Compressed sensing reconstruction of 3D ultrasound data using dictionary learning and line-wise subsampling

Oana Lorintiu*, Hervé Liebgott*, Martino Alessandrini†, Olivier Bernard*, Denis Friboulet*

* Université de Lyon, CREATIS ; CNRS UMR5220 ; Inserm U1044 ; INSA-Lyon ; Université Lyon 1, France

†Cardiovascular Imaging and Dynamics, KULeuven, Leuven, Belgium

***Abstract*—In this paper we present a compressed sensing (CS) method adapted to 3D ultrasound imaging (US). In contrast to previous work, we propose a new approach based on the use of learned overcomplete dictionaries. Such dictionaries allow for much sparser representations of the signals since they are optimized for a particular class of images such as US images. In this study, the dictionary was learned using the K-SVD algorithm on patches extracted from a training dataset and the reconstruction was performed from 3D volumes not included in the training dataset. In each case, CS reconstruction was performed on the non-log envelope data by removing 20% to 80% of the original samples and the accuracy of the reconstruction was evaluated in terms of the normalized root mean square error relative to the original volume. Using numerically simulated data, we evaluate the influence of the training parameters and the influence of the sampling strategy. The latter is done by comparing the two most common sampling patterns, i.e. point-wise and line-wise random patterns. The results show in particular that line-wise sampling yields an accuracy comparable to the conventional point-wise sampling. This indicates that CS acquisition of 3D data is feasible in a relatively simple setting, and thus offers the perspective of increasing the frame rate by simply skipping the acquisition of many lines among the several thousands required in 3D imaging. We then evaluate the approach on US volumes of several *ex vivo* and *in vivo* organs. We first show that the learned dictionary approach yields better performances than conventional sparsifying dictionaries based on fixed transforms such as Fourier or discrete cosine. Finally, we investigate the generality of the learned dictionary approach and show that it is possible to build a general dictionary allowing to reliably reconstruct different volumes of different *ex vivo* or *in vivo* organs. The difference between the reconstruction error obtained with a specific dictionary and the one obtained with the general dictionary is minimal (on the order of 10^{-4}).**

***Index Terms*—Compressed sensing, 3D ultrasound, overcomplete dictionaries, K-SVD, sparse representation**

I. Introduction

Traditional imaging modalities, like ultrasound (US) echography, rely all on Shannon's theorem that fixes the limit for the sampling frequency of a signal to twice its highest frequency component. In order to avoid artifacts and to respect Shannon's theorem, US devices use a sampling rate that is at least four times the central frequency of the emitted pulse. Consequently, when sampling at such rates, the amount of data obtained is large, especially in 3D imaging, and can impair real-time imaging or data transfer. In 3D US imaging, the number of radiofrequency (RF) lines that must be acquired to sweep the whole volume can be extremely high, typically several thousands. The acquisition time of one RF line is related to the speed of sound and the depth of investigation and cannot be compressed, leading to long acquisition times and thus low frame rate (several Hz) which limits acquisitions for dynamic organs, such as the heart. In this context, the recently introduced compressed sensing (CS) theory offers the perspective of reducing the amount of data acquired. CS is based on the idea that it is possible, under certain assumptions, to recover a signal sampled below the Shannon sampling limit [1]–[3]. Its application to medical ultrasound imaging is promising thanks to its capability to reduce the volume of acquired data and thus to speed up the acquisitions and increase the imaging rate of 3D US devices.

The application of compressed sensing techniques to US imaging has motivated research efforts only recently, since the first works were published in 2010 [4]–[6] and therefore few studies have been devoted to this topic to date. A key element in CS is that the data to be reconstructed should have a sparse expansion in some basis or dictionary. As a consequence, one important feature of the existing studies is the choice of the representation where the US data are assumed to be sparse. Several works [7]–[11] assume a sparse distribution of scatterers in the direct, spatial domain. In [12], the authors assume that the US signals themselves are sparse and characterized through α-stable distributions. Most of the previous works however consider that the signal is sparse in a given basis or frame such as Fourier [4], [13]–[17], wavelets [4], [6], [16], [18], Discrete Cosine [6] or wave atoms [4], [16].

Another important feature for ensuring a successful CS reconstruction is the way the measurement of the data is performed. While most of the studies rely on conventional CS performed in the discrete domain, let us mention the works in [8], [9] which are particular inasmuch as they use an analog-to-digital version of CS based on the so-called Xampling mechanism [19] and considering the US data as

Finite Rate of Innovation signals [20]. As far as discrete CS is considered, early works showed that linear projection of the data on sub-Gaussian random matrices (e.g. with i.i.d. Gaussian, Rademacher or Bernoulli entries) provides incoherent measurements allowing for exact reconstruction [21]. In practice however, the use of random sensing matrices is not straightforward, since it implies dedicated acquisition hardware. Such sensing matrices have been thus mostly used in simulation studies, such as [10], [12], [13] with Gaussian matrices and [11] with a Rademacher matrix.

Since using random sensing matrices is difficult in practice, it is often preferred to resort to deterministic, structured sensing matrices. In such case, successful CS reconstruction of US images[1] implies uniform random sampling of the data [21].

Let us end this brief outline by noting that there are very few studies dealing with 3D US data. In [6], [22] point-wise and line-wise sampling patterns have been proposed and compared on one experimental RF image using Daubechies wavelets, finite differences and discrete cosine as sparsifying bases. Very recently Birk et al. [23] extended the analog-to-digital Xampling mechanism to 3D acquisitions.

The objective of this study is to investigate in details the feasibility of compressive sensing 3D ultrasound. Preliminary results [24], [25] were presented in recent conference papers. In the above detailed context, this study introduces three main novelties and contributions :

- First, we address CS reconstruction of 3D US images using learned overcomplete dictionaries since they can be optimized for a particular class of images such as US images and thus allow for sparser representations. We show in particular that the learned dictionary approach yields better performances than conventional sparsifying dictionaries based on fixed transforms such as Fourier or discrete cosine. To the best of our knowledge, this aspect has never been investigated for 3D ultrasound. The only study in the literature that might be compared to our work is reported in [26], which deals with US tomography. The technique that we propose is different as we work on 3D ultrasound echography.
- Secondly, we evaluate the generality of the learned dictionary approach. Indeed this approach may be hampered by overlearning, yielding a learned dictionary providing satisfying reconstruction only for images very close to the learning set. We show that it is possible to build a dictionary allowing to reliably reconstruct different volumes of different *ex vivo* and *in vivo* organs.
- Thirdly, we focus our attention on the line-wise sampling of entire RF lines which results in an increase of the frame rate since we work on the reconstruction of the envelope image. We study the influence of the sampling strategy on simulated data, by comparing in each case the 2 most common sampling patterns, i.e. point-wise and line-wise random patterns. We show in particular that although not fully uniform, the line-wise sampling results in an accuracy comparable to the conventional point-wise sampling, using the dictionary based strategy we propose. As previously mentioned, this indicates that CS acquisition of 3D data is feasible in practice in a relatively simple setting.

The structure of the paper is the following. First, the principle of CS is briefly recalled. Second, the dictionary learning algorithm is described. Section IV presents the application of the proposed approach to 3D simulated and experimental data. Section V presents the results obtained from numerical simulations, section VI describes the results obtained from acquisitions performed *ex vivo* on animal brains, hearts and kidneys and section VII provides the results obtained from *in vivo* liver and kidney data. Finally, section VIII concludes this paper.

