An Optimized Compounding Approach to Ultrasound Imaging

Yair KERNER and Moshe PORAT
Technion, Haifa 32000, Israel
E-mail: yairk@vision.technion.ac.il, mp@ee.technion.ac.il

***Abstract* – We develop an optimized approach to exploit pairs of ultrasound scans of the same image plane with the aim of enhancing the quality of ultrasound imaging. Each image pair is assumed to be co-registered with 90° separation between the two insonification directions. Using multi-channel image restoration, the proposed approach provides significant improvement to the quality of ultrasound imaging. Enhanced images are presented, and the advantages of this new approach over presently available methods are discussed.**

***Index Terms* – Blind Deconvolution, Image Restoration, Medical Imaging, Spatial Compounding, Ultrasound.**

I. INTRODUCTION

In medical ultrasound imaging, short pulses are emitted from an array of resonant elements with a center frequency typically in the range of 2-10 MHz. Such two-dimensional (or B-mode) scans undergo two types of blur before the echo returns to the transducer: In the axial (longitudinal) direction the blur results from the envelope of the acoustic wave and from properties of the tissue through which the wave propagates; In the lateral (transverse) direction the blur is affected by the width and apodization of the transmission and reception apertures as well as the distance of the imaged object from the focus and medium-related distortions. The model that is commonly used in ultrasound imaging research for the relation between the received signal and the tissue reflectivity is of a linear space-invariant system (LSI). Handling space variations of the Point Spread Function (PSF) is usually through partitioning the imaged plane to smaller regions with approximately invariant PSF.

For more than two decades various algorithms have been suggested to sharpen the images, either by deconvolution assuming the PSF is known [1], [2], or through blind deconvolution [3]-[10]. The prevalent approach in blind deconvolution is to use the LSI method of the Wiener filter to recover the reflectivity image, and therefore most of the effort is concentrated in estimating the PSF. It is worth noting that while a few algorithms used the video image as their input [1], [2], in most of the published work deconvolution is applied to the RF image [3]-[10].

In recent years several authors [11]-[14] developed algorithms for compounding of ultrasound images of the same region from different angles. In parallel, researchers who sought ways to overcome blur phenomena of photographic images developed algorithms for multi-channel image restoration. Ghiglia [15] presented a constrained least-squares algorithm for image restoration given several blurred images of the same object, each corresponding with a different PSF. Later, Katsaggelos et al. [16] presented a systematic framework for performing multi-channel image restoration in the frequency domain. In the field of medical imaging this method was applied to dual-radionuclide imaging [17]. Tom et al. [18] established a maximum-likelihood formulation for the general problem of multi-channel image restoration, and utilized the expectation-maximization algorithm to solve it.

To the best of our knowledge no algorithm was developed to exploit both deconvolution and compounding for ultrasound image applications. It should be also noted that the above-mentioned algorithms for multi-channel restoration were developed for photographic images. When considering these algorithms for the field of ultrasound images, one must note fundamental differences in several characteristics of the problem: the blur transfer function has band-pass behavior in one dimension rather than low-pass in both dimensions, the sampling intervals are different in each direction, and non-linear operations are involved in the display procedure.

The research presented in this paper aims at improving the quality of ultrasound images through exploiting pairs of ultrasound scans of the same plane. Each image pair is assumed to be co-registered with 90° separation between the two-insonification directions.

II. MODEL OF THE PROBLEM

A. Notation

A 2-D spatial location in the imaging plane is denoted as (x, z), where x is the coordinate in the lateral direction and z is the coordinate in the axial direction. In the frequency domain (ω_x, ω_z) represents the spatial frequency in radians. Whenever matrix representation is used the axial direction is column-wise unless stated otherwise. When matrix coordinates appear in parentheses the first parameter specifies the horizontal coordinate $\mathbf{A}(m, n) = \mathbf{A}_{nm}$ in order to keep consistency between matrix and 2-D signal or image formulations.

The matrix **R** indicates the unknown tissue reflectivity in the imaging plane. The matrix **S** denotes the observed RF image, which is the collection of sampled signals from all the transducer elements during a single scan.

