3-D contrast enhanced ultrasound imaging of an in vivo chicken embryo with a sparse array and deep learning based adaptive beamforming

Ossenkoppele, Boudewine W.; Wei, Luxi; Luijten, Ben; Vos, Hendrik J.; De Jong, Nico; Van Sloun, Ruud J.G.; Verweij, Martin D.

DOI
10.1109/IUS54386.2022.9957383

Publication date
2022

Document Version
Final published version

Published in
IUS 2022 - IEEE International Ultrasonics Symposium

3-D contrast enhanced ultrasound imaging of an in vivo chicken embryo with a sparse array and deep learning based adaptive beamforming

Boudewine W. Ossenkoppele*, Luxi Wei†, Ben Luijten‡ and Hendrik J. Vos*†, Nico de Jong *†,
Ruud J.G. van Sloun‡§, Martin D. Verweij*†
*Dept. of Imaging Physics, Delft University of Technology, The Netherlands
†Dept. of Cardiology, Erasmus MC Rotterdam, 3015 CN Rotterdam, The Netherlands
‡Dept. of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
§Philips Research, Eindhoven, The Netherlands

2022 IEEE International Ultrasonics Symposium (IUS) | 978-1-6654-6657-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IUS54386.2022.9957383

***Abstract*—3-D contrast enhanced ultrasound enables better visualization of inherently 3-D vascular geometries compared to an intersecting plane. Additionally, it would allow the application of motion correction techniques for all directions. Both contrast detection and motion correction work better on high-frame rate data. However high-frame rate 3-D ultrasound imaging with dense matrix arrays is challenging to realize. Sparse arrays alleviate some of the limitations in cable count and data rate that fully populated arrays encounter, but their increased level of secondary lobes negatively impacts image contrast. Meanwhile the use of unfocused transmit beams needed to achieve high-frame rates negatively impacts resolution. Here we propose to use adaptive beamforming by deep learning (ABLE) to improve the image quality of contrast enhanced ultrasound images acquired with a sparse spiral array. We train the neural network on simulated data and evaluate simulated images and *in vivo* images of an *ex ovo* chicken embryo. ABLE improved resolution compared to delay-and-sum (DAS) and spatial coherence (SC) beamforming on the simulated and *in vivo* data. The qualitative improvements persist after histogram matching, indicating that the image quality improvement of the ABLE images was not purely due to dynamic range stretching.**

***Index Terms*—contrast enhanced ultrasound, spiral array, sparse array, deep learning, beamforming**

I. Introduction

Contrast enhanced ultrasound imaging (CEUS) has enabled the visualization of the microvasculature, which used to be difficult to distinguish from tissue due to the slow blood velocity and the weak scattering of red blood cells [1]. Three-dimensional images are highly desired for contrast enhanced ultrasound, because they enable better visualization of the vasculature compared to 2-D images and allow for motion correction in all directions. High-frame rates are not just desirable in scenarios with a high bloodflow velocity, but also allow more effective application of postprocessing filters, better motion compensation and lower acquisition times.

Sparse arrays are an attractive option to realize high-frame rate volumetric imaging due to their relatively low cable count, data rate and cost. Although their transmit efficiency and receive sensitivity is lower compared to fully populated arrays, the high reflectivity of microbubbles makes this less of a concern in contrast enhanced imaging than in imaging without contrast. However, variation of magnitude and pulse shape within the emitted field is higher than with dense matrix arrays and the high-amplitude secondary lobes of a sparse array can cause artifacts when beamforming with Delay-and-Sum (DAS) while reducing contrast. Meanwhile the unfocused transmit beams needed for high-frame rate imaging negatively impact resolution. Furthermore, a larger PSF size resulting in more overlapping microbubbles and a greater variation in PSF shape also hamper the application of ultrasound localization microscopy techniques.

