Shallow-water seafloor characterization for high-frequency multibeam echosounder: image segmentation using angular backscatter

Xavier Lurton, Jean-Marie Augustin, Samantha Dugelay, Laurent Hellequin, & Michel Voisset

IFREMER Centre de Brest
BP70
29280 Plouzané FRANCE
Email: lurton@ifremer.fr

Abstract

Sonar images from multibeam echosounders are submitted to various degradations, partly due to the sensor itself. Signal artefacts and corresponding corrections are presented, both for signal processing and for array directivity effects. The corrected image is processed using a segmentation method accounting for angular variations of backscattering strength, allowing to compensate effects of specular reflection and bathymetry. These various points are illustrated by experimental data shown at various processing stages.

1. Introduction

A strong analogy exists between works in geology on the continental shelf, and those undertaken onshore; often, main geological structures continue into the sea, in particular on the inner shelf. The main difference between onshore and offshore is sediment dynamics which strongly re-mould the sea floor. For studies of seafloor structures, side-scan sonar imaging has been routinely exploited for years by many scientific teams. More recently, modern multibeam echosounders have appeared allowing deeper investigation thanks to the simultaneous acquisition of bathymetry and reflectivity data by the same sensor.

For the geologist, obtaining a mosaic in which the actual information represents material and interface roughness variations, is an essential tool for a further rigorous interpretation. Typically, information is collected from various distinct sources: geoacoustical sensors (mainly side scan sonar and seismic reflection), or by sampling and photos. At this stage, the geologist analyses the sonar mosaic by making a one-to-one correspondance between a mean grey level and a geological unit. The synthesis document is a map established with methods identical to those used for terrestrial geological maps.

The work presented here deals with two aspects of multibeam echosounder data processing. *Correcting the artefacts* brought by the echosounder itself is an essential requirement, in order to make the image correctly interpretable. *Segmenting the sonar image* relies on an expert geological advice for the training phase; relevant information is used when delimiting the learning zones. The segmentation produces a cartography of different zones whose acoustical characteristics are similar to those estimated in the training stage. This procedure is similar to the one undertaken by the geologist. It produces a mosaic whose parameters are objective because having been rid, as far as possible, of artefacts brought by bathymetry and sensor imperfections. A further stage would consist in *fitting a model* with the backscattered signal characteristics estimated in the learning zones.

The method presented below will be illustrated using experimental data given in Figure 1. These data were obtained on the Atlantic inner shelf in the Bay of Biscay, using a high frequency multibeam echosounder. The extract presented here shows an interesting geological variety, featuring soft sediments alternating with harder deposits and rock areas.

