A Confidence-Based Approach to Enhancing Underwater Acoustic Image Formation

Vittorio Murino, Member, IEEE, and Andrea Trucco, Student Member, IEEE

Abstract—This paper describes a flexible technique to enhance the formation of short-range acoustic images so as to improve image quality and facilitate the tasks of subsequent postprocessing methods. The proposed methodology operates as an ideal interface between the signals formed by a focused beamforming technique (i.e., the beam signals) and the related image, whether a two-dimensional (2-D) or three-dimensional (3-D) one. To this end, a reliability measure has been introduced, called confidence, which allows one to perform a rapid examination of the beam signals and is aimed at accurately detecting echoes backscattered from a scene. The confidence-based approach exploits the physics of the process of image formation and generic a priori knowledge of a scene to synthesize model-based signals to be compared with actual backscattered echoes, giving, at the same time, a measure of the reliability of their similarity. The objectives that can be attained by this method can be summarized in a reduction in artifacts due to the lowering of the side-lobe level, a better lateral resolution, a greater accuracy in range determination, a direct estimation of the reliability of the information acquired, thus leading to a higher image quality and hence a better scene understanding. Tests on both simulated and actual data (concerning both 2-D and 3-D images) show the higher efficiency of the proposed confidence-based approach, as compared with more traditional techniques.

Index Terms—Acoustic imaging, image enhancement, image generation, underwater acoustics.

I. INTRODUCTION

Typically, underwater acoustic imaging systems aim at the investigation of underwater scenes whenever operational conditions do not allow the observation of such scenes by using optical systems. Unfortunately, acoustic images are affected by a low quality, due to a set of problems that make it difficult to understand an observed scenario. Nevertheless, acoustic systems are useful in visualizing submerged objects (whether natural or artificial) and in generating precise seabed maps [1], [2].

Generally, the beamforming (BF) technique [3] is utilized for the formation of acoustic images, both two-dimensional (2-D) and three-dimensional (3-D) ones [1]–[6]. For 2-D imaging, a linear array is sufficient and the resulting image represents a section of a scene, visualized in a direction-distance format (typical of sonar systems [1]; see Fig. 1). For 3-D imaging, a 2-D array is required and the obtained image is similar to a conventional (orthoscopic) optical image with additional precise information about the scene distance, i.e., a range image. A 3-D map representation is also used to visualize this kind of data [2] (see Fig. 2). In 2-D imaging, successive reflections, generated by different scene layers or by different objects located in the same direction, are indispensable to visualize the external surface and the internal structure of a scene [7]–[9]. By contrast, to obtain an understandable 3-D image, typically, only the external surface of a scene (i.e., the profile visible with an optical system) should be considered, neglecting its internal structure [2].

Besides underwater applications, 2-D and 3-D acoustical imageings are used in medical applications and nondestructive testing (NDT); moreover, microwave imaging presents many analogies to acoustical imaging. These applications deal with objects like off-shore petrol structures, shipwrecks, and archaeological objects in underwater imaging, body organs and possible internal lesions in medical imaging, material defects in NDT, and military or civilian targets (e.g., ships, tanks, or aircraft, with their sizes, shapes, and orientations) in microwave imaging. Also in these cases, object detection and recognition are performed by expert operators who try to understand a scenario on the basis of a priori knowledge (related to the sensor and to the scene considered) and experience.

Manuscript received May 3, 1996; revised April 2, 1997. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Hua Lee.

V. Murino is with the Dipartimento Scientifico e Tecnologico, University of Verona, I-37134 Verona, Italy.

A. Trucco is with the Department of Biophysical and Electronic Engineering, University of Genoa, I-16145 Genoa, Italy.

Publisher Item Identifier S 1057-7149(99)00927-6.
In this paper, we consider acoustic imaging processes utilizing the BF technique for image formation. In short, BF deals with the processing of signals acquired by an array of sensors aimed at enhancing signals coming from a fixed steering direction and attenuating, as much as possible, signals coming from other directions. Therefore, the output signal of a BF process is a function of time, called beam signal. In our case, BF is used for imaging purposes and the general procedure adopted for the imaging process lies in insonifying a scene by means of one or more suitable acoustic pulses, in acquiring the backscattered echoes by using an array of sensors, in applying beamforming to the acquired signals, and, finally, in extracting from the beam signals the information necessary for the formation of an image [1]–[5], [7], [10]–[17]. More specifically, after calculating the beam signal in a fixed direction, it is possible to determine the information about the scene (e.g., the distance) in the steering direction, thanks to the spatial selectivity of BF. By repeating this operation for various directions, one can obtain the image of the region of interest.

Unfortunately, acoustic images are affected by severe problems that make it difficult, as compared with optical images, to extract useful information about an observed scenario as well as to understand it. The poor directivity of the used antennas (i.e., main lobe width) produces images with rough resolution, and the nonideal rejection of the signals coming from directions different from the steering one (i.e., presence of side lobes) generates artifacts and a strong blurring effects. Finally, speckle noise (generated by the coherent interference of the backscattered signals) affects such images, giving rise to an inhomogeneous object representation. For these reasons, several postprocessing techniques [2], [18]–[21] are used to restore and enhance computed images, so improving scene understanding. All the above-mentioned applications require that these techniques should be computationally economical in order to satisfy real-time needs. This can be achieved either by using special hardware devices or, more often, by devising simple methods for image improvement (like threshold-based methods). However, due to either the use of high-cost hardware or the simple on-line methods available for image improvement, the overall cost of accomplishing an effective and useful processing of an image is considerably high.

In this paper, we present a new approach to enhancing the formation of acoustic images by improving the information extraction from beam-signal envelopes leading to high-quality (2-D or 3-D) acoustic images. The method proposed and the related results refer to short-range underwater imaging, but the approach can be straightforwardly generalized to other applications (e.g., medical imaging). More precisely, our approach is based on a local comparison between the beam signal envelopes actually obtained and the expected ones, suitably defined on the basis of both the physics of the image formation process and the a priori modeling of a scene. To this end, a reliability measure, called confidence, has been introduced that allows a rapid examination of beam signals, aimed at accurately detecting possible target echoes backscattered from a scene, while disregarding useless information. The confidence measure can be considered as an ideal interface between beam signals and the related image; it operates directly (on line) on the image-formation process, with only a slight increase in the computational load. The confidence-based approach yields better results than other similar methods used in this field. In particular, we describe how the most widely used technique for echo detection (i.e., the matched filter) can be re-enforced by the subsequent use of the confidence-based method. As a result, images of improved quality can be obtained.