B. Tissue reflectivity and degradation

The tissue reflectivity can be considered as resulting from an assembly of reflectors and scatterers [7], [10]. A reflector is an interface, large compared with the wavelength of the ultrasonic pulse, while scatterers are objects, small compared with the wavelength and typically inducing a speckle pattern. For that reason, and following [7], the tissue reflectivity is modeled in this work as the sum of a deterministic function **D** representing the specular reflections and a zero-mean Gaussian stochastic process **U** corresponding with the speckle component:

$$\mathbf{R}(x, z) = \mathbf{D}(x, z) + \mathbf{U}(x, z). \quad (1)$$

The interaction of the ultrasonic pressure field with the tissue is 3-dimensional, but its observation in B-mode imaging is a 2-D space. Subject to customary assumptions, the formation process of the RF image can generally be modeled as a 2-D spatial linear filtering operation with a spatially variant point-spread function [5], [6], [19], [20]. It is possible to segment the image into regions for which the PSF is approximately constant and simple convolution describes with good accuracy the image formation:

$$\mathbf{S}(x, z) = \mathbf{P}(x, z) * \mathbf{R}(x, z) + \mathbf{N}(x, z), \quad (2)$$

where $*$ denotes 2-D convolution, **P** represents the PSF and **N** represents additive noise. The noise term is modeled as a white zero-mean Gaussian random process that is independent of the tissue reflectivity and the PSF [5]-[7].

The system response in the frequency domain $\mathbf{P}(\omega_x, \omega_z)$ is related to the PSF through 2-D Fourier transform. The resolution in the lateral direction is significantly worse than the axial resolution; therefore the PSF has wide extent laterally and narrow support axially. Equivalently, the bandwidth of the system response is much smaller in the lateral dimension than in the axial dimension.

C. Rotation

Rotation is treated as positive when its direction is from the positive ray of the *x*-axis toward the positive ray of the *z*-axis, and as negative when it is in the opposite direction. This implies that if conventional matrix coordinates are employed then positive rotation is clockwise.

Let $\mathbf{A}_1$ and $\mathbf{A}_2$ denote two RF images that were recorded with $\mathbf{A}_2$ taken after the transducer was turned by –90° relative to its direction during the recording of $\mathbf{A}_1$. The superscript $^{\text{rot}}$ is used to indicate 90° rotation, while $^{\text{inv-rot}}$ indicates –90° rotation. Assuming that the tissue reflectivity **R** is the same for both recordings, then:

$$\mathbf{A}_1 = \mathbf{P}_1 * \mathbf{R} + \mathbf{N}_a \quad (3)$$

$$\mathbf{A}_2 = \mathbf{P}_2 * \mathbf{R}^{\text{inv-rot}} + \mathbf{N}_b,$$

where $\mathbf{P}_1$, $\mathbf{P}_2$ denote the system's PSF in each case and $\mathbf{N}_a$, $\mathbf{N}_b$ denote the additive noise. Rotating $\mathbf{A}_2$ in (3) by 90° and substituting: $\mathbf{S}_V \equiv \mathbf{A}_1$, $\mathbf{S}_H \equiv \mathbf{A}_2^{\text{rot}}$, $\mathbf{N}_1 \equiv \mathbf{N}_a$ and $\mathbf{N}_2 \equiv \mathbf{N}_b^{\text{rot}}$, yields:

$$\mathbf{S}_V = \mathbf{P}_1 * \mathbf{R} + \mathbf{N}_1 \quad (4)$$

$$\mathbf{S}_H = \mathbf{P}_2^{\text{rot}} * \mathbf{R} + \mathbf{N}_2.$$

In other words, the blur is treated as rotated by 90° instead of regarding the image as rotated by –90°. Note that for $\mathbf{S}_V$ the axial direction coalesces with the vertical direction, while for $\mathbf{S}_H$ it is horizontal. Therefore, according to the degradation model, **R** is more blurred horizontally to generate $\mathbf{S}_V$ and more blurred vertically to generate $\mathbf{S}_H$.

When dealing with matrices and discrete Fourier transform (DFT), rotation should be given special attention because straightforward 90° rotation results with phase shift in the frequency domain. This is alleviated if the definition of 90° rotation utilizes the periodicity of the DFT:

$$\mathbf{B}(m', n') \equiv \mathrm{T}\mathbf{A} = \mathbf{A}(n', [N - m']_{\text{mod } N}) \quad , \quad (5)$$

$$m' = 0, 1 \ldots N-1$$

where $\mathbf{A}(m, n)$ is a matrix with *N* rows. This is equivalent to column-wise inversion excluding the first row followed by transposition.

