IEEE CVPR-99, pages xx-yy, Colorado, USA

Unifying Boundary and Region-based information for Geodesic Active Tracking*

Nikos Paragios Rachid Deriche
I.N.R.I.A
BP 93, 2004 Route des Lucioles
06902 Sophia Antipolis Cedex, France
e-mail: *{nparagio,der}@sophia.inria.fr*

Abstract

This paper addresses the problem of tracking several non-rigid objects over a sequence of frames acquired from a static observer using boundary and region-based information under a coupled geodesic active contour framework. Given the current frame, a statistical analysis is performed on the observed difference frame which provides a measurement that distinguishes between the static and mobile regions in terms of conditional probabilities. An objective function is defined that integrates boundary-based and region-based module by seeking curves that attract the object boundaries and maximize the a posteriori segmentation probability on the interior curve regions with respect to intensity and motion properties. This function is minimized using a gradient descent method. The associated Euler-Lagrange PDE is implemented using a Level-Set approach, where a very fast front propagation algorithm evolves the initial curve towards the final tracking result. Very promising experimental results are provided using real video sequences.

1 Introduction

The tracking problem has a wide variety of applications in computer vision and motion analysis, including coding, video surveillance, robotics, etc. Besides, it provides a good basis for high level tasks of computer vision like 3-D reconstruction and 3-D representation.

During the last years, a great variety of tracking algorithms have been proposed. They may be classified in two distinct categories:

1. The **Motion-based** approaches that rely on a robust method for grouping visual motion consistencies over time [12].
2. And the **Model-based** approaches that impose high-level semantic representation and knowledge [5, 9].

In both cases the tracking is performed using measurements provided by geometrical or region-based properties of the tracked object. In this direction there are two main approaches: the **boundary-based** (they usually referred to as edge-based approaches) rely on the information provided by the object boundaries (shape properties) [6, 8, 11] while the **region-based** approaches rely on information provided by the entire region [1, 12] (texture and motion-based properties). Besides, there are some efforts to combine the boundary and region-based modules under a common framework [2].

The objective of the proposed work is to provide a general framework that incorporates boundary-based and region-based tracking modules under the geodesic active contour framework. Following our previous work on tracking using geodesic active contours [15], we propose a considerable extension that incorporates region-based information to the existing boundary-based information. This extension increases the robustness and liberate the model from the initialization step.

We assume that the difference frame (between the current and the background reference frame) is composed of two populations, one that corresponds to the static part (static background) and one that corresponds to the mobile part (moving targets). Based on this hypothesis, we model the observed probability density function of the difference frame using two multi-components elements (mixture distribution). Given this analysis, we propose a new method to determine the object boundaries as discontinuities on the motion detection feature space. Additionally, we express the region-based information via conditional probabilities with respect to intensity and motion properties. Finally, we incorporate the two information sources to a coupled geodesic active contour model which leads to a system where the two modules operate simultaneously, namely a **Geodesic Active Region** model [14]. The objective function is minimized using a gradient descent method while the

*This work was funded in part under the VIRGO research network (EC Contract No ERBFMRX-CT96-0049) of the TMR Program.

obtained PDE is implemented using a Level-Set approach where topological changes are naturally handled [13]. Finally, the level-set propagation is performed using a very fast front propagation algorithm [15].

The remainder of this paper is organized as follows. In Section 2, we propose the modeling of the the intensity and the motion-based tracking information using low level statistics, while in Section 3 we introduce Geodesic Active Region model for tracking. The minimization of model appears in Section 4. Finally, Section 5 contains some experimental results, followed by conclusions and discussion.

2 Statistical Modeling

The most common way to perform tracking relies on optical flow estimation, that requires high computational cost if no motion models are considered. This cost is significantly decreased using linear motion models that make certain assumptions related to the type of motion as well as to the type of the moving objects and have difficulty to deal with non-rigid objects. We propose an alternative technique to determine the motion based information that models the intensity space as well as the motion detection/segmentation space with low level statistics.

2.1 Intensity-based Statistics

Using the background reference frame $[\mathbf{R}(x,y)]$, we apply a set of predefined filter operators to extract some intensity-based features and properties. These operators are: the intensity, the x and y directional derivative operators, and the magnitude of the gradient smoothed by a Gaussian $[F = \{I, D_x, D_y, G\}]$.