The objectives that can be attained by the proposed approach are as follows:

1) reduction in artifacts due to the lowering of the side-lobe level;
2) better lateral resolution;
3) precise distance determination;
4) accurate detection (for each beam signal) of target echoes, associated with an estimation of the reliability of the information extracted, and, consequently,
5) a higher image quality that allows a better image recognition and a better understanding.

These results are obtained by the application of a physical model-based approach operating on line on the set of beam signals. The enhancement carried out by this method can be considered as a fast restoration method to be applied at an early stage to one-dimensional (1-D) signals to directly generate improved 2-D and 3-D images, instead of performing demanding postprocessing algorithms. In our opinion, these results and the formulation of the method itself constitute the main original points of this paper.

This paper is organized as follows. In Section II, we outline the characteristics of the most widely used methods for extracting information from beam signals. Section III describes the proposed method, its theoretical principles and practical fea-
tutes. Its performances were quantitatively assessed by means of the beam power pattern computation and a comparative analysis was also made. In Section IV, the results (on 2-D and 3-D images) obtained by using both real and synthetic data are reported. A further comparison is made with the results yielded by a postprocessing restoration algorithm. Finally, conclusions are drawn in Section V.

II. IMAGE FORMATION AND RELIABILITY EVALUATION

In this section, we analyze first the methods based on the beamforming algorithm for the generation of 2-D and 3-D images, then some techniques to evaluate the reliability of the information contained in them. The methods considered are the most widely used by acoustic imaging systems and meet simplicity and speed requirements. Other problems (e.g., time delay estimation, source localization, etc.) can be solved by more complex and more computationally onerous methods that are difficult to apply to 2-D and 3-D imaging. In order to make a comprehensive comparison, we consider methods for underwater acoustic imaging [1], [2], [10]–[14], [22], [23], for medical acoustic imaging [4]–[6], [15], for nondestructive acoustic testing (NDT) [15]–[17], and for microwave imaging [8].

A. Beamforming and Visualization

Beamforming (BF) is a technique widely used in signal processing: its goal is to estimate signals coming from a fixed direction (steering direction), while attenuating those coming from other directions [3]. In essence, it is a spatial filter that linearly combines the temporal signals spatially sampled by a discrete antenna, i.e., an array of sensors located according to a known geometry. One considers an array of M punctiform and omnidirectional sensors constituting the receiving device, and assumes that an acoustic pulse composed of plane waves impinge on the array. (The far-field hypothesis is not a strict requirement, as, under near-field conditions, one can take into account the wave curvature by applying a focused BF [3], [24].) Denoting by \( \mathbf{u}_i \) the position of the \( i \)th sensor, by \( \mathbf{r} \) the versor normal to the wave front and directed toward the wave source, by \( c \) the sound velocity, and by \( \omega \) the angular frequency, the signal \( x_i(t;\mathbf{r}) \) received by the \( i \)th sensor can be expressed (after a base-band demodulation that yields a complex output signal [3]) as

\[
x_i(t;\mathbf{r}) = A(t + \mathbf{u}_i \cdot \mathbf{r} / c) \cdot \exp\left(j\omega \mathbf{u}_i \cdot \mathbf{r} / c \right) \tag{1}
\]

where \( A(t) \) is the temporal envelope of the acoustic pulse.

According to the definition of quadrature beamforming [3], a beam signal steered in the \( \mathbf{u} \) direction (\( \mathbf{u} \) being a versor), \( b_{\mathbf{u}}(t;\mathbf{r}) \), is defined as

\[
b_{\mathbf{u}}(t;\mathbf{r}) = \sum_{i=0}^{M-1} w_i \cdot x_i(t - \tau_i;\mathbf{r}) \cdot \exp(-j\omega \tau_i) \tag{2}
\]

where \( M \) is the number of sensors, \( w_i \) are the weights assigned to each sensor, and \( \tau_i = (\mathbf{u}_i \cdot \mathbf{u})/c \). The net result is that backscattered echoes coming from \( \mathbf{u} \) are emphasized, whereas those coming from other directions are attenuated, as one can deduce by substituting (1) into (2)

\[
b_{\mathbf{u}}(t;\mathbf{r}) = \sum_{i=0}^{M-1} w_i \cdot A(t + \mathbf{u}_i \cdot (\mathbf{r} - \mathbf{u}) / c) \cdot \exp\left(j\omega \mathbf{u}_i \cdot (\mathbf{r} - \mathbf{u}) / c \right). \tag{3}
\]

In general, after fixing the steering direction, the beam pattern is a function that yields the attenuation with which the beamforming system transfers a plane wave to the beam signal, for each possible incidence angle [3]. When the received echoes are not continuous waves but pulses, the beam pattern depends on the envelope shape and duration. In this condition, the beam pattern can be defined as the maximum of the beam signal over time [25], [26]:

\[
BP_{\mathbf{u}}(\mathbf{r}) = \max_t \{|b_{\mathbf{u}}(t;\mathbf{r})|\} \tag{4}
\]

or on the basis of the total energy of the beam signal [25], [26]:

\[
BP_{\mathbf{u}}(\mathbf{r}) = \left( \int_{-\infty}^{\infty} |b_{\mathbf{u}}(t;\mathbf{r})|^2 \, dt \right)^{1/2}. \tag{5}
\]

As suggested in [4] and [26] for imaging applications, the definition of the beam pattern presented in (4) is more suitable than that presented in (5); therefore, in this paper, the beam pattern will be defined as the maximum of the beam signal over time. The beam pattern of an imaging system is very important, as it allows (like the point spread function) an objective evaluation of the system performances. A conventional beam pattern presents a main lobe in the direction in which the array is steered and side lobes of minor, but not negligible, magnitudes in other directions. The width of the main lobe corresponds to the lateral resolution of the imaging system, whereas the generation of artifacts degrading useful information depends on the level of the side lobes. The side-lobe level can be reduced by suitably fixing the values of the coefficients \( w_i \); but, unfortunately, this operation increases the width of the main lobe, thus worsening the lateral resolution of the system.

Images are generated by collecting the beam signals steered in each direction of interest (according to the lateral resolution of the system) and by displaying the information contained in them. Obviously, the acoustic field backscattered by a real scene is not a plane wave but the wave front is a corrugated surface. Therefore, the signal received by the \( i \)th sensor, \( x_i(t) \), cannot be related to a precise direction of arrival \( \mathbf{r} \) any more. Nevertheless, the beam signal steered in the \( \mathbf{u} \) direction, \( b_{\mathbf{u}}(t) \), can always be computed by using (2). Two-dimensional images can be formed by linear arrays, whereas 2-D arrays allow the formation of acoustical 3-D images. In 2-D images, each beam signal contributes to forming a whole image line; consequently, the most widely used imaging method is the direct one: it consists in converting a whole beam-signal envelope into the pixels forming an image line (see Fig. 1). A color or a gray level proportional to the amplitude of the beam-signal envelope at a fixed instant is assigned to the related pixels [1], [4], [5], [7], [11]–[14], [16], [18].
By using suitable palettes—look-up tables (LUT’s)—one can achieve results that are sufficiently understandable. In order to obtain beam signals for which the echo amplitudes do not depend on time positions, received signals are amplified with a gain that increases over time—time-varying gain (TVG)—and is able to compensate for the loss due to the propagation spreading and the medium attenuation [17], [22], [28]. To improve image quality, some authors deem it convenient to visualize the power of a beam signal envelope, rather than its amplitude. This is accomplished on a linear scale [10] or on a logarithmic scale [8], [12]. An interesting comparison of various visualization methods is made in [20].

