Designing a 3-D sonar system using a sparse array and CZT beamforming

Maria Palmese (1), Andrea Trucco (1)

(1) Dept. of Biophysical and Electronic Engineering, University of Genoa, Italy

ABSTRACT

Critical issues in the development of high-resolution 3-D sonar systems are the cost of hardware associated with the large number of sensors composing the planar array and the computational burden of processing in real-time the signals gathered. In this paper, such problems are overcome by the optimized synthesis of an aperiodic sparse array and the efficient processing of the acquired signals carried out in the frequency domain and based on Chirp Zeta Transform (CZT) beamforming. On the one hand, the synthesized sparse array enables the device to operate at frequencies yielding an acceptable side-lobe level and a good tradeoff between the sector of view and the resolution. On the other hand, the CZT beamforming, specifically devised to cope with the requirements of volumetric sonar imaging, allows the processing of wideband signals collected by a planar array and generated by a scene encompassing both near-field and far-field regions. The combination of a very limited number of sensors with the CZT beamforming generates a computational load that is two orders of magnitude lower than that of the delay-and-sum beamforming and one order of magnitude lower than that of the traditional frequency-domain implementation. The reduction of the number of sensors and the computational load produces, in turn, a noticeable reduction of the hardware cost.

I. INTRODUCTION

Despite the introduction of commercial equipment, development of underwater systems capable of generating a real-time three-dimensional (3-D) acoustic video of an investigated environment [1,2] is still a challenge, mainly due to the cost of the planar array required and the computational load associated with the digital beamforming technique [1]. Moreover, 3-D imaging systems should be able to work in the near-field region, imposing a focused beamforming approach [1–3], and this causes additional difficulties, especially when beamforming is implemented in the frequency domain.

The computational burden and hardware requirements of digital beamforming implemented in the time-domain can be mitigated either by using one of the approximate implementations [4–6] proposed in recent years or by adopting a sparse planar array of sensors [7]. A solution based on the adoption of a sparse array makes it possible to strictly contain both the cost of the system front-end and the cost of the architecture to support the computational burden of processing the acquired signals.

Beamforming implementation in the frequency domain represents an attractive solution, which has been extensively investigated [3, 8, 9] in an attempt to exploit the computational advantages of the fast Fourier transform (FFT). In [9], Maranda thoroughly described and compared some frequency-domain beamforming methods with reference to one-dimensional (1-D) array and far-field condition. He divided these into exact and approximate methods.

Traditionally, "direct method" (DM) beamforming is the most widely applied exact method, and its extension to planar arrays and near-field conditions has been introduced and extensively investigated [10, 11].

Although less recognized, Chirp Zeta Transform (CZT) beamforming [9] is a flexible frequency-domain method that can process wideband signals without any error, with a computational load equal to or lower than that of approximated FFT classical beamforming.

The aim of this paper is the preliminary design of a high-resolution 3-D sonar system, working in real-time, with a reduced cost. To do that, the combination of two methods is proposed: the first method is devoted to the synthesis of a sparse planar array, whereas the second one performs the near-field beamforming with a very limited computational load. One of the most promising approaches to reducing the number of array elements (for both linear and planar arrays) is based on the concept of aperiodic arrays. Typically, a fully sampled array is thinned by removing a fraction of the original set of elements, thus obtaining a sparse array. The thinning should be an optimization operation aimed at reducing the number of elements, while maintaining adequate beam pattern properties for the desired application. The authors devised an array synthesis method that is based on the simulated annealing stochastic optimisation [7]. The purpose of the method is to minimize the number of elements of a planar array able to generate a beam pattern that fulfils some *a priori* fixed constraints by acting on the positions and weights of the sensors. Notwithstanding the strong reduction of the sensor number, the processing by a delay-and-sum (D&S) beam-

forming can result still demanding. To overcome this problem, a beamforming implementation based on CZT is proposed. The authors devised a spatial processing method for wideband 3-D imaging, based on CZT beamforming, that is specifically tailored to image a volume of space encompassing the near-field and far-field regions [12].

This paper is organized as follows. Section II introduces the sparse array concept and describes its optimization in the specific case addressed. The CZT formulation for a planar array and near-field conditions is presented in Section III. The concepts for the evaluation of computational load of different beamforming implementations in relation to the generation of a volumetric image are introduced in Section IV. In Section V, the synthesized sparse array is described and the computational load of the CZT beamforming is assessed and compared with those of traditional beamforming implementations. Finally, the conclusions are drawn in Section VI.

II. DESIGNING A SPARSE PLANAR ARRAY

As a starting point, a receiving array composed of 100×100 transducers is hypothesized and, according to a recognized technological constraint in sonar manufacture, an inter-element space $d = 2.5$ mm is allowed. Thus, the resulting spatial aperture is a square with a side of about 25 cm.

On the basis of other similar projects [13-16], a frequency value around 0.5 - 1 MHz seems to be well suited to this kind of sonar. Here it is tentatively assumed that the system works at 600 kHz. The maximum range that one can expect is of several tens of meters [17].

