2Pa2-5

Deep learning-based digital refocusing in acoustic microscope

Himanshu Singh[1†], Kaushik Shukla[2], Azeem Ahmad[3], Prakhar Kumar[2], Frank Melandsø[3] and Anowarul Habib[3]
[1]Dept. of Civil. Eng., Indian Institute of Technology Guwahati, Asam, India
[2]Dept. of Elect. and Comms. Eng., Indian Institute of Technology, Dhanbad, India
[3]Dept. of Phys. and Tech., UiT The Arctic Univ. of Norway, Norway

1. Introduction

Scanning acoustic microscopy (SAM) is a label-free imaging technique which is capable of visualizing the surface and sub-surface structure from materials to biological samples [1-4]. Conventional SAM employs high frequency acoustic waves that propagate into the sample and reflects back at various speeds depending upon the stiffness of the sample. In high frequency SAM, depth of focus depends on frequency and curvature of the transducer element. In high frequency imaging, one of the big challenges is focusing of the samples. Therefore, the diffraction characteristic of the ultrasound signals lowers the lateral resolution of produced images. For a better visual representation of the acquired images, the reduced lateral resolution in the out-of-focus region needs to be addressed. In such a situation, the image quality depends upon how far the ultrasound waves travel above or below the focal plane. A wide range of methods has been adopted to correct the focusing issue of the sample.

In acoustic imaging, synthetic aperture focusing technique (SAFT) that was developed in the 1970's is one of the most versatile techniques to solve the focusing issue [5]. The optical holography and synthetic aperture radar technologies served as inspiration for the ultrasonic SAFT reconstruction. There are some shortcoming in SAFT for reconstructing the SAM images- like it generates line artifacts in the reconstructed images. It also requires defocusing distance of the sample to focal point. To overcome such issue, we proposed a physics based deep learning (DL) approach to effectively refocus the images. Point spread function (PSF) has been obtained from COMSOL Multiphysics, which was used to defocus the images. In this paper we demonstrated 1 mm out-of-focus image can be corrected by employing DL.

Email: anowarul.habib@uit.no

2. Experimental setup:

2D axisymmetric-based FEM simulation of the spherical ultrasonic transducer of 50 MHz was performed in commercially available software COMSOL Multiphysics. The information about the lens diameter, focal distance in water and transducer material was provided by the manufacturer. The radius of curvature R of the concave section of the spherical lens was estimated using the focal distance F in water. The material properties used in the transducer modeling are taken from the literature[2]. To replicate the experimental conditions, the bottom portion of the transducer is considered to be water and the top portion is composed of sapphire (Al_2O_3). A thin layer ($\lambda/4$ thickness) of acoustic anti-reflection glass coating is considered between them because of the significant impedance mismatch between the sapphire and the water, where λ is the wavelength of the acoustic wave in the glass.

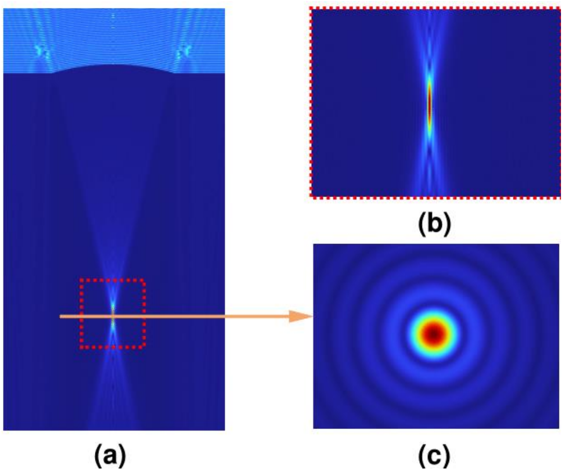


Fig. 1. (a) Simulated beam formation by the 50 MHz ultrasonic transducer showing the point focus at the optical axis (XZ-plane). (b) Enlarged view of the focal zone. (c) 2D view of the point spread function (PSF) at the focus point (XY-plane).

A frequency domain study is performed, and a cut plane is then drawn at the point of focus. The point spread function (PSF) data is then exported at the focus point and at 1 mm on both sides of the focus point.

