Minimum Description Length for shape-based segmentation of seabed images

Laure Amate[(1)], **Maria-João Rendas**[(1)]
[(1)] *Laboratoire I3S, 2000 route des lucioles, Les Algorithmes BP 121, 06903 SOPHIA ANTIPOLIS, FRANCE*
E-mail: `amate@i3s.unice.fr`
E-mail: `rendas@i3s.unice.fr`

SHORT ABSTRACT: *The paper is a contribution to the problem of producing sensor-specific seabed maps, that describe the morphological features of each different region in the area considered. It applies the Minimum Description Length (MDL) principle to identify the dimensionality of a statistical shape model that describes them. Our shape model combines the finite-dimensional representation of continuous curves with the notion of shape space as introduced by Kendall to obtain a finite representation of closed curves with a prespecified set of invariances. The work presented can be considered as an extension of previous work by other authors which use MDL to adjust the complexity of a spline representation to a single contour present in an image, while we use MDL to adjust the complexity of a* statistical model *for the spline representation of a* set *of contours.*
Keywords: *Minimum Description Length, spline representation, ocean mapping.*

RÉSUMÉ COURT: *Cet article est une contribution au problème de production des cartes des fonds marins qui décrivent les informations acquises avec des capteurs perceptuels. Il utilise MDL pour déterminer la dimension d'un modèle statistique de forme qui décrit la morphologie des objets détectés par le capteur. Il combine la représentation de dimension finie de courbes fermées continues par des splines avec la notion d'espace de formes introduite par Kendall pour obtenir une représentation finie avec des proprétés d'invariance. Le travail présenté peut être considéré comme une extension de travaux précédents par d'autres auteurs qui utilisent le MDL pour ajuster la complexité de la représentation par des splines d'un seul contour présent dans une image. Nous utilisons le MDL pour ajuster la complexité du* modèle statistique *d'une représentation par des splines d'un* ensemble de contours.
Mots-clés: *MDL, spline, cartographie sous-marine.*

1 INTRODUCTION

Most Autonomous Underwater Vehicles rely on inertial sensors to navigate and to position themselves. The drift of gyroscopes and accelerometers is in this case a well-known problem that is overcome by ressorting to auxiliary external systems such as GPS or acoustic beacons. These solutions have drawbacks: frequent returns to surface for GPS fixes, previous preparation of the operating area for the installation of acoustic beacons. A different way to overcome positioning drift consists in using a map of the workspace which the robot can use as a reference to estimate its position.
Map-based positioning requires that the map elements be stable (re-observable) and non ambiguous (identifiable). Point-like features such as strong Sonar returns do not usually satisfy these conditions. Alternatively, some authors consider 'map elements' that effectively group several individual echoes, in an effort to increase the robustness of perception to map associations. The work we present here is somehow related to this approach, by considering the local collective characteristics

of the signals returned from the bottom as the stable information on which positioning can be based. Our approach considers long range navigation exploiting the extended geometric features of the environment as the major map elements. More precisely, we identify these macroscopic features with the contours between regions of the seabed of distinct perceptual appearances. Work along these lines using a vision sensor and simplistic shape models has been presented in [1].
In this paper we address the problem of producing this kind of maps from sea bottom (side-scan) images. More precisely, we will be concerned with producing a segmentation of side-scan images, and to associate a statistical geometric model to each region. Previous work on image segmentation (identification of *distinct* regions) has mainly relied on contrast measures for the statistical properties of the *pixels* belonging to each region. Here, we propose to identify a region by the similarity of the shapes of the shadows of the side-scan image inside it, as well as homogeneity of their spatial scattering, as we discuss in detail in section 2.
To adapt the complexity of models used to describe shape properties, we consider description of each contour by splines, and treat the control points of the spline representation as the *landmarks* that describe the object's shape. We use the Minimum Description Length (MDL) principle introduced by Rissanen in [2], to determine the number of knots, in the statistical spline representation, required to appropriately describe the collective properties of the set of shapes found in each region. The derivation of the MDL criterion for the number of spline control points is presented in section 3. Finally, we present results obtained for a set of contours extracted from real side-scan Sonar data.