II. Compressed sensing theory

Compressive sensing (CS) [2] allows the reconstruction of a signal $x \in \mathbb{R}^n$ from a linear combination of a small number of random measurements $y \in \mathbb{R}^m, m < n$. In a general setting, the measurements y may be acquired in the so-called "sensing basis" Φ , which depends on the acquisition device. For example, in MRI, Φ is the Fourier basis and in ultrasound, Φ simply consists in the usual delta functions. We then have:

$$y = R\Phi x \tag{1}$$

where $R\Phi$ is thus a $m \times n$ matrix. The columns of R have an entry one at random positions and zero elsewhere, thereby modeling the random selection of the measurements.

The CS theory assumes that x has a sparse representation in some model basis Ψ, which can be an orthonormal basis, a frame or an overcomplete dictionary, such that:

$$x = \Psi v \tag{2}$$

where v has only $s < m < n$ non zero coefficients. The signal v is then said to be s-sparse. CS theory shows that this sparsity allows an exact recovering of v with overwhelming probability for a certain class of matrices $\Phi\Psi$ [3]. In particular, the sensing basis Φ has to be incoherent with the model basis Ψ [27], which is ensured by the randomness of the non-zero components of $R\Phi$. Finally, the problem can be written as follows:

$$y = R\Phi\Psi v = Av \tag{3}$$

where A is a $m \times n$ full rank matrix (i.e. the m rows of A are independent).

In these settings, the CS problem thus amounts to solve (3) for v , under the constraint that v is sparse. Once v is estimated, the signal x , can then be computed from (2).

For reconstruction of measurements with additive noise, we have:

$$y = Av + e \tag{4}$$

where e represents a noise term of bounded energy $\|e\|_2 \leq \varepsilon$. In this setting, reconstruction can be performed by solving the following minimization problem [28]:

$$P: \ \hat{v} = arg \min_{v \in \mathbb{R}^n} \|v\|_1 \quad \text{subject to} \quad \|y - Av\|_2 \leq \varepsilon \tag{5}$$

[1]More generally, the physics of acquisition often imposes a deterministic sensing matrix, e.g. Fourier in MRI. Since US acquisition is performed in the direct domain, the sensing matrix is simply the identity in this case.

In practical applications the signal is generally not exactly sparse but most of its coefficients in (2) are small. When signal coefficients v decays exponentially in absolute value, the signal is called compressible. The solution found by P (5) gives the approximation of v by keeping its largest entries.

III. Learning overcomplete dictionaries

The CS theory assumes that x has a sparse representation in some model basis which, in this study, will be an overcomplete dictionary rather than a fixed basis. Such dictionaries allow for much sparser representations of the signals since they are optimized for a family of signals that are of interest such as US images.

Dictionary learning uses a set of training samples $X = \{x_i\}_{i=1}^{N}$, which contains N samples, to find an optimal dictionary D of $\mathbb{R}^{p\times K}$ containing K atoms of size p that will best sparsify them. This can be formulated as the following minimization problem:

$$\min_{D,\Theta} \|X - D\Theta\|_F^2 \;\; s.t.\; \forall i, \|\theta_i\|_0 \leq T_0 \tag{6}$$

where $\|\|_F$ is the Frobenius norm. T_0 is the number of non-zero entries, which is expected to be very small, X contains all the training samples as columns and Θ contains the corresponding coefficients.

Solving the dictionary learning problem is also NP-hard and numerous algorithms have been proposed, the literature on this topic being vast and fast growing. Tosic *et al.* give a review of these methods in [29]. In this study we have chosen the K-SVD algorithm due to its efficiency and ease of implementation.

The K-SVD method solves iteratively the optimization problem using two steps: the sparse-coding step and the dictionary update step. In the sparse-coding stage we assume the knowledge of D and we find Θ using any pursuit algorithm. The minimization problem (8) can be decomposed as follows:

$$\min_{\theta_i} \|x_i - D\theta_i\|_2^2 \quad \text{s.t.} \quad \|\theta_i\|_0 \leq T_0, \;\text{for } i = 1, 2, .., N \tag{7}$$

and is usually solved using orthogonal matching pursuit (OMP). In the dictionary update step, both Θ and D are assumed to be fixed, and only the kth column d_k in the dictionary D is updated with its corresponding representation coefficients.

This algorithm is called K-SVD as it obtains the updated dictionary by SVD computations, each determining one column. The algorithm iterates between the two steps until convergence. K-SVD algorithm and variations can be seen in detail in [30], [31]. In this study, the implementation of the approximate K-SVD presented by Rubinstein [32] were used with the improvements proposed in [31] for the learning of the overcomplete dictionary.

IV. Application to 3D US imaging

A. Sampling patterns in 3D US imaging

The sampling protocols in US imaging have to be designed to fulfill both the requirements of CS and of the US devices. On one hand, according to the CS theory the sampling basis has to be incoherent with the sparsifying basis. On the other hand, the US imaging devices have a limited number of sampling strategies that can be adopted due to the physical constraints of the medical devices.

As a result, two main sampling strategies are then possible: point-wise random sampling, which corresponds to the standard CS settings, or line-wise sampling, which has the advantage of yielding a simple practical implementation for ultrasound acquisitions. In the first strategy each of the acquired RF line[2] is sampled according to a distinct random pattern and the full US acquisition (corresponding to the juxtaposition of the N acquired RF lines) then forms the RF image. Overall, this strategy can thus simply be seen as directly sampling the full RF image with a point-wise 2D random pattern and many studies devoted to CS in US uses such scheme [10]–[15]. Let us note however that while this sampling strategy is desirable for CS reconstruction, its practical implementation is far from obvious. First, such implementation implies a dedicated hardware allowing for the desired random sampling. Moreover, RF lines are nowadays obtained by beamforming in the digital domain and the sampling has thus to be performed at the level of raw data. This is taken into account in the study by Liebgott [4], which performs CS reconstruction of pre-beamformed RF signals. The second strategy corresponds to a line-wise random pattern: it simply consists in randomly skipping the acquisition of several RF lines, thus yielding a RF image with missing lines. The main interest of this strategy relies on the fact that it is by design much easier to implement in practice and it has thus been used in several studies [17], [18], [33].

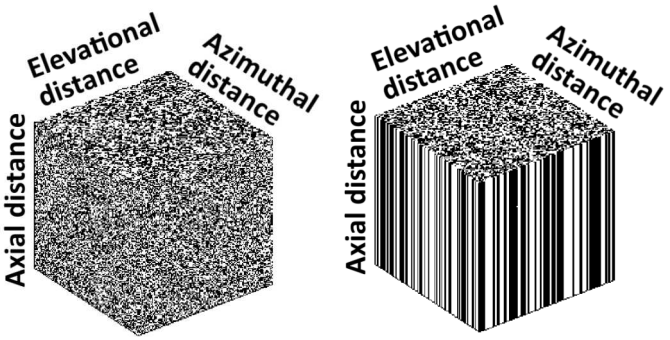
The data acquisition in US imaging is performed in the image space so the acquisition basis Φ is the Dirac basis. In this context, two different sampling schemes R_1 and R_2 are proposed and evaluated. R_1 is a uniform random sampling pattern in the three directions. R_2 consists in uniformly random sampling the same set of RF lines on each consecutive slice of the volume in the axial direction. These two sampling schemes are represented in Figure 1.

B. Reconstruction bases

The quality of the reconstruction depends on the sparsity of its representation in the model basis Ψ. RF data represent a specific challenge, because they do not easily lend themselves to a sparse representation in fixed basis or frames. We thus propose using learned overcomplete dictionaries to represent them. Indeed, such a dictionary is optimized for a given family of signals of interest and allow theoretically for much sparser representations of these signals.