D. Sampling grid

We assume that the ultrasonic B-mode scan is performed with a linear array transducer. Hence, the sampling grid is rectangular, but the sampling intervals are different in each direction. The lateral sampling interval d_L is a consequent of the spacing between the piezoelectric crystals, which is in the region of few hundreds of μm. The axial sampling interval d_A is related to the sampling frequency f_S at the receiver and the speed of sound *v* within the tissue:

$$d_A = \tfrac{1}{2}\, v / f_S. \quad (6)$$

For example, if $v \cong 1{,}540$ m/sec and $f_S = 20$ MHz [2]-[4], then $d_A \cong 38.5$ μm, and assuming $d_L = 500$ μm [1], the ratio between the lateral and axial sampling intervals is about 13:1. If $f_S = 10.5$ MHz [8] and the crystal spacing is approximately 200 μm[12], this ratio drops off to around 3:1.

Taking into account the different sampling intervals and treating the matrices of the reflectivity **R** and the PSF **P** as samples on a square grid with intervals d_A in both directions, the discrete-space LSI model for the generation of the 2 source images $\mathbf{S}_V$ and $\mathbf{S}_H$ from the tissue reflectivity **R** is:

$$\mathbf{S}_V = (\mathbf{P}_1 * \mathbf{R}) \downarrow_H K + \mathbf{N}_1 \quad (7)$$

$$\mathbf{S}_H = (\mathbf{P}_2^{\text{rot}} * \mathbf{R}) \downarrow_V K + \mathbf{N}_2,$$

where $\downarrow_H K$ denotes horizontal decimation by factor K and $\downarrow_V K$ stands for vertical decimation.

E. Frequency domain

DFT of (7) with proper zero-padding gives:

$$\mathbf{S}_V(\omega_x, \omega_z) = \mathbf{P}_1(\omega_x, \omega_z)\, \mathbf{R}(\omega_x, \omega_z) + \mathbf{N}_1(\omega_x, \omega_z) \quad (8)$$

$$\mathbf{S}_H(\omega_x, \omega_z) = \mathbf{P}_2(\omega_z, -\omega_x)\, \mathbf{R}(\omega_x, \omega_z) + \mathbf{N}_2(\omega_x, \omega_z)$$

where $\mathbf{S}_V$, $\mathbf{S}_H$, $\mathbf{P}_1$, $\mathbf{P}_2$, **R**, $\mathbf{N}_1$ and $\mathbf{N}_2$ are respectively the 2-D DFT's of $\mathbf{S}_V$, $\mathbf{S}_H$, $\mathbf{P}_1$, $\mathbf{P}_2$, **R**, $\mathbf{N}_1$ and $\mathbf{N}_2$. The zero padding is performed in a way such that all the elements in (8) are square matrices of size $N_{DFT} \times N_{DFT}$.

Using the vector notation $\mathbf{s}_{k,l} \equiv [\mathbf{S}_V(\omega_x, \omega_z), \mathbf{S}_H(\omega_x, \omega_z)]^T$, $\mathbf{h}_{k,l} \equiv [\mathbf{P}_1(\omega_x, \omega_z), \mathbf{P}_2(\omega_z, -\omega_x)]^T$, $r_{k,l} \equiv \mathbf{R}(\omega_x, \omega_z)$, and

$\mathbf{n}_{k,l} \equiv [\mathbf{N}_1(\omega_x, \omega_z), \mathbf{N}_2(\omega_x, \omega_z)]^T$, with $\omega_x = 2\pi k/N_{DFT}$ and $\omega_z = 2\pi l/N_{DFT}$, (8) can be compactly written as:

$$\mathbf{s}_{k,l} = \mathbf{h}_{k,l}\, r_{k,l} + \mathbf{n}_{k,l} \ . \tag{9}$$

The variable $r_{k,l}$ is stochastic with mean $d_{k,l} \equiv \mathbf{D}(\omega_x, \omega_z)$ and variance λ_u, and is assumed to have Gaussian probability density function (PDF) [7]. The noise term $\mathbf{n}_{k,l}$ is a random vector with zero mean and a diagonal 2 × 2 covariance matrix $\mathbf{\Lambda}$, where the diagonal elements are the variances λ_1 and λ_2 of $\mathbf{N}_1(\omega_x, \omega_z)$ and $\mathbf{N}_2(\omega_x, \omega_z)$ respectively.