2 STATISTICAL (APPEARANCE) MAPS OF THE OCEAN BOTTOM

This section presents the map model underlying our work. The mapped area $\mathcal{A}$ is partitioned in R disjoint regions $\mathcal{A}_r$:

$$\mathcal{A} = \bigcup_{r=1}^{R} \mathcal{A}_r \ , \tag{1}$$

and a set of morphological parameters θ_r is associated to each region. In the context of this paper, $\mathcal{A}$ is the footprint of the sensor (a side-scan sonar) in the sea-bed, see Figure 1, and we will not address here the problem of actual registration of the image regions to a global world coordinate system. Definition of the individual ("homogeneous") regions is thus a problem of *segmenting* the side-scan images. Many approaches have been proposed in the scientific literature for the notion of "homogeneity" of a region (distribution of the pixel intensities, parametric models, apparent motion, ...). We propose here to base homogeneity of two regions in the similarity of the shape and the pattern of spatial scattering of the individual "objects" found inside them. The next section formally presents the mathematical model necessary to correctly express this intuitive notion of similarity, indicating what is the exact meaning of the morphological parameters χ_r that characterize each region in Equation (1).

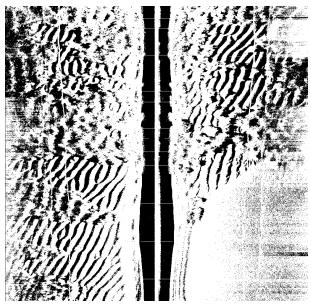

Fig. 1: Example of a side-scan Sonar image.

2.1 Random Closed Set models

Let $\mathcal{S}$ be a compact region in Cartesian m-dimensional space. The basic model of random (binary) fields on $\mathcal{S}$ on which our approach is based relies on the combination of two distinct random processes: a *point* process P , whose realisations P_j are random sets of isolated points in $\mathcal{S}$:

$P_j = \{s_i^j\}_{i=1}^{N_j}, s_i \in \mathcal{S}$, and a *shape* process, whose realisations S_i are random compact sets, see [3] for a detailed presentation of these models. A realisation M_j of the random field model is obtained by placing, at each realisation of the point process, a random shape:

$$M_j = \bigcup_{i=1}^{N_j} S_i^j \oplus s_i^j \ . \tag{2}$$

These models effectively describe the appearance of natural scenes in a broad domain of applications. The simplest Random Closed is the Boolean model, which considers that P is a Poisson process, the realisations S_i^j are independent and identically distributed, and independent of the points s_i^j. Figure 2 illustrates an example of a realisation of the Boolean model, where P is an homogeneous Poisson process, and the shapes are circles of random radius. It should now be obvious what information is contained in the vector of parameters χ_r. For example, if we use a Boolean model to describe region $\mathcal{A}_r$, χ_r contains the information required to specify the intensity function of the Poisson process, λ_r, as well as parameters α_r describing the *distribution of the random shapes* S_i^j, such that

$$\chi_r = [\{\lambda_r(s)\}_{s\in\mathcal{S}} \quad \alpha_r] \, .$$

The next section presents the formal definition of "shape" that we use in this work. As we will see, the set of all possible shapes is not an Euclidean (linear) space, and thus the usual parametric distributions are not possible candidates as statistical shape models.

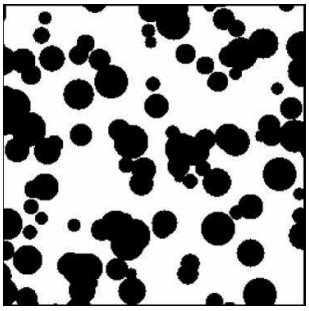

Fig. 2: Example of a RCS model with circle as shape model.

