

A Robust Video Object Tracking by Using Active Contours

Mohand Saïd Allili and Djemel Ziou

Département d'informatique,

Faculté des Sciences,

Université de Sherbrooke,

Sherbrooke (QC), J1K 2R1, Canada.

Emails:{MS.Allili, D.Ziou}@Usherbrooke.Ca.

Abstract

In this paper, we propose a novel object tracking algorithm in video sequences. The formulation of our tracking model is based on variational calculus, where region and boundary information cooperate for object boundary localization by using active contours. In the approach, only the segmentation of the objects in the first frame is required for initialization. The evolution of the object contours on a current frame aims to find the boundary of the objects by minimizing the Kullback-Leibler distance of the region features distribution in the vicinity of the contour to the objects versus the background respectively. We show the effectiveness of the approach on examples of object tracking performed on real video sequences.

Keywords: Object tracking, Kullback-Leibler distance, texture, color, mixture of pdfs, boundary, level sets.

1. Introduction

Object tracking is one of the most important problems in computer vision and it has a variety of applications such as coding, video surveillance, monitoring, and robotics. It has been the focus of numerous research works in the last decades where two main approaches have been investigated. In the first approach, the tracking is resolved by first detecting the objects in every frame. Then, a correspondence between the detected objects is calculated by estimating the motion parameters of the moving objects [3, 9]. In methods relying on this approach, motion models like similarity, affine, projective transformations for the moving objects can be used [9, 15]. Although this approach behaves very well when the objects undergo small deformations, its efficiency may drastically decrease when there are significant structural changes for the objects [12]. In applications where precision is a strict requirement, such as medical imaging, this approach is simply not usable.

In the second approach, the tracking of the moving objects in a current frame is performed by using the previous

frame tracking result as a starting point [4, 14, 16]. Active contour methods are based on this approach [8, 11]. With active contours, non-rigid body motion can be easily tracked. Moreover, having these methods based on the homogeneity characteristics of the objects, there is no need to calculate motion parameters. In [11], the authors used background subtraction to localize the moving parts of the images by assuming a static background. Then, level set curves [10] are evolved to capture these parts and boundary information is used to align the curves with high intensity gradients of the image. However, texture and noise constitute a serious limitation. The approach proposed in [8] was an attempt to track non-rigid objects without computing any motion parameters and without assuming a static background. The author used a probabilistic model based on calculating pixel translations around a circular window to model possible displacements of the pixels inside the moving objects. However, having the tracking based on pixel intensity correspondence, the model is sensible to texture, illumination changes and self-shadowing. In [17], a probabilistic tracking model was proposed for tracking where texture and color features were combined to describe the moving objects. The features distribution is modelled by using non-parametric distributions. However, no boundary information was used to describe the moving objects. This constitutes a big shortage since, usually, boundary information is desirable for accurate boundary localization [1, 5, 8]. To our knowledge, no work in the past combined texture and color features for tracking while using boundary information.

In the present paper, we propose an approach for object tracking by combining region and active contours for tracking. Region information comprises color and texture features while boundary information is formulated by using a multi-band edge detector [1]. A probabilistic model is proposed for the tracking where, having the segmentation of a moving object in the first frame, the tracking aims to correct the

position of the contour in the subsequent frames by minimizing the Kullback-Leibler (KL) distance of both sides of the contour vicinity distributions with the object and background distributions respectively. Moreover, using distribution comparison to assigning pixel neighborhoods to the moving objects versus the background augments the robustness of the tracking to noise, self-shadowing and texture. As stated before, no motion field is calculated a priori for the moving objects. However, in our algorithm the region features of both object and background are modelled by using mixture of pdfs, which gives the probabilistic aspect of our model. The method is implemented by using level set active contours that allow for natural topology changes of the contours [10] and stable numerical schemes.

This paper is organized as follows: In section (2), we present the main ideas of the proposed tracking approach. In section (3), we discuss some implementation issues of the model and the proposed solutions. In section (4), we show a validation of the approach on examples of single and multiple object tracking. We end the paper with some conclusions and future work perspectives.

