ecua 2010 İstanbul
5-9 July, 2010
www.ecua-2010-istanbul.org

Sonar images segmentation and classification fusion

Julien Lengrand-Lambert[1], Arnaud Martin[1], Romain Courtis[2]

[1]ENSIETA, 2 rue Francois Verny, 29200 Brest, France, {lengraju, Arnaud.Martin}@ensieta.fr
[2]GESMA, DT/GESMA/GDM, CC42 - 29240 Brest Cedex 9, France, Romain.Courtis@dga.defense.gouv.fr

Sonar images segmentation and classification fusion

Julien Lengrand-Lambert[1], Arnaud Martin[1], Romain Courtis[2]

[1]ENSIETA, 2 rue Francois Verny, 29200 Brest, France, {**lengraju, Arnaud.Martin**}**@ensieta.fr**
[2]GESMA, DT/GESMA/GDM, CC42 - 29240 Brest Cedex 9, France, **Romain.Courtis@dga.defense.gouv.fr**

This issue handles the ability of processing sonar images in order to automatically draw a map of undersea borders and classes. This study takes part of the project of automatic undersea cartography. Indeed, the current use of experts to analyse those images is very long and costs money. In addition, two experts have almost always different points of view on the same image. This work is based on an older software available of automatically classify images using a learning database. This database is based on the Haralick parameters of slices of images. Methods such as support vector machine give good classification results. However, there is a lack of precision in the border of the different classes. In order to enhance those results, we worked on a different way of segmentation using the "level-sets method". This method do not need any information concerning the classes and search for the highest gradient zones. As both methods have given good results, we tried to fuse them in order to clarify and precise the borders of classes in images. Finally, this work allows us to draw a conclusion concerning the quality of such a method and to explore deeper in this way.

1 Introduction

For the past several years, a great number of researches have been conducted on automatic segmentation of sonar images (*cf.* [4, 3] for example). Despite all those works, no common way has been found in order to segment sonar images automatically. The latter are still usually processed by experts which is both really expensive and not reliable. Experience shows indeed that two experts will most of the time disagree about sonar images interpretation (see figure 1). Finally, one must precise that those effects are consequences of the great constraints usually present on sonar imagery: propagation losses, lack of knowledge of the underwater environment and above all change of the sea floor typography. Sonar imagery is in brief a complex field by itself.

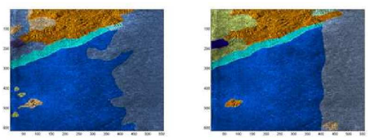


Figure 1: Segmentation processed by two different experts

Underwater environment has a growing interest in both military and industrial world, for applications such as mine hunting. Those fields would actually need a full suite of techniques automatically processing and segmenting sonar images. A first version which is based on texture characterization and includes automatic classification of images has been developed by [2]. However, as every automatic method some errors occur and results must be enhanced. This article suggest a fusion between two automatic segmentation and classification techniques.

Those two techniques are slightly different: both have the same input but give totally different results, hence the need of a step of fusion. Classification aims to divide the image in classes (sand, ripples, cobbles, rock, ...), using a previous knowledge. The segmentation conversely uses the same image but without any *a priori*. It is based on Level-Sets, highlighting the parts with a high gradient level. This technique does not need any information on classes but search for borders only. Thereby, one may try to combine the benefits of each of those two methods to meet the objective.

In this paper, we will firstly point out the two segmentation and classification methods we used. Then, we will present the fusion step and the ideas that came up. We will finally end up with some results.

2 Sonar images classification

A first automatic classification method has been developed by [2]. It allows to process sonar images and output them divided into classes depending on the type of sea floor. The method lies on a set a experts-classified images, which allowed to create a knowledge database. In this way, sediments and seabed types can be reliably characterized.

Feature extraction On sonar images, sediments are characterized using their texture. Three classes will be taken into account: sand, ripples and rocks. The texture is extracted in cutting the images in tiles and calculating parameters on them. Some approaches have been implemented [7]. The results presented in this paper use a method founded on confusion matrix proposed by [1]. According to this work, a sonar image is processed into 6 texture images (*cf.* figure 2).

