International Journal on Magnetic Particle Imaging
Vol 9, No 1, Suppl 1, Article ID 2303033, 5 Pages

Infinite Science Publishing

Proceedings Article

Recovery of fundamental frequency component in magnetic particle imaging Using an attention-based neural network

Zechen Wei[a,b,c] · Yanjun Liu[d,e] · Tao Zhu[a,b,c] · Xin Yang[a,b,c] · Jie Tian[a,b,d,e] · Hui Hui[a,b,c,*]

[a] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Beijing, China
[b] Beijing Key Laboratory of Molecular Imaging, Beijing, China
[c] University of Chinese Academy of Sciences, Beijing, China
[d] Key Laboratory of Big Data-Based Precision Medicine (Beihang University), Ministry of Industry and Information Technology of the People's Republic of China, Beijing, China
[e] School of Engineering Medicine & School of Biological Science and Medical Engineering, Beihang University, Beijing, China
[*] Corresponding author, email: hui.hui@ia.ac.cn

Abstract

Magnetic particle imaging (MPI) is a rapidly developing medical imaging modality, which uses the nonlinear response of superparamagnetic iron oxide nanoparticles to the applied magnetic field to image their spatial distribution. Due to the direct feedthrough of excitation signals, the existing MPI systems directly filter out the fundamental frequency component of the received signal, resulting in the loss of first harmonic information. In this work, we proposed a deep learning (DL) method adopting self-attention mechanism, which can effectively recover fundamental frequency component of the signals in the presence of background noise. At the same time, our method deals with two-dimensional time-frequency spectrum obtaining by short time Fourier transform (STFT) from the time domain signals. The performance of our method is analyzed via simulation experiments, which show that our method can effectively recover first harmonic information and obtain high quality MPI reconstructed images.

I. Introduction

Magnetic particle imaging (MPI), a new tracer based imaging method, has been proposed by B. Gleich and J. Weizenecker [1]. MPI utilizes the nonlinear magnetization response of superparamagnetic iron oxide nanoparticles to generate an image of their spatial distribution with high resolution, contrast, and sensitivity [1-3]. Due to the inherent structure of the MPI system (coaxial assembly of transmit and receive coils), in addition to particle signal, the received signal is coupled with excitation signal, which can be many orders of magnitude larger than the particle signal. The existing systems directly filter out the fundamental frequency component of received signal by analog/digital filtering, resulting in the loss of first harmonic, which breaks the shift-invariant property of the imaging system and affects the quality of reconstructed image [4]. Therefore, recovering the fundamental frequency component of the time domain signal contributes to high-quality image reconstruction.

Currently, there are two common-used reconstruction methods, the X-space method [5] and the system

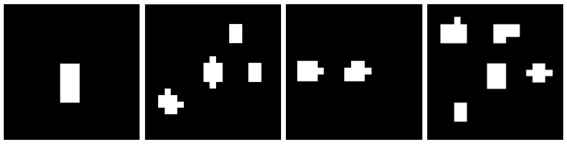

Figure 1: Representative simulated phantoms.

function method [6]. The X-space method maps the received signal directly to the image domain [5]. Therefore, the loss of first harmonic has a great impact on reconstruction. The system function method establishes a system matrix describing the relationship between particle distribution and received signal [6]. The recovery of first harmonic provides more information for the system matrix, which helps to improve the reconstruction quality. Some studies showed that the lost first harmonic corresponds to offsets of the constant (or dc) component of the MPI image and restored it to improve the quality of reconstruction [4, 7, 8].

In this paper, we propose a deep learning (DL) method to **re**cover the **f**undamental **f**requency component (ReFFNet). We fuse the information of time and frequency domain to obtain two-dimensional (2D) time-frequency spectrum, and recover the fundamental frequency component through end-to-end processing of the spectrum. We apply consecutive residual blocks with attention module to learn useful features and suppress the residual features [9, 10]. Specifically, we use channel attention module to exploit the features from different channels and suppress the channels that originate from noise according to the channel weights. Besides, we extract features from the time and frequency dimensions respectively, and recover the fundamental frequency component by fusing the features. The simulation experiments prove the effectiveness of our method, which can recover the first harmonic information in the presence of harmonic interference and Gaussian noise, and obtain high-quality signal and reconstructed image.