The operator responses are modeled using low level statistics, where the conditional probability density functions are expressed directly from the observed histograms. The output of this operation is a tuple of conditional probability density functions

- $\vec{\mathbf{p}}_0(\vec{\mathbf{x}}) = (p_{0I}(x_I), p_{0D_x}(x_{D_x}), p_{0D_y}(x_{D_y}), p_{0G}(x_G))$

where $p_{0I}(.)$ (resp. $p_{0D_x}(.), p_{0D_y}(.), p_{0G}(.)$) corresponds to the intensity (resp. directional derivatives, gradient) operator.

Besides, if we assume that there are N moving objects in our scene and have been temporally tracked, then we can produce the same statistical modules with respect to them:

- $\vec{\mathbf{p}}_i(.) = (p_{iI}(.), p_{iD_x}(.), p_{iD_y}(.), p_{iG}(.)),\ i \in [1, N]$.

2.2 Motion-based Statistics

Let $\mathbf{I}(x,y)$ be the current frame, and

$$\mathbf{D}(x,y) = \mathbf{I}(x,y) - \mathbf{R}(x,y)$$

the current difference frame. If we assume that this frame is a selection of independent pixels, then it is composed of two populations. The *static* that contains the background pixels (with low difference values) and the *mobile* that contains the moving objects pixels [17]. Besides, we assume that the *mobile* population can be decomposed into a sum of sub-populations with respect to the different intensity properties preserved by the moving objects. These assumptions can be easily projected to a statistical model, where the observed density function can be decomposed into two main statistical components, the static and the mobile

$$p_D(d) = P_{static} p_S(d) + P_{mobile} p_M(d)$$

where P_{static} (resp. P_{mobile}) is the *a priori* probability for the *Static* (resp. *Mobile*) case. Additionally, we can consider that the conditional probability density function with respect to the *mobile* component is a collection of subcomponents that express the different illumination properties of the observed objects, given by

$$p_M(d) = \sum_{i=0}^{C_N} P_{i,M} p_M(d|i)$$

where $P_{i,M}$ is the *a priori* probability of the i component. Finally, we assume that these probability density functions follow Gaussian or Laplacian law. As for the unknown parameters of this model, some constraints are imposed by the problem. The differences between background values appear because of the presence of noise, and as a consequence, the conditional probability density function with respect to the static case is zero-mean. Additionally, we can assume that the mobile mixture model contains a zero-mean density function due to the fact that some moving objects may preserve similar intensity properties with respect to the background. The estimation of the unknown parameters of this model $(P_{static}, P_{mobile}, \sigma_{static}, \sigma_0, \{P_{i,M}, \mu_i, \sigma_i\} : i = 1, ..., C_N)$ is done using the maximum likelihood principle [4].

3 Setting the Energy

The geodesic active contour models and the level set methods have been well studied and successfully applied to a rich variety of computer vision applications. These models are based on boundary-based information, and aim at finding the best minimal-length smooth curve that takes into account the desired image properties. Based on work developed in [3, 7, 10], we are going to reformulate the problem of tracking within the framework of curve evolution theory.

In such a case, we view tracking as a frame partition problem, since we would like to incorporate boundary and region-based features which can be easily expressed within this framework. In order to facilitate the notation, let us now make some definitions:

- Let $\mathcal{P}(R)$ be partition of the image domain R into N regions,
- Let R_0 be the static region, and let R_i be the region that corresponds to object O_i,
- Let ∂R_i be the boundary of region R_i, and let $\partial\mathcal{P}(R) = \{\cup_{i=1}^{N} \partial R_i\}$.