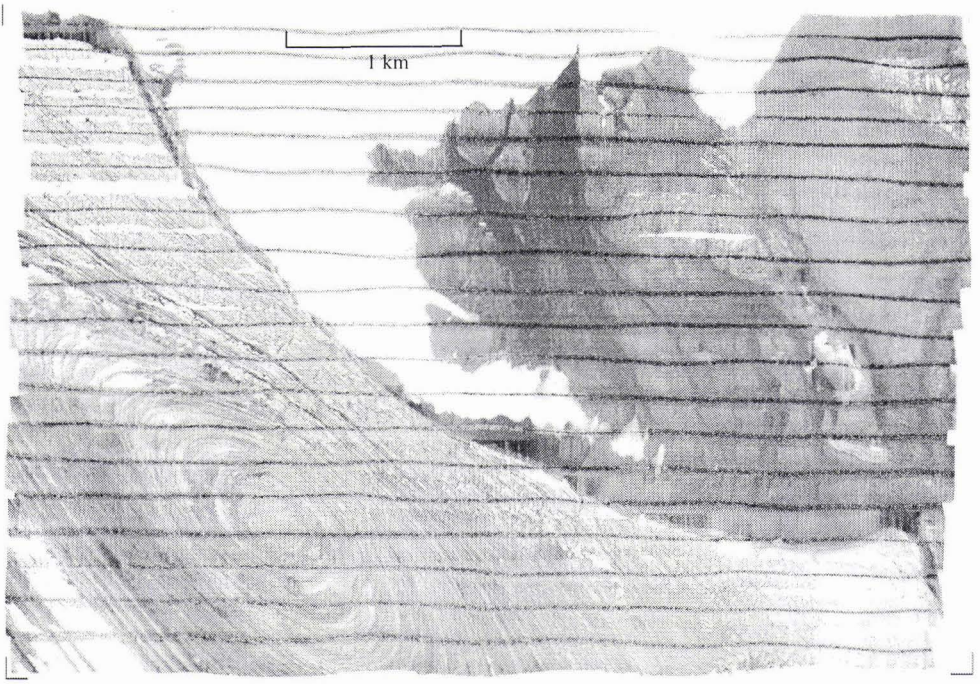


Figure 1. Sonar image from a 95 kHz multibeam echo sounder in shallow water (Bay of Biscay - water depth 15 to 45 m - zone area approx. 5.5 x 4 km)

2. Signal processing and corrections

After a quick presentation of the echosounder, it is shown that its signal processing and array characteristics may lead to undesirable effects in the obtained sonar images. Careful analysis of these effects leads to a correction procedure, indispensable before any image interpretation or processing.

2.1 Basic echo sounder signal processing

We consider in the following signals and data obtained with a Simrad EM950 multibeam echosounder [1]. This system emits pulses of duration 0.2 or 1 ms (according to water depth; the above data correspond to 0.2 ms) in a 150° vertical fan at about 225 dB/μPa/1m level. In reception, beamforming allows to create 120 beams; note that only 60 beams are formed for every other ping. The same array is used for transmission and reception, leading to individual beamwidth of 3.3°x3.3°. The signals inside each formed beam are used for bathymetry measurements: the arrival time is measured directly from the signal envelope for the near-vertical beams; for oblique incidences, every beam is divided into two half-beams for an interferometric precise measurement of the arrival time and angle. The measured time is finally converted into depth using the raypath geometry and the sound speed profile. Individual time signals from each beam are simultaneously exploited to create side-scan sonar images.

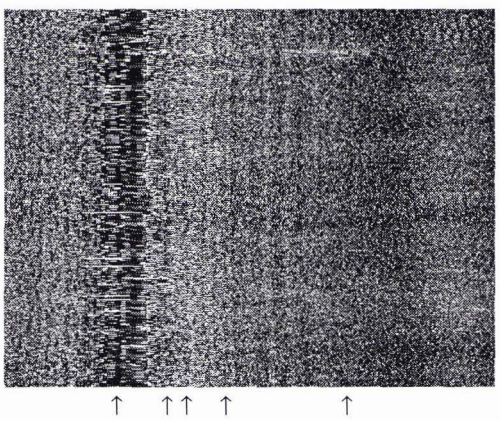
The various processing steps inside the echo sounder are as follows. After reception on the listening transducers, a time-varying gain (*TVG*) is applied on raw signals, aimed at compensating the high dynamics between returns from vertical and lateral incidences, both due to transmission losses in water, size of the insonified area, and backscattering strength (at its higher at the vertical, and strongly decreasing at shallow angles). The compensation law in the echo-sounder accounts for the first two effects using simple formulae. The *BS* angular variations are compensated by a Lambert's law in $\cos^2\theta$ for shallow grazing angles, and a gain varying linearly with angle around the vertical; the value of the latter correction adapts from one ping to the next one, and its amplitude is recorded along with other parameters of the ping. These various corrections, concentrated in a single TVG law, are intended to limit the dynamics of the incoming signal to the processing chain, and to homogeneize its level when presenting it as a sonar image.

Beamforming is processed by summing weighted contributions from rows of elementary transducers along the U-shaped array, the active row sliding along the circumference for angle scanning. In extreme sectors on both sides (60° to 75°), beams are formed using conventional beamforming techniques on a fixed row of transducers. Sidelobes are lowered by using a Dolph-Chebyshev shading. Time signals from the various beams are then processed on the one hand for the bathymetry measurement, and on the other hand they are presented in rows for sonar image display, digitized at a 0.2 ms rate; note that the amplitudes are rounded to integer values in half-dBs.

