Component-Based Modeling and Processing of Medical Ultrasound Signals

Yael Yankelevsky, *Student Member, IEEE*, Zvi Friedman, and Arie Feuer, *Life Fellow, IEEE*

***Abstract*—Component-based modeling and processing can be found in many signal and image processing applications. In this paper, we apply this approach to cardiac medical ultrasound imaging. We show how the raw data generated in the ultrasound scanning process can be modeled as consisting of two distinct components. By decomposing these signals into the two components, we are able to apply separate processing pipelines tailored to each component. The potential benefits of this approach to medical ultrasound imaging are demonstrated in two central issues of the commonly used processing: 1) the very large amounts of data generated by the scanning process, and 2) artifacts in the images and videos generated by standard processing, commonly referred to as side lobe artifacts. We manage to compress the data by a factor of over 20 as well as remove the side lobe artifacts. These benefits are demonstrated on both simulated data and real cardiac scanning data.**

***Index Terms*—Sparse representation, signal modeling, decomposition, biomedical ultrasound, digital compression, beamforming, side lobe suppression.**

I. Introduction

THE component based approach is very common in signal and image processing. Examples include denoising [1], [2], compression [3], [4], frequency decomposition of signals and images [5], [6], image dehazing [7], [8], image inpainting [9], audio signal processing [10], and others.

In many cases, decomposition is the goal itself. However, in this paper, we demonstrate the potential benefit of the component based modeling *and processing* as an alternative to the single processing pipeline used for the combined signal in a specific application - cardiac medical ultrasound imaging. Currently, the standard processing is a compromise between the properties of the various signal components. As an alternative, we suggest decomposing the signal into its distinct components, designing a different processing pipeline for each component, tailored to its properties, and then combining the results if desired. These two processing options are schematically described in Fig. 1.

Manuscript received November 6, 2016; revised May 28, 2017; accepted July 10, 2017. Date of publication July 24, 2017; date of current version September 5, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. D. Robert Iskander. *(Corresponding author: Yael Yankelevsky.)*

Y. Yankelevsky is with the Department of Computer Science, Technion-Israel Institute of Technology, Haifa 3200003, Israel (e-mail: yaelyan@tx.technion.ac.il).

Z. Friedman is with GE Healthcare, Tirat Carmel 39120, Israel (e-mail: zvi.friedman@ge.com).

A. Feuer is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 3200003, Israel (e-mail: feuer@ee.technion.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2017.2731303

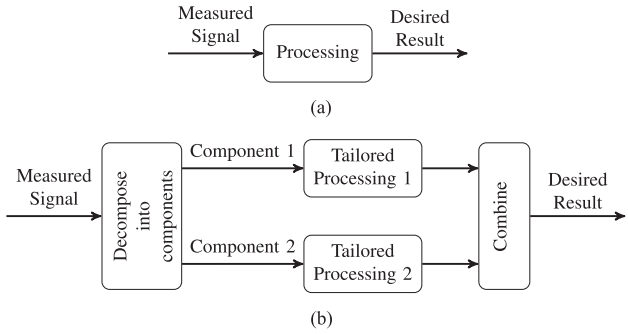


Fig. 1. Alternative processing of a signal consisting of two components: (a) A single processing pipeline applied to the combined signal, (b) Individually tailored processing for each component.

As we discuss in some detail in the following sections, we point to two central issues in the current cardiac ultrasound processing: the very large amounts of data generated by the scanning systems, and side lobe artifacts that affect the image quality. We demonstrate how, by using our approach, an over twenty-fold reduction of the data size can be achieved while retaining the image contents and improving its quality by removing the side lobe artifacts.

In order to familiarize the readers with the ultrasound imaging process and set the foundations for understanding our proposed approach, we begin by introducing the common modeling and processing of ultrasound signals. Section II describes the various models used in the literature and the two components model we propose here, and Section III describes the processing commonly used in current cardiac scanning systems. In Section IV we present our approach which consists of the decomposition step, the individual processing step and the recombining step. Each of these steps is discussed in detail. Then, in Section V we demonstrate, through experiments, the benefits of using our approach in terms of both data compression and image quality, as claimed above.

II. Modeling Ultrasound Data

Medical ultrasound imaging allows visualization of internal body structures by radiating them with acoustic energy and, in most systems, analyzing the returned echoes. The generated two-dimensional image typically comprises of multiple one-dimensional scan lines, each constructed by integrating the data collected by the transducer elements following the transmission of an acoustic pulse along a narrow beam. As the transmitted pulse propagates through the body, echoes are scattered

by acoustic impedance perturbations in the tissue. These backscattered echoes are detected by the transducer elements and combined, after aligning them with the appropriate time delays, in a process referred to as beamforming, which results in Signal-to-Noise Ratio (SNR) enhancement. Each resulting beamformed signal forms a line in the image. This process will be described in more detail in the sequel.

Throughout the paper, we refer to the data collected by the individual sensors as the raw data.

The processing of ultrasound signals or images is often based on an underlying signal model, and modeling this data has been a challenge for some time. Resulting from interferences of randomly spaced scatterers, the ultrasonic echoes have a stochastic nature which is manifested in the resulting image as speckle phenomena. In the context of denoising or despeckling, various statistical models for speckle have been proposed that are inspired by the physical process underlying speckle formation. A commonly used model for the ultrasonic echo assumes that any range cell contains a large number of randomly located scatterers, hence the envelope of the echo obeys a Rayleigh distribution [11]. When the number of scatterers is not large enough or the scatterers are not randomly located, deviations from the Rayleigh model may occur. To account for such deviations, other models have been proposed, such as the Rician distribution [12], [13], K-distribution [14], and others.

Beyond the statistical models, several works (e.g. [15], [16]) are based on parametric spectral analysis and model the ultrasonic radio frequency (RF) signal as an autoregressive process (AR), i.e. as an output of a linear filter driven by a white Gaussian noise.

Another suggested model [17], [18] considers the ultrasound RF image to be a result of the convolution of the tissue reflectivity function with the two-dimensional point-spread function (PSF) of the imaging system.

Some other, less common, models have also been proposed, such as a local polynomial representation of the ultrasonic pulse log-spectrum [19] or modeling of the RF echo as a power-law shot-noise process [20].

A more popular approach is based on a sparsity assumption, i.e. considering the signals to be sparse under some basis or dictionary. Explicitly, one seeks an approximation $\hat{\mathbf{y}} = \mathbf{D}\hat{\mathbf{x}}$ of the signal $\mathbf{y} \in \mathbb{R}^N$ that obeys

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_0 \leq T_0, \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^K$ is the sparse representation vector, $\|\mathbf{x}\|_0$ is the ℓ_0 pseudo-norm counting the non-zero entries in the vector $\mathbf{x}$, and the matrix $\mathbf{D} \in \mathbb{R}^{N \times K}$ is the representation dictionary. This dictionary may be either predetermined or learned from the data itself, in a process referred to as dictionary learning. The underlying idea is an adaptive adjustment of the dictionary to observed signal characterizations, by solving

$$\arg\min_{\mathbf{D},\mathbf{X}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 \quad \text{subject to} \quad \|\mathbf{x_i}\|_0 \leq T_0 \quad \forall i, \quad (2)$$

where $\mathbf{Y} \in \mathbb{R}^{N \times M}$ is a data matrix having the observed signals as its columns, and $\mathbf{x_i}$ is the i-th column of the coefficients matrix $\mathbf{X} \in \mathbb{R}^{K \times M}$.

Based on the sparsity model, Quinsac *et al.* [21], [22] suggested using Fourier or Wavelets as the representation basis. Friboulet *et al.* [23], [24] used a dictionary of directional wave atoms, which shows good properties for sparsely representing warped oscillatory patterns. The authors showed that following random sparse sampling, reconstruction using the wave atoms dictionary yields better results compared with the Fourier or Wavelets bases.

Another set of works [25], [26] assumed sparsity by modeling the ultrasonic echo as a finite stream of strong pulses. This model suggests that the echo signal $y(t)$ consists of L pulses that are replicas of a known-shape pulse $h(t)$ with unknown delays (time-shifts) t_l and amplitudes a_l:

$$y(t) = \sum_{l=1}^{L} a_l h(t - t_l). \quad (3)$$

Few works attempted to select the representation basis for the ultrasound *images (or image patches)* using dictionary learning methods [27]–[29]. To our knowledge, no similar attempts were made to apply dictionary learning approaches to one-dimensional ultrasound raw data.

In this paper, we propose a new model for the raw ultrasonic signals as the sum of two components, both carrying valuable information: the strong reflectors, which are crucial for tracking purposes in cardiac ultrasound imaging [30], [31], and the speckle, also referred to in this work as the background component, that characterizes the microscopic structure of the tissue [11].

We choose to model the strong reflectors using a stream of pulses model, similar to the one proposed by [25], [26] mainly, because of its relative simplicity. Our experiments indicate that indeed, it serves our purpose very well. For the background component, we learn a suitable dictionary using the K-SVD method [32], as we describe in the sequel.

