

---

## **Spatial Statistics of Objects in 3-D Sonar Images: Application to Fisheries Acoustics**

R. Lefort<sup>1, 2\*</sup>, R. Fablet<sup>2</sup>, L. Berger<sup>1</sup>, J.-M. Boucher<sup>2</sup>

<sup>1</sup> Ifremer, Technopol Brest Iroise, 29280 Plouzane

<sup>2</sup> Telecom Bretagne, Technopol Brest Iroise, 29280 Plouzane

\*: Corresponding author : Riwal Lefort, email address : [riwal.lefort@telecom-bretagne.eu](mailto:riwal.lefort@telecom-bretagne.eu)

---

### **Abstract :**

In this letter, we address the characterization of objects in 3-D sonar images of the water column obtained by a multibeam echo sounder. Compared with classic 2-D images from a monobeam echo sounder, these 3-D images provide finer scale observation of the pelagic biomasses and new tools to characterize 3-D distributions. By viewing object patterns as realizations of spatial point processes, we investigate descriptive spatial statistics. This method is then applied to 3-D fisheries acoustics data set for characterization of the distribution of pelagic fish schools. Reported experiments illustrate the relevance of the proposed descriptors. The comparison of our method with 2-D sonar data analysis further demonstrates the information gain from using 3-D sonar imagery.

**Keywords :** Object patterns in images, spatial statistics, point processes, multibeam sensor, fisheries acoustics

## 1. Introduction

---

Monitoring marine ecosystems is a major issue in the current context of global environmental change. Echosounder systems provide a remote sensing device to monitor the distribution of pelagic environment, typically plankton and fish distribution [2]. The echosounder emits an acoustic pulse that is backscattered by objects of the water column and provides an acoustic image of the distribution of the pelagic biomass. Besides these traditional echosounders that form 2D images of the pelagic environment by utilizing depth and vessel displacement information, new multibeam systems offer finer-scale 3D images with an additional transversal dimension. Figure 1 depicts typical examples of the 3D structures which appear more complex with a multibeam echosounder than captured through a 2D echosounder. For instance, a single aggregation in 3D images is typically viewed as a collection of small patches in 2D images. Figure 2 illustrates the variety of aggregation patterns that can be observed, including well-defined schools, patchy aggregations, diffuse layers, etc. Whereas the processing of fisheries acoustics data have mostly dealt with the characterization and classification of fish schools in 2D [3], these 3D observations emphasize the need for new descriptors characterizing both fish schools and their spatial distribution. The main contributions of this paper are therefore to adapt point process statistics to marked point process (section II) and to take into account the trapezoidal geometry of 3D sonar echoes (section III). This method is applied to real fisheries acoustics data (section III). We also carry out a quantitative evaluation of the proposed descriptors for the categorization problem (section IV).

## 2. Descriptors of the spatial distribution of objects in images

---

### A. Spatial patterns formed by objects in images

In our work, we assume that both object extraction and associated object characterization result from an appropriate image pre-processing step. In order to do this, let us consider a set of  $M$  processed images. We define the image index  $m$  such that  $1 \leq m \leq M$ , and the object index  $n$  in image  $m$  such that  $1 \leq n \leq N(m)$ . The position of any object  $x_{mn}$  is defined in the Euclidean space  $\mathbb{R}^L$ , where  $L$  denotes the image dimension (typically  $L = 2$  or  $L = 3$ ). Let us denote the feature vector describing object  $n$  in image  $m$  by  $f_{mn}$ . Object sets  $\{x_{mn}; f_{mn}\}_{m,n}$  are viewed as realizations of random marked point processes. A random point process is defined in [4] as follows: a point process  $\mu$  on  $\mathbb{R}^L$  is a measurable map from a

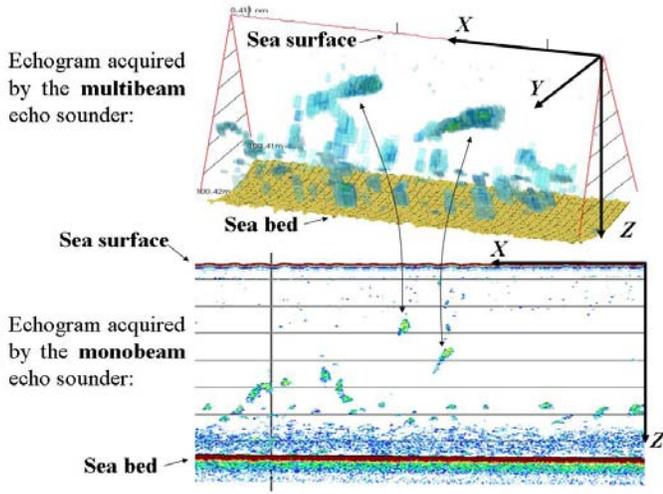


Fig. 1. The same area is viewed by a multibeam echosounder (top) and by a monobeam echosounder (bottom). The 3D image (multibeam echosounder) is at finer scales and provides description of fish schools in the water column.