2.2 Shape spaces

Kendall theory of shape [4] defines a shape by **what is left after removing size translation and scale from an object**. Kendall has shown that this space can be conveniently associated as the set of equivalence classes defined by a rotation operator in multi-variate complex projective space. Kendall's definition was targeted to situations were the labels associated to the contour points have a semantic meaning (they are sets of remarkable points, or landmarks). This is not the case in our mapping application, where the points are obtained by scanning the contour of objects detected in the side-scan images, and whose origin and orientation are arbitrary.
Let X be one "contour", i.e. an ordered set of n points in 2 dimensions, and $\tilde{X}$ its complex representation under the natural isomorphism from $\mathbb{R}^2$ to $\mathbb{C}$. According to our definition (see [5]) the shape of X, $[Z(X)]$, is the set of all sets of n-dimensional complex vectors that can be generated from X by scaling, translation, rotation (complex multiplication) and circular permutations:

$$[Z(X)] = \left\{ W = \rho e^{j\phi} P\tilde{X} + t\mathbf{1}, \ \ \rho \in \mathbb{R}^+, \phi \in [0, 2\pi], P \in \mathcal{P}, t \in \mathbb{C} \right\}, \tag{3}$$

where $\mathcal{P}$ is the set of circular permutations and mirroring operators, and $\mathbf{1}$ is a column vector of ones.
We have seen that the Procrustes, full Procrustes and gap measures proposed in the literature for Kendall's definition of shape (which does not include the action of the permutation group $\mathcal{P}$), can be extended trivially to our definition. Our shape space can thus be made a metric space.

2.3 Distributions in shape space

The space of all shapes composed of k landmarks in m-dimensions is noted Σ_m^k. In our case, $m = 2$. Σ_m^k is not euclidean but we can define a distance in this space, the mean of a set of shapes and the tangent space to a given shape. Let the distance between the shapes of 2 contours D_P, be chosen as the minimum over the set of permutations $\mathcal{P}$ and over the set of 2D rotations, $\mathbf{SO}(2)$, of the angular distance between the centered and normalized versions of the objects (the preshapes) noted Z as summarized in Definition 1. Definition 2 identifies the mean shape $[\mu]$ as the *Karcher mean* of the set of considered shapes. Let $T_{[\gamma]}$, with $[\gamma]$ a particular shape, be the tangent space to Σ_m^k at $[\gamma]$. The projection of a shape onto $T_{[\gamma]}$ is given in Definition 3.

Definition 1 *Let X_1 and X_2 be 2 objects (their preshapes Z_1 and Z_2) and $[Z_1]$, $[Z_2]$ their shapes. The distance between $[Z_1]$ and $[Z_2]$ is*

$$D_P([Z_1],[Z_2]) = \min_{P\in\mathcal{P}} \inf_{\Phi\in[0\,2\pi]} \arg(Z_2, e^{j\Phi}PZ_1). \tag{4}$$

Definition 2 *Given a set of objects $\{X_i\}_{i=1}^M$, the mean shape of this set is*

$$[\mu] = \arg \inf_{[m]\in\Sigma_m^k} \sum_{i=1}^{M} D_P^2\left([Z_i],[m]\right). \tag{5}$$

Definition 3 *Let X be an object, Z its preshape and $[Z]$ its shape. Its projection in the tangent space $T_{[\gamma]}$ noted v is:*

$$v([Z]) = \left(\mathbf{I}_{k-1} - \gamma\gamma^T\right)\widehat{P}e^{j\widehat{\Phi}}Z, \tag{6}$$

with $[\gamma] \in \Sigma_m^k$ (γ its preshape), $\mathbf{I}_{k-1}$, the identity matrix and $\widehat{P}$ and $\widehat{\Phi}$ the minimizers of the distance between $[Z]$ and $[\gamma]$.

$T_{[\gamma]}$ is a Euclidean space and thus we can use probability distributions in it. For each region $\mathcal{A}_r$ of the seabed, we describe the geometrical properties of the objects (the shape process) inside it by multi-variate normal distributions in the tangent space to a mean shape $[\mu_r]$ characteristic of the region, so that the parameter vector α_r of the shape model in $\mathcal{A}_r$ is:

$$\alpha_r = [\,[\mu_r] \quad \Sigma_r\,] \tag{7}$$

where Σ_r is the covariance matrix of the normal distribution in $T_{[\mu_r]}$.