2 Formulation of the tracking model

In what follows, we denote by I^n the n^{th} frame in the image sequence. The aim of the proposed tracking model is to localize one or several moving objects on the background in all the sequence frames $I^n (0 \leq n < \infty)$, having only the segmentation of the objects in the first frame I^0 (reference frame). Without loss of generality, we assume the scene is composed of one moving object (extension to multiple objects is described later in the paper). Let us denote by R_{obj} and R_{bck} the region of the object and the background, and by ∂R_{obj} and ∂R_{bck} their boundaries respectively. We assume that both the region of the object and the background are represented by using a mixture of pdfs which describe their distribution of color and texture features. We assume also that the region features distributions of the object and the background do not vary much in time. Otherwise, the algorithm should require a re-initialization of their respective statistics. Having that, let the probability of observing a value of the feature vector $U(\mathbf{x}) = (u_1(\mathbf{x}), \dots, u_d(\mathbf{x}))$ in the reference frame, where (\mathbf{x}) denotes the coordinates (x, y) , be given as following:

$$f_{in}(U(\mathbf{x})) = \sum_{i=0}^{M_{obj}-1} \pi_i \cdot p(U(\mathbf{x})/\theta_i). \quad (1)$$

$$f_{out}(U(\mathbf{x})) = \sum_{j=0}^{M_{bck}-1} \xi_j \cdot p(U(\mathbf{x})/\theta_j) \quad (2)$$

Here, f_{in} and f_{out} designate respectively the mixture probabilities of the object and background regions. M_{obj} and

M_{bck} designate the number of classes contained in the object and the background, and π_i and ξ_j designate the mixing parameters for the two mixture models. The number of classes in the object and background (M_{obj} and M_{bck}) is calculated by using the MDL information-theory criterion [13]. We choose the formalism of General Gaussian distributions (GGDs) to model the probability density functions of the mixture components. In [1, 2], we have shown that using the formalism GGD brings a better fit to data than the Gaussian while not over-fitting the number of components in the mixture. Having a mixture of GGDs, the unknowns parameters of the distributions are calculated by using the Maximum Likelihood Estimation.

2.1 Region and boundary features

To achieve a robust and accurate tracking, we use a combination of color and texture features to describe the regions. For color, we use CIE- $L^*a^*b^*$ color space which is perceptually uniform. For texture, we use features calculated from the correlogram of a pixel neighborhood [7]. An element of the correlogram matrix $C^{d,\theta}(c_i; c_j)$ should give the probability that, given a pixel \mathbf{x}_1 of color c_i , a pixel \mathbf{x}_2 at distance v and orientation θ from \mathbf{x}_1 is of color c_j . We calculate the correlogram for 4 orientations $(v, 0)$, $(v, \frac{\pi}{4})$, $(v, \frac{\pi}{2})$ and $(v, \frac{3\pi}{4})$ and we derive from each correlogram three typical characteristics that are namely: *Inverse-Difference-Moment*(IDM), *Energy*(E) and *Correlation*(C). E and C measure respectively the homogeneity of the texture. IDM, however, measures the coarseness of the texture. By using two displacements v and combining the texture and color features, the vector of region features $U(\mathbf{x})$ will contain 9 dimensions.

For boundary information, we use a multi-band edge detector that is proposed in [6]. In this approach, the boundary plausibility is approximated by the strength of the strongest first directional derivative of the vector-valued function $I(\mathbf{x}) = U(\mathbf{x})$. In the following, we denote the boundary plausibility for the pixel \mathbf{x} by: $P(\mathbf{x})$.

2.2 Variational model for tracking

Suppose we have the tracking result of the object in the frame I^n of the sequence. Assume that the object, that might undergo a non-rigid motion, takes the location R_{obj} in the next frame I^{n+1} . To track the object in the frame I^{n+1} , we initialize the object contour ∂R_{obj} to its position in the frame I^n . Let us model for each pixel $\mathbf{x} = (x, y)$ belonging to the object contour its neighborhood data by using a mixture of pdfs for the region features: the part inside the contour by using a mixture of pdfs $\tilde{f}_{in}(\mathbf{x})$ and the part outside the contour by using a mixture of pdfs $\tilde{f}_{out}(\mathbf{x})$ (see fig. (1)). The parameters of both mixtures are calculated

by using the Maximum Likelihood estimation. To track the

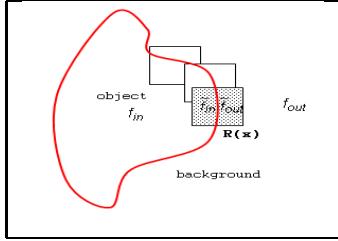


Figure 1: Representation of the proposed approach for tracking by using the information in the vicinity of the contour.