2.2 Beamforming directivity effects

2.2.1 Artefact descriptions

The partial sonar image presented in Figure 1 (left), recorded on a 30 meter depth gravel seabed, allows to identify several echo sounder artefacts, despite image resolution and deceptive grey scale dynamics. This image is presented as corrected by the echosounder, and hence is supposed to reflect only the backscattering strength of the seafloor coupled with bathymetry effects. The corresponding averaged amplitude is presented on the right side, giving a better idea of the order of magnitude of the phenomena.



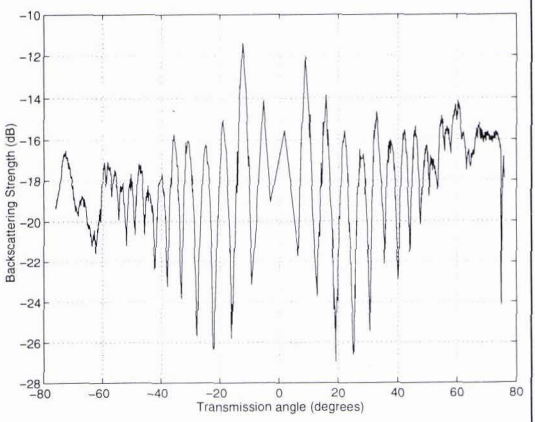


Figure 2: *(Left)* Sonar image obtained on a 30 m water depth, gravel seafloor (note that the whole swath is not presented). *(Right)* Corresponding averaged signal amplitude *vs* angle given by the echosounder

First of all, for emitted angles greater than 50°, slow variations affect the backscattered level: artificial grey tone contrasts appear on the corresponding sonar image. These artefacts are thought to derive from uneven array sensitivity and/or electronic gains applied to the signal emission and reception. Secondly, strong fluctuations (with extinction depths down to -20 dB), whose amplitudes become greater and greater as the angle tends to zero, appear parallel to the vessel track in the sonar image. The stable shape of these artefacts corresponds to the uncompensated directivity diagrams of each one of the 60 reception beams. Finally, the backscattered level is suspicious in the specular region: even after the specular effect compensation has been removed, the backscattered intensity collapses whereas the physics predict an intensity peak. To understand this third phenomenon, we had to detail the TVG law implemented in the sounder, and in particular, the insonified area compensation.

2.2.2 Phenomena analysis.

The estimated backscattering strength $\overline{BS(\theta)}$ is usually obtained from the classical "sonar equation" and basically computed as:

$$\overline{BS(\theta)} = 10\log \iint_{A(\theta)} BS_{nat}(\theta) W^2(\theta) dA \; - \; 10\log \overline{A(\theta)} \qquad (1)$$

where the integral term is the target strength of the actual insonified area $A(\theta)$ at angle θ, featuring $BS_{nat}(\theta)$ the backscattering strength value in natural unit and $W^2(\theta)$ the beam directivity diagram, $\overline{A(\theta)}$ being the estimated insonified area at angle θ. The actual BS estimation in our echosounder is based on the following approximations, which are discussed below:

- the integrand in (1) varies slowly enough to be taken out of the integral;
- the insonified area is approximated by

$$\overline{A(\theta)} = \frac{c\tau}{2}\theta_T \frac{H}{\cos\theta \sin\theta} \tag{2}$$

where τ is the pulse duration, and H the water depth.

For shallow grazing angles, the above approximations are acceptable; only the directivity compensation should take into account a slight angle shift corresponding to half a pulse duration. Apart from this minor correction, no deformation of the observed beam diagrams is expected; this justifies to use a method of linear corrections of directivity diagrams for off-specular emission angles.