The aforesaid methods build an image by starting directly from a beam-signal envelope, possibly amplified with a TVG. A more advanced method, well-suited for low signal-to-noise ratio (SNR) conditions (as frequently occurring in underwater acoustic imaging) consists in applying a matched filter to the beamformer output [14], [22], adopting as an impulse response the pulse used to insonify the scene. Then, an envelope detector is applied, and the resulting signal (i.e., a local energy) is utilized to create an image line. Visualization can be performed on a linear scale [22] or on a logarithmic scale [14]. This method aims at determining the temporal positions of echoes (thus limiting the effect of noise) and at creating energy peaks corresponding to such positions. In the case of a pulse in white Gaussian noise, the matched-filter receiver has proved to be the optimal filter in terms of SNR maximization. Moreover, if the insonification pulse has a time-bandwidth product greater than unity, the coherent gain of the matched filter increases, thus ensuring better performances in very noisy conditions.

In acoustic 3-D imaging, one must extract from each beam signal the distance related to the first reflection. A tern \((x, y, z)\) can be calculated in accordance with this distance and with the steering direction: it represents a point in a 3-D map (see Fig. 2). An analysis of each beam signal is required. In most cases, the selected distance to be represented in an image is the one related to the temporal instant at which the beam signal envelope (possibly, after the matched filtering) reaches an absolute maximum amplitude [2], [4], [13]. As mentioned earlier, to represent the absolute maximum amplitude [2], [4] or a logarithmic measure of the power of such a maximum value [4], colors or gray levels can be used. Of course, this simple procedure does not ensure that the absolute maximum amplitude will be related to the reflection of an object (if any) in the steering direction [2]. The probability of detecting a spurious reflection (due to side lobes) instead of the reflection of an object in the steering direction is not negligible.

In conventional 2-D and 3-D imaging systems, the adopted insonification pulse is often an amplitude-modulated carrier of constant frequency. In the case of a rectangular pulse that lasts for a time equal to \(T\), it is easy to understand that, without the matched filter, the accuracy of the distance determination is equal to \(\pm c \cdot T/2\). To decrease the error of the distance measurement down to zero (in the ideal case), it is necessary to insert a matched filter before the envelope extraction. This improvement is obtained through a local integration performed by the matched filter along the beam signal.

One important problem related to methods based on the amplitude of the beam signal envelope is the sensitivity to the side lobes present in the beam pattern. The practical consequences of such sensitivity are degrading effects consisting in halos visible around the objects contained in an image [1], [7], [8], [15] where no physical obstacles are actually present (i.e., artifacts). The use of a suitable threshold can aid in attenuating such halos [2], [4], [8], [15], [16]. In 2-D imaging, the profile of a beam signal is represented in an image only at time instants when the envelope amplitude exceeds a suitable threshold. In 3-D imaging, each tern \((x, y, z)\) is represented on the 3-D map only if the absolute maximum amplitude of the related beam signal envelope exceeds a suitable threshold. When the value of the maximum amplitude does not exceed the threshold, one can assume the absence of obstacles in the steering direction considered. Unfortunately, it is very difficult to find a suitable value for the threshold, due to the differences in the amplitudes of the echoes generated by different objects, or by different parts of the same object differently located. Reflections, generated by the internal layers of a scene (in 2-D imaging), show marked differences in amplitude, too. Therefore, in order not to lose some details of the targets contained in a scene, one must fix a rather low threshold, resulting in an unsatisfactory reduction in halos. Vice versa, if one fixes a higher threshold, one obtains an acceptable reduction in halos but loses some useful information about the scene.

### B. Reliability Evaluation

In many cases, to evaluate the reliability of the spatial information extracted from each beam signal, the SNR relative to the envelope of the sampled beam signal, \(\text{SNR}_\text{U} = 10 \cdot \log \left[ \frac{1}{K} \sum_{n=1}^{K} (b_\text{U}(n) - \bar{b}_\text{U})^2 \right] \) (possibly after a matched filtering), is considered [2]. Such a parameter is computed through the ratio between the square of the estimate of the mean value \(\bar{b}_\text{U}\) and the variance estimate, as follows:

\[
\bar{b}_\text{U} = \frac{1}{K} \sum_{n=1}^{K} b_\text{U}(n) 
\]

\[
\text{SNR}_\text{U} = 10 \cdot \log \left[ \frac{1}{K} \sum_{n=1}^{K} (b_\text{U}(n) - \bar{b}_\text{U})^2 \right] \]  

where the constant \(K\) is the total number of time samples contained in \(b_\text{U}(n)\). The limitations of such a reliability estimate lie in the too general significance of the SNR, which is an indicator of the overall quality of the imaging process but it does not take into account how specific information is extracted from beam signals.

To improve reliability estimation, only in the case of 3-D images built by starting from the absolute maximum amplitude of each beam signal (denoted by \(b_\text{U}(\hat{n}) = \max_i \{b_\text{U}(n)\}\)),
The calculated parameter, \( \text{SNR}_{pk} \), is called the SNR of the peak. This method aims to assign high reliability values to those beam signals whose absolute maximum amplitudes are quite above the mean value, and that show overall profiles with small variations around the mean value [2]. The limitations of this statistical method lie in its poor ability to discriminate between beam signals with very different real contents. For instance, if the envelope of a beam signal shows a peak at a given time instant and is flat at other instants (i.e., it consists of a sequence of samples that are all equal, except for the peak), the \( \text{SNR}_{pk} \) value is not dependent on the peak value (as would be reasonable) but it is a function of the total number of time samples making up the signal, which results in an undesired characteristic.