Spatial coherence (SC) beamforming has been used to improve image quality in contrast enhanced ultrasound images acquired by a sparse array [2]. Through exploiting the coherence of signals, the level of microbubble clutter was reduced. Others have proposed to use minimum-variance beamforming (MV) [3] instead of delay-and-sum specifically for improving the separation of close-by microbubbles, i.e. resolution. However, minimum-variance beamforming is computational expensive and not well established for volumetric imaging with sparse arrays.

Adaptive beamforming by deep learning (ABLE) [4] has a lower computational cost than minimum variance beamforming and has been applied to improve ultrasound image quality in non-contrast enhanced images for arrays with regularly spaced elements and has been used jointly with deep learning-based localization [5] to improve the localization of microbubbles in contrast enhanced ultrasound. Inspired by this we hypothesize that ABLE can be used to improve contrast enhanced imaging with a sparse array. We trained an adapted version of the original ABLE network on *in silico* data and compared beamformed images of an *in vivo* chicken embryo, acquired with single pulse detection at the fundamental frequency, to DAS and SC beamforming.

This work is part of the 3-D ICE project (STW project 14279) which is financed by the Netherlands Organization for Scientific Research (NWO).

II. Methods

A. Sparse spiral array

We used a 5 MHz prototype sparse array with a 1.6 cm diameter aperture, consisting of 256 piezoelectric elements with an element size of 200 µm x 200 µm. The elements are arranged in a tapered spiral pattern, specifically designed to minimize sidelobe levels of the spiral array by decreasing the element density towards the edge of the array according to a Blackman window [6], [7] (see Fig. 1).

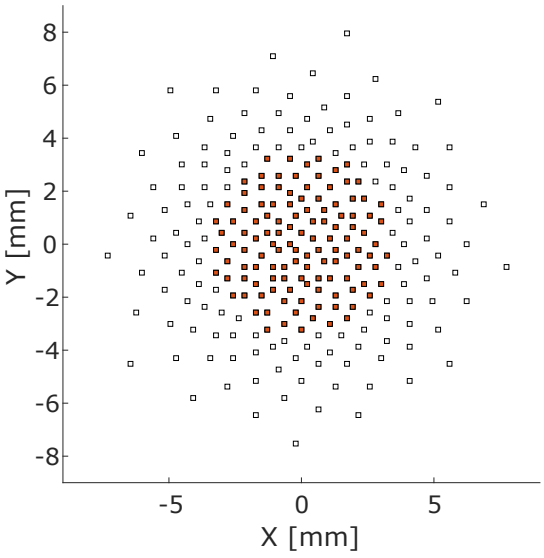

Fig. 1. Layout of the 256 elements of the sparse spiral array. Only the center 120 elements indicated in red are used in transmit. All elements are used in receive.

B. Imaging scheme

Diverging waves with a 30° opening angle were used to insonify a large region of interest. The final volumetric images were formed by angular compounding of five steered diverging waves (angles: [0°, 5°], [5°, 0°], [0°, -5°], [-5°, 0°], [0°, 0°]) to reach a volumetric frame rate of 1000 Hz. To increase the uniformity of the transmit field in both lateral and axial direction, only the center 120 elements were used in transmission (see Fig. 1). Further details on the probe and the design of the transmit and receive strategy can be found in [2].

C. In vivo data acquisition

An *ex ovo* chicken embryo and chorioallantoic membrane (CAM) were used as the *in vivo* model to investigate the effect of the different beamformers on contrast enhanced ultrasound imaging. The five-day-old chicken embryo and CAM were removed from the eggshell and further prepared according to the protocol of [8], after which four μL of custom-made F-type microbubble was injected into the vasculature of the chicken embryo. The chicken embryo was placed in a PBS solution and the ultrasound array was positioned on top of the chicken embryo at a 3 cm distance. More details of the procedure can be found in [2], [8]. The ultrasound recording lasted 3.9 s. Before off-line beamforming of the data, it was pre-processed with an SVD filter along the full frame length of 3900 frames to remove the quasi-stationary tissue and background signal. The RF data of each transmit angle was filtered separately and the twenty lowest ranks were removed.

