BAYESIAN HIGH PRIORITY REGION GROWING FOR CHANGE DETECTION

I. Grinias and G. Tziritas

Department of Computer Science, University of Crete
P.O. Box 2208, 71409, Heraklion, Greece
phone: + (30) 2810 393517, fax: + (30) 2810 393501, email: {grinias, tziritas}@csd.uoc.gr

ABSTRACT

In this paper we propose a new method for image segmentation. The new algorithm is applied to the video segmentation task, where the localization of moving objects is based on change detection. The change detection problem in the pixel domain is formulated by two zero mean Laplacian distributions. The new method follows the concept of the well known *Seeded Region Growing* technique, while is adapted to the statistical description of change detection based segmentation, using *Bayesian* dissimilarity criteria in a way that leads to linear computational cost of growing.

1. INTRODUCTION

Video segmentation is a key step in determining the motion features, as well as the position and 2D shape of the scene objects. Such a description may be used either for coding purposes in order to reduce storage and transmission requirements or for indexing and retrieval purposes in order to improve the content description and storage reduction of visual databases. The development of the corresponding international standards MPEG-4 for coding and MPEG-7 for visual content description, which both rely on the concept of audio/visual objects, has raised the importance of these methods.

Several approaches have been proposed for spatio-temporal video segmentation. A recent overview of segmentation tools as well as the object-oriented video description are presented in [1]. Among them, segmentation of image sequences is achieved using techniques such as *Active Contour Models* [2], *Geodesic Active Contours* and *Level Sets* [3] as well as *Region Growing* [4]. In the framework of European Forum COST-211 [5], the so-called *Analysis Model* (AM) for images and video has been proposed, in which colour based and motion based segmentation are combined with the information extracted by change detection. A detailed overview of change detection algorithms is found in [6]. In [7], we proposed an object localization algorithm in which change detection is based on Bayesian tests that are applied on the inter-frame difference, while object localization is achieved using the object colour information. Change detection segmentation was based on the *Multi-label Fast Marching* algorithm, introduced in [8] and is an application of the general segmentation framework of *Bayesian Level Sets*, found in [9].

In this work, we present a new method, which follows the algorithmic structure of *Seeded Region Growing* (SRG) [4] and is adapted to the statistical properties of the features under consideration, in order to speed up the segmentation task without sacrificing the accuracy of the segmentation result. The method is applied to change detection segmentation and requires a map of initial decisions for each label. Initial decisions are then propagated by the new algorithm using *Bayesian* dissimilarity criteria which are based on the statistical description of the *change detection* problem, in a way similar to [7].

2. INTER-FRAME DIFFERENCE

The segmentation algorithm is mainly based on *change detection*. The two classes of "changed"/"unchanged" pixels are modeled by two *Laplacian* distributions. Let $D = \{d(s), s \in S\}$ denote the gray level difference of each site s in the image grid S. The *change detection* problem consists of determining a binary label $\Theta(s)$ for each pixel s. We associate the random field $\Theta(s)$ with two possible events, $\Theta(s) = static$ ("unchanged" pixel), and $\Theta(s) = mobile$ ("changed" pixel). Let $p_{D|static}(d|static)$ (resp. $p_{D|mobile}(d|mobile)$) be the probability density functions of the observed inter-frame difference under the H_0 (resp. H_1) hypothesis. These probability density functions are zero-mean *Laplacian* for both hypotheses ($l = 0, 1$):

$$p(d(s)|\Theta(s) = l) = \frac{\lambda_l}{2} e^{-\lambda_l |d(s)|} \quad (1)$$

Let P_0 (resp. P_1) be the a-priori probability of hypothesis H_0 (resp. H_1). Then, the probability density function is given by

$$p_D(d) = P_0 \;\; p_{D|0}(d|static) + P_1 \;\; p_{D|1}(d|mobile) \quad (2)$$

In this mixture distribution $\{P_l, \lambda_l, l \in \{0, 1\}\}$ are unknown parameters and are estimated using an EM-like algorithm.