III. Conventional Processing in Ultrasound Imaging

We focus on two problems associated with the described imaging process. The first is the very large amount of data generated on the array of elements at the probe. By way of demonstration, a typical system in cardiac 2D imaging would generate around 3 Gbits/sec. This size of data necessitates compression as we describe in the sequel. The second problem we highlight has to do with the mentioned beamforming process. This process is known to result in artifacts in the image, compromising its quality. We will show the potential benefit of using our proposed approach in addressing both these problems.

In the following, we discuss the common approach which is currently used in relation to the two mentioned problems.

A. Data Compression

According to the Shannon-Nyquist sampling theorem, the sampling rate at each transducer element should be at least twice the bandwidth of the detected signal. Taking into account the high frequency used for ultrasound imaging, the number of

transducer elements and the number of lines in an image, the amount of data needed to be transferred and processed is very large. This motivated methods to reduce the amount of needed data without compromising the reconstructed image quality and its diagnostic credibility. In state-of-the-art ultrasound systems, an initial reduction of data size is already made possible by exploiting the fact that the detected signals are modulated onto a carrier and occupy only a portion of the entire bandwidth. Therefore, by demodulating the received signals, the sampling rate and consequently the data size, may be reduced. Nevertheless, as we will demonstrate, this reduction is far from fulfilling the compression potential.

To further reduce the amount of data, a number of approaches were proposed in the literature. Several works have recently adapted compressed sensing (CS) methods to ultrasound signals (e.g. [21]–[24], [28], [33]–[35]). Other works (e.g. [36]) treat the collection of raw signals detected by the transducer elements for a single scan line as a two-dimensional image, and attempt to compress it using standard techniques such as JPEG compression.

Another work attempting to reduce the amount of sampled data in ultrasound signals, based on complementary ideas arising from the Finite Rate of Innovation (FRI) framework [37], [38], was carried out by Tur *et al.* [25]. The authors, as mentioned earlier, consider the returning echo to be a stream of pulses that are replicas of a known-shape pulse with unknown time-shifts and amplitudes. Assuming that overall L pulses were reflected back to the transducer from the pulse's propagation path, the detected signal is completely defined by $2L$ degrees of freedom, corresponding to the unknown time delays and amplitudes of these pulses. Based on the FRI framework, these $2L$ parameters are estimated and the signal recovered from a minimal subset of $2L$ of the signal's Fourier series coefficients. The needed coefficients are recovered from low rate samples of the analog signals, as the sampling frequency is now determined by the number of pulses L, which is rather small compared to the bandwidth of the transmitted pulse, leading to a substantial sample rate reduction. This work was extended by Wagner *et al.* [26], who applied concepts of the CS framework in order to directly reconstruct the beamformed signals from low rate samples of the individual signals detected by the different transducer elements.

These methods achieve an almost eight-fold reduction of sample rate, however the reconstructed data is partial as it only contains one component of the data, the macroscopic reflections, and does not include the speckle.

A different approach used in radar is based on the CLEAN algorithm [39]. This algorithm performs an iterative search of the strong reflections, where at each iteration the maximal amplitude of the returned signal is located and the signal is then attenuated around this detected maxima using some attenuation function. This process is repeated until the residual signal reaches a predefined threshold, after which the strong reflectors could be reconstructed from their locations and amplitudes collected across iterations. Like the previous methods, the main shortcoming of this approach is that it disregards speckle and only reconstructs the strong targets.

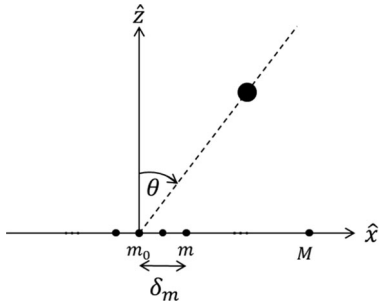


Fig. 2. Imaging setup.

B. Beamforming

Modern ultrasound transducers, as mentioned earlier, are comprised of multiple elements. This inherent redundancy can be utilized by averaging the signals detected by the different sensors after aligning them with the appropriate time delays. This process, referred to as receive beamforming, allows both dynamic control over the focus of the receiving array (and thus better localization of the scattering structures) and improvement of the Signal-to-Noise Ratio (SNR) [40]. In order to achieve dynamic focusing, the focal point is swept along the beam direction, where at each coordinate the appropriate delays can be computed assuming a point reflector exists at that point. For simplicity, let us assume that the transducer lies along the $\hat{x}$ axis, centered at the origin ($\hat{x} = 0$), and comprised of M elements, as depicted in Fig. 2. A reference receiver denoted m_0 is set centered at the origin, and δ_m denotes the distance measured from the origin to the center of the m-th receiver.

The imaging cycle starts at time $t_0 = 0$, when the array transmits an acoustic pulse into the tissue at some direction θ. The pulse propagates through the tissue at velocity c (the speed of sound) such that at time $t \geq 0$ its coordinates are $(\hat{x}, \hat{z}) = (ct \sin\theta, ct \cos\theta)$. Consider a potential reflection, originating at this coordinate and arriving at the m-th element. The distance traveled by such a reflection is:

$$d_m(t,\theta) = \sqrt{(ct\cos\theta)^2 + (ct\sin\theta - \delta_m)^2}. \tag{4}$$

The time this reflection takes to cross that distance is $d_m(t,\theta)/c$, so it reaches the m-th element at time:

$$\tau_m(t,\theta) = t + \frac{d_m(t,\theta)}{c}. \tag{5}$$

Plugging (4) into (5):

$$\begin{aligned}\tau_m(t,\theta) &= t + \frac{1}{c}\sqrt{(ct\cos\theta)^2 + (ct\sin\theta - \delta_m)^2} \\ &= t + \sqrt{t^2 + \left(\frac{\delta_m}{c}\right)^2 - 2\left(\frac{\delta_m}{c}\right)t\sin\theta}.\end{aligned} \tag{6}$$

Echoes are then reflected from density and propagation-velocity perturbations in the radiated medium and detected by the transducer elements. Denote by $\varphi_m(t,\theta)$ the signal detected by the m-th element. In order to compensate the differences in arrival time, an appropriate delay should be applied to $\varphi_m(t,\theta)$ before these echoes could be averaged. Let us denote the delayed signal detected by the m-th element by $\hat{\varphi}_m(t,\theta)$.

It is readily seen that $\tau_{m_0}(t,\theta) = 2t$. Hence, in order to align the echo received by the m-th element with the one received by the reference element, we require:

$$\hat{\varphi}_m(2t,\theta) = \varphi_m(\tau_m(t,\theta),\theta). \tag{7}$$

Therefore the time-aligned signal received by the m-th element is:

$$\hat{\varphi}_m(t,\theta) = \varphi_m\left(\tau_m\left(\frac{t}{2},\theta\right),\theta\right) \tag{8}$$

where using (6):

$$\tau_m\left(\frac{t}{2},\theta\right) = \frac{t}{2} + \frac{1}{2}\sqrt{t^2 + 4\left(\frac{\delta_m}{c}\right)^2 - 4\left(\frac{\delta_m}{c}\right)t\sin\theta}. \tag{9}$$

Finally, the beamformed signal is the average of the individual time-aligned echoes, i.e.

$$\Phi(t,\theta) = \frac{1}{M}\sum_{m=1}^{M}\varphi_m\left(\tau_m\left(\frac{t}{2},\theta\right),\theta\right). \tag{10}$$

Often the beamforming process designates different weights to different elements. This type of weighted averaging is commonly referred to as apodization and is used to influence the beam-profile shape.

In the sequel, to simplify the notation, we will drop the dependence on θ.

So far we have assumed continuous time signals. However, in modern ultrasound systems the receive beamforming process is performed digitally. Hence, the signals at the elements are sampled and we have $\varphi_m[n] = \varphi_m(nT)$ where T is the sampling interval. Using the sampled data for our beamforming raises a new problem, as we are now constrained to signal values on the grid nT. So, if the delay we need to apply is not an integer multiple of T, we approximate by linear interpolation

$$\begin{aligned}\varphi_m[n] &= \varphi_m\left(\tau_m\left(\frac{nT}{2}\right)\right)\\ &\approx (1-\alpha_m[n])\,\varphi_m[\tau_m[n]] + \alpha_m[n]\varphi_m[\tau_m[n]+1]\end{aligned} \tag{11}$$

where

$$\tau_m[n] = \left\lfloor \frac{\tau_m\left(\frac{nT}{2}\right)}{T} \right\rfloor \tag{12}$$

and

$$\alpha_m[n] = \frac{\tau_m\left(\frac{nT}{2}\right)}{T} - \tau_m[n]. \tag{13}$$

Then, the received beamforming in its digital form is

$$\Phi[n] = \frac{1}{M}\sum_{m=1}^{M}\left[\,1-\alpha_m[n]\,,\,\alpha_m[n]\,\right]\begin{bmatrix}\varphi_m[\tau_m[n]]\\ \varphi_m[\tau_m[n]+1]\end{bmatrix}. \tag{14}$$

While the described beamforming process became the standard in all ultrasound imaging systems, it comes with a price. There are a number of known artifacts which deteriorate the resulting image quality. Prominent among them is the side lobe artifact which will be demonstrated in some of the examples later on.

Though aspiring to transmit a narrow beam, the actual transmission is not solely concentrated in the main axis, but rather radiates a wide region according to a sinc-shaped beam profile, such that the beam has maximal energy on the axis but some of the energy is dispersed in side lobes.

