Unsupervised marked point process model for boat extraction in harbors from high resolution optical remotely sensed images

Paula Craciun, Josiane Zerubia

UNSUPERVISED MARKED POINT PROCESS MODEL FOR BOAT EXTRACTION IN HARBORS FROM HIGH RESOLUTION OPTICAL REMOTELY SENSED IMAGES

Paula Craciun, Josiane Zerubia *

AYIN research group, INRIA Sophia-Antipolis Méditerranée
2004 Route des Lucioles, BP 93, 06902 Sophia-Antipolis Cedex (France)
e-mail: Firstname.Lastname@inria.fr

ABSTRACT

Marked point process models have been successfully used in image analysis for feature extraction purposes in high resolution remotely sensed images. The model is usually based on two types of energy terms: a data term, which reflects the configuration's fidelity with respect to the input image, and a prior term, which reflects some knowledge about the objects to be extracted. In this paper, we deal with the problem of elliptical shape extraction. We propose new energy terms for the extraction of boats in harbors, which is a particularly difficult problem and we show results on high resolution optical images.

Index Terms— marked point processes, object extraction, remote sensing, high resolution, optical sensor.

1. INTRODUCTION

The general problem of object extraction in static images is hard to solve, since the system has to distinguish a particular object class from all the others. Thus, the system must process a model which has high inter-class and low intra-class variability. Furthermore, it must deal with situations where the object is unevenly illuminated, rotated in the plane, partially occluded or blended with the background.

In the case of high-resolution images, object-based detection is preferred. In this sense, we devise the use of a Marked Point Process (MPP) model [1, 2] for object extraction. Basically, it is a stochastic process which allows for modeling the scene as a set of unknown number of geometrical shapes. The actual object set can be viewed as a realization of the MPP. The consistency of an object set is given by two types of energy terms. The first term is called the data term, which controls how well the proposed objects fit to the actual image. The second term is called a prior term and reflects prior knowledge about the objects, thus allowing the incorporation of geometrical and spatial constraints into the model. First attempts to object extraction using MPPs of circles, proved to be very encouraging [3, 4]. The model was then extended to fit other shapes like ellipses and rectangles [5]. Herein, we propose a modification of the model, in the particular case of boat detection in harbors, which is a difficult task since the objects have a particular distribution. The paper is divided into two main parts. The first one presents a theoretical study of the proposed unsupervised method by describing the general model for object extraction using MPPs. The second part details the new model used for boat extraction in harbors. We depict and discuss some detection results and propose some perspectives.

2. PROPOSED MODEL FOR OBJECT EXTRACTION

We consider a MPP of ellipses. When seen from above, i.e. satellite images at nadir, the boats tend to have an elliptical shape. Therefore, ellipses best suit the shape of boats as opposed to circles or other geometrical shapes. The object space $\mathcal{W}$, is a bounded set in $\mathbb{R}^5$:

$$\mathcal{W} = [0, X_M] \times [0, Y_M] \times [a_m, a_M] \times [b_m, b_M] \times [0, \pi]$$

where X_M and Y_M are the width and height of the input image, respectively, a_m and a_M are the minimum and the maximum of the semi-major axis of the ellipse, b_m and b_M are the minimum and the maximum of the semi-minor axis of the ellipse and $\omega \in [0, \pi]$ is the orientation of the ellipse. The parameterization of the ellipse is shown in Figure 1 .

We are interested in a particular family of MPPs, namely the Gibbs processes [6]. Denoting the observed image with $\mathbf{y}$, the density of the considered MPP is given by:

$$f_\theta(X = \mathbf{x}|\mathbf{y}) = \frac{1}{c(\theta|\mathbf{y})} \exp^{-U_\theta(\mathbf{x},\mathbf{y})}$$

where $c(\theta|\mathbf{y})$ is a normalizing function given by:

$$c(\theta|\mathbf{y}) = \int_\Omega \exp^{-U_\theta(\mathbf{x},\mathbf{y})\mu(d\mathbf{x})}.$$

Here, θ is a parameter vector which allows the model to be flexible and fit different images. It has to be adjusted according to the given image. $\mu(\cdot)$ is the intensity measure of the

*We thank the French Space Agency (CNES) for providing the data, Saima Ben Hadj from INRIA/I3S for fruitful discussions and Antoine Mangin from ACRI-ST for useful feedback regarding the model.

reference Poisson process, $U_\theta(\mathbf{x}, \mathbf{y})$ is called the energy, Ω is the configuration space and $\mathbf{x}$ denotes the object configuration.

