

Geodesic Active Regions: A new framework to deal with frame partition problems in Computer Vision

Nikos Paragios¹

*Siemens Corporate Research,
755 College Road East, Princeton, NJ 08540, USA*

E-mail: nikos@scr.siemens.com

Rachid Deriche

*I.N.R.I.A. (RobotVis)
2004 Route des Lucioles, 06902 Sophia Antipolis , France*

E-mail: der@sophia.inria.fr

This paper presents a novel variational framework for dealing with frame partition problems in Computer Vision by the propagation of curves. This framework integrates boundary and region-based frame partition modules under a curve-based energy framework, which aims at finding a set of minimal length curves that preserve three main properties: (i) they are regular and smooth, (ii) they are attracted by the boundary points (boundary-based information), (ii) and they create a partition that is optimal according to the expected region properties of the different hypotheses (region-based information). The defined objective function is minimized using a gradient descent method. According to the obtained motion equations, the set of initial curves is propagated towards the best partition under the influence of boundary and region-based forces, and being constrained by a regularity force. The changes of topology are naturally handled thanks to the level set implementation. Furthermore, a coupled multi-phase propagation is proposed that imposes the idea of mutually exclusive propagating curves, and increases the robustness as well as the convergence rate. The proposed framework has been validated using three important applications in Computer Vision, the tasks of image and supervised texture segmentation in low-level vision and the task of motion estimation and tracking in motion analysis.

¹This work has been carried out during the appointment (doctoral research) of the first author with the Computer Vision and Robotics Group (RobotVis) of I.N.R.I.A. Sophia Antipolis from October 1, 1996 to November 1, 1999 and was funded in part under the VIRGO research network (EC Contract No ERBFMRX-CT96-0049) of the TMR Program.

1. INTRODUCTION

During the last twenty years a wide variety of mathematical and computational frameworks have been proposed to deal with Computer Vision problems. Their classification is not an easy task. Most of them, are based on the fact that many of the computer vision applications turn out to be frame partition problems (image segmentation, motion estimation and tracking, depth recovery, 3-D reconstruction). According to this assumption, the goal is to decompose the image domain into regions with “*homogeneous properties*”. This “*homogeneity*” is related to the nature of the examined application.

However, feature-based frame partition is performed using two basic image processing techniques: the **boundary-based** methods (which are often referred as edge-based) rely on the generation of a strength image and the extraction of prominent edges, while the **region-based** methods rely on the homogeneity of spatially localized features and properties.

- Early approaches for **boundary-based** frame partition solutions have used local filtering techniques such as edge detection operators. A step further was the use of Snake/Balloons models [8, 12] which can provide a closed curve as a compromise between regularity of the curve and important boundary features. Recently, the geodesic active contour model has been introduced [4, 13] which combined with the level set theory [17] refers to a very elegant and powerful frame partition tool.

- The **region-based** methods are more suitable approaches to deal with frame partition problems and can be roughly classified into two categories: The region-growing techniques [1, 26] and the Markov Random Fields based approaches [2, 9, 10]. The region growing methods are based on split-and-merge procedures using “feature-based” homogeneity tests. Another powerful region-based tool, which has been widely investigated for frame partitioning, is the Markov Random Fields (MRF) [11].

- Finally, there is a significant effort to **integrate boundary-based with region-based frame partition approaches** [6, 26].

Finally, many researchers more recently started to pay attention to novel ways of analyzing, formulating and solving computer vision problems via variational methods including several frame partition problems [3, 14, 16].

Here, we will present a unified approach, namely the **Geodesic Active Region** [18] model to deal with frame partition problems in Computer Vision that incorporates boundary and region information sources under a curve-based minimization framework. This framework assumes that

- i. The considered application refers to a frame partition problem of the image domain into N classes (the number of classes is not fixed),
- ii. Some information regarding the real boundaries between the different frame regions (one for each class) is available or can be estimated in the context of the specific application,
- iii. Some information regarding the expected properties of each class is also available or can be estimated in the context of the specific application (e.g. intensity, motion, etc.).

Then, the desirable frame partition (set of image regions) is obtained by minimizing a curve-based objective function. This function aims at finding a set of regular curves (one for each class) attracted by pixels with important boundary information. At the same time the proposed model aims at finding a set of curves that define regions which preserve the expected properties of the associated class. This function is minimized with respect to the different region boundaries (multiple curves) using a gradient descent method, where the obtained motion equations for the curves propagations are implemented using the level set methods [17]. Moreover, the level set equations are coupled by demanding a non-overlapping set of curves since each pixel of the image cannot belong to more than one region. The resulting model deals automatically with the changes of topology thereby allowing either several sub-regions with the same class properties to be the output of a single initial curve, or a single curve to be the output of multiple initial curves. Various experimental results on different computer vision tasks demonstrate the performance of the proposed framework.

The reminder of this paper is organized as follows. In section 2, the most closely work related with the proposed approach is presented, while in section 3 the main contribution of this paper namely the Geodesic Active Region model is presented. Then, the coupling between the different curves propagations which imposes the concept of mutually exclusive propagations is going to be presented in section 4. Three applications are used for the validation of the proposed framework in section 5. Finally, conclusions and discussion appear in section 6.

2. RELATED WORK

Caselles, Kimmel and Sapiro in [4, 5] and Kichenassamy, Kumar et al in [13] have proposed the *geodesic active contour* model for image segmentation as a geometric alternative for snakes which might be considered as an “extension” of the classic snake since it overcomes an important number of handicaps implied by the snake model. A geometry-based model with similar properties was proposed in [3, 15].

According to this framework the optimal frame partition is obtained by finding the set of minimal length geodesic curves that are attracted by the real region boundaries. Moreover, the same authors have proposed the implementation of the obtained motion equation using the level set methods resulting on a paradigm that can deal automatically with topological changes. Summarizing, the geodesic active contour model is a more elaborated simple scheme for boundary-based frame partition applications that is compared favorably with the classical snake model. However, this model also encounters some weak points: (i) It only makes use of very local information (like the snake model) and is very sensitive to local minima, (ii) Due to the fact that the geodesic active contour framework relies on a non-parameterized curve, and evolves mainly an initial curve towards one direction (constrained by the curvature effect), it demands a *specific* initialization step, where the initial curve should be completely exterior or interior to the real object boundaries.

Many efforts have been made to overcome these shortcomings by introducing some region-based features to snake-based partition methods with objective to make them free from the initial conditions and more robust.