III. BLUR ESTIMATION

A. The algorithm

The blind deconvolution problem is solved in two steps:

1. Estimation of the blur function.
2. Image reconstruction assuming this estimated function is the true blur and image compounding.

The estimation of the blur function can be viewed as an optimization problem where we search for the unknown parameters of a PDF and for which the maximum-likelihood (ML) approach can be applied. Employing optimization of the direct likelihood function yields a difficult minimization problem since the unknown quantities of reflectivity and blur are coupled through multiplication. The expectation-maximization (EM) algorithm is an iterative technique that greatly simplifies the ML problem.

According to the EM method, the *complete data* **y** is not observed directly, but only by means of the *observed data* **s**, which is related to **y** through a non-invertible linear mapping. Applying to the problem at hand, we define: $\mathbf{y}_{k,l} = [\mathbf{s}_{k,l}^T, r_{k,l}]^T$, so the mapping is: $\mathbf{s}_{k,l} = [\mathbf{I}, \mathbf{0}]^T\, \mathbf{y}_{k,l}$. The unknown quantity $r_{k,l}$ within the complete data is referred to as the *hidden data*. The PDF of the complete data is $f_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the set of unknown parameters of the PDF: $\boldsymbol{\theta} = [\mathbf{h}^T, d, \lambda_u, \lambda_1, \lambda_2]^T$.

According to the complete-probability formula:

$$f_{\mathbf{y}}(\mathbf{y}_{k,l}|\boldsymbol{\theta}) = f_r(r_{k,l}|\boldsymbol{\theta})\, f_{\mathbf{s}}(\mathbf{s}_{k,l}|r_{k,l},\boldsymbol{\theta}). \tag{10}$$

Due to the assumption that the stochastic terms are white, $f_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})$ is the product of $f_{\mathbf{y}}(\mathbf{y}_{k,l}|\boldsymbol{\theta})$ over all possible combinations of k and l. Hence the likelihood of the complete data is:

$$L(\boldsymbol{\theta}) = \ln\{f_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})\} = \sum_k \sum_l L_{k,l}(\boldsymbol{\theta}) \tag{11}$$

$$L_{k,l}(\boldsymbol{\theta}) = -\tfrac{3}{2}\ln(2\pi) - \tfrac{1}{2}\ln(\lambda_u\lambda_1\lambda_2) - \tfrac{1}{2}\lambda_u^{-1}|r_{k,l} - d_{k,l}|^2 - \tfrac{1}{2}(\mathbf{s}_{k,l} - \mathbf{h}_{k,l}\, r_{k,l})^\dagger\, \mathbf{\Lambda}^{-1}\, (\mathbf{s}_{k,l} - \mathbf{h}_{k,l}\, r_{k,l}).$$

In the EM algorithm each iteration is composed of two steps: expectation (E step) and maximization (M step). In the E step the conditional expectation of $\ln\{f_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})\}$, using the current estimates of the parameters $\boldsymbol{\theta}^{[n]}$ and conditioned upon the observed data, is calculated:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = E[\ln\{ f_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta}) \}|\mathbf{s},\boldsymbol{\theta}^{[n]}]. \tag{12}$$

In the M step the expectation $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ is maximized with respect to $\boldsymbol{\theta}$ to provide a new estimation of the parameters:

$$\boldsymbol{\theta}^{[n+1]} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}). \tag{13}$$

B. Steps of the algorithm

After (11) is substituted for $f_{\mathbf{y}}(\mathbf{y}|\boldsymbol{\theta})$ in (12), it follows that maximization of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ is equivalent to minimization of:

$$J(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_k \sum_l J_{k,l}(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) \tag{14}$$

$$J_{k,l}(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \ln(\lambda_u\lambda_1\lambda_2) + \lambda_u^{-1}\, |\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}] - d_{k,l}\, |^2 + (\mathbf{s}_{k,l} - \mathbf{h}_{k,l}\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}])^\dagger\, \mathbf{\Lambda}^{-1} (\mathbf{s}_{k,l} - \mathbf{h}_{k,l}\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]) + Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]\, (\lambda_u^{-1} + \mathbf{h}_{k,l}^\dagger\, \mathbf{\Lambda}^{-1}\, \mathbf{h}_{k,l}).$$