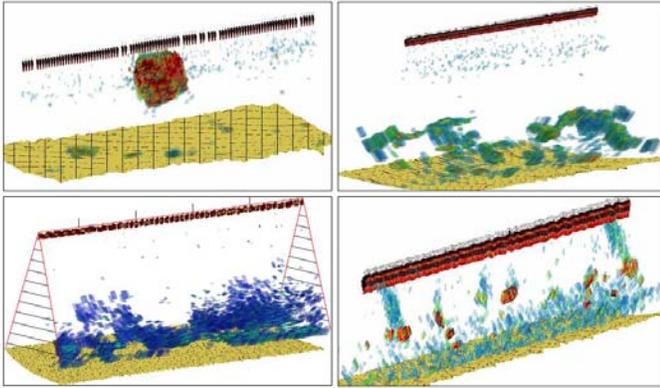


Fig. 2. Examples of multibeam acoustic observations depicting different types of distributions of fish aggregations: dense ball-like sardina aggregation (top left), patchy anchovy distributions (top right and bottom right), and diffuse blue whiting with low-density structures (bottom left).

probability space  $\{\Omega, \mathcal{F}, \mathcal{P}\}$  to a measurable space  $\mathcal{R}$ . For instance, a realisation of point process  $\xi$  in image  $m$  can be represented as:

$$\xi_m = \sum_{n=1}^N \delta_{X_{mn}} \quad (1)$$

where  $\delta_X$  denotes the Dirac measure centered on  $X$ ,  $N$  is an integer-valued random variable and  $X_{mn}$  are random elements in  $\mathbb{R}^L$ .

Our objective is to characterize the probability space  $\{\Omega, \mathcal{F}, \mathcal{P}\}$  from observation sets  $\{x_{mn}\}_{mn}$ .

We rely on descriptive statistics computed as moments of the random processes; more particularly second-order moments associated to Ripley's  $K$ -function [5].

### B. Descriptive statistics of point processes

To characterize a point process, the first-order moment, that describes the intensity of a homogeneous point process, is defined along the lines of [5] as:

$$K = \int_B \rho(v) dv \quad (2)$$

where  $\rho(v)$  is the probability density function of the number of points in an infinitesimal volume  $dv$  and  $B$  denotes the support of analysis. This first-order moment does not however characterize the spatial interactions between points. Hence, second-order moments are used [6] [7]. In our case, we consider the covariance structure of the count variable, i.e. descriptive statistics of pairs of points of finite random sets which is given by:

$$K = \frac{1}{V} \int_V \int_B \rho^{(2)}(x_1, x_2) dx_1 dx_2 \quad (3)$$

where  $\rho^{(2)}(x_1, x_2)$  is the probability density function of the pairs of points in infinitesimal volumes  $dx_1$  and  $dx_2$ , and  $V$  is the whole volume of interest.

In the case of an isotropic and stationary process, the density  $\rho^{(2)}(x_1, x_2)$  only depends on the distance between points ( $\|x_1 - x_2\|$ ). Defining regions  $B_r$  as spheres parameterized by their radius  $r$ , we use the Ripley's  $K$  function whose empirical estimation from a given realization of the point process is given by [6] [7] as:

$$K(r) = \frac{1}{V} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \mathbb{1}_{B_r}(x_i, x_j) \quad (4)$$

where  $\mathbb{1}_{B_r}(x_i, x_j) = 1$  if  $\{x_i, x_j\} \in B_r$ , and  $\mathbb{1}_{B_r}(x_i, x_j) = 0$  if not. Note that  $B_r$  is centered in  $x_i$ .

### C. Marked point process

The second-order moment defined above can be extended to marked point process. Let  $\{x_n, y_n\}_{1 \leq n \leq N}$  be a marked point process, where  $\{x_n\}_n$  refers to point positions and  $\{y_n\}_n$  to discrete marks ( $y_n \in \mathbb{N}$ ) encoding relevant features of the objects. We consider that a mapping from

object feature vector to object categories exists and that these categories define the discrete marks. In this work, this is achieved using unsupervised clustering in the feature space using the k-means algorithm [8] [9].