II. Datasets

In this study, we obtained 100,000 dot images with size of 21×21 as simulated phantoms using built-in Shepp-Logan function of MATLAB (R2020a) (90,000 as training dataset and 10,000 as testing dataset). The dots in the images simulated the particle distribution. Thus, the number (1-5), position, and size of the dots were all random. Fig. 1 shows the simulated phantom.

Next, we generated the time domain signal based on the phantoms. We simulated a field-free point (FFP) MPI scanner with Cartesian trajectory. In the simulation, the Langevin function was used to describe the nonlinear magnetization of the magnetic particles $M(r,t)$. The diameter of the magnetic nanoparticle was set to be 25

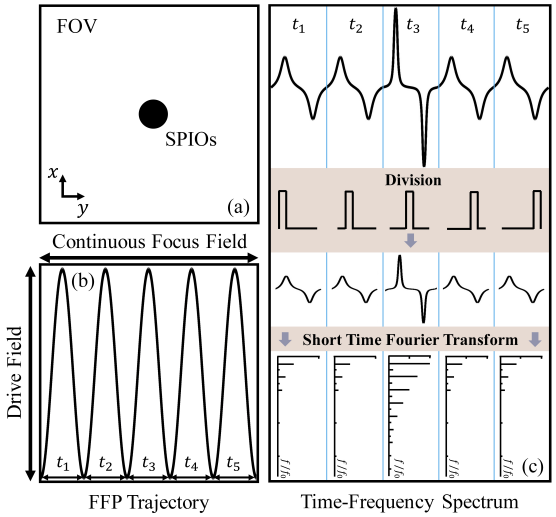

Figure 2: Schematic diagram of scan sequence and time-frequency spectrum generation. (a) Field of view (FOV). (b) FFP trajectory corresponding to the FOV in (a). A continuous focus field is used to move the FFP at a constant speed along the y direction, which can be achieved by applying an alternating magnetic field with a triangular wave function. And the drive field oscillates rapidly along the x direction and is used to excite the response signal of the particles. t_i represents one drive period. (c) The acquisition of time-frequency spectrum. We divide the time domain signal into frames of equal length by rectangular window function based on drive period t_i, followed by STFT to obtain the 2D spectrum.

nm. A selection field with a gradient of (-0.5, 1) T/m in (x, y) directions was utilized. The drive fields were simulated as a 5kHz cosine wave along the x direction, and a continuous focus field at a rotation rate of 7.94 T/s (f_F=39.68Hz) was simulated along the y direction (see Fig. 2(b)). Besides, we simulated a 2D field of view (FOV) with a size of 100×100 mm^2 and pixel spacing of 5mm.

In order to better simulate the received signal collected in the MPI scanner, we added a degradation model to the generated signal, including harmonic interference and Gaussian noise, with signal-to-interference ratio at 5dB and signal-to-noise ratio at 10dB [8]. Besides, we applied a digital band-stop filter to simulate the process of removing the fundamental feedthrough signal in practice, and obtained the signal with noise and without fundamental frequency component.

Further, we obtained the 2D time-frequency spectrum by STFT of the time domain signal (see Fig. 2(c)). Compared with the time domain or frequency domain signal, the 2D spectrum is composed of the frequency spectrum of different time periods. Thus, it is dominated by the frequency domain information and integrates the particle distribution information of time domain, that is, the trajectory information of FFP. Specifically, the signal generated contains one Cartesian cycle of the FFP for the entire FOV and includes 63 drive periods. Thus, the

10.18416/ijmpi.2023.2303033

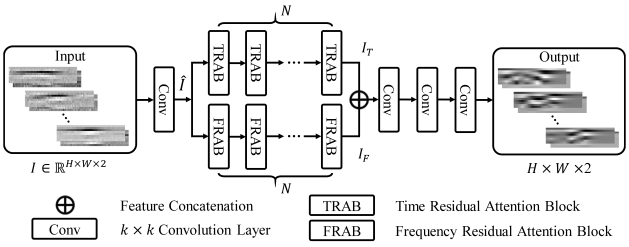

Figure 3: Overall architecture of the proposed ReFFNet.