In order to find the conditional expectation and variance of $r_{k,l}$ given $\mathbf{s}$ and $\boldsymbol{\theta}^{[n]}$, we need to look at the conditional probability density. Using the complete-probability formula $f_r(r_{k,l}\ |\ \mathbf{s},\boldsymbol{\theta}^{[n]}) = f_{\mathbf{y}}(\mathbf{y}_{k,l}\ |\ \boldsymbol{\theta}^{[n]}) \,/\, f_{\mathbf{s}}(\mathbf{s}_{k,l}\ |\ \boldsymbol{\theta}^{[n]})$, it follows:

$$E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}] = [\lambda_1^{-1}\, \mathbf{P}_1(\omega_x, \omega_z)^*\, \mathbf{S}_V(\omega_x, \omega_z) + \lambda_2^{-1}\, \mathbf{P}_2(\omega_z, -\omega_x)^*\, \mathbf{S}_H(\omega_x, \omega_z) + \lambda_u^{-1}\, d_{k,l}] \,/\, [\lambda_1^{-1}\, |\mathbf{P}_1(\omega_x, \omega_z)|^2 + \lambda_2^{-1}\, |\mathbf{P}_2(\omega_z, -\omega_x)|^2 + \lambda_u^{-1}]. \tag{15}$$

$$Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}] = 1/ [\lambda_1^{-1}\, |\mathbf{P}_1(\omega_x, \omega_z)|^2 + \lambda_2^{-1}\, |\mathbf{P}_2(\omega_z, -\omega_x)|^2 + \lambda_u^{-1}]. \tag{16}$$

Note the expression in (15) is similar to the vector Wiener filter [16]. For each of the parameters in (15) and (16) the current estimation is substituted, though the superscript $^{[n]}$ is suppressed for brevity.

Now, we take from (14) only the terms that involve $\mathbf{h}_{k,l}$ and use the identity $(\mathbf{v}^\dagger\, \mathbf{A}\, \mathbf{w}) = \mathrm{tr}(\mathbf{A}\, \mathbf{w}\, \mathbf{v}^\dagger)$, where $\mathbf{A}$ is a matrix of size N × N, $\mathbf{v}$, $\mathbf{w}$ are vectors of size N × 1 and tr{$\mathbf{A}$} denotes the trace of $\mathbf{A}$:

$$J_{\mathbf{h}}(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_k \sum_l \mathrm{tr}\{\mathbf{\Lambda}^{-1}\, (Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]\, \mathbf{h}_{k,l}\, \mathbf{h}_{k,l}^\dagger + (\mathbf{s}_{k,l} - \mathbf{h}_{k,l}\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}])\, (\mathbf{s}_{k,l} - \mathbf{h}_{k,l}\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}])^\dagger)\}. \tag{17}$$

Since $\mathbf{\Lambda}$ is a diagonal matrix, we obtain from the expression in (17):

$$\mathbf{h}_{k,l}^{[n+1]} = E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]^*\, \mathbf{s}_{k,l} \,/\, (|E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]|^2 + Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]). \tag{18}$$

Explicitly, the update of the estimation of the blur is:

$$\mathbf{P}_1(\omega_x, \omega_z)^{[n+1]} = E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]^*\, \mathbf{S}_V(\omega_x, \omega_z) \,/ \tag{19}$$

$$E[|r_{k,l}|^2|\mathbf{s},\boldsymbol{\theta}^{[n]}] \qquad / \left(E[|r_{k,l}|^2|\mathbf{s},\boldsymbol{\theta}^{[n]}] + E[|r_{-l,k}|^2|\mathbf{s},\boldsymbol{\theta}^{[n]}]\right)$$

$$\mathbf{P}_2(\omega_z, -\omega_x)^{[n+1]} = E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]^* \, \mathbf{S}_H(\omega_x, \omega_z) / E[|r_{k,l}|^2|\mathbf{s},\boldsymbol{\theta}^{[n]}]$$

where we used for $E[|r_{k,l}|^2|\mathbf{s},\boldsymbol{\theta}^{[n]}]$ the identity:

$$E[|r_{k,l}|^2|\mathbf{s},\boldsymbol{\theta}^{[n]}] = |\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]\,|^2 + Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]. \tag{20}$$

