Program Digest, 2019 IEEE International Ultrasonics Symposium (IUS)
Glasgow, Scotland, October 6-9, 2019

High Contrast and High Signal-to-noise Ratio Minimum Variance Beamforming Combined with Deep Neural Networks

Renxin Zhuang
School of Software Engineering
South China University of Technology
Guangzhou, Guangdong 510006, China
renxinzhuang@gmail.com

Junying Chen*
School of Software Engineering
South China University of Technology
Guangzhou, Guangdong 510006, China
jychense@scut.edu.cn

***Abstract*—This work combined neural network and traditional minimum variance (MV) beamforming to produce ultrasound images with high resolution, high contrast and relatively high speckle signal-to-noise ratio (SNR). The received channel data are in time domain and the neural network is a deep convolutional neural network. The network is responsible for suppressing off-axis scattering signals and the apodization weights of MV beamforming guarantee the resolution performance. Experimental results showed that the proposed method significantly improved the contrast performance of MV beamforming while reserving high lateral resolution and satisfactory speckle SNR.**

***Index Terms*—high contrast, minimum variance beamforming, convolutional neural network, off-axis suppressing, ultrasound imaging**

I. Introduction

In medical ultrasound B-mode imaging, delay-and-sum (DAS) beamforming [1] is the most widely used beamforming technique due to its real-time performance. It consists of two steps, the first of which is applying time delays to the received channel data to dynamically focus to imaging pixels and the second of which is summing the channel measurements to determine the pixel amplitude. The resolution performance of DAS beamforming is poor. Hence, minimum variance (MV) beamforming was proposed to improve the quality of ultrasound images [2]. MV beamforming complies to the framework of DAS beamforming. However, MV calculates the channel apodization weights and then applies weighting summation to the delayed channel data. The resolution performance of MV is impressive but the contrast and the robustness needs improving. Several techniques have been proposed to improve the contrast and the robustness of MV. Spatial smoothing devides the receiving aperture into sub-apertures, respectively calculates the covariance matrix of each sub-aperture and then averages all the covariance matrices. It enhances the estimation accuracy of the covariance matrix. Diagonal loading adds a proportion of Gaussian white noise to the estimated covariance matrix. These two techniques improve the robustness of MV.

In recent years, the combination of neural network and ultrasound beamforming has been a research hotspot. In 2018, Ye *et al.* utilized deep neural network to interpolate the missing channel data in the sub-sample channel data and then applied DAS beamforming to the interpolated channel data [3]. The neural network successfully learned the pattern of redundancy in channel data and achieved comparable performance for the full-sample case. Luchies and Byram designed a neural network beamformer to suppress off-axis scattering signals [4]. They first converted channel data into frequency domain using short-time Fourier transform and then operated the frequency domain channel data with a five-layer fully-connected neural network. Their method achieved much higher contrast than DAS beamformer and simutaneously reserved high speckle signal-to-noise(SNR) ratio. In 2019, Dahl *et al.* utilized convolutional neural network to significantly reduce the speckle and thus achieved high speckle SNR and high contrast [5].

In this paper, we proposed a high contrast high speckle signal-to-noise ratio minimum variance beamforming method combined with neural network. The neural network was trained to suppress the sidelobe and the apodization weights of MV guaranteed the resolution performance. The neural network operated delayed channel data in time domain, getting rid of the annoying Fourier transform which is necessary in [4]. The method we proposed significantly improved the contrast of images while maintaining the resolution performance of traditional MV and reserving the speckle SNR.

The rest of this paper is organized as followed. Section II introduces the detail of our method, including the network training and the combination of neural network and traditional MV. The experimental setup and result analysis will be presented in Section III. Finally, Section IV concludes the paper.

This work was supported in part by “National Natural Science Foundation of China” (No. 61802130), “Guangdong Natural Science Foundation” (No. 2018A030310355), and “Guangzhou Science and Technology Program” (No. 201707010223).

*Corresponding author.