object R_{obj} on the frame I^{n+1} , we minimize the following energy functional over the boundary of the object ∂R_{obj} :

$$\begin{aligned} & E(\partial R_{obj}) \\ &= \alpha \underbrace{\oint_{\partial R_{obj}} (g(P(s))) ds}_{E_b} \\ &+ \beta \underbrace{\left(\iint_R D(\tilde{f}_{in}(\mathbf{x}) \| f_{in}(\mathbf{x})) \cdot \chi(R_{obj}(\mathbf{x})) d\mathbf{x} \right)}_{E_r(obj)} \\ &+ \underbrace{\left(\iint_R D(\tilde{f}_{out}(\mathbf{x}) \| f_{out}(\mathbf{x})) \cdot \chi(R_{bck}(\mathbf{x})) d\mathbf{x} \right) ds}_{E_r(bck)} \quad (3) \end{aligned}$$

where $\chi(R(\mathbf{x}))$ is the characteristic function that gives 1 if a pixel lies inside the region which is given as its parameter, and 0 otherwise. In all the integrals, s represents the arc-length parameter. In the functional (3), g is a strictly decreasing function of the boundary plausibility: P . This is formulated by using the following function: $g(P(\mathbf{x})) = \frac{1}{|P(\mathbf{x})| + \varepsilon}$, where ε is a constant parameter. The parameters α and β weight the contribution of the region and boundary information in the energy functional. Note that inside the integrals, we used symmetric KL-distance between the distributions. Having two distributions f and g , this distance is given by:

$$\begin{aligned} D(f \| g) &= \frac{1}{2} \int_x f(x) \log \left(\frac{f(x)}{g(x)} \right) \\ &+ \frac{1}{2} \int_x g(x) \log \left(\frac{g(x)}{f(x)} \right) \quad (4) \end{aligned}$$

where the above integrals are calculated over the features space domain. Here, for simplicity we assume the different bands of the color and texture features are independent. This hypothesis seems somewhat suboptimal but, in most of our tests, the model showed good results which demonstrates that the performance of the algorithm is not much affected for a range of real-world scenes.

Our goal by proposing the functional (3) is to partition a given frame such that the local image statistics within the object (resp. background) are "close" to the global statistics within the same object (resp. background). This is expressed by the symmetric KL-distance which is an information-theory distance measure between two distributions. In order to interpret the criterion (3), let us assume for a moment that for a pixel \mathbf{x} , the neighborhood part inside the contour R_{in} has the distribution \tilde{f}_{in} and the part outside the contour R_{out} has the distribution \tilde{f}_{out} . Minimizing (3) then amounts to minimizing the description length of the object code: A minimal code for R_{in} (resp. R_{out}) has average length $H(\tilde{f}_{in})$ (resp. $H(\tilde{f}_{out})$), H denoting Shannon entropy. Encoding R_{in} (resp. R_{out}) using the model of the object or the background requires a code of average length $H(\tilde{f}_{in}) + D(\tilde{f}_{in} \| f_{in/out})$ (resp. $H(\tilde{f}_{out}) + D(\tilde{f}_{out} \| f_{in/out})$). In order to minimize (3), we should assign R_{in} to the object and R_{out} to the background.

2.3 Curve evolution

The functional (3) is minimized by using Euler-Lagrange equations according to the object contour ∂R_{obj} . We use the level set formalism [10] to implement the curve evolution and we obtain the following motion equation for the object contour:

$$\frac{d\Phi}{dt} = (\alpha V_b - \beta V_r) |\nabla \Phi| \quad (5)$$

with

$$\begin{cases} V_b &= g(P)\kappa + \nabla g(P) \cdot \frac{\nabla \Phi}{|\nabla \Phi|} \\ V_r &= (E_r(obj) - E_r(bck))\kappa \\ &+ \nabla(E_r(obj) - E_r(bck)) \cdot \frac{\nabla \Phi}{|\nabla \Phi|} \end{cases} \quad (6)$$

where the value of $E_r(obj)$ and $E_r(bck)$ are those given in the functional (3). In equation (5), $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a level set function, where ∂R_{obj} is represented by its zero level set. The symbol κ in the equation (2.3) stands for the curvature of the zero level set. The term V_b represents the boundary velocity that aligns the curve with pixels having high boundary plausibility. The term V_r represents the region velocity. In the interior of a region, the boundary term is supposed to vanish, and the contour is driven only by the region velocity in the direction that minimizes the KL-distance between the distributions of each pixel neighborhood side with the interior and exterior of the object contour respectively.