Chakraborty, Staib and Duncan in [6] proposed a model for medical image segmentation that integrates edge and region-based information within a deformable boundary finding framework. This framework introduces a prior knowledge about the shape form (the final contour should be close to this shape), a boundary term that propagates the curve towards high gradient values points (edges), and a region term incorporates some region-based information into the boundary finding framework. However, according to it the boundaries are parameterized using Fourier descriptors, which present some important limitations with respect to the shapes that they can describe. Additionally, the region information is expressed **via intensity homogeneity which limits the model applicability**, even if it is not a strong constraint. Moreover, due the Lagrangian implementation of the curve propagation, the topological changes cannot be solved naturally. Furthermore, this approach requires a pre-segmentation map $[I_{\mathbf{R}}]$. Finally, some of the probabilistic assumptions that have been done to determine the model may be not valid if the current boundary is far away from the true boundary. Thus, the model is sensitive to the **initial conditions**. However, this is a pioneering effort that combines successfully boundary and region-based information within a snake minimization framework.

Zhu and Yuille in [26] proposed a statistical variational approach for image/texture segmentation which combines the geometrical features of a snake/balloon model and the statistical techniques of region growing. In this approach, although a snake/balloon model is employed for the segmentation process, **the boundary-based information is ignored** since the snake/balloon model is used only to impose a regularity constraint to the segmentation process. Besides, due to the implementation of the curve propagation using a Lagrangian approach, the changes of topological are not naturally handled. However, this approach solves this problem by introducing a region growing step. Hence, regions (curves) that have common boundaries are merged if the entropy of the resulting region is inferior to the sum of the regions entropies before merging. In any case, this step can be performed only to region growing approaches and cannot deal with the “splitting” topology change.

More recently, Chan and Vese in [7], Samson et al in [23] and Yezzi, Tsai and Whittsky in [24] have proposed different level set variational frameworks for image segmentation/classification. The work of [7, 24] is constrained to bi-modal and three-modal image segmentation while the work of [23] is based on the propagation of mutually exclusive curves [25] and is constrained to supervised image classification with a predetermined number of regions. However, these frameworks do not make any use of any boundary-based information and are very sensitive to the initial conditions since statistics are generated and dynamically updated over regions dynamically. Furthermore, they are constrained to bi-modal/three-modal or supervised image segmentation/classification cases.

3. GEODESIC ACTIVE REGIONS

The Geodesic Active Region model has been initially introduced in [19, 21] for supervised texture segmentation, has been extended to deal with the un-supervised image segmentation case in [22] and has been successfully exploited in [20] to provide an elegant solution to the motion estimation and the tracking problem.

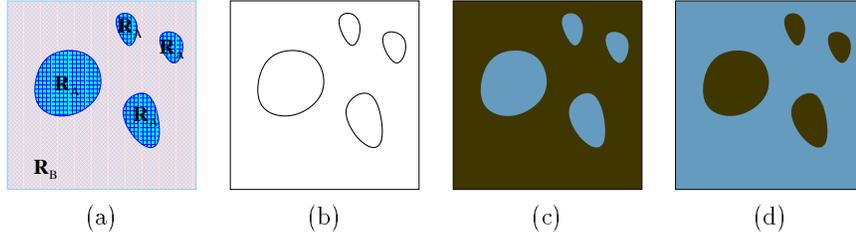


FIG. 1. Geodesic Active Region Model: (a) the input, (b) the boundary-based information, (c) the region-based information corresponding to hypothesis h_A , [the information is proportional to the frame intensities] (d) the region-based information corresponding to hypothesis h_B [the information is proportional to the frame intensities].

In order to facilitate the presentation of the most closely related boundary/region frame partition approaches and the introduction of the new model, the bi-modal case will be considered.

3.1. Notation

Let us make some definitions as well as some assumptions regarding the *a priori* knowledge that are going to be used to introduce the Geodesic Active Region model,

- Let I be the input image composed of two classes (h_A, h_B) ,
- Let $\mathcal{P}(\mathcal{R}) = \{\mathcal{R}_A, \mathcal{R}_B\}$ be a partition of the image domain into two non-overlapping regions [fig. (1.a)],
- Let $\partial\mathcal{P}(\mathcal{R}) = \{\partial\mathcal{R}_A, \partial\mathcal{R}_B\}$ be the boundaries of the $\mathcal{P}(\mathcal{R})$ partition.
- Let us make the assumption that some knowledge regarding the expected positions of real region boundaries is available, the boundary probabilities $p_{b,A}(), p_{b,B}()$ (the b stands for boundary) which measure the likelihood of a given pixel being at the real boundaries of the considered classes (h_A, h_B) [fig. (1.b)],
- Finally, let us also make the assumption that some knowledge regarding the expected region properties of the classes h_A, h_B is available, the region probabilities $p_{r,A}(), p_{r,B}()$ (the r stands for region) which measure the likelihood of a given pixel preserving the expected region properties of the considered classes (h_A, h_B) [fig. (1.c), (1.d)].

3.2. Setting the Boundary Module

Thus, according to this framework, the frame partition task can be viewed initially as the problem of accurately extracting the boundaries of the regions \mathcal{R}_A and \mathcal{R}_B . This, can be done within the geodesic active contour framework, thus minimizing

$$E(\partial\mathcal{R}) = \sum_{X \in \{A, B\}} \int_0^1 g \left(\underbrace{p_{b,X}(I(\partial\mathcal{R}_X(c_x)))}_{X \text{ boundary probability}} \right) \underbrace{|\partial\mathcal{R}_X(c_x)|}_{X \text{ regularity}} dc_x. \quad (1)$$

$X \text{ boundary attraction}$

3.3. Setting the Region Module

At the same time, the examined problem is equivalent to creating a consistent frame partition between *the observed data, the associated hypothesis and their expected properties*. This partition can be viewed as an optimization problem with respect to the *a posteriori* frame partition probability, given the observation set.

Let $[p_S(\mathcal{P}(\mathcal{R})|I)]$ be the *a posteriori* frame partition density function with respect to the different partitions $\mathcal{P}(\mathcal{R})$ given the input image I . This density function is given by the Bayes rule as:

$$p_S(\mathcal{P}(\mathcal{R})|I) = \frac{p(I|\mathcal{P}(\mathcal{R}))}{p(I)}p(\mathcal{P}(\mathcal{R}))$$

where

- $p(I|\mathcal{P}(\mathcal{R}))$ is the *a posteriori* segmentation probability for the image I , given the partition $\mathcal{P}(\mathcal{R})$,
- $p(\mathcal{P}(\mathcal{R}))$ is the probability of the partition $\mathcal{P}(\mathcal{R})$ among the space of all possible partitions of the image domain,
- and $p(I)$ is the probability of having as input the image I among the space of all possible images.

If we assume that all the partitions are *a priori* equally possible $[p(\mathcal{P}(\mathcal{R})) = \frac{1}{Z}]$ where Z is the number of possible partitions, then we can ignore the constant terms $p(I)$, $p(\mathcal{P}(\mathcal{R}))$ and we can rewrite the density function as:

$$p_S(\mathcal{P}(\mathcal{R})|I) = p(I|\{\mathcal{R}_A, \mathcal{R}_B\})$$

Besides, since normally there is no correlation between the regions labeling, and the region probabilities depend only on their observation set (within the region), we obtain the following form

$$p_S(\mathcal{P}(\mathcal{R})|I) = p([I|\mathcal{R}_A] \cap [I|\mathcal{R}_B]) = p(I|\mathcal{R}_A) p(I|\mathcal{R}_B)$$

where $p(I|\mathcal{R}_A)$ is the *a posterior* probability for the region \mathcal{R}_A given the corresponding image intensities (*resp.* $p(I|\mathcal{R}_B)$).