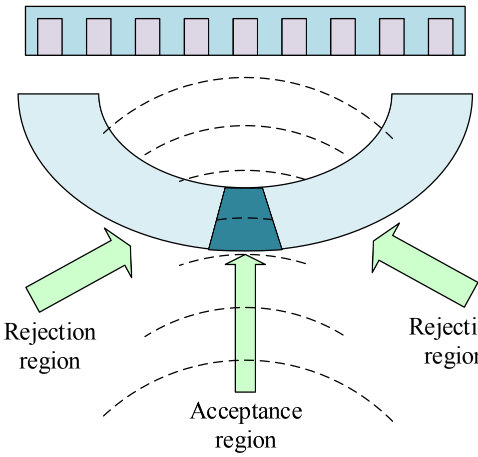


Fig. 1: Acceptance region and rejection region.

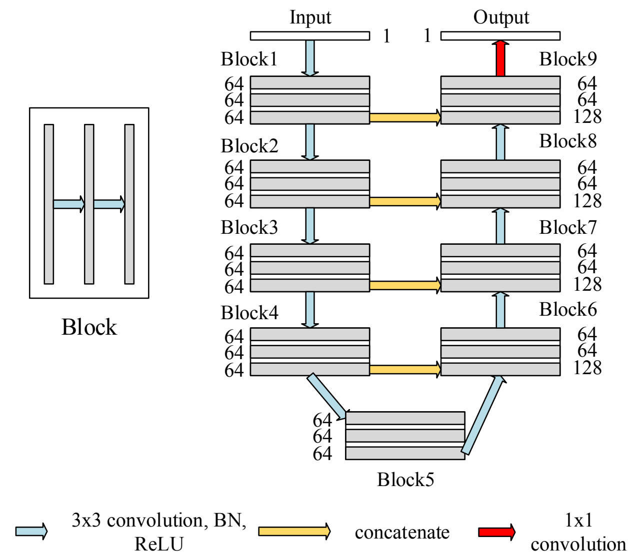


Fig. 2: The network architecture. A block is consisted of two sets of 3×3 convolution, batch normalization and ReLU.

II. Method

A. Neural Network Training

In order to train the neural network to distinguish off-axis scattering signals and on-axis scattering signals, a dataset with these two kinds of signals must be created. We use Field II to simulate the ultrasound channel data [6]. In the simulation phantom, the acceptance region and the rejection region are defined as shown in Fig. 1. The acceptance region is defined accrording to the beam width at the focal point. The width of the annular sector is 2 × wavelength. For the training dataset, 15,000 scatters are randomly positioned in the acceptance region and 15,000 scatters are randomly positioned in the rejection region. For the validation dataset, 1,500 scatters are randomly positioned in the acceptance region and 1,500 scatters are randomly positioned in the rejection region. For each scatter, a focused beam whose axis is in the middle of the aperture insonifies the imaging phantom and then the echo signals are recorded. For scatters in the acceptance region, the echo signals are on-axis signals so the output is exactly the same as input. But for scatters in the rejection region, the echo signals are off-axis signals so the output is zeros.

The network employed in this paper is a fully convolutional neural network, as shown in Fig. 2 . We expect that a convolutional network can capture the pattern of off-axis signals in time domain. Different from [4], no Fourier transform is executed so a five-layer fully-connected network may not be capable of suppressing off-axis signals. Our network is composed of 28 convolution layers and each layer consists of a 3×3 convolution, batch normalization and ReLU except for the last layer, which merely contains a 1×1 convolution operation. The learning rate is set to be 1.0×10^{-5} and the decay is 1.0×10^{-7}. We employ an early-stopping strategy that is when the validation loss does not improve after 20 epochs, the training will be terminated.

B. MV Beamforming with Neural Network

In focused ultrasound imaging, channel data are acquired scanline by scanline as the active aperture slides across the transducer surface. After applying time delays to the channel data, we can get a datacube of shape $P \times M \times SL$ where SL represents number of scanlines, M represents number of receiving channels and P represents number of imaging pixels of each scanline. The delayed channel data of each scanline are then used to compute apodization weights of all imaging pixels.