Ripley's  $K$ -function is extended to marked point process as follows. For any pair of object categories, an associated Ripley's  $K$ -function is computed, i.e. the expected number of points of one category in a sphere centred at a point of the realization of another category is found. Formally, for the object categories  $p$  and  $q$  and the analyzing radius  $r$ , the considered descriptive statistics is given by:

$$\Gamma_{p,q}(r) = \sum_i^N \frac{1}{V_i(r)} \sum_{j=1, j \neq i}^N \mathbb{1}_{B_r}(x_i, x_j | y_i = p, y_j = q) \quad (5)$$

$V_i(r)$  is a normalization coefficient that expresses intersection between the volume  $V$  and the sphere  $B$ .

These second-order spatial statistics can be viewed as spatial cooccurrence statistics of object sets in images. Parameterized by a spatial radius and object category pairs, they provide a joint characterization of a set of objects in an image, both in terms of object characteristics and of spatial organization.

### III. APPLICATION TO FISHERIES ACOUSTICS DATA

Data processing tools for fisheries acoustics data have mainly focused on school detection, characterization and classification [10] [11] [12] [13] [2]. The development of school extraction software made the data processing easier [14]. Other analysis scales, especially school clusters, are also of key interest to understand and characterize fish distribution. However, they remain weakly explored. Along the lines of [15] [16], we address here the characterization of school clusters at a typical scale of one-nautical-mile-long echograms [2] [17]. We evaluate the feasibility of the proposed spatial statistics of fish schools w.r.t. previous work, especially [16], for an echogram-classification task.

#### A. Proposed approach

In the reported experiment, the extraction of fish schools is issued from a thresholding-based

detection scheme. Let us denote the position of the extracted schools by  $x_{mn}$  and the associated feature vectors by  $f_{mn}$ . Extracted features are the length, height, width, volume, and backscattering strength. They comprise morphological features (size and water column position) and energy features (back scattering strength). These features are categorized by an unsupervised k-means method to obtain feature marks  $y_{mn}$ . Spatial statistics of school sets are then computed for each echogram. The analysis of the distribution of the distance between fish schools revealed a bimodal distribution. For properly incorporating the two modes, two radius values were chosen:  $2.5m$  and  $10m$ .

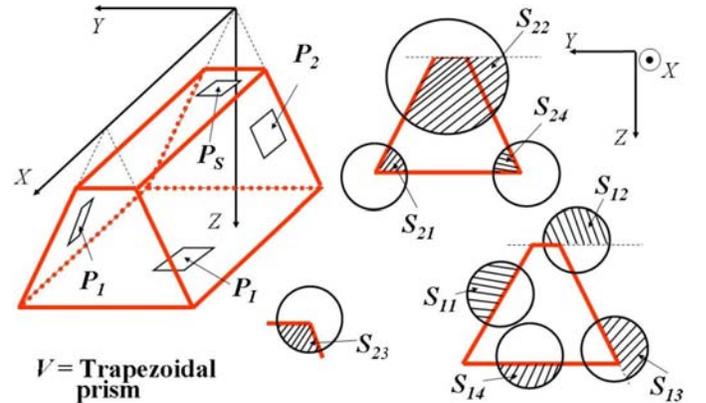


Fig. 3. The intersection volume  $V_i(r)$  is calculated between the trapezoidal prism  $V$  (left) due to multibeam geometry and sphere  $B(r)$  localized in  $x_i$  with radius  $r$ . The definition of elementary area  $S_{uv}$  (right) allows to generalize all situations.

#### B. Corrections for echosounder geometry

In Eq.(5),  $V_i(r)$  is a normalization coefficient that expresses the intersection between the volume  $V$  and sphere  $B$ , the center of sphere  $B$  being inside  $V$ . We should take here into account the geometry of the images acquired by echosounders [6] (refer Figure 2). In 3D, volume  $V$  is a trapezoidal prism which leads to complex formulation for coefficient  $V_i(r)$ . One could set  $V_i(r)$  to constant value but this causes underestimation for  $\Gamma_{p,q}(r)$ . Objects close to image boundaries would be weighted low because of the empty area outside the trapezoidal prism  $V$ . Consequently, the mean number of points in the sphere  $B$  close to boundaries of volume  $B$  would be lower than the mean number of points closer to the center. Next, we propose a general expression

for the computation of volume  $V_i(r)$  according to the geometry of multibeam echosounder.

As illustrated in figure 3, we define the following variables:

- $P_1, P_2, P_S, P_I$ : planes that define the trapezoidal prism.
- $S_{11}, S_{12}, S_{13}, S_{14}$ : exterior surfaces of sphere  $B$  relative to planes.
- $S_{21}, S_{22}, S_{23}, S_{24}$ : interior surface of sphere  $B$  relative to plane intersection.

