MODELING AND EQUALIZATION IN RADAR IMAGING FOR 3-D RECONSTRUCTION

Ricardo Merched

Dept. of Electrical and Computer Engineering
Universidade Federal do Rio de Janeiro, Brazil
merched@lps.ufrj.br

ABSTRACT

We revisit the fundamental problem of 1-D pulse compression in a multistatic MIMO focused scenario, from a communications viewpoint, and extend it to a more general 4-D pulse compression. The process of focusing and scanning over a 3-D object is interpreted as a MIMO 4-dimensional convolution between a reflectivity tensor and a space-varying system. This implies that several well established equalization methods employed in communications can be easily extended to a 3-D scenario with the purpose of achieving exact reconstruction of a given reflectivity volume. Assuming that no multiple scattering occurs, resolution is only limited in range by the sampling device in the unfocused case, while unlimited in case of focusing at multiple depths. Reconstruction under zero-forcing or least-squares criterion depends solely on the amount of diversity induced by sampling in both space and time, which further allows for a tradeoff between range and cross-range resolution.

1. INTRODUCTION

In radar imaging, resolution is generally dictated by its corresponding system point spread function, the response to a point source as a result of an external excitation (see [1] and the references therein). This notion of resolution turns out to be questionable, as the interpretation of echoes received from a range of continuous targets according to a linear model [2],[3] allows one to cast the imaging problem as a communication system that maps the target reflectivity function onto measurements, which in turn suggests that by virtue of sampling and equalization, one can achieve unlimited spatial resolution.

An interesting fact with all current radar-based approaches is that resolution in range seems to be treated quite differently from resolution in cross-range. That is, while the former requires "pulse compression", the latter is tackled via "beam compression", which is achieved by increasing the carrier frequency and/or antenna dimensions. Note that because in general we have a collection of signals transmitted and received from several directions, the pulse is in reality a 3-D function whose shape is given by the combined effect of the transmitted pulse and the MIMO medium's Green function. This motivates us to think of compression as one that is performed in all three dimensions in space, in a way that the term "pulse compression" is more properly applied to achieve volumetric resolution, as opposed to only range resolution. Moreover, methods for recovering resolution in radar imaging systems have always assumed a continuous model for the reflectivity information. This is perhaps the major reason why the solution to the inverse problem could only be approximated, normally done so via matched filtering. The fact is, after sampling, reflectivity can be seen as a discrete information whose granularity is dictated by the required target resolution, *for both range and cross-range* imaging.

In this paper, we recast the imaging problem into one that maps a reflectivity volume into measurements collected by illuminating a target object with a focused beam. As a result, the common wisdom related to minimum range resolution for target detection can be considered irrelevant, once the pulse transmission problem is restated as a 3-D communication problem through the medium's Green function $g(x, z, t)$.

2. IMAGING IN LIGHT OF COMMUNICATIONS

In the MIMO case the transmitting and receiving beams can be generally unfocused, each antenna element transmitting a different waveform, where a particular patch of the antenna transmits omnidirectionally from each element illuminating the target. For a given point z in the volume, the received signal at point y on the antenna, given the transmitted pulse $p(t, x)$ at point x only, can be compactly written as

$$\begin{aligned} y(t,x,z) &= J_s(x)J_r(y)[g(x,z,t) \star \partial_t^2[g(z,x,t) \star p(t)]s(z) \\ &= J_s(x)J_r(y)[g(y,z,t) \star g(z,x_m,t) \star p''(t)]s(z) \end{aligned} \quad (1)$$

where $s(z)$ is the reflectivity function, $p''(t)$ denotes the second derivative with respect to t, and $J_s(x)$ and $J_r(y)$ are time derivatives of the current density at the transmitting and receiving points on the antenna.

Our goal is to reconstruct the entire object volume, or equivalently, to recover all the cross sections that form a 3-D image. Here is where the geometry of the problem comes into picture. First, it is assumed that the aspect angle of all antennas elements are such that they see the same radar cross section (RCS). Second, we would like to preserve the geometry of the volume we want to reconstruct, in the sense that each cross section is "unrolled" to form a 3-D tensor representing

Thanks to CNPq, the Fulbright Comission, and CAPES for funding.