2.4 Extension to multiple object tracking

Here, we assume the image sequence contains N different moving objects: $R_{obj}^1, \dots, R_{obj}^N$ lying on the background R_{bck} . The number of objects is given initially by the user. In the proposed model, we track each object by competing its

region with the background. The region of the background is defined as the intersection of the complementary of all the region objects to the image. The N-objects tracking model is formulated by using the following energy functional:

$$\begin{aligned}
& E(\partial R_{obj}^1, \dots, \partial R_{obj}^N) \\
= & \alpha \sum_{k=1}^N \underbrace{\int_{\partial R_{obj}^k} (g(P(s))) ds}_{E_b^k} \\
+ & \beta \sum_{k=1}^N \underbrace{\int_{\partial R_{obj}^k} \left(\iint_R D(\tilde{f}_{in}^k(\mathbf{x}) \| f_{in}^k(\mathbf{x})) \cdot \chi(R_{obj}^k(\mathbf{x})) d\mathbf{x} \right)}_{E_r^{k(obj)}} \\
+ & \underbrace{\iint_R D(\tilde{f}_{out}^k(\mathbf{x}) \| f_{out}^k(\mathbf{x})) \cdot \chi(R_{bck}(\mathbf{x})) d\mathbf{x}}_{E_r(bck)} ds \quad (7)
\end{aligned}$$

where $\chi(R_{bck}) = \prod_{k=1}^N (1 - \chi(R_{obj}^k))$, which formulates the intersection of the complementary of each object region to the image domain. In a word, the above model is equivalent to tracking an object on the domain of the image having the other objects excluded from the background. Otherwise, two objects may have similar statistical properties for the region features, and the tracking could be badly affected if one of one object is considered as part of the background of the other. This allows for tracking independently objects having similar region properties. The minimization of the functional (7) leads to a vector of Euler-Lagrange equations that give the curve evolution for each object contour. By using the level set formalism, these equations are given by:

$$\text{For } (k \in \{1, \dots, N\}) : \frac{d\Phi_k}{dt} = (\alpha V_b^k - \beta V_r^k) |\nabla \Phi_k| \quad (8)$$

where:

$$\begin{cases} V_b^k &= g(P)\kappa + \nabla g(P) \cdot \frac{\nabla \Phi_k}{|\nabla \Phi_k|} \\ V_r^k &= (E_r^{k(obj)} - E_r(bck))\kappa \\ &\quad + \nabla(E_r^{k(obj)} - E_r(bck)) \cdot \frac{\nabla \Phi_k}{|\nabla \Phi_k|} \end{cases}$$

3 Implementation issues

The proposed model aims to track an recover the position of N different objects that undergo non-rigid motion from a frame of the sequence to the next. Some issues could be faced, however, during the implementation of the method which can be summarized as follows. Firstly, the calculation of the kullback-Leibler distances requires calculating integrals over the domain of each components of the region features vector. This calculation is performed for each pixel in the vicinity of an object contour and makes the computation of these distances a real issue. To alleviate the computation expense, we calculate the distance integrals for all

the pixels in the same time. We perform this by previously changing the dynamic of the features into one range in order to have the same boundaries for the integrals. Then, the calculation of integrals is performed in the same time as the re-initialization of the level set distance functions. This permits to visit once the neighborhood of a pixel and perform the computations related to the level sets and KL-distance respectively. Practically, it takes few seconds to perform these computations when the tracking contains two or three moving objects.

