Multiresolution 3-D Reconstruction From Side-Scan Sonar Images

Enrique Coiras, Yvan Petillot, *Member, IEEE*, and David M. Lane, *Member, IEEE*

Abstract—In this paper, a new method for the estimation of seabed elevation maps from side-scan sonar images is presented. The side-scan image formation process is represented by a Lambertian diffuse model, which is then inverted by a multiresolution optimization procedure inspired by expectation–maximization to account for the characteristics of the imaged seafloor region. On convergence of the model, approximations for seabed reflectivity, side-scan beam pattern, and seabed altitude are obtained. The performance of the system is evaluated against a real structure of known dimensions. Reconstruction results for images acquired by different sonar sensors are presented. Applications to augmented reality for the simulation of targets in sonar imagery are also discussed.

Index Terms—Augmented reality, expectation–maximization, multiresolution, shape from shading, side-scan sonar, 3-D reconstruction.

I. INTRODUCTION

ACOUSTIC sensing is the imaging modality of choice for the analysis of underwater environments. Acoustic waves propagate well in water—as opposed to electro-magnetic waves, which are quickly attenuated—and several acoustic sensors (*sonars*) are routinely used to study the underwater environment, with side-scan and multibeam [1] sonars being the most common. To obtain images of the seafloor, side-scan is the most widely used system, while multibeam is the preferred choice for the production of seabed elevation maps (*bathymetry*). Side-scan sonars provide high resolution images (typically 10×5 cm resolution) and reasonably large coverage (up to 70-m range) at high frequency (500 KHz). Emerging synthetic aperture sonar (SAS) technology will provide in the near future even better resolution ($\sim$15 mm) at longer ranges ($\sim$300 m). However, side-scan sonars remain, and should remain some years for economic reasons, the sensors of choice for seabed imagery. Although they do not provide direct measures of seabed elevation, the images they produce are related to seabed topography—as well as seafloor composition and direction of observation. Furthermore, the visual analysis of side-scan imagery suggests that it should be possible to establish a correspondence between image pixel intensities and the local surface slope at any given point of the seafloor. Of more interest to the image processing and computer graphics communities is the fact that the side-scan sonar imaging process is well approximated by the Lambertian image model [2]–[4], which permits the use of standard digital image processing techniques both for synthesis and analysis of sonar images.

Bathymetry maps are usually constructed using multibeam systems, which are expensive and require powerful sensor platforms. Side-scan systems are comparatively cheap, easier to deploy, and provide wider area coverage, which makes the possibility of employing them as a substitutive mean for acquiring elevation maps an extremely interesting proposition.

This is particularly true for offshore and military applications where autonomous underwater vehicles are becoming widely spread for inspection tasks (offshore) and Mines and Counter Measure Tasks (military). The ability to simulate the appearance of underwater structures (pipelines, mines) in real environments (augmented reality) will be the key to the development of reliable autonomous systems able to interpret the environment robustly. This has been recently demonstrated in the mine warfare context [5].

II. BACKGROUND

Efforts oriented to the utilization of side-scan sonar for the indirect determination of seabed topography have been scarce [2], [6]–[8], [10], the main reasons being the complexity of the full mathematical projection model and the high number of procedures required for preprocessing the original source data before it can be effectively used. In most cases where acquisition of seabed topography is important, attention is driven to more straightforward solutions such as multibeam bathymetric systems.

Most existing work on seabed reconstruction from side-scan has been mainly qualitative and oriented to obstacle avoidance for underwater vehicles [2], [7]–[9]. Other works focus on seabed texture classification or object recognition [10]–[13]. In all these situations precise descriptions of the seabed topography are not critical.

In general, the fundamental idea behind the reconstruction methods is to determine a model for the ensonification process that is simple enough for the image formation problem to be inverted, obtaining an approximation to the surface gradients, which can be globally described as shape-from-shading methods [3], [4], [6], [14]. Our goal is to apply and extend these methods using a multiresolution statistical approach to determine the most probable configuration of the seabed topography compatible with the side-scan image actually observed.

The authors are with the Ocean Systems Laboratory, Heriot-Watt University, EH14 4AS Edinburgh, U.K. (e-mail: coiras@nurc.nato.int; ecoiras@gmail.com; y.r.petillot@hw.ac.uk; d.m.lane@hw.ac.uk).
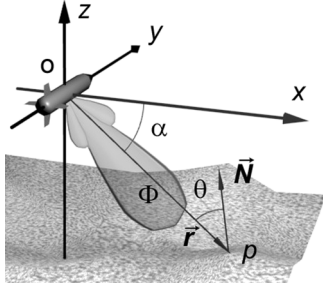
Fig. 1. Side-scan image formation.

The special case of side-scan images is particularly advantageous when attempting shape from shading. The two main reasons are the position of the light source (here transducer) is known (it coincides with the position of the camera, here the receivers), and the Lambertian model is a very good approximation for the sonar scattering produced by the materials usually found on the seabed (rocks, sand, etc.) [15]–[17]. Radiosity effects (interreflections) on sonar images are also much smaller than those found in visual images, since the illuminating sonar pulses are time gated, as opposed to standard light sources.

Shape from shading is generally an underconstrained problem in the sense that, usually, just a single sample for every scene point is available (in the form of a line-scan pixel) and that most of the model parameters cannot be directly inferred from that observation alone. Additional constraints and an optimization method have to be used in order to solve the problem. Our particular approach uses a multiresolution implementation, starting from a coarse subsampled version of the image, and then working out more details at finer scales. At each scale, a statistical framework is used to select the most probable set of parameters from the solution space and, in our case, is implemented as an iterative procedure inspired by the expectation–maximization method [18], [19]. At every iteration, the estimated solution is used to simulate a side-scan image of the observed scene, which is then compared to the side-scan image actually observed; solution parameters are finally refined using gradient descent [20] as the optimization method. The process is repeated until convergence is achieved up to a given degree.

III. IMAGE FORMATION

A. Side-Scan Sonar

The side-scan image formation process is briefly sketched in Fig. 1. The sensor's acoustic source at o produces an ensonification pulse that illuminates the seafloor. Some of the acoustic energy reaching any seabed point p is scattered back and can be measured by the sensor. The intensity of the corresponding pixel of the side-scan image will depend on the amount of energy scattered back from the surface point. The pulse is not isotropic, but follows a particular beam-profile Φ that depends on the grazing angle α subtended by the vector $\vec{r}(p)$ from o to p. The amount of energy scattered back also depends on the seabed reflectivity R at the point. Statistically, side-scan backscatter images present a Rayleigh distribution [21].