On the other hand, when the emission angle tends to vertical, the insonified area $A(\theta)$ comes to be limited by beam aperture instead of pulse duration τ; hence its actual value becomes much smaller than its estimation $\overline{A(\theta)}$ based solely on pulse duration limitation. The measured $\overline{BS(\theta)}$ should be increased by an amount of:

$$I = 10\log\left(\frac{\theta_T}{c\tau} H \tan\theta\right) \tag{3}$$

whenever this term is positive. Moreover, in this case, the increase of $A(\theta)$ precludes the simplification of the integral term in (1) and leads to quite intricate effects. These will be detailed in a further paper [10], with signal simulations to evaluate their importance and to determine the validity limits of the correction procedures.

2.3 Artefacts compensation method.

To get rid of the sonar image defects, since no array calibration measurements were available to us, we developed a post-processing correction method using data from a "learning" seabed chosen as flat and homogenous as possible, such as those presented in Figure 2.

1. The various echosounder processing algorithms are compensated as described above. The Lambert's law correction and the specular effect processing provided by the echosounder are removed, and the insonified area is corrected using (3), in order to recover physical *BS* values. At this stage one should have a correct estimation of $BS(\theta)$ mixed with the array directivity effects.

2. Assuming that beam directivity patterns are only slightly deformed by backscattering, an ideal $BS(\theta)$ model is fitted on the beam central samples. We use for this purpose a simple functional form $BS_f(\theta)$, Gauss-like near vertical, and Lambert-like at oblique incidence:

$$BS_f(\theta) = 10\log\left(A\exp(-\alpha\theta^2) + B\cos^\beta\theta\right) \tag{4}$$

This implies that a form such as (4) correctly describes *BS* variations with angle; although this assumption is usually correct, this may be not always the case. The result of such a fitting is given in Figure 3.

3. If one substracts $BS_f(\theta)$ from the measured $BS(\theta)$, the resultant pattern is expected to be the actual array directivity diagrams and should be independent of the seabed nature. In a further step, one may use a realistic directivity model according to the actual array shape and processing, and fit it to the estimated array directivity, beam after beam; this allows to smooth the obtained directivity patterns.

4. The whole image can now be corrected from the directivity effects, using the emission angle and beam number of each pixel. This leads on the one hand to sonar images free of directivity artefacts, on the other hand it allows a correct estimation of $BS(\theta)$ from new homogeneous zones.

Results are presented in Figure 3: the left side shows the "raw" $BS(\theta)$ corrected from Lambert's law, specular effect and insonified area, and fitted with the ideal $BS(\theta)$ given by (4); the right side presents the directivity pattern obtained from three different seafloors (gravel, rock and sand). It is clear from these results that this approach provides remarquably stable directivity patterns, thus justifying *a posteriori* the method validity.

Shortcomings appear nevertheless for beams close to vertical: details of the oscillations are not correctly compensated, and should result in residual fluctuations in the central part of the sonar image; note that this problem is restricted to the [-10°, 10°] sector around the vertical, hence concerning a minor part of the image. Moreover, and more penalising for further identification operations, specular level still remains unexpectedly low after the various compensations have been applied, meaning that correction (3à has not been sufficient to get a correct estimation of the *BS* level. Most presumably, this is due to time undersampling: close to the vertical and in shallow water, signals associated with a given beam may last not more than one time sample. Further analysis of the central beams are presently under study [10], but it seems awkward to get fully informative data inside this angle sector.