The second problem is related to object occlusions and split. For occlusions, if two objects are characterized by different region features, no problem would arise as each contour would continue to track the non-occluded part of its object. However, when the two objects have similarity of color and texture content, pixels of one object may be attributed to another object. In our implementation, we deal with this issue by first detecting if an occlusion has occurred. Thanks to the level set formalism, we know exactly the area that forms the interior of each object at a given time. We detect if there is a contact or and overlapping between an object R_{obj}^k and R_{obj}^h ($k \neq h$) if some pixels have negative distance value for the functions Φ_k and Φ_h respectively. In this case, knowing the distributions of the two objects f_{in}^k and f_{in}^h , we decide to them into one object if the following conditions are satisfied:

$$\begin{cases} \exists(\mathbf{x}), \Phi_k(\mathbf{x}) \leq 0 \text{ and } \Phi_h(\mathbf{x}) \leq 0 \\ D(f_{in}^k \| f_{in}^h) \leq \delta \end{cases} \quad (9)$$

The first condition indicates that the zero level set functions of the two objects are in contact or may be overlapping, in which case possible occlusion would occur. The second condition meets the requirement that the two objects should be fused if their respective distributions are similar; here, the threshold δ is determined experimentally. In the same time, we have a testing connected component procedure that detect if an object has split between two successive frames of the sequence. If a new object is created, it is added to the list of objects and a new level set function is initialized to track it separately by using equation (8). We note that after each object fusion and splitting operation, we recalculate the mixture of pdfs of the new objects and perform the appropriate modification for the tracking by removing (resp. adding) an object from (resp. to) the list of moving objects. Finally, a new moving object may enter the scene, in which case it is added to the list of objects by the user. The above operations are performed for each frame of the sequence.

4 Experiments and discussion

In the experiments that we have conducted, we firstly illustrate the advantage of the choices that we adopted to our

model. First, to show the benefit of combining boundary and region information for object tracking, we present on fig. (2) an example where the tracking algorithm is run with and without boundary model, i.e: we put $\beta = 0$. We used as initialization the previous frame tracking result (first picture). Notice the leak of the contour where boundary information is absent (middle picture). The color of this part of the object is similar to the background and the boundary between them is not really apparent although it exists. Having the boundary information in our model permitted to correct the situation by aligning the contour with the real object boundaries (third picture).

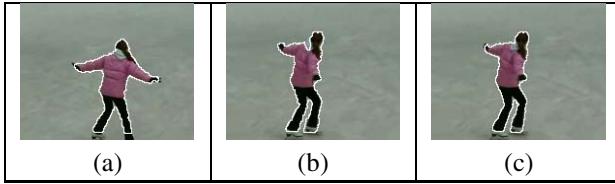


Figure 2: Fig. (a) represents the contour initialization. Fig. (b) and (c) represent respectively the result of tracking without and with boundary information for the tracking.

In a second test, we show the advantage of using KL-distance for tracking over using mixture probabilities classification. In fig. (3), we show on the first and second row an example of segmentation on a synthetic image degraded by noise. On the top right, the segmentation is performed by classifying each pixel by using the maximum a posteriori probability. On the second row, we show a classification based on KL-distance applied with several neighborhood sizes m . We notice the increasing robustness of segmentation by using KL-distance, which justify its use to reduce the sensibility of tracking to noise. In the last row of the figure, we show the result of tracking of two consecutive frames of an image sequence. The second frame has been added a Gaussian noise ($\mu = 0, \sigma = 3$). We show the result of tracking by using respectively: (middle picture) a classification of the pixels to the tracked object versus the background based on mixture probabilities, (third picture) our tracking model. We note that by using mixture probabilities, the classification of the pixels was severely affected by noise, where the object contour was deviated from the real boundaries of the object. Using this result as an initialization for the next frame would diverge the contour far from the object. By using our model with a neighborhood size ($m = 15 \times 15$), the contour was aligned with the real boundaries of the object, which demonstrates the efficiency of the method.