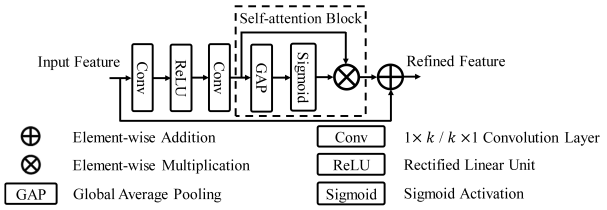

Figure 4: Structure of residual attention block.

signal is divided into L frames of equal length according to the drive period t_i by rectangular window function $w(\bullet)$, and each frame includes one complete drive period. Then, STFT is applied on each frame:

$$\hat{U}_l(k) = \sum_{n=0}^{N-1} \hat{u}^l_{MNP} e^{-j2\pi nk/N} \tag{1}$$

where $\hat{u}^l_{MNP}$ is the l-th frame of the original signal, $\hat{U}_l(k)$ is the k-th spectrum component after STFT of the l-th frame, and N is the frame length. The frequency spectrum obtained by each frame is sorted by column to obtain the 2D time-frequency spectrum $S \in \mathbb{C}^{K\times L}$:

$$S = \left[\hat{U}_1(k)' \quad \ldots \quad \hat{U}_L(k)'\right] \tag{2}$$

The horizontal axis represents the time domain, containing the trajectory information of FFP, and the vertical axis represents the frequency domain, containing the information of different harmonics of the signal.

The 2D spectrum of the original generated signal and the processed signal are obtained as the label and input respectively. In addition, by balancing the contribution of different harmonics and the size of our network model, we believed that the signal composed of the first eight harmonics contained most of the information of the original signal. Thus, we further extracted the first eight harmonics of the time-frequency spectrum as the label and input of our network model.

III. Methods

An overview of the proposed ReFFNet is presented in Fig. 3. Concretely, given a processed spectrum $I \in \mathbb{R}^{H\times W\times 2}$ with H harmonics, W periods, and 2 channel (real part and imaginary part), we first utilized a convolution layer to generate compact feature maps $\hat{I} \in \mathbb{R}^{H\times W\times C}$ capturing time and frequency information. Then, to minimize the interaction of time and frequency dimension, we applied consecutive residual attention blocks to extract local features. Specifically, two parallel processing paths are designed to deal with time and frequency domain features respectively. Each path contains N continuous residual attention blocks. The two paths extract local features I_T in time dimension and I_F in frequency dimension respectively. Then, we concatenate I_T and I_F to get the fusion representations.

The structure of residual attention block in the two paths is basically the same, composed of two convolution layers, one rectified linear unit, one residual skip connection, and one self-attention block (see Fig. 4). The self-attention block we used here is global average pooling (GAP), which computes the channel-wise statistic $\tilde{z} \in \mathbb{R}^{1\times 1\times C}$ of the given input feature map $\tilde{I}$ by averaging across each $\hat{K}\times L$ spatial dimension. The c-th channel statistic $\tilde{z}_c$ is calculated as follows:

$$\tilde{z}_c = F_{GAP}(\tilde{I}_c) = \frac{1}{\hat{K}\times L}\sum_{\hat{k}=1}^{\hat{K}}\sum_{l=1}^{L}\tilde{I}_c(l,\hat{k}) \tag{3}$$

where $\tilde{I}_c \in \mathbb{R}^{\hat{K}\times L}$, $\forall c \in \{1,\ldots,C\}$ refers to the c-th feature of the input feature map. Given an input feature map $\tilde{I}$, the self-attention block deals with it as follow:

$$H_{SAB}(\tilde{I}) = \tilde{I} \bigotimes \delta\left(GAP\left(\tilde{I}\right)\right) \tag{4}$$

where GAP is the channel-wise attention, δ is sigmoid activation, and $\bigotimes$ donates element-wise multiplication. The difference between time residual attention block and frequency residual attention block lies in the size of the convolution kernel. In time path, the size of the convolution kernel is 1×3, focusing on the features of distribution of particles, while in frequency path, the size of the convolution kernel is 3×1, extracting the features between the harmonics. Finally, three convolutional layers were used to reconstruct the feature and recover the spectrum with fundamental frequency component after removing the background noise.

