STATISTICAL REGION-BASED ACTIVE CONTOURS WITH EXPONENTIAL FAMILY OBSERVATIONS

François Lecellier[a], Stéphanie Jehan-Besson[a], Jalal Fadili[a], Gilles Aubert[b], Marinette Revenu[a]

[a] GREYC UMR 6072 CNRS
ENSICAEN-Université de Caen
14050 Caen, France

[b] Laboratoire J.A. Dieudonné UMR 6621 CNRS
Université de Nice Sophia-Antipolis
06108 Nice, France

ABSTRACT

In this paper, we focus on statistical region-based active contour models where image features (e.g. intensity) are random variables whose distribution belongs to some parametric family (e.g. exponential) rather than confining ourselves to the special Gaussian case. Using shape derivation tools, our effort focuses on constructing a general expression for the derivative of the energy (with respect to a domain) and derive the corresponding evolution speed. A general result is stated within the framework of multi-parameter exponential family. More particularly, when using Maximum Likelihood estimators, the evolution speed has a closed-form expression that depends simply on the probability density function, while complicating additive terms appear when using other estimators, e.g. moments method. Experimental results on both synthesized and real images demonstrate the applicability of our approach.

1. INTRODUCTION

In image segmentation, the main issue is to extract one or several regions according to a given criterion. Region based active contours have proven their efficiency for such a task. The evolution equation of the deformable curve is deduced from a functional to minimize that may take benefit of statistical properties of the image features. In this case, the functional depends on the probability density function (pdf) of the image feature within the region. For instance, the minimization of the - log-likelihood of the pdf has been widely used for the segmentation of homogeneous regions in noisy images [1, 2, 3]. In these works, the pdf is parametric, i.e. it follows a prespecified law (Gaussian, Rayleigh). The pdf is then indexed by one or more parameters (e.g. mean and variance for Gaussian laws) describing the distribution model. These parameters depend on the region and must be estimated at each evolution step. The estimation may be performed using different methods (moments method, maximum likelihood (ML)). However, to the best of our knowledge, the influence of the estimation method on the computation of the evolution equation has never been investigated.

In this paper, using shape derivative tools [4, 5, 6], our effort focuses on constructing a general expression for the derivative of a functional depending on parametric pdfs, and hence on the corresponding evolution equation. Our contribution is twofold. Firstly, we show that the estimator of the distribution parameters is crucial for the derived speed expression. Secondly, we propose to give general results for the evolution equation within the framework of multi-parameter exponential family and ML estimation theory. Therefore, in many image processing problems, one can consider a minimization problem with any distribution if the statistical features of the underlying process are known, e.g. Poisson for a counting process, Rayleigh for ultrasound images, and so on.

This paper is organized as follows: the region-based variational framework for segmentation is introduced in Section 2. Influence of the parameters estimator is investigated in Section 3. Our new theoretical results are detailed in Section 4. Experimental results are discussed in Section 5. We finally conclude and give some perspectives.

2. PROBLEM STATEMENT

Let $\mathcal{U}$ be a class of domains (open, regular bounded sets, i.e. C^2) of $\mathbb{R}^k$, and Ω_i an element of $\mathcal{U}$ of boundary $\partial\Omega_i$. The image domain is noted Ω_I. A region-based segmentation problem aims at finding a partition of Ω_I in n regions $\{\Omega_1, .., \Omega_n\}$ of respective boundaries $\{\partial\Omega_1, .., \partial\Omega_n\}$ that minimizes the following criterion:

$$E(\Omega_1, .., \Omega_n, \Gamma) = \sum_{i=1}^{n} E_{ri}(\Omega_i) + \lambda E_b(\Gamma) \qquad (1)$$

where E_{ri} describes the homogeneity of the region Ω_i, and $\Gamma = \bigcup_{i=1}^{n} \partial\Omega_i$. The energy term E_b is a regularization term balanced with a positive real parameter λ.

