COMPRESSED SENSING RECONSTRUCTION OF 3D ULTRASOUND DATA USING DICTIONARY LEARNING

O. Lorintiu*, H. Liebgott*, M. Alessandrini†, O. Bernard*, D. Friboulet*

*Université de Lyon, CREATIS ; CNRS UMR5220 ; Inserm U1044 ; INSA-Lyon ; Université Lyon 1, France
†Lab. on Cardiovascular Imaging and Dynamics, Dept. of Cardiovascular Sciences, KU Leuven, Belgium

ABSTRACT
In this paper we propose 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. We will investigate two undersampling patterns of the 3D US imaging: a spatially uniform random acquisition and a line- wise random acquisition. The latter being extremely interesting for 3D imaging: it would indeed allow skipping the acquisition of many lines among the several thousands required in 3D acquisitions, thus, speeding up the whole acquisition process and incrementing the imaging rate. In this study, the dictionary was learned using the K-SVD algorithm on patches extracted from a training dataset constituted of simulated 3D non-log envelope US volumes. Experiments were performed on a testing dataset made of a simulated 3D US log-envelope volume not included in the testing dataset. CS reconstruction was performed by removing 20% to 80% of the original samples according to the two undersampling patterns. Reconstructions using a K-SVD dictionary previously trained dictionary indicate minimal information loss, thus showing the potential of the overcomplete dictionaries.

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

1. INTRODUCTION
In 3D US imaging, the number of RF lines that must be acquired to sweep the whole volume can be extremely high, typically several thousands. The acquisition time of one such a line is related to the speed of sound and the depth of investigation and cannot be compressed thus leading to long acquisition times. In this context, the recently introduced 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 Nyquist sampling limit [1, 2, 3]. Its application to medical ultrasound imaging is promising due its capability to reduce the volume of acquired data and thus to speed up the acquisitions and increase the imaging rate of 3D US data.

The success of the CS reconstruction lies on two principles: sparsity and incoherence. In this paper we focus mainly on the sparsity aspect and in contrast to previous work, we propose a new approach based on the use of learned overcomplete dictionaries. Our choice is motivated by the fact that the existing studies mostly focus on fixed sparsifying transforms [4, 5, 6, 7, 8, 9]. Since sparsity is an important prerequisite to achieve accurate CS reconstruction, we have chosen to investigate learned overcomplete data-driven dictionaries that are optimized for a particular class of images such as US images.

The main purpose of this paper is twofold: first, show the feasibility of reconstructing 3D US volumes from sampled data using the CS theory with overcomplete dictionaries and second, study a sampling protocol designed to speed up 3D acquisitions. 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 simulated data and section V presents the results.

2. COMPRESSIVE 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” $\Phi$, which depends on the acquisition device. For example, in MR imaging $\Phi$ corresponds to the Fourier basis and in conventional ultrasound imaging, $\Phi$ simply consists in the usual delta functions. We then have:

$$y = R\Phi x$$

where $R\Phi$ is 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 $\Psi$, which can be an orthonormal basis, a
frame or an overcomplete dictionary, such that:

\[ x = \Psi v \]  

(2)

where \( v \) has only \( s < m < n \) non zero coefficients. The signal \( v \) is called \( s \)-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 \( \Phi \) has to be incoherent with the model basis \( \Psi \) [10]. Finally, the problem can be written as follows:

\[ y = R\Phi \Psi v = Av \]  

(3)

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).

3. 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 \( Y = \{y_i\}_{i=1}^N \) to find an optimal dictionary \( D \) that will best sparsify them. This can be formulated as the following minimization problem:

\[
\min_{D,X} \|Y - DX\|_F^2 \quad \text{s.t.} \quad \forall i, \|x_i\|_0 \leq T_0
\]  

(4)

where \( T_0 \) is the number of non-zero entries, which is expected to be very small, \( Y \) contains all the training samples as columns and \( X \) 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 [11]. 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 \( X \) using any pursuit algorithm. The minimization problem (4) can be decomposed as following:

\[
\min_{x_i} \left\| y_i - D x_i \right\|_2^2 \quad \text{s.t.} \quad \|x_i\|_0 \leq T_0, \quad \text{for } i = 1, 2, \ldots, N
\]  

(5)

and is usually solved using orthogonal matching pursuit (OMP). In the second stage, the algorithm updates \( D \) one atom at a time and its corresponding non-zero coefficients in \( X \), thus providing an update of both \( D \) and \( X \). The algorithm iterates between the two steps until convergence. K-SVD algorithm and variations can be seen in detail in [12, 13].

4. APPLICATION TO 3D US IMAGING

4.1. 3D ultrasound images simulation setup

We evaluate our algorithm on simulated 3D echocardiographic images generated according to the framework recently proposed in [14]. In synthesis they are obtained by combining an electromechanical model of cardiac contraction [15] with an ultrafast ultrasound simulator (COLE, [16]). The framework was used to simulate the ultrasound scan of one cardiac cycle of an healthy heart.

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

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 RF signals were sampled at 50 MHz and a symmetric transverse two-way beam prole was assumed, focusing at 80 mm when transmitting and dynamically focusing on receive.

The simulated images consisted of \( 107 \times 80 \times 80 \) lines in azimuth and elevation direction over an angle of \( 80 \times 80 \) degrees, resulting in a frame rate of 30 Hz due to the use of parallel beam forming. More details on the ultrasound model can be found in [17]. After non-log envelope calculation and downsampling the final data sets consisted of \( 107 \times 80 \times 290 \) voxels.

Using this protocol we simulated two data sequences for this paper. The data was divided in two groups: the training group consisting of five volumes coming from the first simulated sequence and a testing group including 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.