Besides, if we assume that the pixels within each region are independent which is a widely used assumption, then we can replace the region probability by the joint probability among the region pixels:

$$p(I|\mathcal{R}_X) = \prod_{s \in \mathcal{R}_X} p_X(I(s)) \quad (2)$$

where $X \in \{A, B\}$.

Taking all these into account, the *a posteriori* segmentation probability for a partition $\mathcal{P}(\mathcal{R})$ given the observed image I is determined by

$$p_S(\mathcal{P}(\mathcal{R})|I) = \prod_{s \in \mathcal{R}_A} p_A(I(s)) \prod_{s \in \mathcal{R}_B} p_B(I(s)).$$

The maximization of the *a posteriori* segmentation probability is equivalent to the minimization of the $[-\log(\cdot)]$ function of this probability,

$$E(\partial\mathcal{P}(\mathcal{R})) = - \underbrace{\iint_{\mathcal{R}_A} \log \left[\underbrace{p_A(I(x,y))}_{h_A \text{ probability}} \right]}_{\mathcal{R}_A \text{ fitting measurement}} dx dy - \underbrace{\iint_{\mathcal{R}_B} \log \left[\underbrace{p_B(I(x,y))}_{h_B \text{ probability}} \right]}_{\mathcal{R}_B \text{ fitting measurement}} dx dy. \quad (3)$$

Let us now try to interpret this region-based term:

- Suppose that a pixel s is well classified and the true case is h_A ; then this pixel appears to the \mathcal{R}_A fitting measurement. The corresponding probability for the true case $[p_{r,A}(I(s))]$ is higher than the one for the opposite case $[p_{r,B}(I(s))]$, resulting on a minimum contribution in the objective function $[-\log(p_{r,A}(I(s)))]$.
- On the other hand if this pixel is badly classified, then it appears for example to the \mathcal{R}_A fitting measurement while the true case is h_B . It is obvious to see that this pixel will charge more the objective function compared by the charging introduced by the true case.

Summarizing, this region-based energy term is defined using the partition determined by the curve and aims at maximizing the *a posteriori* segmentation probability given the input image.

3.4. Geodesic Active Region Objective function

Then, the two different frame partition modules are integrated by defining the Geodesic Active Region objective function as

$$E(\partial\mathcal{R}) = - \sum_{X \in \{A,B\}} \alpha \underbrace{\iint_{\mathcal{R}_X} \log \left[\underbrace{p_{r,X}(I(x,y))}_{h_X \text{ probability}} \right]}_{\mathcal{R}_X \text{ fitting measurement}} dx dy \quad (4)$$

$$\underbrace{\sum_{X \in \{A,B\}} (1 - \alpha) \int_0^1 g \left(\underbrace{p_{b,X}(I(\partial\mathcal{R}_X(c_x)))}_{X \text{ boundary probability}} \right)}_{X \text{ Boundary Term}} \underbrace{\left| \partial\mathcal{R}_X(c_x) \right|}_{X \text{ regularity}} dc_x.$$

where α is a positive constant that balances the contributions of the two terms $[0 \leq \alpha \leq 1]$.

The interpretation of the defined objective function is following.

A set of curves is demanded that:

- [**Boundary Term**] are *regular* [eqn. (4): regularity], of minimal length, and are attracted by the real regions boundaries [eqn. (4): boundary attraction],

ii. [**Region Term**] define a partition of the image which maximizes the *a posteriori* frame partition probability.

The minimization of the objective function is performed using a gradient descent method. If $u_A = (x_A, y_A)$ is a point of the initial curve $\partial\mathcal{R}_A$ (resp. $u_B = (x_B, y_B)$) and we compute the Euler-Lagrange equations using the Stokes theorem [5, 26], then we should deform the curves ($\partial\mathcal{R}_A, \partial\mathcal{R}_B$) using the following equations:

$$\left\{ \begin{array}{l} \frac{\partial u_A}{\partial t} = \alpha \log \left(\frac{\overbrace{p_{r,B}(I(u_A))}^{h_B \text{ probability}}}{\underbrace{p_{r,A}(I(u_A))}_{h_A \text{ probability}}} \right) \mathcal{N}_A(u_A) + \\ \quad \underbrace{(1 - \alpha) (g(p_{b,A}(I(u_A))) \mathcal{K}_A(u_A) - \nabla g(p_{b,A}(I(u_A))) \cdot \mathcal{N}_A(u_A))}_{A \text{ boundary-based force}} \mathcal{N}_A(u_A) \\ \frac{\partial u_B}{\partial t} = \alpha \log \left(\frac{p_{r,A}(I(u_B))}{p_{r,B}(I(u_B))} \right) \mathcal{N}_B(u_B) + \\ \quad (1 - \alpha) (g(p_{b,B}(I(u_B))) \mathcal{K}_B(u_B) - \nabla g(p_{b,B}(I(u_B))) \cdot \mathcal{N}_B(u_B)) \mathcal{N}_B(u_B). \end{array} \right. \quad (5)$$

The obtained PDE motion equations have two kind of *forces* acting on the propagating curves, both in the direction of the normal.

- **Region force:** This force aims at shrinking or expanding the curve to the direction that maximizes the *a posteriori* frame partition probability. Let us now try to interpret this force for a given curve pixel u_A by reminding that this pixels is associated to the h_A class,

- if the true state of u_A is h_B , then the conditional density function that accounts for this class [$p_{r,B}(I(u_A))$] should support the true case [$p_{r,B}(I(u_A)) > p_{r,A}(I(u_A))$], resulting on a positive force that aims at shrinking the curve to pass through this pixel:

$$\begin{aligned} p_{r,B}(I(u_A)) > p_{r,A}(I(u_A)) &\Rightarrow \frac{p_{r,B}(I(u_A))}{p_{r,A}(I(u_B))} > 1 \\ &\Rightarrow \log \left(\frac{p_{r,B}(I(u_A))}{p_{r,A}(I(u_A))} \right) > 0 \Rightarrow \alpha \log \left(\frac{p_{r,B}(I(u_A))}{p_{r,A}(I(u_A))} \right) > 0 \end{aligned}$$

- On the other hand, if the true state of u_A is the h_A , then this force aims at expanding the curve to include this pixel.

However, it is important to note that the region-based forces have opposite sign in the motion equations of the classes $\{h_A, h_B\}$.

- The **boundary force** contains information regarding the boundaries of the different regions and is composed of two sub-terms; one that shrinks or expands the curve constrained by the curvature effect towards the object boundaries and one that attracts the curve to the objects boundaries (refinement term).