The angular resolution attainable by the depicted system (measured at –3 dB, in the broadside direction), without the application of any window to abate the level of the side lobes [1,17], is 0.50°. (However, it is important to remember that the reduction of the side lobe level, the highest of these being –13 dB, is always accompanied by a worsening of the angular resolution.) The corresponding lateral resolution at a distance of 1 m from the array is 9 mm. The range resolution can be evaluated by defining the type and the bandwidth of the insonifying pulses [17]. Assuming a transmission of simple burst pulses with a 25% fractional bandwidth, one obtains a range resolution of 5 mm.

However, the λ/2-condition has not been respected as the actual inter-element spacing, d, is equal to λ. Consequently, the related beam pattern (i.e., the spatial responses of the array at the fixed frequency) shows several grating lobes. To avoid any ambiguity effect, it is necessary to limit both the insonification and steering operations to a limited angular sector [1]. Specifically, the maximum steering angle (both in azimuth and elevation) to avoid ambiguities is: $\theta_{max} = \pm\arcsin(\lambda_{min}/2d)$. Considering the minimum wavelength contained in the hypothesized band, the allowed field of view will be 52°×52° at 600 kHz (roughly corresponding to the field of view of the lens of a 35-mm optical camera with focal lengths of 35 mm).

Undoubtedly, the above-mentioned features characterize a very high performance 3-D sonar. However, the cost and the complexity associated with a 100×100 receiving array are absolutely prohibitive, thus imposing a reduction in the number of transducers.

One of the most promising approaches to reducing the number of array elements (for both linear and planar arrays) is based on the concept of aperiodic arrays [18-20]. Typically, a fully sampled array (i.e., a dense array that respects the λ/2-condition) is thinned by removing a fraction of the original set of elements, thus obtaining a sparse array. Aliasing effects (i.e., the grating lobes) are avoided because there are no periodicities at the positions of the sparse array elements. The main drawback of the thinning operation is the often unacceptably high level of the side lobes present in the beam pattern. Therefore, the thinning should be an optimization operation aimed at reducing the number of elements, while maintaining adequate beam pattern properties for the desired application.

However, in the case treated here the dense 100×100 array does not respect the λ/2-condition, due to technological constraints, and some grating lobes are present in the beam pattern of such an array. If a portion of the initial array elements is removed to create an aperiodic sparse array, the side lobes increase and the grating lobes keep their position in the beam pattern. Consequently, to avoid any ambiguity effect, the limitation of the field of view is still mandatory, and the aim of the thinning operation is to reduce the number of array elements while maintaining an acceptable level for the side lobes.

As the distribution and the height of the side lobes depend on the positions of the sparse-array elements and on the weight coefficients assigned to such elements, the optimization problem can be approached from different points of view. To optimize both positions and weights is the most ambitious and effective procedure for sparse-array synthesis. Here, the array synthesis technique is proposed that is based on a stochastic algorithm known as simulated annealing.

The purpose of the method is to minimize the number of elements of a planar array able to generate a beam pattern (BP) that fulfils some *a priori* fixed constraints by acting on the positions and weights of the sensors. The use of simulated annealing to synthesize a 2-D sparse array requires the choice of an energy function, $f(\mathbf{X},\mathbf{W})$, that depends on the vector of the element positions, $\mathbf{X}$, and on the vector of the weights, $\mathbf{W}$. Once a desired normalized BP is fixed, the energy function must be able to penalize array configurations that both yield a great difference between the current BP and the desired one and are composed of a large number of active elements. To this end, one can choose:

$$f(\mathbf{X},\mathbf{W}) = \left[k_1 \iint\limits_{(u,v)\in S} \left(\frac{p(u,v)}{Q} - p_d(u,v) \right) du\,dv + k_2 M \right]^2 \quad (1)$$

where $u = \sin\alpha - \sin\alpha_0$, $v = \sin\beta - \sin\beta_0$, the pairs (α, β) and (α₀, β₀) indicate the arrival direction and the scan direction, respectively, S is the set of values of (u, v) satisfying the relation $p(u, v)/Q > p_d(u, v)$, Q is the sum of all the sensor weights, $p(u, v)$ is the current BP, $p_d(u, v)$ is the desired BP, M is the number of current active elements, and k_1 and k_2 are two constants denoting the relative importance assigned to the discrepancy between the actual BP and the desired one and to the number of array elements. The method, by exploiting the simulating annealing minimization scheme, provides a final configuration that is valid for each possible steering direction.

The thinning of the 100×100 array was carried out assuming, for a while, a λ/2 spacing. First of all, the array is inscribed inside a circle, thus giving 7854 active elements. This operation does not significantly affect the array performance. Then, the fully sampled array was optimally thinned and weighted by the method proposed in [20] and briefly recalled

in the previous paragraph. In order to constrain the side lobe peak under a fixed target, equal to –22 dB, the desired beam pattern was defined as a constant level of –22 dB for each pair (u, v) that was out of a circle of radius 0.02, i.e., $u^2+v^2 > 0.02^2$. The latter condition is imposed to exclude the main lobe from the constrained side lobe region. The variables u and v are defined as $\sin\alpha - \sin\alpha_0$ and $\sin\beta - \sin\beta_0$, respectively, where the pairs (α, β) and (α_0, β_0) indicate the arrival direction and the steering direction, respectively. The adoption of the u-v space is very useful as it makes it possible to synthesize an array configuration that produces a beam pattern showing a desired profile for whatever steering direction. The variables u and v can only assume real values between -2 and 2, for any combination of the arrival and steering directions. However, according to [20], thanks to some symmetry-related properties of the beam pattern, the domain of the beam pattern can be restricted to $u \in [-1,1]$, $v \in [0,1]$, during the optimization stage and to assess the obtained results without any loss of information. Finally, to further reduce the computational burden, the symmetry of the sparse array layout around the x- and y-axes (thus acting only on a quarter of the array elements) has been imposed.