Next, we take from (14) only terms related to $\boldsymbol{\Lambda}$:

$$J_\lambda(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_k \sum_l \ln(\lambda_1\lambda_2) + J_{\mathbf{h}}(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) \tag{21}$$

Substituting (18) for $\mathbf{h}_{k,l}$, the values of λ_1 and λ_2 that minimize (21) are:

$$\lambda_1^{[n+1]} = N_{DFT}^{-2} \sum_k \sum_l \{Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]\, |\mathbf{P}(\omega_x, \omega_z)^{[n+1]}|^2 + |\mathbf{S}_V(\omega_x, \omega_z) - E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]\, \mathbf{P}(\omega_x, \omega_z)^{[n+1]}|^2\} \tag{22}$$

$$\lambda_2^{[n+1]} = N_{DFT}^{-2} \sum_k \sum_l \{Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]\, |\mathbf{P}(\omega_z,-\omega_x)^{[n+1]}|^2 + |\mathbf{S}_H(\omega_x, \omega_z) - E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]\, \mathbf{P}(\omega_z, -\omega_x)^{[n+1]}|^2\}.$$

Then, for the minimization $J(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ with respect to $d_{k,l}$, we take from (14) only the terms that depend on $d_{k,l}$:

$$J_d(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \lambda_u^{-1} \,|\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}] - d_{k,l}\,|^2 \tag{23}$$

It is simple to see that the value of $d_{k,l}$ that minimizes (23) is:

$$d_{k,l}^{[n+1]} = E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}] \tag{24}$$

Finally, we substitute (24) for $d_{k,l}$ in (14) then take only the terms that depend on λ_u:

$$J_u(\theta|\theta^{[n]}) = \sum_k \sum_l (\ln(\lambda_u) + \lambda_u^{-1}\, Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]). \tag{25}$$

The minimum of (25) is achieved when the value of λ_u is:

$$\lambda_u^{[n+1]} = N_{DFT}^{-2} \sum_k \sum_l Var[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]. \tag{26}$$

If we assume that the noise variance is the same in both RF images, then the average of the 2 expressions in (22) should be used to update the variance. If we assume identical blur in both scans $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}$, (19) should be replaced with:

$$\mathbf{P}(\omega_z, \omega_x)^{[n+1]} = (\, E[r_{k,l}|\mathbf{s},\boldsymbol{\theta}^{[n]}]^* \, \mathbf{S}_V(\omega_x, \omega_z) + E[r_{-l,k}|\mathbf{s},\boldsymbol{\theta}^{[n]}]^* \, \mathbf{S}_H(-\omega_z, \omega_x)\,) \tag{27}$$

C. Initialization and constraints

Being highly non-linear the likelihood function $L(\boldsymbol{\theta})$ has multiple maxima, and therefore the initial conditions $\boldsymbol{\theta}^{[0]}$ have a great effect on the ability of the EM algorithm to converge to a good estimation. We found that $\mathbf{P}_1^{[0]}$ and $\mathbf{P}_2^{[0]}$ should have the same value at all frequencies, that is they should be the DFT of an impulse at the origin. Also, a plausible initialization for $d^{[0]}$ is the average of $\mathbf{S}_V$ and $\mathbf{S}_H$. For the noise variance the initial estimate was higher than the true value, as it was found in [18] to produce better results. In the initial iterations the estimation of the variances is unreliable, thus the estimation much improves if λ_u in (15) and (16) is limited such that λ_1/λ_u and λ_2/λ_u are not too large or small relatively to $\max|\mathbf{P}_1(\omega_x, \omega_z)|^2$ and $\max|\mathbf{P}_2(\omega_x, \omega_z)|^2$ respectively.

IV. IMAGE COMPUNDING

According to the approach of multi-channel image restoration the restored image is computed from (15). However, in ultrasound imaging the resulting image would contain oscillations, as is the case with RF images due to the band-pass character of the ultrasonic blur. Consequently there would be required envelope detection.