In fig. (4), we show one single object tracking examples in different sequences by using the proposed method. The object and the background are constituted of multiple

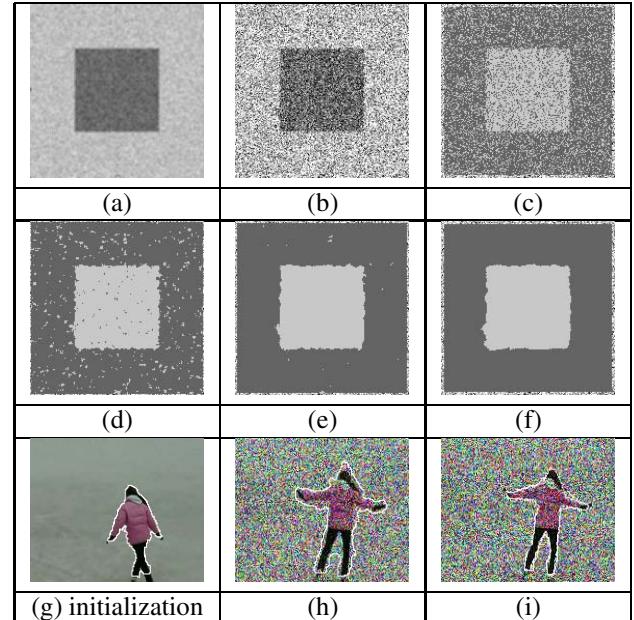


Figure 3: Fig. (a) and (b) show resp. the original image and the same image added a Gaussian noise ($\mu = 0, \sigma = 3$). Fig. (c) shows the result of segmentation by classifying individual pixels by using a mixture of pdfs. Fig. (d), (e) and (f) show the result of classifying the pixels by using the KL-distance with neighborhood sizes: ($m = 5 \times 5$), ($m = 7 \times 7$) and ($m = 11 \times 11$). Finally, fig. (h) and (i) show the result of tracking on a real example by using respectively a mixture of pdfs and the KL-distance.

classes of color and texture. We note that the background and the object both underwent motions, which makes detecting and tracking of moving objects by using background subtraction like in [11] or motion segmentation like in [9] a non-feasible task. For the examples on fig. (4), we run the tracking by setting the neighborhood band around the contour to 30 pixels. We set $\alpha = \beta = 0.5$ and $\varepsilon = 0.5$ for the function g . In example (a), both the moving object (skating child) and the background are composed of 2 classes. In example (b), the moving object (car) is composed of one class and the background is composed of 4 classes. In example (c), the moving object (skating woman) is composed of 3 classes and the background of one class. Finally in example (d), the moving object (roller-skating woman) is composed of 3 classes and the background of 2 classes. In all the sequences, the moving objects have been tracked successfully by the proposed method.

In fig. (5), we show three examples of multiple object tracking in different image sequences. In example (a), two moving objects composed both of 3 classes are tracked. An object contact is detected in the third frame but, having their

statistics different, the objects have not been fused. In example (b) the moving objects have similar distributions of color and texture, the objects have been fused into one object after they contacted each other (second frame). Until the 6th frame, the two objects remained together and tracked as one object. In example (c) we show a tracking of two objects (a woman and a boy) in contact. The boy and woman regions are composed respectively of three and 4 classes. The background is composed of 3 classes. The two objects remained separated because of the difference of their statistics. Remark the boy region has split in the 3rd frame and the formed small region (shoes) has disappeared in the next frames. Finally, in example (d) (from the thermal-IR imagery pedestrian dataset in the conference web-site), two objects composed of one class each are tracked on a background composed of 3 classes. The two objects are fused in the 4th frame and split in the 5th frame.

For computation time, by comparison to methods like [8, 17] on a Pentium-4 2.4 GHZ on the above examples, our approach takes a slightly more time (in the order of 2 or 3 seconds). This difference in time decreases with the size of the object especially comparing to [8] that involves a dense estimation of the pixels displacements.

5. Summary and Conclusions

In this paper, we proposed a novel method for tracking of multiple moving objects in image sequences. The method combines region and boundary for object description and use level set active contour for its implementation. The algorithm has proven efficiency for tracking accurately objects composed of several classes. We showed also the robustness of the method to noise over tracking by using mixture probabilities. For future work, we should investigate, firstly, the fusion the image data of different sensors for an accurate tracking and, secondly, the use of shape descriptors to handle object contour preservation under object occlusions.

Acknowledgments

This work was supported by National Sciences and Engineering Research Council (NSERC).