In order to quantify the interest of learned dictionnary for CS reconstruction of US data, it will be informative to compare this reconstruction with the one obtained with standard bases. In this study, we have selected two bases for this comparison: the first is the Fourier basis (i.e. standard DFT), since it is maximally incoherent with the sensing Dirac

[2]Let us recall here that US acquisition starts with a series of raw channel signals available from the probe elements, which are then beamformed to form the radio-frequency (RF) signals. Detection of the envelope of each RF signal and juxtaposition yields the so-called envelope image. Log transformation finally provides the usual B-mode image.



(a) R_1: 3D random sampling (b) R_2: 3D line-wise sampling

Fig. 1. Sampling masks R_1 and R_2 adapted to a spatial sampling of 3D US volumes. The black pixels correspond to the samples used for CS. The proportion of samples here is 50% of the original volume.

basis and it has been shown to provide reasonnable results [13]-[17]. The second basis is the discrete cosine transform (DCT), since it has better coefficients concentration properties than the Fourier transform and thus should provide a more compressible representation.

C. Reconstruction scheme

In each case, the testing dataset consisted of one 3D volume and formed the original data, x. CS reconstruction was then performed on a sub-sampled dataset by removing varying amounts of samples. In the results presented, $20 - 80\%$ of the original samples were removed. Both the training and CS reconstruction were performed on the 3D envelope of normalized volumes before log-compression. In both experiments, where sectorial probes are used, we work only on the polar data before scan conversion to the Cartesian data.

CS reconstruction using an overcomplete dictionary was performed using a block-wise approach. The patches extracted from the volume may be overlapping with an overlap rate r and are of size p. Let x_i of $\mathbb{R}^p, p \ll n$, be a 3D patch of the volume to be reconstructed and D of $\mathbb{R}^{p \times K}$ be the overcomplete dictionary, where K is the number of atoms of size p, $p < K$. For this patch (2) then writes $x_i = Dv_i$ where v_i thus contains the coefficients of the expansion of x_i in D. By solving P_2, we can then recover x_i from the linear measurements y_i of $\mathbb{R}^m$. The 3D CS reconstruction problem P_2 was solved in our case through the l_1 minimization using the spectral projected-gradient algorithm SPGL1 [34].

The accuracy of the results was quantified by comparing the CS reconstruction to the original data through the normalized root mean square error (NRMSE).

V. Simulation results

A. Simulation setup

We evaluate our algorithm on simulated 3D echocardiographic images generated according to the framework recently proposed in [35], [36]. In synthesis they are obtained by combining an electromechanical model of cardiac contraction [37] with an ultrafast ultrasound simulator (COLE, [38] and Field II, [39]). The framework was used to simulate the ultrasound scan of one cardiac cycle of an healthy heart. [3]

We note that the so obtained sequences, despite synthetic, look extremely realistic and in terms of image properties are fully representative of what is expected from real ultrasound recordings (see Figure 6(a)).

The simulated US system was equipped with a cardiac phased array transducer of center frequency 3.3 MHz transmitting a Gaussian weighted pulse with a -6 dB relative bandwidth of 65%.The raw RF signals were sampled at 50 MHz and a symmetric transverse two-way beam profile was assumed, focusing at 80 mm when transmitting and dynamically focusing on receive.

The simulated images consisted of 107×80 beamlines in azimuthal and elevational direction and 6782 samples in axial direction over an angle of 80×80 degrees, resulting in a frame rate of 30 Hz due to the use of parallel beamforming. More details on the ultrasound model can be found in [36]. After envelope calculation and down-sampling the final data sets were downsampled at 2.2 MHz and consisted of 296 samples in axial direction and 107×80 beamlines .

Using this protocol we simulated two data sequences for this paper with two different scatterer distributions which produce two different speckle patterns. The data was divided in two groups: the training group consisting of five volumes coming from the first simulated sequence and a testing group consisting of one volume of the second simulated sequence. The training data was used for the learning of the overcomplete dictionary while the testing data was reconstructed using the CS theory.

TABLE I
Dictionary learning parameters

Parameters	Default value	Tested values
n	$\mathbf{8^3}$	$4^3, 6^3, 10^3, 12^3$
r (%)	**50**	0, 25, 75
q	**4**	1, 2, 3

B. Results: Influence of dictionary learning parameters

Our first objective is to evaluate the performance of our method by acting on the following parameters of the dictionary learning algorithm: patch size (n), overcompleteness (q, where the size of the dictionary is given by $K = q \times n$), and overlap of patches (r, overlapping percentage rate between two consecutive patches).

We measure the influence of each parameter separately, by making one of them vary in the range specified in the Table I "Tested values" column, while all others are set to their default value. The performance is measured by comparing the CS reconstruction to the original data through the NRMSE as a function of the number of removed samples. For brevity sake, we present in the following the results obtained using the sampling pattern R_2 (line-wise sampling) since the sampling

[3]The corresponding data is made publicly available at the following link: http://bit.ly/3dstraus.

mask R_1 is of limited interest for US acquisitions. The influence of the sampling patterns R_1 and R_2 is then tested on the reconstruction of simulation data using a learned dictionary with default parameters.

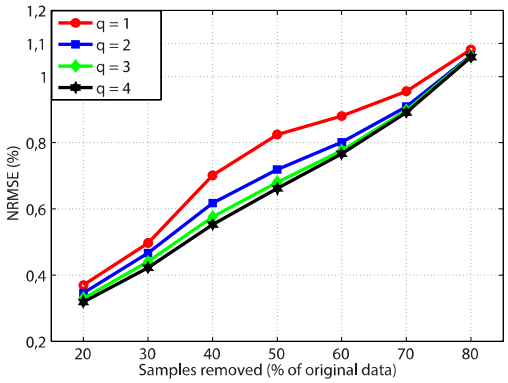


Fig. 2. Influence of the overcompleteness q on the resulting NRMSE as a function of the number of removed samples using the sampling mask R_2. Other parameters are $n = 8^3, r = 50\%$.

1) Influence of the overcompleteness: In Figure 2 we present the reconstruction errors obtained for four tested redundancy values $q = 1, 2, 3, 4$. As can be seen, for any subsampling rate the increase of the number of dictionary elements generally improves the results, although this improvement becomes small when the dictionary is at least three times overcomplete.

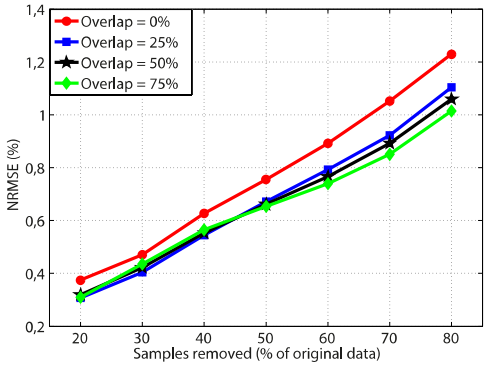


Fig. 3. Influence of the overlap percentage on the resulting NRMSE as a function of the number of removed samples using the sampling mask R_2. Other parameters are $n = 8^3, q = 4$.

2) Influence of the overlapping percentage: As shown in Figure 3, a high overlapping rate means that more patches y_p contribute to the reconstruction of the same patch, which leads to a lower statistical reconstruction error. However, having an overlapping rate that is too close to 1 is not favorable in terms of computational load: indeed increasing the overlap rate increases the number of patch reconstructions. Moreover, we can observe from Figure 3 that the improvement is very small when the overlap is larger than 25%.