D. Network architecture, training procedure and training data

The neural network architecture is shown in Fig. 2. For each of the five transmit angles apodization weights were calculated from the time-of-flight (TOF) corrected RF data. After pixel-wise multiplication of the apodization weights with the TOF data and subsequently summing across the channel dimension, the resulting five voxel values were compounded. Unlike in the original ABLE beamformer the voxels were then envelope detected and finally processed by a final 1-D convolutional layer extending in the axial direction.

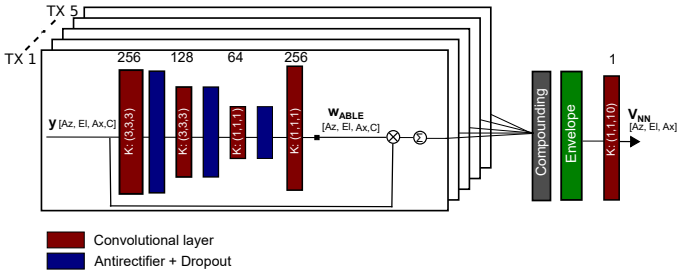

Fig. 2. Network architecture. The number of filters for each convolutional layer are indicated above the layer, the kernel size is indicated in the layer. Five parallel branches are used to calculate apodization weights for the TOF corrected RF data of each transmit angle.

Field II [9] was used to simulate microbubble data acquired by the sparse array. Point scatterers with an amplitude between 0.5 and 1 were placed in a region spanning from -40° to 40° in azimuth and elevation direction. The mean concentration was 0.24 microbubbles per mm^3. For a single image five angled diverging waves were transmitted in the same directions as in the *in vivo* recording. The true location and amplitude of the point scatterers were used to form a target image. During training image patches of 11 $\times$11$\times$100 voxels were randomly selected from a volume at 0.15 to 0.45 cm distance from the transducers extending -20° to 20° in azimuth and elevation. The 78 frames used resulted in 647 distinct image patches. The network was trained with the Adam [10] optimizer to minimize the signed-mean-squared-logarithmic-error between the network output images and a Gaussian filtering of the images containing the true point locations:

$$\mathcal{L}_{\text{SMSLE}}(y,x,\sigma) = \frac{1}{I}\sum_{i=1}^{I}\Big(\frac{1}{2}||\log_{10}(V_{\text{NN}}^{(i)})^{+} - \log_{10}(V_{\text{Target}}^{(i)})^{+}||_2^2 + \frac{1}{2}||\log_{10}(V_{\text{NN}}^{(i)})^{-} - \log_{10}(V_{\text{Target}}^{(i)})^{-}||_2^2\Big). \tag{1}$$

Here $V_{\text{NN}}^{(i)}$) refers to the network output for voxel i and $V_{\text{Target}}^{(i)}$ to the target voxel.