When handling single RF images, where the oscillations are along just one axis, the envelope can be detected through demodulation, followed by absolute value calculation. But, the image of (15) has oscillations along both axes, thus demodulation or Hilbert filtered cannot be utilized. In addition, the frequency support regions of $\mathbf{P}_1(\omega_x, \omega_z)$ and $\mathbf{P}_2(\omega_z, -\omega_x)$ are far from overlapping. This inhibits the benefit of (15) over a single-channel Wiener filter, which is the ability to get weighted average in the intersection of the support regions while retaining the best of the single-channel restorations at any other frequency.

The conclusion is that the compound image should be generated through the following steps:

1. Calculation of the 2 separate Wiener filter solutions using (15) and substituting 0 for λ_1^{-1} or λ_2^{-1} respectively.
2. Envelope detection of each resulting image.
3. Computation of the average of the 2 envelope detected images [11], [12].

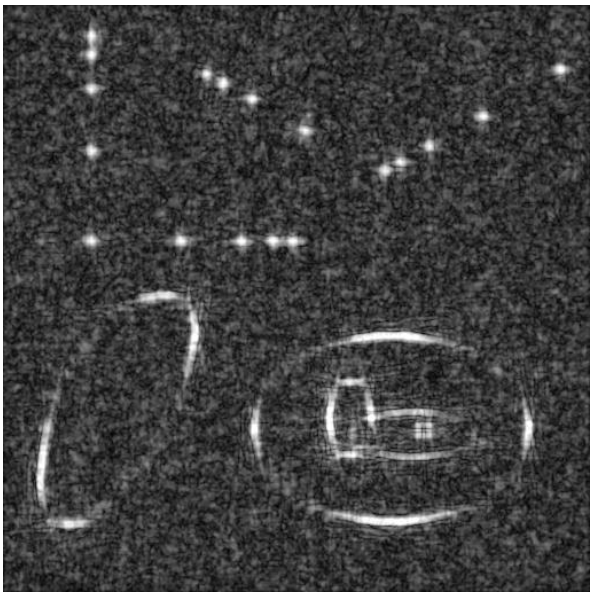
V. EXPERIMENTAL RESULTS

Since the estimation of the blur can be accurate up to a gain factor, it is required to restore the scaling before comparing the estimation with the true blur. We used the least squares method for this purpose:

$$J_g = \|g\hat{\mathbf{P}} - \mathbf{P}\|^2. \tag{28}$$

We used normalized mean squared error (NMSE) as a quality metric to evaluate the PSF estimation. The MSE is the right-hand side of (28) and the normalization is with respect to $\|\mathbf{P}\|^2$. When substituting for g the value that minimizes (28) the quality measure becomes:

$$\mathrm{NMSE}\{\mathbf{P}\} = 1 - \mathrm{Re}(\hat{\mathbf{P}}^\dagger \mathbf{P})^2 / (\|\hat{\mathbf{P}}\|^2 \, \|\mathbf{P}\|^2) \tag{29}$$



a)

b)

Figure 1. The output: compound images. a)– with the proposed algorithm; b)– with averaging the 2 envelope-detected images

We tested the algorithm with several artificial reflectivity maps. Their size was 512 × 512 pixels, which corresponds approximately with 20 × 20 mm for a sampling interval of about 40 μm. In the example that is presented in this paper there are 4 series of bright dots simulating wire targets that are separated by 0.5, 1, 2 and 3 mm.

The blur in the example in this paper was according to the non-separable model that was given in [20]. The parameters for this blur function corresponded with a linear-array transducer having width of 25 mm without apodization and focus depth of 50 mm, using 4 MHz ultrasonic pulses with envelope having full width at half maximum (FWHM) of about 0.3 μsec. Compounding with the suggested algorithm is compared in Figure 1 to compounding with only averaging the envelope-detected and log-compressed images.

VI. CONCLUSIONS

We developed an algorithm for blind identification of ultrasonic blur and for ultrasound image compounding. The approach of multi-channel image restoration was adopted for the identification task, but Wiener restoration was found to be inappropriate for compounding in the case of ultrasound imaging with 90° separated views. The algorithm in this work is based on registration of the two input images, and exploits both deconvolution and compounding to provide an enhanced output image. The simulation experiments that we conducted show that the quality of the compound image can be enhanced compared to the usage of algebraic average without Wiener filtering. Further research is planned so as to adapt the algorithm for angles different than 90° and for multi-angle compounding [21].