3) Influence of the patch size: Figure 4 shows the accuracy obtained for different patch sizes. Note that this experiment was made independent of the overlap factor by setting r=0%

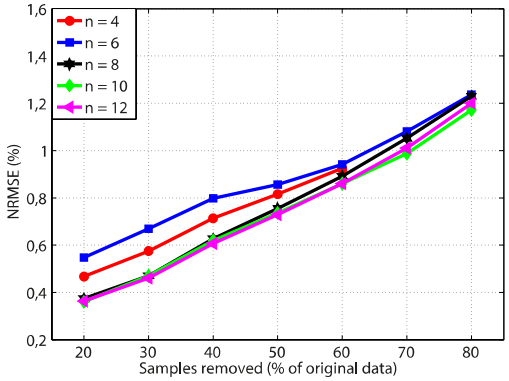


Fig. 4. Influence of the patch size n on the resulting NRMSE as a function of the number of removed samples using the sampling mask R_2. Other parameters are $q = 4, r = 0\%$.

(i.e. no overlap). We notice on Figure 4 that the reconstruction NRMSE is close for patches of size equal to or larger than 8, thus we might be tempted to choose a high patch size in order to reduce the number of total patch reconstructions. However, the higher the patch, the larger the reconstruction computational load. A patch size of 8 seems to be a good compromise between a good NRMSE and a reasonable amount of computations.

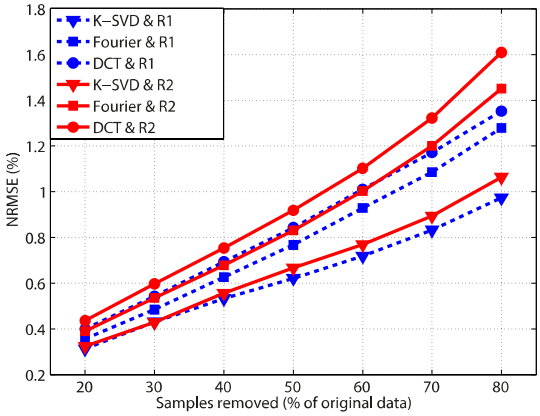
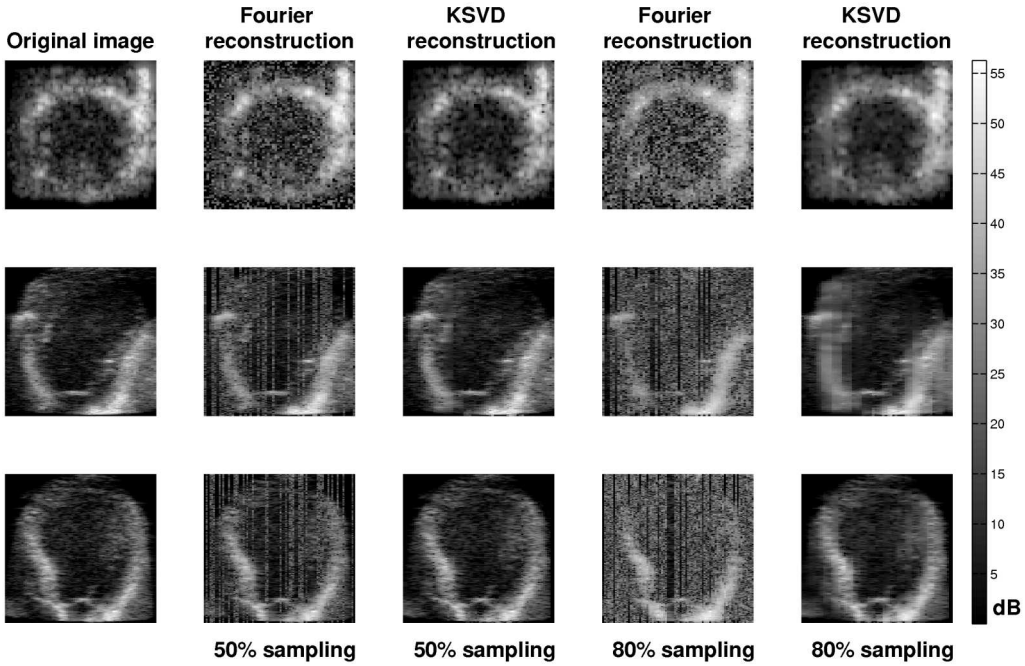


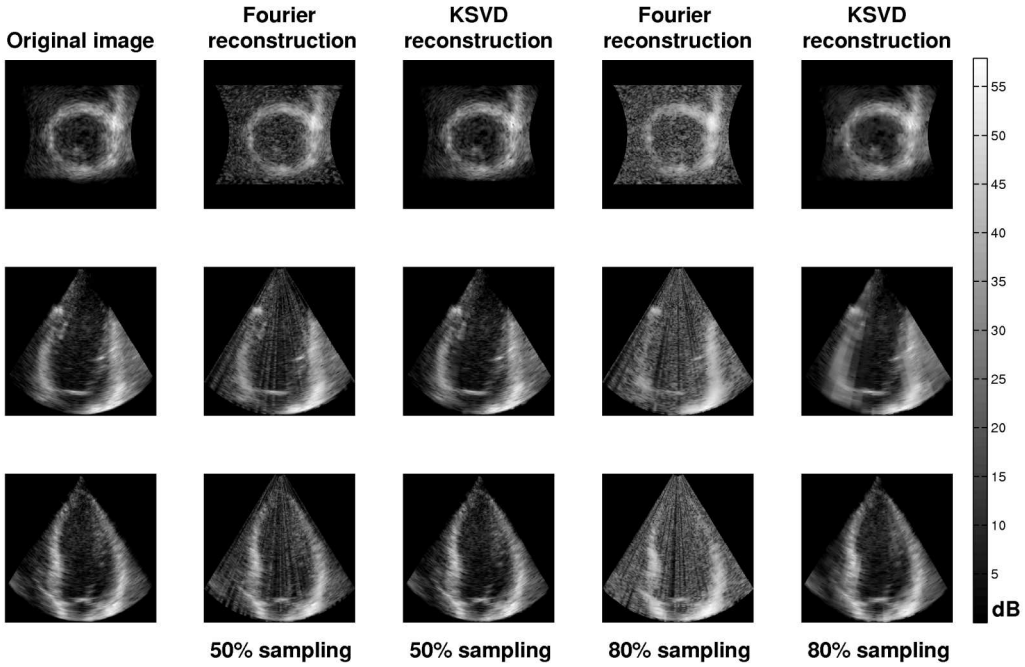
Fig. 5. NRMSE as a function of the number of removed samples using the sampling mask R_1 and R_2. The error is computed on the envelope of the 3D US volume after CS reconstruction using K-SVD dictionary, Fourier basis and DCT.