The beamforming process concentrates the received beam by assuming that all the received reflections originated from the main axis. Consequently, echoes from a scatterer in the side lobe pathway are erroneously perceived coming from the main beam, resulting in image artifacts, manifested as smeared out echoes. These artifacts not only degrade the contrast of the image and impair its visual quality, but may also compromise its diagnostic credibility [41]. Side lobes suppression is usually treated by apodization, which degrades the lateral resolution, or by adaptive beamforming [42].

As the side lobe artifacts result from the presence of strong reflectors during the beamforming process, we are further motivated to apply the decomposition approach and perform beamforming only on the background component, after separating the strong reflectors.

IV. Component-Based Processing

There are four parts to an attempt of utilizing component based processing. First is identifying the distinct components in the signal, second is devising a method for decomposing the signal into these components, third is tailoring a processing chain for each component, and finally, demonstrating that the effort is worthwhile. Namely, that the result of the component based processing is superior to a direct processing of the combined signal.

As stated in Section II, our model assumes each raw signal to be composed of a background signal and a strong reflectors component. Having identified the distinct components, we next relate to each of the remaining parts in the context of ultrasound imaging.

A. Signal Decomposition

Next, we address the problem of decomposing each such signal to its individual components, the background signal and the strong reflectors, which are denoted y_b and y_s, respectively. The decomposition algorithm is based on a greedy detection of the strong reflectors followed by their separation from the original signal. We here propose two methods for performing the decomposition: one in the time-frequency domain and the other using the in-phase and quadrature representation.

1) Strong Reflectors Removal via Short-Time Fourier Transform: Modeling the strong reflectors component, we mostly adopt the "stream of pulses" model given by (3), according to which this component is composed of a limited number of strong pulses, that are amplified and delayed replicas of a known-shape pulse. This pulse has the form of a sinusoid signal oscillating at the transmission frequency f_0 in a Gaussian envelope.

As an extension to the previously proposed model, we suggest that the returning pulse shape is somewhat corrupted with

Algorithm 1: STFT-Based Decomposition.

Task: Decompose a given signal y into the strong reflectors and background components (y_s, y_b respectively)
Inputs: The signal's STFT $\mathbf{Y}(t,\omega)$, the STFT signature of the pulse model $\mathbf{H}(t,\omega)$ centered such that $\arg\max_{(t,\omega)}|\mathbf{H}(t,\omega)| = (0,0)$, the maximal number of pulses L, and an error threshold ϵ_0.
Initialization: Set the initial residual $\mathbf{r^0} = \mathbf{Y}(t,\omega)$
Main Iteration: for $k = 1, ..., L$ perform the following:

- Locate the strongest reflection (t_k, ω_k) with magnitude a_k:

$$(t_k,\omega_k) = \arg\max_{(t,\omega)}|\mathbf{Y}(t,\omega)| \quad ; \quad a_k = \mathbf{Y}(t_k,\omega_k)$$

- Residual update: $\mathbf{r^k} = \mathbf{r^{k-1}} - a_k\mathbf{H}(t-t_k,\omega-\omega_k)$
- Stopping Rule: If $\|\mathbf{r^k}\|_\infty < \epsilon_0$, stop. Otherwise, apply another iteration.

Output: $y_s(t) = ISTFT\left(\sum_{j=1}^{k} a_j\mathbf{H}(t-t_j,\omega-\omega_j)\right)$, $y_b(t) = ISTFT(\mathbf{r^k})$.

respect to the transmitted pulse. This corruption may be manifested in either a frequency shift, resulting from the frequency dependent attenuation [43], or a phase shift formed between the carrier wave and the Gaussian envelope.

In order to account for these possible corruptions, we have previously proposed in [4] to represent the strong reflectors in a time-frequency domain, using the Short-Time Fourier Transform (STFT). This will allow simultaneous optimization of both the time-delay and frequency shift.

Denoting the STFT of $y(t)$ by $\mathbf{Y}(t,\omega)$, the STFT decomposition is

$$\begin{aligned}\mathbf{Y}(t,\omega) &= \mathbf{Y_b}(t,\omega) + \mathbf{Y_s}(t,\omega)\\ &= \mathbf{Y_b}(t,\omega) + \sum_{k=1}^{L} a_k\mathbf{H}(t-t_k,\omega-\omega_k). \qquad (15)\end{aligned}$$

The proposed algorithm for STFT-based separation of the strong reflectors is described in Algorithm 1. Upon reaching the stopping condition, we are left with two separate signals in the time-frequency domain. By applying inverse STFT, the time-domain components can be reconstructed.

In the presented algorithm, the strong reflectors are detected as the maximal magnitude peaks of the STFT. In practice, the detection can be improved by matching the known pulse pattern in a narrow region around each peak. The maximal number of pulses L and error threshold ϵ_0 are chosen empirically.

It can be observed that the strong reflectors are naturally compressed as only the pulse model parameters $\{a_k, t_k, \omega_k\}_{k=1}^{L}$ are saved along the decomposition process. This may be thought of as sparse coding over a very large dictionary whose atoms represent all the possible time and frequency shifts of the known pulse. However, due to the high sampling rate of the signals and the enormous dimensions of such dictionary, standard pursuit techniques like OMP [44] are not feasible and an alternative amplitude-based pulse matching was performed here.

2) Strong Reflectors Removal in I/Q: The ultrasonic RF signal can be presented as a complex, low-frequency baseband signal, modulated by a carrier wave that oscillates at the much higher transmission frequency $\omega_0 = 2\pi f_0$. The resulting RF signal takes the form

$$y_{RF}(t) = I(t)\cos(\omega_0 t) - Q(t)\sin(\omega_0 t) \qquad (16)$$

and the baseband signal is assumed to be of the form

$$y_{IQ}(t) = I(t) + jQ(t) \qquad (17)$$

where $I(t), Q(t)$ are slowly varying (in comparison to f_0) Inphase and Quadrature components, and t is the time along the beam. I/Q demodulation is therefore the process of extracting the I/Q signal (17) from the measured physical RF signal (16).

Since the baseband components I/Q, which contain all the relevant information, change slowly, they can be sampled at a lower frequency, thus reducing the amount of data to be saved. Utilizing this fact, modern ultrasound systems sample the complex baseband signal after I/Q demodulation instead of directly sampling the received RF signal. Beside the advantage of a lower sampling rate, identifying the strong reflections at signals of lower frequency is less susceptive to errors such as time inaccuracies or phase shifts. Consequently, it would be beneficial to carry out the proposed decomposition and detect the strong reflectors at the I/Q components.

Working in baseband, there is no longer a need to match the full pulse model (including the frequency and phase estimation) and only a Gaussian envelope should be discovered, with the appropriate delay and amplitude.

The algorithm performed for each raw signal independently is generally similar to the one implemented in STFT, with few differences: The inputs are the baseband signal $\varphi_{\mathbf{IQ}}(\mathbf{t})$ and the modeled pulse envelope $\mathbf{g}(\mathbf{t})$, instead of the STFT of the signal and pulse. Respectively, the residual is initialized with $\mathbf{r^0} = \varphi_{\mathbf{IQ}}(\mathbf{t})$ and updated by subtracting the adjusted envelope from I and Q individually:

$$\mathbf{r^k_{I,Q}} = \mathbf{r^{k-1}_{I,Q}} - a^k_{I,Q}\mathbf{g}(\mathbf{t}-\mathbf{t_k}).$$

Additionally, the saved parameters for the strong reflectors now contain only the time delays $\{t_j\}_{j=1}^{k}$ and respective amplitudes, and the final residual $\mathbf{r^k}$ represents the baseband background component.

3) Data Separation Modification to Enhance Side Lobe Artifact Suppression: When evaluating the ability of our decomposition methods to reduce side lobe artifacts, we note that the success of these methods depends on the gain of the side lobe reflections with respect to its surroundings. Recall that both proposed decomposition methods are based on an amplitude threshold as a stopping criterion. Therefore, if the side lobe reflection is weak compared with speckle reflections along the same beam, no amplitude threshold will enable its detection as a strong reflector.

In order to account for such cases as well, we propose the following modification to our decomposition methods. A greedy search similar to the one previously described is applied to each of the raw signals. In each iteration, the best match for the strong pulse that is above the amplitude threshold is detected

Algorithm 2: Modified Decomposition Algorithm for the l-th scan line (example given for the STFT-based method).

for each sensor $m = 1, ..., M$
- run Algorithm 1 and obtain the background $\mathbf{r^k}$ and strong reflectors parameters $\{t_j, \omega_j, a_j\}_{j=1}^{k}$.
- for each identified pulse, adjust its time delay to all sensors and to adjacent scan lines:

for pulse $j = 1, ..., k$
for line $q = l-1, l, l+1$
for sensor $s = 1, ..., M$
– compute t_s^j using (20)
– if t_s^j is a local maxima for line q and sensor s, with amplitude $a_{q,s}$, subtract $a_{q,s}\mathbf{H}(t - t_s, \omega - \omega_j)$ from the relevant residual

and subtracted using either the STFT of I/Q based method. Unlike the previous methods, before proceeding to the next iteration, the time-delay of the detected pulse is here adjusted for all the other sensors and for the adjacent scan lines. For these other signals, a pulse will be matched and subtracted if a local maxima occurs at the adjusted expected time, regardless of the predefined threshold. The time adjustment is calculated similarly to the delay computed for the purpose of receive beamforming (see Section III-B).