The geodesic active contour framework consists of minimizing:

$$E(\partial\mathcal{P}(R)) = \sum_{i=1}^{N}\left\{\underbrace{\beta\int_0^1 |\partial\dot{\mathrm{R}}_i(p_i)|^2 dp_i}_{Regularity} + \underbrace{(1-\beta)\int_0^1 g^2(|\nabla\mathbf{I}(\partial\mathbf{R}_i(p_i))|)dp_i}_{Image}\right\}$$

where $\partial R_i(p_i) : [0,1] \rightarrow \mathbf{R^2}$ is a parameterization of the region boundaries R_i in a planar form, the dot operator denotes the partial derivative with respect to time, and β is a real positive constant $\{0 \leq \beta \leq 1\}$. Besides, the regularity component accounts for the expected spatial properties (*i.e. smoothness*) of the curve, while the image component stands for the *attraction* energy term of the curve towards the object boundaries. Finally, $g(.)$ is a monotonically decreasing function such that $g(r) \rightarrow 0$ as $r \rightarrow \infty$ and $g(0) = 1$ (*e.g*, Gaussian).

The minimization of the above objective function leads to a geodesic curve with a new metric,

$$E(\partial\mathcal{P}(R)) = \sum_{i=1}^{N}\int_0^1 g(|\nabla\mathbf{I}(\partial\mathbf{R}_i(p_i))|)|\partial\dot{\mathrm{R}}_i(p_i)|dp_i$$

3.1 Setting the Boundary Module

An obvious way to provide the boundary-based information is to seek high gradient values in the current or in the difference frame. The use of an edge-based detector on the input frame could provide reliable boundary-based information for the moving objects, but it will provide similar information for the high contrast parts of the static background. On the other hand the use of the difference frame will eliminate the static background edges but it will provide high boundary information inside the moving objects due to the fact that their intensities are not homogeneous.

In order to deal with these problems we propose an alternative method for the extraction of boundary-based information. Let s be a point of the frame domain, $N_R(s)$ and $N_L(s)$ be the regions associated with a neighborhood partition, and let $\mathbf{D}(N(s))$ be the corresponding difference data. Using the Bayes rule, the boundary probability $p(B|\mathbf{D}(N(s)))$ is given by :

$$\begin{aligned} p(B|\mathbf{D}(N(s))) &= \frac{p(\mathbf{D}(N(s))|B)}{p(\mathbf{D}(N(s)))}p(B) \\ &= \frac{p(\mathbf{D}(N(s))|B)}{p(\mathbf{D}(N(s))|B\cup\bar{B})}p(B) \end{aligned}$$

where $p(\mathbf{D}(N(s))|B)$ (*resp.* $p(\mathbf{D}(N(s))|\bar{B})$) is the conditional boundary (*resp.* non-boundary) probability. Besides, $p(B)$ is the *a priori* boundary probability which can be ignored due to the fact that it is a constant scale factor.

The conditional boundary probability can be easily estimated since if s is a boundary point there is a partition $[N_L(s), N_R(s)]$ where the most probable assignment for

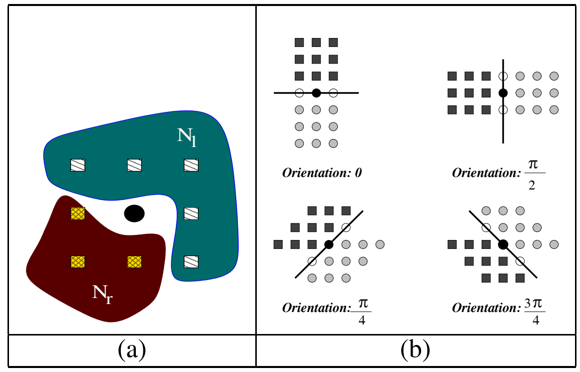

Figure 1. (a) Neighborhood partition that indicates a boundary point, (b) Possible partitions.

$N_L(s)$ is the *static* hypothesis and for $N_R(s)$ is the *mobile* hypothesis or the opposite [**fig.** (1.a)]) . Besides, if s is not a boundary point, then for every possible neighborhood partition the most probable assignment for $N_L(s)$ as well as for $N_R(s)$ is either the *static* or the *mobile* hypothesis. Thus, the conditional boundary/non-boundary probabilities can be estimated using the following formulas [14]:

$$\begin{gathered} p(\mathbf{D}(N(s))|B) = \\ p_S(\mathbf{D}(N_R(s)))p_M(\mathbf{D}(N_L(s))) + p_M(\mathbf{D}(N_R(s)))p_S(\mathbf{D}(N_L(s))) \end{gathered}$$

$$\begin{gathered} p(\mathbf{D}(N(s))|\bar{B}) = \\ p_S(\mathbf{D}(N_R(s)))p_S(\mathbf{D}(N_L(s))) + p_M(\mathbf{D}(N_R(s)))p_M(\mathbf{D}(N_L(s))) \end{gathered}$$