REFERENCES

[1] E.E. Hundt and E.A. Trautenberg, "Digital processing of ultrasonic data by deconvolution", *IEEE Trans. Sonics. Ultrason.*, vol. 27, no. 5, pp. 249-252, Sep. 1980.

[2] C.N. Liu, M. Fatemi and R.C. Waag, "Digital processing for improvement of ultrasonic abdominal images", *IEEE Trans. Med. Imaging.*, vol. 2, no. 2, pp. 66-75, June 1983.

[3] J.A. Jensen and S. Leeman, "Nonparametric estimation of ultrasound pulses", *IEEE Trans. Biomed. Eng.*, vol. 41, no. 10, pp. 929-936, Oct. 1994.

[4] J.A. Jensen, "Estimation of in vivo pulses in medical ultrasound", *Ultrasonic Imaging*, vol. 16, pp. 190-203, 1994.

[5] T. Taxt, "Restoration of medical ultrasound images using two-dimensional homomorphic deconvolution", *IEEE Trans. Ultrason., Ferroelect.*, vol. 42, no. 4, pp. 543-554, July 1995.

[6] U.R. Abeyratne, A.P. Petropulu and J.M. Reid, "Higher order spectra based deconvolution of ultrasound images", *IEEE Trans. Ultrason., Ferroelect.*, vol. 42, pp. 1064-1075, 1995.

[7] J.H. Hokland and P.A. Kelly, "Markov models of specular and diffuse scattering in restoration of medical ultrasound images", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 43, pp. 660-669, 1996.

[8] T. Taxt and G.V. Frolova, "Noise robust one-dimensional blind deconvolution of medical ultrasound images", *IEEE Trans. Ultrason., Ferroelect.*, vol. 46, no. 2, pp. 291-299, 1999.

[9] T. Taxt and J. Strand, "Two-dimensional noise-robust blind deconvolution of ultrasound images", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 48, no. 4, pp. 861-866, 2001.

[10] D. Adam and O. 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.

[11] P. He, K. Xue and Y. Wang, "Effects of spatial compounding upon image resolution", *Int. Conf. IEEE Eng. Medic. Bio. Society*, pp. 598-600, Oct. 1997.

[12] S. K. Jespersen, J. E. Wilhjelm and H. Sillesen, "Ultrasound spatial compound scanner for improved visualization in vascular imaging", 1998 *IEEE Ultrason. Symposium*, pp. 1623-1626, 1998.

[13] D. F. Leotta and R. W. Martin, "Three-dimensional spatial compounding of ultrasound scans with incidence angle weighting", *IEEE Ultrason. Symposium*, pp. 1605-1608, 1999.

[14] R. E. Shenhar and M. Porat, "A Dual Transducer Approach to Ultrasound Imaging and Spatial

Deformations", *IEEE International Conference on Image Processing*, 2000.

[15] D.C. Ghiglia, "Space-invariant deblurring Given *N* independently blurred images of a common object", *J. Opt. Soc. Amer. A*, vol. 1, no. 4, pp. 398-402, April 1984.

[16] A.K. Katsaggelos, K.T. Lay and N.P. Galatsanos, "A general framework for frequency domain multi-channel signal processing", *IEEE Trans. Image Proc.*, vol. 2, no. 3, pp. 417-420, 1993.

[17] J.M. Links, J.L. Prince and S.N. Gupta, "A vector Wiener filter for dual-radionuclide imaging", *IEEE Trans. Med. Imaging*, vol. 15 , pp. 700-709, Oct. 1996.

[18] B.C. Tom, K.T. Lay and A.K. Katsaggelos, "Multichannel image identification and restoration using the expectation-maximization algorithm", *Opt. Eng.*, vol. 35, no. 1, pp. 241-254, Jan. 1996.

[19] D. Iraca, L. Landini and L. Verrazzani, "Power spectrum equalization for ultrasonic image restoration", *IEEE Trans. Ultrason., Ferroelect.*, vol. 36, no. 2, pp. 216-222, March 1989.

[20] W.F. Walker and G.E. Trahey, "The application of K-space in pulse echo ultrasound", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 45, no. 3, pp. 541-558, May 1998.

[21] H. Kirshner and M. Porat, "On the Role of Exponential Splines in Image Interpolation", IEEE Trans. on Image Processing, vol. 18, pp. 2198-2208, 2009.