3. Results and discussions

Autoencoder introduced by Vincent, is an unsupervised deep learning algorithm that leverages deep neural networks for dimensionality reduction and feature extraction [6]. It learns to compress the input representation and learns the subsequent reconstruction of the input. Adversarial Autoencoder (AAE) is a clever idea of blending the autoencoder architecture with the adversarial loss concept introduced by Generative Adversarial Network (GAN) [7]. A VGG-16 network inspired autoencoder with skip connections from input to all subsequent layers, and between symmetrical layers in encoder and decoder, is used as the generator model (Fig. 2). We have used the same discriminator model as used in SRGAN super resolution GAN model with

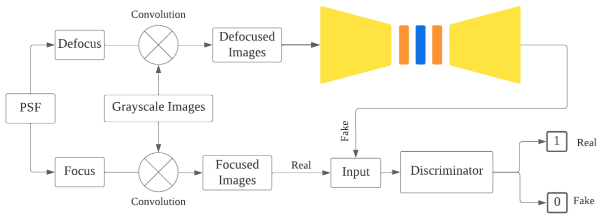
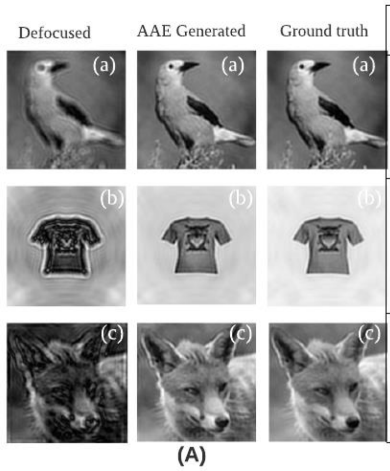


Fig. 2. Flowchart of the dataset generation to model training

80,000 grayscale images of different real-life objects, each of size 256 × 256 pixels is normalized in the range [0,1] [8].



	PSNR	SSIM
(a)	28.56	0.9570
(b)	29.43	0.9703
(c)	24.06	0.9331

(B)

Fig. 3A. Defocused, AAE generated focused and ground truth focused images of three different objects. **Fig. 3B.** PSNR and SSIM values of AAE generated focused images.

Each image is further convolved with the PSFs at defocus to produce defocused images and at focus to obtain the corresponding focused image (ground truth). Thus, the final dataset D consists of 80,000 defocus-focus image pairs.

An AAE is trained in the batches of 8 defocus-focus image pairs, sampled randomly from D for 10000 iterations, with VGG loss as perceptual loss using Adam optimizer at learning rate of 0.0001[9]. The model is validated over a separate test set containing 20,000 defocus-focus image pairs obtained using the same methodology as the training set. PSNR and SSIM are used as evaluation metrics for the quantitative assessment. Both metrics provide different aspects of requirement. PSNR provides higher visual interpretation, whereas the SSIM can be used to measure much finer similarity. Fig. 3 summarizes the result with PSNR higher than 25 & SSIM greater than 90% on average.

4. Conclusion

In this study, deep learning-based adversarial autoencoders have been exploited to refocus the images obtained through acoustic microscopy. Defocus-focus image pairs obtained through PSF convolution are used to train the model. PSNR and SSIM metrics show that the model performs exceptionally well, achieving state-of-the-art results and manifesting AI's immense potential in digital refocusing.

Acknowledgment
UiT funded Tematisk Satsinger project VirtualStain (Cristin Project ID: 2061348)

References

1) A. Habib, J. Vierinen, A. Islam, I.Z. Martinez and F. Melandsø: IEEE Int. Ultrason. Symp., 2018, p. 1.
2) M. Hofmann, R. Pflanzer, A. Habib, A. Shelke, J. Bereiter-Hahn, A. Bernd, R. Kaufmann, R. Sader and S. Kippenberger: Transl. Oncol. **9** (2016)179.
3) S. Wagle, A. Habib and F. Melandsø: Jpn. J. Appl. Phys. **56** (2017)07JC05.
4) A. Habib, A. Shelke, M. Vogel, U. Pietsch, X. Jiang and T. Kundu: Ultrasonics. **52** (2012)989.
5) J.J. Flaherty, K.R. Erikson and M. Van Lund: (1970)
6) P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol and L. Bottou: J. Mach. Learn. Res. **11** (2010).
7) I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio: NeurIPS Proceedings. **27** (2014).
8) C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz and Z. Wang: CVPR, 2017, p. 4681.
9) J. Johnson, A. Alahi and L. Fei-Fei: ECCV, 2016, p. 694.