III. DEVISING AN EFFICIENT BEAMFORMING ALGORITHM

Let us describe the CZT beamforming assuming to work with a dense planar array. A sparse array (synthesized starting from a dense planar array) can been seen as a dense array where the weight coefficient assigned to some elements have a value equal to zero.

Thus, let us consider a planar array placed on the plane xy, composed of $M{\times}N$ sensors. The sensor, identified by the indices (m, n), is placed at position (x_m, y_n) and generates the signal $s_{m,n}(t)$. The steering direction is identified by the azimuth and elevation angles (θ_a, θ_e) which, in contrast to the conventional formalism [2, 21], are defined as shown in Fig. 1. The choice of such angles is useful for easily extending the CZT beamforming to a planar array. According to such a notation, the unit vector of the steering direction, $\hat{\mathbf{u}}$, can be expressed as:

$$\hat{\mathbf{u}} = \left[\sin\theta_a, \sin\theta_e, \sqrt{\cos^2\theta_a - \sin^2\theta_e} \right] \qquad (2)$$

where θ_a is the angle between the vector $\hat{\mathbf{u}}$ and its projection on the plane yz, and θ_e is the angle between the vector $\hat{\mathbf{u}}$ and its projection on the plane xz.
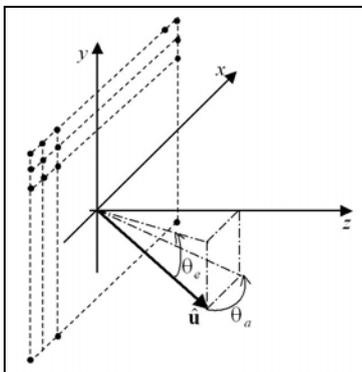
Figure 1. Notation and geometry for a planar array

According to D&S beamforming [1], if the far-field condition is still valid, the beam signal, $b(t, \theta_a, \theta_e)$, steered in the direction (θ_a, θ_e), can be expressed as follows:

$$b(t,\theta_a,\theta_e) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} w_{m,n}\, s_{m,n}\!\left(t - \frac{x_m\sin\theta_a + y_n\sin\theta_e}{c} \right), \qquad (3)$$

where $w_{m,n}$ is the apodizing weight assigned to the (m,n) sensor. The extension of the CZT beamforming to an equispaced planar array working in far-field conditions has been introduced and discussed in [12, 22].

This paper focuses on near-field conditions, where the curvature of the wavefront cannot be neglected, and a focusing distance r_0 is introduced in the delay term [1, 3, 21], which, in turn, becomes much more complex. The D&S beamforming expression in Eqn. (3) is modified as follows:

$$b(t,\theta_a,\theta_e,r_0) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} w_{m,n}\, s_{m,n}\!\left(t - \tau_{m,n}(\theta_a,\theta_e,r_0) \right) \qquad (4)$$

$$\tau_{m,n}(\theta_a,\theta_e,r_0) = \frac{r_0 - \sqrt{r_0^2 + x_m^2 + y_n^2 - 2x_m r_0\sin\theta_a - 2y_n r_0\sin\theta_e}}{c}$$

$$(5)$$

The computation of the delay terms given in the last equation is very significant; therefore, it is necessary to compute and store them a priori. The number of possible combinations among m, n, θ_a, θ_e and r_0 requires a very large memory. To simplify such a delay and reduce the memory needs, the Fresnel approximation can be applied [1, 3, 21], obtaining the following term:

$$\tau_{m,n}(\theta_a,\theta_e,r_0) = \frac{x_m\sin\theta_a}{c} + \frac{y_n\sin\theta_e}{c} - \frac{x_m^2 + y_n^2}{2r_0 c} \qquad (6)$$

The validity region of the Fresnel approximation (i.e., the region where approximation errors are negligible) has been discussed by Ziomek [21].

The beamforming implementation in the frequency domain requires, as in the linear-array case, the temporal sampling of the signals gathered by the sensors, producing the time series $s_{m,n}(l)$ followed by the segmentation into partially overlapped blocks of length K and the transformation into their frequency versions, $S_{m,n}(k)$, by the FFT. If the Fresnel approximation is adopted, the DFT coefficients $B(k, \theta_a, \theta_e, r_0)$ of the beam signal $b(t, \theta_a, \theta_e, r_0)$ are given by the following expression:

$$B(k,\theta_a,\theta_e,r_0) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} w_{m,n}\exp\!\left(j2\pi f_k\!\left(\frac{x_m^2 + y_n^2}{2r_0 c} \right) \right) S_{m,n}(k)\cdot$$

$$\cdot \exp\!\left(-j2\pi f_k\!\left(\frac{x_m\sin\theta_a}{c} + \frac{y_n\sin\theta_e}{c} \right) \right) \qquad (7)$$

which represents the DM for a near-field, frequency-domain beamforming.