Since the probability that point s lies on the boundary of a moving object is defined, the next problem is to define the neighborhood partition. We consider four different partitions of the neighborhood (the vertical, the horizontal and the two diagonals) [**fig.** (1.b)]. These partitions can be obtained by assuming four different orientations $\left[\theta = \left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}\right]$. We estimate the boundary probability and we generate the vector of the conditional probabilities:

$$P_{\mathsf{EDGE}}(s,\theta) = \begin{bmatrix} p\left(B|\mathbf{D}(N(s)), 0\right) \\ p\left(B|\mathbf{D}(N(s)), \frac{\pi}{4}\right) \\ p\left(B|\mathbf{D}(N(s)), \frac{\pi}{2}\right) \\ p\left(B|\mathbf{D}(N(s)), \frac{3\pi}{4}\right) \end{bmatrix}$$

where the lines correspond to the neighborhood partitions (with respect to θ). The boundary probability $p_{\mathbf{B}}(.)$ is provided by the highest element of the vector $P_{EDGE}(s,\theta)$ $[p_{\mathbf{B}}(s) = \mathbf{max}\{P_{\mathsf{EDGE}}(s,\theta)\}]$ and the boundary-based features are then captured using a Gaussian function: $\left[g(s,\sigma_B) = \frac{1}{\sigma_B\sqrt{2\pi}}e^{-\frac{p^2_{\mathbf{B}}(s)}{2\sigma^2_B}}\right]$.

Thus the geodesic curve that is attracted by the moving objects boundaries is given by,

$$E(\partial\mathcal{P}(R)) = \sum_{i=1}^{N}\int_0^1 g(\partial R_i(p_i), \sigma_B)\,|\partial\dot{\mathrm{R}}_i(p_i)|dp_i$$

3.2 Setting the Region Module

The main goal of region-based tracking methods is to create a partition of the image domain over the time that is consistent with the the intensity and motion properties of the objects as well as the background. In our case these properties are well determined using the statistical modeling described in Section 2.

This partition can be viewed as a maximization problem with respect to the *a posteriori* segmentation probability. Let $p(\mathcal{P}(R)|\mathbf{I})$ be the *a posteriori* segmentation density function for the different partitions $\mathcal{P}(R)$ given the input data $\mathbf{I}$. This density function is given by the Bayes rule as:

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

If we assume that all the partitions are *a priori* equally possible then the constant terms $p(\mathbf{I}),\ p(\mathcal{P}(R))$ can be ignored and the density function is rewritten as:

$$p(\mathcal{P}(R)|\mathbf{I}) = p\left(\cap_{i=0}^{N}\left[\mathbf{I}(R_i)|i\right]\right) = \prod_{i=0}^{N} p_i\left(\mathbf{I}(R_i)\right)$$

where $p_i(.)$ is the associated probability density function with respect to the properties of object assigned to the region R_i and $\mathbf{I}(R_i)$ is the data associated with R_i. Besides, if we assume that the pixels within each region are independent, we can replace the region probability with: $\left[p_i\left(\mathbf{I}(\mathbf{R}_i)\right) = \prod_{s\in\mathbf{R}_i} p_i\left(\mathbf{I}(s)\right)\right]$. The maximization of *a posteriori* probability is equivalent with the minimization of the negative [**log**()] function of this probability,

$$\begin{aligned} E(\partial\mathcal{P}(R)) &= -\,\mathbf{log}\left[\prod_{i=0}^{N}\prod_{s\in R_i} p_i\left(\mathbf{I}(s)\right)\right] \\ &= -\sum_{i=0}^{N}\iint_{\mathbf{R}_i} \mathbf{log}\left[p_i(\mathbf{I}(x,y))\right] dxdy \end{aligned}$$

A moving object is well tracked if the corresponding curve at different time instants refers to a region with constant intensity properties, that correspond to the object intensity properties. Since in our case the intensity properties are expressed under a multi-modal framework (*intensity, derivatives, magnitude of the gradient*) we can define a multi-module energy component with respect to this module, that assumes independence between the different filter responses. In that case the energy expression for the intensity-based region module is given by:

$$E(\partial\mathcal{P}(R)) = -\sum_{o\in F} w_o \sum_{i=0}^{N}\iint_{\mathbf{R}_i} \mathbf{log}\left[p_{io}(\mathbf{I}(x,y))\right] dxdy$$

where w_i are positive weights that take into account the operator uncertainties.