the object image. That is, we shall map each point of $s(z)$, onto the corresponding point of the unrolled object, defined as continuous tensor $S(\boldsymbol{r})$, where $\boldsymbol{r} = [r_1 \;\; r_2 \;\; r]^\intercal$. Its discrete-time version is denoted by the $R_1 \times R_2 \times R_3$ tensor $\boldsymbol{S}$, which is obtained by sampling $S(\boldsymbol{r})$ at $[n_0 T_0 \;\; n_1 T_1 \;\; kT_r]^\intercal$. The quantities $\{R_1, R_2\}$ define the target resolution in cross range, while R_3 is the desired resolution in range. We shall denote by $\boldsymbol{S}(k)$, $k = 0, 1, \ldots, R_3 - 1$ the $R_1 \times R_2$ matrix containing a rectangular lattice of points within a cross section of the object at range k, and by $\boldsymbol{S}(\boldsymbol{n})$ a vector of reflectivities within the tensor $\boldsymbol{S}$, $\boldsymbol{n} = \boldsymbol{n}_0, \boldsymbol{n}_1, \ldots, \boldsymbol{n}_{R_1 R_2}$. Moreover, we denote by s_k the reflectivity at point $\boldsymbol{r}_k$ for $k = 0, 1, \ldots, \tilde{R} - 1$, $\tilde{R} = R_1 R_2 R_3$. We further define $R = R_1 R_2$. Let us spatially sample the antenna surface, so that the transmitting patch contains $Q_t = Q_1 Q_2$ antenna elements, which transmit Q_t waveforms, i.e., $p_\ell(t)$, $\ell = 0, 1, \ldots, Q_t - 1$. These in turn are received at Q_r receiving elements, at $y_i(t), i = 0, 1, \ldots, Q_r - 1$. The complete model comprises forward and backward MIMO channels denoted by $G_f(t)$ and $G_b(t)$, respectively. Let $u_k(t)$, $k = 0, 1, \ldots, \tilde{R}$ be the k-th output of $G_f(t)$, at the corresponding point $\boldsymbol{r}_k$,

$$u_k(t) = \sum_{\ell=0}^{Q_t - 1} p''_\ell(t) \star g_k(x_\ell, t) \tag{2}$$

and note that we can instead express the problem as a mapping from reflectivities to measurements. We thus see that our main problem is how to recover the tensor $\boldsymbol{S}$, given the measurements $y_i(t)$, at the desired resolution $\{R_1, R_2, R_3\}$. Here, we analyze the case of a focused (transmitted) beam.

2.1. Focused Imaging

Focusing is achieved by integration of delayed copies of the transmitted signal, weighted by the time derivative of the current density over the antenna, giving rise to a beam pattern. Assuming that the antenna elements are closely located, and $p(t)$ is narrowband, the transmitter in this case becomes a bidimensional *phased array*, modeled as a SIMO system represented by a steering vector. The receiving beam, however, can be generally unfocused, so that the signals are collected at all antenna elements and processed by a MIMO system.

Let the m-th entry of the undelying SIMO system be given $e^{j\omega_c t_m}$, where t_m is the delay applied to the signal point m relative to the signal at the center of the antenna $\boldsymbol{x}_o$, with respect to the focal point, which we denote by $\boldsymbol{r}_f$. We shall omit the dependence on $\boldsymbol{r}_f$ for simplicity of notation. Let $u_k(t)$ be the signal at $\boldsymbol{r}_k$, defined within the depth of field corresponding to $\boldsymbol{r}_f$. Equation (2) in this case is given by

$$u_k(t) = -\omega_o^2 \sum_{\ell=0}^{Q_t - 1} J(\ell) p(t) e^{j\omega_c t_\ell} \star g_k(\boldsymbol{x}_\ell, t), \tag{3}$$

where the term $J(\ell)$ consists of a 2-D window that provides *apodization*, so as to shape the transmitted beam pattern. Let $\boldsymbol{a}_\ell = \boldsymbol{x}_\ell - \boldsymbol{x}_o$ be a vector from the center of the antenna to an arbitrary point on the antenna, and define $\kappa = \omega_o / c_o$ the wave number. Thus, expressing $g_k(\boldsymbol{x}_\ell, t)$ explicitly as $g_k(\boldsymbol{x}_\ell, t) = \bar{g}_k(\boldsymbol{x}_\ell, t) \star \frac{\delta(t - |\boldsymbol{r}_k - \boldsymbol{x}_\ell|/c)}{4\pi |\boldsymbol{r} - \boldsymbol{x}_\ell|}$, and using the *Fraunhofer* condition at the focal plane, we have

$$\begin{aligned} u_k(t) \;=\; & -\omega_o^2 \frac{p(t - |\boldsymbol{r}_k - \boldsymbol{x}_o|/c)}{4\pi |\boldsymbol{r}_k - \boldsymbol{x}_o|} \sum_{\ell=0}^{Q_t - 1} J(\ell) e^{j(\omega_c t_\ell + \kappa \boldsymbol{a}_\ell^\intercal \widehat{\boldsymbol{r}_k - \boldsymbol{x}_\ell})} \\ & \star \;\; \bar{g}_k(\boldsymbol{x}_\ell, t) \\ = & \underbrace{\frac{-\omega_o^2}{4\pi |\boldsymbol{r}_k - \boldsymbol{x}_o|} \int w(\widehat{\boldsymbol{r}_k - \boldsymbol{x}}_o, t - \tau) p(\tau - |\boldsymbol{r}_k - \boldsymbol{x}_o|/c) d\tau}_{\beta_k(\boldsymbol{r}_f, t)} \end{aligned} \tag{4}$$

where $w(\widehat{\boldsymbol{r}_k - \boldsymbol{x}}_o, t)$ is the 2-D discrete time Fourier transform of the time varying tapering function $J(\ell) e^{j\omega_c t_\ell} \bar{g}_k(\boldsymbol{x}_\ell, t)$ at $\boldsymbol{r}_k$ at time t, and $\beta_k(\boldsymbol{r}_f, t)$ beam pattern at time t, with f denoting the focal point. The complete system is illustrated by the spatial multirate structure of Fig. 1(a).

