A deep learning approach to ultrasound image recovery

Dimitris Perdios[1], Adrien Besson[1], Marcel Arditi[1], Jean-Philippe Thiran[1,2]; [1]Signal processing laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne, Lausanne, VD, Switzerland, [2]Department of Radiology, University Hospital Center (CHUV) and University of Lausanne (UNIL), Lausanne, VD, Switzerland

Background, Motivation and Objective

Compressed sensing (CS) has drawn many interest in the field ultrasound (US) image recovery. It has demonstrated promising results in the recovery of radio-frequency element raw-data [Liebgott *et. al.* ULTRAS13, Besson *et. al.* SPARS17]. The objective of such approaches is to recover the raw-data from undersampled random measurements. It is achieved by means of convex optimization or greedy methods which require a high number of iterations to converge and whose accuracy highly depends on hyper-parameters fine-tuning. Recently, deep-neural networks (DNN) has redefined the paradigm of signal recovery, leading to remarkable results for CS reconstruction of natural images [Mousavi *et. al.* ALBERTON15].

Statement of Contribution/Methods

Inspired by this success, we propose to use a DNN based on stacked denoising autoencoders for US image recovery. The network is composed of 4 layers (*i.e.* encoding - decoding - encoding - decoding). Each layer acts as $\mathbf{y} = F(W\mathbf{x} + \mathbf{b})$, where $\mathbf{x}$ is the input, $\mathbf{y}$ the output, W the weight matrix, $\mathbf{b}$ the bias and F a non-linearity function. For each encoding (decoding) layer, W is a $M \times N$ ($N \times M$) matrix with $M < N$. We use tanh as non-linearity function since US raw-data are zero-centered. During the training, the weights of the 4 layers are learned by minimizing a l_2-loss function on the training set. Once trained, the 1st layer is used to compress the US raw-data, and the remaining layers are used for the recovery. The dataset is simulated using the open source k-Wave toolbox [Treeby *et. al.* JBO10] on a configuration mimicking an acquisition system. The phantoms and medium properties are randomly generated from typical tissue zones (*i.e.* speckle, bright reflectors, inclusions). The simulation accounts for the attenuation and element directivity.

Results/Discussion

We generated the datatest based on a Verasonics system and a L12-5 probe (128 elem., 195µm pitch, 5MHz US freq., 31.2MHz sampl. freq.) with 1 plane wave insonification. We trained the network (100 epochs) over 800000 raw-data traces using Adam optimizer with a learning rate of 0.01. While requiring a fraction of the reconstruction time, Fig. (a) shows that the proposed SDA is competitive and in some cases outperforms state-of-the-art CS algorithm in terms of PSNR on both *in-vitro* and *in-vivo* acquisitions. Fig. (b) shows the reference phantom image (100% of the data) and (c) the image recovered from 15% of the data.