3. CHANGE DETECTION GROWING

3.1 Initialization

The algorithm requires some initial correctly labeled sets of pixels. The initialization of label map is achieved by statistical tests of high confidence as it is described in detail in [7]. As in [7], pixels that may be considered "changed" with high confidence are determined using the decision threshold:

$$T_1 = \frac{1}{\lambda_0} \ln \frac{1}{P_{FA}},$$

where P_{FA} is a given small false alarm probability, while a series of tests for each of the remaining pixels is performed in order to determine "unchanged" sites with a small probability of non-detection.

3.2 Growing

The initial decisions are then propagated towards the space of unlabeled image pixels. In SRG, the contour of each initial region is propagated towards the space of unlabeled image pixels, according to dissimilarity criteria which are based on the label and the segmentation features. Contour pixels are sorted according to their dissimilarity from the regions they adjoin and at each step, the contour pixel of minimum dissimilarity is set to the label that most probably belongs. Contour pixels sorting is accomplished by a data structure, traditionally referred to as *Sequentially Sorted List* (SSL).

Although we follow the principle of growing introduced in SRG, a label-dependent term is set according to the *a-posteriori* probability principle, as is the case in [9]. If growing is meant to be based only on the change detection statistics, the dissimilarity of a site s from a label l could be measured as

$$DIS_l(s) = \frac{\sum_{k \neq l} p(d(s)|k(s))Pr(k(s))}{p(d(s)|l(s))Pr(l(s))}$$

or equivalently

$$\begin{aligned} dcd_l(s) &= \ln DIS_l(s) \\ &= \ln \frac{\sum_{k \neq l} Pr(k(s)|d(s))}{Pr(l(s)|d(s))} \\ &= \ln(\sum_{k \neq l} p(d(s)|k(s))Pr(k(s))) \\ &\quad - \ln(p(d(s)|l(s))Pr(l(s))) \end{aligned}$$

In our case of change detection the metric for label 0 becomes

$$dcd_0(s) = \ln(p(d(s)|1)P_1) - \ln(p(d(s)|0)P_0)$$

and under the assumption of Laplacian distributions this gives

$$dcd_0(s) = -\ln \frac{\lambda_0 P_0}{\lambda_1 P_1} + (\lambda_0 - \lambda_1)|d(s)| \qquad (3)$$

and $dcd_1(s) = -dcd_0(s)$. Since P_l, $(l = 0, 1)$ are only estimates and not a-priori knowledge, they have been set to 0.5 in the current implementation of criterion dcd_0. Apparently,

$$dcd_0(s) = -\alpha + \beta|d(s)| \qquad (4)$$

where $\alpha = \ln \frac{\lambda_0}{\lambda_1}$ and $\beta = \lambda_0 - \lambda_1$. In Fig. 1, dcd_l is plotted against $|d|$ for $\lambda_0 = 1.5$ and $\lambda_1 = 0.05$ respectively, for $l = 0, 1$. As we see, for $|d| = \frac{\alpha}{\beta}$ we get the decision point $dcd_0 = dcd_1 = 0$ between the two classes.

The fundamental principle of the new algorithm is that it labels groups of yet unlabeled pixels, at each execution step. Group labeling refers to pixels, which are placed on region contours at an instance of the propagation progress and their metric dcd_l against label l is quite the same. This fact, implies the *quantization* of dissimilarity metric according to the specific characteristics of change detection driven propagation. Towards this direction, by Eq. (4) the random variable $dcd_0(s)$ is a linear function of $|d|$ and by Eq. (2) the probability density function $p(dcd_0(s))$ is given by the mixture

$$\begin{aligned} p(dcd_0(s) = y) &= P_0 p(dcd_0(s) = y|static) + \\ &\quad P_1 p(dcd_0(s) = y|mobile) \end{aligned} \qquad (5)$$

where $y \geq -\alpha$. Using Eq. (1), holds that

$$p(dcd_0(s) = y|\Theta(s) = l) = \frac{\lambda_l}{\beta} e^{-\frac{\lambda_l}{\beta}(y+\alpha)}, \quad y \geq -\alpha \qquad (6)$$