C. Results: Simulation data reconstructions

Using the default parameters, we compare the reconstruction results that we obtain using the K-SVD dictionary with compressed sensing reconstructions using the Fourier basis and the DCT. This comparison is performed for the two proposed sampling patterns R_1 and R_2. Figure 5 shows the reconstruction NRMSE error as a function of the subsampling rate and for each of the transforms used for reconstruction and using the sampling patterns R_1 and R_2 (Figure 5). It can be observed that the error increases with the number of removed samples, for every transform and whatever the sampling pattern. The error corresponding to the DCT takes the largest values, and the distance between the error curves corresponding to each



(a) Polar coordinates



(b) Cartesian coordinates

Fig. 6. Visualization of 3D CS reconstructions of a simulated US volume using the sampling mask R2 in (a) Polar coordinates and (b) Cartesian coordinates. Original data, Fourier based reconstruction using a 50% subsampling rate, K-SVD based reconstruction using a 50% subsampling rate, Fourier based reconstruction using a 80% subsampling rate and K-SVD based reconstruction using a 80% subsampling rate.

transform is relatively constant. The K-SVD dictionary clearly gives the smallest error, whatever the subsampling rate and performs particularly well for high subsampling rates. We note that for any transform, the reconstruction error associated with the line-wise sampling pattern R_2 is close and slightly larger than the error obtained with the point-wise subsampling scheme R_1, which is from a theoretical point of view the most incoherent with the representation basis/dictionnary. This is interesting since it shows that CS acquisition of 3D data is feasible in a relatively simple setting. Overall, these results show that the default parameters previously chosen for the K-SVD dictionary perform well on simulated volumes.

Figure 6 shows the log-envelope images corresponding to the reconstructed non-log envelope 3D US volume in Polar and Cartesian coordinates. The reconstructions have been performed for the subsampling rates 50% and 80%, and allow

to visually evaluate the significance of the NRMSE plotted in Figure 5. We focus our attention on the sampling mask R_2 since the sampling mask R_1 is of limited interest for US acquisitions. We do not show the results obtained with the DCT since they are similar to the ones obtained with the Fourier basis and the reconstruction error is higher. Figure 6(a) represents the original volume and the reconstruction at 50% and 80% subsampling in Polar coordinates and 6(b) in Cartesian coordinates. It can be observed that the K-SVD CS reconstruction provides the best results using 50% of the initial data while with the Fourier CS reconstruction the missing sample lines can be observed on the reconstructions. Even though the reconstruction produced by the K-SVD dictionary using a 80% subsampling rate are better than with the Fourier basis, the results deteriorate and the missing sample lines can be seen on the reconstructions. In the case of the Fourier reconstruction the missing lines are clear distinguishable in both Polar or Cartesian coordinates while for the K-SVD reconstruction we can see the effect of the reconstruction only on the 80% subsampling case.

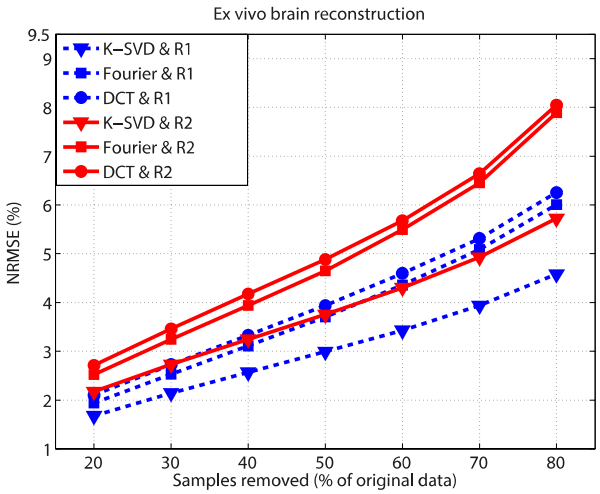


Fig. 7. NRMSE as a function of the number of removed samples using the sampling masks R_1 and R_2. The error is computed for the CS reconstruction of the *ex vivo* brain using organ-dependent K-SVD dictionary, Fourier basis and DCT.

VI. Ex vivo experimental results

A. Ex vivo experimental acquisition setup

For the acquisition of the volumes we use the Ultrasonix MDP research platform equipped with the 4DC7-3/40 Convex 4D transducer. Using this imaging system we imaged the following *ex vivo* organs purchased from the store: 3 pig brains, 3 sheep hearts and 2 sheep kidneys. The central frequency of the probe was of 5 MHz for the 3 organs and the signals were collected using a 40 MHz sampling rate. The transmitted beams, as well as the received signals, were focused at 56 mm depth. The frame rate of the acquired volumes is between 1 and 2 MHz depending or the acquisition. The field of view (FOV) varied between 55 and 65 degrees. For each organ we acquired the non-log envelope. As previously, we only use the non-log envelope volumes to perform the CS reconstruction.

The acquired images consisted of 128×47 beamlines (azimuthal and elevational direction) for the brains, 96×43 beamlines for the hearts and 96×41 for the kidneys. The number of samples in the axial direction was 1792, 2304 and 1792 samples respectively. After non-log envelope calculation the final data sets were downsampled at 5 MHz and consisted of $128 \times 47 \times 224$, $96 \times 43 \times 288$ and $96 \times 41 \times 224$ voxels for the brains, hearts and kidneys respectively.

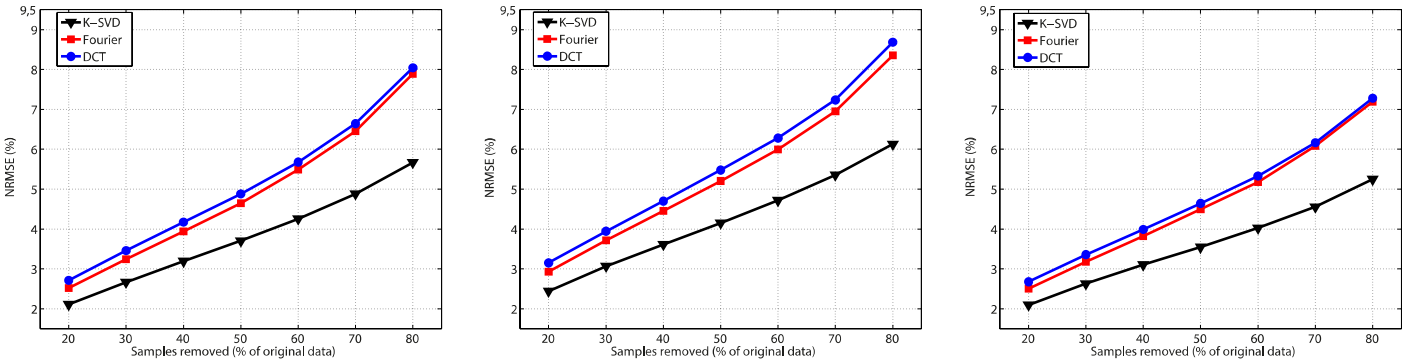
B. Ex vivo reconstruction results using organ-dependent dictionaries

1) Ex vivo brain and heart US images reconstruction: The CS reconstruction using organ-specific dictionaries was tested on sets composed of US acquisitions of two different organs: *ex vivo* brain and heart. In both cases, the training data consisted of two acquired volumes of two distinct brains or hearts and the testing data corresponded to the acquisition of a third brain or heart. The training and testing data thus consisted of different acquisitions of the same kind of organ. According to the reconstruction scheme and the previously chosen default parameters the training data was composed of patches extracted from the volumes belonging to the training dataset.

Figure 7 shows the reconstruction error (NRMSE) computed on the non-log envelope volumes as a function of the percentage of removed data for the three bases tested. In order to ease the assessment of the error, the original data acquired were normalized to unity before reconstruction in this experiment. Quite consistently, it can be observed that the error increases with the number of samples removed, whatever the reconstruction basis. The error corresponding to the DCT takes the largest values. The K-SVD dictionary gives the smallest error, whatever the subsampling rate and the distance between the K-SVD NRMSE and the other basis reconstruction errors is rather large. The Fourier basis and the DCT give errors that remain close to each other with the same tendency and deteriorate faster with the increase of the subsampling rate, as compared to the K-SVD dictionary. We notice that, as in the case of simulated data the reconstruction error associated with the line-wise sampling pattern R_2 is close and slightly larger than the error obtained with the point-wise subsampling scheme R_1. This suggests that CS acquisition of 3D data is feasible in the line-wise sampling setting without an important increase of the reconstruction error.

