Moving object Segmentation

using optical flow with active contour model

Youssef ZINBI[(1)], Youssef CHAHIR[(1)]
GREYC - URA CNRS 6072 , Université de Caen
Campus II - BP 5186, 14032 Caen France
yzinbi@info.unicaen.fr
ychahir@greyc.ensicaen.fr

Abder ELMOATAZ
GREYC - URA CNRS 6072 Université de Caen
Campus II - BP 5186, 14032
Caen France
elmoataz@greyc.ensicaen.fr

***Abstract*—The paper presents an object segmentation approach that combines optical flow and active contour model to characterize objects and follow them in video sequences. Our aims is to discriminate moving objects from a static background. The approach is based on a minimization of a functional of energy (E) which uses perceptual information in regions of interest (ROI) in an image, in conjunction with a mixture of Gaussian to model voxels of the background image and those of the visual objects. In this work, we compute the optical flow then we use the result of the optical flow as an input in an active contour model. Experiments with a number of test sequences are promising and extend the numerous works on this subject.**

Keywords : Optical Flow, Active Contour, Tracking, Image and Video Segmentation

INTRODUCTION

A fundamental step for video/image retrieval by content is the calculation of the visual features. In contexts, the most relevant content consists in represented subjects rather than the whole scene. The distinction between focused objects and background permits to calculate the visual features of each object alone, making more effective the retrieval task. Segmentation of moving objects from a video sequence is an important task whose applications cover domains such like video compression, video surveillance or object recognition. Image segmentation, widely employed in medical imaging, computer vision, production quality control, etc, is the process to extract meaningful regions from an image. However, the definition of meaningful region is dependent from the applicative context. Generally, an image segmentation process should capture image parts that are perceptually relevant. The definition of what is perceptual relevant characterizes any segmentation method. In this work, the meaningful regions are these related to the focused objects that are in motion. Many image segmentation techniques can be found in literature. They can be roughly classified in region-based and contour-based approaches. More recently [1-2], it was suggested that it should be possible to follow edges in images by suggesting a curve in an image, and then letting the curve itself move to a suitable shape and position. This curve should have physical properties like elasticity and rigidity, and be attracted by edges in the image. For the contour to be attracted to edges in the image an energy image is created, which has high values where the original image has edges and low values otherwise. The attractor image gives the contour a potential energy by summing the energy in the points the contour passes. The contour itself has an internal energy level determined by its shape and by minimizing the total energy one aims at a smooth contour that follows the original image's edges well.

Since we can distinguish moving objects from static elements of a scene by analyzing norm of the optical flow vectors, this one is incorporated in a region-based active contour model in order to attract the evolving contour to moving objects contours. Optical flow aims to measure motion field from the apparent motion of the brightness pattern. Computation of the optical flow can be achieved by many different methods, among a large literature [3-4].

Our objective is to construct a segmentation method able to identify what properties characterize objects and distinguish them from other objects and from the background. The use of optical flow procedures together with other segmentation techniques has been already exploited in other works [5], but to the end to get a more accurate estimation of the motion field.

In this work we have used a modified Active Contour Model, in order to fit contours to shapes in 3D images. This includes finding good ways of representing the active contour as well as how to iterate and control the contour. The operator suggests an initial contour, which is quite close to the intended shape. In 3D the image, force is defined in the same way as in 2D, but the internal energy has to be calculated in a slightly different way, which also leads to a modification of the Euler-Lagrange-equation. To improve the detection stage, the energy criterion and the evolution equation are defined in n-dimension and we combine active contours with an implementation of optical flow. There are different ways to perform moving objects segmentation , using different mathematical techniques. Our aims is to discriminate moving objects from a static background. In this work, we compute the optical flow then we use the result of the optical flow as an input in an active contour model. We used optical flow to detect zones where there is real activity. Then, one carries out the detection of objects of interest in these zones. A preprocessing of derivative of the optical flow field is used to delimit zones in what the energy function is minimized, to segment out the three dimensional objects from video sequence.

The first task of the algorithm exploits the cues from motion analysis for moving area detection. Contour based segmentation is a difficult task if we use the optical flow as an input of our video segmentation. Such a function is more related to a region information as consequence the idea to incorporate region information from the inside and outside seems more natural than using this function as boundary potential.

The proposed method has been developed as a pre-processing stage to be used in methodologies and tools for video/image indexing and retrieval by content.