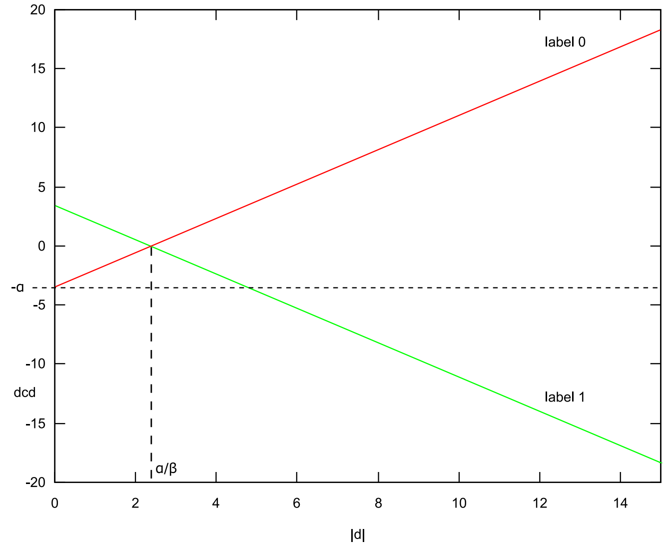

Figure 1: Dissimilarity, change detection based, criteria.

is an exponential distribution for $l = 0, 1$ and Eq. (5) becomes

$$p(dcd_0(s) = y) = P_0 \frac{\lambda_0}{\beta} e^{-\frac{\lambda_0}{\beta}(y+\alpha)} + P_1 \frac{\lambda_1}{\beta} e^{-\frac{\lambda_1}{\beta}(y+\alpha)} \qquad (7)$$

for $y \geq -\alpha$. Similarly, for $y = dcd_1(s)$, we get

$$p(dcd_1(s) = y) = P_0 \frac{\lambda_0}{\beta} e^{-\frac{\lambda_0}{\beta}(-y+\alpha)} + P_1 \frac{\lambda_1}{\beta} e^{-\frac{\lambda_1}{\beta}(-y+\alpha)} \qquad (8)$$

for $y \leq \alpha$. However, since we refer to contour sites, P_k $(k = 0, 1)$ vary during growing and are not the same for the two labels. Thus, for each label l and at growing step T, P_k in Eqs. (7), (8) are dynamically replaced by the percentage $P^T_{l,k}$ of sites which are placed on the contour of regions of label l and according to change detection statistics belong to label k, for $l, k \in \{0, 1\}$. This fact does not affect the statistical analysis for dissimilarity criteria dcd_l $(l = 0, 1)$ that follows, since it is based only on the exponential probability density functions of mixtures, without using percentages $P^T_{l,k}$. In Fig. 2, the probability density functions of the mixtures of Eq. (7) and Eq. (8) respectively, are depicted graphically for $l = 0, 1$ and $\lambda_0 = 1.5$, $\lambda_1 = 0.1$.

As a consequence, if for a contour pixel s of a “changed” region holds that $dcd_1(s) < -\alpha$, then s can be labeled as “changed” with high confidence, since unlikely belongs to an “unchanged” region. Indeed, since

$$P_F(t) = e^{-\frac{\lambda_0}{\beta}(t+\alpha)}, \quad t \geq -\alpha$$

is the *false alarm* probability of labeling a pixel as “changed” while it is “unchanged”, then if we set $t_1 = -\alpha$, for $t \geq -t_1$ holds that

$$P_F(t) \leq e^{-2\alpha\frac{\lambda_0}{\beta}},$$

which is a fairly small false alarm probability. In Fig. 1, certainty limit $y = -\alpha$ is depicted by the horizontal dashed line. By analogy, a threshold t_0 for labeling contour sites s of “unchanged” regions is set, according to equation