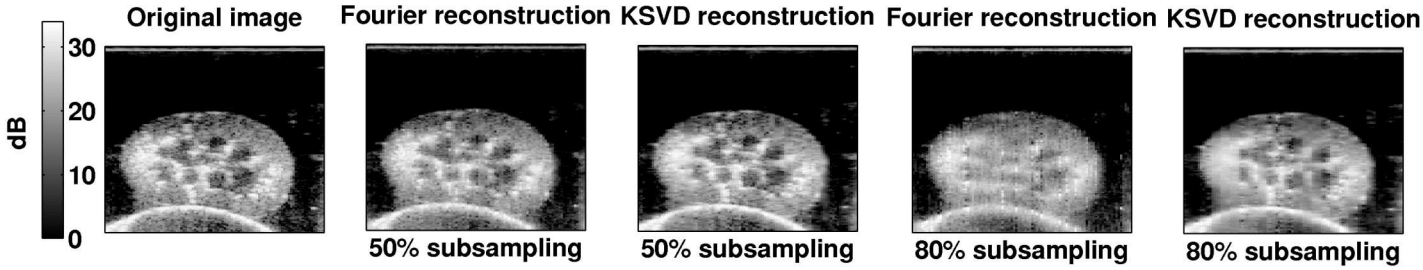
C. Generality: Reconstruction results using a single dictionary

In order to test the generality of the CS reconstruction using an overcomplete dictionary learned with the K-SVD algorithm, we trained the dictionary on a set composed of US acquisitions of three different organs. The training data is thus composed of US volumes of two brains, two hearts and one kidney. The testing dataset is composed of one brain, one heart and one kidney that are not included in the training dataset. The dictionary learning parameters remain the same as previously

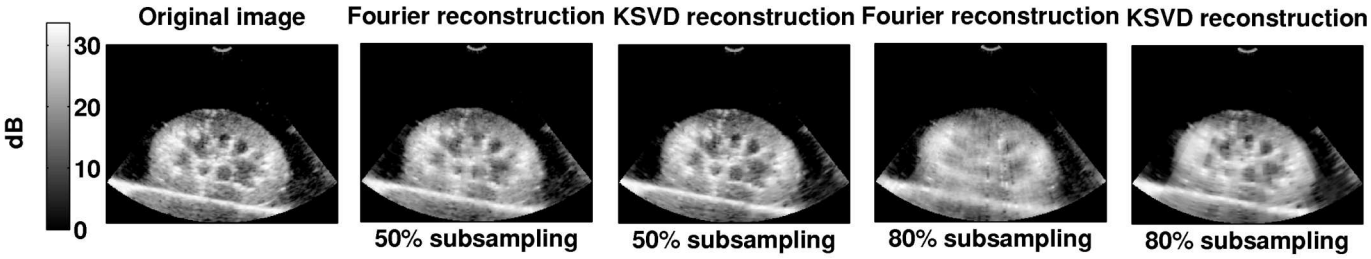


(a) brain (b) heart (c) kidney

Fig. 8. NRMSE as a function of the number of removed samples using the line-wise sampling mask R_2. The error is computed for the CS using the general K-SVD dictionary, Fourier basis and DCT for the CS reconstruction of: (a) brain, (b) heart and (c) kidney.



(a) Polar coordinates



(b) Cartesian coordinates

Fig. 9. Visualization of 3D CS reconstructions of an *ex vivo* kidney US volume using the sampling mask R_2 in (a) Polar coordinates and (b) Cartesian coordinates. Reconstructions using the K-SVD dictionary for 50 and 80% subsampling rates.

and we focus our attention on the sampling mask R_2 since the sampling mask R_1 is of limited interest for US acquisitions and yields similar results.

Figure 8 shows the reconstruction error NRMSE for the beamformed non-log envelope images as a function of the subsampling rate, for each of the transforms used for reconstruction and for each of the reconstructed US organ volumes i.e.: brain, heart and kidney. We can notice that the results obtained for the three organs are very close and the curves exhibit the same tendencies. Interestingly, the differences between the reconstructions performed using the specific dictionaries of the previous section (Figures 7) and the ones performed using the general dictionaries are very small: in the range $[4.8-6.4]\times10^{-4}$ for the brain reconstruction and $[1.1-7.4]\times10^{-4}$ for the heart reconstruction. This indicates that the general dictionary performs as well as the organ-specific dictionaries.

Figure 9 shows the log-envelope images corresponding to the reconstructed non-log envelope 3D US volume of the *ex vivo* kidney in Polar and Cartesian coordinates. We show only one axial-azimuthal slice on which we can see the effects of the sampling and the quality of the reconstruction. We do not show the results obtained with the DCT since they are similar to the ones obtained with the Fourier basis and the reconstruction error is higher. We do not show either the reconstructions of the *ex vivo* heart and brain since the results in terms of reconstruction errors are similar. The left image represents the original data before subsampling and reconstruction followed by the CS reconstruction using the K-SVD dictionary and the CS reconstruction using the Fourier

basis at 50% and 80% subsampling.

The first observation to make is that for all the subsampling rates, the CS method provided good reconstructions of the whole volume except in the case of the 80% subsampling where we can clearly start to see the missing columns, particularly in the case of the Fourier-based reconstruction. It can be observed that the K-SVD CS reconstruction provides the best results, as previously, while with the Fourier CS reconstruction the missing sample lines can always be seen on the reconstructions. Even though the reconstruction produced by the K-SVD dictionary using 20% of the initial data are better than with the Fourier basis, the quality of the results decreases when removing more samples and the missing samples start to be visible on the reconstructions. At 50% subsampling CS with the dictionary reconstructs better the fine details such as the vessel ramifications in the kidney. At 80% subsampling, even though the reconstruction produced by the K-SVD dictionary are better than with the Fourier basis, the quality of the results is lower and the missing sample channels start to be seen on the reconstructions. At such subsampling rate, the CS reconstruction with the Fourier basis loses almost completely the fine details and the resulting image is much deteriorated while with the K-SVD these structures are still preserved.

Overall we observe that the reconstructions performed with experimental data are consistently worse than with the simulated data. In terms of NRMSE the error is indeed five times higher in the case of the experimental volumes. This is easily explained by the quality of the original samples. The comparison of Figure 6 (simulated data) and the Figures 9 (experimental data) shows clear differences: the simulated images have a better resolution and less speckle regions than the experimental volumes.