Let us now assume that the planar array placed on the plane xy is centered on the coordinate origin and equispaced with an intersensor spacing d in both directions. Let us consider the computation of a pyramid of beams composed of $M_b{\times}N_b$ signals, identified by the indices (p, q). For each of these beams, the steering direction is $(\theta_{a,p}, \theta_{e,q})$, where $\theta_{a,p}$ and $\theta_{e,q}$ are the two angles chosen from a set of M_b azimuth angles and a set of N_b elevation angles, respectively. The predefined angles are equispaced in the sine domain, ranging from an initial to a final angle. After some mathematics, the kth Fou-

rier coefficient of the beam signal (p, q), focused at a distance r_0, can be written as follows:

$$B(k,\theta_{a,p},\theta_{e,q},r_0) = W_a^{\frac{-p^2}{2}} W_e^{\frac{-q^2}{2}} \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} v_{m,n} S_{m,n}(k) A_a^m A_e^n \cdot$$
$$\cdot W_a^{\frac{-m^2}{2}} W_e^{\frac{-n^2}{2}} W_a^{\frac{(p-m)^2}{2}} W_e^{\frac{(q-n)^2}{2}} \qquad (8)$$

The terms W_a, A_a, W_e, and A_e are defined as follows:

$$A_a = \exp\left(-j2\pi f_k \frac{d}{c}\sin\theta_{ai}\right) \qquad (9)$$

$$W_a = \exp\left(j2\pi f_k \frac{d}{c}\Delta s_a\right) \qquad (10)$$

$$A_e = \exp\left(-j2\pi f_k \frac{d}{c}\sin\theta_{ei}\right) \qquad (11)$$

$$W_e = \exp\left(j2\pi f_k \frac{d}{c}\Delta s_e\right) \qquad (12)$$

where

$$\Delta s_a = \frac{\sin\theta_{af}-\sin\theta_{ai}}{M_b - 1}, \quad \begin{array}{l}|\theta_{ai}|\leq\dfrac{\pi}{2}\\[4pt]|\theta_{af}|\leq\dfrac{\pi}{2}\end{array} \qquad (13)$$

$$\Delta s_e = \frac{\sin\theta_{ef}-\sin\theta_{ei}}{N_b - 1}, \quad \begin{array}{l}|\theta_{ei}|\leq\dfrac{\pi}{2}\\[4pt]|\theta_{ef}|\leq\dfrac{\pi}{2}\end{array} \qquad (14)$$

θ_{ai} and θ_{af} are the initial and final steering angles for the azimuth respectively, and θ_{ei} and θ_{ef} are the initial and final steering angles for the elevation respectively. The steering angles $\theta_{a,p}$ and $\theta_{e,q}$ can be derived from the following relations:

$$\sin\theta_{a,p} = \sin\theta_{ai} + \Delta s_a\, p \quad , \qquad p = 0,\dots,M_b-1 \qquad (15)$$

$$\sin\theta_{e,q} = \sin\theta_{ei} + \Delta s_e\, q \quad , \qquad q = 0,\dots,N_b-1 \qquad (16)$$

The term $v_{m,n}$ is defined as follows:

$$v_{m,n} = w_{m,n}\exp\left\{j2\pi f_k\left[\left(m-\frac{M-1}{2}\right)^2 + \left(n-\frac{N-1}{2}\right)^2\right]\frac{d^2}{2r_0 c}\right\} \qquad (17)$$

In fact, when moving from Eqn. (7) to Eqn. (8), a phase term linear with f_k, which does not depend on the sensor indices, has been neglected. This is reasonable as such a term simply introduces a known delay in the beam signal that can be easily fixed a posteriori.

Eqn. (8) expresses $B(k,\theta_{a,p},\theta_{e,q},r_0)$ as a two-dimensional (2-D) discrete convolution of two $M\times N$ matrixes, $\mathbf{C}(k)$ and $\mathbf{D}(k)$, whose elements of indices (f, g) are defined as follows:

$$C_{f,g}(k) = v_{f-1,g-1}\, S_{f-1,g-1}(k)\, A_a^{f-1} A_e^{g-1} W_a^{\frac{-(f-1)^2}{2}} W_e^{\frac{-(g-1)^2}{2}} \qquad (18)$$

$$D_{f,g}(k) = W_a^{\frac{(f-1)^2}{2}} W_e^{\frac{(g-1)^2}{2}} \qquad (19)$$

where $f \in [1, M]$ and $g \in [1, N]$, thereby allowing FFT methods to perform the 2-D "fast convolution" [23] effectively. In Section IV, an evaluation of the specific computational load is presented. Here, it is worth noting that at a single frequency f_k, the convolution FFT-based method simultaneously computes all the values of the $M_b\times N_b$ beams. After all the segments of the beams have been computed over the desired frequency band, $M_b\times N_b$ inverse FFTs are necessary to obtain the beam time series.