The interpretation of this scheme is as follows. Assume we have the antenna beam focused at $\boldsymbol{r}_f$. Because the beam is narrow at the focal plane, and due to a finite extent depth of field, the illuminated portion of the object can be represented by a $K_1 \times K_2 \times K_3$ parallelepiped, which contains the corresponding significant reflectivity samples. Their positions relative to the point $\boldsymbol{r}_f$ are set by the spatial shifts $z(-\boldsymbol{m}_i) \triangleq z_0^{m_i,0} z_1^{m_i,1} z_2^{m_i,2}$, $i = 0, \ldots, K - 1$, where $K \triangleq K_1 K_2 K_3$. We further define $\tilde{K} \triangleq K_1 K_2$. The columns of the downsampling matrix $\boldsymbol{B}$ define the geometry in which these spatial blocks are processed, as the object is scanned at a given rate. That is, let $\bar{\boldsymbol{s}}_{\boldsymbol{r}_a}$ and $\bar{\boldsymbol{s}}_{\boldsymbol{r}_b}$ be two parallelepipeds (tensors), possibly intersecting, at focal points $\boldsymbol{r}_a$ and $\boldsymbol{r}_b$ respectively. For instance, choosing $K = |\det \boldsymbol{B}|$ (i.e., the volume of the fundamental parallelepiped defined by $\boldsymbol{B}$), implies that nonoverlapping blocks are being processed each time. This is actually the lowest spatial rate for which the object can be processed without loss of useful information. On the other hand, when choosing $\boldsymbol{B} = \boldsymbol{I}$ the object is being scanned at the highest possible rate, which is given by the desired spatial resolution. For simplicity, we shall assume rectangular parallelepipeds, so that the beam region of support coincides with a rectangular parallelepiped as well. Each tensor block is then (continuous-time) convolved with the $K_1 \times K_2 \times Q_r$ tensor $\boldsymbol{\mathcal{F}}_{\boldsymbol{r}_f}(t)$, the resulting impulse response of the cascade of the $\beta_k(\boldsymbol{r}_f, t) p(t)$ with $\boldsymbol{G}_b(t)$, for $f = \{a, b\}$.

More specifically, assume $J_r(x_i) = 1$, and let $\boldsymbol{\mathcal{F}}_{\boldsymbol{r}_j, q}(t)$, $q = 0, 1, \ldots, Q_r - 1$, be the $K_1 \times K_2$ matrices that constitute $\boldsymbol{\mathcal{F}}_{\boldsymbol{r}_f}(t)$ at the focal point $\boldsymbol{r}_j$, $j = 0, 1, \ldots, \mathcal{D} - 1$, where $\mathcal{D}$ is number of foci. The $Q_r \times 1$ received vector is defined as $\boldsymbol{y}_{\boldsymbol{r}_j}(t) = [y_0(t) \;\; y_1(t) \;\; \cdots \;\; y_{Q_r - 1}]^\intercal$ and expressed as

$$\boldsymbol{y}_{\boldsymbol{r}_j}(t) = \boldsymbol{F}_{\boldsymbol{r}_f}(t) \star \boldsymbol{s}_{\boldsymbol{r}_f}(\ell) \;+\; \boldsymbol{v}_{\boldsymbol{r}_f}(t), \quad j = 0, 1, \ldots, \mathcal{D} - 1, \tag{5}$$

where

$$\boldsymbol{F}_{\boldsymbol{r}_j}(t) = \begin{bmatrix} vec^\intercal(\boldsymbol{\mathcal{F}}_{\boldsymbol{r}_j, 0}(t)) \\ \vdots \\ vec^\intercal(\boldsymbol{\mathcal{F}}_{\boldsymbol{r}_j, Q_r - 1}(t)) \end{bmatrix} \quad (Q_r \times \tilde{K}) \tag{6}$$

and $\boldsymbol{s}_{\boldsymbol{r}_j}(\ell) = vec(\bar{\boldsymbol{s}}_{\boldsymbol{r}_j}(\ell))$, $\ell = 0, 1, \ldots, K_3 - 1$. The quantity $\boldsymbol{v}_{\boldsymbol{r}_j}(t)$ is the corresponding vectorized noise samples.