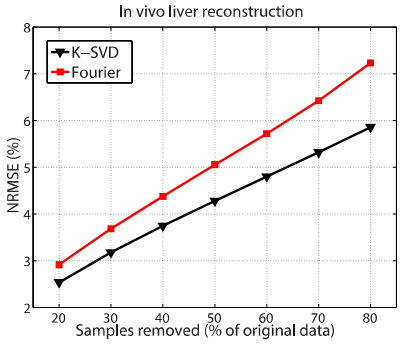
VII. In vivo experimental results

A. In vivo experimental acquisition setup

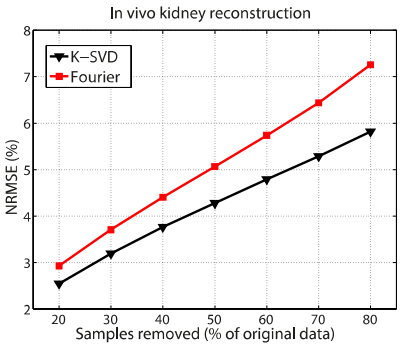
For the acquisition of the volumes we use the Ultrasonix MDP research platform equipped with the 4DC7-3/40 Convex 4D transducer. Using this imaging system we imaged the liver and the kidney of two healthy subjects. All subjects provided written informed consent. The central frequency of the probe was of 4 MHz for all the acquisitions. The transmitted beams, as well as the received signals, were focused at 76 mm depth. Other parameters of the probe were: field of view (FOV) = 45 degrees, frame rate = 2 MHz and gain = 50% for all the acquisitions. For each subject and organ we acquired the non-log envelope. As previously, we only use the non-log envelope volumes to perform the CS reconstruction. After non-log envelope calculation and down-sampling, the final data sets were sampled at 5MHz and consisted of volumes of $33 \times 128 \times 512$ voxels.

B. In vivo reconstruction results non organ-dependent dictionaries

The CS reconstruction of *in vivo* US organs was tested in a general context using the dictionary described in the previous section, i.e. learned from a set of 3 different *ex vivo* organs (two brains, two hearts and one kidney). The testing dataset is composed of one liver and one kidney that were acquired *in vivo*. We focus our attention on the sampling mask R_2 and on the Fourier basis.



(a) liver

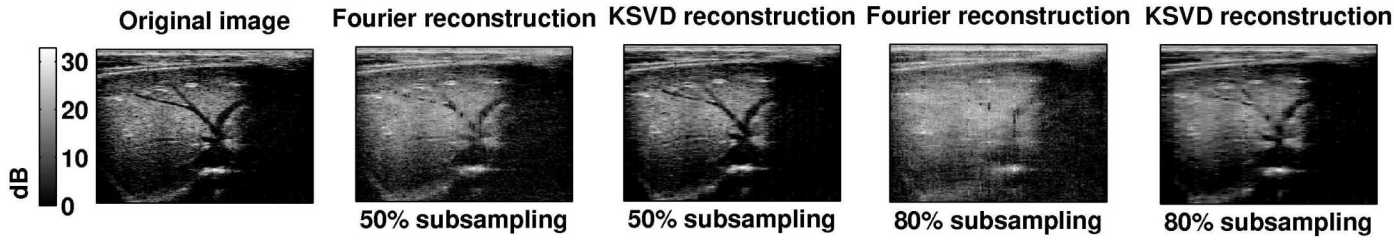


(b) kidney

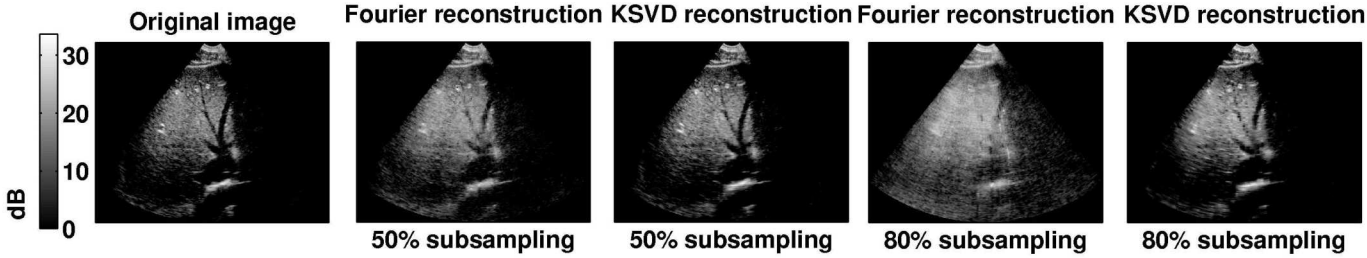
Fig. 10. NRMSE as a function of the number of removed samples using the line-wise sampling mask R_2. The error is computed for the CS using the general K-SVD dictionary and the Fourier basis for the CS reconstruction of: (a) liver and (b) kidney.

Figure 10 shows the reconstruction error NRMSE for the non-log envelope images as a function of the subsampling rate, for each of the transforms used for reconstruction and for each of the reconstructed US organ volumes i.e. liver and kidney. We can notice that the results obtained for the two organs are very close and the curves exhibit the same tendencies. The differences between the reconstructions performed using the same dictionary for the reconstruction of *ex vivo* volumes (Figure 8) and the ones performed on *in vivo* US volumes are very small: in the range $[4.5 - 7.6] \times 10^{-3}$ for the kidney reconstruction. This indicates that the general dictionary performs well for both *ex vivo* or *in vivo* 3D acquisitions.

Figure 11 shows the log-envelope images corresponding to the reconstructed non-log envelope 3D US volume of an *in vivo* liver, in Polar and Cartesian coordinates for better visibility. We show only one axial-azimuthal slice on which we can see the effects of the sampling and the quality of the reconstruction. We do not show the results obtained with the DCT since they are similar to the ones obtained with the Fourier basis and the reconstruction error is higher. The left image represents the original data before subsampling and reconstruction followed by the CS reconstruction using the *ex vivo* K-SVD dictionary and the CS reconstruction using the Fourier basis at 50% and 80% subsampling. We observe



(a) Polar coordinates



(a) Cartesian coordinates

Fig. 11. Visualization of 3D CS reconstructions of an *in vivo* liver US volume using the sampling mask R_2. Reconstructions using the K-SVD dictionary for 50 and 80% subsampling rates.

that we do not see anymore the subsampled beamlines on the reconstruction. The Fourier reconstruction still has visually a poorer quality than the K-SVD reconstruction, the images are covered with noise and the finer details are lost. At 50% subsampling, CS using the dictionary gives satisfying results and both small and big structures are preserved. At 80% subsampling, even though the reconstruction produced by the K-SVD dictionary is better than with the Fourier basis, the quality of the results is lower and some regions are deteriorated.

VIII. Conclusion

This study demonstrates that the CS theory using overcomplete learned dictionaries can be applied to 3D ultrasound imaging to reduce the volume of data needed for the reconstruction and speed up the acquisitions. We showed in particular that the learned dictionary approach yields lower errors (NRMSE) than conventional sparsifying dictionaries based on fixed transforms such as Fourier or discrete cosine. Experiments performed on simulated and experimental 3D US volumes with the K-SVD based CS reconstruction using a subsampling rate up to 80% resulted in US volumes close to the original, with minimal loss of information and seem to confirm our sparsity hypothesis.

In addition, a sampling protocol suited to US imaging was proposed here by randomly subsampling full RF lines. The obtained results using this line-wise sampling strategy showed that we can recover 3D ultrasound signals of high quality using only 50% of the initial data. This shows in particular that although not fully uniform, the line-wise sampling yields results with an accuracy comparable to the conventional point-wise random subsampling. This indicates that CS acquisition of 3D data is thus feasible in practice in a relatively simple setting.

Finally, we evaluated the generality of the learned overcomplete dictionary by training it on 3D US volumes of different *ex vivo* and *in vivo* organs. This approach showed that the dictionary is not hampered by overlearning and provides satisfying reconstructions for images coming from different organs. The difference between the reconstruction error obtained with a specific dictionary and the one obtained with the general dictionary is minimal (on the order of 10^{-4}).

In this paper, we showed the powerful potential of CS to reduce data volume and speed up acquisitions with the proposed sampling protocol. However, since the image is treated as overlapping blocks, the overall computational time is highly increased. Considering that the dictionary is learned in advance, using dedicated circuits (GPU type) for the CS reconstruction could allow a great improvement in processing times and overall increase the imaging rate, keeping the real-time nature of US imaging.