B. Scattering Model

In order to model the scattering process, we use the traditional Lambertian model [2]–[4], which permits us to derive the returned intensity from the parameters defining the observed scene. This simple model for diffuse scattering assumes that the returned intensity depends only on the angle of incidence of the illuminating sound pulse, and not on the angle of observation or on the frequency of the pulse. Under this assumption, the intensity I returned from a seabed point p can be represented by the following expression:

$$I(p) = K\Phi(p)R(p)\left|\cos\left(\theta(p)\right)\right| \tag{1}$$

where Φ represents the intensity of the illuminating sound wave at point p, R is the reflectivity of the seafloor, θ is the incidence angle of the wave front, and K is a normalization constant. Since most logged side-scan images already include a time-varying gain (TVG) correction [22], [23] for compensation of the intensity decay with distance and grazing angle, no dependence on radial decay has been included in the model (this would otherwise appear as a term on $\|\vec{r}(p)\|^{-2}$). Therefore, in order to simplify the model, all the intensity variations caused by the sensor's beam profile, the radial decay and the corrections are supposed to be grouped under the beam-pattern Φ.

The dependence on the seafloor's elevation is implicit in the incidence angle $\theta(p)$, which depends on the grazing angle α from the acoustic source and the orientation of the surface normal $\vec{N}(p)$. This dependence can be made explicit by first expanding the cosine in (1) as follows:

$$\cos\left(\theta(p)\right) = \frac{\vec{r}(p)\cdot\vec{N}(p)}{|\vec{r}(p)|\cdot\left|\vec{N}(p)\right|} \tag{2}$$

and then by representing $\vec{N}$ and $\vec{r}$ on a coordinate system relative to the sensor (Fig. 1). Expressing p as $(x, y, Z(x, y))$—with x being the across distance from the sensor and y pointing along its direction of movement—gives

$$\vec{r}(p) = (x, 0, Z(x, y))$$
$$\vec{N}(p) = \left(-\frac{\partial Z}{\partial x}(x, y), -\frac{\partial Z}{\partial y}(x, y), 1\right) \tag{3}$$

where the y coordinate in $\vec{r}(p)$ is 0 because the side-scan sonar pulse Φ is shaped so that only the points contained in the $x - z$ plane are illuminated.

A combination of expressions (1)–(3) gives the forward model for the computation of the intensity I at any point p, given the model parameters R, Z, and Φ

$$I(x, y) = -K\Phi(x, y)R(x, y)$$
$$\cdot\frac{Z(x, y) - x\cdot\frac{\partial Z}{\partial x}(x, y)}{\sqrt{x^2 + Z^2(x, y)}\cdot\sqrt{\left(\frac{\partial Z}{\partial x}(x, y)\right)^2 + \left(\frac{\partial Z}{\partial y}(x, y)\right)^2 + 1}} \tag{4}$$

where the gradients $\partial Z/\partial x$ and $\partial Z/\partial y$ can be approximated by finite differences, yielding an expression that depends directly on Z.

C. Determination of K

The values of the model parameters are limited to particular ranges, which have to be observed during the optimization process.

Reflectivity values, by definition, have to lie between 0 and 1, but to avoid the model from collapsing in the darker areas, a lower bound $R_{\min}$ greater than 0 is chosen on initialization (with a typical value of 0.1). Intensity values for the incident sound wave are assumed to lie between 0 and 1, which would normally mean bounding Φ within that range. However, because Φ includes the unknown applied TVG, values greater than 1 are allowed. In practice, this amounts to just a little overshoot for the bigger angles, which naturally correspond to points of the seabed farther away from the sensor and, therefore, requiring higher TVG corrections.

Completion of the forward model (4), therefore, requires a normalization constant K to account for the fact that most of the source image intensity values are being affected by some kind of scaling (because of the TVG, the radial decay, etc.). Our approach to determine K is to use the maximum achievable returns at every surface point to normalize the computed pixel values. Under the diffuse reflectance model and the selected parameter ranges, the maximum intensity return $I_{\max}$ at any surface point is 1 (for reflectivity and incident intensity values both 1 and for the optimal surface orientation with respect to the incident illumination). Assuming the part of the surface normal depending on y is kept constant, the surface orientation with maximal return can be obtained by differentiation of (4) with respect to $\partial Z/\partial x$ and equating to 0, checking the second derivative is negative to ensure the extremum is a maximum. Solving for $\partial Z/\partial x$ and substituting in (4) gives the following expression for the computation of the maximal intensity return at any surface point:

$$I_{\max}(x,y) = \frac{-Z + \frac{x^2}{-Z} \cdot \left(1 + (\partial_y Z)^2\right)}{\sqrt{x^2 + Z^2} \cdot \sqrt{1 + (\partial_y Z)^2 + \frac{x^2}{Z^2} \cdot (1 + (\partial_y Z)^2)^2}} \tag{5}$$

where explicit dependences of Z on (x,y) have been removed for clarity. Note that $I_{\max}$ is always positive because the sensor is always above the seafloor, which means $Z(x,y)$ is always negative according to our choice of coordinates (Fig. 1). It is also clear from the expression that $I_{\max}$ is always between 0 and 1.

Expression (5) results in the following normalization for the forward model (1):

$$K(x,y) = \frac{1}{I_{\max}(x,y)}. \tag{6}$$

IV. Sonar Inversion

A. Methodology

Equation (4) provides a direct formula for estimating the returned intensity given the model parameters. However, the inverse problem—obtaining the model parameters from the observed intensities—is clearly underdetermined, since we only have one observation (of I) at each point to compute the values of the three model parameters.

In order to solve this problem, we propose to take a least mean-squares (LMS) approach and try to minimize the sum E of the squared differences between the points of the observed image $I(x,y)$ and those of the model $\hat{I}(x,y)$

$$E = \sum_{x,y} E(x,y) = \sum_{x,y} \left(I(x,y) - \hat{I}(x,y)\right)^2. \tag{7}$$

Therefore, the following optimization problem needs to be tackled:

$$(Z, R, \Phi) = \arg\min(E). \tag{8}$$

There are a number of ways to approach such an optimization problem. For convex problems, gradient descent or conjugate gradient techniques are the natural choices. For nonconvex problems with possibly a large number of local minima, other techniques such as stochastic techniques (simulated annealing, stochastic expectation maximization) or evolutionary techniques (genetic algorithms, particle swarm optimization) are preferable [24]. These techniques are broadly performing two main tasks: exploration of the parameter space and exploitation of the promising areas. For these techniques to perform well, the exploration phase needs to be effective (both in computation time and precision) at locating promising areas. Our functional is likely to yield a large number of local minima due to the underconstrained nature of the problem. This would favor the use of global optimization techniques with strong exploratory component such as genetic algorithms or evolutionary computing. However, those techniques can be prohibitive in terms of convergence time if the parameter space is large. This is the case here for two reasons:

First, we are dealing images which are typically 1000×1000 pixels. For each pixel, three parameters have to be determined. The overall search space is, therefore, in $\Re^{3*100000}$. This is a huge parameter space which cannot be realistically tackled by global techniques if true convergence criteria (stochastic convergence for simulated annealing [25]) are to be met.