Here is the crucial point in understanding the imaging process in general. First, it is important to note the difference between the range resolution defined by the spatial sampling T_r, and the time separation between two consecutive received pulses (here, more generally we speak of MIMO pulses), which leads to the sampling rate $T_s = 2\Delta_{\min}/c$, where $\Delta_{\min}$ is the minimum target separation required for a certain application, in the colocated radar. This coincides with the spatial sampling T_r in the far field scenario, since in this case we are allowed only one focal point in range. Therefore, $T_r = \Delta_{\min}$, that is, we are limited by the achievable time sampling. Focusing on the other hand allows sampling at the desired spatial resolution, so that $\Delta_{\min}$ is only defined within the incoming tensors of reflectivities. That is, sampling the received vectors at the rate nT_s yields the discrete model

$$\boldsymbol{y}_{\boldsymbol{r}_j}(n) = \sum_{\ell=0}^{B-1} \boldsymbol{F}_{\boldsymbol{r}_j}(n-\ell)\boldsymbol{s}_{\boldsymbol{r}_j}(V\ell) + \boldsymbol{v}_{\boldsymbol{r}_j}(n), \qquad (7)$$

for $j = 0, 1, \ldots, \mathcal{D}-1$, where $\boldsymbol{F}_{\boldsymbol{r}_j}(0) \ldots \boldsymbol{F}_{\boldsymbol{r}_j}(N-1)$ is the sampled version of $\boldsymbol{F}_{\boldsymbol{r}_j}(t)$ at T_s, while $\boldsymbol{s}_{\boldsymbol{r}_j}(V\ell)$ corresponds to the downsampled version of $\boldsymbol{s}_{\boldsymbol{r}_j}(\ell)$, with $V = \Delta_{\min}/T_r$. Figure 1(b) illustrates the difference between T_r and $\Delta_{\min}$.

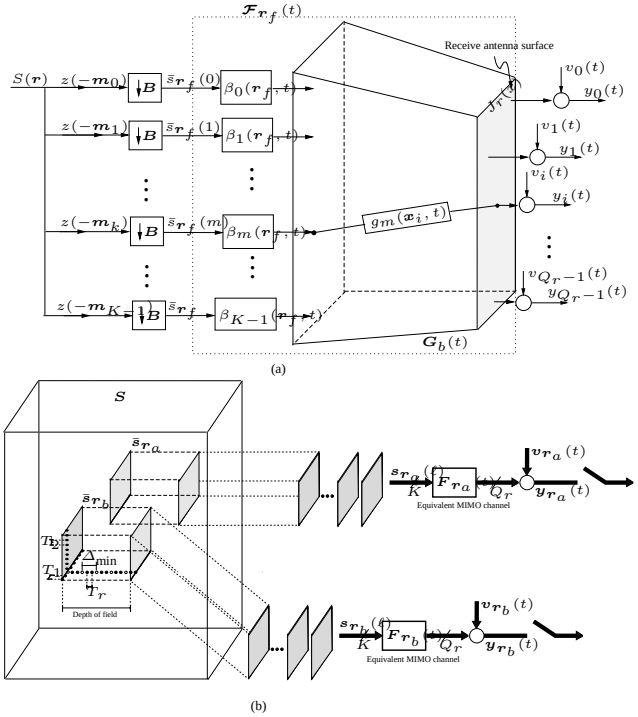

Fig. 1. *Equivalent mappings (a) and (b) from reflectivity to measurements.*

3. MODELING FOR REFLECTIVITY ESTIMATION

Suppose we are to estimate a tensor arriving from one particular focal point. Essentially, the tensor imaging can be cast like any other MIMO estimation problem, formulated linearly as either a Bayesian or classical estimation (Minimum Variance Unbiased Estimator-MVUE). Now, assume the output signal is sampled at T_s. In formulating the convolution model for the received echoes, two common structures can be induced, depending on the receiver design:

3.1. Block processing

Let $\mathcal{L}$ be the received signal window length, and define the vectorized tensor of of reflectivities and its corresponding received vector around $\boldsymbol{r}_j$, respectively:

$$\boldsymbol{x}_{\boldsymbol{r}_j}(k) = vec([\, \boldsymbol{s}_{\boldsymbol{r}_j}(V(k+B'-1)) \;\; \cdots \;\; \boldsymbol{s}_{\boldsymbol{r}_j}(V(k+1)) \;\; \boldsymbol{s}_{\boldsymbol{r}_j}(Vk)\,]) \qquad (8)$$

$$\bar{\boldsymbol{y}}_{\boldsymbol{r}_j} = vec([\, \boldsymbol{y}_{\boldsymbol{r}_j}(\mathcal{L}-1) \;\; \cdots \;\; \boldsymbol{y}_{\boldsymbol{r}_j}(1) \;\; \boldsymbol{y}_{\boldsymbol{r}_j}(0)\,]) \qquad (9)$$