2.1. Statistical energy

Let us denote $\mathbf{Y}(\mathbf{x})$ a vector of random variables (VRV) at location $\mathbf{x}$ where $\mathbf{x} \in \Omega_I$. A region Ω_i is considered to be homogeneous according to the image feature $\mathbf{y}(\mathbf{x}) \in \chi \subset \mathbb{R}^d$ if

the associated VRV $\mathbf{Y}(\mathbf{x})$ follows a prespecified probability distribution with pdf $p(\mathbf{Y}(\mathbf{x}), \boldsymbol{\theta}_i)$ of parameters $\boldsymbol{\theta}_i$ in the region Ω_i. The image feature may be for example the intensity of the region (a 3-dimensional vector for color regions $d = 3$, or a scalar for grey level regions $d = 1$). We then consider functions of the form:

$$E_{ri}(\Omega_i) = \int_{\Omega_i} \Phi(p(\mathbf{y}(\mathbf{x}), \boldsymbol{\theta}_i)d\mathbf{x} \tag{2}$$

with Φ at least C^1 and Lebesgue integrable function.

Since the above integrals are with respect to domains, and also that the parameters $\boldsymbol{\theta}_i$ are region-dependent, optimization (and hence derivation) of (2) is not straightforward. The derivation of such a functional is performed using domain derivation tools as in [4]. The following theorem from [4] will be useful for region-based terms derivation:

Theorem 1 *The Gâteaux derivative of the functional $J(\Omega) = \int_{\Omega} f(\mathbf{x}, \Omega)\, d\mathbf{x}$ in the direction of $\mathbf{V}$ is the following:*

$$< J'(\Omega), \mathbf{V} > = \int_{\Omega} f_s(\mathbf{x}, \Omega, \mathbf{V})d\mathbf{x} - \int_{\partial\Omega} f(\mathbf{x}, \Omega)(\mathbf{V}\cdot\mathbf{N})d\mathbf{a}(\mathbf{x})$$

where $\mathbf{N}$ is the unit inward normal to $\partial\Omega$, $d\mathbf{a}$ its area element and f_s the shape derivative of f.

2.2. Towards a geometrical PDE

From the shape derivative, we can deduce the evolution equation that will drive the active contour towards a minimum of the criterion. In order to fix ideas, let us consider the case of a partition of an image in two regions $\{\Omega_1, \Omega_2\}$. The curve Γ stands for the interface between the two regions. The boundary energy term E_b may be chosen as the curve length and derived classically using calculus of variation [7] or shape derivation tools [4].

Let us suppose that the shape derivative of each region Ω_i may be written as follows:

$$< E'_{ri}(\Omega_i), \mathbf{V} > = - \int_{\partial\Omega_i} f_i(\mathbf{x}, \Omega_i)(\mathbf{V}(\mathbf{x}) \cdot \mathbf{N}(\mathbf{x}))d\mathbf{a}(\mathbf{x}) \tag{3}$$

In this paper, we focus our attention on the computation of the shape derivative of the criterion (2). Let us note that when $\Phi(t) = -\log(t)$ the function (2) is known as the log-likelihood score function, used to describe the homogeneity of a region according to a prescribed noise model. It has been used in [1, 2, 3]. Under some conditions, we are able to prove that minimizing this function with the ML estimator gives a simple evolution speed whose closed-form expression depends only on the pdf. However, one must be aware that using a more general function Φ or different hyperparameter estimator will generally yield complicating additive terms in the shape derivative and hence in the evolution speed expression.

3. INFLUENCE OF THE ESTIMATOR ON THE EVOLUTION EQUATION

This section is devoted to illustrate two examples that support our previous claims on the influence of the parameters estimator. More precisely, we state two derivative results in the case of the Rayleigh distribution with two different estimators (ML and moments).