III. THE CONFIDENCE-BASED APPROACH

A novel approach based on confidence measures has been developed that allows one to determine the correct distance of an obstacle (or of the various layers of a scene) by estimating the reliability with which, at every sampling instant, the beam signal examined can be compared with a possible (model-based computed) echo of the transmitted pulse. Unlike the methods described in the previous section, the proposed approach makes it possible to determine the distance and, at the same time, to estimate the reliability by a single mechanism. To carry out such operations, the developed confidence-based method requires the calculation of the mean square distance between the beam signal envelope at the beamformer output and an expected beam signal based on the modeling of the scene. The concept of an expected signal has recently been developed, even though for different purposes, in the medical and NDT fields [5], [16].

A. Estimation of Expected Beam Signals

For given structures of the transmitter, of the scene, and of the receiving array, a beam signal is defined as the signal that a single sensor, placed in the coordinate origin, would receive if its reception pattern were identical with the beam pattern generated by the array. Therefore, due to the side lobes in the beam pattern, an \emph{a priori} definition of the expected envelope requires the knowledge of the shape and of the structure of the whole scene. As such a knowledge is not available, it is necessary to resort to some approximations.

A first suitable approximation is to assume the beam pattern to be rectangular in shape, flat at 0 dB in the area within the width of the main lobe (measured on the real beam pattern by using the points at \(-3\) dB) and null outside this area. Therefore, the only scene portion that contributes to the beam signal is the small area within the width of the main lobe. As a result, it is not necessary any more to know the whole structure of the scene: thanks to this approximation, the scene portion not covered by the main lobe gives no contribution. The second approximation is to assume the material of the scene to be diffusive so that the expected pulse can be considered independent of the steering direction of the beam signal. For low steering angles, the width of the main lobe does not change very much, and backscattering is nearly independent of the incidence angle between the beam signal and the scene [14], [29], [30]. Under these conditions, the surface of a small area depends only on the distance between the array and the scene; consequently, the expected envelope depends only on this distance and the material in the small area. The latter approximation is not too restrictive as, typically, the wavelengths used in short-range underwater imaging are short with respect to the roughness of recurrent objects (i.e., natural or encrusted man-made objects); therefore, a diffusive behavior is often a reasonable assumption.

With reference to Fig. 3, let us denote by \( A(n) \) the sampled envelope of the acoustic pulse transmitted, and by \( b_{\text{exp}}(n, m) \) the sampled envelope of the beam signal expected, where \( m \) is the arrival time corresponding to the distance of the
scene from the array. One can see that the amplitude of the expected envelope decreases and its deformation increases as the distance increases or, equivalently, as the arrival time $m$ increases. However, for not too different distances, the shape and the amplitude of the expected envelope vary only slightly, hence, for the sake of simplicity, we assume a constant shape and a varying amplitude. On the basis of all the above considerations, one can easily make a more or less refined collection of expected envelopes, according to the distance of a possible target, to the material the target is made of, and, if necessary, to the characteristics of the propagation medium.

B. Direct Comparison

The first confidence measure proposed, named $CL1$ (i.e., confidence level one), is used to make a direct comparison between the envelopes of two beam signals, i.e., the expected beam signal $b_{\text{exp}}(n,m)$ and the real one $b_\mathbf{u}(n)$, for each direction of interest $\mathbf{u}$. The comparison is made in terms of the quadratic distance, as follows:

$$CL1_{\mathbf{u}}(m) = 10 \cdot \log \left[ \frac{\sum_{n=m-N}^{m+N} b_\mathbf{u}(n)^2}{\sum_{n=m-N}^{m+N} [b_\mathbf{u}(n) - b_{\text{exp}}(n,m)]^2} \right]$$

where $2N+1$ is the number of sampling instants containing the expected envelope. If we let $m$ assume all possible sampling instants for a beam signal, we can evaluate, at each instant, the similarity between the envelopes of the real beam signal and those of the expected one. The amplitude and the shape of the expected beam signal may vary just as a function of $m$. The larger the value of $CL1_{\mathbf{u}}(m)$, the closer the similarity of the expected envelope to the envelope present at about the instant $m$. On the basis of the temporal profile of $CL1_{\mathbf{u}}(m)$, it is possible to build an acoustic image in a very advantageous way, as will be shown in the following.

In (9), $CL1_{\mathbf{u}}$ makes it possible to compare the energy related to the difference between the two beam signals with the energy of the real beam signal in order that the values assumed by the confidence level may not depend on the energy of the transmitted signal, nor on the distance of the scene from the array. In terms of energy, the confidence approximates a local SNR, where the signal is represented by the real envelope and noise is represented by the difference between the expected and real envelopes. The definition of an SNR measure that is not based on the statistical characteristics of the beam signal [as (7) and (8) are] is made possible by the use of the model-based signal concept. Therefore, we can regard, instant by instant, the value of the confidence level as being equal to the SNR computed over a time window that is as long as the insonification pulse.

The defined confidence measure requires the $a$ priori knowledge (even though approximate) of the amplitude of the pulse that is expected to be present in a beam signal (expected envelope), possibly after the matched filtering operation. However, such a knowledge is not always characterized by a sufficient accuracy [8], mainly due to the effects of speckle and scattering peculiarities. For instance, in 2-D imaging, the presence of successive reflections makes it difficult to forecast the amplitudes of the various echoes. The same problem is faced in 3-D imaging, too, when the material of a scene object is not known, or when scene objects have nonnegligible reflecting behaviors. As a consequence, a direct comparison turns out not to be flexible enough to meet some practical requirements.

C. Indirect Comparison

To overcome the problem of an inaccurate knowledge of expected amplitudes, the previous confidence level has been modified by separating the information about the envelope shape from that about the envelope amplitude. To this end, both the envelope of the expected beam signal and the envelope of the real signal must be split into the product of two terms: an amplitude coefficient $a$ and a shape function $s$ with a normalized amplitude. The adopted notation is the following:

$$b_{\text{exp}}(n,m) = a_{\text{exp}}(m) \cdot s_{\text{exp}}(n,m)$$

$$s_{\text{exp}}(m,m) = 1$$

(10)

where $a_{\text{exp}}(m) = b_{\text{exp}}(m,m)$ is the expected amplitude of the reflected pulse, coming from a distance corresponding to the instant $m$. Analogously, the envelope of the real beam signal is normalized to the value assumed at the sampling instant $m$:

$$b_\mathbf{u}(n) = a_\mathbf{u}(m) \cdot s_\mathbf{u}(n,m)$$

$$s_\mathbf{u}(m,m) = 1$$

(11)

where $a_\mathbf{u}(m) = b_\mathbf{u}(m)$ is the amplitude of the envelope of the beam signal at the instant $m$.