$$t_0 = -\alpha - \frac{\beta}{\lambda_1} \ln(1 - P_{ND})$$

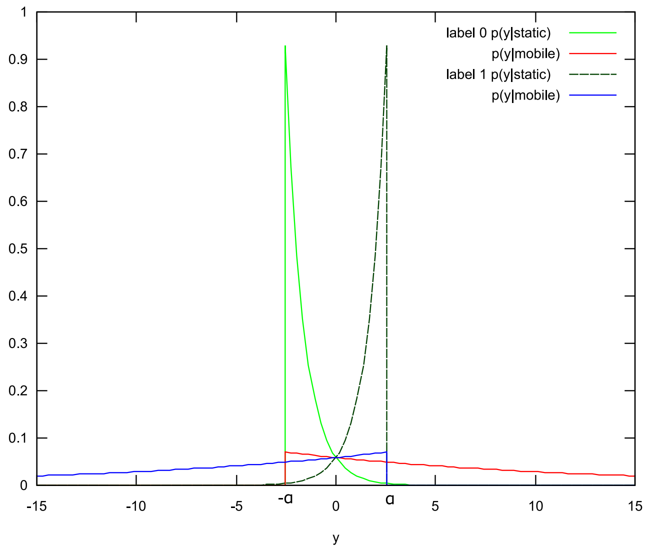

Figure 2: Probability density functions for dcd_0, dcd_1.

where P_{ND} is a given small probability of not detecting a "changed" pixel. Thus, contour sites s, which are placed on the border of an "unchanged" region, with $dcd_0(s) < t_0$, can be labeled as "unchanged" with high confidence. However, a memory term has to be introduced in the dissimilarity criterion of the "changed" class, to cope with the disadvantages in propagating "changed" regions contours, due to possible uniform areas and imperceptible motion of moving objects. After all, the dissimilarity metric of contour pixel s against the "changed" class becomes

$$\delta_1(s) = dcd_1(s) - \rho(s) \tag{9}$$

where $\rho(s)$ is a non-decreasing function of the distance of point s from the border of the previous "changed" mask, computed as in [8]. For clarity, we also set $\delta_0(s) = dcd_0(s)$.

In the implementation of the algorithm, pixels s that satisfy inequalities $\delta_l(s) < t_l$ $(l = 0,1)$ are inserted in *high priority* simply connected lists (HPL_l) once they are scanned. By contrary, the rest of contour pixels are inserted in *normal priority* simply connected lists, denoted as $NPL_{l,i}$ $(l = 0,1)$, each one corresponding to the i-th quantization interval of dissimilarity criteria values, according to a given quantization step t_q:

$$i = Q(\delta_l) = \lfloor \frac{\delta_l - \min\{\delta_l\}}{t_q} \rfloor \tag{10}$$

At each step, first the items of HPL_l for $l = 0,1$ (if any) are popped and assigned to the corresponding label. Otherwise, if high priority lists are empty, the items of lists $NPL_{l,i}$ of minimum i are popped and get labeled.

The description of the new algorithm in pseudo-code is as follows:

S1 Label the initial pixels of classes 0 and 1 in order to form spatially connected labeled regions of pixels R (initialization stage).

S2 Insert all the unlabeled spatial neighbors s of initial regions R into $HPL_{l(R)}$ or $NPL_{l(R),i}$ according to their dissimilarity value $\delta_{l(R)}(s)$, where $l(R)$ denotes the label of region R. Insert s in $HPL_{l(R)}$ if $\delta_{l(R)}(s) < t_l$, otherwise insert it into $NPL_{l(R),i}$, where $i = Q(\delta_{l(R)}(s))$.