Second, it is important to note that in the case of our functional, several combinations (Z, R, Φ) will yield the same error metric while possibly being in very different places in the parameter space. Therefore, the functional will present numerous local minima and possibly several global minima. Therefore, even the application of global techniques would not guarantee the convergence to the true global minimum (out of all the potential possible ones).

To alleviate the problem one solution would be to reduce the parameter space by representing the problem differently (spline surfaces for Z and R, polynomials for Φ). This solution has been envisaged and discarded as the level of details that needs to be recovered (high frequencies) would require a very large number of parameters for the models making the models difficult to use and probably unstable.

The solution we have chosen is to use a standard gradient descent, thus converging to the next local minimum. To ensure convergence to a realistic minimum, great care is taken in the initial starting point by introducing prior knowledge on the scene and image formation process. The cost function is also smoothed by using a multiresolution approach, hence avoiding
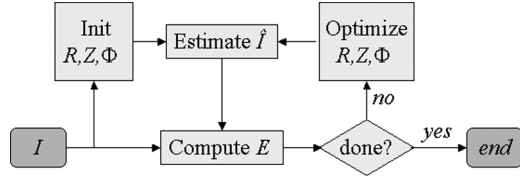
Fig. 2. Outline of the optimization procedure.

obvious local minimum in the vicinity of the starting point. Finally, we introduce an iterative process where regularization conditions are enforced globally on the parameter space after each local optimization phase. The details of the process are described in the next section.

B. Optimization Process

In order to solve this optimization problem, we propose the use of a method inspired by expectation–maximization [18], [19] (Fig. 2), which will iteratively converge to an optimal set of modeling parameters given a source side-scan image I.

In the expectation stage, the current estimates for the model parameters are used to compute an estimation of the intensity $\hat{I}(x,y)$. This is achieved by substituting the parameters $R(x,y)$, $\Phi(x,y)$, and $Z(x,y)$ from the previous iteration in the forward model presented in expression (4).

In the maximization stage, a straightforward gradient descent approach [20] is used to minimize E by updating the model parameters as follows:

$$R(x,y) \leftarrow R(x,y) - \lambda \cdot \frac{\partial E}{\partial R}(x,y)$$
$$\Phi(x,y) \leftarrow \Phi(x,y) - \lambda \cdot \frac{\partial E}{\partial \Phi}(x,y)$$
$$Z(x,y) \leftarrow Z(x,y) - \lambda \cdot \frac{\partial E}{\partial Z}(x,y) \qquad (9)$$

where λ is a small constant value used to control the rate of change. The expressions are iterated until the variation in the error E is below a given threshold. The particular value chosen for lambda does not seem to be very relevant for the final 3-D reconstruction results, but we found $\lambda = 0.25$ to be a good compromise between final optimization error and computation time.

Direct operation in (9) yields the following expressions for updating the model parameters:

$$R(x,y) \leftarrow R(x,y) + 2\lambda \cdot \frac{\hat{I}(x,y)}{R(x,y)}\left(I(x,y) - \hat{I}(x,y)\right)$$
$$\Phi(x,y) \leftarrow \Phi(x,y) + 2\lambda \cdot \frac{\hat{I}(x,y)}{\Phi(x,y)}\left(I(x,y) - \hat{I}(x,y)\right)$$
$$Z \leftarrow Z - 2\lambda\hat{I}(I - \hat{I})$$
$$\cdot\left(\frac{-\partial Z/\partial y - \partial Z/\partial x}{1 + (\partial Z/\partial y)^2 + (\partial Z/\partial x)^2}\right.$$
$$\left. + \frac{1+x}{x(\partial Z/\partial x) - Z} + \frac{Z}{x^2 + Z^2}\right) \qquad (10)$$

where the explicit dependence of the parameters on (x,y) has been removed in the last equation for clarity.

C. Regularization

As our method is pixel based, a regularization scheme is needed to maintain coherence between neighboring surface points. Regularization is performed at the end of every iteration by filtering the estimated reflectivity and beam-pattern maps. Reflectivity values for the points in shadowed areas are set to that of their nearest illuminated neighbors by hexadecagonal dilation [26] of nonshadowed areas. Whereas values of Φ for all the points subtending the same angle α to the sensor are set to their median value, since the beam profile of the sensor is supposed to be constant

$$\Phi(x,y) = \text{Median}\left\{\Phi(x_0,y_0)|\alpha(x_0,y_0) = \alpha(x,y)\right\}. \quad (11)$$

Other simple regularization techniques were tested. For instance, inspired on simulated annealing methods, we tried smoothing reflectivity and elevation maps on logarithmic iterations (less frequent as numbers of iterations increased). This slightly reduced the convergence error in exchange for slightly longer convergence times, but did not produce clearly noticeable differences in the resulting 3-D reconstructions.

D. Initialization

The optimization procedure starts by initialization of the R, Z, and Φ maps. The reflectivity map is set to a constant value (typically 0.9, although the particular value does not seem to be especially relevant), and the elevation of every point (x,y) is set to that of the first return at $(0,y)$—which is equivalent to the traditional assumption of a flat seabed. Then, the beam-pattern Φ is computed by setting it to the original image values and then regularizing it as shown in (11).

Other initialization approaches have been tested. In particular we have tried to use the shadows in the original image to obtain a coarse approximation to Z, and also to use a variation of the original image as an initialization for R. All approaches generated very similar end results and final convergence errors. After extensive tests, we were able to conclude that the most important factor driving the final optimization result is the multiresolution implementation (see the following section), which provides a better initialization for the final full-resolution stage of the optimization procedure.