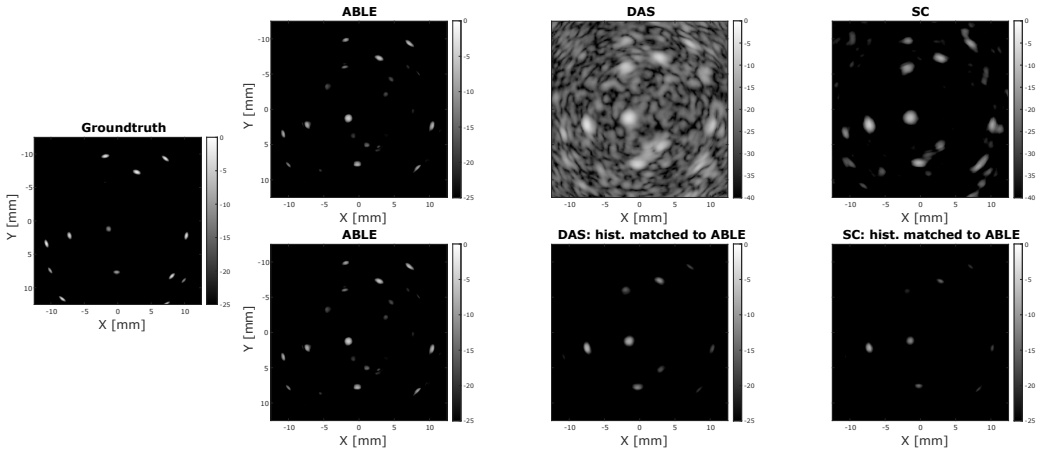

Fig. 3. C-planes of a volume of simulated point scatterers. *Left)* The ground truth is visualized by a Gaussian filtering of the point scatterers. *Right)* The beamformed images are shown for ABLE, DAS and SC. *Top row)* No further processing is applied. *Bottom row)* Beamformed images after matching the histograms of the SC and DAS beamformed images towards that of ABLE.