3.1. Shape derivative for the Rayleigh distribution

We consider the Rayleigh distribution of scalar parameter θ_i:

$$p(\mathbf{y}(\mathbf{x}), \theta_i) = \frac{\mathbf{y}(\mathbf{x})}{\theta_i^2} \exp\left(\frac{-\mathbf{y}(\mathbf{x})^2}{2\theta_i^2}\right) \tag{4}$$

Classically, one can compute an estimate of the parameter θ_i using the moment method. In this case the estimator is given by :

$$\widehat{\theta}_{iMO} = \sqrt{\frac{2}{\pi}} \frac{1}{|\Omega_i|} \int_{\Omega_i} \mathbf{y}(\mathbf{x})d\mathbf{x} = \sqrt{\frac{2}{\pi}} \overline{\mathbf{y}(\mathbf{x})}_i \tag{5}$$

where $|\Omega_i| = \int_{\Omega_i} d\mathbf{x}$ and $\overline{\mathbf{y}(\mathbf{x})}_i$ denotes the sample mean inside the region Ω_i.

Alternatively, one can also compute an estimate using the ML estimator given by:

$$\widehat{\theta}_{iML} = \sqrt{\frac{1}{2|\Omega_i|} \int_{\Omega_i} \mathbf{y}(\mathbf{x})^2 d\mathbf{x}} \tag{6}$$

3.1.1. Shape derivative with the moment estimator

Theorem 2 *The Gâteaux derivative, in the direction of $\mathbf{V}$, of the functional $E_{ri}(\Omega_i) = -\int_{\Omega_i} \log(p(\mathbf{y}(\mathbf{x}), \widehat{\theta}_{iMO})d\mathbf{a}(\mathbf{x})$ with p a Rayleigh distribution, is the following:*

$$< E'_{ri}(\Omega_i), \mathbf{V} > = \int_{\partial\Omega_i} (\log{(p(\mathbf{y}(\mathbf{x}), \widehat{\theta}_{iMO})}$$
$$+ A(\mathbf{y}(\mathbf{x}), \Omega_i)(\mathbf{V} \cdot \mathbf{N})d\mathbf{a}(\mathbf{x})$$

$$A(\mathbf{y}(x), \Omega_i) = \left(2 - \frac{\pi}{4}\frac{\overline{\mathbf{y}^2(\mathbf{x})}_i}{\overline{\mathbf{y}(\mathbf{x})}_i^2}\right)\left(1 - \frac{\mathbf{y}(\mathbf{x})}{\overline{\mathbf{y}(\mathbf{x})}_i}\right)$$

where $\overline{\mathbf{y}^k(\mathbf{x})}_i = \frac{1}{|\Omega_i|}\int_{\Omega_i} \mathbf{y}^k(\mathbf{x})d\mathbf{x}$.

3.1.2. Shape derivative with the ML estimator

Theorem 3 *The Gâteaux derivative, in the direction of $\mathbf{V}$, of the functional $E_{ri}(\Omega_i) = -\int_{\Omega_i} \log(p(\mathbf{y}(x), \widehat{\theta}_{iML})d\mathbf{x}$ with p a Rayleigh distribution, is the following:*

$$< E'_{ri}(\Omega_i), \mathbf{V} > = \int_{\partial\Omega_i} \log\left(p(\mathbf{y}(\mathbf{x}), \widehat{\theta}_{iML})\right)(\mathbf{V} \cdot \mathbf{N})d\mathbf{a}(\mathbf{x})$$

One can then see that the parameter estimator has a clear impact on the evolution speed expression, as the additive term $A(\mathbf{y}(x), \Omega_i)$ appears in the shape derivative when the moment estimator is used.

4. GENERAL RESULTS FOR THE EXPONENTIAL FAMILY

Let us now consider the shape derivative within the framework of multi-parameter exponential family. This family includes Poisson, Rayleigh, Gaussian...

4.1. Exponential families

4.1.1. Definition

The multi-parameter exponential families are naturally indexed by a k-real parameter vector and a k-dimensional natural statistics $T(\mathbf{Y})$. A simple example is the normal family when both the location and the scale parameters are unknown ($k = 2$).