2.4 B-splines representation of closed contours

As Kendall's definition of shape, we consider the description of objects with a *fixed* number of points. However, in our mapping framework, the number of points that represents the contour of each detected object is dependent on the observation condition, which determine the resolution of the digital images of the observed terrain. When identifying the statistical models that describe each region, we must be able to adjust the complexity of the statistical model that describes the shapes of the objects found inside it, i.e., the dimension of the vectors and matrices in α_r. To relate representations of the same contour with distinct number of points, we use *B-splines*. B-splines curves (of order $k+1$) $c(t)$ are piece-wise polynomial functions:

$$c(t) = Q_i(t), t \in [t_i, t_{i+1}[,$$

that have k continuity at the joining points t_i and where $(k+1)$ is the maximum degree of the polynomials $Q_i(t)$: $\quad Q_i(t) = \sum_{n=1}^{k} a_n^i t^n$. The limits t_i of the intervals in the above expression are called the knots of the spline representation. Equivalently, the spline curve can be expressed as a linear combination of a set of basic spline functions (the B-splines) $\beta_i(t)$ whose support is limited to the interval $[t_i, t_{i+k+1}]$:

$$c(t) = \sum_{i=1} P_i\beta_i(t), \tag{8}$$

where summation extends over the entire set of knot points. Let $\theta_k = \{P_0, P_1, \ldots, P_{k-1}\} \in \mathbb{R}^2$ be the set of control points of the spline representation in Equation (8).
Fitting a B-spline onto a contour $v_i(t)$ consists in finding the control points $\theta^i_{(k)}$ that best describe the contour. We do not consider here the *free knots* problem - that is more interesting but much harder to solve - and restrict ourselves to the special case of uniform splines. In this case the knots are taken such that $\{t_j = j,\ j = 0, \ldots, k-1\}$.
Thus we represent the observed contours $\{v_i(t)\}_{i=1}^M$ with:

$$v_i(t) = \mathbf{B}^i_{(k)} \theta^i_{(k)}, \tag{9}$$

where the elements of $\mathbf{B}^i_{(k)}$ are:

$$\left[\mathbf{B}^i_{(k)}\right]_{nj} = \beta_j \left(t^i_0 + \frac{(t_{k-1} - t_0)n}{N} \right). \tag{10}$$

Then the parameter vector is estimated using:

$$\widehat{\theta^i_{(k)}} = (\mathbf{B}^i_{(k)})^\dagger v_i, \tag{11}$$

where $(\mathbf{B}^i_{(k)})^\dagger$ is the pseudo-inverse matrix of $\mathbf{B}_{(k)^i}$.

2.5 Spline Shape model

For each contour $v_i(t)$, we observe a discretized version noted v_i that depends on the parameter t^i_0, the first observed point on the continuous curve. As seen in Equation (10), $\mathbf{B}^i_{(k)}$ depends on this first observed point. We assume that $t^i_o = 0 \forall i$ so that $\mathbf{B}^i_{(k)} = \mathbf{B}_{(k)}$ for all $i = 1 \ldots M$. We model the observed contours $v_i(t)^M_{i=1}$ as:

$$v_i = \mathbf{B}_{(k)} \left([\mu_r] + \xi_i \right) + \eta_i, \tag{12}$$

where

- $[\mu_r] \in \mathbb{R}^k$ is the mean shape of the control points of the model.
- η_i is a sample of a white Gaussian noise: $\eta_i \sim \mathcal{N}(0, \Sigma_N)$, $\Sigma_N = \sigma^2 \mathbf{I}_N$
- ξ_i is a sample of a Gaussian distribution in the tangent space to Σ^k_2 at $[\mu_r]$: $\xi_i \sim \mathcal{N}(0, \sigma^2_r \mathbf{I}_k)$

We define $r = \sigma^2_r / \sigma^2$. $r \in \mathbb{R}^+$ which can be interpreted as a signal to noise ratio.