References

- [1] M.S. Allili and D. Ziou, An Automatic Segmentation Combining Mixture Analysis and Adaptive Region Information: A level Set Approach, *IEEE CRV*, pp. 73-80, 2005.
- [2] M.S. Allili, N. Bouguila and D. Ziou, Generalized Gaussian Mixture and MML, *Technical Report*.
- [3] J. Barron and D. Fleet and S. Beauchemin, Performance of Optical Flow Techniques, *IJCV*, Vol. 12, pp. 43-77, 1994.
- [4] A. Blake and M. Isard, *Active Contours*, Springer-Verlag, 1998.
- [5] A. Chakraborty and J.S. Duncan, Game-Theoretic Integration for Image Segmentation, *IEEE Trans. on PAMI*, Vol. 21 n. 1, pp. 12-30, 1999.
- [6] C. Drewniok, Multispectral Edge Detection: Some Experiments on Data From Landsat-TM, *Int. Jour. on Remote Sensing*, Vol. 15 n. 18, pp. 3743-3765, 1994.
- [7] J. Huang and S. R. Kumar and M. Mitra and W.-J. Zhu and R. Zabih, Image Indexing Using Color Correlograms, *IEEE CVPR*, pp. 762-768, 1997.
- [8] A.R. Mansouri, Region Tracking Via Level Set PDEs Without Motion Computation, *IEEE Trans. on PAMI*, Vol. 24 n. 7, pp. 947-961, 2002.
- [9] A.R. Mansouri and J. Konrad, Multiple Motion Segmentation With Level Sets, *IEEE Trans. on IP*, Vol. 12 n. 2, pp. 201-220, 2003.
- [10] S. Osher and J. Sethian, Fronts Propagating With Curvature-Dependant Speed: Algorithms Based on Hammilton-Jacobi Formulations, *Jour. of Comp. Physics*, Vol. 22, pp. 12-49, 1988.
- [11] N. Paragios and R. Deriche, Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects, *IEEE Trans. on PAMI*, Vol. 22 n. 3, pp. 266-280, 2000.
- [12] N. Paragios and R. Deriche, Geodesic Active Regions and Level Set Methods for Motion Estimation and Tracking, *CVIU*, pp. 259-282, 2004.
- [13] J. Rissanen, Modelling by Shortest Data Discription, *Automatica*, Vol. 14, pp. 65-471, 1978.
- [14] C. Stauffer and W. Grimson, Learning Patterns of Activity Using Real Time Tracking, *IEEE Trans. on PAMI*, Vol. 22 n. 8, pp. 747-767, 2000.
- [15] C. Stiller and J. Konrad, Estimating Motion in Image Sequences: A Tutorial on Modelling and Computation of 2D Motion, *IEEE Signal Processing Magazine*, pp. 70-91, 1999.
- [16] C.R. Wren and A. Azarbayejani and T. Darrell and A.P. Pentland, Pfinder: Real-Time Tracking of the Human Body, *IEEE Trans. on PAMI*, Vol. 19 n. 7, pp. 780-785, 1997.
- [17] A. Yilmaz and X. Li and M. Shah, Contour-Based Object Tracking With Occlusion Handling in Video Acquired Using Mobile Cameras, *IEEE Trans. on PAMI*, Vol. 26 n. 11, pp. 1-6, 2004.