Definition 1 *The family of distributions of a VRV* $\mathbf{Y}$ $\{\mathcal{P}_{\boldsymbol{\eta}} : \boldsymbol{\eta} \in \mathcal{E} \subseteq \mathbb{R}^k\}$*, is said a* k*-parameter* **canonical** *exponential family, if there exist real-valued functions* $\eta_1, \ldots, \eta_k : \Theta \mapsto \mathbb{R}$ *and* $A(\boldsymbol{\eta})$ *on* $\mathcal{E}$*, and real-valued functions* $h, T_1, \ldots, T_k : \mathbb{R}^d \mapsto \mathbb{R}$*, such the pdf* $p(\mathbf{y}, \boldsymbol{\eta})$ *of the* $\mathcal{P}_{\boldsymbol{\eta}}$ *may be written:*

$$p(\mathbf{y}, \boldsymbol{\eta}) = h(\mathbf{y}) \exp[\langle \boldsymbol{\eta}, \mathbf{T}(\mathbf{y}) \rangle - A(\boldsymbol{\eta})], \mathbf{y} \in \chi \subset \mathbb{R}^d \quad (7)$$

where $\mathbf{T} = (T_1, \ldots, T_k)^T$ *is the* natural sufficient statistic, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)^T$ *and* $\mathcal{E}$ *are the* natural parameter vector *and space.* $\langle \boldsymbol{\eta}, \mathbf{T} \rangle$ *denotes the scalar product.*

We draw the reader's attention to the fact than $\boldsymbol{\eta}$ is a function of θ which is the parameter of interest in most applications.

For example, for the Rayleigh distribution : $\boldsymbol{\eta} = -\frac{1}{2\theta^2}$, $T(\mathbf{y}) = \mathbf{y}^2$, $h(\mathbf{y}) = \frac{\mathbf{y}}{\theta^2}$ and $A(\boldsymbol{\eta}) = -\log(-2\boldsymbol{\eta})$.

4.1.2. Properties

The following theorem establishes the conditions of strict convexity of A, and then those for ∇A to be 1-1 on $\mathcal{E}$. This is very useful result for optimization (derivation) purposes:

Theorem 4 *Let* $\mathcal{P}$ *a full rank (i.e.* $Cov[T(\mathbf{Y})]$ *is a positive-definite matrix)* k*-parameter canonical exponential family with natural sufficient statistic* $\mathbf{T}(\mathbf{Y})$ *and open natural parameter space* $\mathcal{E}$ *[8].*

(i) $\nabla A : \mathcal{E} \mapsto \mathcal{S} \subseteq \mathbb{R}$ *is 1-1. The family may be uniquely parameterized by* $\boldsymbol{\mu}(\boldsymbol{\eta}) \equiv E(\mathbf{T}(\mathbf{Y})) = \nabla A(\boldsymbol{\eta})$.

(ii) The -log-likelihood function is a strictly convex function of $\boldsymbol{\eta}$ *on* $\mathcal{E}$.

These results establish a 1-1 correspondence between $\boldsymbol{\eta}$ and $E(\mathbf{T}(\mathbf{Y}))$ such that:

$$\mathcal{S} \ni \boldsymbol{\mu} = \nabla A(\boldsymbol{\eta}) = E(\mathbf{T}(\mathbf{Y})) \Leftrightarrow \mathcal{E} \ni \boldsymbol{\eta} = \psi\left(E(\mathbf{T}(\mathbf{Y}))\right) \quad (8)$$

holds uniquely with ∇A and ψ continuous.

4.2. General results

In the sequel, for the sake of simplicity, we will invariably denote $\boldsymbol{\eta}$ for the natural parameter and its finite sample estimate over the domain (without a slight abuse of notation, this should be $\hat{\boldsymbol{\eta}}$).