The energy function is divided into two parts: the data term $U_\theta^d(\mathbf{x}, \mathbf{y})$ which represents the configuration's fidelity with respect to the input image and the prior term $U_\theta^p(\mathbf{x})$ that reflects some knowledge about the objects to be extracted. Using the MAP criterion, the most likely configuration which allows the extraction of objects corresponds to the global minimum of the total energy $U_\theta(\mathbf{x}, \mathbf{y})$:

$$x \in \underset{\mathbf{x}\in\Omega}{Argmax}\ f_\theta(X = \mathbf{x}/\mathbf{y}) = \underset{\mathbf{x}\in\Omega}{Argmin}[U_\theta(\mathbf{x}/\mathbf{y})].$$

Parameter estimation techniques are not presented here, for more details, see [3, 5]. Once the parameter vector θ is determined, a solution to this optimization problem can be found by means of simulated annealing combined with a sampling algorithm, e.g. Reversible Jump MCMC [7].

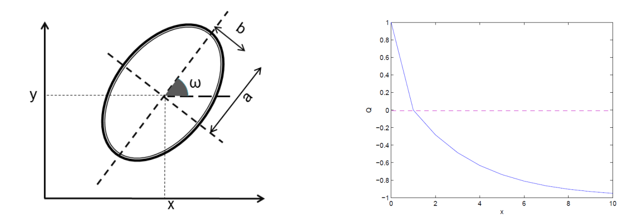


Fig. 1. left: Parameterization of an ellipse; right: Representation of the quality function $\mathcal{Q}(x)$

2.1. Data energy term

Two main approaches have been previously devised for computing the data energy term [3]. The first one, carried out in a Bayesian framework, does not take sufficiently the spatial features of the objects into account. Therefore, a second approach was proposed, the idea behind it being to model the data energy term locally, for each object in the configuration. Thus, the total data energy term of the configuration $\mathbf{x}$ is:

$$U_{\theta_d}^d(\mathbf{x}, \mathbf{y}) = \gamma_d \sum_{u\in\mathbf{x}} U^d(u),$$

where γ_d represents the weight which has to be estimated and θ_d is the parameter vector.

The definition of the local data energy term relies on the computation of a distance measure between the distribution of the set of pixels belonging to the object u and the set of pixels belonging to its border $\mathcal{F}^\rho(u)$. The local data energy is built as a qualification of this contrast measure, thus favoring well positioned objects (i.e. high contrast between the interior and the border of the object) while penalizing misplaced ones:

$$U^d(u) = \mathcal{Q}(\frac{d(u, \mathcal{F}^\rho(u))}{d_0})$$

where $d(u, \mathcal{F}^\rho(u))$ is the Bhattacharya distance [8] between the object u and its boundary $\mathcal{F}^\rho(u)$ of width ρ and d_0 is a threshold for that distance. $\mathcal{Q} : \mathbb{R}^+ \rightarrowtail [-1, 1]$ is a quality function and attributes a negative value to well placed objects (e.g. those objects u for which $d(u, \mathcal{F}^\rho(u))$ is higher than the threshold d_0) and a positive value to misplaced objects. $\mathcal{Q}(x)$ is defined using a cubic root which allows a moderate penalization when the distance is close to the threshold, as shown in Figure 1:

$$\mathcal{Q}(x) = \begin{cases} 1 - x^{1/3} & \text{if } x < 1 \\ exp(-\frac{x-1}{3}) - 1 & \text{if } x \geq 1 \end{cases}$$