The remainder of the paper is organized as follows. The next section summarizes the optical flow estimation. Section 3 describes tracking and the active contour model. In section 4, results and evaluation are provided, and finally, concluding remarks are offered in section 5.

OPTICAL FLOW BASED 2D+T ACTIVE CONTOUR MODEL

Optical Flow

The optical flow intrinsic image is a vector flow field that records, point wise, the instantaneous velocity of pattern displacements within the plane of image formation. It is widely recognized that the optical flow function must be subjected to further processing in order to recover surfaces and /or shapes whose identity, location, size, attitude (slat ant tilt), rotation, and translation parameters are canceled within the spatiotemporal patterns of the time varying optical flow field. At least three criteria are generally judged to be important for rating the performance of optical flow methods. First, optical flow methods should lead to flow fields that have high resolution in both space and time. The optical flow vector obtained at each spatiotemporal point should accurately represent the velocity within a small volume dV=(dx, dy, dt) of image (space, time) rather than representing the averaged velocity over a more extensive volume of space and time. Second, flow derivation methods should be sufficiently general to be applicable to a wide range of natural imagery rather than require strong restrictions on the image formation process and scene content. Third, optical flow methods should not be overly sensitive to noise, which is introduced, for example, by censor electronics [6].

We describe here the optical flow: This is the apparent motion of the brightness patterns across the image plane. This is based on the assumption that if an object changes its position across frames then its intensity pattern remains the same. Given a continuous image f(x,y,t), one can form the Taylor series expansion of this function as:

$f(x+dx,y+dy,t+dt)=f(x,y,t)+f_x dx+f_y dy+f_t dt$ +higher order terms

where $f_x=\partial f/\partial x$.

According to the optical flow constraint (OFC), object movement in spatial-temporal domain will generate brightness patterns with certain orientations. We suppose that the optical flow is constant on a neighbourhood of each pixel (local neighbourhood).We define the brightness constancy constraint equation (OFC) for standard 2D optical flow as the form:

$$f_x \frac{dx}{dt}+f_y \frac{dy}{dt}+f_t =0 \Leftrightarrow (f_x ,f_y).(v_x,v_y)=-f_t \Leftrightarrow \nabla f.v+f_t=0$$

Horn and Schunk [7] suggest an alternative method for disambiguation based on relaxation. The spatiotemporal gradient method yields $\nabla f.v=-f_t$ where $\nabla f=(f_x,f_y)$ and $v=(v_x,v_y)$. The ambiguity problem is solved subject to a smoothness constraint given in terms of measured Laplacians, $\nabla^2(v_x)$ and $\nabla^2(v_y)$

The laplacians can be approximated as $\nabla^2(v_x)=v_x-\bar{v}_x$ and $\nabla^2(v_y)=v_y-\bar{v}_y$ are be the average velocity components for some image neighborhood. The derivation of an unique solution results from the minimization of a specific energy function given as:

$$E^2(x,y)=(\nabla f.v+f_t)^2+\lambda^2[(\nabla^2(v_x))^2+(\nabla^2(v_y))^2]$$

The first term ensures that the solution is a close approximation to the spatiotemporal equation, while the second term introduced by a Lagrange multiplier ensures vector field solution smoothness. To minimize the energy function E, one has to set the derivative pair $(\frac{\partial E^2}{\partial v_x},\frac{\partial E^2}{\partial v_y})$ equal to zero. This yields the following equation:

$$(\lambda^2+f_x^2)v_x+f_x f_y v_y=\lambda^2\bar{v}_x-f_x f_t$$
$$f_x f_y v_x+(\lambda^2+f_y^2)v_y=\lambda^2\bar{v}_y-f_y f_t$$

Assuming that $\partial^2[\nabla^2(v_x)]/\partial v_x\equiv 1$, the solution to the system of equation is,

$$v_x=\bar{v}_x-f_x\frac{P}{D} \quad \text{and} \quad v_y=\bar{v}_y-f_y\frac{P}{D}$$

The iterative algorithm for multiple frames is :

$t=0$

Initialize $v_x(x,y,0)$ and $v_y(x,y,0)$

For t=1 until max frames do

$$v_x(x,y,t)=\bar{v}_x(x,y,t-1)-f_x\frac{P}{D}$$

$$v_y(x,y,t)=\bar{v}_y(x,y,t-1)-f_y\frac{P}{D}$$

EndFor

Active Surfaces

Deformable statistical models have proven to be highly useful in many applications. Especially with respect to segmentation tasks, they can be very effective and therefore they form an active field of research.