Theorem 5 *The Gâteaux derivative, in the direction of* $\mathbf{V}$*, of the functional* $E_{ri}(\Omega_i) = \int_{\Omega_i} \Phi(p(\mathbf{y}(\mathbf{x}), \boldsymbol{\eta}(\Omega_i)))d\mathbf{a}(\mathbf{x})$ *where* $p(.)$ *belongs to the multi-parameter exponential family and is expressed by (7) and* $\boldsymbol{\eta}$ *the natural hyperparameter vector, is:*

$$< E'_{ri}(\Omega_i), \mathbf{V} >= -\int_{\partial\Omega} \Phi(p(\mathbf{y}))(\mathbf{V} \cdot \mathbf{N})d\mathbf{a}(\mathbf{x})$$

$$+ \int_{\Omega_i} p(\mathbf{y})\Phi'(p(\mathbf{y}))\langle \nabla_{\mathbf{V}}\boldsymbol{\eta}, \mathbf{T}(\mathbf{y}) - \nabla A(\boldsymbol{\eta})\rangle d\mathbf{x} \quad (9)$$

with $\nabla_{\mathbf{V}}\boldsymbol{\eta}$ *the Gâteaux derivative of* $\boldsymbol{\eta}$ *in the direction of* $\mathbf{V}$*, and* $\langle \mathbf{x}, \mathbf{y} \rangle$ *the scalar product of vectors* $\mathbf{x}$ *and* $\mathbf{y}$.

In a finite sample setting, when using the ML estimator, we can replace $\nabla A(\boldsymbol{\eta})$ by $\overline{\mathbf{T}(\mathbf{Y})}$ (the 1st order sample moment of $\mathbf{T}(\mathbf{Y})$). Thus, when using the -log-likelihood function, the second term becomes equal to $\int_{\Omega_i} \langle \nabla_{\mathbf{V}}\boldsymbol{\eta}, \mathbf{T}(\mathbf{y}) - \overline{\mathbf{T}(\mathbf{Y})}\rangle\rangle d\mathbf{x}$, and hence vanishes. The following corollary follows:

Corollary 1 *The Gâteaux derivative, in the direction of* $\mathbf{V}$*, of the functional* $E_{ri}(\Omega_i) = -\int_{\Omega_i} \log(p(\mathbf{y}(\mathbf{x}), \hat{\eta}_{\mathbf{ML}}(\Omega_i))d\mathbf{a}(\mathbf{x})$ *when* $\hat{\eta}_{\mathbf{ML}}$ *is the ML estimate, is the following:*

$$< E'_{ri}(\Omega_i), \mathbf{V} >= \int_{\partial\Omega_i} (\log\left(p(\mathbf{y}(\mathbf{x}), \hat{\eta}_{\mathbf{ML}}(\Omega_i))\right)(\mathbf{V}\cdot\mathbf{N})d\mathbf{a}(\mathbf{x})$$

This provides an alternative proof to the result of [1, 3]. Nonetheless, we here point out that, in the work of [1, 3], the role of the parameters estimator was not elucidated.

5. EXPERIMENTAL RESULTS

This section presents some experimental results on noisy images. The initial noise-free image is shown in Fig.1. For four different Battacharya distances (BD), we have systematically corrupted this image with two types of noise: Poisson and Rayleigh. The Battacharya distance is used as a measure of "contrast" between objects and background. It is defined as :

$$\mathcal{D}(p_f(\mathbf{y}), p_o(\mathbf{y})) = -\log \int_{\mathbb{R}^d} \sqrt{p_f(\mathbf{y})p_o(\mathbf{y})}d\mathbf{y}$$