In focused beamforming (independent of the time or frequency implementation), the depth of field (DOF) [1, 24] is defined as the range interval around the focusing distance r_0 inside which the performance, evaluated in term of angular resolution and amplitude gain, only marginally degrades. The extension of the DOF depends on the specific value of r_0: the smaller the focusing distance r_0, the smaller the DOF extension [24]. Unfortunately, the range extension of the volume to be imaged typically exceeds that of the DOF. The problem is commonly solved by segmenting the received time signals into subsequent blocks that are processed using different focusing distances, which increase with time. In other words, the volume to be imaged is subdivided into multiple, adjacent focal regions. Each focal region is centered on a specific focusing distance and is contained inside the DOF corresponding to such a focusing distance. As in our method the time signals are already segmented into blocks of length K, so it is easy to apply different focusing distances to each block. However, in setting the value of K, it is essential that the spatial extension corresponding to the K time samples does not exceed the shortest of the extensions of the focal regions. The general method and the specific solutions described above are used to extend the CZT beamforming to the planar array case and to allow correct focusing over an extended near-field volume. The accuracy of computing wideband beam signals by the proposed method is equal to that of the traditional delay-and-sum (D&S) beamforming, [1, 3] except for the errors introduced by the Fresnel approximation. However, as discussed and quantified in [25] such errors are really negligible inside the validity region [21] of the approximation. Therefore, the described method represents a computationally convenient way to compute beam signals that, inside the Fresnel validity region, negligibly differ from those computed by D&S beamforming. Below, two original solutions devoted to further reducing the computational burden will be briefly introduced:

(1) It can generally be observed that it is computationally convenient to increase the length K of the signal blocks. This is in contrast to the need for short blocks close to the array, where the DOF extension is more limited. To overcome this problem the length of the signal blocks can vary, increasing in synchrony with the DOF extension.

(2) The angular resolution is fixed and so the lateral resolution worsens with the distance. Instead, the range resolution depends on the bandwidth and does not generally vary with the distance. However, the generation of a cubic resolution cell is often welcome, [1] and the worsening of the range resolution with the distance is perfectly acceptable. This makes it possible to save many operations by reducing the number of frequency bins considered (i.e., the bandwidth) with the distance. In other words, Eq. (8) can be computed for a number of indexes k that decreases block after block, according to the desired range resolution.

IV. COMPUTATIONAL LOAD EVALUATION

The aim of this section is to provide information useful for evaluating the minimum number of operations needed to compute a whole 3-D image. The following beamforming techniques are considered: (1) time-domain D&S beamforming; (2) frequency-domain DM beamforming; and (3) frequency-domain CZT beamforming.

In this paper, the number of operations considered includes both real additions and real multiplications without distinction. Only the operations that need to be performed strictly online have been taken into account, disregarding the operations that can be performed offline (i.e., before starting the 3-D image generation), with the result stored in the system memory. It is assumed that the dense planar array is square and made of $N \times N$ elements, the sparse array is made of N_s total elements and that the volumetric image is computed through a square pyramid of beams (i.e., $M_b = N_b$).

With regard to D&S beamforming, according to [3], the oversampling of the input signals through an interpolation filter is necessary before the signal sum, to precisely implement the time delays. Here, it is assumed that the desired degree of oversampling is achieved by using an interpolation FIR filter consisting of H stages. The rate of real operations [3] necessary to generate N_b^2 beams is the following:

$$\mathrm{TD}_1 = \left[N_s \left(H + N_b^2 \right) + N^2 (H+1) \right] f_s \tag{20}$$

The computation of delays is considered an offline operation.

Now, let us consider the frequency-domain implementations. After an offline stage comprising all operations that can be performed only once for a given parameter configuration, the first step consists of computing the DFT coefficients for each sensor signal of the sparse array (N_s elements). By exploiting the FFT [3], the number of real operations necessary to process a real K-length (K being a power of two) time sequence is equal to:

$$FD_1 = 2.5K \log_2 (K/2) + 7K \tag{21}$$

The second step consists of spatial processing. DM beamforming is defined in Eqn. (7), which can be interpreted as a complex dot product. If the pair of indices (m,n) is replaced by a single index h ranging from 0 to $N_s - 1$, then Eqn. (7) can be considered as the dot product between two vectors, each of size N_s. The computation of each dot product requires N_s complex multiplications and $N_s - 1$ complex additions. Therefore, the generation of N_b^2 beams at a single frequency f_k by DM beamforming requires the following number of real operations:

$$FD_2 = (8N_s - 2) N_b^2 \tag{22}$$

If the generation of cubic resolution cells is enabled, only the frequency bins f_k included in the bandwidth Q should be considered.

In contrast to DM beamforming, Eqn. (8) allows one to interpret the spatial processing of CZT beamforming as a discrete convolution. The following are the steps of the algorithm:

1) Creation of the matrix $\mathbf{C}(k)$ by the complex product of each DFT coefficient $S_{f-1,g-1}(k)$ and the corresponding coefficients, computed offline, as shown in Eqn. (18).

2) 2-D convolution of the matrix $\mathbf{C}(k)$ with the matrix $\mathbf{D}(k)$. The latter is computed offline.

3) Complex product of each frequency sample with the corresponding coefficient, $W_a^{-p^2/2} W_e^{-q^2/2}$, computed offline, as shown in Eqn. (8).

With reference to step 2), the matrices $\mathbf{C}(k)$ and $\mathbf{D}(k)$ have a size $N \times N$, while the output matrix should have a size $N_b \times N_b$. To prevent wrap-around from contaminating the computation of the linear convolution, the matrix $\mathbf{D}(k)$ should have a size $L_n \times L_n$, where L_n satisfies the relation $L_n \geq N + N_b - 1$. To fully exploit the FFT advantage, L_n should be a power of two; therefore, it is convenient to set L_n equal to the first power of two, greater than or equal to $N + N_b - 1$.