where $B' = N + \mathcal{L} - 1$, for $k = 0, 1, \ldots, B-1$. This choice of $\mathcal{L}$ induces a block-by-block transmission with interblock interference (IBI), where each block has size $\mathcal{L}$. That is, let $\mathcal{D}$ be the number of foci set to cover the object space. Then,

$$\bar{\boldsymbol{y}}_{\boldsymbol{r}_j}(k) = \boldsymbol{H}_{\boldsymbol{r}_j}\boldsymbol{x}_{\boldsymbol{r}_j}(k) + \bar{\boldsymbol{v}}_{\boldsymbol{r}_j}(k), \quad j = 0, 1, \ldots, \mathcal{D} \qquad (10)$$

where

$$\boldsymbol{H}_{\boldsymbol{r}_j} = \begin{bmatrix} \boldsymbol{F}_j(0) & \cdots & \boldsymbol{F}_j(N-1) & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \boldsymbol{F}_j(0) & \cdots & \boldsymbol{F}_j(N-1) \end{bmatrix} \qquad (11)$$

is $Q_r\mathcal{L} \times \tilde{K}(N+\mathcal{L}-1)$, and $\bar{\boldsymbol{v}}_{\boldsymbol{r}_j}(k)$ is the vectorized noise samples. This can be the case in a monostatic scenario, where at the time the radar starts listening, part of the signal has already returned to the radar site, illustrated in Fig. 2.

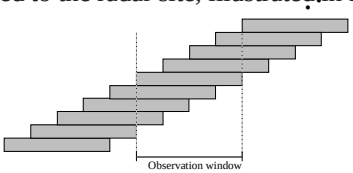

Fig. 2. *Overlapped pulses – short window*

Assume that the observation window is sufficiently long so that all incoming MIMO pulses can be completely observed. In this case, transmission becomes memoryless, meaning that at the time the radar starts listening, no information of near targets is present. Defining the input and output tensors as $\boldsymbol{x}_{\boldsymbol{r}_j} = vec([\, \boldsymbol{s}_{\boldsymbol{r}_j}(V(B-1)) \;\; \cdots \;\; \boldsymbol{s}_{\boldsymbol{r}_j}(V) \;\; \boldsymbol{s}_{\boldsymbol{r}_j}(0)\,])$ and $\bar{\boldsymbol{y}}_{\boldsymbol{r}_j} = vec([\boldsymbol{y}_{\boldsymbol{r}_j}(N+B-2) \;\; \cdots \;\; \boldsymbol{y}_{\boldsymbol{r}_j}(1) \;\; \boldsymbol{y}_{\boldsymbol{r}_j}(0)\,])$, where $\boldsymbol{H}_{\boldsymbol{r}_j}$ has dimensions $(N+B-1)Q_r \times B\tilde{K}$, we have

$$\bar{\boldsymbol{y}}_{\boldsymbol{r}_j}(k) = \boldsymbol{H}_{\boldsymbol{r}_j}\boldsymbol{x}_{\boldsymbol{r}_j}(k) + \bar{\boldsymbol{v}}_{\boldsymbol{r}_j}(k), \quad j = 0, 1, \ldots, \mathcal{D} \qquad (12)$$

When $\boldsymbol{H}_{\boldsymbol{r}_j}$ is a tall full rank matrix, a solution is found rather independently for every focal point. Two important solutions can be envisioned: (*i*) *Matrix Inversion*. In this case, the solution is given by $\hat{\boldsymbol{x}}_{\boldsymbol{r}_j}(k) = (\boldsymbol{I}_\delta^T\boldsymbol{H}_{\boldsymbol{r}_j})^{-1}\boldsymbol{I}_\delta^T\bar{\boldsymbol{y}}_{\boldsymbol{r}_j}(k)$ where

$$\boldsymbol{I}_\delta^T = [\, \boldsymbol{0}_{B\tilde{K}\times\delta} \;\; \boldsymbol{I}_{B\tilde{K}\times B\tilde{K}} \;\; \boldsymbol{0}_{B\tilde{K}\times NQ_r-\delta}\,] \qquad (13)$$

The optimal delay $\delta_{opt} \in [0, NQ_r]$ is chosen such that it minimizes the output noise power, by minimizing the matrix norm $\|(\boldsymbol{I}_\delta^T \boldsymbol{H}_{\boldsymbol{r}_j})^{-1}\|$ (see [5], for the case of $\boldsymbol{H}_{\boldsymbol{r}_j}$ Toeplitz); (ii) *Minimum variance (least-squares)*. This is given by:

$$\hat{\boldsymbol{x}}_{\boldsymbol{r}_j}(k) = (\boldsymbol{H}^*_{\boldsymbol{r}_j}\boldsymbol{H}_{\boldsymbol{r}_j})^{-1}\boldsymbol{H}^*_{\boldsymbol{r}_j}\bar{\boldsymbol{y}}_{\boldsymbol{r}_j}(k)\ , \quad j = 0, 1, \ldots, \mathcal{D}. \quad (14)$$