There are two well-known active contour segmentation methods: Geodesic Active Contours by Caselles et al. [8] and

Region Competition by Zhu and Yuille [9]. In both methods, the evolving estimate of the structure of interest is represented by one or more contours. An evolving contour is a closed surface C(x,y,t) parameterized by variables x, y and by the time variable t. The contour evolves according to the following partial differential equation (PDE):

$$\frac{\partial}{\partial t}C(x,y,t)=F\overrightarrow{N}$$

N is the unit normal to the contour, and F represents the sum of various forces that act on the contour in the normal direction. These forces are characterized as internal and external: internal forces are derived from the contour's geometry and are used to impose regularity constraints on the shape of the contour, while external forces incorporate information from the image being segmented. Active contour methods differ in the way they define internal and external forces. Mean curvature of C is used to define the internal force.

We use a regularized function in the energy. We suppose that the statistical parameters can be estimated on respective region. Then , the functional energy exhibit the following form,

$$E(C)=\iiint_{Cin} k_{in}(x,y,t)dxdydt+\iiint_{Cout} k_{out}(x,y,t)dxdydt$$

where the integrands k_{in} and k_{out} depend on f.

Domain Cin (inside C) represent objects to detect while Cout (outside C) represent background.

For example, the piecewise-constant segmentation model of Chan Vese [10], utilized for the experiments favors a curve which yields the least total squared error approximation of the image by one constant inside the curve and another constant outside the curve:

$k_{in}=(f-\mu_{in})^2$ and $k_{out}=(f-\mu_{out})^2$ where μ_{in} and μ_{out} denote the mean values of f f inside and outside C.

In our case, we computed voxel probabilities P. The probability that f(x,y,t) is drawn from the spherical Gaussian distribution $\{\mu,\sigma\}$ is denoted $P(f|\mu,\sigma)$ and can be estimated using:

$$P(f|\mu,\sigma)=\frac{1}{(\sigma\sqrt{2\pi})^3}e^{-\frac{1}{2\sigma^2}\|f-\mu_{in}\|}$$

In many cases, the basic RGB components may provide very valuable information about the environment. However, the perceptual models, such as HSV or HIS, are more intuitive and therefore enable the extraction of characteristics according to the model of human perception. Our goal is the partition of the image sequence into homogeneous regions with colour or texture properties in its interior. We consider each 3D region modelled by a three-variant Gaussian distribution. So the mean vector and the covariance matrix characterises the colour region behaviour and the probability of a voxel j of belonging to a 3D region Ri is given by:

$$P(f_j|\mu,\sigma)=\frac{1}{\sqrt{(2\pi)^3|\Sigma_i|}}e^{-\frac{1}{2}(\vec{f}_j-\vec{\mu}_i)^T\Sigma_i^{-1}(\vec{f}_j-\vec{\mu}_i)}$$

where $\vec{f}_j$ is the voxel colour vector, $\vec{\mu}_i$ is the colour mean vector of the region i and $\sum_i$ its covariance matrix.

In [11] we proposed a method which combines active contours with an implementation of optical flow. We extended the energy formulation of active contour by including an additional force resulting from the calculation of the optical flow.

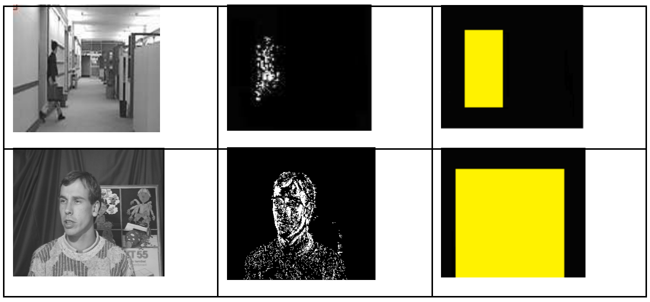
In this work, we propose to use the result of the optical flow as an input in an active contour model. Motion detection produces a coarse motion mask as described in figure 1-b. Each pixel of an image frame f(x,y,t) is classified as moving or stationary by exploiting the result of the corresponding optical flow and a motion blob mask is obtained. Obtained object masks are further refined using mathematical morphology and holes in the motion are filled. Then, active contours are started from the motion blob boundaries and evolved toward moving object boundaries. Since the active contour segmentation relies on the color or intensity differences between background and foreground in the motion mask , the method is more stable and robust across very different appearances, non-homogeneous backgrounds and foregrounds.