The FFT implementation of such a convolution can be performed according to the following routine:

- The matrixes $\mathbf{C}(k)$ and $\mathbf{D}(k)$ are zero-padded to attain a size equal to $L_n \times L_n$;

- The 2-D FFTs of both the zero-padded matrixes are performed;

- Each coefficient of the first DFT is multiplied by the relative coefficient of the second DFT, requiring $L_n \times L_n$ complex multiplications;

- A 2-D IFFT of the result obtained in the previous step is performed.

The zero-padding of the matrix $\mathbf{D}(k)$ and its 2-D DFT are performed offline. Therefore, the generation of N_b^2 beams at a single frequency f_k by CZT beamforming requires the following number of real operations:

$$FD_3 = 6 \left[N^2 + N_b^2 + L_n^2 \right] + 20 L_n^2 \log_2 (L_n). \tag{23}$$

Finally, as in the case of DM beamforming, if the generation of cubic resolution cells is enabled, only the frequency bins f_k included in the bandwidth Q should be considered.

After spatial processing, the final step for both DM and CZT beamforming involves the computation of time samples of each beam signal by performing an inverse FFT. If the dynamic range resolution is enabled, the number of considered frequency bins is equal to ζ, centered around the central frequency of the acoustic pulse. It is convenient to shift these bins around $f = 0$, thus obtaining the equivalent low-pass spectrum [3, 23] of each beam signal. This frequency-shift operation allows one to reduce the size of the inverse FFT, as the latter will be performed on a sequence of length ζ instead of K. For each block of beam signal, the number of real operations is equal to:

$$FD_4 = 5\zeta_2 \log_2 (\zeta_2) \tag{24}$$

where ζ_2 is the first power of two greater than ζ. Working in the described way, the final beam signal is characterized by a sampling interval that varies as a function of time. Therefore, some care is necessary in providing continuity of processing across blocks. In particular, a look-up table could be arranged to directly link the index of the time samples (of any beam signal, for a given setting of the minimum and maximum ranges) to the corresponding distance, providing useful information for the visualization stage.

V. RESULTS AND COMPARISONS

Starting from a receiving array composed of 100×100 transducers, a sparse array has been synthesized.

One of the best results obtained from the running of the synthesis algorithm was an array of 584 elements, with a current taper ratio (i.e., the ratio between the maximum and minimum weight coefficients [20]) of 3.28. Figure 2 shows the obtained array layout and the related weight coefficient profile. The resulting beam pattern is shown in Fig. 3. Despite the dramatic thinning carried out by the optimization process (i.e., 92.5% of the array elements were removed), and the very limited current taper ratio, one can observe that the obtained beam pattern perfectly meets the imposed side lobe constraint.
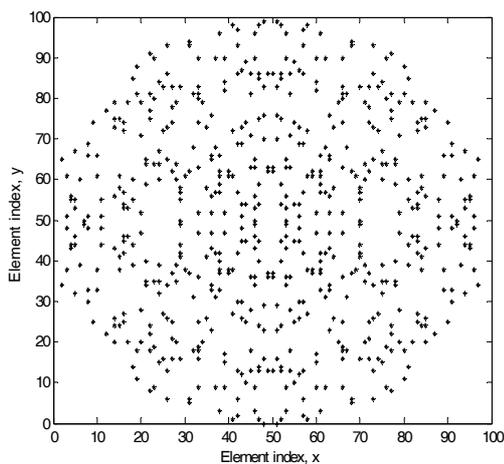
(a)
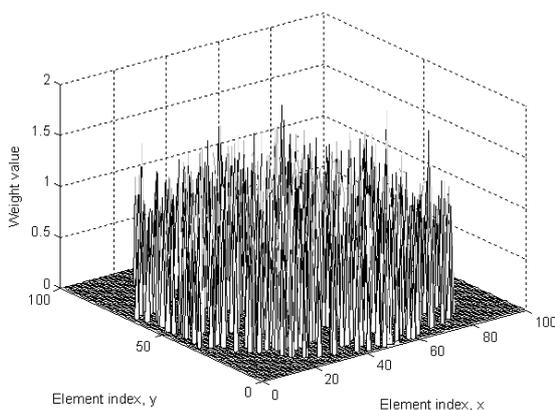
(b)

Figure 2. Optimized design of the aperiodic sparse array. (a) Array layout showing the placement of the 584 sensors. (b) Weight coefficients associated with the array sensors.

If the frequency at which the synthesized array is used is increased until it reaches an inter-element spacing equal to λ (i.e., 600 kHz), the side lobes remain under the −22 dB constraint, although some grating lobes arise in the beam pattern. Consequently, the field of view should be restricted as described in the Section II. Moreover, the thinning and weighting operations typically produce an enlargement of the main lobe with respect to that of a dense, uniformly weighted ar-

ray. For the optimized array, the angular resolution (measured at −3 dB, in the broadside direction) becomes 0.64°. Therefore, in this case, the resolution loss with respect to the dense array case is limited and compatible with the specific application envisaged.