Image classification [2] proposed different approaches of supervised classification. Some of them are support vector machines or simpler techniques as k-nearest neighbors. The latter already gives good results by itself. Classes are

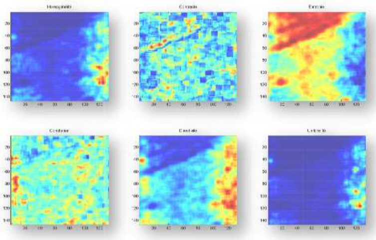


Figure 2: Example of parameters matrix on sonar image

globally reliable for simple images, as soon as the learning database is complex enough. However, the main drawback is its lack of accuracy, due to the division of the whole image in tiles. Thus, the parameters are mean values and cause loss in spatial resolution. In addition, this classificatio approach does not allow precise bordering between classes.

3 Automatic segmentation of sonar images

In order to enhance the previous results, we developed a new method of seabed segmentation in [5]. This method is not correlated with the presented classification and is based on a 'region approach'. The latter won't give any information on the seabed type, but only on the borders between the seabed types. Though, the information on the borders should be more accurate and give new information compared to the classificatio method. The automatic segmentation takes its source in the level sets method developed by [10]. The algorithm may be divided into three main parts:

- **Settings:** Here are chosen some of the 6 parameters to be used. It is the only human interaction in the processing chain.
- **Data preprocessing:** The firs step uses tile based images, *i.e.* image containing the calculated Haralick parameters. The images are modified and only the chosen parameters are extracted.
- **Segmentation:** The main part of this method applies the level-sets algorithm on the chosen matrix of Haralick parameters. Some processing is done to get images of borders with the same size as the input.

Here will be described the two last steps.

3.1 Data preprocessing

As it has a great impact of the following, this step is important for the method. If data are badly processed, the algorithm won't be able to clearly detect borders which will lead to poor quality of results. The preprocessing is divided into different parts, sequentially performed.

Choice of the parameters The very beginning is the choice of Haralick parameters to use for the processing. The results have different levels of quality, and some could not even be used. Roughly speaking, correlation and contrast are noisy. In addition, uniformity and homogeneity give the same kind of information in matter of segmentation, but uniformity seems attenuated. In order to optimize the time of processing, this parameter won't be taken into account. Finally the three chosen parameters during our study will be: homogeneity, entropy and directivity. One can of course choose other parameters presented in [7].

Segmentation of the chosen matrix Each image is firstl segmented into 4 levels using a Fisher segmentation (*cf.* [9]). This allows to greatly reduce the number of level-sets to be created during the segmentation, and enhance the accuracy in highlighting. Fisher's algorithm is based on the histogram. Four levels are performed in each image, with minimization of sum of inertia for each class as a criteria. Given an histogram and a number of classes to create, the algorithm automatically fin the separator between classes.

Formatting results In order to perform the segmentation, the amplitudes of matrix must be found in the same zones. It is indeed reverse for some parameters, and homogeneity and contrast matrix have high levels where entropy matrix is minimum. The outputs are thus modifie in order to give the maximum in the same areas. This step is mandatory as level sets will later fi on high levels areas.

Segmentation

Here is the main part of our segmentation method. It contains the level-sets algorithm, and is processed step by step.

Level-set processing Level-sets are used to search for forces evolution in images, and especially discontinuities. It allows to automatically fin zones of minimal discontinuity and creates close shapes as showed in [10]. Assuming that this shape mark out two zones with different evolutions some forces may be existing: a force in the normal direction of the curve, an external vector fiel force based on the curvature of the curve. It can be explicated with the following differential equation:

$$\frac{\partial f}{\partial t} + \overrightarrow{S}.div(f) + V_n.|div(f)| = V_n.|div(f)|$$

where the firs term is vector fiel based, the second one is normal direction based and the last one curvature based where a force f is applied on the curvature. In this paper, we use only the normal direction contribution, which can correspond to curvatures as following:

$$\frac{\partial f}{\partial t} = V_n.|div(f)|$$

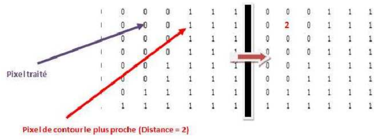


Figure 3: Flowchart on the building of the Gaussian distance matrix

It means that we are searching for high gradient zones in the image.