3 ESTIMATION OF THE NUMBER OF KNOTS

We present in this section the derivation of the criterion used to determine the dimensionality k of the shape model that is fitted to a set of contours representative of a given sea-bes region.
The MDL principle is a generic criterion for choosing amongst a set of competing models $\{H_k\}_{k=1}^K$ as descriptors of a given data set, where each model corresponds to a family of probability distributions, most typically a parametric distribution families, where the dimensionality of the parameter indexing the individual distributions in each H_i can depend on i. It exploits the relation between probability distributions and optimal codes, and can, in very intuitive terms, be expressed as saying that one should always choose the model k that leads to the shortest data description DL_k:

$$\widehat{k} = \arg\min_k DL_k \tag{13}$$

While many authors have presented generic and unified expressions for DL_k, a simple and commonly used version consists in considering "two-part" codes, such that the overall description length for data using model H_i, DL_i is the sum of two terms:

$$DL_k = \mathcal{L}(data|\hat{\theta}_k) + \mathcal{L}(\hat{\theta}_k) \qquad k = 1, \ldots, K\ , \tag{14}$$

where we considered that $H_k = \{p(\cdot|\hat{\theta}_k), \rho_k \Theta\}$ is a parametric model, and θ_k is the vector that indexes the elements of H_k. This description length is the sum of two terms, one that codes the data

using the best model in H_k (the one indexed by $\hat{\theta}_k$) and the second term is the length of the code that specifies $\hat{\theta}_i$. In our case, the different models correspond to distinct dimensions of the spline representation, i.e., k is the number of knots or control points.

The best code $p(\cdot|\hat{\theta}_i)$ is the one that minimizes the description length of the data and can thus be trivially identified with the Maximum Likelihood (ML) estimator of the model parameters: $\hat{\theta}_k = \arg\max_{\theta_k} p(data|\theta_k)$, and, as we presented in Equation (12), the parameters of our statistical models are the mean $[\mu_r]$, the variance σ^2 and the ratio r, such that

$$\mathcal{L}(data|\hat{\theta}_k) = -\log p\left(data|[\hat{\mu_r}]^k, \hat{\sigma}^{2k}, \hat{r}^k\right) , \tag{15}$$

where the super index k stresses the fact that the estimates depend on the dimension of the spline model that is fitted to the contours.
The second term of the description length is the number of bits required to code the optimal parameter. We use the simple asymptotic expression that depends only on the dimension of the parameter vector and the length of the observed sequence:

$$\mathcal{L}(\hat{\theta}_k) = \frac{K_T^k}{2} \log N , \tag{16}$$

with K_T^k the total number of model parameters (the dimension of H_k as a probabilistic manifold) and N the number of observations:

$$K_T^k = 2 \times (k+1), \qquad N = 2M \times n. \tag{17}$$

3.1 ML estimates

We assess now the determination of the estimates $[\hat{\mu_r}]^k$, $\hat{\sigma}^{2k}$ and $\hat{r}^k$ required to compute the description length of our data, i.e., of the M contours v_i represented each by n points in $\mathbb{R}^2$ or $\mathbb{C}$. To simplify the presentation, we omit the dependency on k in the presentation below. Our statistical models assume normal distributions, such that

$$p\left(\{v_i\}_{i=1}^M|[\mu_r], r, \sigma^2\right) = \prod_{i=1}^{M} \frac{1}{(2\pi)^{n/2}\,|\Sigma_{model}|^{1/2}} \exp\left(-\frac{1}{2}\left\{(v_i - \mathbf{B}[\mu_r])^T\,\Sigma_{model}^{-1}\,(v_i - \mathbf{B}[\mu_r])\right\}\right) , \tag{18}$$

where, as we discussed before

$$\Sigma_{model} = \sigma^2 \mathbf{I}_n + \sigma_r^2 \mathbf{B}\mathbf{B}^{\mathbf{T}} . \tag{19}$$