Therefore, the new confidence level, $CL2_{\mathbf{u}}$, can be defined as

$$CL2_{\mathbf{u}}(m) = K_1 \cdot 10 \cdot \log \left[ \frac{\sum_{n=m-N}^{m+N} s_\mathbf{u}(n,m)^2}{\sum_{n=m-N}^{m+N} [s_\mathbf{u}(n,m) - s_{\text{exp}}(n,m)]^2} \right] + K_2 \cdot 10 \cdot \log \left[ \frac{(a_\mathbf{u}(m))^2}{(a_\mathbf{u}(m) - a_{\text{exp}}(m))^2} \right]$$

(12)

where $K_1$ and $K_2$ are two constants, real and positive, that rule the degrees of importance that one would like to lend to the comparison between the envelope shapes and to the comparison between the envelope amplitudes. The former is a local comparison and the latter is a point-wise comparison; both are made through the square distance. The larger the value of $K_1$, the larger the decrease in the confidence level caused by the difference between the expected and real shapes. The same holds true for $K_2$, with regard to the amplitudes. By suitably tuning the two constants, $K_1$ and $K_2$, one can obtain an acceptable trade-off between amplitude and shape requirements. Strategies for the regulation of $K_1$ and $K_2$ can be devised according to the $a$ priori knowledge of the scene (i.e., target objects) and of the medium properties.
D. Performance Evaluation

Performances of confidence levels are difficult to evaluate in an analytical way as they depend on the goodness of the simulation performed, i.e., on the similarity of the expected echoes to the received ones. To face this problem, performances can be assessed for different degrees of similarity by varying the SNR. In other words, starting from a perfect simulation (i.e., the expected echo is equal to the real echo), the expected echo is gradually degraded by lowering the SNR to emulate realistic situations.

As mentioned in Section II, in underwater acoustics the beam pattern of an imaging system is recognized to be a reliable measure of performance. Consequently, in this subsection, the computation and the comparison of the beam patterns before and after applying the confidence-level will be presented for realistic working conditions.

It was assumed that a plane wave incoming from the direction \( \mathbf{r} \) (with a frequency equal to 1 MHz, a duration equal to 30 \( \mu s \), and a Gaussian-shaped envelope) impinged on a linear array composed of 64 \( \lambda/2 \)-equispaced sensors. As usual, the reasoning was restricted to a plane where the steering direction \( \mathbf{u} \) was indicated by the angle \( \rho \) and the arrival direction \( \mathbf{r} \) was indicated by the angle \( \theta \). After fixing \( \rho = 0^\circ \) (the direction perpendicular to the array baseline), the beam signals \( bs_p(n; \theta) = bs_u(n; \mathbf{r}) \) were computed for any possible angle \( \theta \) and the beam pattern was evaluated by using (4), rewritten as follows:

\[
BP_p(\theta) = \max_n \{ |bs_p(n; \theta)| \}.
\] (13)

The conventional beam pattern (after transformation into a logarithmic scale and normalization) is shown in Fig. 4(a): it follows the theory of wide-band beamforming [26].

An expected beam envelope was defined equal to the envelope of the beam signal obtained when the steering and the arrival directions coincide (\( \theta = \rho \), i.e., the condition in which the confidence level should be the highest). Then, the expected envelope was corrupted by an additive Gaussian noise of tunable power and the goodness-of-fit level was measured by using the related SNR. In Fig. 5(a), for instance, one can see the expected envelope and the real envelope (of the beam steered in the arrival direction) when the SNR is equal to 20 dB.
Following (12) and after fixing $K_1 = K_2 = 1$, the profile of the confidence $CL2$ was computed for each beam signal, $\delta_s(n;\theta)$, of the above-mentioned collection. In this way, the confidence was dependent on the arrival angle like the beam signal; therefore, we denoted it by $CL2_p(n;\theta)$. The beam pattern after the application of the confidence was computed by using a new version of (4) in which the profile of the confidence replaced the envelope of the beam signal:

$$BP_p(\theta) = \max_n\{CL2_p(n;\theta)\}. \quad (14)$$

It is important to note that this equation yields values that are already on a logarithmic scale.

The normalized beam power pattern was evaluated when the SNR of the expected envelope was equal to 20 dB [see Fig. 5(a) and 4(b)], 10 dB [see Fig. 5(b) and 4(c)], and 5 dB [see Fig. 5(c) and 4(d)]. The beam pattern in Fig. 4(b) shows a dramatic decrease in the side-lobe level together with a sharp narrowing of the main lobe, as compared with the conventional beam pattern in Fig. 4(a). However, the close agreement between the expected and real envelopes could be judged too optimistic for real conditions. If the SNR is reduced to 10 dB (more realistic condition) and 5 dB (pessimistic condition), the advantages are always present but their importance is proportionally decreased. In any case, a comparison of Fig. 4(d) with Fig. 4(a) shows that at an SNR $\approx 5$ dB the side-lobe reduction is greater than 10 dB and the main lobe width is less than a half, despite the poor similarity between the real and expected envelopes [see Fig. 5(c)].

To complete the analysis, it is necessary to study what happens when a great difference in shape between the expected and real envelopes is associated with a wrong amplitude prediction. For instance, we computed the beam pattern when the expected envelope had an amplitude that was twice the real one, as shown in Fig. 6(a). We fixed $K_1 = 2$ and $K_2 = 1$ in order to stress the shape comparison, and we obtained the result shown in Fig. 6(b). The shape comparison is very important, independently of the amplitude comparison. When the wave has an arrival angle different from the steering one, the pulse in the beam signal has a shape that is sharply enlarged, as compared with the expected one. To point out this effect, Fig. 7(a) shows the pulse envelope when the beam signal is steered in the arrival direction, whereas Fig. 7(b) shows the pulse envelope when the beam signal is steered 15° from the arrival angle. Thanks to this effect, the tuning of the two constants allowed us to achieve a good result despite a wrong amplitude prediction: the obtained beam pattern is very similar to those in Fig. 4(c), computed at an SNR of 10 dB.

From the above consideration, one can deduce that the confidence level could be a flexible way to improve the beam pattern of an imaging system when a knowledge (even though if approximate) of the backscattered pulse envelope is available. The beam patterns showed that an improvement in the lateral resolution and a decrease in the interfering signals can be achieved.

### E. Comparative Analysis

An additional way to evaluate the performances of confidence levels is to compare the output signals from different echo-detection systems when the input signal is the same, i.e., a given beam signal in which a couple of replicas of the insonification pulse is present. Fig. 8 shows the plot of such signals: $f(n)$ is a basic signal containing two ideal replicas (of amplitudes 1 and 0.5, respectively) of a rectangular pulse composed of 5 carrier cycles. A realistic beam signal was obtained by adding Gaussian noise to $f(n)$ at an SNR of 0 dB: $z(n)$ is the output signal from a common envelope detector preceded by a matched filter; $CL1(n)$ is the signal obtained...
Fig. 6. (a) Example of the expected (solid line) and received (dotted line) echo envelopes when an inaccurate amplitude estimation is performed. (b) The related beam power pattern computed by applying the CL2 confidence measure.