In our method, spatial smoothing and diagonal loading are utilized to guarantee the robustness of MV. A full aperture with M channels is divided into $M - L + 1$ sub-apertures with L channels. The channel data of each sub-aperture $\boldsymbol{x}_i, i = 1, 2, ..., M - L + 1$ contain signals from ith channel to $(i + L - 1)$th channel. Then, the data covariance matrix $\boldsymbol{R}_{cov}$ is calculated accroding to (1):

$$\boldsymbol{R}_{cov} = \frac{1}{M - L + 1} \sum_{i=1}^{M-L+1} \boldsymbol{x}_i \cdot \boldsymbol{x}_i^H. \tag{1}$$

Diagonal loading is subsequently executed. It adds a certain proportion of Gaussian white noise to covariance matrix $\boldsymbol{R}_{cov}$, as illustrated in (2):

$$\boldsymbol{R}_{cov} = \boldsymbol{R}_{cov} + \sigma \cdot trace(\boldsymbol{R}_{cov} \cdot \boldsymbol{I}), \tag{2}$$

where I is a unit matrix and σ is the diagonal loading factor which is $1/(100L)$. The apodization weight $\boldsymbol{w}$ is a vector with length L, calculated as (3):

$$\boldsymbol{w} = \frac{\boldsymbol{R}_{cov}^{-1} \cdot \boldsymbol{a}}{\boldsymbol{a}^H \cdot \boldsymbol{R}_{cov}^{-1} \cdot \boldsymbol{a}}, \tag{3}$$

where $\boldsymbol{a}$ is a vector of ones.

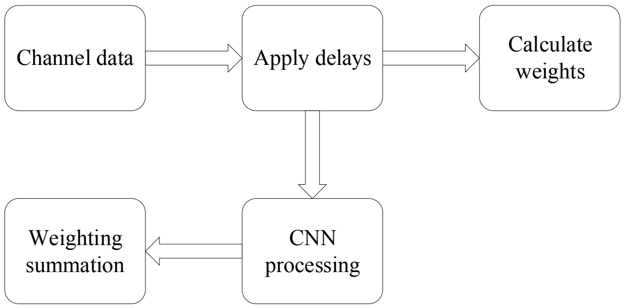


Fig. 3: The flowchart of our method.

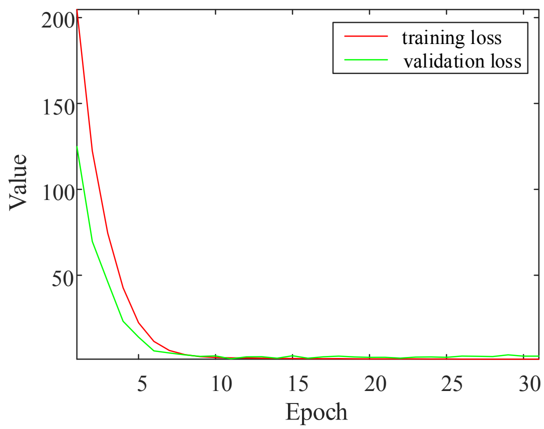


Fig. 4: The training results.

The delayed channel data of each scanline are then processed by the trained neural network to suppress off-axis scattering signals. Then, the pixel amplitude is computed by multiplying apodization weight $\boldsymbol{w}$ with every sub-aperture data processed by network $\boldsymbol{x}'_i, i = 1, 2, ..., M - L + 1$ and then the results are averaged:

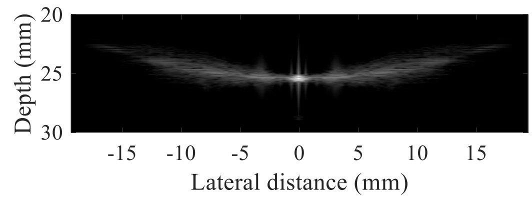
$$v = \frac{1}{M - L + 1} \sum_{i=1}^{M-L+1} \boldsymbol{w}^H \cdot \boldsymbol{x}'_i. \quad (4)$$

The overall procedure is illustrated in Fig. 3. Now beamforming is completed and after post-processing procedure such as envelop detecting and log compressing, an ultrasound image is produced.