S3 While HPL_l or $NPL_{l,i}$, $(l = 0,1)$ are not empty:

- S3.1 if HPL_l are not empty remove (pop) all of their items, otherwise remove (pop) the items of lists $NPL_{l,i}$ of minimum i.
- S3.2 For all the popped items y:
  - S3.2.1 if y is still unlabeled, assign y to the corresponding label k.
- S3.3 For all the popped items y:
  - S3.3.1 add to HPL_k or to $NPL_{k,i}$ pixels s which are neighbors of y and are not already labeled, according to their value of dissimilarity metric $\delta_k(s)$: if $\delta_k(s) < t_k$, add s to HPL_k, otherwise add it to $NPL_{k,i}$, where $i = Q(\delta_k(s))$.

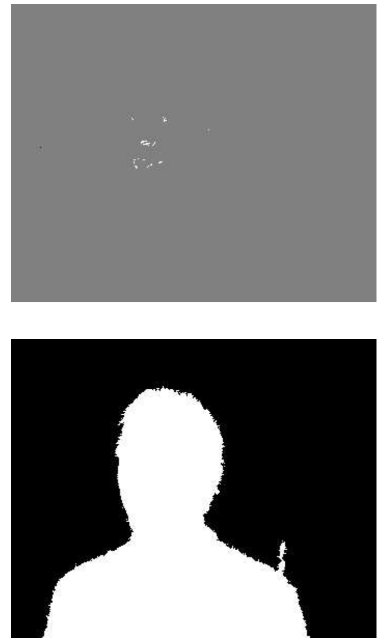

Figure 3: Minimum initial labeling map (up) and propagation result (down) for frame 26 of "erik" sequence.

3.3 Analysis of Growing

The computational cost of the new algorithm is (subject to a constant factor)

$$C = ((1-\xi)I + \xi)n,$$

where n is the number of initially unlabeled pixels, I is the number of quantization intervals and ξ is the percentage of initially unlabeled pixels that are inserted in *high priority* lists. From Eq. (10) I is a function of t_q and since by Eq. (4) the quantization of dcd_l according to $t_q = \beta$ is equivalent to the quantization of $|d|$ in integer steps, a reasonable choice is to set t_q equal to β. By keeping I fixed, the growing time C becomes a linear function of n only. Hence, the cost strongly relates to the decisions made for the high priority thresholds t_l, which in turn, as it is evident from the analysis given in the previous subsection, are adjusted according to the ability of change detection statistics to discriminate the two classes.

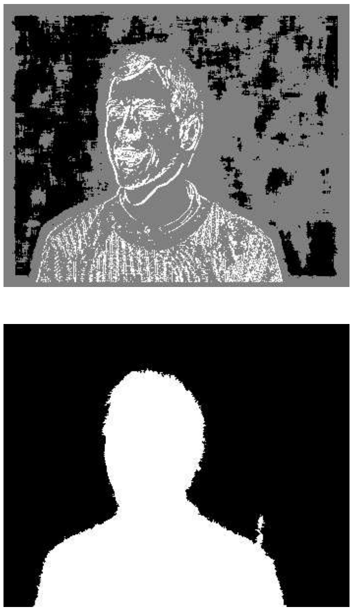

Figure 4: Initialization (up) and propagation result (down) for frame 26 of “erik” sequence.

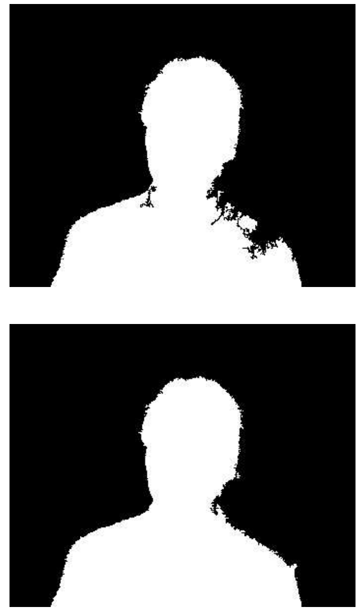

Figure 5: Change detection based propagation result (up) and memory term improvement (down) for frame 40 of “erik”.