Future work will include an investigation of adaptive sparse learning routines providing on the fly dictionaries as well as an optimization of the processing and reconstruction times. Various applications will also be considered (3D imaging using matrix arrays and duplex ultrasonography (B-mode/Doppler)).

Acknowledgment

This work was performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). The work of Martino Alessandrini was supported by the Fonds Wetenschappelijk Onderzoek (1263814N). The authors would also like to thank Adeline Bernard for her help with the experimental acquisitions.

REFERENCES

[1] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] E. Candés and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[3] E. Candés, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, 2006.

[4] H. Liebgott, R. Prost, and D. Friboulet, "Pre-beamformed rf signal reconstruction in medical ultrasound using compressive sensing," *Ultrasonics*, vol. 53, no. 2, pp. 525–533, 2013.

[5] C. Quinsac, A. Basarab, J. Girault, and D. Kouamé, "Compressed sensing of ultrasound images: sampling of spatial and frequency domains," in *IEEE Workshop on Signal Processing Systems*, 2010, pp. 231–236.

[6] C. Quinsac, A. Basarab, J.-M. Gregoire, and D. Kouamé, "3d compressed sensing ultrasound," in *IEEE International Ultrasonics Symposium*, 2010, pp. 363–366.

[7] M. Schiffner, T. Jansen, and G. Schmitz, "Compressed sensing for fast image acquisition in pulse-echo ultrasound," *Biomedical Engineering*, vol. 57, no. SI-1, pp. 192–195, 2012.

[8] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1827–1842, 2011.

[9] N. Wagner, Y. C. Eldar, and Z. Friedman, "Compressed beamforming in ultrasound imaging," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4643–4657, 2012.

[10] Q. Zhang, B. Li, and M. Shen, "A measurement-domain adaptive beamforming approach for ultrasound instrument based on distributed compressed sensing: Initial development," *Ultrasonics*, vol. 53, p. 255264, 2013.

[11] X. Zhuang, Y. Zhao, Z. Dai, H. Wang, and L. Wang, "Ultrasonic signal compressive detection with sub-nyquist sampling rate," *Journal of Scientific and Industrial Research*, vol. 71, no. 3, pp. 195–199, 2012.

[12] A. Achim, B. Buxton, G. Tzagkarakis, and P. Tsakalides, "Compressive sensing for ultrasound rf echoes using a-stable distributions," in *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2010, pp. 4304–4307.

[13] A. Achim, A. Basarab, G. Tzagkarakis, P. Tsakalides, and D. Kouamé, "Reconstruction of compressively sampled ultrasound images using dual prior information," in *IEEE International Conference on Image Processing (ICIP)*, 2014, p. in press.

[14] A. Basarab, H. Liebgott, O. Bernard, D. Friboulet, and D. Kouamé, "Medical ultrasound image reconstruction using distributed compressive sampling," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2013, pp. 624–627.

[15] N. Dobigeon, A. Basarab, D. Kouamé, and J.-Y. Tourneret, "Regularized bayesian compressed sensing in ultrasound imaging," in *European Signal and Image Processing Conference (EUSIPCO)*, 2012.

[16] D. Friboulet, H. Liebgott, and R. Prost, "Compressive sensing for raw rf signals reconstruction in ultrasound," in *IEEE International Ultrasonics Symposium*, 2010, pp. 367–370.

[17] C. Quinsac, N. Dobigeon, A. Basarab, D. Kouamé, and J.-Y. Tourneret, "Bayesian compressed sensing in ultrasound imaging," in *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2011, pp. 101–104.

[18] Y. Chuo, T.-H. Chan, and M.-L. Li, "Ultrasound compressed sensing: Performance study of reconstruction on different ultrasound imaging data," in *IEEE International Ultrasonics Symposium*, 2013, pp. 903–905.

[19] M. Mishali, Y. C. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-nyquist rates," *IET Circuits, Devices & Systems*, vol. 5, no. 1, pp. 8–20, 2011.

[20] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *Signal Processing, IEEE Transactions on*, vol. 50, no. 6, pp. 1417–1428, Jun 2002.

[21] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.

[22] C. Quinsac, A. Basarab, and D. Kouamé, "Frequency domain compressive sampling for ultrasound imaging," *Advances in Acoustics and Vibration*, vol. 12, pp. 1–16, 2012.

[23] M. Birk, A. Burshtein, T. Chernyakova, A. Eilam, J. Choe, A. Nikoozadeh, P. Khuri-Yacoub, and Y. C. Eldar, "Compressed 3d ultrasound imaging with 2d arrays," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014.

[24] O. Lorintiu, H. Liebgott, O. Bernard, and D. Friboulet, "Compressive sensing ultrasound imaging using overcomplete dictionaries," in *IEEE International Ultrasonics Symposium*, 2013, p. Accepted.

[25] O. Lorintiu, H. Liebgott, M. Alessandrini, O. Bernard, and D. Friboulet, "Compressed sensing reconstruction of 3d ultrasound data using dictionary learning," in *IEEE International Conference on Image Processing (ICIP'14)*, 2014, p. accepted.

[26] I. Tosic, I. Jovanovic, P. Frossard, M. Vetterli, and N. Duric, "Ultrasound tomography with learned dictionaries," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, March 2010, pp. 5502–5505.

[27] E. Candés and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, p. 969, 2007.

[28] E. J. Candés, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 910, pp. 589 – 592, 2008.

[29] I. Tosic and P. Frossard, "Dictionary learning," *Signal Processing Magazine, IEEE*, vol. 28, no. 2, pp. 27–38, 2011.

[30] M. Aharon, M. Elad, and A. Bruckstein, "k -svd: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.

[31] L. Smith and M. Elad, "Improving dictionary learning: Multiple dictionary updates and coefficient reuse," *Signal Processing Letters, IEEE*, vol. 20, no. 1, pp. 79–82, 2013.

[32] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the k-svd algorithm using batch orthogonal matching pursuit." [Online]. Available: http://www.cs.technion.ac.il/%7Eelad/software/

[33] A. Basarab, A. Achim, and D. Kouamé, "Medical ultrasound image reconstruction using compressive sampling and lp norm minimization," in *SPIE Medical Imaging*, 2014.

[34] E. van den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.

[35] M. Alessandrini, M. De Craene, O. Bernard, S. Giffard-Roisin, P. Allain, J. Weese, E. Saloux, H. Delingette, M. Sermesant, and J. D'hooge, "A pipeline for the generation of realistic 3d synthetic echocardiographic sequences: Methodology and open-access database." *Medical Imaging, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

[36] M. D. Craene and S. M. *et al.*, "3d strain assessment in ultrasound (straus): A synthetic comparison of five tracking methodologies," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 9, pp. 1632–1646, 2013.

[37] M. Sermesant, Y. Coudière, H. Delingette, N. Ayache, and J. Désidéri, "An electro-mechanical model of the heart for cardiac image analysis," in *Medical Image Computing and Computer-Assisted Intervention MICCAI*, 2001, vol. 2208, pp. 224–231.

[38] H. Gao, H. F. Choi, P. Claus, S. Boonen, S. Jaecques, G. Van Lenthe, G. Van der Perre, W. Lauriks, and J. D'hooge, "A fast convolution-based methodology to simulate 2-dd/3-d cardiac ultrasound images," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 56, no. 2, pp. 404–409, 2009.

[39] J. Jensen and N. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 39, no. 2, p. 262267, 1992.