Explicitly, if a point reflector is placed at distance r and direction θ from the array center (see Fig. 2), then its distance from the m-th sensor, positioned along $\hat{z} = 0$ at $\hat{x} = \delta_m$, is:

$$r_m = \sqrt{(r\cos\theta)^2 + (r\sin\theta - \delta_m)^2} \tag{18}$$

and the echo reflected from it arrives to the m-th sensor at time $t_m = t_{TX} + \frac{r_m}{c}$, where $t_{TX} = \frac{r}{c}$. From t_m we can extract:

$$r = \frac{c^2 t_m^2 - \delta_m^2}{2(ct_m - \delta_m \sin\theta)}. \tag{19}$$

Then, the expected time of arrival to the k-th sensor is:

$$t_k = t_{TX} + \frac{r_k}{c} = \frac{r}{c} + \frac{1}{c}\sqrt{(r\cos\theta)^2 + (r\sin\theta - \delta_k)^2}. \tag{20}$$

The modification of the decomposition algorithm as applied to the l-th scan line is summarized in Algorithm 2.

The modified method here presented has a significant impact in cases where the strong reflectors are placed in a speckled region and have visible side lobe artifacts, as discussed in the synthetic cyst phantom example described in Section V-B.

B. Tailored Processing

Having separated the two components, we next want to adjust their individual processing schemes. We focus on data compression and beamforming, as described earlier in Section III.

The strong reflectors are already compressed by saving only the pulse model parameters along the decomposition process presented in Section IV-A. Furthermore, this component does not need to go through any SNR enhancement in the form of beamforming.

As for the speckle (background) component, we propose to compress each such signal by sparsely representing it over an optimized dictionary. First, we train the dictionary offline using available prototype examples of background signals. Then, we apply this dictionary to obtain the sparse representation for the online background data. It should be emphasized that this training process is only performed once for every imaging system settings, and need not be repeated for every analyzed signal or even for every imaged frame. However, clearly, the resulting dictionary does depend on the system used and its particular setting.

In the following, we elaborate on the compression and beamforming steps as tailored for the background component.

1) Learning a Dictionary for the Background Data: For the purpose of dictionary learning, we shall use the K-SVD algorithm [32]. We chose the training set to be a subset of the signals constituting a single frame of real cardiac imaging data. Although our goal is to compress raw signals, i.e. signals detected by each sensor prior to receive beamforming, the training set signals are chosen to be beamformed scan lines, since those were shown to have improved SNR [41].

The input frame to the K-SVD consists of 120 scan lines, out of which 10 scan lines were randomly chosen for the training set. Each of these signals contains 3328 samples and was divided into one-dimensional, non-overlapping patches of 100 samples each. Initializing the dictionary with a random matrix, the training process was limited to 10 iterations.

It should be noted that the chosen training set is relatively small. Moreover, overlapping patches are commonly used due to the fact that the different sparse approximations could eventually be averaged at the overlapping segments, which could lower the representation error and smooth the overall signal reconstruction. Despite this apparent advantage, we chose to reduce the amount of patches (and thus the runtime of both training and representation) by using non-overlapping patches and achieved very satisfactory results. As we show later, by training a dictionary with a small subset of *beamformed* signals of a single frame, one can get a good representation for the *raw echoes* of all the sensors and all the scan lines in the frame, as well as of raw signals from other frames obtained with the same system (for the same settings).

2) Sparse Representation of the Background Data: Having the dictionary trained offline, let us return to the online processing cycle. After decomposition, the separated background component undergoes further processing, starting with compression.

Due to computational limitations of the K-SVD algorithm, the dictionary is learned for small one-dimensional patches. Thus each background signal is first divided into non-overlapping patches, similar to the division of the training set signals described in the previous section. Using the Orthogonal Matching Pursuit (OMP) algorithm, each patch is then sparsely represented over the trained dictionary.

Denote the patch length by Q and the trained dictionary by $\mathbf{D} \in \mathbb{R}^{Q\times K}$, then using OMP we solve for each patch

$$\arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2 \leq \epsilon \tag{21}$$

TABLE I
FIELDII SIMULATION PARAMETERS

Parameter	Value
c (Speed of Sound)	1540 m/sec
f_0 (Central Frequency)	3.5 MHz
f_s (Sampling Frequency)	16 MHz
Element Width	0.22 mm
Element Height	5 mm
Kerf	0.055 mm
Number of Elements	64
Apodization	none
TX Focus	70 mm

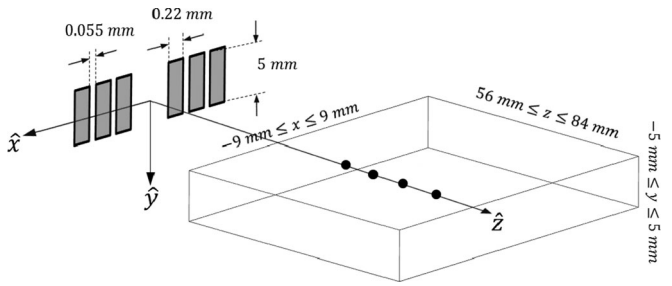


Fig. 3. FieldII simulation setup for the point reflectors phantom.

where $\mathbf{y}$ is the $Q \times 1$ extracted patch and $\mathbf{x}$ is its corresponding $K \times 1$ sparse coefficients vector.

Having obtained the sparse representations of all the individual patches of a given RF signal, the signal could be reconstructed. For that matter, each patch is recovered from its sparse coefficients by $\hat{\mathbf{y}} = \mathbf{D}\mathbf{x}$. The recovered patches are then plugged back to the locations from which they were extracted, to reassemble the full signal. This process is repeated for each of the raw signals detected by the different transducer elements for the same scan line. Afterwards, standard beamforming techniques can be applied to generate the beamformed background signal, which could then be further processed to form the image. However, we propose an alternative reconstruction scheme based on performing the beamforming in the representation domain, thus constructing the beamformed background component signal directly from the sparse coefficients vectors of all the individual patches, rather than reconstructing each of the sensor echoes separately. The formulation of the representation domain beamforming is presented in the next section.

3) Beamforming in the Representation Domain: As previously mentioned, each sampled data vector from each element has been divided into one dimensional patches (sub-vectors) and for each we have found its sparse representation over the learned dictionary.

Let us denote by $\boldsymbol{\varphi}_m \in \mathbb{R}^N$ the background component of the signal received by the m-th sensor. Each such component is divided into P patches $\mathbf{y}_{m,p} \in \mathbb{R}^Q$ of length Q:

$$\begin{aligned}\boldsymbol{\varphi}_m &= \left[\varphi_m[0] \quad \varphi_m[1] \quad \dots \quad \varphi_m[N-1]\right]^T \in \mathbb{R}^N \\ &= \left[\mathbf{y}_{m,1}^T \quad \mathbf{y}_{m,2}^T \quad \dots \quad \mathbf{y}_{m,P}^T\right]^T. \end{aligned} \tag{22}$$

For simplicity here we assume that the patches do not overlap. However, we note that a similar derivation can be made for overlapping patches.

Let $\mathbf{D} \in \mathbb{R}^{Q \times K}$, $K > Q$ denote the learned dictionary, then using OMP we solve for each patch ($\forall 1 \le m \le M$, $1 \le p \le P$):

$$\arg\min_{\mathbf{z}_{m,p}} \|\mathbf{z}_{m,p}\|_0 \quad \text{subject to} \quad \|\mathbf{y}_{m,p} - \mathbf{D}\mathbf{z}_{m,p}\|_2 \le \varepsilon_0 \tag{23}$$

where $\mathbf{z}_{m,p} \in \mathbb{R}^K$ but $\|\mathbf{z}_{m,p}\|_0 << Q$ (namely, the number of non-zero entries of all $\mathbf{z}_{m,p}$ is considerably smaller than the patch size). Following, each patch is reconstructed by $\hat{\mathbf{y}}_{m,p} = \mathbf{D}\mathbf{z}_{m,p}$. Using (22) we can write

$$\boldsymbol{\varphi}_m = \begin{bmatrix} \mathbf{D} & 0 & \dots & 0 \\ 0 & \mathbf{D} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{m,1} \\ \mathbf{z}_{m,2} \\ \vdots \\ \mathbf{z}_{m,P} \end{bmatrix} = \widetilde{\mathbf{D}}\mathbf{z}_m \tag{24}$$

where $\widetilde{\mathbf{D}} \triangleq \mathbf{I}_P \otimes \mathbf{D}$ and $\otimes$ denotes the Kronecker tensor product, namely $\mathbf{A} \otimes \mathbf{B} = [a_{ij}\mathbf{B}]_{i,j}$ for matrices $\mathbf{A}$ and $\mathbf{B}$.