We evaluated the result of our network from three aspects. Firstly, the enhanced spectrum was obtained by our network, and we adopted the mean absolute error (MAE) to calculate the average per-pixel differences between the label and the result. Besides, the spectrum was restored to original size by zero padding and transformed to the time domain signal by Inverse Fast Fourier Transform. Then, the MAE of temporal signal was calculated. Finally, the time domain signal was mapped to image by standard X-space algorithm. Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) were employed to evaluate the image quality.

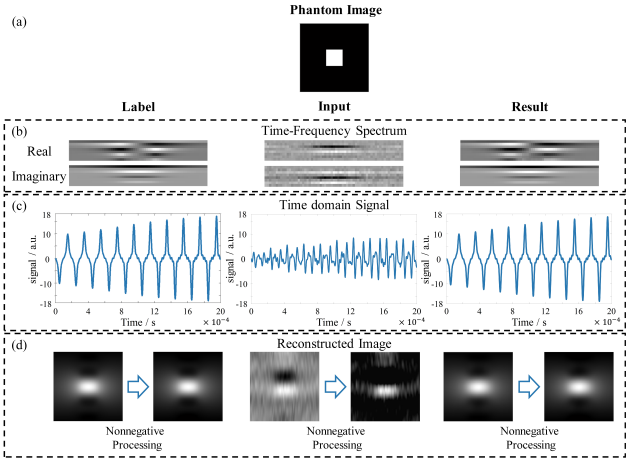

Figure 5: The time-frequency spectrum, time domain signal, and reconstructed image of the ground truth label, input of our model, and result of our model. (a) is the phantom image, (b) is the real and imaginary part of the spectrum, (c) shows ten consecutive periods in the temporal signal obtained from the corresponding spectrum, and (d) is the corresponding reconstructed image and the nonnegative processed image.

IV. Results

To validate the effectiveness of our method, we tested our model on the simulation dataset. The experimental results are summarized in Table 1 and Fig. 5. From Table 1 and Fig. 5, due to the loss of fundamental frequency component, the amplitude and phase of the input signal are affected. At the same time, the existence of noise further distorts the time domain signal (MAE=523.25×10^{-2}). Besides, the degradation from temporal signal creates artifacts on the reconstructed image, which seriously affect the image quality (PSNR=8.40, SSIM=4.69×10^{-2}). And from the time-frequency spectrum of input, we can find that the first harmonic is missing and the spectrum is interfered by noise. From the result, our method can recover the first harmonic while removing the background noise. Thus, it reduces the errors of both real and imaginary part of the time-frequency spectrum. From the time domain signal, our method recovers the amplitude and phase information of the signal, reducing the error by two orders of magnitude (MAE=6.28×10^{-2}). Besides, the image reconstructed from the enhanced signal can also accurately describe the particle distribution (PSNR=44.47, SSIM=99.20×10^{-2}).

V. Discussion

The loss of fundamental frequency component is an inevitable problem due to the structure of MPI system, and it is still a challenging task to recover the signal in the presence of various background noise. In this work, we designed a network model, which extracted local features from time and frequency domain to get only one kind of information, so as to model the relationship between different dimension and recover the fundamental frequency component through high-frequency harmonics. We used residual blocks to increase the modeling power. This learned only the difference between the input and the target output and transformed the distorted-spectrum to a recovered-spectrum efficiently. Besides, with GAP, the importance of different channels was learned, and the value of the feature map was selectively attenuated or amplified according to their importance for more effective training. In addition, GAP increased the modeling capacity of the network to effectively recover the fundamental frequency component in the presence of noise. Moreover, we added background noise and applied a digital filter to the dataset, to reduce the gap between the simulated data and the real data. Thus, the trained model is expected to be effective on the real data from the MPI scanner.

Table 1: Quantitative results of our method.

			Input	**ReFFNet**
Spectrum	MAE ↓	Real	56.77±16.92	**0.23±0.10**
		Imag	144.15±32.67	**4.97±2.78**
Time Domain Signal	MAE ↓($\times10^{-2}$)		469.37±7.82	**11.29±0.19**
Reconstructed Image	PSNR ↑		8.59±1.37	**43.03±3.82**
	SSIM ↑($\times10^{-2}$)		6.86±4.71	**98.91±0.63**

Besides, we dealt with 2D time-frequency spectrum obtained by short time Fourier transform (STFT). The 2D spectrum contains the trajectory information of FFP in different time periods, which directly reflects the distribution of particle concentration, and different harmonic information (properties of particles and system information). For the network model, the relationship between different harmonics can be fitted from time and frequency domain to recover the fundamental frequency component. From the simulation results, we have shown that ReFFNet provides a satisfactory performance on the recovery of fundamental frequency component.