The intersection area  $A_i(r)$  (in the plane  $\{Y, Z\}$ ) between disk localized in  $x_i$  with radius  $r$  and trapezoid (in the plane  $\{Y, Z\}$ ), can be expressed as a function of  $\{S_{uv}\}_{1 \leq u \leq 2, 1 \leq v \leq 4}$ :

$$A_i(r) = -3A(r) + \sum_{u=1}^2 \sum_{v=1}^4 S_{uv} \quad (6)$$

where  $A(r) = \pi r^2$  denotes the disk area with radius  $r$ . If the sum of  $\{S_{uv}\}$  is computed, by considering all the possible intersections that depend on the disk localization and on the disk radius,  $3A(r)$  must be subtracted to find the intersection area. The integral over the third dimension  $X$  gives the final intersection volume:

$$V_i(r) = \int_X A_i(r(x)) dx \quad (7)$$

Equation (7) and the definition of  $A_i(r)$  provides a general expression for the computation of Ripley's  $K$  statistics for any sphere  $B$  intersecting the trapezoidal prism  $V$  and for any radius  $r$ .

## IV. EXPERIMENTS

### A. Dataset

The considered fisheries acoustics dataset involves three classes of echograms: (a) 63 echograms comprising large and dense schools of sardines, (b) 72 echograms depicting a mixture of anchovy and horse mackerel with complex school shapes, (c) 87 echograms depicting scattered schools involving a mixture of anchovy and horse mackerels. All echograms were obtained from a scientific survey carried out by Ifremer<sup>1</sup> in June 2008 in the Bay of Biscay. This dataset was acquired with both a 3D and a 2D echosounder, so that we can evaluate the relevance of the proposed descriptors of school clusters for both 3D and 2D echograms.

<sup>1</sup>French Research Institute for Exploitation of the Sea.

### B. Echogram classification

In this section we evaluate the relevance of various echogram descriptors for the echogram classification task. We use 50-fold cross-validation, and in each run we split the dataset into a class-balanced training set and a class-balanced test set. Next, we compute classification statistics (mean and variance) for the positive class. In this work, we use random forests [1] as the baseline classification tool.

We compare the following echogram descriptors:

- *Sea bed depth* (referred to as "Depth"). All descriptors are correlated to seabed depth. We propose to use seabed depth as a baseline reference to check for the relevance of the improvement brought by additional features.
- *Global descriptors proposed by Burgos and Horne* [16] (referred to as "Burgos"): the school density in the image, the percentage of spatial occupancy, the fragmentation index, the median distance, the relative school area in the image, the 10th-50th-90th percentiles of the area, the depth, and of the acoustic density (Sv) of the schools.
- *Cooccurrence Ripley's statistics* (referred to as "Ripley"). We compute the spatial statistics detailed in Section II.

For the experiments, we considered two thresholds for the extraction of fish schools (-60dB and -54dB) based on the work of experts. We concatenate the descriptors computed for the school sets associated with each threshold to form the feature vector which is subsequently used for echogram classification.

Results are reported in Table I for both multibeam (3D) and monobeam (2D) data. The mean correct classification rate is reported as a function of the number of training instance.

Overall, the correct classification rate increases with the increase in the size of the training set. This shows that a larger number of annotated images leads to better echogram classification. These experiments also point out the advantage of using the multibeam technology. For all descriptors, an increase in correct classification rates is observed (7% mean improvement with Burgos to 12% mean improvement with the proposed spatial statistics). Multibeam technology offers an actual 3D visualization of school structures whereas monobeam echosounder provides a coarser 2D transversal observation of 3D structures. In particular, schools

observed in the 2D echograms often refer to cluster of small 3D schools that can not be perceived due to the lower resolution of the monobeam echosounder. Then, the 3D echogram achieves better characterization and discrimination of school clusters. Though echogram categorization is partially correlated to seabed depth, we show that the proposed echogram characterization significantly outperforms the geographical categorization of echograms from seabed depth (typically the correct classification rate is improved from 10% to 25%).

Regarding the different feature sets, those proposed by Burgos and Horne [16] outperform spatial statistics of school clusters when considering 2D echograms. That might be explained by the rather low mean number of fish schools. We had about 16 fish schools per 2D echogram compared to 168 fish schools per 3D echogram. This makes the computation of spatial school statistics less robust. By contrast, experiments with 3D echograms show that the best performances are obtained using the proposed spatial statistics (91% vs 94%). This validates the utility of the proposed descriptors.