Substituting in (14) we obtain for the beamformed signal

$$\begin{aligned}\Phi[n] &= \frac{1}{M}\sum_{m=1}^{M} \begin{bmatrix} 1-\alpha_m[n] & \alpha_m[n] \end{bmatrix} \begin{bmatrix} \mathbf{e}_{\tau_m[n]}^T \\ \mathbf{e}_{\tau_m[n]+1}^T \end{bmatrix} \widetilde{\mathbf{D}}\mathbf{z}_m \\ &= \frac{1}{M}\sum_{m=1}^{M} \mathbf{b}_m[n]^T \mathbf{z}_m \end{aligned} \tag{25}$$

where $\mathbf{e}_l$ is the l-th column of the $N \times N$ identity matrix $\mathbf{I}_N$ and $\mathbf{b}_m[n] \in \mathbb{R}^{KP}$, and we recall that $\tau_m[n]$ are indices from $(1,\ 2, ..., N)$.

Finally, denote:

$$\mathbf{B}_m = \left[\mathbf{b}_m[1] \quad \mathbf{b}_m[2] \quad \dots \quad \mathbf{b}_m[N]\right]^T \in \mathbb{R}^{N \times PK}, \tag{26}$$

$$\mathbf{H} = \frac{1}{M}\left[\mathbf{B}_1 \quad \mathbf{B}_2 \quad \dots \quad \mathbf{B}_M\right], \tag{27}$$

$$\mathbf{Z} = \left[\mathbf{z}_1^T \quad \mathbf{z}_2^T \quad \dots \quad \mathbf{z}_M^T\right]^T. \tag{28}$$

Then the beamformed signal is given by

$$\boldsymbol{\Phi} = \mathbf{H}\mathbf{Z}. \tag{29}$$

It is readily seen that $\mathbf{H}$ depends solely on the imaging settings and the learned dictionary. This result indicates that beamforming in the representation domain is manifested as a weighted combination of the representation coefficients, where the weight matrix $\mathbf{H}$ is data independent and can be computed apriori for each scan line direction θ. Therefore, only the sparse coefficients should be transferred to the beamformer, leading to a much simplified computation of the beamformed signal.

V. SIMULATION AND EXPERIMENT RESULTS

We are at the fourth part of utilizing the data decomposition concept in medical ultrasound, which is showing that it is worthwhile. We have conducted experiments using simulated data and real cardiac data. In the following, we describe in some detail the setup and information pertaining to these experiments.

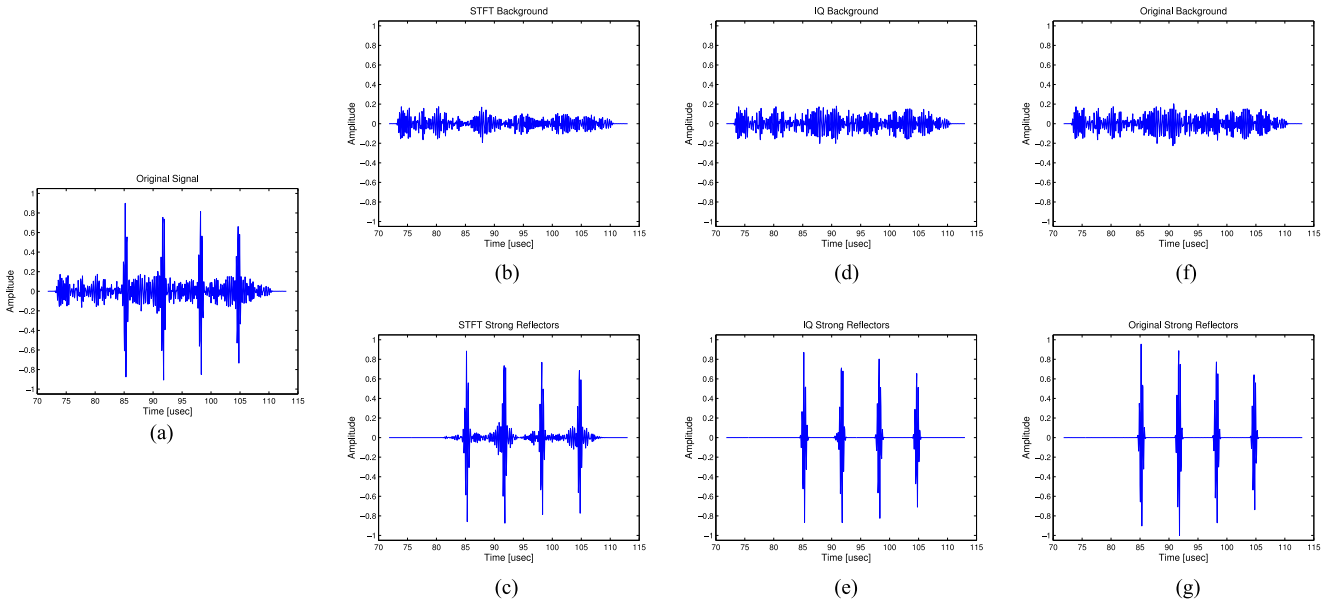


Fig. 4. Point reflectors RF signal with its decomposition: (a) Received RF signal, (b) Background component (STFT decomposition), (c) Strong reflectors (STFT decomposition), (d) Background component (I/Q decomposition), (e) Strong reflectors (I/Q decomposition), (f) Background component (ground truth), (g) Strong reflectors (ground truth).

A. 1D Simulation Results

To evaluate the decomposition algorithms, both the STFT and I/Q based methods for removing the strong reflectors were first tested on a single scan line simulated using the FieldII simulation program [45].

We created an aperture comprising 64 transducer elements, with central frequency $f_0 = 3.5$ MHz. The width of each element, measured along the $\hat{x}$ axis, was $\frac{1}{2}\lambda = \frac{c}{2f_0} = 0.22$ mm, and its height, measured along the $\hat{y}$ axis, was 5 mm. The elements were arranged along the $\hat{x}$ axis, with a 0.055 mm inter-element spacing (kerf). The transmitted pulse was simulated by exciting each element with two periods of a sinusoid of frequency f_0, where the delays were adjusted such that the transmission focal point was at depth $r = 70$ mm. No apodization was used.

The simulation parameter settings are summarized in Table I.

The simulation setup is illustrated in Fig. 3.

A single scan line was simulated for a steering angle $\theta = 0^\circ$. The simulated phantom consists of speckled background with 4 strong reflectors positioned 5 mm apart along the $\hat{z}$ axis ($\theta = 0^\circ$). The speckled region was constructed by randomly drawing 10^5 point reflectors distributed uniformly in the three dimensional region given by $\{(x, y, z) : |x| \leq 9 \text{ mm}, |y| \leq 5 \text{ mm}, |z - 70| \leq 14 \text{ mm}\}$. The corresponding amplitudes were also drawn at random according to a Normal distribution with zero-mean and unit-variance. Then, the strong reflectors were added along the $\hat{z}$ axis at depths 65 mm, 70 mm, 75 mm, 80 mm. The strong reflectors amplitudes were set to be 50 times the variance of the speckle reflectors amplitudes.

Fig. 4 depicts the RF echo received by one of the transducer elements with its decomposition results using both STFT and I/Q approaches. These decomposition results are compared with a ground truth obtained by individual simulation of the two components.

TABLE II
PSNR (IN [DB]) AND MSE RESULTS FOR THE POINT REFLECTORS PHANTOM. RESULTS ARE RELATIVE TO THE GROUND TRUTH BACKGROUND COMPONENT 4.

	STFT 4b	I/Q 4d	Original Signal 4a
PSNR	30.231	34.204	18.714
MSE	3.79e-40	1.52e-40	5.37e-39

Comparing the two approaches, it appears that the I/Q based method gives a better separation of the two components, where less of the background component is removed with the strong reflectors. The visual superiority is also reflected by the Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) calculated with respect to the ground truth decomposition, as summarized in Table II.

This result is consistent with our formerly mentioned expectations, since using lower frequencies and relieving the need for phase optimization makes the estimation less susceptible to errors.

Recall that PSNR is defined as $PSNR = 10\log_{10}\left(\frac{I_{\max}^2}{MSE}\right)$ where $I_{\max}$ is the maximal possible value of the compared signals or images (for images spanning the full 8-bit gray-scale, $I_{\max} = 255$) and $MSE = \mathbb{E}[(\mathbf{I} - \hat{\mathbf{I}})^2]$ with $\mathbf{I}$ and $\hat{\mathbf{I}}$ being the compared signals or images.