For each combination of BD value and noise type, 50 noisy images were generated. Each noisy image was then segmented using four different energy functionals, namely Chan-Vese [9], and our method with -log-likelihood and ML estimator with three assumed noise models: Gaussian, Rayleigh and Poisson. For each segmented image with each method at each
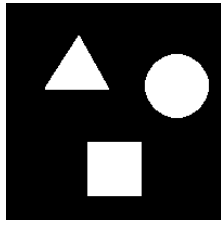
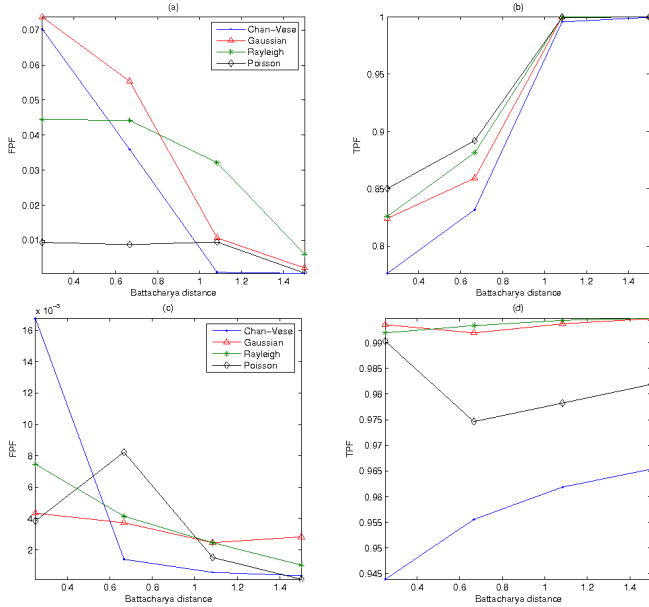**Fig. 1**. Initial noise-free image.

Fig. 2. FPF and TPF as a function of BD.

BD value, the average false positive fraction (FPF) and true positive fraction (TPF), over the 50 simulations were computed. The bottomline of these experiments is to show that using the appropriate noise model will yield the best performance in terms of compromise between specificity (over-segmentation as revealed by the FPF) and sensitivity (under-segmentation as revealed by the TPF).

Fig.2 depicts the average FPF (left) and TPF (right) as a function of the BD for Poisson ((a)-(b)) and Rayleigh ((c)-(d)) noises. As expected, the FPF exhibits a decreasing tendency as the BD increases, while the TPF increases with BD, which is intuitively acceptable. More interestingly, the best performance in terms of compromise between FPF and TPF is reached when the contaminating noise and the noise model in the functional are the same. This behaviour is more salient at low BD values, i.e. high noise level. One can also point out that the Chan-Vese functional is very conservative at the price of less sensitivity. Clearly this method under-segments the objects.

6. CONCLUSION

In this work, we proposed a novel statistical region-based active contours method, where the region descriptor is written as a function Φ of some pdf belonging to the exponential family. The case of the likelihood score, which describes the homogeneity of the regions, is obtained as a special case of our setting. We shedded light on the influence of the noise parameters estimator, and on the adequacy between the noise in the observations and the one incorporated in the active contours functional. Our ongoing work is directed towards extending our approach to n regions by adapting the multiphase method [9]. Furthermore, encouraged by our preliminary application results, a deeper experimental work is currently carried out to validate our approach on real data sets.

7. REFERENCES

[1] S. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and bayes/MDL for multiband image segmentation," *IEEE PAMI*, vol. 18, pp. 884–900, September 1996.

[2] N. Paragios and R. Deriche, "Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision," *J. Vis. Comm. and Im. Rep.*, vol. 13, pp. 249–268, 2002.

[3] P. Martin, P. Réfrégier, F. Goudail, and F. Guérault, "Influence of the noise model on level set active contour segmentation," *IEEE PAMI*, vol. 26, no. 6, pp. 799–803, june 2004.

[4] M.C. Delfour and J.P. Zolésio, *Shape and geometries*, Advances in Design and Control, SIAM, 2001.

[5] S. Jehan-Besson, M. Barlaud, and G. Aubert, "DREAM^{2}S: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation," *Int. J. of Comp. Vis.*, , no. 53, pp. 45–70, 2003.

[6] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, "Image segmentation using active contours: Calculus of variations or shape gradients ?," *SIAM Applied Mathematics*, vol. 63, no. 6, pp. 2128–2154, 2003.

[7] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. of Comp. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.

[8] P.J. Bickel and K.A. Docksum, *Mathematical statistics: basic ideas and selected topics*, vol. I, Prentice-Hall, London, 2nd edition, 2001.

[9] L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *Int. J. of Comp. Vis.*, vol. 50, no. 1, pp. 271–293, 2002.