The level-set algorithm was almost not modifie during our study: it takes as input one of the three chosen matrix, and a given number of iterations, which will have a threshold role. The more iterations, the sharper the borders. This number must thus wisely be chosen, and may be considered as a filterin step. It allows to accurately and quickly determine the zones of great changes in the image. This smoothing will then be used as input to get the fina borders.

Gaussian distance matrix creation After having processed the level-set algorithm on each of the three matrix, one must now search for the fina borders. A smoothing and thresholding is performed in order to be as precise as possible and to keep only the most important borders. Those three matrix are then merged, using a gaussian distance matrix (see figur 3). The principle is pretty simple: on each point of a matrix is calculated the distance of the nearest borders. A 'distance matrix' is then created. Once the whole image is done, a gaussian distribution is applied to it. The three results are finall summed. In this way, one can easily detect the retrieval zones corresponding to correlated borders on several Haralick images.

Retrieval of the region matrix This is the last step of the automatic segmentation method. The gaussian distance matrix contains all the fina borders. One must now modify it to be correctly used later for fusion. The borders are firstl transformed into regions. The last part is to get exact borders. In fact, the latter are still several pixels broad because of the gaussian distribution.

This segmentation method is totally uncorrelated with the former classificatio and gives new information concerning the seabed. Given the results, the fina borders are more accurate than with the classification However, no information is given concerning the type of seabed. After manual checking, the results seem to be pretty good for sand and rock bottoms, but ripples tend to be hidden by the algorithm.

Besides, the apparent quality of the two presented methods are a good argument in favor of a fusion. The latter will now be described.

4 Fusion method and algorithm

Given the results of classificatio and segmentation, we worked on a method which would allow to both have the benefit of 'region' based and 'class' based methods. In addition, it had to stay full automatic. In this third part, we suppose classificatio and segmentation already processed. The aim is to get a fina image as close as possible of reality only with the two last outputs. The following process may be divided into two main parts: fina borders retrieval and class determination of the fina regions.

4.1 Final borders retrieval

The firs step of data fusion aims to take the segmentation and classificatio results and merge them to come up with the best borders as possible. Once the borders are found, the job will be to search for classes of the created regions.

In order to perform this step, the same method as for the segmentation will be used. Thus, the two images are thresholded using Fisher's algorithm and level-sets, followed by the creation of the gaussian distance matrix. The fina borders may finall be found searching for the highest levels in the sum of the gaussian distance matrix. In this way, the use of our previous work allows to directly highlight zones that will be classifie next.

4.2 Regions classification

We ended up in the last part with a matrix whose borders are finall smooth enough. But the regions have no label yet, and we still have no concrete information concerning the sea floo type. The Haralick parameters will once again be useful to classify the regions, using a supervised classification Two main steps are needed to perform this stage: searching a vector characterizing each region, and the fina processing of findin the corresponding class for this vector.

Search for a characterization vector Each pixel of a given region is fully represented by a vector of Haralick parameters. The spatial mean is processed in order to fin the corresponding vector or this region. This mean vector is considered as a representation of the given texture of this part of the image.

Regions classification In order to efficientl classify the region using the mean vector, some approaches may be considered. Two of them will be presented here:

- The k nearest neighbors method, which is searching for the minimum distance between the vector and a learning base
- The prototypes method, a special conception of the latter, using means of vectors for each class.

The firs tests have been processed using the simple k nearest neighbors method (*cf.* [8]). Finally, it came up that this algorithm gave the most reproducible and precise results. Thus, the k nearest neighbors vector are searched among the learning database. The k parameter has been chosen between 3 and 5 during our study. The distance used for calculating neighborhood is euclidean distance, which fit well with our space of work. After having found the neighbors, the last part consists in assigning the most represented class in the sample.

The prototype method is closely related with the previous one. Instead of representing each learning vector in the space, a mean is done between vectors of the same class. In this way, a full class is only represented with a single sample. the region is given the class which distance is the closest to the sample.