Observe that either in the unfocused scenario, or in case each tensor block $\boldsymbol{s}_{\boldsymbol{r}_j}$ is to be recovered independently from other blocks, its 3-D resolution can only rely on the feasibility of the sampling rate device. That is, while lateral resolution in these cases depends on the spatial grid density, resolution in range will be limited to $\Delta_{\min}$. The point is thus how accurate one can estimate $\boldsymbol{s}_{\boldsymbol{r}_j}$ by using these techniques, specially when $\boldsymbol{H}_{\boldsymbol{r}_j}$ is not a tall full rank matrix[1]. As we have mentioned, the beam foci and their positions relative to the antenna will define the convolution model employed. For this reason, we shall consider the beam motion in 2 separate steps, first with respect to lateral motion, covering all azimuths and elevations, which is then followed by depth motion.

a) *Lateral Motion*

Consider the signals in (7), and assume that the object is scanned at a fixed depth $r = d_k$ across all possible azimuths and elevations such that the object is still within the beam. This corresponds to picking $\boldsymbol{r}_j = \boldsymbol{r}_{l,k} = [\ r_1 \quad r_2 \quad d_k\ ]$, where $\{r_1, r_2\}$ are varied in order to cover $\mathcal{D}_1 = (R_1+K_1-1)(R_2+K_2-1)$ focal points, for $l = 0, 1, \ldots, \mathcal{D}_1 - 1$. Thus, we recognize the sequence $\{\boldsymbol{y}'_{\boldsymbol{r}_j}(n) = \boldsymbol{F}_{\boldsymbol{r}_j}(n-\ell)\boldsymbol{s}_{\boldsymbol{r}_j}(V\ell)\}$ as the output of a MIMO 2-D space varying convolution between the constituting matrices $\boldsymbol{\mathcal{F}}_{\boldsymbol{r}_j,q}(n)$ defined in (6), and the tensor slice at range $V\ell$ corresponding to depth d_k, which we denote by $\boldsymbol{S}_{d_k}(V\ell)$. This convolution at depth d_k results in Q_r images, or equivalently, a matrix with vector entries, each one having Q_r coefficients, which we denote by $\boldsymbol{\mathcal{Y}}'_{d_k,n}$. Note that similarly to the definition of a time window, which sets the number of received images from a focal point, one can also set a spatial window length, with dimension smaller than $\mathcal{D}_1$. However, in order to exploit full space and time diversity, we shall continue to assume a full spatial convolution model as well. We thus have

$$\begin{aligned} \boldsymbol{y}'_{d_k}(n) &= vec(\boldsymbol{\mathcal{Y}}'_{d_k,n}) && (15)\\ &= vec([\ \boldsymbol{y}'_{\boldsymbol{r}_{0,k}}(n)\ \ \boldsymbol{y}'_{\boldsymbol{r}_{1,k}}(n)\ \ \cdots\ \ \boldsymbol{y}'_{\boldsymbol{r}_{\mathcal{D}_1-1,k}}(n)\ ]), && (16)\end{aligned}$$

so we can replace (7) by a more compact form as

$$\tilde{\boldsymbol{y}}_{d_k}(n) = \sum_{\ell=0}^{B-1} \boldsymbol{\mathcal{H}}_{d_k}(n-\ell)vec(\boldsymbol{S}_{d_k}(V\ell)) + \boldsymbol{V}_{d_k}(n), \quad (17)$$

where $\boldsymbol{\mathcal{H}}_{d_k}(n-\ell)$ is a two-level block banded matrix of size $Q_r(R_1+K_1-1)(R_2+K_2-1) \times R_1R_2$, representing a MIMO 2-D convolution. Defining

$$\begin{aligned} \boldsymbol{\mathcal{S}}_{d_k} &= [\ vec(\boldsymbol{S}_{d_k}(V(B-1)))\ \ \cdots\ \ vec(\boldsymbol{S}_{d_k}(V))\ \ vec(\boldsymbol{S}_{d_k}(0))\ ] \\ \boldsymbol{Y}_{d_k} &= [\ \tilde{\boldsymbol{y}}_{d_k}(N+B-2)\ \ \cdots\ \ \tilde{\boldsymbol{y}}_{d_k}(1)\ \ \tilde{\boldsymbol{y}}_{d_k}(0)\ ] \quad (18)\end{aligned}$$

and denoting $\boldsymbol{\mathcal{X}}_{d_k} = vec(\boldsymbol{\mathcal{S}}_{d_k})$ and $\bar{\boldsymbol{y}}'_{d_k} = vec(\boldsymbol{Y}_{d_k})$, we have

$$\bar{\boldsymbol{y}}'_{d_k} = \boldsymbol{\mathcal{H}}_{d_k}\boldsymbol{\mathcal{X}}_{d_k} + \boldsymbol{v}_{d_k} \quad (19)$$

where $\boldsymbol{\mathcal{H}}_{d_k}$ is a $Q_r(N+B-1)(R_1+K_1-1)(R_2+K_2-1) \times R_1R_2B$ block toeplitz matrix, with block banded blocks. The three-dimensional model (19) fully describes the far-field transmission scenario for a single depth at $d_0 = \infty$; we see that while resolution is unlimited in cross-range, we are still limited to estimate B out of K_3 images within the tensor $\boldsymbol{S}$.