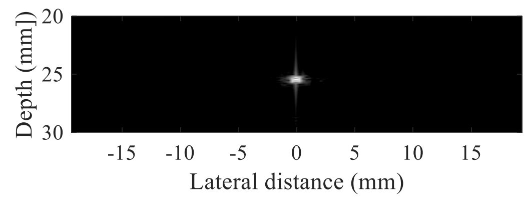
III. Experiments

A. Neural Network Training

The neural network was trained on an NVIDIA GeForce RTX 2080 Ti GPU. The training loss curve and the validation loss curve were plotted in Fig. 4. As shown, the training process ended at the 31th epoch and did not suffer from overfitting. The optimal validation loss was 1.38 achieved at the 11th epoch.



(a) Traditional MV



(b) CNNMV

Fig. 5: Point target images (60dB dynamic range) of traditional MV and the proposed CNNMV.

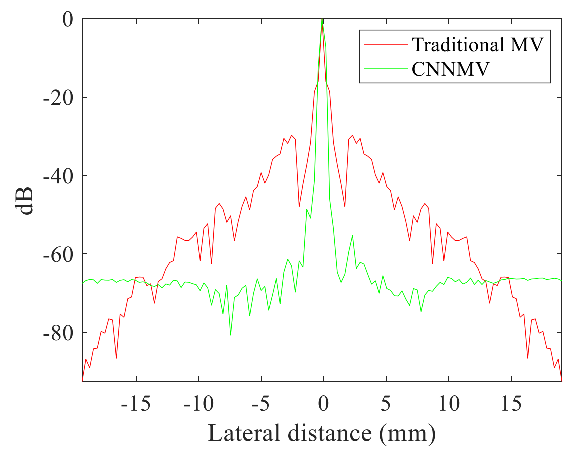


Fig. 6: Point spread function comparison of Traditional MV and the proposed CNNMV.

B. Imaging Scenario Simulation

We first tested the network performance on point target simulation. Fig. 5 showed point target images of traditional MV and the proposed CNNMV. It was clear that the sidelobe of the point target was largely suppressed in image of CNNMV. From point spread function comparison in Fig. 6, we could know that the lateral resolution performance of CNNMV suffered a little yet the sidelobe was greatly suppressed. This indicated that the network successfully learned the pattern of off-axis scattering and suppressed off-axis signals while reserving on-axis signals.

In cyst simulation, we quantified the imaging results using contrast (CR) and speckle signal-to-noise ratio (SNR). The cyst was centered in the focal depth that was 25mm and the radius of the cyst was 3mm. The phantom was filled with scatters with density of 100 scatters per cubic millimeter and

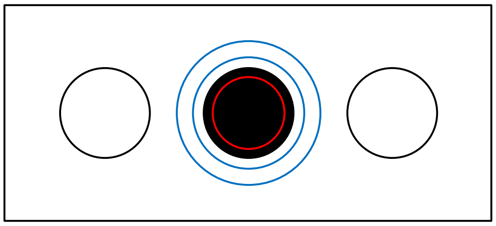


Fig. 7: The illustration of the calculation of contrast and speckle SNR.