3.5. Model Generalization

The main assumption that has been made to provide the proposed framework relies on the fact that all partitions are equally probable. Generally, this assumption is not valid, but this does not constrain the proposed model since the same framework can be recovered by replacing the *a posteriori* frame partition probability with the *joint* probability. However, is important to note that the maximization of the *joint* segmentation probability is a widely used optimization criterion.

The proposed model can be generalized very easily by assuming for a specific application the existence of functions that capture the boundary and the region properties of the different regions. Thus, if we consider

- A frame partition problem with N classes,
- A set of functions $[b_i : \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R}, i \in [1, N]]$ that “capture” the boundary features of the different classes (**for the real boundary pixels this function returns minima values**),
- A set of functions $[r_i : \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R}, i \in [1, N]]$ that “capture” the region features of the different classes (**for the real class pixels this function returns minima values**),

then, the geodesic active region framework consists of minimizing:

$$\begin{aligned}
 E(\partial\mathcal{R}) = & \sum_{i=1}^N \alpha \underbrace{\int\int_{\mathcal{R}_i} r_i(I(x, y)) dx dy}_{i \text{ Region Term}} \\
 & + \sum_{i=1}^N (1 - \alpha) \underbrace{\int_0^1 \underbrace{b_i(I(\partial\mathcal{R}_i(c_i)))}_{i \text{ boundary attraction}} \underbrace{|\partial\mathcal{R}_i(c_i)|}_{i \text{ regularity}} dc_i}_{\text{Boundary Term}}.
 \end{aligned} \tag{6}$$

The minimization of the generalized objective function is performed using a gradient descent method and leads to a system of N motion equations (one for each class/curve) given by,

$$\left\{ \begin{array}{l} \forall i \in [1, N] \\ \frac{\partial u_i}{\partial t} = \alpha \underbrace{(r_i(u_i) - r_{o_i}(u_i)) \mathcal{N}_i(u_i)}_{i \text{ region force}} + \\ \quad \underbrace{(1 - \alpha) (b_i(u_i) \mathcal{K}_i(u_i) - \nabla b_i(u_i) \cdot \mathcal{N}_i(u_i)) \mathcal{N}_i(u_i)}_{i \text{ boundary force}}. \end{array} \right. \tag{7}$$

where the assumption that the pixel u_i lies between the regions \mathcal{R}_i and \mathcal{R}_{o_i} has been made implicitly. According to the above equations the a given curve is propagated along his normal direction under the influence of two forces:

- A region-based force that moves the curves towards the direction that creates the optimal frame partition (according to “*region properties*”) result using the the observation set and the expected properties of the different classes,

- A boundary-based force that shrinks the curve under the influence of a regularity constraint (curvature effect) towards the different region boundaries (according to the “*boundary properties*”).

Then, the proposed framework is employed as follows: Initially a set of random curves is used to initialize the region positions. Then, each region is deformed according to the corresponding motion [eq. (7)] towards the final frame partition. The interaction between the regions positions is obtained through the region-based force since for a given pixel that is attributed to two different regions, forces with opposite signs appear to the corresponding motion [PDE] equations.

3.6. Level Set Methods

The obtained motion equations can be implemented using a difference approximation scheme (Lagrangian approach). However, for this approach the evolving model is not capable to deal with topological changes of the moving front. This can be avoided by introducing the pioneering work of *Osher and Sethian* [17], the level set theory where the central idea is to represent the moving front $\partial\mathcal{R}(c, t)$ as the zero-level set

$$\phi(\partial\mathcal{R}(c, t), t) = 0$$

of a function ϕ . This representation of $\partial\mathcal{R}(c, t)$ is implicit, parameter-free and intrinsic. Additionally, it is topology-free since different topologies of the zero level-set do not imply different topologies of ϕ . It can be shown easily that if the moving front evolves according to

$$\frac{\partial}{\partial t}\partial\mathcal{R}(c, t) = F(\partial\mathcal{R}(c, t)) \mathcal{N}$$

for a given function F , then the embedding function ϕ deforms according to

$$\frac{\partial}{\partial t}\phi(c, t) = F(c) |\nabla\phi(c, t)|$$

For this level-set representation, it is proved that the solution is independent of the embedding function ϕ which in most of the cases is initialized as a signed distance function.

Thus, the system of motion equations that drives the curve propagation for the Generalized Geodesic Active Region framework is transformed into a system of multiple surfaces evolution given by,

$$\left\{ \begin{array}{l} \forall i \in [1, N] \\ \frac{\partial\phi_i}{\partial t}(u) = \underbrace{\alpha (r_i(u) - r_{o_i}(u)) |\nabla\phi_i(u)|}_{i \text{ region force}} + \\ \underbrace{(1 - \alpha) \left(b_i(u)\mathcal{K}_i(u) + \nabla b_i(u) \cdot \frac{\nabla\phi_i(u)}{|\nabla\phi_i(u)|} \right) |\nabla\phi_i(u)|}_{i \text{ boundary force}} \end{array} \right. \quad (8)$$

In order to demonstrate the proposed model, the task of image segmentation is considered for a synthetic frame that is composed of two classes [fig. (2)].

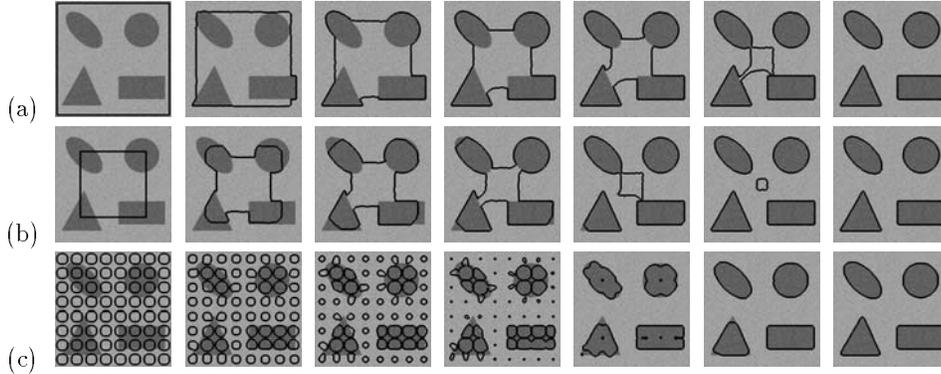


FIG. 2. Geodesic Active Regions: A new framework to deal with frame partition problems in Computer Vision. The proposed framework is used for image segmentation where the independence from the initial conditions and the ability to deal with topological changes are demonstrated.

- The first $[h_B]$ refers to the background and is composed from pixels with intensities that follow a Gaussian distribution with a mean value equal to 150 and a standard deviation equal to 10 $[p_B() \sim G(150, 10)]$,
- The second $[h_A]$ refers to a region composed of four components, and pixels with intensities that follow Gaussian distribution with a mean value equal to 90 and a standard deviation equal to 10 $[p_A() \sim G(90, 10)]$,

For this case, the boundary information is determined using a Gaussian edge detector, applied on the norm of the gradient values space. The curve evolution for the class h_A with respect to three different initialization steps is illustrated in [fig. (2)].