E. Multiresolution

A multiresolution implementation of the method described in the paragraphs above results in better convergence and improved results. The main reason being the better initialization for the final (full resolution) scale, as compared to the simple initialization proposed in Section IV-D. The subsampling and resampling steps performed in the multiresolution implementation also contribute to the regularization procedure, which is of extreme importance due to the point-wise nature of the optimization method described in Section IV-B, which operates on a per-pixel level. The simple regularization stage described in Section IV-C is able to restore some of the interdependence of the pixel neighborhoods, but bigger seabed features, such as slowly varying slopes, cannot be fully recovered just by it. Effects of this limitation are shown in the top part of Fig. 3, where
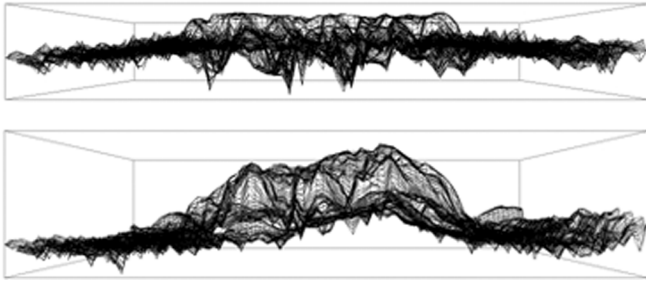
Fig. 3. Front view of the reconstruction of a rock using (top) single-stage and (bottom) three-level multiresolution implementations of the proposed reconstruction method. The single stage is not capable of recovering shape features of a bigger scale.
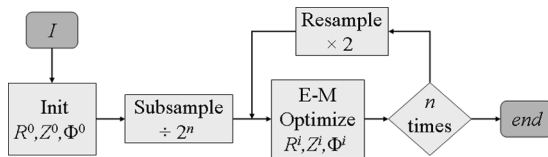

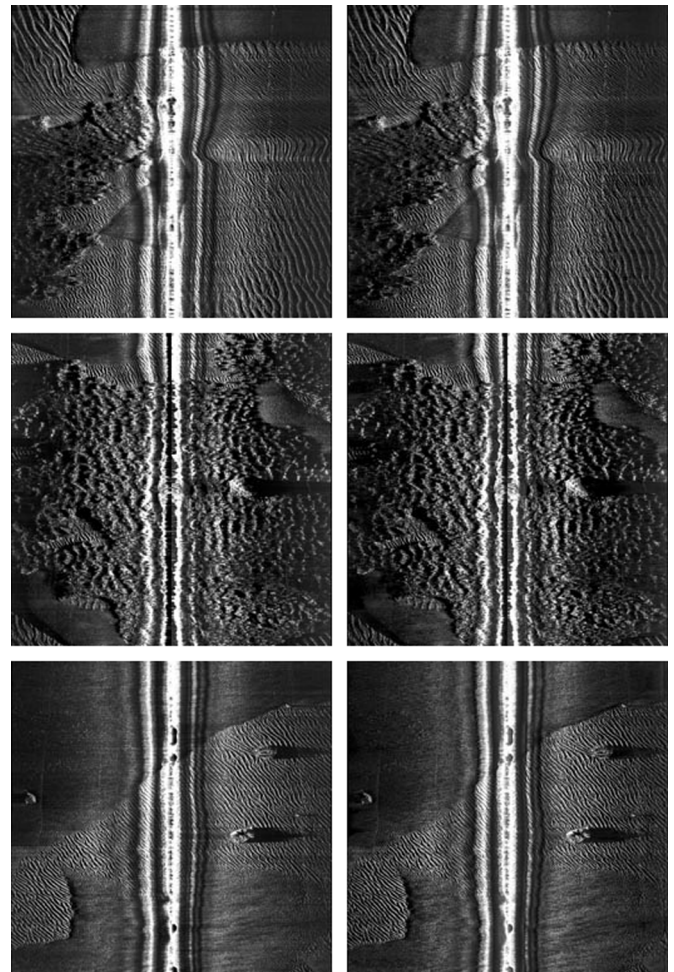Fig. 4. Outline of the multiresolution implementation for the optimization procedure.

Fig. 5. (Left column) Ground-range images and (right column) synthetic models after convergence, for three different source side-scan images of the same Remus mission. Image sizes are 1024×1000 pixels with a resolution of 0.058 m/pixel horizontally and 0.12 m/pixel vertically, which result in an actual size of 59.39×120 m.

the full shape of the rock has not been completely recovered on the single-stage implementation of the proposed method.

Results improve notably when using a multiresolution version of the same algorithm, which is able to recover the seafloor scenes in a more natural way, as well as reducing the overall error at convergence. The bottom part of Fig. 3 shows the results of this approach, where the shape of the rock is better estimated, once that more of the spatial frequencies involved are taken into account.

Implementation of the multiresolution version starts by the construction of a multiresolution pyramid by iterated subsampling of the source side-scan image. Processing starts at the smallest level (coarser resolution), using the initialization and regularization procedures described in the previous sections. The resulting R, Z, and Φ maps from one level are used as initial maps for the next resolution level. The process finishes when the final stage—corresponding to the full resolution image—is processed. An outline of this multiresolution implementation is presented in Fig. 4.

V. RESULTS

Datasets from different systems have been processed in order to reconstruct the bathymetry from their side-scan images. Results for Remus, Ocean Explorer and Geosub datasets are presented below. Later, the limitations of the proposed reconstruction method and their influence on the results are discussed. Typical running time for a 1024×1000 pixel image using three multiresolution levels is about 5 min on a Pentium III 2.0 GHz.

A. Remus Results

The BP02 survey was performed off the coast of Italy using a Remus autonomous underwater vehicle (AUV) equipped with a

Marine Sonics sensor. Three images from the survey have been processed as shown in Fig. 5. Ground-range corrected side-scan images are shown in the left column; the corresponding synthetic models after convergence of the three-level multiresolution implementation are shown in the right column. On convergence (to a possible local minimum), the resulting synthetic models are extremely similar to the ground-range images, suggesting that the data are well fitted by the solutions found by the proposed reconstruction approach.

The full set of outputs of the reconstruction process for the image in the bottom row of Fig. 5 is shown in Fig. 6. The effect of the beam profile permeates the reflectivity and elevation maps in the region right under the sensor path, where the Lambertian model cannot properly approximate the nondiffuse reflections. Results are nonetheless consistent with the observed features on the source image. The sand ripples and the structure of the rock are clearly visible in the perspective view of the 3-D surface.