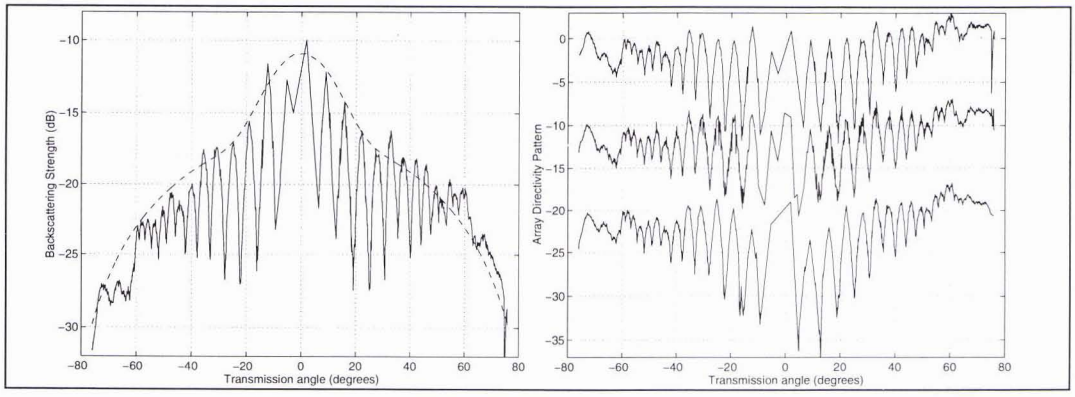


Figure 3. *(Left)* Averaged $BS(\theta)$ after compensation of echosounder processing *(full)* and fitted to an ideal $BS(\theta)$ *(dashed)* - *(Right)* Array directivity pattern obtained from gravel *(upper)*, rock *(middle)* and sand *(lower)* seafloors; the two lower curves have been shifted downwards 10 and 20 dB.

3. Sonar image segmentation

The purpose of image segmentation is to delimit homogeneous zones according to acoustical signal properties, namely backscattering strength level. Due to the fact that *BS* is both strongly dependent on the signal incident angle and on the seafloor properties, and that data are biased by the sounder characteristics, this precludes dealing directly with the raw sonar image; hence removing the sonar artefacts and estimating $BS(\theta)$ on representative zones were the first tasks to be conducted. We propose in the following a sonar image processing method accounting for the physics of the problem.

3.1 Measurement of angular *BS*

The practical measurements of *BS* have been partly described in the previous section. The output signals from the various beams have first to be corrected for propagation losses, and for array directivity effects. The actual incident angle on the seafloor has then to be computed, accounting for two distinct effects:

- refraction inside the water column: raypath angles (related to vertical) at the array and on the seabottom are related by the classical Snell-Descartes' law,
- local topography: data from the local Digital Terrain Model, obtained from bathymetry measurements, is to be used for computing the seafloor slope, both in the acrosstrack and alongtrack directions.

The exact insonified area may then be computed, limited either by the pulse duration or by the beamwidth, and taking into account the local seafloor slope. Finally, for a given signal sample, the obtained *BS* value is affected to the corresponding effective incident angle.

3.2 Image pre-processing

The EM950 signal is sampled at a constant time rate, producing an irregularly spaced sampled signal on the seabed. Mosaicking then consists in projecting this signal onto a regular grid. At shallow grazing angles, all raw sonar samples located in a same mosaic pixel are averaged; alternatively, near vertical, sonar samples are duplicated to compensate for the poor spatial resolution of the echosounder [3]. From this construction process, three levels of representation are produced:

- a chronological mosaic where pings are piled up without taking into account ship navigation, heading or interping distance, but where lateral projection is applied,
- a cartographic representation, where pings are relocated in a cartographic mosaic,
- an interpolated representation obtained from the previous mosaic.

The following segmentation method has been developed for chronological mosaics; on the one hand, neighbourhood computation is easier with this representation than with the cartographic representation, on the other hand, interpolation introduces pixel correlation therefore rending the model of the interpolated image complex [6].

3.3 Image segmentation

3.3.1 Segmentation using Markov Field theory

Segmenting means associating a label to each pixel in order to partition the image into homogeneous regions conditionally to the observed image. Mathematically, this involves maximising the *a posteriori* probability $P(X = \omega / Y = y)$, where X is the estimated label image, and Y is the observed mosaic.

Analysis of the backscattered signal shows that data is noisy, that the signal contains angular acoustic variations and at the same time contains bathymetric effects [3]. Since images are produced from a projection of the signal onto a regular grid, mosaics are in fact, a degraded view of the seabottom. Hence, the deformation that leads from X to Y is known and is expressed by the conditional probability $P(Y = y / X = \omega)$. We also have before hand knowledge of pixel interactions or existing correlations. This is described by the *a priori* probability $P(X = \omega)$. Using Baye's rule, the *a posteriori* probability is then expressed by:

$$P(X = \omega / Y = y) = \frac{P(Y = y / X = \omega) P(X = \omega)}{P(Y = y)} \tag{5}$$

where $P(Y = y)$ is a constant.