RESULTS

To validate the method that we have described, we implemented a moving and tracking object segmentation system. The figure 1 shows optical flow results for sample frames in our test bed sequences. As can be seen from Figure1-a, certain objects can contain holes due to aperture problem. We refined the detection stage using mathematical morphology to fill holes and eliminate noises in the motion. Then, the motion mask is deduced .

Generally, pixel movement is sensitive to moving shadows and illumination changes that can alter object shapes and can result in false detections, making detection and tracking of real object nearly difficult. The parameters for the optical flow method use a neighborhood size W = 9x9x9 and alpha 0.9.

The 2D +t optical flow method produces less noisy and more compact foreground masks compared to pixel based background subtraction methods or 2d , since it integrates temporal information from isotropic spatial neighbourhoods.



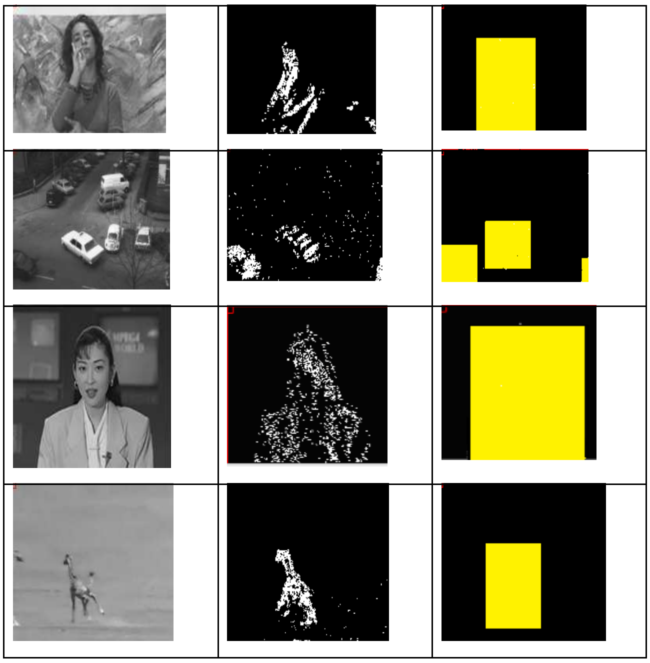


Figure 1: 2d+t Optical flow results and Motion mask

Figure 2 shows the resulting object segmentation at the end of active contour evolution in the image. Active contour evolution moves the contour from initial contours inwards object boundaries. The method refers to the background estimation method by mixture of Gaussians.

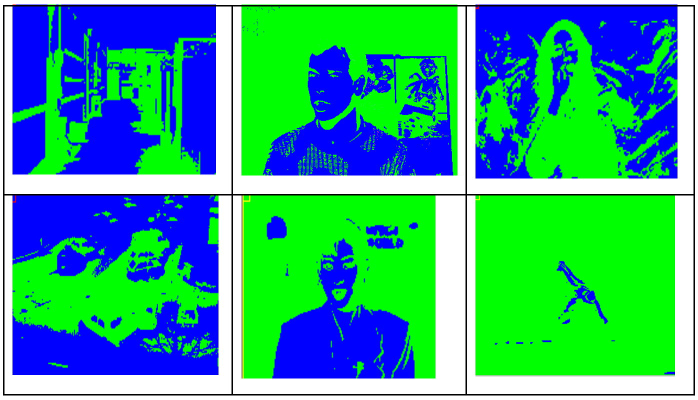


Figure 2: Active contour extraction for our benchmark sequences:

While image sequences particularly sequence 2-a contain non-moving regions such as part of the ground, 2d+t optical flow successfully identifies only the moving objects. Then, active contour evolution moves the contour from the motion mask boundaries inwards toward object boundaries.

Figure 3 illustrates the effect of contour refinement inside the motion mask. Active contours can separate objects due to high color contrast of regions of interest (ROI) compared to background. The approach refine object boundaries and segment objects into individual masks. In figure 4 the input image sequences and its segmented representation are shown, the segmented blobs reflect the semantic objects.