Our goal is to find the values of $[\mu_r]$, σ^2 and r that maximize Equation (18). Computing the partial derivatives and equating to zero, we obtain the following equations

$$[\hat{\mu_r}] = RD^{-1}Q^T\bar{v} \tag{20}$$

$$0 = \sum_{j=1}^{k}\left\{\frac{d_j^2}{(1+\hat{r}d_j^2)}\left(M - \frac{1}{\hat{\sigma^2}}\frac{\alpha_j}{1+\hat{r}d_j^2}\right)\right\} \tag{21}$$

$$\hat{\sigma}^2 = \frac{1}{Mn}\left[\sum_{j=1}^{k}\frac{\alpha_j}{1+\hat{r}d_j^2} + Y\right] \tag{22}$$

where

- $\bar{v} = \frac{1}{M}\sum_{i=1}^{M} v_i$ is the average of the observed contours
- Q and D are defined by the SVD decomposition of $\mathbf{B}$: $\mathbf{B} = QDR^T$

- Q_c is a unitary $n \times (n-k)$ matrix such that $Q_c^T Q = \mathbf{0}$ and $Q_c^T Q_c = I_{n-k}$
- $x_i \triangleq Q^T(v_i - \bar{v}) \in \mathbb{C}^n$
- $\alpha_j \triangleq \sum_{i=1}^{M} x_{ji}^2 \in \mathbb{C}$
- $Y \triangleq \sum_{i=1}^{M} v_i^T Q_c Q_c^T v_i \in \mathbb{R}$.

Equation (20) shows that $[\hat{\mu_r}]$ is independent of both r and σ^2, requiring only with knowledge of $\mathbf{B}$. On the contrary, the estimates of r and σ^2 are coupled by equations (21) and (22).

If we consider that $r \gg 1$, i.e., that the variability of the observed shapes should not be associated to observation noise but instead captured by the statistical shape model, we can obtain approximate analytical expressions for $\hat{\sigma}_\theta^2$ and $\hat{\sigma}^2$:

$$(\hat{\sigma}^2)^k = \frac{Y}{M(n-k)}, \qquad \hat{r}^k = \frac{n-k}{k}\frac{1}{Y}\sum_{j=1}^{k}\frac{\alpha_j}{d_j^2}\ , \tag{23}$$

where we explicitly indicate the dependency of these estimates on the dimension k of the spline model. Use of Equations (23), (18) and (17) in the general description length in Equation (14) finally yields the criterion that must be minimized over k.

3.2 Simulations

We applied the results above to sets of contours extracted from real side-scan images[1]. The contours have been partitioned into two classes C_1 and C_2 using the algorithm presented in [5]. Figure 3 illustrates the identified classes in one test image, using a different color for each class.

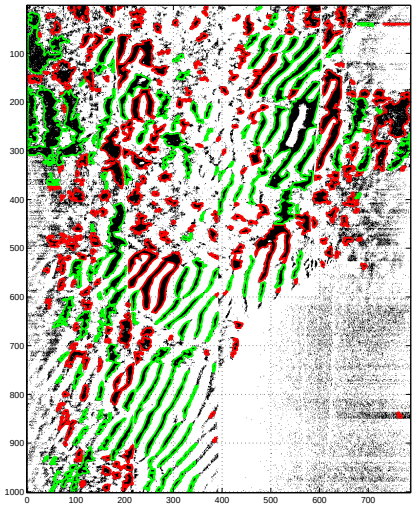

Fig. 3: Binary partition of the shadows detetected in a side-scan image).