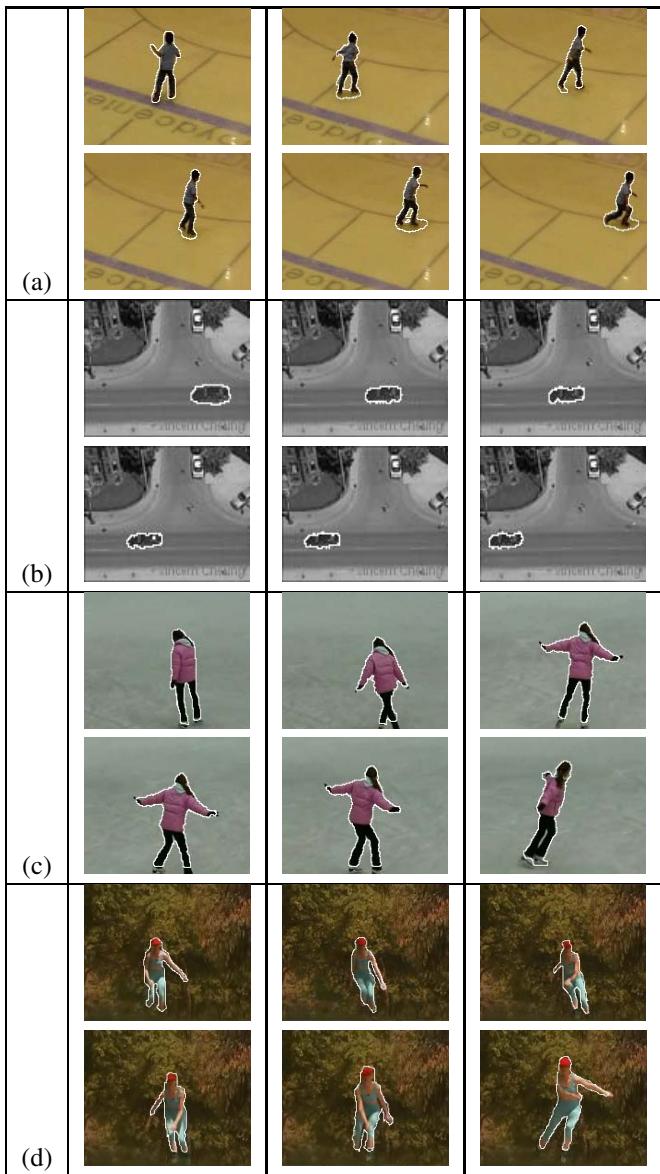


Figure 4: Representation of three tracking sequences by using the proposed method. Six frames are shown for each tracking example.

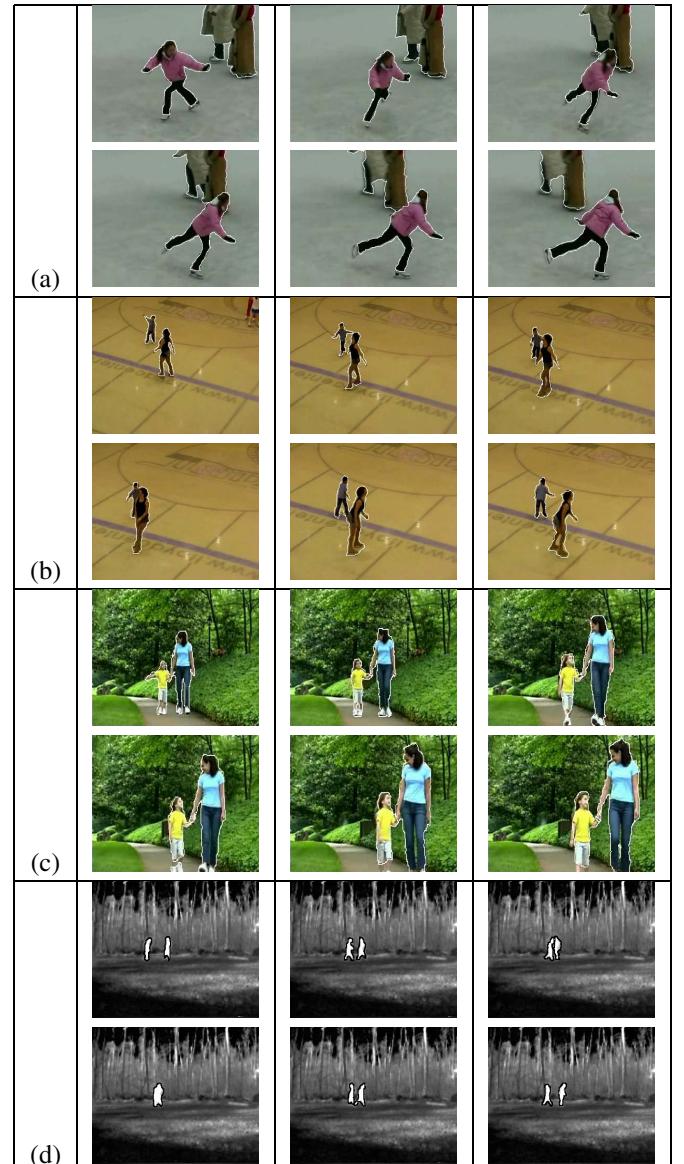


Figure 5: Example of multiple object tracking on a video sequence by using the proposed method. Example (a) represents an example of occlusion without object fusion. Example (b) represents an example of occlusion with object fusion and splitting. Example (c) represents an example of object splitting.