Before presenting the fina results, one must precise that the full processing is quite fast, excepting the creation of the learning database which is time consuming for the firs tests. The full processing is indeed performed in average 30 second with a common computer. One can imagine the results that may be achieved in using such a system directly on a sonar. The results of the two previous methods, and their parameters will be developed in the next part.

5 First results

In this section, we will present one after the other every progress we achieved using our methods. The quality of the results may be appreciated looking at the confusion matrix which compares the fina fusion with the segmentation coming from an expert.

The 42 sonar images have been produced by the GESMA (Groupe des études sous-marines de l'Atlantique) during a campaign offshore of Bretagne. They come from a Klein 5000 sonar and their resolution is up to 30cms azimut and 3cms range. The seabed is 15 to 40 meters deep.

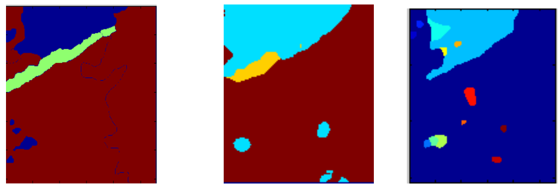
Using a supervised classificatio laid us to divide the dataset into two parts: the learning database, which will allow us to apply our algorithm and have information about the sea floo and the test database which data will be classified The learning database contains 39 sonar images. All of them have been manually segmented, which allowed us to generate about 6700 tiles. The latter have a squared shape and are 32 pixels wide. One must pay attention that the bottom of the sea is not fully equally distributed between the different sediment layers. In this way, some of the classes are over represented in front of the others. In order to get rid of this effect, we chose to use only the same number of tiles for each class. Finally, each of the three classes (ripples, sand, rocks) hold 2034 tiles. This number seems big enough to lead to reliable results, but may later be completed.

Here is the color code for the following images:

- **Rocks**: blue color
- **Ripples**: green color
- **Sand**: red color

5.1 Segmentation and classification

The firs step of the test phase is the validation of the automatic segmentation and classificatio steps.



Expert segmentation Classificatio Segmentation

Figure 4: Automatic segmentation results

Automatic classification The results on figur 4 have been produced using the automatic classificatio method previously described. One can obviously see similarities between the expert segmentation and the automatic classification The shapes are correlated, and the three main classes are there. Those results may be completed in having a look at the confusion matrix (*cf.* table 1). The percentage of good detection for rocks and sand are really good (up to 80%). However, the error rate is highly raising for the ripple class (more than 60% of error). Thereby, the detection method is not fully reliable concerning the ripple class. it is however good enough for sand and rocks to process the fusion.

	ripples	rocks	sand
ripples	1424	2243	298
rocks	298	13336	2243
sand	414	11158	56075

Table 1: Confusion matrix for the automatic classificatio

The confusion matrix for the automatic classificatio given in Table 1 lead to the following percentages:

- **Rocks**: 84% of good detection
- **Ripples**: 64% of error (the ripple zone is cut off in the classifie image)
- **Sand**: 82.9% of good detection.

Automatic segmentation Before performing the fusion step, the automatic detection still has to be checked. However, statistical results cannot be achieved in this part as a validation tool. As previously described, the region approach does not give any information concerning the seabed type and only borders are represented. A validation approach of the segmentation step has however been presented in [6].

The fina image can be found in figur 5. Each different color on the image is associated with a different region. As expected,the regularity of the borders is less present than is the previous part. In addition, they seem to fi with the expert segmentation. Our approach may thus be validated.

In an ideal way, one may merge the two methods in order to have the benefit of both borders accuracy and seabed information.

5.2 Fusion of automatic segmentation and classification

Expert segmentation k nearest neighbors prototypes

Figure 5: Results for the fusion of the two previous approaches

k nearest neighbors approach The results using this method are presented on the second figure The general bearing of the fusion matrix stays correlated with the expert segmentation. The three classes are present, and seem well placed. Some differences may however be highlighted. Rocks are fully represented in the top left corner, but some layers in the bottom are missing. The algorithm failed to detect them. In addition, Ripples are represented, but only with a little blob. The major information loss is contained in this ripple zone that is represented as sand.