B. Cyst Phantom Results

The four methods for identifying and removing the strong reflectors (STFT, I/Q, modified STFT and modified I/Q) presented here, have been tested on a synthetic phantom simulated

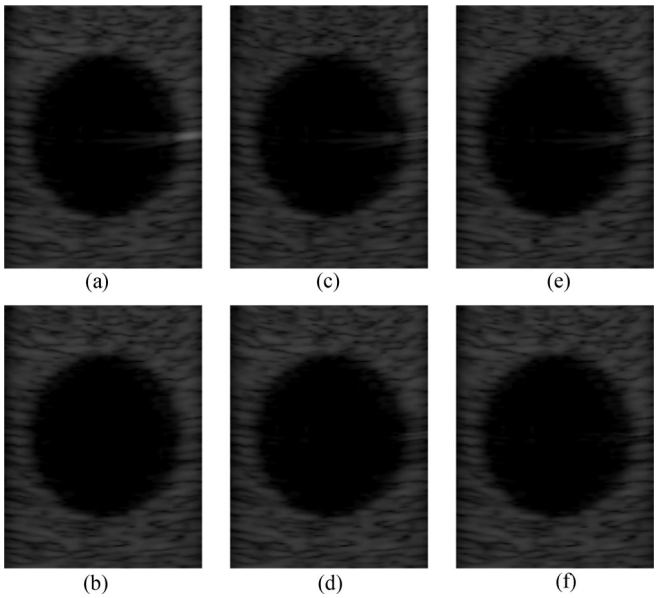


Fig. 5. Cyst phantom background estimation results: (a) Standard processed image, (b) Standard image without strong reflectors (ground truth), (c) STFT background image, (d) Modified STFT background image, (e) I/Q background image, (f) Modified I/Q background image.

using the FieldII program [45]. The data acquisition setup is summarized in Table I.

During this simulation, a 24° sector was imaged using 48 scan lines in Single Line Acquisition mode (SLA), i.e. a single reception line was computed for each transmission.

The phantom comprises of a large cyst in a speckled background, with a single strong reflector placed in the speckled region right besides the cyst.

The speckled region was constructed by randomly drawing 10^5 point reflectors distributed uniformly in the three dimensional region given by $\{(x, y, z) : |x| \leq 9$ mm, $|y| \leq 5$ mm, $|z - 70| \leq 14$ mm$\}$. The corresponding amplitudes were also drawn at random according to a Normal distribution with zero-mean and unit-variance. The cyst was then created by removing all point reflectors from a circle of radius 8.5 mm centered at $(x, y, z) = (0,0,70$ mm$)$. Finally, a single strong reflector was added at $(x, y, z) = (8.6$ mm$, 0,70$ mm$)$, i.e. at a depth of $r = 70.5$ mm at an angle of 7°, such that it lies exactly along the main direction of the 39-th scan line. The strong reflector's amplitude was set to be 100 times the variance of the speckle reflectors amplitudes.

The reflector was intentionally placed very close to the cyst (0.1 mm from its boundary). While the beam focus is aimed at the cyst, side lobes might insonify the strong reflector that is positioned outside it, and its reflection would be interpreted as if it originated from within the cyst. This causes undesired artifacts in the resulting image, as previously mentioned in Section IV-A3. Had the strong reflector been placed in a speckled region, its reflections might have blended with other weak reflections in its speckled surroundings, but in our example, due to the low echogenicity of the cyst, the side lobe reflections are expected to be easily observed.

TABLE III
PSNR (IN [DB]) AND SSIM RESULTS FOR THE CYST PHANTOM. RESULTS ARE RELATIVE TO THE GROUND TRUTH IMAGE 5 B.

	Orig. Image 5a	STFT 5c	Modified STFT 5d	I/Q 5e	Modified I/Q 5f
PSNR	34.56	33.28	40.08	36.03	43.27
SSIM	0.9925	0.9848	0.9947	0.9901	0.9957

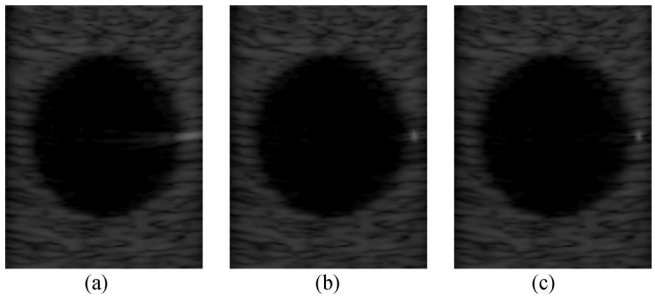


Fig. 6. Cyst phantom combined image results: (a) Standard processed image, (b) Combined image using modified STFT decomposition, (c) Combined image using modified I/Q decomposition.

Using this phantom we show that our methods, operating at the sensor level, are able to remove the side lobe reflections, so that these artifacts no longer exist, or are significantly decreased, in the output image.

We have generated a background image by repeating the simulation without the strong reflector, thus obtaining what we refer to as the ground truth image (Fig. 5(b)).

The background estimation results obtained for the cyst phantom are depicted in Fig. 5. Corresponding PSNR and SSIM[1] [46] values compared with the ground truth image are presented in Table III.

The reconstructed combined images, obtained after adding the estimated strong reflectors to the beamformed background signals, are depicted in Fig. 6. These results show that both our modified decomposition methods (based on either STFT or I/Q) yield good quality images, in which the smearing of the strong reflector due to side lobe artifacts is removed and replaced with the actual point reflector, such that the cyst boundary is no longer obscured.

Based on these results, the modified decomposition method presented in Section IV-A3 seems better in terms of removing side lobes. Nevertheless, this modification has its down-side. Depending on adjacent lines, this method can no longer run in parallel for all the sensor echoes of all the scan lines, but should rather run on a collection of such lines, thus dictating a longer runtime compared to the independent removal methods.

It should be emphasized that this simulated example represents the worst-case scenario. In some cases, a suitable amplitude threshold, i.e. one that separates the side lobe reflections from the background, may exist for the non-modified methods as well. Additionally, in more typical cases the strong reflectors are placed within echogenic tissues, such that the side lobe

[1]The Structural Similarity Index Measure (SSIM) is a commonly used measure for comparing the similarity between two images.

TABLE IV
CARDIAC IMAGING PARAMETERS

Parameter	Value
c (Speed of Sound)	1540 m/sec
f_0 (Central Frequency)	3.423 MHz
f_s (Sampling Frequency)	16 MHz
Element Width	0.29 mm
Number of Elements	64
Number of Scan Lines	120
Sector Size	75 degrees

reflections themselves, or any minor errors in their detection, are within the standard deviation of the surrounding speckle and thus unidentifiable in the resulting image. Our chosen example is therefore less tolerant to errors.

Another source of difficulty arises from our underlying assumption of the pulse shape. The assumed pulse model was developed for the main lobe reflection of the pulse, yet for more distant lines, representing side lobe echoes, the pulse shape is not guaranteed to remain undistorted.

Despite these limitations, our methods are successful at suppressing the side lobe artifacts.

C. Cardiac Data Results

In addition to synthetic phantoms, our methods were tested on several sets of consecutive frames of cardiac ultrasound data provided by GE Healthcare.

In this section we examine the results obtained by applying our methods (with $L = 10$) to raw RF data acquired and stored for cardiac images of a healthy consenting volunteer. The acquisition was performed using a breadboard ultrasonic scanner employing a 64-element phased array probe. Operating in second harmonic imaging mode, pulses were transmitted at 1.7 MHz, and the corresponding second harmonic signal, centered at 3.4 MHz, was then acquired. Data from all acquisition channels was sampled at 16 MHz and collected along 120 beams, forming a 75° sector. The maximal imaging depth was $z = 16$ cm, implying a cycle time of $T = 208$ μsec. The imaging settings are summarized in Table IV.

Noting that the typical frame rate is 30 fps and that 16 bits are allocated for each of the 3328 samples, combined with the information in Table IV, one can readily calculate the data rate is $\sim$3 Gbits/sec, as mentioned earlier.

In Fig. 7 we present, for two different frames, the images generated by standard processing as compared to the results of using our proposed approach.

First we want to evaluate the decomposition results. For this purpose, we present for each frame its strong reflectors image and background estimation as obtained using the STFT based decomposition method (the results obtained using the I/Q based decomposition are very similar and so we do not present them here). The images reconstructed for the separated background components (without compression), illustrated in Fig. 7(c) and (d), indicate that our proposed decomposition methods successfully detect and remove the strong reflections, thus producing a background image with relatively homogeneous regions.