b) *Depth Motion*

The lateral scanning can be generally performed at several depths, giving rise to a full 4-D convolution model. In other words, (17) can be defined for $k = 0, 1, \ldots, \mathcal{D}_2$, where $\mathcal{D}_2 = (R_3+K_3-1)$, so that overall $\boldsymbol{r}_j$ covers $\mathcal{D} = (R_1+K_1-1)(R_2+K_2-1)(R_3+K_3-1)$ focal points. There is, however, a subtle difference in the way convolution is performed in the axial direction. While the 2-D convolution sum can be computed at the minimum resolution step $\{T_1, T_2\}$, the volumetric convolution is obtained by combining images at ranges $V\ell$, $\ell = 0, 1, \ldots, B-1$, and at times $0, 1, \ldots, N+B-1$, for each depth d_k. Let $\boldsymbol{\mathcal{C}}_{d_k,c}$, with $c = 0, 1, \ldots, B-1$ be the c-th block column of $\boldsymbol{\mathcal{H}}_{d_k}$ and define

$$\begin{aligned} \boldsymbol{Y} &= [\ \bar{\boldsymbol{y}}'_{d_0}\ \ \bar{\boldsymbol{y}}'_{d_1}\ \ \cdots\ \ \bar{\boldsymbol{y}}'_{d_{\mathcal{D}_2-1}}\ ], && (20)\\ \boldsymbol{X} &= [\ vec(\boldsymbol{S}(R_3-1))\ \ \cdots\ \ vec(\boldsymbol{S}(1))\ \ vec(\boldsymbol{S}(0))\ ] && (21)\end{aligned}$$

Then, denoting $\boldsymbol{y} = vec(\boldsymbol{Y})$ and $\boldsymbol{x} = vec(\boldsymbol{X})$, we write

$$\boxed{\boldsymbol{y} = \boldsymbol{\mathcal{H}}\boldsymbol{x} + \boldsymbol{v}} \quad (22)$$

where $\boldsymbol{\mathcal{H}}$ is now a 4-level block banded convolution matrix, with dimensions $Q_r(N+B-1)(R_1+K_1-1)(R_2+K_2-1)(R_3+K_3-1) \times BR_1R_2R_3$.

Again, we can proceed in finding an estimate of the reflectivity tensor $\boldsymbol{S}$ in (22) via matrix inversion or LS, with $\boldsymbol{H}_{\boldsymbol{r}_j}$ replaced by $\boldsymbol{\mathcal{H}}$. Note that the tall full rank condition imposed for separate blocks becomes much more relaxed in this case due to the redundancy of reflectivity in adjacent blocks.

[1]The fact is that, in addition to the 1-D time convolution model for each focal point, scanning at an arbitrary resolution rate yields overlapping of successive blocks, which naturally characterizes a volumetric convolution between the object, and the space-varying tensor defined by the beam pattern for different focal points. The spatial model for this operation will depend on the number directions and depths the beam is moved to, whereby, similarly to a block processing in time, motion can induce a particular processing structure in space (e.g., block processing, with or without overlap). The overall process is in general a 4-D MIMO space (time)-varying convolution model. For instance, choosing $\mathcal{D} = (R_1+K_1-1)(R_2+K_2-1)(R_3+K_3-1)$ implies scanning at the the spatial resolution $\{T_0, T_1, T_r\}$, and corresponds to a complete 4-D MIMO (space-varying) convolution. On the other hand, for the fastest scanning, we can pick $\mathcal{D} = R_1R_2R_3/|\det \boldsymbol{B}|$ foci, and the whole space is scanned with nonoverlapping blocks. This tells us that unlike the unfocused scenario, or in case information received from around the focal point is not sufficient for independent estimations of each block, focusing allows for further diversity; in this sense, resolution is unlimited, and ultimately dictated by the desired spatial sampling of $\boldsymbol{S}$, and not only by the sampling device. As a result, estimation must be accomplished jointly, after data received from around all foci have been properly acquired.

