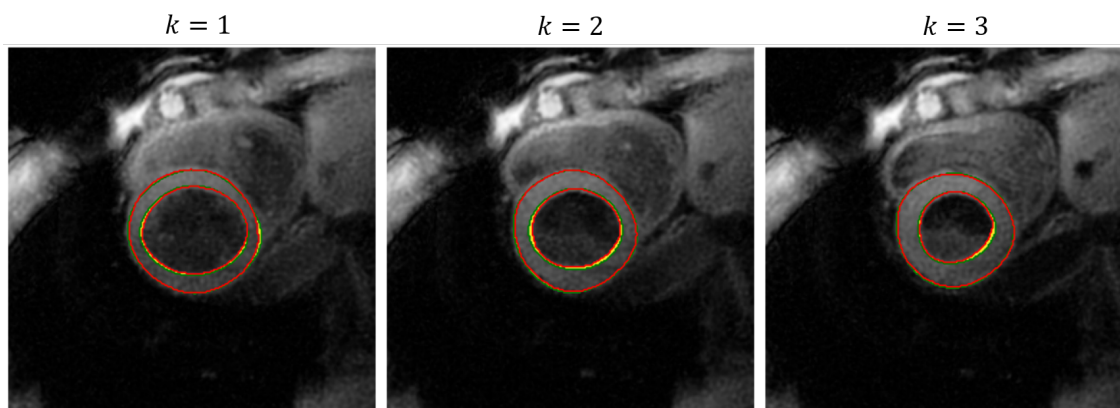
Figure 3. Segmentation results of one subject with $\tau = 0$ and $k \in 1, 2, 3$. Green contours are the ground truth segmentation result, while red contours are the predictions. The absolute difference between ground truth and prediction is highlighted with yellow.
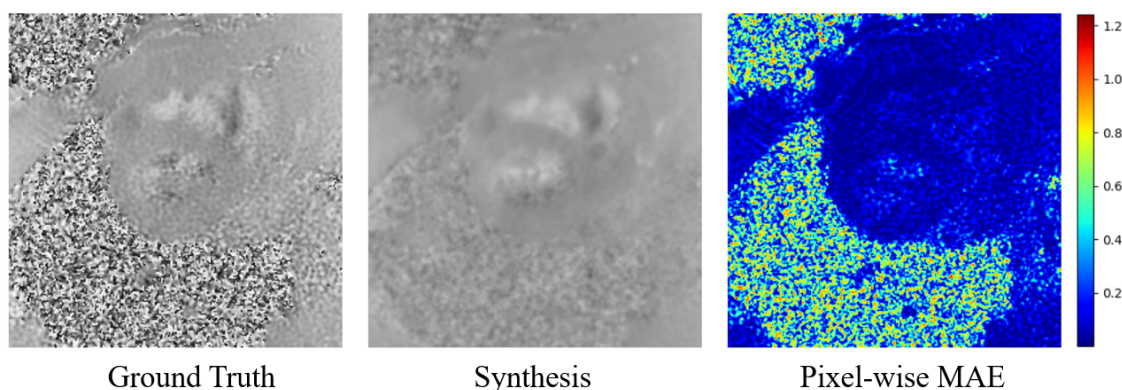

Ground Truth | Synthesis | Pixel-wise MAE

Figure 4. The phase synthesis result of one subject with $\tau = 0$ and $k = 1$.

Table 1. Comparison methods.

Modality	Magnitude			Phase		
Methods	MAE	PSNR	SSIM	MAE	PSNR	SSIM
Linear Inter-polation	0.037	76.638	0.841	0.299	58.696	0.139
Horn-Schunck optical flow	0.027	79.246	0.871	0.363	56.916	0.158
Ours	**0.017**	**83.505**	**0.929**	**0.166**	**63.727**	**0.225**

Table 2. Ablation studies.

Independent encoder	Independent decoder	Shared hidden layer	Weighted loss function	Magnitude PNSR	Phase PSNR	Dice coeffi-cient	Velocity coefficient
✓	✓	×	✓	73.368	59.982	0.940	0.841
×	×	✓	✓	76.654	61.746	0.945	0.868
✓	✓	✓	×	76.487	59.272	0.945	0.832
✓	✓	✓	✓	**83.505**	**63.727**	**0.959**	**0.889**

4. CONCLUSIONS

In this study, we have proposed a multi-task attention UNet structure to synthesize intermediate 3Dir MVM frames using their existing neighbors. The experimental results have shown that our algorithm has not only improved the temporal resolution of 3Dir MVM data by synthesizing both magnitude and phase images, but has also generated corresponding left ventricle segmentation automatically and accurately.

It is of note that although our algorithm has achieved promising performance on magnitude synthesis and segmentation generation, its performance on the synthesis of phase images still needs to be improved. Phase images lack structural visual information, or edges, compared to magnitude images, whereas convolution operations focus on edges of objects. In the future, we will focus on improving the performance of phase image synthesis. Our future work will also include investigation of the application of our algorithm from a single modality synthesis into cross-modality syntheses, such as synthesizing phase images from magnitude images.

REFERENCES

[1] Lotz, J., Meier, C., Leppert, A., and Galanski, M., "Cardiovascular flow measurement with phase-contrast mr imaging: basic facts and implementation," *Radiographics* **22**(3), 651–671 (2002).

[2] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., et al., "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999* (2018).

[3] Simpson, R., Keegan, J., and Firmin, D., "Efficient and reproducible high resolution spiral myocardial phase velocity mapping of the entire cardiac cycle," *Journal of cardiovascular magnetic resonance* **15**(1), 1–14 (2013).

[4] Kervadec, H., Bouchtiba, J., Desrosiers, C., Granger, E., Dolz, J., and Ayed, I. B., "Boundary loss for highly unbalanced segmentation," in [*International conference on medical imaging with deep learning*], 285–296, PMLR (2019).