We used the MDL criterion to find the dimension of the statistical spline/shape model that represents each of these classes. Figure 4 plots the evolution of DL_k for both classes (C_1 on the left and C_2 on the right). The optimal model dimensions indicated by our criterion are $\hat{k_1} = 36$ and $\hat{k_2} = 25$. As we should expect, the dimension of the model fitted to the less regular shapes in C_1 is larger than the one fitted to C_2.
Our experiments show that the description length is dominated by the likelihood term, the standard (BIC) penalty term inducing very small corrections. We conclude that a better approximation of the

[1]Images have been provided to us by the NATO Undersea Research Center (La Spezia, Italy) and were pre-processed with an algorithm based on cooccurence matrices developed by Y. Petillot, Heriot-Watt Univ. (Edinburgh, UK).

model complexity should be used here, for instance the first order term that depends on the Fisher information matrix. We will assess derivation of this term for the assumed models in the future.

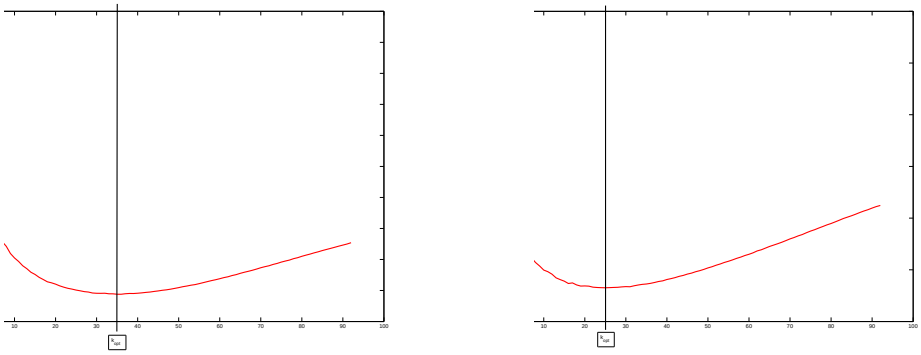

Fig. 4: DL_k (left: C_1, right: C_2).

Using the optimal model dimensions above, obtained using the MDL criterion, we computed the mean and variances of the normal distributions that approximate the true shape distribution of the control points in the tangent space to the shape space at each class mean as summarized in Figure 5.

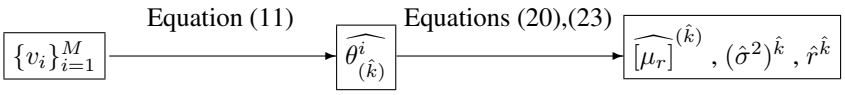

Fig. 5: Block diagram of the model estimation process.

Figure 6 displays in the left the two mean shapes, and on the right a color representation of the corresponding covariance matrices.

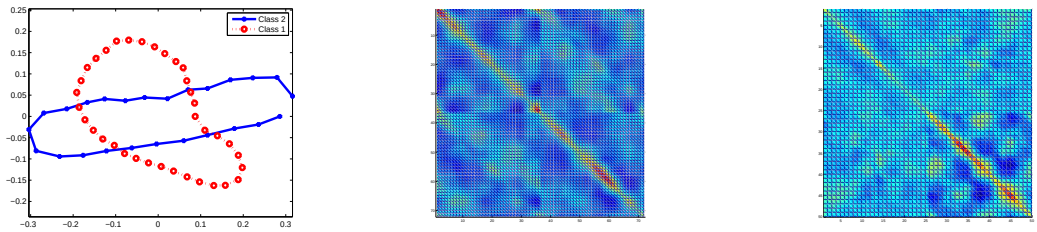

Fig. 6: Mean shape of the control points θ_i of each class (red: C_1, blue: C_2) and covariance matrices (center: C_1, right: C_2).

References

[1] S. Rolfes, "Stochastic geometry : an approach to featureless perception-based robot navigation," *PhD thesis*, December 2002.

[2] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, 1978.

[3] J. M. D. Stoyan, W.S. Kendall, *Stochastic geometry and its applications*. John Wiley and sons, 1985.

[4] D. G. Kendall, "Shape manifolds, procrustean metrics and complex projective space," *Bull. London Math. Soc.*, pp. 81–121, 1984.

[5] L. Amate and M. Rendas, "A contribution to the problem of mapping seabed transition zones," *Proc. of Oceans'06, Boston, USA*, September 2006.