The Markov Random Field theory (MRF) [4] is particularly interesting in the sense that only a local model is sufficient, which means that there is no need to take into account the whole image when analysing a specific pixel, a set of local neighbouring pixels will suffice. It is then clear that defining an adapted neighbourhood system is an important step in the Markov modelling. In addition, the Hammersley-Clifford theorem states that the probability law of a Markov Random Field is a Gibbs distribution:

$$P(\omega) = \frac{1}{Z} \exp\left(-U(\omega)\right) \tag{6}$$

where $U(\omega)$ is the energy function defined over the neighbourhood θ and Z the partition function, a normalizing constant. X, X/Y and Y/X are all MRF's and therefore

$$P(X = \omega / Y = y) = \frac{1}{Z} \exp\left(-U_1(y / \omega) - U_2(\omega)\right) \tag{7}$$

where $U_1(y / \omega)$ is the energy function translating the deformation from X to Y, and $U_2(\omega)$ the component due to the spatial interaction of the labels. So, the maximization problem has now become the minimization of the energy functions. An extensive computation of all configurations is impossible due to the large number of possibilities. Therefore, in order to find the optimum configuration, we use a standard optimization algorithm (Iterated Conditional Mode) which computes the marginal probabilities for each pixel of the image and iterates over the image until satibilization of the process. The advantage of the algorithm is its speed, which is an important feature when dealing with images as large as we do. But, on the other hand, this process needs a correct initialisation in order to converge to a suitable solution. Finally, we work in a supervised framework; training zones delimited before the segmentation are essential to determine the number of classes in the image, and characteristic features of each label. These learning zones must be carefully chosen and sufficiently distinguishable so that the process may achieve a satisfactory segmentation.

3.3.2 Characteristic features of the labels

Because of angular variations and bathymetric effects, a homogeneous seafloor will not present a uniform grey level image. In this paper, we have considered the characteristic feature of a homogeneous zone is the estimated angular reverberation [2]. In this manner, we are substracting bathymetric effects from the signal and we are looking to segment according to seafloor nature.

3.3.3 Modelling the interaction energy

The segmentation process is applied to the chronological mosaic. The interaction energy is described by a classical function, which counts the number of similar labels in the pixel neighbourhood. In order to compensate for the interping distance and geometry which have been dropped in the representation, information is re-injected into the interaction energy .

- a rectangular neighbourhood in the mosaic is taken so that in reality a square neighbourhood is considered on the seabed,
- the influence of a pixel is weighted by the inverse of its relative distance on the seabed.

Interaction energy is then described by the following equation:

$$U_2(x) = \alpha \sum_{(s,t)} c_{st} \delta(x_s, x_t) \tag{8}$$

where c_{st} is the weighting coefficient and $\delta(a,b)=1$ if $a=b$, and $\delta(a,b)=0$ if $a \neq b$.

3.3.4 Modelling the deformation energy

Analysis of pixel distributions show that we do not have a homogeneous noise over the entire image. Probability distributions vary with the position of the pixel according to vertical incidence. Previous works with a 13 kHz multibeam system have proven that raw amplitude data follow a truncated Rayleigh law [5], [6]. Equivalently, intensities follow a truncated $\chi^2(2)$ probability distribution while averaged observations are modelled by a weighted χ^2 law where degrees of freedom depend on the number of averaged observations. Consequently, the deformation model is described by:

$$U_1(y/\omega) = \sum_{s\in S} -\ln(\exp(-y_s/x_s) - \exp(-10^{1/20} y_s/x_s))\,\delta(n_s = 1)$$
$$+\left[-(n_s-1)\ln y_s + n_s y_s/x_s - n_s \ln(n_s/x_s) + \ln\Gamma(n_s)\right]\delta(n_s > 1) \tag{9}$$

where y_s is the pixel intensity, x_s is the label intensity, and n_s the number of averaged observations.