2.2. Prior energy term

This energy term corresponds to a penalization of overlapping objects, avoiding the detection of the same object several times. The proposed model uses a *hard core* process to handle object overlapping. Thus, denoting by $A(x_i, x_j) = \frac{Area(x_i \cap x_j)}{\min(Area(x_i), Area(x_j))}$ the area of intersection between the objects x_i and x_j, the prior energy can be defined as: $U^p(\mathbf{x}) = \sum_{1<i\neq j<n(\mathbf{x})} t_s(x_i, x_j)$, where:

$$t_s = \begin{cases} 0 & \text{if } A(x_i, x_j) < s \\ +\infty & \text{otherwise} \end{cases}$$

where $s \in [0, 1]$ corresponds to the amount of overlapping allowed by the model and $n(\mathbf{x})$ is the number of objects in the configuration $\mathbf{x}$. Accordingly, all configurations containing at least two objects that overlap to a higher ratio than specified by s are prohibited.

3. EXTRACTION MODEL FOR BOATS IN HARBORS

As stated before, boats in harbors tend to have a particular distribution. Previous work has been done in devising a model for this difficult case [5]. The energy term has been modified, by considering only a section of the exterior border for computing the distance measure as shown in Figure 2, since boats in harbors tend to be close to each other.

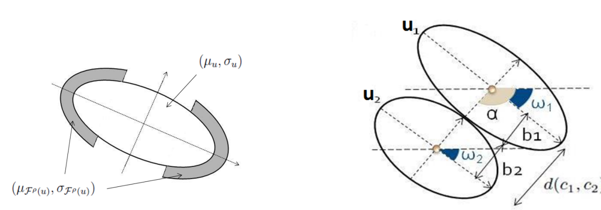


Fig. 2. left: Exterior border considered for the computation of the data energy term; right: Alignment constraints on the prior energy term

Furthermore, the prior energy term has been modified to incorporate additional constraints. This model implied that the objects are very close to each other, they are not shifted wrt each other and have the same, global orientation. These constraints are graphically represented in Figure 2 (right). One can observe that this model is far too restrictive and is not applicable for the general case. Therefore, we have tried to relax these constraints in order to better fit the general case.

3.1. Modification of the data energy term

We will keep the modification of the data energy term proposed in [5] and introduce an additional key element. Considering the entire interior of the objects when computing the contrast measure might not yield the best results, since this area is not always homogeneous. Therefore, we propose to introduce an additional term, by taking into account the contrast measure between the interior border of the objects and the exterior one, as shown in Figure 3 .

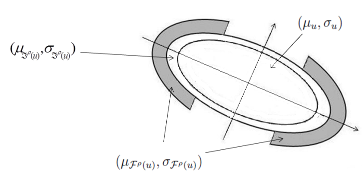


Fig. 3. Exterior and interior border considered for the computation of the new data energy term.

The data term, thus becomes:

$$U^d_{\theta_d}(u) = \mathcal{Q}(\frac{d(u, \mathcal{F}^\rho(u))}{d_0}) + \gamma_c \mathcal{Q}(\frac{d(\mathcal{I}^\rho(u), \mathcal{F}^\rho(u))}{d_0})$$

where γ_c is the weight of the contrast measure between the interior border, $\mathcal{I}^\rho(u)$, and the exterior border, $\mathcal{F}^\rho(u)$, of the object u. Border width, ρ, and contrast threshold, d_0, are kept the same for both parts of $U^d_{\theta_d}(u)$ in order to keep the number of parameters as low as possible.