The beam profiles recovered from the three source images of Fig. 5 are compared in Fig. 7. Apart from the inaccuracies corresponding to the region right below the sensor, the shape of the main lobes is consistent across the three profiles, suggesting
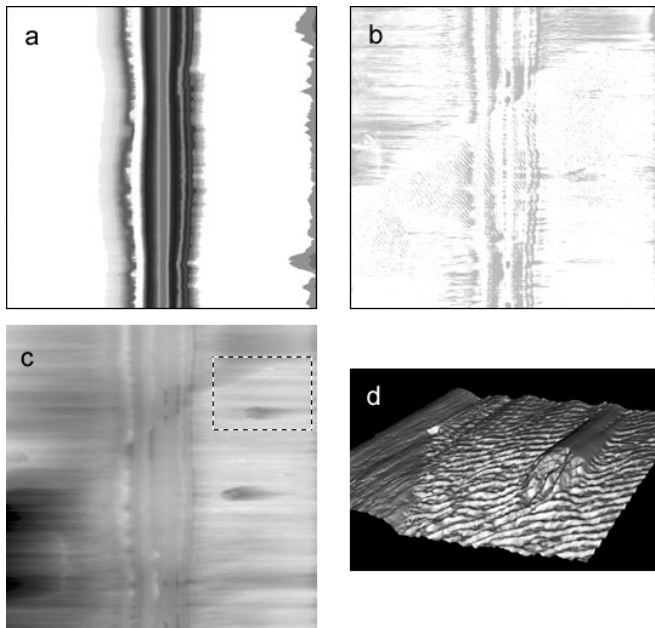
Fig. 6. Full set of model parameters and 3-D view for the Remus image shown in the last row of Fig. 5. (a) Beam-pattern, (b) reflectivity, and (c) elevation maps. (d) Perspective view of a 3-D flat-shaded surface constructed from the highlighted region (328 × 247 pixels, corresponding to an actual size of 19.02 × 29.64 m) of the elevation map.
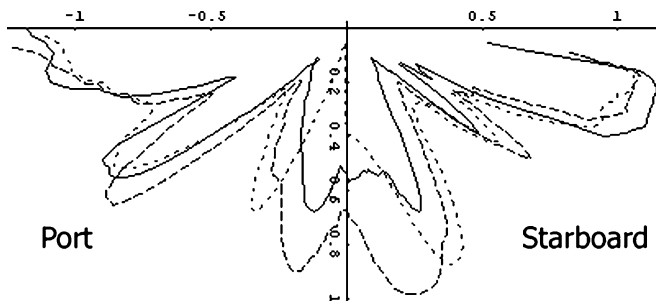

Fig. 7. Beam profiles recovered from the three source images of Fig. 5.

that, in effect, the same sensor and TVG settings have been used for the acquisition of the three source images.

Fig. 8 shows the evolution of the overall error E at the final full-resolution stage for the source image corresponding to the bottom row of Fig. 5. The three- and five-level multiresolution versions perform better than the single stage (no multiresolution), due to better initialization.

B. Ocean Explorer Results

The GOATS2000 survey was conducted off the coast of Italy using the Ocean Explorer AUV equipped with an Edgetech dual frequency DF-1000 sonar. Fig. 9 shows a region of one of the side-scan images from the survey. Posidonia formations, as well as heavy sand ripples can be appreciated. The result of the reconstructions clearly shows the 3-D structure of these features.

C. Geosub Results

One of the Autotracker project missions was performed off the Orkney Islands in Scotland. The Geosub AUV was used to track an underwater pipe of 76.2 cm in diameter, and images
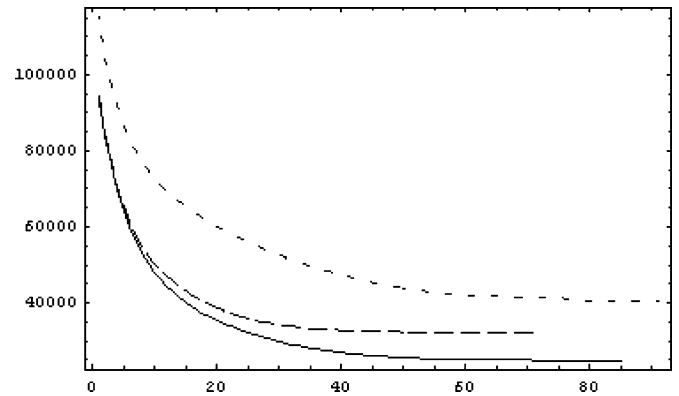

Fig. 8. Evolution of the overall error at the last full-resolution stage with the number of iterations for: (short dash) no multiresolution, (long dash) three-level multiresolution, and (continuous line) five-level multiresolution.
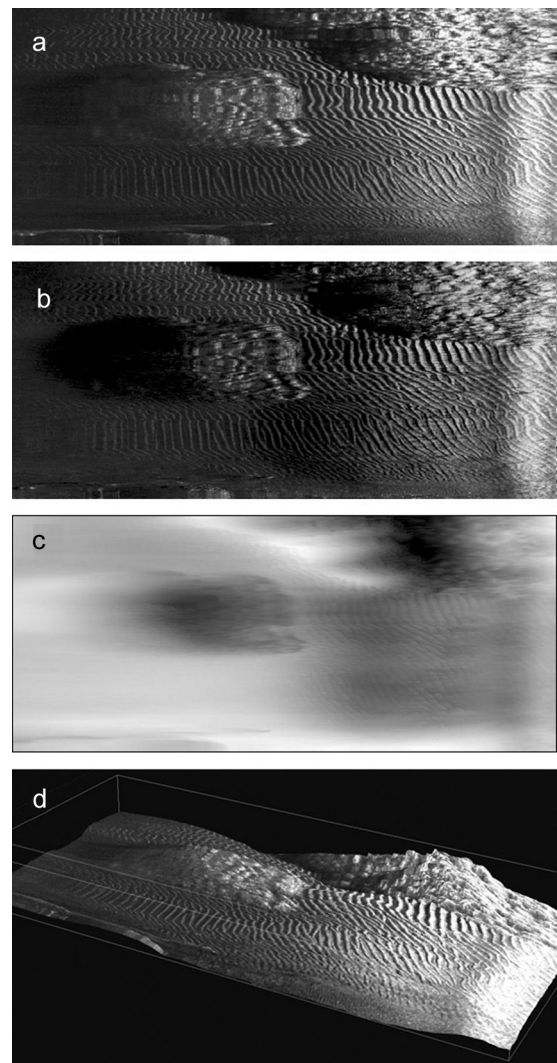