However, our method remains highly dependent on data. Therefore, the design of digital filters can be optimized and a more complicated noise model can be created to distort the generated signal, making our network model more robust and generalized to different MPI systems. Besides, we will further study the effectiveness of different attention mechanisms and upgrade our network model. And we will consider the relationship between the two dimensions of time and frequency, and further model the long distance dependency in the global space. In addition, the acquisition method of spectrum will be studied to integrate more information, which is benefit for network training.

10.18416/ijmpi.2023.2303033

VI. Conclusions

We introduced a DL method for enhancing the MPI signal. Our model adopted the self-attention mechanism to model relationship between different harmonics, and used high-frequency harmonics information to recover the first harmonic. At the same time, we dealt with 2D time-frequency spectrum integrating the particle distribution information in time domain with the harmonics information in frequency domain. Our method was proved to be effective in recovering fundamental frequency component in the presence of background noise, and based on the restored spectrum, we can obtain high-quality MPI images.

Acknowledgments

The authors would like to acknowledge the instrumental and technical support of Multimodal Biomedical Imaging Experimental Platform, Institute of Automation, Chinese Academy of Sciences. Funding: This work was supported in part by the National Key Research and Development Program of China under Grant: 2017YFA0700401; the National Natural Science Foundation of China under Grant: 62027901, 81827808, 81227901; Beijing Natural Science Foundation JQ22023; CAS Youth Innovation Promotion Association under Grant 2018167 and CAS Key Technology Talent Program.

Author's statement

Conflict of interest: Authors state no conflict of interest.

References

[1] B. Gleich and J. Weizenecker. Tomographic imaging using the nonlinear response of magnetic particles. Nature, 435(7046):1214–1217, 2005.

[2] J. Weizenecker, J. Borgert, and B. Gleich. A simulation study on the resolution and sensitivity of magnetic particle imaging. Phys Med Biol, 52(21):6363-6374, 2007.

[3] P. W. Goodwill, E. U. Saritas, L. R. Croft, T. N. Kim, K. M. Krishnan, D. V. Schaffer, and S. M. Conolly. X-Space MPI: Magnetic Nanoparticles for Safe Medical Imaging. Adv Mater, 24(28):3870-3877, 2012.

[4] K. Lu, P. W. Goodwill, E. U. Saritas, B. Zheng, and S. M. Conolly. Linearity and Shift Invariance for Quantitative Magnetic Particle Imaging. IEEE transactions on medical imaging, 32(9):1565-1575, 2013.

[5] P. W. Goodwill and S. M. Conolly. The x-space formulation of the magnetic particle imaging process: 1-d signal, resolution, bandwidth, snr, sar, and magnetostimulation. IEEE transactions on medical imaging, 29(11):1851–1859, 2010.

[6] J. Rahmer, J. Weizenecker, B. Gleich, and J. Borgert. Analysis of a 3-D System Function Measured for Magnetic Particle Imaging. IEEE transactions on medical imaging, 31(6):1289-1299, 2012.

[7] J. Rahmer, J.Weizenecker, B. Gleich, and J. Borgert. Signal encoding in magnetic particle imaging: Properties of the system function. BMC medical imaging, 9(1):1–21, 2009.

[8] S. Kurt, Y. Muslu, and E. U. Saritas. Partial FOV Center Imaging (PCI): A Robust X-Space Image Reconstruction for Magnetic Particle Imaging. IEEE transactions on medical imaging, 39(11):3441-3450, 2020.

[9] Z. Wei, X. Wu, W Tong, S. Zhang, X. Yang, J. Tian, and H. Hui. Elimination of stripe artifacts in light sheet fluorescence microscopy using an attention-based residual neural network. *Biomed Opt Express*, 13(3):1292-1311, 2022.

[10] Y. Ko, S. Moon, J. Baek, and H. Shim. Rigid and non-rigid motion artifact reduction in X-ray CT using attention module. *Med Image Anal*, 67, 2021.