3.2. Modification of the prior energy term

Considering only the non-overlapping constraint does not yield satisfactory results. One can see that indeed objects tend to be locally aligned, where alignment should be interpreted as neighboring objects sharing the same, or similar, orientation. For this purpose, we define an alignment interaction between two neighboring ellipses $u1$ and $u2$ in the following way:

$$u_1 \sim_{al} u_2 \Leftrightarrow d_\omega(u_1, u_2) \leq d_{\omega_{max}}$$

where, $d_\omega = |\omega_1 - \omega_2|$ is the difference of the orientations of the two ellipses u_1 and u_2 and $d_{\omega_{max}}$ is the maximum angle allowed between two neighboring objects. Then, we append a prior energy that promotes this alignment in the following way:

$$U_{al}(u_1, u_2) = \begin{cases} \delta\varpi(d_\omega(u_1, u_2), d_{\omega_{max}}) \\ \qquad \text{if } u_1 \sim_{al} u_2 \\ 0 \qquad \text{otherwise} \end{cases}$$

where $\varpi(x, x_{max}) = -\frac{1}{x^2_{max}}[\frac{1+x^2_{max}}{1+x^2} - 1],\ for\ x \leq x_{max}$ is a reward function that favors aligned frames. The prior energy of a configuration $\mathbf{x}$ corresponding to the alignment constraint is the sum over all the object pairs of the configuration $\mathbf{x}$:

$$U^p_{al}(\mathbf{x}) = \gamma_{al} \sum_{1<i<j<n(\mathbf{x})} U_{al}(u_i, u_j).$$

The weight γ_{al} is also a parameter that has to be estimated.

A method proposed in [9] for road extraction, is used in order to locally determine the orientation of the keys using the water, in a preprocessing step, thus obtaining a set of relevant pixels, i.e. pixels that incorporate data about the orientation.

Finally, we search for the local orientation of the objects, i.e. at each creation of an object, we search for a relevant pixel in its neighborhood. If no relevant pixel is found, no assumptions are made on the orientation of the object. If a relevant pixel is found, its position wrt the semi-axes of the ellipse is determined. We will favor ellipses perpendicular to the orientation indicated by the relevant pixel. Thus, the final prior energy term that refers to alignment becomes:

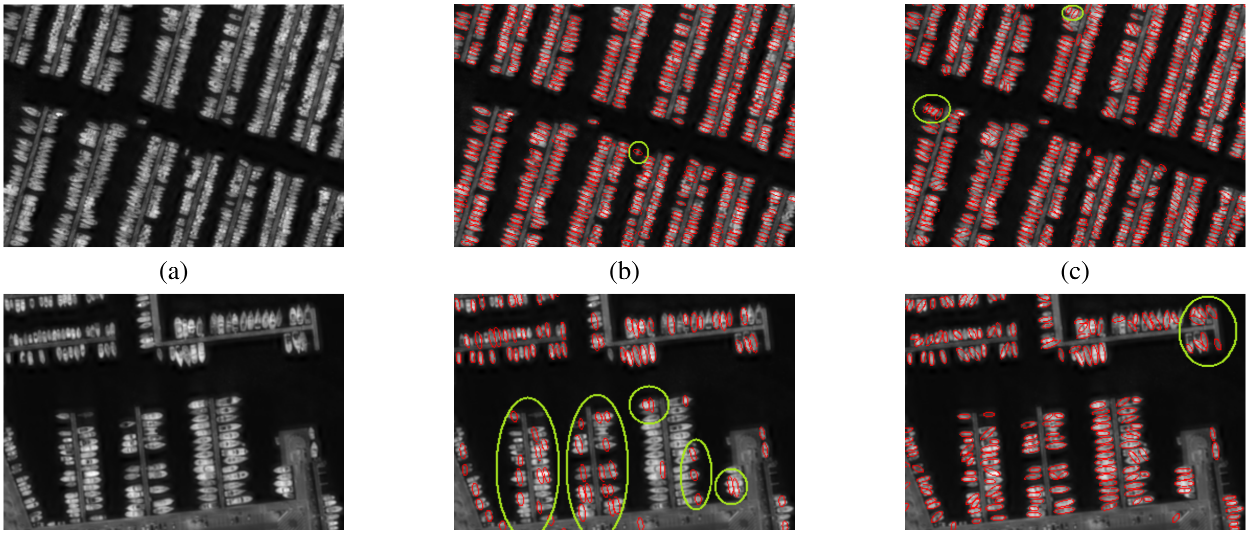
$$U^p_{al_{\omega_l}}(u) = \begin{cases} \gamma_{al} \sum_{v \in \mathbf{x}} U_{al}(u, v) \\ \qquad \text{if relevant pixel is found and} \\ \qquad ||\omega_u - \omega_l|| \leq d_{\omega_{max}}\text{, or} \\ \qquad \text{if no relevant pixel is found} \\ 0 \qquad \text{if a relevant pixel is found and} \\ \qquad ||\omega_u - \omega_l|| > d_{\omega_{max}} \end{cases}$$