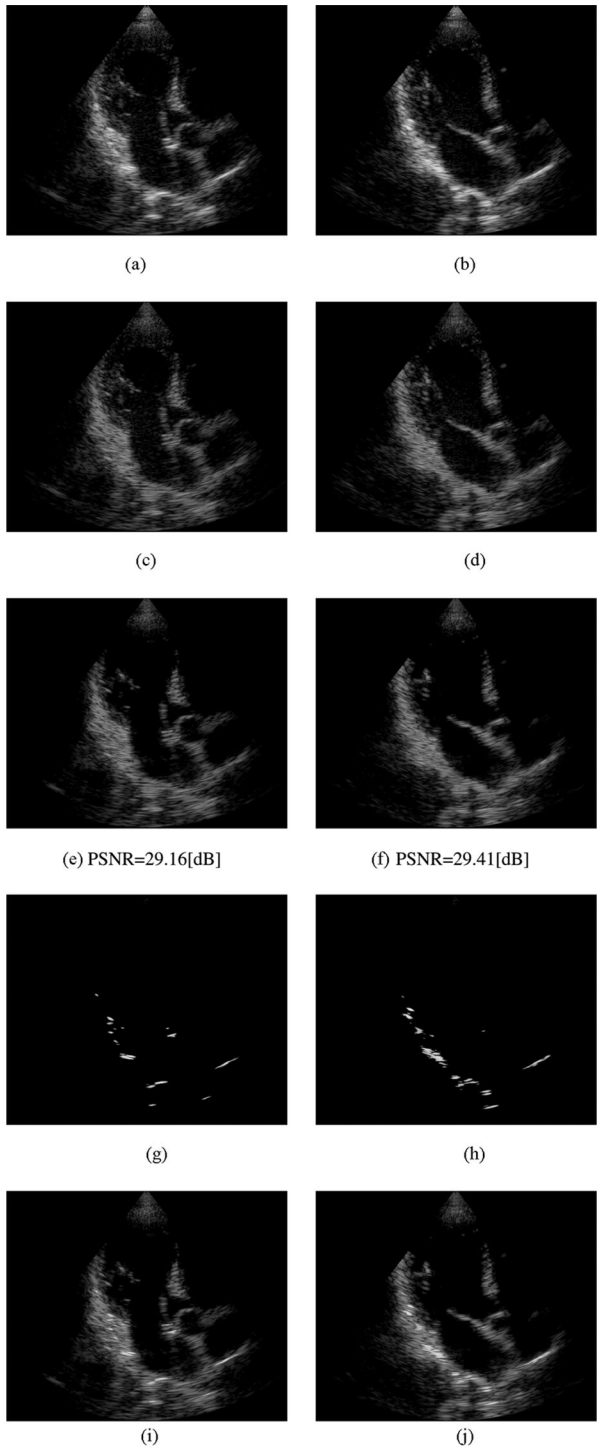


Fig. 7. Cardiac data results. The first column corresponds to frame 1, and the second corresponds to frame 2. (a),(b) Standard processed image. (c),(d) Estimated background image (STFT). (e),(f) Compressed background. (g),(h) Strong reflectors. (i),(j) Combined reconstruction.

Following decomposition, background compression was conducted according to the algorithm described in Section IV-B. Non-overlapping patches of length $Q = 100$ were used, the number of sensors is $M = 64$ and the RF signal length is $N = 3328$ samples. It should also be noted that in this experiment, apodization was applied to each raw signal before

summing them up, which is not reflected in the theoretical formulation of representation domain beamforming presented in Section IV-B3.

The 24-fold compressed versions of the background images for frame 1 and frame 2 are depicted in Fig. 7(e) and (f) (respectively). The corresponding PSNR values are 29.16[dB] and 29.41[dB]. Note that while data from frame 1 was used as the training set, we got comparable results when compressing frame 2, in terms of both visual quality and PSNR. This indicates that although the dictionary was trained based on signals from a single frame, it is suitable for representing data of other frames as well. Similar results were also obtained for other sets of cardiac ultrasound frames acquired using the same imaging settings.

In Section IV-B1 we pointed out a few challenges to the background compression process, namely the small size of the dictionary training set, the use of non-overlapping patches and the fact that training was performed using beamformed signals while intended to represent raw signals. Observing the compressed images, our proposed compression scheme seems to produce visually good images that preserve even the subtle image features. These results were obtained despite the formerly mentioned challenges, and while achieving an average compression ratio of 24, implying that the number of coefficients needed for reconstruction was only 4% of the number of time samples in the received RF signal.

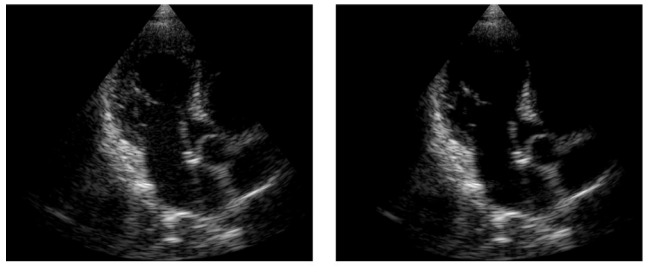
Finally, Fig. 7(i) and (j) present the combined images, which are the final outcomes of our suggested processing chain. While created using only 1/20 of the data (approximately), our outcome successfully retains the content of the standard processed image and improves its quality (note the reduced smear of the high intensity regions).

To further demonstrate the superiority of our chosen compression scheme, let us return to our initial conjecture motivating the raw signals decomposition, which was that individually compressing each component may lead to better compression ratios compared with a direct compression of the complete raw signal. In order to prove this hypothesis, we conducted a dictionary learning process using the same methods previously applied for the background signals, now using the original beamformed signals as the training set, without removing the strong reflectors. We then used this new dictionary for sparsely representing each of the detected raw signals.

For the sake of comparison, the same error threshold was used for the dictionary learning process, and the training set also consisted of 10 beamformed signals, taken from the same frame dataset.

The compression results obtained for the original frame data (without decomposition) over this specifically trained dictionary, are presented in Fig. 8. The resulting image is of comparable quality to the outcome of our decomposition based scheme (Fig. 7(i)) in terms of the visible amount of saved features. The benefit though is in the achieved compression ratio, which is over 2-fold better when utilizing decomposition.

Since the image generated by standard processing contains the strong reflectors information as well, a fair comparison of the compression ratios demands that the amount of coefficients needed for representing the strong reflectors is added to those used for representing the background signal. When the strong reflectors are considered, a slightly reduced compression ratio of 21.4 is achieved. Nonetheless, this achieved compression ratio is still twice as high as the one achieved for the original raw data.



(a) (b)

Fig. 8. Compression results for frame 1 without decomposition: (a) Original image, (b) Compressed image.

TABLE V
COMPARISON OF THE ACHIEVED COMPRESSION RATIOS

	% of coefficients	compression factor
background	4.07	24.56
background + strong reflectors	4.67	21.40
full raw signal	9.10	10.99

The comparison of achieved compression ratios is summarized in Table V. The first line presents information for the compression of the background only, the second presents information combined from separate compression of the background and the strong reflectors, and the third line presents information for the direct compression of the raw signals without decomposition.

It is readily seen that the individual compression of the background and strong reflectors components outperforms the direct compression of the raw data, in accordance with our predictions.

In this regard, it should be pointed out that the compression ratio considering both the background signal and strong reflectors, is relevant for analyzing the total amount of saved data. However, in terms of the data needed to be employed in beamforming computations, the higher compression factor (that only considers the background) is still applicable, as our proposed processing scheme suggests that the strong reflectors need not undergo beamforming for their reconstruction. Moreover, recall that the beamforming process itself was simplified by conducting it in the representation domain, thus achieving further reduction of the computational load besides the use of a smaller amount of coefficients.

VI. CONCLUSION

In this work, we have promoted the concept of data decomposition based processing in medical ultrasound imaging and demonstrated its potential benefits. Specifically, we first recognized the two components in the ultrasound raw data. In doing so, we extended previous models proposed in [25], [26] by integrating the speckle reflections and assembling a direct sum of two components, each of which carries valuable information

and could be characterized by a limited amount of parameters. Next, we devised algorithms to separate the data into these components, the strong reflections and the background data.

After the signal decomposition stage, the separated background component undergoes further processing. First, it is sparsely represented over a suitable dictionary, that was trained offline from background signal examples. Afterwards, the compressed background signals are integrated in a modified beamforming process, applied in the representation domain.

Finally, the strong reflectors could be reconstructed from their sparse coefficients directly into the beamformed signal, and combined with the reconstructed background component in order to construct the complete signal.

While the component based processing is not new, the novelty of our paper is in the specific application of medical ultrasound imaging. An important application that gains from this approach and the derived processing scheme, is the reduction of the amount of data needed to be transferred from the system front-end and processed by the beamformer. Applying our processing schemes to real cardiac ultrasound data, we successfully reconstruct the image contents while achieving over twenty-fold reduction of the data size. For comparison, this compression rate is twice as high as the one achieved by an equivalent compression of the raw signal without decomposition.

We point out that designed for the raw signals, the online algorithm may be applied at the sensors immediately after sampling, and does not require the full frame to be acquired.

Moreover, it is clearly desirable to compress the data as early in the processing chain as possible. As far as digital compression is concerned, our approach operates on raw signals "close to the source", i.e. immediately after sampling. Though not yet attempted in the scope of our work, we believe that utilizing the proposed two-component model and learned dictionary, a low rate sampling scheme can be established, such that our algorithm may be extended to the compressed sensing framework. Doing so, our results could be compared with other ultrasound compression techniques currently employed in the analog domain.

An additional contribution of our work relates to the resulting suppression of side lobe artifacts. We show that by separating the strong reflectors at early stages of the imaging cycle, before the receive beamforming, side lobe artifacts are significantly reduced alongside the data size reduction, thus improving the contrast of the reconstructed image and its diagnostic value.

Finally, we note that the component-based modeling may open more possibilities for analyzing ultrasonic signals. While we identified two main components, other decomposition ideas may be investigated, such as separating the first- and second-harmonic echoes, or detecting more than two components related to various artifacts which require special processing.

REFERENCES

[1] H. Krim, D. Tucker, S. Mallat, and D. Donoho, "On denoising and best signal representation," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2225–2238, Nov. 1999.

[2] G. Ghimpeţeanu, T. Batard, M. Bertalmío, and S. Levine, "A decomposition framework for image denoising algorithms," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 388–399, Jan. 2016.

[3] F. G. Meyer, A. Z. Averbuch, and R. R. Coifman, "Multilayered image representation: Application to image compression," *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 1072–1080, Sep. 2002.