The confusion matrix gives precise statistics of the results(*cf.* table 2).

	ripples	rocks	sand
ripples	374	560	52
rocks	4	2489	1101
sand	30	639	14023

Table 2: Confusion matrix for the knn algorithm

The confusion matrix for the knn algorithm given by the table 2 gives the following error rates:

- **Rocks**: 69.2% of good detection
- **Ripples**: 62% of error (big parts are still missing)
- **Sand**: 95.4% of good detection

Prototypes approach The results using this method are presented on figur 5. The results are completely non-sense. This is caused by the variance intra-class of the Haralick parameters. Performing the mean of the vectors implies a loss of information about space dispersion. Thus, the results are totally unusable. A simple look at the image (*cf.* figur 5) speaks by itself.

	ripples	rocks	sand
ripples	0	392	3202
rocks	106	134	746
sand	220	332	14140

Table 3: Confusion matrix for the prototypes approach

The confusion matrix for the prototypes approach given by the table 3 gives the following error rates:

- **Rocks**: 24.3% of good detection (the zones are present but misplaced)
- **Ripples**: 100% of error (no ripples on the image).
- **Sand**: 96.2% of good detection.

After having displaying those results, the k nearest neighbor approach seems to fi better with the fusion method. The results are indeed consistent and seem close enough to reality. The confusion matrix for this method is far better than the one with prototypes. However, if sand and rocks have good statistics, some progress have to be made with the ripple class.

6 Conclusions

In this paper, we both presented an automatic segmentation approach based on level sets and a new method for fusion of segmentation and classificatio of sonar images. Those preliminary results show the progress that can be performed in this field The bottom of the sea may thus be fully automatically classified In addition, considering the real-time aspect of this method, one may hope an easy implementation directly on a sonar system. Those results are not good enough yet to constitute a fina system, but they may be opening gates for future researches. One may hope to be able to both merge the good performance of sea floo characterization with accuracy of its borders.

This work gives a preview for further researches. One can as an example imagine more complex approaches: belief functions could be a good pretender. Finally, we would conclude saying that an expert is still mandatory for the learning phase. Those experts are giving advices about their level of confidenc on images, and that could be used as a weight for supervised approaches. Non-supervised classificatio methods could be investigated too in the future in order to get rid of the expert.

References

[1] R. Haralick. Statistical and textural approaches to textures. *Proceedings of the IEEE*, 67(5):786–804, 1979.

[2] Hicham Laanaya. *Classification en environnement Incertain : Application la Caractérisation de Sédiments Marins*. PhD thesis, Université de Bretagne Occidentale, ENSIETA, Brest, 14 December 2007.

[3] Gilles Le Chenadec and Jean-Marc Boucher. Sonar image segmentation using the angular dependence of backscattering distributions. In *IEEE Oceans'05 Europe*, Brest, France, 2005.

[4] Isabelle Leblond, Michel Legris, and Bassel Solaiman. Use of classificatio and segmentation of sidescan sonar images for long term registration. In *IEEE Oceans'05 Europe*, Brest, France, 2005.

[5] Julien Lengrand-Lambert, Arnaud Martin, and Romain Courtis. Fusion de segmentation et classifica tion automatique d'images sonar. In *Atelier Fouille de données complexes, conférence Extraction et Gestion de Connaissances*, Hammamet, Tunisia, January 2010.

[6] Arnaud Martin, Hicham Laanaya, and Andreas Arnold-Bos. Evaluation for uncertainty image classificatio and segmentation. *Pattern Recognition*, 39(11):1987–1995, Novembre 2006.

[7] Arnaud Martin, Gwénola Sévellec, and Isabelle Leblond. Characteristics vs decision fusion for seabottom characterization. In *Journée d'Acoustique Sous-Marine*, Brest, France, 2004.

[8] Sylvain Pasini and Bertrand Grandgeorge. Image segmentation. *Projet en Digital Photography - Image Segmentation*, 1:5–9, 2003.

[9] John C. Russ. *The image processing handbook*. CRC Press, Cleveland, 2002.

[10] Baris Sumenger. Level set method presentation. Technical report, Vision Research Lab, UCSB, 2005.