Fig. 7. Envelope of a beam signal when the beam is steered: (a) in the arrival direction; (b) 15° from the arrival angle.

by applying the first confidence level (i.e., direct comparison) to the signal \( x(n) \); finally, \( CL2(n) \) is the signal obtained by fixing \( K_1 = 2 \) and \( K_2 = 0.5 \) and applying the second confidence level (i.e., indirect comparison) to the signal \( x(n) \). For both confidence levels, the expected envelope was defined

as a triangle of unitary amplitude (i.e., \( a_{\text{exp}} = 1 \)) and appropriate duration.

By comparing the output signals of the three echo detection systems that were aimed at processing the beam signal in Fig. 8, it is possible to point out some considerations.

- Owing to the coherent gain of the matched filter, the envelope \( x(n) \) shows a peak for each echo, despite the low SNR. Unfortunately, the behavior of the signal around the peak is not very steep and the peak height is always related directly to the echo amplitude.
- The confidence-based processing of the signal \( x(n) \) by the direct comparison produces a very steep peak when the amplitude of the triangle is close to the expected one, and
completely rejects the echo having an amplitude that is half the expected one. Sometimes, when the amplitudes of the expected echoes can be computed with high precision, this behavior can be very useful to discard echoes due to side lobes.

- The confidence-based processing of the signal $x(n)$ by the indirect comparison produces two steep peaks with a sufficient independence of the echo amplitude. These two peaks are not so steep as in the direct-comparison case, but, from an imaging point of view, the overall behavior of the signal $C(x(n))$ is much better than that of the signal $x(n)$. Moreover, the dependence of the peak on the echo amplitude can be regulated by tuning the value of the coefficient $K_2$, thus allowing the exploitation of this method even when the a priori amplitude information is not so precise, as very often occurs.

- As both the confidence levels and the matched filter try to locate the exact position of a reflected pulse independently of the duration $T$ of the insonification pulse, an accurate determination of the distance does not depend any more on the duration of the transmitted pulse, but it tends to approach the ideal case $\pm c \cdot T_s/2$, where $T_s$ is the sampling period [2].

IV. RESULTS

When the confidence approach is applied, an image is not built any more by starting from the beam-signal envelope but from the profile of the selected confidence level over time. This results in an image of higher quality in which target objects can be reliably detected and the related range information is more accurate, so allowing one to avoid the use (or improve the performances) of restoration/reconstruction algorithms (e.g., [31], [32]) to be applied before scene understanding.

Thanks to the flexibility of the indirect comparison, the peaks in the profile of the confidence level are narrower in time and of more uniform amplitude (see Fig. 8). Therefore, in 2-D imaging, it is not necessary to fix a threshold and to visualize only the points at which the confidence level exceeds the threshold. In 3-D imaging, in order to obtain the external surface of a scene, only the point closest to the observer, among the points at which the confidence level exceeds an attention threshold, can be visualized for each beam signal.

This section is divided into three subsections: the first considers 2-D imaging starting from real data and the second deals with 3-D imaging starting from simulated data. In these experiments, the confidence measures were computed without the use of a matched filtering, as the SNR’s of the received signals were quite high. The third subsection concerns the load of the confidence-level computation.

A. 2-D Real Images

To verify the actual efficiency of the proposed confidence measures, a real experiment was developed and performed in the laboratories of the O. M. Corbino Institute of Acoustics, National Research Council, Rome, Italy. The array was obtained by the synthetic-aperture technique by translating a single transducer, along a straight line, to 250 different 1.5 mm-spaced positions. At each step, the transducer emitted a fixed acoustic pulse and received the pulse echoes coming from a scene. The scene consisted of a set of parallel nylon strings stretched between two plane surfaces and arranged in such a way that the perpendicular section formed the word “SEA,” as shown in Fig. 9. Thanks to the large number of strings used, more than ten successive reflections occurred in some beam signals, due to the strings placed in the steering direction, or to those present in the directions of the side lobes. As can be deduced, it was a severe testbed to assess the performances of the confidence-based approach.

The values of the parameters of the array and the scene configuration were the following: sound velocity in the medium: 1500 m/s (assumed constant); carrier frequency: 460 kHz (wavelength $\lambda = 3.26$ mm); rectangular pulse envelope; pulse duration $T = 15$ $\mu$s; sampling frequency in quadrature [3], [33] equal to 1.024 MHz; number of acquired time samples.
equal to 202; diameter of nylon strings equal to 0.6 mm; scene distance from the array: between 52 and 64 cm (i.e., 12 cm-high characters), width of the word “SEA” equal to 35 cm.

The expected envelope was computed by assuming that the insonification pulse was backscattered from a single target similar to a nylon string, and the echo was received by one omnidirectional sensor placed in the coordinate origin. The scattering was simulated by the method described in [29]. A physics-based model of the expected target was built under the hypothesis that it was not a string but a particle whose dimensions were very small in comparison with λ, and that it had the same radius, density, and bulk modulus as an actual string. This model permitted the use of the scattering transfer function as defined in [34]. Moreover, the simulation was not repeated at all the time instants, but the beam was partitioned into about ten adjacent time intervals and an expected envelope was computed for each interval.

To process the acquired signals, we used the beamforming in quadrature in the time domain [3],[33], with a dynamic focusing [27] and unit weight coefficients. Images were obtained by means of parallel beams perpendicular to the array [23], by simply shifting the application point for each beam (i.e., the coordinate origin), as shown in Fig. 9. To calculate each beam, only the signals received from 80 sensors were utilized: 40 to the right and 40 to the left of the application point for the beam.