The possibility of generating a sequence of 3-D images in real time is assured by the limited number of active elements in comparison with those used in a commercial device (i.e., 584 transducers against 1600 transducers [17]) currently performing real-time 3-D underwater imaging with similar features. The sparse array allows *per se* a drastic reduction in both cost and processing load. A further reduction of the processing load can be obtained applying the proposed CZT beamforming.
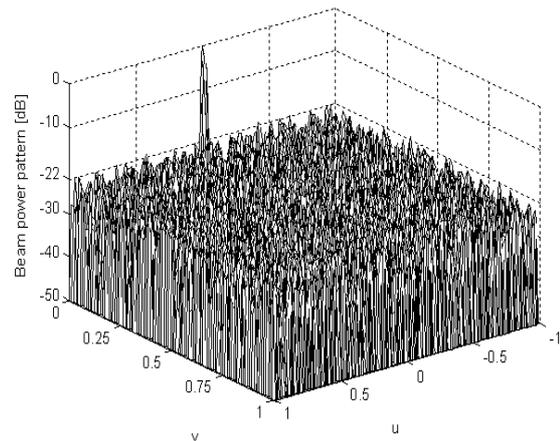
Figure 3. Beam pattern of the aperiodic sparse array composed of 584 sensors, with the positions and the weights shown in Fig. 2.

The equations introduced in the previous section are useful to compute the total number of operations needed to generate a whole 3-D image as well as to compare the computational loads of the three beamforming techniques considered. To make such a comparison possible, it is necessary to set some of the variables that influence the computational burden.

The acoustic pulse is assumed to have a center frequency of 600 kHz and a bandwidth of 150 kHz. The spacing of the array grid is 2.5 mm, the sector of view is 52°×52° and is covered by 200×200 beams, the extension of the volume to be imaged ranges from 1 m to 50 m. Finally, the sampling frequency is 1.8 MHz. For the time-domain beamforming, an interpolating filter with $H = 100$ stages is assumed [3]. For both the frequency-domain implementations (i.e., DM and CZT), the solution described in Section III (i.e., the dynamic range resolution) has been assumed enabled in the evaluation of the computational burden.

The count of the on-line real operations necessary to generate a whole 3-D image with the three beamforming methods when the signals are received by a dense array (100×100 sensors) and by the synthesized sparse array (584 sensors) is shown in Fig. 4. The computational analysis demonstrated that the advantage of CZT beamforming with respect to D&S time-domain beamforming and DM frequency-domain beamforming is significant. For both array layouts, the advantage of CZT beamforming with respect to DM and D&S methods is of about one order and two orders of magnitude, respectively. In particular, in the sparse array case the number of on-line operations is adequately decreased for all the beam-

forming methods achieving the minimum for the CZT algorithm implementation.

(a)

(b)
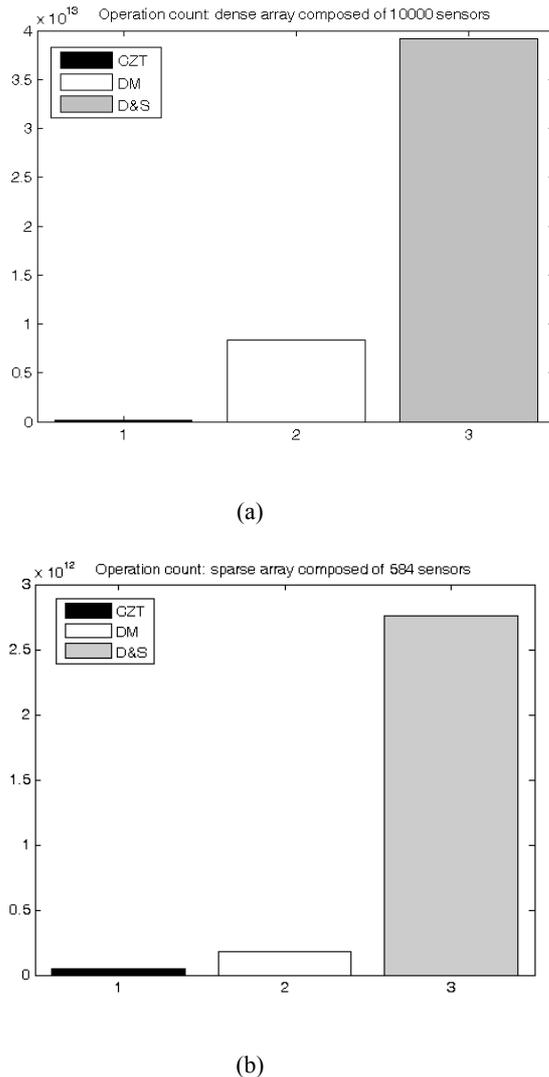
Figure 4. Operation count for the three beamforming methods (i.e., CZT, DM, D&S) necessary to create one 3-D image using (a) a dense array with 100×100 elements; (b) the sparse array with 584 elements.

VI. CONCLUSIONS

In this paper, the preliminary design of a high-resolution 3-D acoustic imaging system, based on a sparse planar array of sensors and Chirp Zeta Transform beamforming, has been proposed. The hypothesized sonar system could be realized with a limited hardware cost thanks to the reduction of the number of sensors and the reasonable computational load made possible by the combination of the two devised methods. In particular, starting form a dense planar array with 100×100 elements a sparse array composed of about 600 sensors has been syhnthesized. CZT beamforming implementation allows one to generate a 3-D image with a computational advantage with respect to D&S time-domain beamforming and DM frequency-domain beamforming which is of about one order and two orders of magnitude, respectively.