[4] Y. Yankelevsky, A. Feuer, and Z. Friedman, "Decomposition-based compression of ultrasound raw-data," in *Proc. Int. Conf. Latent Variable Anal. Signal Separation*, 2015, pp. 522–529.

[5] S. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 2091–2110, Dec. 1989.

[6] C. P. Lamarche, P. Paultre, J. Proulx, and S. Mousseau, "Assessment of the frequency domain decomposition technique by forced-vibration tests of a full-scale structure," *Earthquake Eng. Struct. Dyn.*, vol. 37, no. 3, pp. 487–494, 2008.

[7] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. no. 72.

[8] Z. Li and J. Zheng, "Edge-preserving decomposition-based single image haze removal," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5432–5441, Dec. 2015.

[9] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 882–889, Aug. 2003.

[10] T. Shoham, D. Malah, and S. Shechtman, "Quality preserving compression of a concatenative text-to-speech acoustic database," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 3, pp. 1056–1068, Mar. 2012.

[11] C. B. Burckhardt, "Speckle in ultrasound B-mode scans," *IEEE Trans. Sonics Ultrason.*, vol. 25, no. 1, pp. 1–6, Jan. 1978.

[12] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," *IEEE Trans. Sonics Ultrason.*, vol. 30, no. 3, pp. 156–163, May 1983.

[13] M. F. Insana, R. F. Wagner, B. S. Garra, D. G. Brown, and T. H. Shawker, "Analysis of ultrasound image texture via generalized Rician statistics," *Proc. SPIE*, vol. 556, pp. 153–159, Nov. 1985.

[14] E. Jakeman and R. J. A. Tough, "Generalized K distribution: A statistical model for weak scattering," *J. Opt. Soc. Amer. A*, vol. 4, no. 9, pp. 1764–1772, Sep. 1987.

[15] J. Girault, F. Ossant, A. Ouahabi, D. Kouame, and F. Patat, "Time-varying autoregressive spectral estimation for ultrasound attenuation in tissue characterization," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 45, no. 3, pp. 650–659, May 1998.

[16] O. V. Michailovich and D. Adam, "A novel approach to the 2-D blind deconvolution problem in medical ultrasound," *IEEE Trans. Med. Imag.*, vol. 24, no. 1, pp. 86–104, Jan. 2005.

[17] T. Taxt, "Restoration of medical ultrasound images using two-dimensional homomorphic deconvolution," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 42, no. 4, pp. 543–554, Jul. 1995.

[18] O. V. Michailovich and A. Tannenbaum, "Despeckling of medical ultrasound images," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 53, no. 1, pp. 64–78, Jan. 2006.

[19] D. Adam and O. V. Michailovich, "Blind deconvolution of ultrasound sequences using nonparametric local polynomial estimates of the pulse," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 118–131, Feb. 2002.

[20] M. Kutay, A. Petropulu, and C. Piccoli, "On modeling biomedical ultrasound RF echoes using a power-law shot-noise model," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 48, no. 4, pp. 953–968, Jul. 2001.

[21] C. Quinsac, A. Basarab, J.-M. Girault, and D. Kouame, "Compressed sensing of ultrasound images: Sampling of spatial and frequency domains," in *Proc. IEEE Workshop Signal Syst.*, 2010, pp. 231–236.

[22] C. Quinsac, A. Basarab, and D. Kouame, "Frequency domain compressive sampling for ultrasound imaging," *Adv. Acoust. Vib.*, vol. 12, pp. 1–16, 2012.

[23] D. Friboulet, H. Liebgott, and R. Prost, "Compressive sensing for raw RF signals reconstruction in ultrasound," in *Proc. IEEE Ultrason. Symp.*, 2010, pp. 367–370.

[24] 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.

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

[26] N. Wagner, Y. C. Eldar, A. Feuer, and Z. Friedman, "Compressed beamforming applied to B-mode ultrasound imaging," in *Proc. 9th IEEE Int. Symp. Biomed. Imag.*, 2012, pp. 1080–1083.

[27] I. Tosic, I. Jovanovic, P. Frossard, M. Vetterli, and N. Duric, "Ultrasound tomography with learned dictionaries," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Mar. 2010, pp. 5502–5505.
[28] O. Lorintiu, H. Liebgott, O. Bernard, and D. Friboulet, "Compressive sensing ultrasound imaging using overcomplete dictionaries," in *Proc. IEEE Int. Ultrason. Symp.*, Jul. 2013, pp. 45–48.
[29] O. Lorintiu, H. Liebgott, M. Alessandrini, and O. Bernard, "Compressed sensing reconstruction of 3D ultrasound data using dictionary learning," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 1317–1321.
[30] G. E. Trahey, J. W. Allison, and O. T. Von Ramm, "Angle independent ultrasonic detection of blood flow," *IEEE Trans. Biomed. Eng.*, vol. BME-34, no. 12, pp. 965–967, Dec. 1987.
[31] L. N. Bohs and G. E. Trahey, "A novel method for angle independent ultrasonic imaging of blood flow and tissue motion," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 3, pp. 280–286, Mar. 1991.
[32] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[33] A. Basarab, H. Liebgott, O. Bernard, D. Friboulet, and D. Kouame, "Medical ultrasound image reconstruction using distributed compressive sampling," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 628–631.
[34] H. Liebgott, A. Basarab, D. Kouame, O. Bernard, and D. Friboulet, "Compressive sensing in medical ultrasound," in *Proc. IEEE Int. Ultrason. Symp.*, Oct. 2012, pp. 1–6.
[35] M. F. Schiffner and G. Schmitz, "Fast pulse-echo ultrasound imaging employing compressive sensing," in *Proc. IEEE Int. Ultrason. Symp.*, Oct. 2011, pp. 688–691.
[36] Y. F. Li and P. C. Li, "Ultrasound beamforming using compressed data," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 308–313, May 2012.
[37] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1417–1428, Jun. 2002.
[38] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 31–40, Mar. 2008.
[39] L. T. Chira, J.-M. Girault, T. Mateo, S. Ménigot, and C. Rusu, "Ultrasound medical image deconvoluion using CLEAN algorithm," in *Proc. Acoustic Conf.*, Nantes, France (Societe Francaise d'Acoustique), Apr. 2012.
[40] J. A. Jensen, "Linear description of ultrasound imaging systems," *Notes for the International Summer School on Advanced Ultrasound Imaging*, Tech. Univ. Denmark, Kgs. Lyngby, Denmark, Jul. 1999.
[41] T. L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*. Amsterdam, The Netherlands: Elsevier, 2004.
[42] F. Vignon and M. R. Burcher, "Capon beamforming in medical ultrasound imaging with focused beams," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 55, no. 3, pp. 619–628, Mar. 2008.
[43] H. Azhari, *Basics of Biomedical Ultrasound for Engineers*. Hoboken, NJ, USA: Wiley, 2010.
[44] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. 27th Annu. Asilomar Conf. Signals, Syst. Comput.*, 1993, vol. 1, pp. 40–44.
[45] J. A. Jensen, "FIELD: A program for simulating ultrasound systems," in *Proc. 10th Nordicbaltic Conf. Biomed. Imag.*, 1996, vol. 34, pp. 351–353.
[46] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

Yael Yankelevsky (S'16) received her B.Sc. and M.Sc. degrees from the Department of Electrical Engineering at the Technion-Israel Institute of Technology, in 2006 and 2013, respectively. She is currently pursuing her Ph.D. at the Department of Computer Science at the Technion. Her research interests include signal and image processing, computer vision in medical imaging, sparse representations and inverse problems.

Zvi Friedman received the B.Sc. degree in physics from the Technion-Israel Institute of Technology, Haifa, in 1966, the M.Sc. degree from the Weizmann Institute of Science, Rehovot, in 1970, and the Ph.D. degree in solid-state physics from the Tel Aviv University, Israel, in 1974. He joined the ultrasound division of Elscint, Haifa, as a Senior Scientist in 1991. Since 1998, this division became part of the cardiovascular activity in GE Healthcare. His main areas of interest include application of advanced signal processing methods to ultrasonic imaging in general and to beamforming in particular. He is also working on the development of advanced cardiac applications.

Arie Feuer (M'76–SM'93–F'04–LF'14) has been with the Department of Electrical Engineering at the Technion - Israel Institute of Technology, since 1983 where he is currently a Professor Emeritus. He received his B.Sc. and M.Sc. from the Technion in 1967 and 1973 respectively and his Ph.D. from Yale University in 1978. From 1967 to 1970 he was in industry working on automation design and between 1978 and 1983 with Bell Labs in Holmdel, N.J. Between the years 2013 and 2015 he worked for a startup company developing algorithms for medical ultrasound imaging. Between the years 1994 and 2002 he served as the president of the Israel Association of Automatic Control and was a member of the IFAC Council during the years 2002-2005. Arie Feuer is a Life Fellow of the IEEE. In the last 25 years he has been regularly visiting the EE&CS department at the University of Newcastle. In 2009 he received an honorary doctorate from the University of Newcastle. His current research interests include medical imaging, in particular ultrasound and CT, resolution enhancement of digital images and videos, 3D video and multi-view data, sampling and combined representations of signals and images, and adaptive systems in signal processing and control.