

Object of Interest segmentation and Tracking by Using Feature Selection and Active Contours

Mohand Saïd Allili and Djemel Ziou

Department of computer science,

University of Sherbrooke,

J1K 2R1, Sherbrooke (QC), Canada.

Email:ms.allili,d.ziou@usherbrooke.ca.

Tel: +1(819) 821 8000. ext(3247/2859), Fax: +1 (819) 821 8200.

Abstract

Most image segmentation algorithms in the past are based on optimizing an objective function that aims to achieve the similarity between several low-level features to build a partition of the image into homogeneous regions. In the present paper, we propose to incorporate the relevance (selection) of the grouping features to enforce the segmentation toward the capturing of objects of interest. The relevance of the features is determined through a set of positive and negative examples of a specific object defined a priori by the user. The calculation of the relevance of the features is performed by maximizing an objective function defined on the mixture likelihoods of the positive and negative object examples sets. The incorporation of the features relevance in the object segmentation is formulated through an energy functional which is minimized by using level set active contours. We show the efficiency of the approach on several examples of object of interest segmentation and tracking where the features relevance is used.

Keywords: Segmentation, object of interest (OOI), feature relevance, positive & negative examples, mixture model, active contours.

1 Introduction

Recently, we have seen the emergence of applications requiring an object-based segmentation. Among them, we find the content-based image retrieval (CBIR) using objects [4, 9, 12, 20], where the user seeks for images in a database having the same object content as his/her query image. Therefore, the performance of the image retrieval will depend largely on how the segmentation is capable of discerning the objects in the image [3, 8, 13]. Object tracking and coding in video sequences are also applications that

require object-based segmentation [2, 6, 11]. In principle, the more information we have, the better a segmentation algorithm is expected to perform. In practice, however, this is not the case, especially for the segmentation of objects of interest (OOI). In fact, some features may be more relevant to distinguish a specific OOI from the rest of the image, while others can be just "noise", thus not contributing to (or even degrading) the segmentation of the object. Generally, a human subject can detect and recognize an OOI in an image after seeing many instances of the same object (may be in different contexts) and attributing to it some relevant features that make straightforward the task of detecting and recognizing the object. For example, a "tiger" will have almost the same appearance in a lot of natural images. Based on this, knowing the features characterizing a specific OOI, e.g., the "tiger", is of a key importance toward its successful segmentation in new images.

The subject of the present work is the integration of the user knowledge to learn the relevance of the features characterizing an OOI, in order to enhance the accuracy of the OOI segmentation. We learn the feature relevance through a set of positive and negative examples of the same OOI. By positive example, we mean a segmentation of an image where one or several OOI sought by the user are well-localized. On the other hand, a negative example could be another set of objects, or simply the OOI background in the image. According to this, the feature relevance is calculated for the feature discriminative power between the defined positive and negative examples. In the sequel, the feature relevance is integrated in a segmentation framework based on active contours. We show on examples of OOI segmentation and tracking that the accuracy of segmentation is significantly enhanced by using the feature relevance.

This paper is organized as follows: In Section 2, we present the model for feature relevance calculation and how it is incorporated in the segmentation. In Section 3, we show

some experiments related to OOI segmentation and tracking.

2 Learning the segmentation model

We learn the features characterizing a given OOI using a set of images segmented manually or semi-automatically (using a segmentation algorithm and then correcting the segmentation result to form examples of the OOI). In these examples, the user indicates the good segmentation examples, where the OOI is well segmented, as well as the bad examples, which can be the background of the OOI or other segmented objects different from the one of interest. We call these two sets the sets of *positive examples* and *negative examples*, respectively. We suppose that, through his assessment, the user designated N_p positive examples and N_n negative examples. In what follows, let (u_1, u_1, \dots, u_d) be a set of features associated with each pixel of the image (e.g., color, texture, edge orientation). Based on the examples, the aim is to extract the relevance of the features that influenced the user's assessment. In other words, we seek for the features which maximize the similarity within the sets of positive and negative examples and the dissimilarity between the sets of positive and negative examples.