Fig. 9. Reconstruction results for the Ocean Explorer vehicle. (a) Ground range section (1920 × 657 pixels, resolution 0.026 m/pixel horizontally and 0.11 m/pixel vertically) of the original side-scan image; (b) model after convergence of the three-level multiresolution reconstruction; (c) resulting elevation map; (d) corresponding perspective view of the textured 3-D surface.

were acquired with an Edgetech DW-2200 sonar. Fig. 10 shows one of the images, where the acoustic shadow casted by the pipe
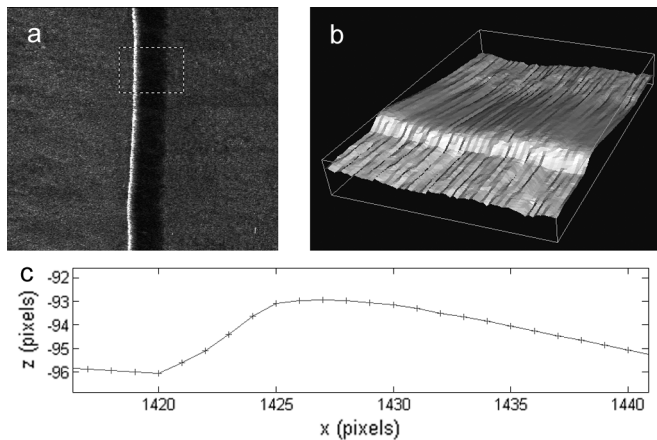
Fig. 10. Results for the Geosub vehicle. (a) Section (383 × 346 pixels) of a side-scan image showing an underwater pipe; (b) flat-shaded perspective view of the resulting 3-D reconstruction for the highlighted region (92 × 65 pixels); (c) profile view of the reconstructed pipe. Diameter of the pipe is 76.2 cm, and ground-range image resolution is 8.7 cm per pixel.

Fig. 11. Application of the proposed reconstruction method to augmented reality. Left: Region of the seabed showing sand ripples (242 × 153 pixels, resolution 0.058 m/pixel horizontally and 0.12 m/pixel vertically). Right: Results after embedding three spherical objects (2-m diameter). Notice how the shape and shadow of the objects realistically integrate with the rippled seafloor.
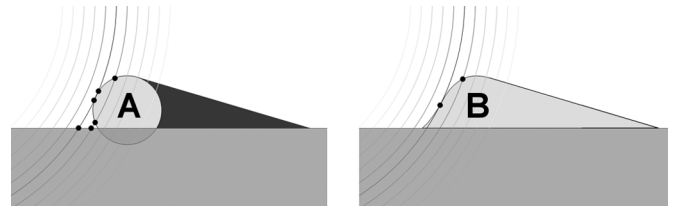
Fig. 12. Two seafloor configurations that would generate the same side-scan image. On the left, pipe A sits on the seabed, resulting in a multivalued elevation function $Z(r)$ for every range r, as shown by the black dots. On the right, an equivalent shape B that generates the same signal return for every range r is constructed so that the elevation function $Z(r)$ is single valued.

can be clearly appreciated. The shadows have traditionally been used for the estimation of the height of targets, and in principle could be used to estimate the pipe radius. However, the fact that the pipe might be partially buried into the seabed makes this approach less suitable for this type of target. The 3-D reconstruction method proposed in this paper, however, can be used to obtain the profile of the pipe, which can be later used to estimate its radius independently of its burying depth, as discussed in Section V-F of the paper.

D. Augmented Reality Results

One direct application of the reconstruction procedure described in this paper is the generation of augmented reality sonar images. Inversion of an original side-scan image yields its elevation, reflectivity and beam-profile maps, which can then be locally modified to account for the presence of various synthetic objects or structures. Use of the forward model (4) on these modified maps will create a variation of the original image which contains realistic depictions of those synthetic objects.

An example of application is shown in Fig. 11, where three spherical objects have been artificially embedded into a side-scan image. This is of direct interest for training and benchmarking of detection and classification algorithms in a realistic environment, as getting real data of underwater objects is a complicated and expensive task (logistics, localization of targets). In practice, very few and limited databases of real object exist, constraining the types of algorithms that can be used. The technique presented in this paper opens the possibility to train and benchmark algorithms using synthetic images of the targets but in a very realistic setup where real environments can be preserved. More complex AR implementations could take into account the interactions between the synthetic objects and the real environment, given a proper model for these interactions is devised.

E. Reconstruction Limits

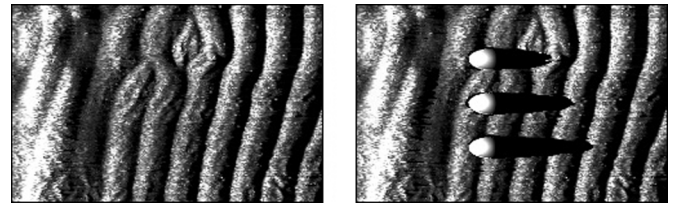The performance of the proposed reconstruction procedure is limited by two main factors. The first is the set of assumptions used for the development of the model and its implementation (Lambertian scattering, steepest descent implementation, etc.). However, the second, and most important, is the performance of the imaging sensor itself. Since the side-scan sonar integrates all the returns received from the surface at a given time, different surface configurations might produce exactly the same side-scan image. An example is shown in Fig. 12, which presents two equivalent cases. In Fig. 12 (left), a pipe is sitting on the seabed; each acoustic wave front intersects both the pipe and the seafloor in one or more points, whose contribution will be integrated when the acoustic return is measured; surface points under the acoustic shadow produced by the pipe (colored dark) are effectively hidden from the sensor's view and cannot be resolved. In Fig. 12 (right), a shape B is constructed so that each wave front intersects the shape at only one point and returns the same energy as A and the seafloor do. Targets A and B will generate the same image when scanned by a side-scan sonar.

Generally speaking, for every seafloor configuration where wave fronts might intersect different points at each sampling interval, there exists another seafloor configuration—which results in exactly the same side-scan image—such that the wave fronts intersect only one point at each sampling interval. This latter seafloor configuration is the one that would be reconstructed by the method proposed in this paper, since we implicitly assume that $Z(r)$ is a single-valued function for each range r.