Besides, to demonstrate the independence of the proposed framework from the initial conditions, we show the evolution of the energy terms over time for the class A . Thus, in [fig. (3)] the boundary-based energy term, the region-based energy term as well as the corresponding values of the total energy for the first two initialization steps [fig. (2.a),(2.b)] are shown.

4. MUTUALLY EXCLUSIVE PROPAGATING CURVES

The use of the level set methods provides a very elegant tool to propagate curves where their position is recovered by seeking for the zero level set crossing points. Moreover, the state of given pixel with respect to a region hypothesis can be easily determined since if it belongs to the region, then the corresponding level set value is negative. On the other hand if it does not belong to it, then the corresponding value is positive. Additionally, since we consider signed distance functions for the level set implementation, a step further can be done by estimating the distance of the given pixel from each curve. This information is valuable during the multi-phase curve propagation cases where the overlapping between the different curves is prohibited.

However, the overlapping between the different curves is almost an inevitable situation at least during the initialization step. Moreover, the case where an image pixel has not been attributed to any hypothesis may occurs. Let us now assume

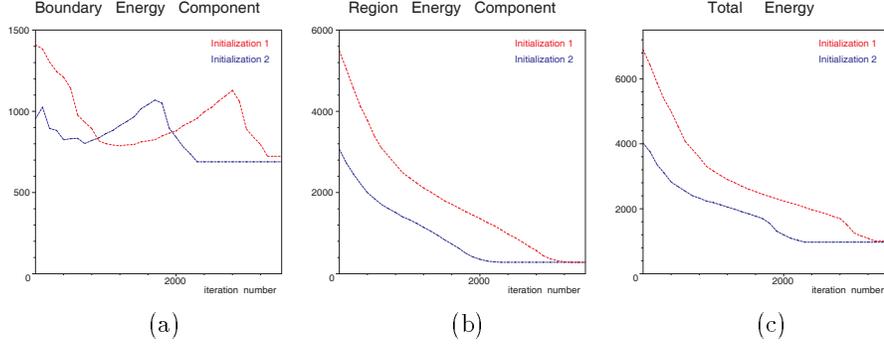


FIG. 3. Evolution of the Geodesic Active Region Energy for the [Fig. (2)] example. (a) The evolution of the boundary energy term, (b) the evolution of the region energy term, (c) the evolution of the total energy. The different colors of the plotted function correspond to the different initial curve conditions: the red color refers to [fig. (2.a)], while the blue color to [fig. (2.b)].

that a pixel is attributed initially to two different regions (there are two level set functions with negative values at it). Then, a constraint that discourages a situation of this nature can be easily introduced, by **adding an artificial force (always in the normal direction)** to the corresponding level set motion equations that penalizes pixels with multiple labels (they are attributed to multiple regions). Moreover, a similar force can be introduced to discourage situations where pixels are not attributed to any regions. Inspired by the work in [25], this can be done by modifying the motion equations [eqn. (7)] as,

$$\left\{ \begin{array}{l} \forall i \in [1, N] \\ \frac{\partial \phi_i}{\partial t}(u) = \beta \underbrace{(r_i(u) - r_{o_i}(u)) |\nabla \phi_i(u)|}_{i \text{ region force}} + \\ \gamma \underbrace{\sum_{j \in [1, N]} H_i(i, \phi_j(s)) |\nabla \phi_i(u)|}_{i \text{ coupling force}} + \\ \delta \underbrace{\left(b_i(u) \mathcal{K}_i(u) |\nabla \phi_i(u)| + \nabla b_i(u) \cdot \frac{\nabla \phi_i(u)}{|\nabla \phi_i(u)|} \right)}_{i \text{ boundary force}} \end{array} \right.$$

where β, γ, δ are positive constants [$\beta + \gamma + \delta = 1$], and the function $H_i(j, \phi())$ is given by

$$H_i(m, \phi_n(s)) = \begin{cases} 0, & \text{if } m = i \\ -\text{sign}(\phi_j(u)), & \text{if } m \neq i \end{cases}$$

The introduction of this new force is inspired by the work presented in [25]. Let us now interpret the new artificial force that has been added to the motion equation i , for a given pixel u :

Expanding Effect:

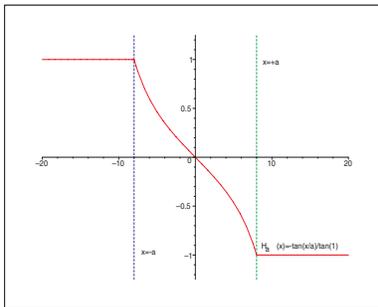


FIG. 4. The trigonometric basis of the level set coupling function.

If this pixel does not belong to any region, (the corresponding level set values at s are positive for all level set functions), then the new force is negative, equal to $f_c = -(N - 1)|\nabla\phi_i|$ and aims at expanding the region \mathcal{R}_i to occupy this pixel (appearance of non-attributed pixels is discouraged).

Shrinking Effect:

On the other hand, if this pixel has been already attributed to another region $[\mathcal{R}_k]$, then the corresponding level set function $[\phi_k]$ will contribute with a positive force that aims at shrinking locally the region \mathcal{R}_i (the overlapping is discouraged).

To summarize, this coupling force for a given class i is negative (*expanding*) to a considered pixel u if it is attributed to any region. On the other hand this force is positive (*shrinking*) if u is already attributed to another region j .

Although the selection of the function $[H_i(\phi())]$ seems to be proper, it introduces some problems. First, the not-attributed pixels are penalized with the same manner similarly with the ones that have been attributed to multiple regions. Second, their distance values of a given pixel from the other curves are not considered and hence valuable information is lost. Finally, the defined coupling function is discontinuous which is a not desirable property since it creates stability problems during the level set evolution.

To summarize, the coupling function has to be redefined by taking into account the following considerations:

i. A pixel that is already attributed to a region j and is far away from $\partial\mathcal{R}_j$, should strongly discourage the evolution of the level set $\phi_i()$ to include this pixel in \mathcal{R}_i (the coupling force should be in terms of absolute value proportional to the distance from the region boundaries),

ii. On the other hand, a pixel which belongs to the region \mathcal{R}_j and is close to its boundaries (small level set value $\phi_j()$), can be reached or be liberated by $\partial\mathcal{R}_j$ during the next few iterations, and hence, the coupling force in i propagation because of the j level set function (region) should be less powerful and “tolerate” a temporal overlapping.