We have considered that this is still a reasonable assumption in the case of our high frequency echosounder. Obviously, the model can be improved by taking into account specific aspects linked to the EM950, such as probability laws including textural diffraction (K distributions), texture features, or spectral parameters.

4. Real case application results

The leg we are presenting here (Figure 1 and 4) is an extract from the survey PLABAS where a EM950 echosounder was employed aboard the IFREMER *R/V Thalia* off the Basque coast (France). We have also acquired supplementary data on the same zone: side scan sonar, VHR seismics, samples, photos. The substratum is is a layer cake of calcareous layers and marl; each layer is less than 1 meter thick. These ancient rocks were folded, faulted and then eroded. The fold fault forms a cliff approximately 10 meters high visible in F, the full line AB is very close to its axis. Recent sedimentary deposits partially recover the substratum. These sediments are mainly constituted of fine to coarse sands and gravels. The spatial organization of these sediments is controlled by storm currents and tides which are very strong in this area (dunes, macrowaves). In the yellow region (S), fine sands have been deposited in a relatively shallow depression. The ocher zone corresponds to to coarse sands, and the green zone includes cobbles and gravels.

The segmented image was considered to be quite satisfactory by the geologists, and correctly reflecting their interpretation of the physical reality. Some particular points are to be highlighted:

- Most of the specular effect present on Figure 1 has been removed on Figure 4. Some reliquates are found on the upper part of the fine sediment area; note that they correspond to very strong and wide specular stripes on the original image.
- The various local features present on Figure 1 have been respected by segmentation, e.g. small sediment areas on the extreme left of the image; although the lighter area on the upper right corner may have been insufficiently identified by segmentation (it might have been interesting to introduce one more seafloor class).
- The distinction between fine and coarser sediments has been clearly established, although it was not at all evident on the original image.
- The stratified texture of the lower left corner of the image was the most challenging part of the process; it is comforting to see that it was most correctly segmented, globally respecting the concentric textural features of the geological structure.

Together with the segmented image, the $BS(\theta)$ curves obtained from the various test zones are presented in Figure 5. Their general shapes and levels are in good accordance with other observations found in the litterature. However the pattern shapes for the two soft sediments are remarkable, featuring an unusually slow decay in the [10-30°] sector. Obviously this is in contradiction with the general shape proposed in (4); note that our directivity pattern estimation was not affected by this effect since it proved to be the same for three different seafloors described using (4). Further studies are under progress for identifying characteristic features of the $BS(\theta)$ curves, and fitting them with physical models.

5. Conclusions

We showed in this paper that, provided that a correct compensation of echosounder artefacts is applied, sonar images can be efficiently segmented using the physical phenomena knowledge. Results obtained in a complex shallow water configuration appear quite satisfactory, leading to a segmented image in excellent accordance with the geologist perception, although it is clear that validation of information revealed with this method can only be fully achieved selectively with sample analyses and photos. This was obtained despite somehow penalising performances of the echosounder, especially around the vertical.

The emphasis should be stressed on a careful and relevant selection of the test zones as a preliminary to segmentation processing. All the above was presented with backscattering strength as the only characteristic physical feature to be used for segmentation; other signal characteristics are known to be potentially usable for high frequency echosounders used in shallow water (amplitude statistics, spectral and textural features...), and this should be checked in the future. Another important direction to explore, and also presently under study, is the identification of the measured $BS(\theta)$ with a theoretical model, and the possibility of extracting objective seafloor parameters from such a fit.

References

[1] SIMRAD, EM950 Product description.

[2] R.M. Gott, Remote seafloor classification using multibeam sonar, Doctoral Thesis 1995.

[3] J.M. Augustin et al., Contribution of the multibeam acoustic imagery to the exploration of the sea-bottom, Marine Geophysical Researches, 18, pp 459-486, 1996.