![Fig. 1](image_url)  

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

We note that the so obtained sequences, despite synthetic, look extremely realistic and in terms of image properties are fully representative of what expected from real ultrasound recordings.
4.2. Sampling patterns in 3D US imaging

The sampling protocols in US imaging are 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. From a theoretical point of view, the ideal sampling mask is a uniform random sampling in the 3 spatial directions. However, this is only of a limited interest in US imaging since this is probably as difficult as just acquiring the whole US data set. An interesting approach would be to remove entire columns of the data set. This would indeed consist in skipping the acquisition of entire RF lines, resulting in an increase of the frame rate since we work on the reconstruction of the envelope image. The data acquisition in US imaging is performed in the image space so the acquisition basis $\Phi$ 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 of uniformly random sampling the same set of RF lines on each consecutive slice of the volume in the azimuthal direction. These two sampling schemes are represented in Figure 1.

4.3. Reconstruction scheme

The testing dataset consisted of one simulated 3D volume and formed the original data, $x$. CS reconstruction was then performed on a subsampled dataset where 20% to 80% of the original samples were removed. Both the training and CS reconstruction were performed on the 3D envelopes of normalized simulated volumes before log-compression.

CS reconstruction using an overcomplete dictionary was performed using a block-wise approach. Let $x_p$ of $\mathbb{R}^n$, $n < N$, be a 3D patch of the volume $x$ and $D$ of $\mathbb{R}^{n \times K}$ be an overcomplete dictionary, with $n < K$, such that $x_p = Dv_p$. Application of CS implies learning $D$ such that $v_p$ is a sparse representation of the patch $p$ in $D$. We will then be able to recover the original image patches from the linear measurements $y$. The implementation of the approximate K-SVD presented by Rubinstein [18] were used with the improvements proposed in [13] for the learning of the overcomplete dictionary. The 3D CS reconstruction was solved through the $l_1$ minimization using the spectral projected-gradient algorithm SPGL1 [19].

The accuracy of the results was quantified by comparing the CS reconstruction to the original data through the normalized root mean square error (NRMSE). Besides the K-SVD dictionary, we include two other transforms in our experiments for comparison: the Fourier basis and the discrete cosine transform (DCT).

5. RESULTS ON SIMULATED 3D US IMAGES

5.1. Dictionary learning

The training data consisted of $8 \times 8 \times 8$ patches extracted from the volumes belonging to the training dataset. We use typical values for the algorithm parameters: sparsity $T_0 = 8$; patch dimension $n = 512$ (i.e. $8 \times 8 \times 8$); and overcompleteness of $K = 4n$ (i.e. $4 \times 512 = 2048$).

5.2. Reconstruction results

Figure 2 shows the reconstruction NRMSE error as a function of the subsampling rate and for each of the transforms used for reconstruction. It can be observed that the error increases with the number of removed samples, whatever the reconstruction transform. The error corresponding to the DCT takes the largest values, and the distance between the errors of reconstruction is relatively constant. The K-SVD dictionary gives the smallest error, whatever the subsampling rate. It is interesting to notice that the error of reconstruction using the K-SVD dictionary and a line-wise sampling mask $R_2$ is smaller than the error obtained with the Fourier basis and random subsampling scheme $R_1$, which is from a theoretical point of view the most incoherent.

The error corresponding to the DCT transform with the sampling mask $R_2$ takes the highest values (range $[0.59 - 1.61] \times 10^{-2}$), the error associated with the DCT and the sampling mask $R_1$ as well as the error of the Fourier basis with the sampling masks $R_1$ and $R_2$ show intermediate values (range $[0.54 - 1.35] \times 10^{-2}$, $[0.48 - 1.45] \times 10^{-2}$ respectively $[0.54 - 1.28] \times 10^{-2}$). The K-SVD dictionary with the sampling masks $R_1$ and $R_2$ yield the smallest errors (range...
Fig. 3. Visualization of 3D CS reconstructions of a simulated US volume using the sampling mask $R_2$. (a) original data, (b) Fourier based reconstruction using 50% of the samples, (c) K-SVD based reconstruction using 50% of the samples (d) Fourier based reconstruction using 20% of the samples, (e) K-SVD based reconstruction using 20% of the samples.

Figure 3 shows the log-envelope images corresponding to the reconstructed non-log envelope 3D US volume, for better visibility. 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 3(a) represents the original volume, Figures 3(b)-(c) show the CS reconstructions obtained from 50% of the original samples and the mask $R_2$ using Fourier and dictionary-based reconstruction. Figures 3(d)-(e) show the measurements obtained from 20% of the original samples and the mask $R_2$ obtained with the Fourier transform and the dictionary. The first observation to make is that for all the sampling masks, the CS method provided good reconstructions of the whole volume. The plan that was best reconstructed in each case was always the axial-lateral plan, where the corresponding 2D sampling schemes applied were random.

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 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 results are lower and the missing sample channels can be seen on the reconstructions.

6. CONCLUSION

This study demonstrates that the CS theory can be applied to 3D ultrasound imaging to reduce the volume of data needed for the reconstruction and speed up the acquisitions. One important criteria for the success of the CS theory is the sparsity. We dealt here with the fact that the overcomplete dictionaries allow for much sparser representations of the signals. Experiments performed on simulated 3D US volumes with the K-SVD based CS reconstruction using only 20% to 80% of the initial samples resulted in US volumes close to the original, with minimal loss of information and seem to confirm our sparsity hypothesis.

In addition, a sampling protocols suited to US imaging was proposed here where the US acquisitions can be sampled at random times. The obtained results using this sampling strategy showed that we can recover 3D ultrasound signals of high quality using only 50% of the initial data. Future work will include an investigation of other sparse dictionary learning routines providing better sparsity 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)).
7. REFERENCES