Fig. 10 presents the images obtained by the most widely used methods of visualization (described in Section II). Each image column was formed starting from the behavior of a given beam signal: the output values provided by the method considered were assigned to the image point by point. The visualized area represented a 19 cm × 38.4 cm rectangle, 50 cm far from the array. The number of formed beam signals was equal to 250; the application points for the beam signals corresponded to the positions of the sensors in the synthetic array. Fig. 10(a) was obtained by using the amplitudes of the beam-signal envelopes [23], compensated by TVG’s. Fig. 10(b) was obtained according to the behaviors of the compensated square envelopes (envelope power). Fig. 10(c) shows, point by point, the value of the power represented on a logarithmic

---

**Fig. 10.** Two-dimensional real images of 65 nylon strings. The images were obtained by using: (a) the amplitudes of the envelopes of compensated beam signals; (b) the envelope power; (c) the logarithm of the power (measured in decibels); (d) the envelopes of the output signals of the matched filter.
scale in dB (denoted by \( q \)), calculated through the following transformation [18]:

\[
q = 10 \cdot \log \left( \frac{p - p_{\min}}{p_{\max} - p_{\min}} \right)
\]

(15)

where the actual power value is denoted by \( p \), and \( p_{\min} \) and \( p_{\max} \) are the minimum and maximum power values when all the beam signals are considered. At the points at which \( p = p_{\min} \), the value of \( q \) is indefinite; it is set equal to the minimum \( q \) value among the \( q \) values obtained at the other points. Fig. 10(d) was obtained by means of the envelopes representing the outputs of the matched filter, whose inputs were the beam signals compensated by TVG’s. Finally, Fig. 11 was obtained by using the outputs of the confidence method based on an indirect comparison (CL2), after fixing \( K_1 = 2 \) and \( K_2 = 1.5 \). In both Figs. 10 and 11, 256 gray levels were used, which were assigned according to the linear law

\[
gl = 255 \cdot \frac{\chi - \chi_{\min}}{\chi_{\max} - \chi_{\min}}
\]

(16)

where \( gl \) is the gray level of the pixel associated with the value \( \chi \), which represents, depending on the method used, the amplitude, the power, the power logarithm, the output of the matched filter, or the confidence level. The values \( \chi_{\min} \) and \( \chi_{\max} \) are the minimum and maximum values of \( \chi \) in a whole image.

The images obtained by the matched filter and the confidence-based method (both are local procedures) are very uniform, unlike those obtained by the described point-wise methods (based on amplitude, power, and the power logarithm). However, one can notice that, among the methods used, only the one based on confidence levels makes it possible to clearly separate the target points from the background by eliminating the “gray” background and sharply reducing the halos around the points. Moreover, in Fig. 11, the size of the points representing the nylon strings is smaller than that obtained by the other methods, including the matched filter. This filter improves the distance determination, but the slow slopes of a signal around peaks (see Fig. 8) limit the visual effectiveness of such determination, providing a lower visual quality than that obtained by the confidence-based method.

### B. 3-D Simulated Images

To verify the efficiency of the confidence measures in 3-D imaging, an experiment was performed by the simulation method described in [29], generalized to 3-D images [30]. The simulation required a square array placed on the plane \( z = 0 \) (Fig. 2) and made up of receiving sensors that acquired the echoes coming from a scene. The echoes were produced by an omnidirectional insonification performed by a single transmitter placed in the coordinate origin. The medium was assumed to be linear and homogeneous. To evaluate the robustness of the confidence-based method, white Gaussian noise was added to the signals received by the sensors; the noise was such as to obtain an SNR equal to 10 dB.

The configuration used for the experiment was characterized by the following parameters: sound velocity in the medium: 1477 m/s, carrier frequency: 500 kHz (\( \lambda = 2.95 \) mm), triangular acoustic-pulse envelope, pulse duration \( T = 200 \) µs, array made up of 50 × 50 1-cm-spaced sensors, sampling frequency in quadrature equal to 50 kHz. The scene consisted of three continuous surfaces:

1) a rectangle of dimensions 20 cm × 30 cm, parallel to the array and 5 m far from it [upper left portion of Fig. 12(a)];
2) a 20 cm × 20 cm square whose distance from the array plane \( z = 0 \) ranged from 3.5–3.62 m, according to the equation \( z = 3.5 + 0.6y \), where \( 0 \leq y \leq 0.2 \) [middle portion of Fig. 12(a)];
3) a spherical cap of 30 cm diameter whose distance from the array ranged between 6.50 and 6.57 m [lower right portion of Fig. 12(a)].

To process the signals acquired, we used a phase-shift BF implemented through the FFT, with a dynamic focusing and unit weight coefficients [24]. This type of beamforming is 2-D and beams have the same application point (i.e., the coordinate origin) but different steering directions.

Since, in this case, the targets represented continuous objects, the expected envelope was computed by assuming that the insonification pulse was backscattered from a surface parallel to the array, and the echo was received by one omnidirectional sensor. The dimensions of the simulated surface were equal to those of the region covered by the main-lobe width at the various distances (since the beam pattern was assumed to be rectangular, as described in Section III-A). By analogy to the simulation of the three objects in the scene, the small surface was simulated by a random dispersion of ideal punctiform scatterers at an average distance smaller than \( \lambda/4 \) [29]. The simulation was repeated for all the distances of interest, changing the surface dimensions and storing the resulting expected envelopes. In real cases, the surface should be modeled by using the available \( a \) priori information about the scattering characteristics of the material of the objects to be detected in a scene (e.g., target strength, etc.).

The images shown in Fig. 12 are the projections of the obtained 3-D maps on the array plane. The area of each
image is a square of 1.8 m side. In Fig. 12(a), the darkness represents the value of the absolute maximum amplitude (compensated for according to the distance) for each beam signal. In Fig. 12(b), the darkness indicates the value of the first significant maximum value of the behavior of the confidence level $CL_2$ (calculated for each beam signal) for $K_1 = 2$ and $K_2 = 1.5$. The transformation defined by (16) was applied to these two images, too, in order to use all the 256 gray levels available. Of course, from these images one can derive only the shapes of the objects, not their distances from the array. As can be noticed, the image obtained by the confidence levels is better defined and more homogeneous than that obtained by the maximum amplitudes, as the spurious halos were sharply reduced. If one tries to eliminate the halos from Fig. 12(a) by means of suitable thresholds, one loses, at the same time, many elements of the scene, thus degrading the shapes of the target objects and making the scene understanding more difficult.

Fig. 12. Projections of 3-D maps on a plane parallel to the array: (a) the darkness is proportional to the absolute maximum amplitudes of beam signals; (b) the darkness is proportional to the $CL_2$ confidence measure.

A postprocessing restoration algorithm can also be applied to recover original information. Fig. 13 shows the result of a restoration algorithm applied to the image in Fig. 12(a). This algorithm is based on the Markov random fields (MRF’s) approach [32]. A cost function to be minimized iteratively has been designed that makes it possible to embed criteria useful for image restoration. In this case, as typically occurs in restoration problems, smoothness and closeness to input observations were chosen as constraints imposed on the cost function. One can notice that the final result obtained by the MRF-based algorithm is quite similar to that obtained by the confidence-based technique [Fig. 12(b)], but the latter was achieved with a much lower computational load.