In what follows, the relevance of the features is expressed using a vector of weights $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$, where $\alpha_i > 0, \forall i$. Each component i of this vector expresses the relevance of a feature u_i . As mentioned before, we consider a feature is relevant if it has a good discrimination between the positive and negative examples, which is reflected by a high weight. We assume that the features are conditionally independent given the component labels, and that the sets of positive and negative examples can be composed of one or multiple classes. This assumption enables us to consider for each feature two mixture models, one for the positive examples that we denote by \mathcal{M}_i and one for negative examples that we denote by \mathcal{M}'_i . We suppose that M and M' are the numbers of components in the mixture models \mathcal{M}_i and \mathcal{M}'_i . The probability distributions of the feature u_i in \mathcal{M}_i and \mathcal{M}'_i are given by:

$$p(u_i|\Theta_i) = \sum_{k=1}^M \pi_{ik} p(u_i|\theta_{ik}) \quad (1)$$

$$p(u_i|\Theta'_i) = \sum_{l=1}^{M'} \pi'_{il} p(u_i|\theta'_{il}) \quad (2)$$

where $\Theta_i = (\pi_{i1}, \dots, \pi_{iM}, \theta_{i1}, \dots, \theta_{iM})$ and $\Theta'_i = (\pi'_{i1}, \dots, \pi'_{iM'}, \theta'_{i1}, \dots, \theta'_{iM'})$, with $\sum_{k=1}^M \pi_{ik} = 1$ and $\sum_{l=1}^{M'} \pi'_{il} = 1$, designate respectively the parameters of the mixture models fitting the positive and negative examples data for the feature u_i .

2.1 Estimation of the feature weights

To estimate the weighting parameters for the features, we minimize an objective function defined on the set of positive and negative examples according to the weights $\alpha_1, \dots, \alpha_d$. For each feature, a sample of data is extracted from the positive and negative examples as follows. We subdivide the region of the objects and the background in each image as illustrated in Fig. 1, where the feature vectors are calculated for each square (i.e., each square is associated a vector representing the mean of the feature).

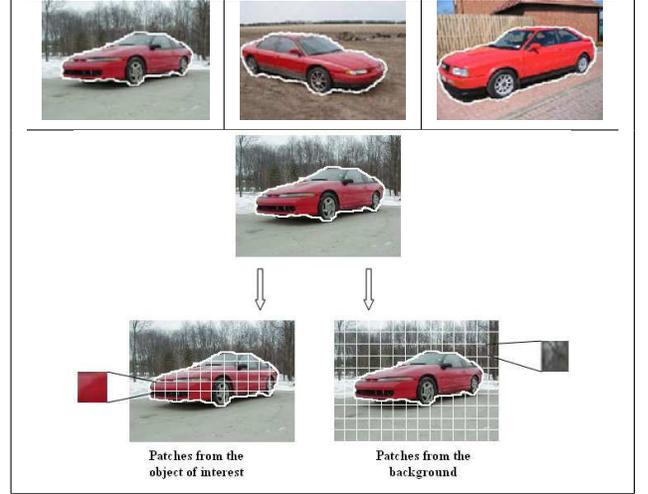


Figure 1. Example of features extraction from positive and negative example as defined by the user.

In past work, the variance of the features was used as an index for determining the relevance of a feature in classification problems [18, 19]. However, the variance is inefficient for choosing the best features that determine the clusters in multi-class data [14]. In our case, we must find the features that best discriminate between the sets of positive and negative examples, each of which may be composed of several classes. To this end, we define the following functions: The *inner likelihood* $\mathcal{L}_I(i)$ for a feature u_i , $i = 1, \dots, d$, which gives the product of the likelihoods of the positive examples with respect to \mathcal{M}_{ip} and the negative examples with respect to \mathcal{M}_{in} :

$$\mathcal{L}_I(i) = \prod_{p=1}^{N_p} p(u_{ip}|\Theta_i) \prod_{n=1}^{N_n} p(v_{in}|\Theta'_i) \quad (3)$$