Since, as it is argued at the end of this entity, the memory term ρ, used in metric δ_1, improves the discrimination ability of growing, ξ relates to the memory term ρ too. In practice, ξ ranges from 0.2 up to 0.7 leading to an extremely rapid classification method. On the contrary, the cost of SRG does not depend at all on the segmentation task and when SSL is implemented using the optimum choice of Priority Queues (or that of AVL trees) is $O(N \log_2 N)$, where N is the number of image pixels.

Furthermore, an admirable property of the new algorithm is that it is *robust* against the portion of pixels that have been initialized by the label initialization stage. This fact is demonstrated in Figs. 3 and 4, for frame 26 of sequence “erik”. The upper images of the figures, represent the initial label map, while the propagation result is shown in the corresponding bottom image of each figure. In those images, black colour depicts the pixels of “unchanged” class, white colour represents the “changed” pixels and the unlabeled pixels are shown in gray. As we see, although initialization maps differ a lot in the number of pixels that have been initially labeled to one of the classes, the segmentation result is almost identical.

Finally, in Fig. 5 we see the effect of the memory term introduced in metric δ_1, in order to improve the change detection based, segmentation result. Results refer to frame 40 of “erik” sequence, where the change detection statistics are not enough to discriminate all the parts of the moving object from its stationary background, due to low motion activity. The upper image of Fig. 5 depicts the segmentation result which is obtained by setting $\delta_l = dcd_l$ for the two labels, while the bottom image shows the correction that is introduced to the propagation result when Eq. (9) is used for δ_1. Hence, memory term permits the correct classification of pixels to the “changed” class with high confidence, by *tracking* the “changed” mask of the previous segmentation map and using this information as *prior knowledge* in the current growing process.

4. EXPERIMENTAL RESULTS

In what follows, we present change detection segmentation results for well known test image sequences, which are included in the COST [5] data set for testing video segmentation algorithms. However, it should be noticed that the results given hereafter, may contain errors due to occlusions caused by objects motion, or because of shadows. These errors can be corrected by the color based segmentation process which is described in [7] and is applied on the output map of the change detection based segmentation. All the label growing results were extracted using $t_1 = -\alpha$ and $t_q = \beta$. Background is painted in light blue in the images of results, while moving objects appear in their physical color.

In the upper image of Fig. 6 we see the result that has been obtained for frame 26 of sequence “erik”. Change detection statistics give $\lambda_0 = 1.57$, $\lambda_1 = 0.12$ and we set $P_{ND} = 0.01$ in the computation of t_0, while the memory term is used in δ_1. Propagation refers to the initial map shown in Fig. 3. In that, only 56 pixels are initially labeled and thus, more than 99.9% of the pixels are initially unlabeled. During the execution of the growing algorithm 41% of the initially unlabeled pixels were inserted in one of the two high priority lists ($\xi = 0.41$), while the remaining 59% were labeled after their insertion to one of the lists of normal priority, according to their dissimilarity value. However, if t_0 is increased by setting $P_{ND} = 0.05$, we get $\xi = 0.57$, since high priority list HPL_0 is used by more contour pixels of regions belonging to label 0. Thus the computational cost of growing decreases, while the segmentation result remains the same. In the bottom image of Fig. 6, we see the result of growing for frame 40 of “erik”, using $P_{ND} = 0.05$ and the memory term ρ in δ_1. Change detection statistics for this frame, give $\lambda_0 = 1.43$ and $\lambda_1 = 0.085$. As for the results of frame 26, only 189 sites are initially labeled and thus more than 99.9% of image

Figure 6: Change detection based segmentation result for frames 26 (up), 40 (down) of sequence “erik”.