Figs. 14 and 15 present the 3-D elevation maps obtained, respectively, by the maximum absolute amplitudes of the beam-signal envelopes and by the confidence levels. In both cases, each beam signal results in a point $(x, y, z)$: if the value of the maximum amplitude or of the confidence level related to this point does not reach a given attention level (threshold), the point is translated to the background. To fix such a threshold, the values of the maximum amplitudes and of the confidence levels related to the whole image were normalized to a value between zero and one; then the value 0.1 was chosen as a threshold to degrade the overall information only slightly. The value on the $z$ axis indicates the height (in meters) with respect to the background, whose distance from the array center was equal to 6.76 m. By analogy to the previous images, in these pictures, too, the darkness indicates the value of the maximum amplitude (Fig. 14) and the value of the confidence level (Fig. 15). The quality and the definition of the map obtained by the confidence approach are higher than those of the map obtained by the absolute maximum amplitudes.

The confidence-based approach improves the quality of a map through an accurate echo detection, while avoiding subsequent applications of restoration and/or reconstruction algorithms. This result should be considered not only from a visual point of view (actually, in a map obtained by using only amplitude values, a human operator could detect the objects...
present in a scene and infer their shapes in some way, even though with more difficulty), but especially for the purpose of automatic scene recognition. Recognition systems have not yet reached such a level of development to face all possible real situations (in particular, in underwater environments), hence input images of good quality can facilitate the automatic recognition/understanding of an observed scene.

Thanks to the perfect knowledge of the scene simulated, one can evaluate the quality of the obtained 3-D maps by means of objective error criteria [8], [16]. To this end, we built an ideal 3-D map of the scene considered and used the mean square error (MSE) to assess the qualities of the maps shown in Figs. 14 and 15. If we denote by $z_i(x, y)$ the value of the $z$ coordinate at the point $(x, y)$ on the ideal map and by $z_m(x, y)$ the corresponding value on the map calculated by means of the maximum amplitudes or by the confidence level, we can define the MSE as

$$\text{MSE} = \frac{1}{R^2} \sum_x \sum_y [z_i(x, y) - z_m(x, y)]^2$$

where $R^2$ is the total number of possible pairs $(x, y)$. Moreover, we also evaluated the percent relative mean square error (RMSE), defined as

$$\text{RMSE} = \frac{100}{R^2} \sum_x \sum_y \frac{[z_i(x, y) - z_m(x, y)]^2}{z_i(x, y)^2}$$

The map built by means of the maximum amplitudes (Fig. 14) shows an MSE equal to 898 cm$^2$ and an RMSE equal to 0.1963%. The map built by means of the confidence level (Fig. 15) shows an MSE equal to 608 cm$^2$ and an RMSE...
equal to 0.1376%. One can notice that both errors decrease, thanks to the confidence level, by a factor equal to about 1.45.

C. Computational Remarks

The computational load required by confidence levels is not very high, as compared with that necessary to calculate a beam signal [2], [3], [33]. However, it can be further reduced by efficient implementation methods. For instance, one can try to limit the number of times the confidence value has to be evaluated by computing it only at those time instants in which the envelope is sharply greater than zero.

For both 2-D and 3-D images, the calculations of the confidence levels on the basis of indirect comparisons were performed at the end of the beamforming process. Such calculations increased the computational load by a percentage ranging from 5%–20%, as compared with the load required by a simple beamforming.

V. CONCLUSIONS

In this paper, we have proposed a confidence-based technique for image formation that requires a local comparison between each beam-signal envelope and the expected one, and that is suitably defined on the basis of the modeling of the scene observed and in accordance with the physics of the signals involved. It is a 1-D signal processing technique that can be applied at an early processing stage to improve the formation of 2-D or 3-D images.

The proposed confidence measure acts as an interface between output beamforming signals and an image. As a result, one can overcome the problems arising when acoustic imaging is performed directly on the basis of the point-wise behavior of beam signals. A direct consequence is that echo detection is performed directly on the basis of the point-wise behavior of beam signals. A direct consequence is that echo detection is performed directly on the basis of the confidence measures could be developed to estimate the resolution beamforming is characterized by a narrow depth of field [35], suboptimal automatic focusing algorithms using the confidence measures could be developed to estimate the correct focusing distance on the basis of the beam signal obtained.

ACKNOWLEDGMENT

The authors wish to thank Prof. G. B. Cannelli, E. D’Ottavi, L. Pitolli, and G. Pontucale of the O. M. Corbino Institute of Acoustics for kindly providing the real data used in this paper and for the useful advice concerning the handling of such data.

REFERENCES


**Vittorio Murino** (S’92–M’93) received the Laurea degree in electronic engineering and the Ph.D. degree in electronic engineering and computer science, both from the University of Genoa, Genoa, Italy, in 1989 and 1993, respectively.

He was a Post-Doctoral Fellow at the University of Genoa, working in the Signal Processing and Understanding Group, Department of Biophysical and Electronic Engineering, as Supervisor of Research Activities on signal and image processing in underwater environments. He worked on several national and European projects funded by the European Commission (EC), especially in the context of the Marine Science and Technology (MAST) program, concerned with the investigation of underwater scenes by visual and acoustical sensors. He is an Evaluator for the European Commission of project proposals related to the MAST III, Industrial and Material Technology (BRITE-EURAM III), and Long Term Research (LTR) programs. In 1995, he became an Assistant Professor with the Department of Mathematics and Computer Science, University of Udine, Italy, where he supervised research activities in an EC-funded project (VENICE, Brite-Euram III program) regarding multisensorial underwater vision for object recognition and virtual reality applications. He is currently an Associate Professor of image processing at the University of Verona, Italy. His main research interests include computer vision and pattern recognition, acoustic and optical underwater vision, data fusion, probabilistic techniques for image processing (Markov random fields, Bayesian networks), and neural networks. Recently, his research has focused on the integration of image analysis and synthesis methodologies for object recognition and virtual reality modeling.

Dr. Murino is a reviewer for several international journals and a member of IAPR.

**Andrea Trucco** (S’91) was born in Genoa, Italy, in 1970. He received the Laurea (M.Sc.) degree in electronic engineering from the and the Ph.D. degree in electronic engineering and computer science, both from the University of Genoa in June 1994 and June 1998, respectively.

He is responsible for the research activities in Acoustics and Underwater Signals, Signal Processing and Telecommunications Group, Department of Biophysical and Electronic Engineering, University of Genoa. His main research interests are array synthesis, coherent and noncoherent algorithms for acoustic imaging, acoustic image improvement, interferometric sonar and radar, and simulation methodologies. His is also involved in several research projects in the context of the Marine Science and Technology program of the European Commission.

Dr. Trucco was three times a finalist for the Philips Awards for European Young Scientists from 1987 to 1990, and ranked third twice. In 1995, he won the Student Paper Competition organized by the 9th International Symposium on Unmanned Untethered Submersible Technology. In 1997, he won the Student Paper Competition organized by the MTS/IEEE Oceans’97 International Conference. He is a member of IAPR.