Thus, inspired by the properties of the trigonometric functions, the following function that is going to be used as basis for the new coupling force [fig. (4)] is

defined

$$H_a(x) = - \begin{cases} +1, & \text{if } x > a \\ -1, & \text{if } x < -a \\ \frac{1}{\tan(1)} \tan(x/a), & \text{if } |x| \leq a \end{cases} \quad (9)$$

with $a \in \mathcal{R}^+$. Then the coupling force is defined as,

$$H_i(j, \phi_j(u)) = \begin{cases} 0, & \text{if } j = i \\ H_a(\phi_j(u)), & \text{if } j \neq i \text{ and } \phi_j(u) \geq 0 \\ \frac{1}{N-1} H_a(\phi_j(s)), & \text{if } j \neq i \text{ and } \left[\bigcap_{k=1, k \neq i}^N \phi_k(u) > 0 \right] \end{cases} \quad (10)$$

To interpret this force via the new function, a level set function $[\phi_i(\cdot)]$ and a pixel location $[u]$ are considered,

i. If u is already attributed to another region, then there is an hypothesis j for which $\phi_j(u) \leq 0$ which will contribute with a positive value (shrinking effect) to the coupling force that is proportional to the distance of this pixel from the boundaries of \mathcal{R}_j ,

ii. A similar interpretation can be done if this pixel is not attributed to any region (expanding effect). However, for this case the coupling force has to be normalized because is not appropriate to penalize with the same way the situation of overlapping and the case in which the given pixel is not attributed to one of the regions. At the same time this force is plausible if and only if this pixel is not attributed to any region $\left[\bigcap_{k=1, k \neq i}^N \phi_k(u) > 0 \right]$.

More details on the interpretation/justification of the above selection can be found in [18, 22]. Another related approach dealing with the ideas of mutual exclusiveness, introduced in [25], can also be found in [23].

5. APPLICATIONS

The proposed framework has been used as basis to provide original solutions to three important applications in Computer Vision, the tasks of image and supervised texture segmentation in low level vision and the task of motion estimation and tracking in low level vision.

5.1. Image Segmentation

In [22], a two stage approach for image segmentation has been proposed that is exploited directly from the **Geodesic Active Region** model and incorporates boundary and region information sources.

The first stage refers to a modeling phase where the observed histogram is approximated using a mixture of Gaussian components. This analysis denotes the regions number as well as their statistics, since a Gaussian component is associated with each region.

Then, the segmentation is performed by employing the Geodesic Active Region model. The different region boundaries are determined using a probabilistic module by seeking for local discontinuities on the statistical space that is associated with

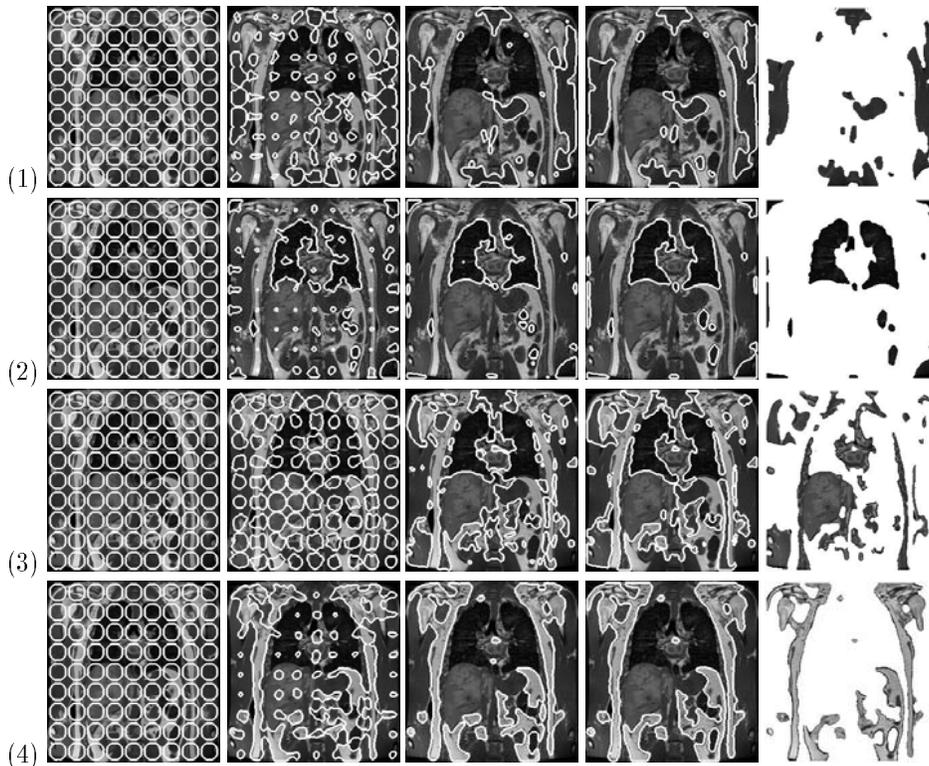


FIG. 5. Geodesic Active Regions for Image Segmentation. A medical image is considered which is composed of four classes that are multi-component (each class contains more than one connected regions). (1) Curve Propagation for Class 1, (2) Curve Propagation for Class 2, (3) Curve Propagation for Class 3, (4) Curve Propagation for Class 4.

the image features. This information is combined with the region one, which is expressed directly from conditional probabilities, resulting in a geodesic active region based segmentation framework. The defined objective function is minimized with respect to the different region boundaries (multiple curves) using a gradient descent method, where the obtained equations are implemented using the level set theory. Moreover, the evolutions of the level set functions are coupled by demanding a non-overlapping set of curves since each pixel of the image cannot belong to more than one region. The resulting model deals automatically with the changes of topology thereby allowing either several sub-regions with the same intensity properties to be the output of a single initial curve, or a single curve to be the output of multiple initial curves.

Some experimental results obtained with proposed framework are shown in [fig. (5)], while more details can be found in [22].

5.2. Supervised Texture Segmentation

In [19, 21] a new paradigm for supervised texture segmentation has been proposed. This framework is composed of a learning and a segmentation phase.