Additionally, the regions that are associated with moving objects should provide high *a posteriori* segmentation probability with respect to the mobile case. On the other hand, the background region should provide high *a posteriori* segmentation probability with respect to the static case. Taking this into account, the energy expression for the motion-based region module is defined as:

$$E(\partial\mathcal{P}(R)) = -\sum_{i=0}^{N}\iint_{R_i} \mathbf{log}\left[p_i(D(x,y))\right] dxdy$$

where $p_i(.)$ is given by:

$$p_i(D(x,y)) = \begin{cases} p_S(D(x,y)) & \text{, if } i = 0 \\ p_M(D(x,y)) & \text{, if } i \neq 0 \end{cases}$$

3.3 Geodesic Active Region Tracking

Following [1, 14, 15, 18] we fuse the two different modules, by defining a Geodesic Active Region Tracking model:

$$\begin{aligned} E(\partial\mathcal{P}(R)) = {} & \alpha\sum_{i=1}^{N}\int_0^1 g(\partial R_i(p_i),\sigma_D)\,|\partial\dot{\mathrm{R}}_i(p_i)|\,dp_i \\ & -\beta\sum_{o\in F} w_o \sum_{i=0}^{N}\iint_{\mathbf{R}_i} \mathbf{log}\left[p_{io}(\mathbf{I}(x,y))\right] dxdy \\ & -\gamma\sum_{i=0}^{N}\iint_{\mathbf{R}_i} \mathbf{log}\left[p_i(\mathbf{D}(x,y))\right] dxdy \end{aligned}$$

where $\{\alpha,\beta,\gamma\}$, are positive normalized constants that balance the contribution of the different tracking modules. Let us now try to interpret the above functional. The tracking is obtained by minimizing three kinds of "energy terms" for each moving object $[R_i]$. The first, is *boundary-based* and gives the best geodesic minimal-length smoothed curve that attracts the object boundaries, the second is a *region-based* and maximizes the *a posteriori* segmentation probability with respect to the intensity properties of object associated with the region defined by the curve, while the third maximizes the *a posteriori* segmentation probability with respect to the mobile h hypothesis inside the region defined by the curve. Also, we would like to maximize the *a posteriori* segmentation probability with respect to the background intensity properties and to the static case hypothesis for the region R_0 that is associated with the static background.

4 Minimizing the Energy

The minimization of the objective function is obtained using a gradient descent method. Let $\vec{u} = (x,y)$ be a pixel of the initial curve that is assigned to the object i. If we compute the Euler-Lagrange equations then we should deform each pixel $\vec{u}$ of the initial curve towards the minima of the objective function using the following equation:

$$\begin{aligned} \frac{du}{dt} = \Bigg(& \underbrace{\alpha\left[g(u,\sigma_D)\mathcal{K}_i(u) - \vec{\nabla}g(u,\sigma_D)\cdot\vec{\mathcal{N}}_i(u)\right]}_{boundary\ force} \\ & \underbrace{-\beta\sum_{o\in F} w_o\,\mathbf{log}\left[\frac{p_{io}(\mathbf{I}(u))}{p_{0o}(\mathbf{I}(u))}\right]}_{intensity-based\ region\ force} - \underbrace{\gamma\,\mathbf{log}\left[\frac{p_M(\mathbf{D}(u))}{p_D(\mathbf{D}(u))}\right]}_{motion-based\ region\ force}\Bigg)\vec{\mathcal{N}}_i(u) \end{aligned}$$

where $\mathcal{K}_i$ is the Euclidean curvature of ∂R_i and $\vec{\mathcal{N}}_i$ is the unit inward normal to ∂R_i. The obtained PDE motion equation has three kind of "forces" acting on the contour, all in the direction of the normal.