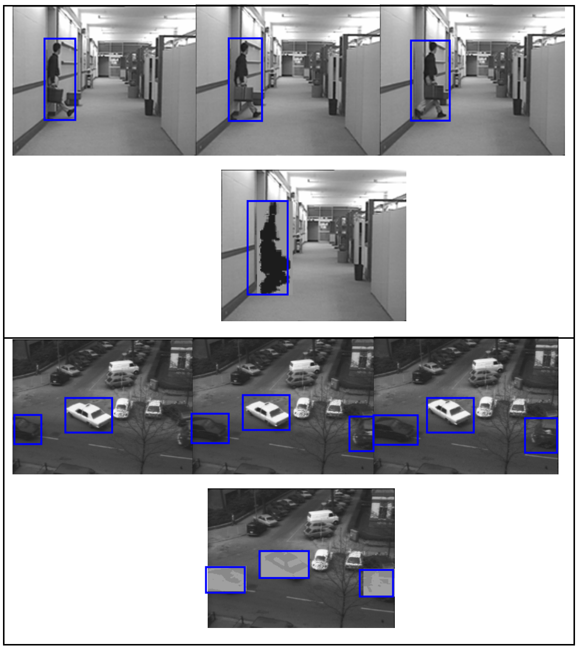


Figure 3:Images sequences

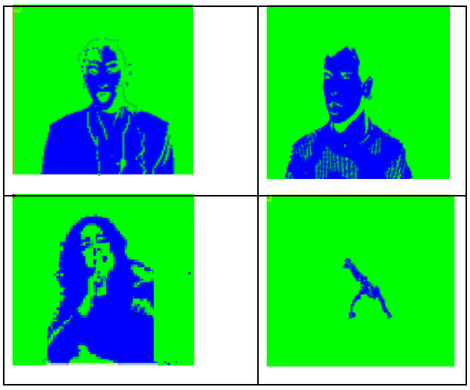


Figure 4:Active contour results with optical flow mask

If the sequence collects noise like movement, it is a critical for persistent object segmentation and tracking.

Discussion and Conclusion

In this paper, we have proposed a method of segmenting 2d+t images based on optical flow and active contour model. This allows bridging the semantic gap from the low level description. The evaluation showed that the proposed approach gives good results according to our test. Nevertheless, sometimes two problems with motion blob detection are:

(1) holes: motion detection produces holes inside slow moving homogeneous objects, because of the aperture problem.

(2) inaccurate object boundaries: motion blobs are larger than the corresponding moving objects, because these regions actually correspond to the union of the moving object locations in the temporal window, rather than the region occupied in the current frame.

The combined approach produce dense and accurate optical flow fields. It provides a powerful and closed representation of the local brightness structure. We construct a segmentation method able to identify what properties characterize objects and distinguish them from other objects and from the background. Also, the method integrate perceptual informations of the color and texture properties in order to better detect the object that has to be segmented.

REFERENCES

[1] L. Haiying, C. Rama, R. Azriel, “Accurate dense optical flow estimation using adaptive structure tensors and a parametric model”, IEEE Image Processing,vol.12, pp.1170-1180,2003,

[2] T. Amiaz and N. Kiryati, “Dense discontinuous optical flow via contour-based segmentation”, in Proc. ICIP 2005, Genova, Italy, September 2005, Vol. III, pp. 1264-1267.

[3] L. Haiying, C. Rama, R. Azriel, “Accurate dense optical flow estimation using adaptive structure tensors and a parametric model”, IEEE Image Processing,vol.12, pp.1170-1180,2003

[4] Weickert, J. and C. Schnorr: 2001, ‘Variational Optic flow Computation with a Spatio-Temporal Smoothness Constraint’. Journal of Mathe. Imaging and Vision 14 245-255

[5] Aubert G., M. Barlaud, and S. Jehan-Besson, ‘Video Object Segmentation Using Eulerian Region-Based Active Contours’. In: ICCV Proceedings. Vancouver, Canada,2001

[6] H. Wechsler, “Computational vision” Academic Press, Inc. pp. 218-219, ISBN 0-12-741245-X

[7] B. K. Horn and B. G. Schunck, “Determining optical flow”, Artificial Intelligence. 1981.

[8] V. Caselles, R. Kimmel et G. Sapiro. Geodesic active contours. IEEE International Conference on Computer Vision, pages 694–699, Boston, MA, USA, 1995.

[9] S. C. 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, pp. 884–900, September 1996.

[10] L. A. Vese and T. F. Chan, “A multiphase level set framework for image segmentation using the Mumford and Shah model”, Int.Journal of Comp. Vision 50(3), 271-293, 2002.

[11] Y. Zinbi, Y. Chahir and A. Elmoataz, Extraction d'objets vidéo : une approche combinant les contours actifs et le flot optique, EGC, Lille, France 2 Volumes. Revue des Nouvelles Technologies de l'Information RNTI-E-6 Cépaduès-Éditions, ISBN 2-85428-718-5, pp. 41-46 , 2006