	2D				3D			
	30	60	90	120	30	60	90	120
Training	30	60	90	120	30	60	90	120
Test	159	129	99	69	159	129	99	69
Depth	0.71	0.72	0.72	0.71	0.71	0.72	0.72	0.71
[16]	<b>0.79</b>	<b>0.83</b>	<b>0.85</b>	<b>0.87</b>	0.89	0.90	0.91	0.91
This work	0.76	0.79	0.83	0.84	<b>0.91</b>	<b>0.93</b>	<b>0.94</b>	<b>0.94</b>

TABLE I

Mean correct classification rates for both multibeam sensor (3D images) and moonbeam sensor (2D images). The row "Training" contains the total number of training instances. The row "Test" contains the total number of test instances. Results are obtained with different descriptors: seabed depth, [16], proposed spatial statistics (This work).

## V. CONCLUSION

The development of global image descriptors is a key topic for remote sensing applications. In this context, we have proposed global descriptors for object distributions in an image. They rely on spatial statistics of the patterns formed by object sets extracted in images.

This approach has been applied to fisheries acoustics datasets involving 2D and 3D images of clusters of fish schools. We have evaluated the proposed descriptors for an image classification task and have

proven its relevance to previous work. Reported experiments also stress that finer characterization of fish aggregation patterns can be achieved upon using multibeam 3D echosounder instead of the classical 2D echosounder. These results open door for further investigation of the spatial organization of fish aggregations.

## REFERENCES

- [1] L. Breiman, "Random forest", *Machine Learning*, 45, 5-32, 2001.
- [2] J.E. Simmonds, and D.N. MacLennan, "Fisheries acoustics: theory and practice", *Oxford: Blackwell Science Ltd*, 2005.
- [3] R. Lefort, R. Fablet, and J.M. Boucher, "Object recognition using proportion-based prior information: application to fisheries acoustics", *Pattern Recognition Letters*, vol. 32(2), pp. 153-158, 2011.
- [4] M. Jacobsen, "Point processes theory and applications", *Springer, Probabilities and its Applications*, 2006.
- [5] B.D. Ripley, "Spatial statistics", *John Wiley, New York*, 1981.
- [6] F. Goreaud, and R. Plisserier, "On explicit formulas of edge effect correction for ripley's k-function", *Journal of Vegetation Science*, vol. 10, pp. 433-438, 1999.
- [7] M. Schlather, "On the second-order characteristics of marked point processes", *Bernoulli*, vol. 7, pp. 99-117, 2001.
- [8] S.P. Lloyd, "Least square quantization in PCM", *IEEE Transactions on Information Theory*, vol. 28(2), pp. 129-137, 1982.
- [9] N. Dala, and B. Triggs, "Histograms of oriented gradients for human detection", *International Conference on computer Vision and Pattern Recognition*, vol. 2, pp. 886-893, 2005.
- [10] G.A. Rose, and W.C. Leggett, "Hydroacoustic signal classification of fish schools by species", *Canadian Journal of Fisheries and Aquatic Sciences*, pp. 597-604, 1988.
- [11] C. Scalabrin, and X. Lurton, "Fish schools amplitude analysis", *European Conference on Underwater Acoustic*, vol. 2, pp. 807-814, 1994.
- [12] C. Scalabrin, N. Diner, and A. Weill, "Automatic shoal recognition and classification based on Movies-B software", *Oceans '94*, vol. 2, pp. 319-324, 1994.
- [13] D. Reid, "Report on echo trace classification", *ICES Cooperation Research Report 238*, 2000.
- [14] A. Weill, C. Scalabrin, and N. Diner, "Movies-B: an acoustic detection descriptor software, Application to shoal species' classification", *Aquatic Living Resources*, vol. 6, pp. 255-267, 1993.
- [15] P. Petitgas, and J.J. Levenez, "Spatial organisation of pelagic fish: echogram structure, spatio-temporal condition and biomass in Senegalese waters", *ICES Journal of Marine Science*, vol. 53, pp. 147-153, 1996.
- [16] J.M. Burgos, and J.K. Horne, "Characterization and classification of acoustically detected fish distribution", *ICES Journal of Marine Science*, vol. 65, pp. 1235-1247, 2008.
- [17] M. Woillez, J.C. Poulard, J. Rivoirard, P. Petitgas, and N. Bez, "Indices for capturing spatial patterns and their evolution in time, with application to European hake (*Merluccius merluccius*) in the Bay of Biscay", *ICES Journal of Marine Science*, vol. 64, pp. 537-550, 2007.