• The first is boundary-based that 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).

• The second is an intensity-based force. Let as interpret this term; suppose that u is a background pixel. In such a case the observed intensity should support the static hypothesis; $[p_{0o}(\mathbf{I}(\vec{u})) > p_{io}(\mathbf{I}(\vec{u})), \forall o \in F]$ and the curve is going to be compressed. On the other hand the curve will be expanded if it is located inside a moving object.

• Finally, the third term is motion-based and aims at shrinking the curve when it is located at the background $[p_S(D(\vec{u})) > p_M(D(\vec{u}))]$ and expanding it is located inside the moving object $[p_S(D(\vec{u})) < p_M(D(\vec{u}))]$.

4.1 Level Set Formulation

The obtained PDE can be implemented using a Lagrangian approach, that is limited since it cannot deal with topological changes of the moving front and suffers from instability in the domain of numerical approximations.

This can be avoided by introducing the work of Osher and Sethian [13] in our scheme. The central idea is to represent the moving front $\partial R(t)$ as the zero-level set $\{\Phi = 0\}$ of a function Φ. This representation of $\partial R(t)$ is implicit, parameter-free and intrinsic. Additionally, it is topology-free.It is easy to show, that if the embedding function Φ deforms according to

$$\frac{d}{dt}\Phi(p,t) = F(p)\,|\nabla\Phi(p,t)|$$

then the corresponding moving front evolves according to:

$$\frac{d}{dt}C(p,t) = F(p)\vec{\mathcal{N}}$$

Thus, the minimization of the proposed geodesic active region objective function is equivalent to searching for a steady-state solution of the following equation:

$$\frac{d}{dt}\Phi(u) = \left(\alpha \left[g(u,\sigma_D)\mathcal{K}_i(u) + \vec{\nabla}g(u,\sigma_D)\cdot\frac{\vec{\nabla}\Phi(u)}{|\nabla\Phi(u)|}\right]\right.$$
$$\left. -\beta\sum_{o\in F} w_o \,\mathbf{log}\left[\frac{p_{io}(\mathbf{I}(u))}{p_{0o}(\mathbf{I}(u))}\right] - \gamma\,\mathbf{log}\left[\frac{p_M(\mathbf{D}(u))}{p_D(\mathbf{D}(u))}\right]\right)|\nabla\Phi(u)|$$

where the $\partial R(p,t)$ is represented by a level-set of Φ and the value of $\mathcal{K}$ is estimated on Φ ($\mathcal{K} = div(\nabla\Phi/|\nabla\Phi|)$).

The Level Set Equation is implemented using the **Hermes** algorithm [15] that proposes a fast way to deform the initial curve towards the minimum of the objective function.

In our case the equation which deforms the initial curve can be rewritten in a more general form as:

$$\Phi^{t+1}_{(x,y)} = \Phi^{t}_{(x,y)} + \mathcal{V}(x,y,\Phi)dt$$

where $\mathcal{V}(x,y,\Phi)$ is the propagation speed function, depending on geometric features and image features. Since the speed $\mathcal{V}(x,y,\Phi)$ is basically estimated according to image characteristics, there are some points for which the front evolves faster compared to the others. The key idea on which the Hermes approach is based is to evolve the front locally according to the speed values of its points. The algorithm at each step selects the point with the highest absolute propagation speed from a set of actual curve points, and deforms the level-set image locally.

4.2 Implementation Issues

The proposed algorithm is self-sufficient and works as follows: Given the *first*, the *reference* frame and an initial curve, the Geodesic Active Region model is activated, and detects the moving objects using the boundary and the motion detection module $\{\alpha = \gamma = 0.5, \beta = 0\}$. Then, each object is associated with an intensity-based descriptor. Besides, the contour initialization from frame to frame is performed based on the displacement of the gravity center of the object, which is estimated using a block matching gradient descent method and we proceed to the next frame.