TABLE I: Contrast and SNR Comparison Results

Metrics	Traditional MV	CNNMV
Contrast	0.5156	0.6607
Speckle SNR	1.1444	1.0858

the amplitudes of scatters within the cyst were set to be zeros. As shown in Fig. 7, the region of interest (ROI) of cyst was selected as a circle with a radius of 0.8 times the cyst radius. The ROI of background was selected as an annular with an inner radius of 1.1 times the cyst radius and an outer radius of 1.5 times the cyst radius. We calculated the mean pixel values of the cyst and the background and then computed the contrast value according to (5):

$$CR = \frac{M_{out} - M_{in}}{M_{out}}, \tag{5}$$

where M_{out} denotes mean pixel value of the background and M_{in} denotes mean pixel value of the cyst. Two circle background regions with the same area as the cyst were selected on the left side and the right side of the cyst. Then the speckle SNRs of these two regions were calculated according to (6), respectively:

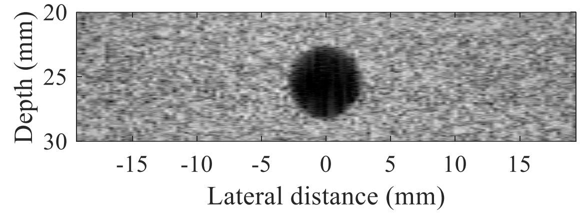
$$SNR = \frac{\mu_{background}}{\sigma_{background}}, \tag{6}$$

where $\mu_{background}$ denotes the mean value of uncompressed envelop signals and $\sigma_{background}$ denotes the standard deviation value of uncompressed envelop signals in the selected background region. Finally, the SNRs of the left circle and the right circle were averaged to get the desired SNR value.

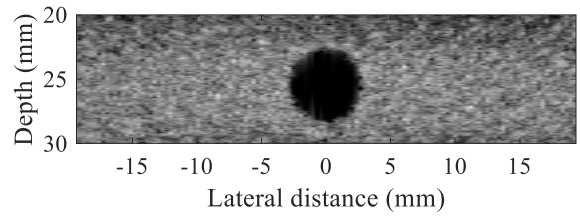
Cyst images of traditional MV and the proposed CNNMV were shown in Fig. 8. It was observed that the cyst of CNNMV was more darker, indicating that signals within the cyst region were suppressed. As shown in Table I, the CR value of traditional MV was 0.5156 while the CR value of CNNMV was 0.6607. The SNR of CNNMV was 1.0858, which was little lower than that of traditional MV, which was 1.1444. This phenomenon indicated that reconstrcuting signals was harder to learn than suppressing signals.

IV. Conclusion

In this work, we explored suppressing off-axis scattering signals in channel data in time domain using convolutional neural network. Experimental results showed that without any Fourier transform, convolutional neural network was capable of capturing the pattern of off-axis scattering in time domain. By combining neural network and traditional MV beamforming, the produced ultrasound images can be of high quality in terms of lateral resolution, contrast and speckle SNR. In future work, SNR should be further improved by enlarging the amount of scatters in the training dataset. Also, using neural network to suppress other sources of image degradation such as reverberation deserves further researching.



(a) Traditional MV



(b) CNNMV

Fig. 8: Cyst images (60dB dynamic range) of traditional MV and the proposed CNNMV.

References

[1] J. F. Havlice and J. C. Taenzer, "Medical ultrasonic imaging: An overview of principles and instrumentation," *Proceedings of the IEEE*, vol. 67, no. 4, p. 620–641, 1979.

[2] J. F. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 8, pp. 1606–1613, 2007.

[3] Y. H. Yoon, S. Khan, J. Huh, and J. C. Ye, "Efficient b-mode ultrasound image reconstruction from sub-sampled rf data using deep learning," *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 325–336, 2019.

[4] A. C. Luchies and B. C. Byram, "Deep neural networks for ultrasound beamforming," *IEEE Transactions on Medical Imaging*, vol. 37, no. 9, pp. 2010–2021, 2018.

[5] D. Hyun, L. L. Brickson, K. T. Looby, and J. J. Dahl, "Beamforming and speckle reduction using neural networks," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 5, pp. 898–910, 2019.

[6] J. A. Jensen, "Field: A program for simulating ultrasound systems," in *10th Nordicbalttc Conference on Biomedical Imaging*, 1996, pp. 351–353.