4. PRELIMINARY EXPERIMENTS

We assume a free-space scenario where a beam pattern is generated by a 1m $\times$ 1m antenna with uniform tappering function and current density, producing two separable *sinc* functions $b_0(n)$ and $b_1(n)$. We estimate a tensor, with resolution $T_1 = T_2 = \Delta_{\min} = 0.1$. The pulse length is $T_p = 10^{-7}$, and the carrier frequency is set to $f_c = 4$ GHz. In the A-scan imaging, the beam is swept over object, and for each received signal, the estimator is computed based solely on the transmitted pulse, while the beam pattern is usually designed in order to minimize the blurring effect in this direction. In the 3D modeling, however, whatever the beam pattern in cross range is, we make use of its shape when estimating along this direction as well. Observe that a tall matrix can always be induced by moving the beam over the object so as to perform a complete spatial convolution. Hence, improved results may be expected compared to the case of A-scans.

We consider an 80 $\times$80 scene image 3(a) at a fixed range and an SNR of 10dB, for a randomly generated pulse with Gaussian samples. The optimal minimum-norm matrix inverse and 3D-LS estimation are illustrated in 3(b) and 4(c) respectively. These can be compared with the corresponding A-scan schemes in 4(d) and 5(e). We further compare these recovered scenes with the ones obtained with a Chirp transmission, in 6(a) and 6(b). We clearly see the effect of conditioning in the matrix inverse solutions, and superiority of the full convolution information based receivers

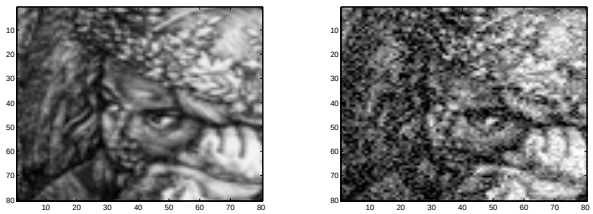

Fig. 3. (left) Original scene; (right) Optimal minimum norm matrix inverse.

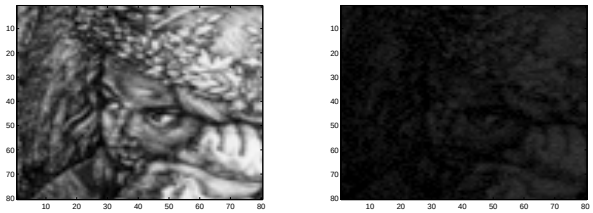

Fig. 4. (left) 3D-LS solution; (right) Optimal minimum norm matrix inverse for A-scan processing.

In this paper, we have presented preliminary simulations considering the effect of the transmitted signal in A-scan and

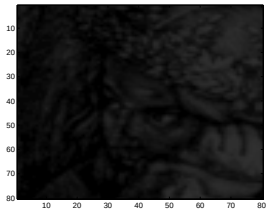

Fig. 5. 3D-LS solution for A-scan processing.

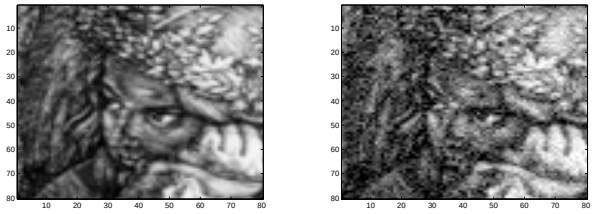

Fig. 6. Chirp transmit pulse. (left) 3D-LS solution; (right) Optimal minimum norm matrix inverse for A-scan processing.

3D estimation scenarios. The behavior of equalizers are well established in a 1D scenario, and the challenge in their implementations lies in cases where the conditioning of $\boldsymbol{\mathcal{T}}$, $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$ becomes too high. A thorough comparison with other methods in fair equivalent scenarios is still to be pursued in a separate future work.

5. REFERENCES

[1] R. Merched, "Superresolution and superfast receivers in free-space radar imaging," *ICASSP 2011, Prague, Czech Republic*, pp. 1241 - 1244, May 2011.

[2] R. J. Zemp, C. K. Abbey, M. F. Insana, "Linear system models for ultrasonic imaging: application to signal statistics,"*IEEE Trans. Ultrason. Ferro. Freq. Contrl.*, vol. 50, no. 6, pp. 642–654, June 2003.

[3] T. Taxt and J. Strand, "Two dimensional noise-robust blind deconvolution of Ultrasound Images," *IEEE Trans. Ultrasonics, Ferroelect., and Freq. Control*, 48(4), pp. 555–566. July 2001.

[4] C.-Y. Chen, and P. P. Vaidyanathan, "MIMO Radar Waveform Optimization With Prior Information of the Extended Target and Clutter," *IEEE Trans. Signal Processing*, vol. 57, no. 9, pp. 3533–3544, Sep. 2009.

[5] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers — Part II: Blind Channel Estimation, Synchronization, and Direct Equaization," *IEEE Trans. on Signal Processing*, vol. 47, pp. 2007–2022, July. 1999.