Finally, we have to define the weights $\{\alpha,\beta,\gamma\}$ of the different tracking modules. Experimentally, it has been found that the intensity-based module fails constantly, since it is based on global statistics and the quality of our sequences is pure. On the other hand the motion detection module is quite reliable. The boundary module has an unaccountable behavior since it presents negative values only due to the curvature effect, thus the curve is propagated towards one direction under a regularity constraint. Thus, if the initial curve has a part inside the object and a part outside the object, then this term has a beneficial contribution for the exterior part, while it discourages the interior part to evolves towards the correct direction (outwards). On the other hand, this term is very important since it ensures the regularity of the curve. We take these remarks into consideration and we determine the modules contributions as follows $\{\alpha = 0.40, \beta = 0.10, \gamma = 0.50\}$.

5 Conclusions

Very promising experimental results have been obtained using the proposed framework for real indoor and outdoor video-surveillance sequences [**fig.** (2)][1]. The main drawback of the model is the cases of occlusion where the output of our model is a single curve for both objects. This can be confronted by incorporating other tracking modules like optical flow-based features [16] as well geometric-based features (object representations).

Summarizing, we presented new ideas concerning the integration of boundary-based and region-based approaches for tracking. The main contribution of this work consists

[1] The computational cost of our approach is strongly related with the position of the initial curves and the number of objects. Approximately, we can say that for a 256x320 sequence (Corridor), we need about 3 seconds per frame using an ULTRA 10, 299 MHz.

of *creating a tracking model that integrates different type (boundary and region) and sources (intensity and motion detection) of information*. This leads to a system where the boundary and the region module operate simultaneously, while the contour propagation is guided by *regularity*, *boundary*, and *region-based* forces. The changes of topology can be easily achieved using a Level-Set approach. Thus several moving objects can be tracked simultaneously.

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

http://www.inria.fr/robotvis/personnel/nparagio/demos/

References

[1] B. Bascle and R. Deriche. Region Tracking through Image Sequences. In *ICCV*, pages 302–307, Boston, USA, 1995.

[2] T. Birchfield. Elliptical Head Tracking Using Intensity Gradients and Color Histograms. In *CVPR*, Santa Barbara, USA, 1998.

[3] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22:61–79, 1997.

[4] R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, Inc., 1973.

[5] D. Gravila and L. Davis. 3-D Model-based Tracking of Humans in Action: a Multi-view Approach. In *CVPR*, San Fransisco, USA, 1996.

[6] M. Isard and A. Blake. Contour Tracking by Stochastic Propagation of Conditional Density. In *ECCV*, volume I, pages 343–356, Cambridge, UK, 1996.

[7] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. In *ICCV*, pages 810–815, Boston, USA, 1995.

[8] F. Leymarie and M. Levine. Tracking deformable objects in the plane using an active contour model. *PAMI*, 33:617–634, 1993.

[9] D. Lowe. Robust model based motion tracking through the integration of search and estimation. *International Journal of Computer Vision*, 8:113–122, 1992.

[10] R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: A level set approach. *PAMI*, 17:158–175, 1995.

[11] D. Metaxas and D. Terzopoulos. Constrained deformable superquadrics and nonrigid motion tracking. In *CVPR*, pages 337–343, 1991.

[12] F. Meyer and P. Bouthemy. Region-based tracking using affine motion models in long image sequences. *CVGIP: Image Understanding*, 60:119–140, 1994.

[13] 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.

[14] N. Paragios and R. Deriche. Geodesic Active Regions for Texture Segmentation. Research Report 3440, INRIA, France, 1998. http://www.inria.fr/rapports/sophia/RR-3440.html.

[15] N. Paragios and R. Deriche. A PDE-based Level Set approach for Detection and Tracking of moving objects. In *ICCV*, pages 1139–1145, Bombay, India, 1998.

[16] N. Paragios and R. Deriche. Geodesic Active Regions for Motion Estimation and Tracking. Research Report 3631, INRIA, France, 1999. http://www.inria.fr/rapports/sophia/RR-3631.html.

[17] N. Paragios and G. Tziritas. Adaptive Detection and Localization of Moving Objects in Image Sequences. *Signal Processing: Image Communication*, 14:277–296, 1999.

[18] S. Zhu and A. Yuille. Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. *PAMI*, 18:884–900, 1996.

(a)

(b)

(c)

Figure 2. Unifying Boundary and Region-based information for Geodesic Active Tracking. (a) Jumber Sequence, (b) Oxford Sequence, (c) Corridor Sequence.