REFERENCES

1   V. Murino, A. Trucco, "Three-Dimensional Image Generation and Processing in Underwater Acoustic Vision," *Proceedings of the IEEE*, **88**, 1903-1946, (2000)

2   T.L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out* (Elsevier Academic Press, Amsterdam, 2004)

3   R.O. Nielsen, *Sonar Signal Processing* (Artech House, Boston, 1991)

4   V. Murino, C.S. Regazzoni, A. Trucco, G. Vernazza, "A Non-Coherent Correlation Technique and Focused Beamforming for Ultrasonic Underwater Imaging: A Comparative Analysis", *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, **41**, 621-630, (1994)

5   K. Ranganathan, M.K. Santy, T.N. Blalock, J.A. Hossack, W.F. Walker, "Direct Sampled I/Q Beamforming for Compact and Very Low-Cost Ultrasound Imaging," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, **51**, 1082-1093, (2004)

6   A. Agarwal, Y.M. Yoo, F.K. Schneider, C. Gao, L.M. Koh, Y. Kim, "New Demodulation Method for Efficient Phase-Rotation-Based Beamforming," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, **54**, 1656-1667, (2007)

7   A. Trucco, M. Palmese, S. Repetto, "Devising an Affordable Sonar System for Underwater 3-D Vision", *IEEE Trans. Instrument. and Measurement*, **57**, 2348-2354, (2008)

8   R.A. Mucci, "A Comparison of Efficient Beamforming Algorithms," *IEEE Trans. Acoust., Spec., Signal Proc.*, vol. ASSP-32, 548-558, (1984)

9   B. Maranda, "Efficient Digital Beamforming in the Frequency Domain," *Journal of Acoustical Society of America*, **86**, 1813-1819, (1989)

10 F. Zhang, A. Bilas, A. Dhanantwari, K.N. Plataniotis, R. Abiprojo, S. Stergiopoulos, "Parallelization and Performances of 3D Ultrasound Imaging Beamforming Algorithms on Modern Clusters", *Proc. of the 16th ACM Int. Conf. On Supercomputing*, New York (USA), (2002)

11 A.C. Dhanantwari, S. Stergiopoulos, L. Song, C. Parodi, F. Bertora, P. Pellegretti, A. Questa, "An efficient 3D beamformer implementation for real-time 4D ultrasound systems deploying array probes," *2004 IEEE Ultrasonics Symposyum*, Toronto, Ontario (Canada), 1421-1424, (2004)

12 M. Palmese, A. Trucco, "3-D Acoustic Imaging by Chirp Zeta Transform Digital Beamforming," *IEEE Trans. Instrument. and Measurement*, **58**, no. 7, 2080-2086, (2009)

13 R.K. Hansen, P.A. Andersen, "The application of real time 3D acoustical imaging", *IEEE/OES Int. Conf. Oceans '98*, Nice (France), 738-741, (1998)

14 J.S. Jaffe, "Target localization for a three-dimensional multibeam sonar imaging system", *Journal of Acoustical Society of America*, **105**, no. 6, 3168-3175, (1999)

15 A.M. Chiang, S.R. Broadstone, J.M. Impagliazzo, "A Portable, Electronic-Focusing Sonar System for AUVs using 2D Sparse-Array Technology," *Int. Conf. Oceans 2002 MTS/IEEE*, Biloxi, Mississippi (USA), 2129-2134, (2002)

16 J. Impagliazzo, M. Medeiros, S. Kay, "Imaging with a 2 MHz Sparse Broadband Planar Array," *Int. Conf. Oceans 2001 MTS/IEEE*, Honolulu, HA (USA), 22-25, (2001)

17 X. Lurton, *An introduction to Underwater Acoustics: principles and applications* (Springer, Praxis Publishing, Chichester, UK, 2002)

18 R.M. Leahy, B.D. Jeffs, "On the Design of Maximally Sparse Beamforming Arrays," *IEEE Trans. Antennas and Propagation*, **39**, 1178-1187, (1991)

19 S. Holm, B. Elgetun, G. Dahl, "Properties of the Beampattern of Weight- and Layout-Optimized Sparse Arrays," *IEEE Trans. Ultrason., Ferroelect., and Freq. Contr.*, **44**, 983-991, (1997)

20 A. Trucco, "Thinning and Weighting of Large Planar Arrays by Simulated Annealing," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, **46**, no. 2, 347-355, (1999)

21 L.J. Ziomek, *Fundamentals of Acoustic Field Theory and Space-Time Signal Processing* (CRC Press, Boca Raton, Florida, USA, 1995)

22 B.O. Odelowo, "A Fast Beamforming Algorithm for Planar/Volumetric Arrays," *Thirty-Ninth Asilomar Conference on Signals, Systems, and Computers,* 1707-1710, (2005)

23 A.V. Oppenheim, R.W. Schafer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, USA, 1973)

24 R. C. Hansen, "Focal region characteristics of focused array antennas," *IEEE Trans. Antennas Propag.* **33**, 1328–1337 (1985)

25 A. Trucco, "A least-squares approximation for the delays used in focused beamforming," *J. Acoust. Soc. Am.* **104**, 171–175 (1998).