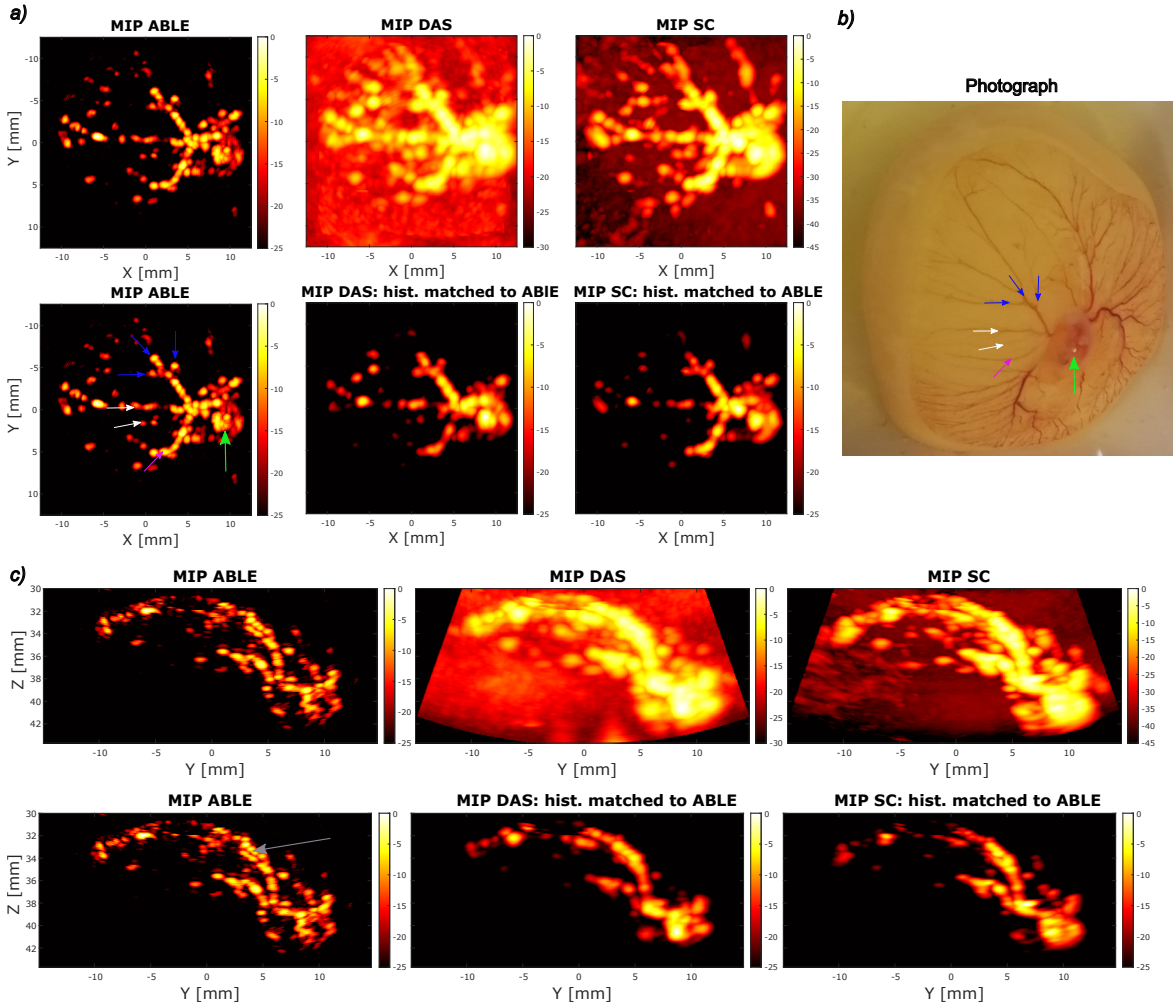

Fig. 4. *a)* The first row shows the maximum intensity projection (MIP) image in the XY plane for the ABLE, DAS and SC beamformer. The second row shows the MIP image in the XY plane for the ABLE, DAS and SC beamformers after the histograms of the SC and DAS images were matched to the histogram of the ABLE image. The colored arrows indicate vessels (blue, white, purple) and the heart (green) that are also indicated in the photograph shown in *b)*. *c)* The first row shows the MIP image in the YZ plane for the ABLE, DAS and SC beamformers. The second row shows the MIP image in the YZ plane for the ABLE, DAS and SC beamformers after the histograms of the SC and DAS image were matched to the histogram of the ABLE image.

E. Comparison to other beamformers

1) Beamformers: The DAS results were calculated with rectangular apodization. The average spatial coherence value was calculated on the TOF corrected RF data of each voxel, using all transducer channel pairs according to:

$$V_{SC} = \frac{\sum_{c=1}^{C-1} \sum_{l=1}^{C-c} y(c) y^*(c+l)}{\sum_{c=1}^{N} |y(c)|^2} \tag{2}$$

Here $y(c)$ is the TOF corrected signal of element c, C is the total number of elements in the array, l is the lag in number of elements and * indicates the complex conjugate.

2) Histogram matching: To make a fair visual and quantitative comparison between the performance of different beamformers the effect of dynamic range compression and displayed dynamic range must be separated from the information content in the image. Bottenus *et al.* [11] have proposed to separate the structural changes from the images respective embeddings by placing the images all under the same embedding through histogram matching. The Matlab implementation made available by [11] was used to implement full histogram matching of both DAS and SC images towards the reference ABLE method. Full histogram matching was used since this resulted in a better match of overall appearance than partial histogram matching. The number of bins was set to 256.

III. Results

A. Simulated data

Fig. 3 shows a C-plane slice of a volumetric image of simulated data from the test set. Without histogram matching the DAS image shows more background clutter compared to the ABLE and SC image. The ABLE image has smaller PSFs than the DAS and SC images. After histogram matching the DAS and SC image to ABLE and displaying them in the same dynamic range (see Fig. 3), several simulated microbubbles are not clearly visible in the DAS and SC image where they are visible in the ABLE image. However, the ABLE image also shows some bubbles that are not present in the groundtruth image slice. This is largely the result of the axial PSF size of ABLE being larger than that of the groundtruth image. Thus, purely transforming the dynamic range with which the DAS and SC image are displayed does not result in the same improvements seen for the ABLE beamformer.

B. In vivo chicken embryo

Fig. 4 shows the maximum intensity projection (MIP) of 800 subsequent frames that were beamformed with DAS, SC and ABLE and subsequently averaged over time. After matching the histograms of the DAS and SC images to that of the ABLE image, visually the contrast between the vessels and the background is similar. However, not all details that are present in the ABLE image are visible in the DAS and SC images (see arrow and caption). Furthermore, the resolution is higher in the ABLE images, such that there is an increased separation of the vessels.