It must also be noted that the reconstruction procedure is performed in the sensor's coordinate space, and—as is the case with the side-scan images themselves—geo-referencing of the reconstruction results is required in order to recover the proper along-track scaling. The examples shown in this paper are presented as they are obtained from the reconstruction, before any geo-referencing has been performed.
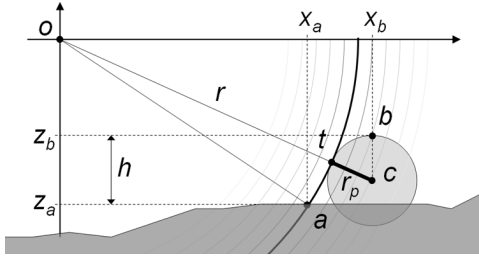
Fig. 13. Geometry of the reconstruction validation problem, showing key points a and b, from which the radius r_p of the proud or partially buried pipe can be determined.

F. System Validation

In order to determine the validity of the reconstruction method proposed here, an experiment on a target whose configuration is known is required. It must be remarked here that availability of information on the real dimensions of underwater structures is extremely limited, which makes this type of validation unfeasible in most of the occasions. We have used one of the Geosub side-scan images showing an underwater pipe of known diameter, and then computed its 3-D reconstruction. Part of the resulting 3-D model is presented in Fig. 10, where the pipe shape is clearly recognizable in the reconstruction. A profile view for the reconstruction of one of the side-scan lines is also shown.

It should also be noted that this section is not concerned with general detection, tracking or measurement of underwater pipes. For approaches to those problems see the specialized literature on the subject [10], [27]. Note, as well, that the approach taken in this section probably will not work if the section of the pipe is not parallel to the direction of movement of the sonar.

Due to the limitations described in the previous section, it is not possible to just fit the computed 3-D points to a cylinder in order to estimate the diameter of the reconstructed pipe. The reconstructed shape is actually the single-valued equivalent shape to that of the pipe, and an indirect validation method is required.

Our approach is to determine unique points, which can be identified both in the original configuration and in the reconstructed single-valued $Z(x)$ contours and from which the radius of the pipe can be estimated. As shown in Fig. 13, the points of interest are the top of the pipe b and the point a where the first wave front that touches the pipe intersects the seafloor, which results in a steep increase in the slope of the reconstructed shape. Using these two points alone, it is possible to compute the original pipe radius even if the pipe is partially buried, as long as the tangent point t first touched by the wave front is above the seafloor and the wave front intersecting the pipe at b does not touch any other seafloor or pipe point. Application of Pythagoras theorem to right-angled triangles $o - a - x_a$ and $o - c - x_b$ (Fig. 13) yields the following equation for r_p:

$$r_p = \frac{x_b^2 + z_b^2 - x_a^2 - (|z_b| + h)^2}{-2|z_b| + 2\sqrt{x_a^2 + (|z_b| + h)^2}}. \tag{12}$$

The quantization of the values of the x coordinate in units of 1 pixel (equivalent to 0.087 m) affects the distribution of the resulting values computed for r_p, as can be seen in Fig. 14. The
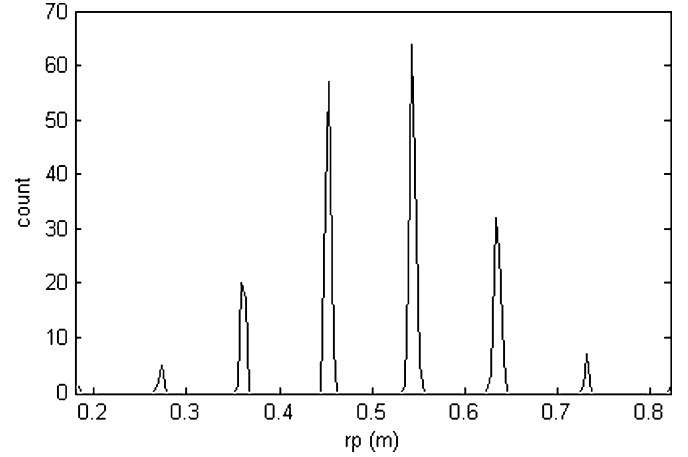

Fig. 14. Effects of the quantization of the x coordinate on the distribution of the computed values for pipe radius r_p are clearly visible in its histogram.

error dr_p associated to the quantization of x can be evaluated by computing

$$dr_p = dx_a \left| \frac{\partial r_p}{\partial x_a} \right| + dx_b \left| \frac{\partial r_p}{\partial x_b} \right| \tag{13}$$

where $dx_a = dx_b = 1$ pixel $= 0.087$ m.

Using (12) and (13), an estimate of the pipe radius and the corresponding quantization error can be computed for each reconstructed side-scan line of the pipe image section shown in Fig. 10. Averaging the results from the 346 lines gives an estimated pipe radius of $r_p = 0.51 \pm 0.19$ m. This has to be contrasted with the real value $r_p = 0.381$ m, which falls within the computed error range.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have presented a new method for the estimation of seabed elevation from side-scan images. The method uses a Lambertian model for the sonar scattering process, which is then used by a multiresolution optimization procedure to determine the seabed features ultimately responsible for the observed side-scan image. Examples of the type of results that can be obtained by the procedure have been presented. The performance of the reconstruction method has also been evaluated by analysis of the reconstruction results for a pipe of know size, showing that the estimated radius is within the computed tolerance error of the real value. It should be noted that the quantization error is quite large compared to the size of the object being reconstructed, which is ultimately caused by the small size of the pipe in the original side-scan image; more experiments should be performed with sensors of higher resolution or bigger targets to obtain more accurate performance measures for the reconstruction procedure presented in this paper. The composition of the target should also be taken into account, as some materials adjust better to the Lambertian acoustic scattering model than others.

Applications of the proposed 3-D reconstruction method are numerous and include accurate mosaic construction, detail improvement on existing bathymetry maps, obtention of 3-D models of underwater structures, generation of augmented reality sonar images, etc.

REFERENCES

[1] P. Blondel and B. J. Murton, *Handbook of Seafloor Sonar Imagery*. New York: Wiley, 1997.

[2] D. Langer and M. Hebert, "Building qualitative elevation maps from side scan sonar data for autonomous underwater navigation," in *Proc. IEEE Int. Conf. Robotics and Automation*, vol. 3, pp. 2478–2483.

[3] R. Zhang, P. Tsai, J. Cryer, and M. Shah, "Shape from shading: a survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 8, pp. 690–706, Aug. 1999.

[4] H. Ragheb and E. R. Hancock, "Surface radiance correction for shape from shading," *Pattern Recognit.*, vol. 38, pp. 1574–1595, 2005.

[5] Y. Petillot, S. Reed, and V. Myers, "Mission Planning and Evaluation for Minehunting AUVs With Sidescan Sonar: Mixing Real and Simulated Data," NATO Undersea Research Centre, La Spezia, Italy, Rep. SR-447, Dec. 2005.