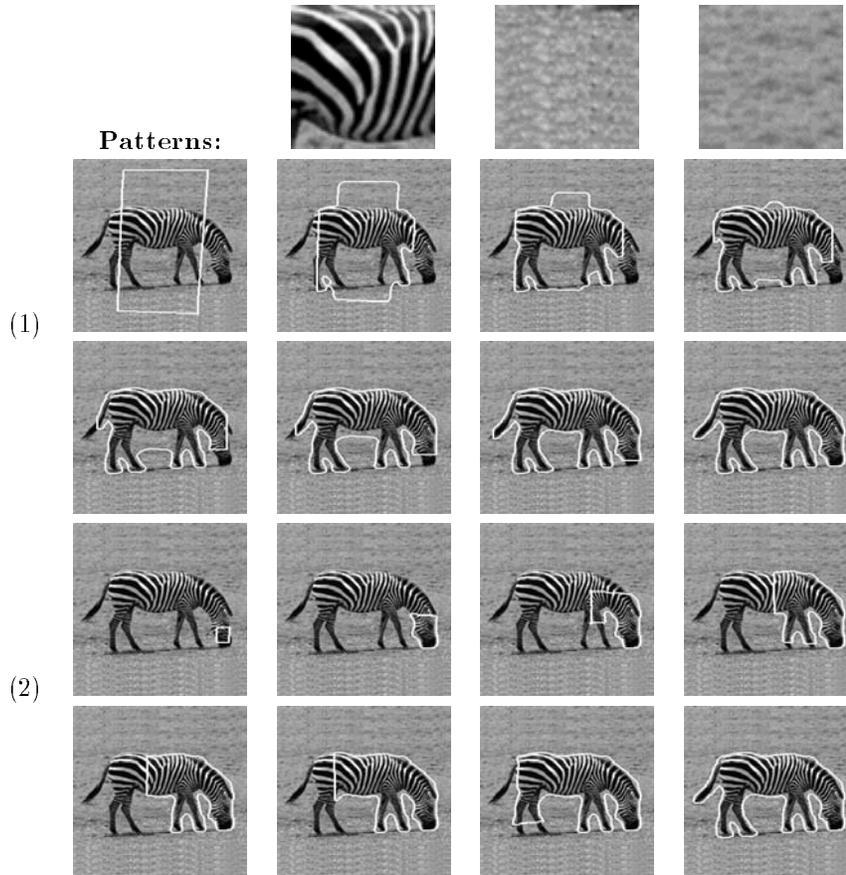


FIG. 6. Geodesic Active Regions for Supervised Texture Segmentation. Demonstration for a real image where the goal is to distinguish the region of interest (zebra) from the other texture regions (the interior curve area corresponds to the zebra). Two different initialization steps are demonstrated.

Initially, an off-line step is performed that creates multi-component probabilistic texture descriptors for the given set of texture patterns, where the multi-modal data is derived using a set of filter operators. The texture analysis/modeling phase consists of three steps. A set of optimally selected predefined filters (mainly Gabor) is applied to each texture pattern (*learning*). Then, each filter response is modeled statistically using a multi-component conditional probability density function (*modeling*). These conditional probabilities are assumed to be Gaussian mixtures with two or three components. Based on this analysis, each data vector from the multi-modal filter response space has an associated tuple of probabilities, that express the probability that this vector belongs to a given texture pattern. Furthermore, the different filter operators are associated with some reliability measurements (*validating*). These measurements are estimated according to their discrimination power. The discrimination power of a filter operator is inversely proportional to the misclassification error that is provided by this operator during the learning phase.

Then, given the input image, we apply the same operators and derive an observation set that is coherent with the texture descriptors. Then, for each pixel we estimate the probability of being on the boundaries between two different texture regions. Since we deal with multi-modal data, a probability vector is obtained. The components of this vector (*e.g.* boundary probabilities) are combined to a single frame using some reliability measurements and provide the **boundary-based** texture information. Besides, using the texture descriptors and the observation set we determine the **region-based** information that is derived from the most probable temporal texture assignment. Then, the segmentation problem is stated under the Geodesic Active Region model, that aims at finding best minimal length geodesic curves that preserve high boundary probabilities, and creates regions with maximum *a posteriori* segmentation probability with respect to the associated texture hypothesis. The defined objective function is minimized using a gradient-descent where the initial curves are propagated by means of velocities that contain three terms: one that shrinks or expands the curve toward the region boundaries, one that supports the homogeneity of the interior curve region given the associated texture hypothesis, and one that expresses the expected curve spatial properties (*i.e.* *smoothness, regularity*). These motion equations are implemented using the level set methods, where a coupled multi-phase propagation is considered to impose the idea of mutually exclusive propagating curves.

Some experimental results obtained with proposed framework are shown in [fig. (6)], while more details can be found in [19, 21].

5.3. Motion Estimation and Tracking

In [20] a general motion estimation and tracking framework is proposed derived from the Geodesic Active Region model that

- i. *Deals simultaneously with the motion estimation and tracking problem,*
- ii. *Makes use in a generic form of different boundary and region-based tracking modules,*
- iii. *Deals with the case of non-rigid as well as occluding moving objects,*
- vi. *And, can be easily extended to cases involving a mobile observer.*

Thus, within this framework the motion estimation and the tracking problem are coupled, resulting in a two direction multi-module front propagation tracking system, where the motion parameters and the exact object position are simultaneously recovered.

To summarize, the motion estimation and the tracking problems are reformulated within the Geodesic Active Region framework, where boundary and region-based modules are involved. The boundary-based information is determined using a combination of a gradient and a probabilistic edge detector while the region-based one using three different modules; a **motion detection/segmentation**, an **intensity-based** and a **visual consistency** module. The observed difference density function (histogram) is analyzed as a mixture density of Gaussian or Laplacian elements that correspond to the static pixels (background) and the the mobile pixels (objects). The output of this statistical modeling phase is used to define the **Motion Detection Module**. Besides, using some knowledge regarding the background as well as the moving intensity properties (it is acquired during the process) we define

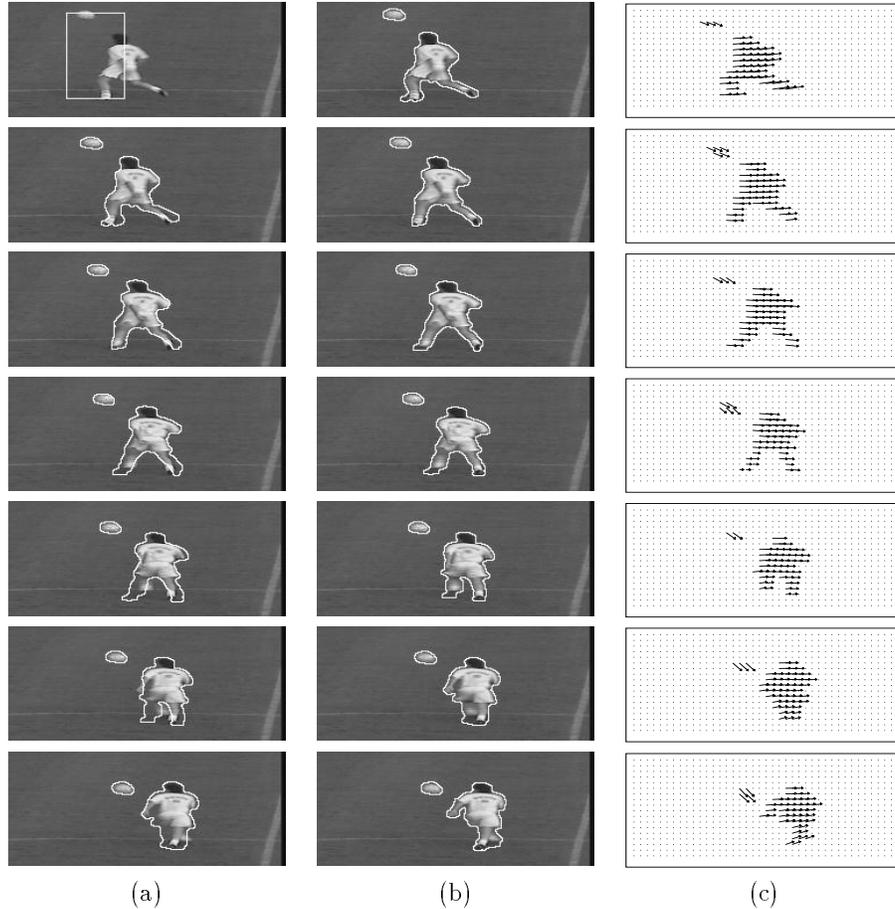


FIG. 7. Geodesic Active Regions for Motion Estimation and Tracking. Soccer Sequence:(a) Initial Curve , (b) Final Curve, (c) Motion Estimation.