IV. Conclusion

In this work we proposed to use an adaptive beamforming by deep-learning, to improve image quality of contrast enhanced ultrasound images acquired with a sparse spiral array. A neural network was used to calculate data-adaptive apodization weights for the TOF corrected RF data received from each transmit angle and then compute the final voxel values after compounding and envelope detection. The neural network was trained on simulations that used point scatterers to emulate microbubbles. ABLE improved resolution compared to DAS and SC beamforming on the simulated data and on *in vivo* data. The improvements of the ABLE images could not be replicated in the SC and DAS image by matching their histograms to the ABLE images. This indicates that it is the beamforming method that results in the image improvements, rather than a change in the dynamic range used for image display.

Acknowledgment

The authors want to acknowledge Bram Meijlink and Klazina Kooiman for their critical work on the chicken embryo setup.

References

[1] D. Cosgrove and N. Lassau, "Imaging of perfusion using ultrasound," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 37, no. SUPPL. 1, pp. 65–85, 2010.

[2] L. Wei, G. Wahyulaksana, B. Meijlink, A. Ramalli, E. Noothout, M. D. Verweij, E. Boni, K. Kooiman, A. F. V. D. Steen, P. Tortoli, N. D. Jong, and H. J. Vos, "High frame rate volumetric imaging of microbubbles using a sparse array and spatial coherence beamforming," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 68, pp. 3069–3081, 2021.

[3] K. Diamantis, T. Anderson, M. B. Butler, C. A. Villagomez-Hoyos, J. A. Jensen, and V. Sboros, "Resolving ultrasound contrast microbubbles using minimum variance beamforming," *IEEE Transactions on Medical Imaging*, vol. 38, pp. 194–204, 2019.

[4] B. Luijten, R. Cohen, F. J. D. Bruijn, H. A. Schmeitz, M. Mischi, Y. C. Eldar, and R. J. V. Sloun, "Adaptive ultrasound beamforming using deep learning," *IEEE Transactions on Medical Imaging*, vol. 39, pp. 3967–3978, 2020.

[5] J. Youn, B. Luijten, M. Schou, M. B. Stuart, Y. C. Eldar, R. J. V. Sloun, and J. A. Jensen, "Model-based deep learning on ultrasound channel data for fast ultrasound localization microscopy," *IEEE International Ultrasonics Symposium, IUS*, pp. 21–24, 2021.

[6] A. Ramalli, E. Boni, A. S. Savoia, and P. Tortoli, "Density-tapered spiral arrays for ultrasound 3-d imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 62, pp. 1580–1588, 2015.

[7] H. J. Vos, E. Boni, A. Ramalli, F. Piccardi, A. Traversi, D. Galeotti, E. C. Noothout, V. Daeichin, M. D. Verweij, P. Tortoli *et al.*, "Sparse volumetric pzt array with density tapering," in *2018 IEEE International Ultrasonics Symposium (IUS)*. IEEE, 2018, pp. 1–4.

[8] B. Meijlink, I. Skachkov, A. F. van der Steen, N. de Jong, and K. Kooiman, "The preparation of chicken ex ovo embryos and chorioallantoic membrane vessels as in vivo model for contrast-enhanced ultrasound imaging and microbubble-mediated drug delivery studies," *Journal of Visualized Experiments*, vol. 2021, pp. 1–27, 2021.

[9] J. Jensen, "Field: A program for simulating ultrasound systems," *Citeseer*, vol. 34, pp. 351– 353, 1996. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.50.4778&rep=rep1&type=pdf

[10] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15, 2015.

[11] N. Bottenus, B. C. Byram, and D. Hyun, "Histogram matching for visual ultrasound image comparison," vol. 68, pp. 1487–1495, 2021.