where u_{ip} and v_{in} denote, respectively, the values of the feature u_i for the p^{th} positive example and n^{th} negative example. The terms $p(u_{ip}|\Theta_i)$ and $p(v_{in}|\Theta'_i)$ are those given

in eqs. (1) and (2). The *outer likelihood* $\mathcal{L}_O(i)$ for the feature u_i gives the product of the likelihoods of the positive examples with respect to \mathcal{M}'_i and the negative examples with respect \mathcal{M}_i :

$$\mathcal{L}_O(i) = \prod_{n=1}^{N_n} p(v_{in}|\Theta_i) \prod_{p=1}^{N_p} p(u_{ip}|\Theta'_i) \quad (4)$$

It is clear that a feature u_i has a good discrimination between the sets of positive and negative examples if the value of the likelihood factor $\frac{\mathcal{L}_I(i)}{\mathcal{L}_O(i)}$ is high ($\gg 1$). Conversely, the feature u_i is not discriminative if the value of the factor $\frac{\mathcal{L}_I(i)}{\mathcal{L}_O(i)}$ is low ($\rightarrow 1$). Formally, let us define the weights $\alpha_i, i = 1, \dots, d$, that represent the relevance of the features $u_i, i = 1, \dots, d$, subject to $\sum_{i=1}^d \alpha_i = 1$. To determine the weights of the features, we propose to minimize the following objective function according to each $\alpha_i, \forall i$:

$$\mathcal{F}(\alpha_1, \dots, \alpha_d) = \prod_{i=1}^d \left(\frac{\mathcal{L}_I(i)}{\mathcal{L}_O(i)} \right)^{1/\alpha_i} \quad (5)$$

subject to $\sum_{i=1}^d \alpha_i = 1$. Note that since the logarithm is an increasing function, minimizing $\log(\mathcal{F})$ is equivalent to minimizing \mathcal{F} . We then define the following function:

$$\varphi(\alpha_1, \dots, \alpha_d) = \sum_{i=1}^d \frac{1}{\alpha_i} (A_i - R_i) - \lambda \left(\sum_{i=1}^d \alpha_i - 1 \right) \quad (6)$$

where λ is a lagrange multiplier to take into account the constraint on the α_i 's to sum to 1. The terms A_i and R_i are given by $A_i = \log(\mathcal{L}_I(i))$ and $R_i = \log(\mathcal{L}_O(i))$. After minimizing the function (6) according to the weighting parameters and straightforward manipulations, we obtain the following optimal weight α_i for the feature u_i :

$$\alpha_i = \frac{\sqrt{f_i}}{\sum_{j=1}^d \sqrt{f_j}} \quad (7)$$

with $f_i = A_i - R_i$. Let us analyze the meaning of the equation (7). The term f_i is proportional to the factor between the *inner* and *outer likelihoods* of the positive and negative examples with respect to the feature u_i , defined respectively in eqs. (3) and (4). Eq. (7) tells that a weight α_i for a feature should be high if its likelihood factor is higher than the likelihood factors of the other features. This corresponds exactly to our objective of obtaining a relevance of a feature by its power of discrimination between the positive and negative examples.

2.2 The model for OOI segmentation

The objective of this section is to build a model for segmentation that takes into account the relevance of the features calculated in the previous section. Suppose an image

I containing the OOI in the focus-of-attention that we want to segment using the d features for which the relevance is calculated for the object. In what follows, we denote the domain of the image by Ω and by R the domain of the object to be segmented. We denote by $u_i(\mathbf{x}), \forall i$, the value of the feature u_i at the pixel $\mathbf{x} = (x, y)$. We consider a modeling of the OOI and the background by using two mixture models \mathcal{M}_i and \mathcal{M}'_i for each feature u_i , with their set of parameters Θ_i and Θ'_i . The optimal segmentation of the OOI, which takes into account the features relevance, is the one that builds a partition $\mathcal{P} = \{R, R^c\}$ of the image, made of the object region R and the background R^c , with $R \cup R^c = \Omega$ and $R \cap R^c = \emptyset$, which maximizes the following function:

$$\mathcal{J} = \prod_{i=1}^d \left[\prod_{(\mathbf{x}) \in R} p(u_i(\mathbf{x})|\Theta_i) \prod_{(\mathbf{x}) \in R^c} p(u_i(\mathbf{x})|\Theta'_i) \right]^{\alpha_i} \quad (8)$$