[6] R. Li and S. Pai, "Improvement of bathymetric data bases by shape from shading technique using side-scan sonar images," in *Proc. Oceans Conf.*, pp. 320–324.

[7] S. Tiwari, "Mosaicking of the ocean floor in the presence of three-dimensional occlusions in visual and side-scan sonar images," in *Proc. Symp. Autonomous Underwater Vehicle Technology*, 1996, pp. 308–314.

[8] J. M. Cuschieri and M. Hebert, "Three-dimensional map generation from side-scan sonar images," *Trans. ASME J. Energy Resources Technol.*, vol. 112, no. 2, pp. 96–102, 1990.

[9] A. E. Johnson and M. Hebert, "Seafloor map generation for autonomous underwater vehicle navigation," *Auton. Robots*, vol. 3, no. 2–3, pp. 145–168, 1996.

[10] Y. R. Petillot, S. R. Reed, and J. M. Bell, "Real time AUV pipeline detection and tracking using side scan sonar and multi-beam echosounder," presented at the MTS/IEEE Oceans Conf., 2002.

[11] V. Murino and A. Trucco, "Three-dimensional image generation and processing in underwater acoustic vision," *Proc. IEEE*, vol. 88, no. 12, pp. 1903–1948, Dec. 2000.

[12] J. M. Bell, E. Dura, S. Reed, Y. R. Petillot, and D. M. Lane, "Extraction and classification of objects from sidescan sonar," presented at the IEE Workshop on Nonlinear and Non-Gaussian Signal Processing, 2002.

[13] B. Zerr and B. Stage, "Three-dimensional reconstruction of underwater objects from a sequence of sonar images," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, pp. 927–930.

[14] T. Okatani and K. Deguchi, "Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center," *Comput. Vis. Image Understand.*, vol. 66, no. 2, pp. 119–131, 1997.

[15] K. V. Mackenzie, "Bottom Reverberation for 530 and 1030 cps sound in deep water," *J. Acoust. Soc. Amer.*, vol. 33, no. 11, pp. 1498–1504, 1961.

[16] H. Boheme and N. P. Chotiros, "Acoustic backscattering at low grazing angles from the ocean bottom," *J. Acoust. Soc. Amer.*, vol. 84, no. 3, pp. 1018–1029, 1988.

[17] S. Stanic, K. B. Briggs, P. Fleischer, R. I. Ray, and W. B. Sawyer, "Shallow water high frequency bottom scattering off Panama City, Florida," *J. Acoust. Soc. Amer.*, vol. 83, no. 6, pp. 2134–2144, 1988.

[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. B*, vol. 39, no. 1, pp. 1–38, 1977.

[19] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, Nov. 1996.

[20] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C. Second Edition*. Cambridge, U.K.: Cambridge Univ. Press, 1994, p. 421.

[21] J. M. Bell and L. M. Linnett, "Simulation and analysis of synthetic sidescan sonar images," *Proc. IEE Radar, Sonar, Navig.*, vol. 144, no. 4, pp. 219–226, 1997.

[22] G. Shippey, A. Bolinder, and R. Finndin, "Shade correction of side-scan sonar imagery by histogram transformation," in *Proc. IEEE Oceans Conf.*, vol. 2, pp. 439–443.

[23] S. Anstee, "Removal of Range-dependent Artifacts from Sidescan Sonar Imagery," Defense Sci. Technol. Org. TN-0354.

[24] C. R. Reeves, Ed., *Modern Heuristic Techniques for Combinatorial Problems*. New York: McGraw-Hill, 1995.

[25] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, 1983.

[26] E. Coiras, J. Santamaria, and C. Miravet, "Hexadecagonal region growing," *Pattern Recognit. Lett.*, vol. 19, pp. 1111–1117, 1998.

[27] L. Tao, U. Castellani, A. Furiello, and V. Murino, "3D acoustic image segmentation by a RANSAC-based approach," in *Proc. MTS/IEEE Oceans Conf.*, vol. 2, pp. 1098–1101.

Enrique Coiras received the degree in physics, with a specialization in computer science, in 1993 from the Universidad Complutense de Madrid, Madrid, Spain, and the Ph.D. degree for his work on automatic image registration in 1999 from the Department of Optics, Universidad Complutense de Madrid.

He has vast experience in image processing, computer graphics, remote sensing, and software engineering, and has worked for several companies developing and implementing systems for digital image processing, video streaming, and simulation. He is currently a Research Associate with the Ocean Systems Laboratory, Heriot-Watt University, Edinburgh, U.K. His current interests include sonar image processing, 3-D reconstruction, and multiresolution registration.

Yvan Petillot (M'03) received the Engineering degree in telecommunications with a specialization in image and signal processing, the M.Sc. degree in optics and signal processing, and the Ph.D. degree in real-time pattern recognition using optical processors from the Université de Bretagne Occidentale, Ecole Nationale Supérieure des Télécommunications de Bretagne (ENSTBr), Brest, France.

He is a Specialist in sonar data processing (including obstacle avoidance) and sensor fusion. He is currently a Lecturer at Heriot-Watt University, Edinburgh, U.K., where he leads the Sensor Processing Group of the Oceans Systems Laboratory, focusing on image interpretation and mine and counter measures.

Dr. Petillot acts a reviewer for various IEEE Transactions.

David M. Lane (M'92) received the B.Sc. degree in electrical and electronic engineering in 1980 and the Ph.D. degree for robotics work with unmanned underwater vehicles in 1986 from Heriot-Watt University, Edinburgh, U.K.

He was a Visiting Professor with the Department of Ocean Engineering, Florida Atlantic University, Boca Raton, and is Cofounder/Director of See-Byte, Ltd., Edinburgh, U.K. Previously, he was with the U.K. Defence and Offshore Industries. He currently is a Professor with the School of Engineering and Physical Sciences, Heriot-Watt University, where he also is the Director of the Ocean Systems Laboratory. He leads a multidisciplinary team that partners with U.K., European, and U.S. industrial and research groups supporting offshore, Navy, and marine science applications. Major sponsors include oil companies, the United States Navy, the European Union, and the U.K. Ministry of Defence. He has published over 150 journal and conference papers on tethered and autonomous underwater vehicles, subsea robotics, computer vision, image processing, and advanced control. He is an Associate Editor for the *International Journal of Systems Science*.

Dr. Lane is a Member of Institution of Electrical Engineers (IEE), London, U.K.; the Professional Network on Robotics; and the U.K. Society for Underwater Technology Underwater Robotics Committee.