[4] S. Geman, D. Geman, Stochastic relaxation, Gibbs Distribution and the Bayesian Restoration of Images, IEEE transactions on pattern analysis and machine intelligence, vol. PAMI-6, NO. 6, 1984

[5] S. Dugelay, X. Lurton, J.M. Augustin, A new method for seafloor characterization with multibeam echosounders: image segmentation using angular backscattering, Proc. 3rd European Conference on Underwater Acoustics, vol. I, pp 439-444, 1996.

[6] S. Dugelay, Caractérisation des fonds marins à partir de données sondeur EM12, Mémoire de thèse, Université Paris-Sud-Orsay, 1997.

[7] L.Hellequin, Postprocessing and signal corrections for multibeam echosounders, unpublished

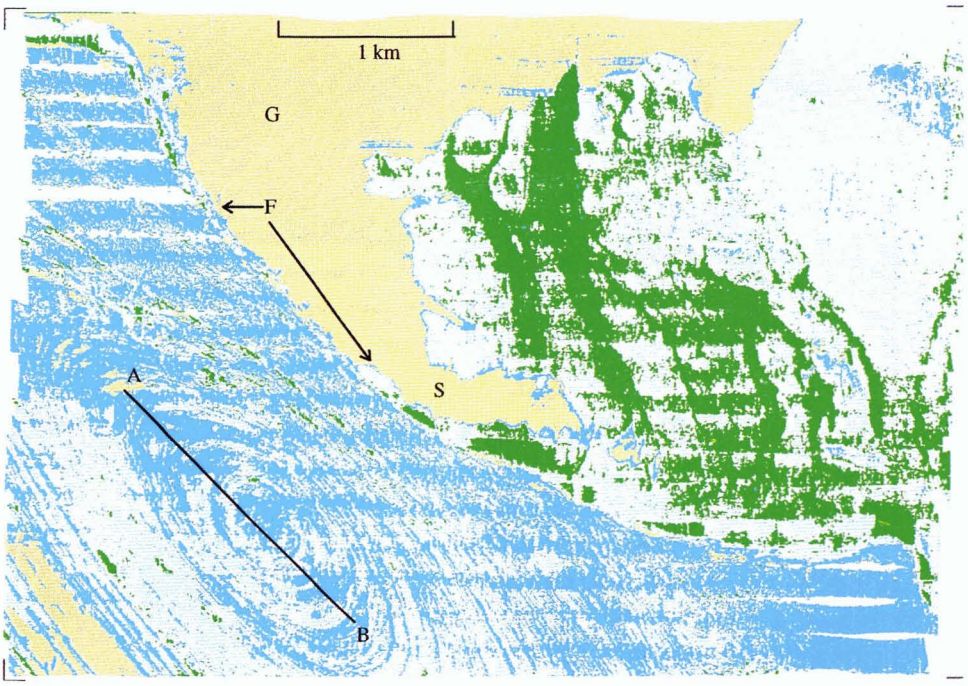


Figure 4. Segmented image corresponding to the sonar image of Figure 1.

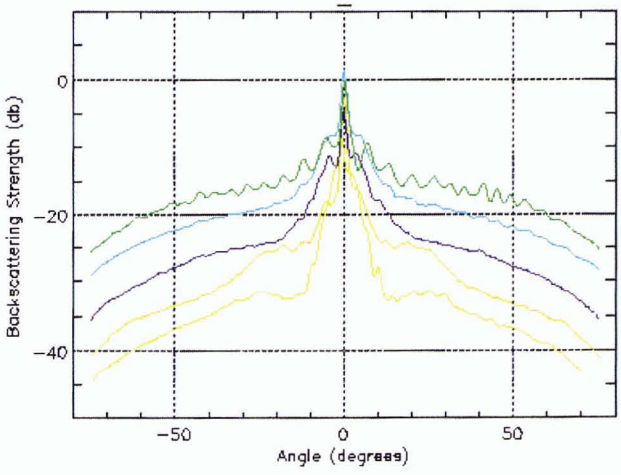


Figure 5. Backscattering Strength *vs* incident angle, for the five test zones used in segmentation