Putting to logarithm the function (8), the maximization of the above function is equivalent to the minimization of the following function $E = -\log(\mathcal{J})$ given by:

$$E = - \sum_{i=1}^d \alpha_i \left[\sum_{(\mathbf{x}) \in R} \log[p(u_i(\mathbf{x})|\Theta_i)] + \sum_{(\mathbf{x}) \in R^c} \log[p(u_i(\mathbf{x})|\Theta'_i)] \right] \quad (9)$$

To interpret the functions (8) and (9), note that the weights α_i have the inverse effect than in the function (5). To explain this, let us look at the behavior of the functions (8) and (9) according to the feature weights. A partition \mathcal{P} that does not separate well the OOI from the background would give low probabilities for the most relevant features (with high values of α). Putting these probabilities to a high power α will penalize highly the partition \mathcal{P} with regard to these features. This is reflected in function (9) by increasing the energy with regard to these features, thereby emphasizing the features contribution to find the best partition that decreases the energy.

To implement the above segmentation model, we use level set formalism [17]. The goal is to initialize a contour $\vec{\gamma}$ in the image and deform it in a direction that minimizes the function (9), according to the contour $\vec{\gamma}$ and the mixture parameters Θ_i and Θ'_i . To this end, we can re-write the function (9) in a continuous form by replacing the sums by integrals and adding a regularization term that smoothes the object contour. The final energy functional is given as follows:

$$E = \nu \int_0^{L(\vec{\gamma})} ds - \sum_{i=1}^d \alpha_i \left[\iint_R \log(p(u_i(\mathbf{x})|\Theta_i)) d\mathbf{x} \right]$$

$$- \left[\iint_{R^c} \log(p(u_i(\mathbf{x})|\Theta'_i)) d\mathbf{x} \right] \quad (10)$$

where $L(\vec{\gamma})$ is the length of the curve $\vec{\gamma}$ and s is the arc-length parameter. The constant parameter ν controls the contribution of the smoothing term in the functional. To minimize the functional (10), according to the contour $\vec{\gamma}$ and the parameters Θ_i and Θ'_i , we use Euler-Lagrange equations. Let $\Phi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be a level set distance function where $\vec{\gamma}$ is represented by its zero level set, i.e.: $\vec{\gamma} = \{\mathbf{x}/\Phi(\mathbf{x}) = 0\}$. The minimization of the functional (10) according to $\vec{\gamma}$ will lead to the following level set motion equation:

$$\frac{d\Phi}{dt} = \kappa \|\nabla\Phi\| - \sum_{i=1}^d \alpha_i \log \left(\frac{p(u_i(\Phi)|\Theta_i)}{p(u_i(\Phi)|\Theta'_i)} \right) \|\nabla\Phi\| \quad (11)$$

where κ stands for the curvature of the zero level set. For the minimization of the functional (10) according the mixture parameters, remark that the integrals given in the function (10) amounts to calculate the maximum likelihood estimation of the mixture parameters inside and outside the contour $\vec{\gamma}$ given the data inside and outside the contour. By assuming each feature's distribution modeled by a mixture of Gaussian pdfs, the updating equations for the mixture parameters are given as following for the object and the background regions, $\forall i = 1, \dots, d, \forall j = 1, \dots, M_{ip}, \forall l = 1, \dots, M_{in}$:

$$\hat{\pi}_{ij} = \frac{\iint_R p(\theta_{ij}|u_i(\mathbf{x})) d\mathbf{x}}{\iint_R d\mathbf{x}} \quad (12)$$