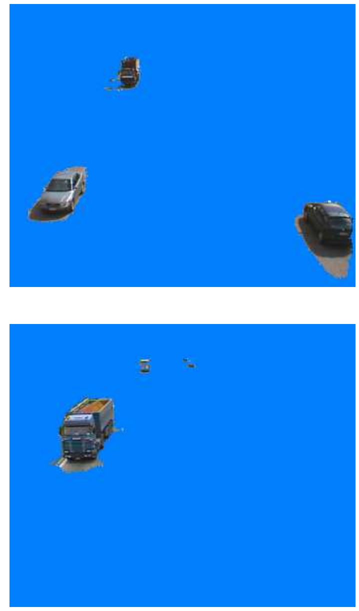

Figure 7: Growing results for frames 195 (up), 240 (down) of sequences “road1” and “road2” respectively.

pixels are unlabeled. High priority lists are used by the 60% of initially unlabeled pixels. On the contrary, when memory term is not used, the result is getting worse, as it is shown in the upper image of Fig. 5, while ξ drops to 45% and consequently the cost of growing increases.

Finally, in images of Fig. 7 we see the results of change detection based growing for sequences “road1”, “road2” respectively, of COST data set. Although, compared to “erik”, for these sequences holds that $\lambda_0 < 0.7$, the change detection growing results remain satisfactory, without the need of further tuning of algorithm ’s parameters. The results shown here, were extracted using $P_{ND} = 0.05$ and the memory term in δ_1, i.e using the growing parameters that were used for the results of “erik”. The average time that is consumed by the growing process for the 300 images of each sequence, is less than 0.1 sec. per frame, on an Intel Centrino 1.6GHz machine with 1GB RAM, running under Linux 2.6.12 OS.

5. CONCLUSION

A new method for image segmentation has been proposed. The new algorithm has been applied to the task of video segmentation. The detection and localization of moving objects was based on change detection. The change detection problem in the pixel domain was formulated by two zero mean Laplacian distributions. The new algorithm is adapted to the statistical description of the change detection problem by introducing the usage of *Bayesian* dissimilarity criteria in the well known *Seeded Region Growing* method, in a way that leads to linear computational cost.

Demonstration: The new method is demonstrated at www.csd.uoc.gr/∼grinias/DEMOS/bhprg_chd/index.html.

Acknowledgement: This work is funded under Greek PENED-2003 project.

REFERENCES

[1] P. Salembier and F. Marques, “Region-based representations of image and video: segmentation tools for multimedia services”, IEEE Trans. on Circuits and Systems for Video Technology, 9:1147-1169, December 1999.

[2] M. Kass, A. Witkin and D. Terzopoulos, “Snakes: Active contour models”, Intern. Journal of Computer Vision, 4:7-38, 1990.

[3] N. Paragios and R. Deriche, “Geodesic active contours and level sets for the detection and tracking of moving objects”, IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-22:266-280, March 2000.

[4] R. Adams and L. Bischof, “Seeded region growing”, IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-16:641-647, June 1994.

[5] A.A. Alatan et al., “Image sequence analysis for emerging interactive multimedia services- the European COST 211 framework”, IEEE Trans. on Circuits and Systems for Video Technology, 8:802-813, November 1998.

[6] R. J. Radke et al., “Image change detection algorithms: a systematic survey”, IEEE Trans. on Image Processing, 14(3):294-307, March 2005.

[7] E. Sifakis, I. Grinias, and G. Tziritas, “Video segmentation using fast marching and region growing algorithms”, EURASIP Journal on Applied Signal Processing, pp. 379-388, April 2002.

[8] E. Sifakis and G. Tziritas, “Moving object localisation using a multi-label fast marching algorithm”, Signal Processing: Image Communication, 16(10): 963-976, 2001.

[9] E. Sifakis and G. Tziritas, “Fast Marching techniques for Visual Grouping and Segmentation”, Chapter in Geometric Level Set Methods in Imaging, Vision and Graphics, edited by S. Osher and N. Paragios, Springer Verlag, July 2003.