the intensity-based tracking module. Finally, given the current curve position, an affine motion model that creates a visual consistency between the object intensities in the current and the previous frame is estimated. The designed objective function is minimized with respect to the curve position as well as to the motion parameters using a gradient descent method. The obtained partial differential motion equations [PDEs] that deform the initial curves, are composed of *boundary*, *intensity* and *motion-based* forces and are constrained by a *regularity* force. These PDEs are implemented using the level set theory while an incremental method is used to update the estimates of the motion parameters. To deal with occlusions, a multi-phase curve propagation is also considered that couples the propagation of curves with respect to the different objects.

Some experimental results obtained with proposed framework are shown in [fig. (7)], while more details can be found in [20].

6. DISCUSSION, SUMMARY

To summarize, in this paper a new energy minimization-based variational framework is proposed to deal with frame partition problems. This framework, namely the **Geodesic Active Region** lies on a very flexible model that integrates boundary and region-based information modules under a generic objective function. The optimization of this framework leads to the curve propagation theory, where a set of initial curves is propagated towards the regions boundaries being constrained by internal forces, while at the same time it is used to determine a coherent partition of the image domain according to some region-based properties. Moreover, we have proposed the implementation of the curve propagation using the level set methods, a very elegant tool that enables a large number of very nice properties, and makes the propagation of curves easier. These methods combined with the proposed model gives a paradigm with a very large applicability set, free from the initial conditions, and of extreme power.

This paradigm was introduced for a very general frame partition case and it can be used for a large variety of computer vision applications, the ones that can be re-formulated as frame partition problems. The efficiency and the performance of the proposed framework have been validated using three important applications in computer vision, the tasks of image and supervised texture segmentation in low level vision and the task of motion estimation and tracking in motion analysis.

The main contributions of this paper consists of

- A curve-based variational framework capable of dealing with frame partition problems that integrates boundary and region-based information modules,
- The connection between the minimization procedure of the objective function with the curve propagation theory,
- The implementation of this framework using the level set theory resulting on a model of extreme power and free from the initial conditions.

As far the future directions of this work, the incorporation to the model of a term that accounts for some shape prior knowledge with respect to expected frame partition map is a challenge (self-constrained geodesic active regions) as well as the extension of the use of the proposed framework to deal with application in three dimensions.

Various experimental results (in MPEG format), including the ones shown in this article, can be found at:

<http://www.inria.fr/robotvis/personnel/der/demos>

REFERENCES

1. R. Adams and L. Bischof. Seeded Region Growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16:641–647, 1994.
2. C. Bouman and M. Shapiro. A multiscale random field model for bayesian image segmentation. *IEEE Transactions on Image Processing*, 3:162–177, 1994.
3. V. Caselles, F. Catte, T. Coll, and F. Dibos. A Geometric Model for Active Contours. *Numerische Mathematik*, 66:1–31, 1993.
4. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *IEEE International Conference on Computer Vision*, Boston, USA, 1995.
5. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22:61–79, 1997.

6. A. Chakraborty, H. Staib, and J. Duncan. Deformable Boundary Finding in Medical Images by Integrating Gradient and Region Information. *IEEE Transactions on Medical Imaging*, 15(6):859–870, 1996.
7. T. Chan and L. Vese. An Active Contour Model without Edges. In *International Conference on Scale-Space Theories in Computer Vision*, pages 141–151, 1999.
8. D. Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, 53:211–218, 1991.
9. H. Derin, H. Elliott, R. Cristi, and D. Geman. Bayes smoothing algorithms for segmentation of binary images modeled by markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:1186–1191, 1984.
10. D. Geiger and A. Yuille. A common framework for image segmentation. *International Journal of Computer Vision*, 6:227–243, 1991.
11. S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
12. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–332, 1988.
13. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. In *IEEE International Conference on Computer Vision*, pages 810–815, Boston, USA, 1995.
14. G. Koepfler, C. Lopez, and J.-M. Morel. A multiscale algorithm for image segmentation by variational framework. *SIAM Journal on Numerical Analysis*, 31:282–299, 1994.
15. R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:158–175, 1995.
16. D. Mumford and J. Shah. Aptimal Approximations by Piecewise Smooth Functions and Variational Methods. *Communication on Pure and Applied Mathematics*, XLII:577–685, 1988.
17. S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed : algorithms based on the hamilton-jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.
18. N. Paragios. *Geodesic Active Regions and Level Set Methods: Contributions and Applications in Artificial Vision*. PhD thesis, School of Computer Engineering, University of Nice/ Sophia Antipolis, Jan. 2000. <http://www-sop.inria.fr/robotvis/personnel/nparagio/papers/thesis.ps.gz>.
19. N. Paragios and R. Deriche. Geodesic Active Regions for Texture Segmentation. Technical Report 3440, INRIA, France, 1998. <http://www.inria.fr/RRRT/RR-3440.html>.
20. N. Paragios and R. Deriche. Geodesic Active regions for Motion Estimation and Tracking. In *IEEE International Conference on Computer Vision*, pages 688–674, Corfu, Greece, 1999.
21. N. Paragios and R. Deriche. Geodesic Active regions for Supervised Texture Segmentation. In *IEEE International Conference on Computer Vision*, pages 926–932, Corfu, Greece, 1999.
22. N. Paragios and R. Deriche. Coupled Geodesic Active Regions for Image Segmentation: A Level Set Approach. In *European Conference in Computer Vision (to appear)*, Dublin, Ireland, 2000.
23. C. Samson, L. Blanc-Feraud, G. Aubert, and J. Zerubia. A Level Set Model for image classification. In *International Conference on Scale-Space Theories in Computer Vision*, pages 306–317, 1999. <http://www.inria.fr/RRRT/RR-3662.html>.
24. A. Yezzi, A. Tsai, and A. Willsky. A Statistical Approach to Snakes for Bimodal and Trimodal Imagery. In *IEEE International Conference on Computer Vision*, pages 898–903, Corfu, Greece, 1999.
25. H.-K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multi-phase motion. *Journal of Computational Physics*, 127:179–195, 1996.
26. S. Zhu and A. Yuille. Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:884–900, 1996.