where ω_u is the orientation of object u and ω_l is the orientation perpendicular to the one retained in the relevant pixel.

4. RESULTS AND DISCUSSIONS

Figure 4 shows the final results. Figure 4 (a) and (d) represent two test images of boats in harbors. Figure 4 (b) shows the result by applying the model presented in [5]. The detection results are very good, counting a total number of 523 ellipses, compared to a number of 518 boats counted by an expert. The computations lasted $1h$ and $38min$ on a $1.86GHz$ processor only for parameter estimation. Figure 4 (c) shows the result by applying the model we propose in this paper. The detection results are a little bit less accurate than in [5] resulting in a total number of 501 ellipses, but the computation time decreased to $45min$ and $3sec$ for parameter estimation and $55min$ and $21sec$ for the detection, on a $2.20GHz$ processor.

The strength of the proposed model is visible in Figure 4 (f). Where the object extraction of the model presented in [5] fails utterly due to the different orientations of the boats



(a) (b) (c)

(d) (e) (f)

Fig. 4. (a), (d) Image of boats in a harbor ©CNES; (b),(e) Extraction results using the model proposed in [5]; (c),(f) Extraction results using the proposed model.

Boat Image	Detection error for old model [5]	Detection error for new model
Figure 4 (a)	$\sim 1\%$	$\sim 10\%$
Figure 4 (d)	$> 45\%$	$< 30\%$

Table 1. Comparison between detection errors

in the image (see Figure 4 (e)), our model proves to be more efficient, see Table 1. We highlight some problematic areas in green, but the overall performance is very good.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a new method for automatic object extraction based on a marked point process. We applied this method to extracting boats in harbors. We have shown why previous attempts to solve this problem, resulted into a very restrictive model. We have, therefore, relaxed the constraints for such models. This resulted in the modification of both, the data energy term and the prior term. Future work will include an automatic detection of the keys in harbors, in order to better extract the local orientation of the boats.

6. REFERENCES

[1] C. J. Geyer and J. Møller, "Simulation procedures and likelihood inference for spatial point processes," *Scandinavian Journal of Statistics*, vol. 21, pp. 359–373, 1994.

[2] C. J. Geyer, "Likelihood inference for spatial point processes," in *Stochastic geometry, likelihood and computation*. CRC Press, 1999.

[3] F. Chatelain, X. Descombes, and J. Zerubia, "Parameter estimation for a marked point process. application to object extraction from remote sensing images," *Proc. EMMCVPR*, vol. 5681, pp. 221–234, 2009.

[4] G. Perrin, X. Descombes, and J. Zerubia, "A marked point process model for tree crown extraction in plantations," *Proc. ICIP*, vol. 1, pp. 661–664, 2005.

[5] S. Ben Hadj, F. Chatelain, X. Descombes, and J. Zerubia, "Parameter estimation for a marked point process within a framework of multidimensional shape extraction from remote sensing images," *Proc. of ISPRS Conference on Photogrammetry Computer Vision and Image Analysis*, vol. XXXVIII, 2010.

[6] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, John Wiley and Sons, 1987.

[7] P. Green, "Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.

[8] F. Goudail, P. Réfrégier, and G. Delyon, "Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images," *Journal of Optical Science of America A*, vol. 21, no. 7, pp. 1231–1240, 2004.

[9] Y. Li and R. Briggs, "Automatic extraction of roads from high resolution aerial and satellite images with heavy noise," *Proceedings of the 6th International Conference on Geographic Information Systems*, 2009.