$$\hat{\pi}'_{il} = \frac{\iint_{R^c} p(\theta'_{il}|u_i(\mathbf{x})) d\mathbf{x}}{\iint_{R^c} d\mathbf{x}} \quad (13)$$

$$\hat{\mu}_{ij} = \frac{\iint_R p(\theta_{ij}|u_i(\mathbf{x})) u_i(\mathbf{x}) d\mathbf{x}}{\iint_R p(\theta_{ij}|u_i(\mathbf{x})) d\mathbf{x}} \quad (14)$$

$$\hat{\mu}'_{il} = \frac{\iint_{R^c} p(\theta'_{il}|u_i(\mathbf{x})) u_i(\mathbf{x}) d\mathbf{x}}{\iint_{R^c} p(\theta'_{il}|u_i(\mathbf{x})) d\mathbf{x}} \quad (15)$$

$$\hat{\sigma}_{ij} = \frac{\iint_R p(\theta_{ij}|u_i(\mathbf{x})) (u_i(\mathbf{x}) - \hat{\mu}_{ij})^2 d\mathbf{x}}{\iint_R p(\theta_{ij}|u_i(\mathbf{x})) d\mathbf{x}} \quad (16)$$

$$\hat{\sigma}'_{il} = \frac{\iint_{R^c} p(\theta'_{il}|u_i(\mathbf{x})) (u_i(\mathbf{x}) - \hat{\mu}'_{il})^2 d\mathbf{x}}{\iint_{R^c} p(\theta'_{il}|u_i(\mathbf{x})) d\mathbf{x}} \quad (17)$$

where the posterior probabilities are given by: $p(\theta_{ij}|u_i(\mathbf{x})) = \frac{\pi_{ij} p(u_i(\mathbf{x})|\theta_{ij})}{\sum_{k=1}^M \pi_{ik} p(u_i(\mathbf{x})|\theta_{ik})}$ and $p(\theta'_{il}|u_i(\mathbf{x})) = \frac{\pi'_{il} p(u_i(\mathbf{x})|\theta'_{il})}{\sum_{k=1}^{M'} \pi'_{ik} p(u_i(\mathbf{x})|\theta'_{ik})}$. The updating of the mixture parameters is performed iteratively until convergence, i.e.: $\|\hat{\Theta}_i^{(old)} - \hat{\Theta}_i^{(new)}\| < \epsilon$ and $\|\hat{\Theta}'_i^{(old)} - \hat{\Theta}'_i^{(new)}\| < \epsilon$, where $(\hat{\Theta}_i^{(old)}, \hat{\Theta}'_i^{(old)})$ and $(\hat{\Theta}_i^{(new)}, \hat{\Theta}'_i^{(new)})$ represent, respectively, the vector of parameters for the mixtures \mathcal{M}_i and \mathcal{M}'_i for two successive iterations in the mixture

updating using eqs. (12)-(17). The threshold ϵ is determined experimentally. The final algorithm for the proposed model for features relevance calculation and segmentation is summed up in the following script:

Algorithm:

Input(positive and negative examples, input image)

Output (segmentation of the OOI)

1-Calculate the features weighting parameters α_i using eq. (7).

2-Object contour initialization.

3-For each feature u_i , initialize the parameters for \mathcal{M}_i and \mathcal{M}'_i .

4-Repeat until convergence:

- Propagate the level set function Φ using eq. (11).

- Update the mixture parameters using eqs. (12)-(17) until $\|\hat{\Theta}_i^{(old)} - \hat{\Theta}_i^{(new)}\| < \epsilon$, and $\|\hat{\Theta}'_i^{(old)} - \hat{\Theta}'_i^{(new)}\| < \epsilon$.

-End of the algorithm.

3 Experiments

3.1 Application to image segmentation

In our experiments, we applied the proposed model for the segmentation of several OOI in natural images, including cars, tigers, zebras and faces. For each OOI, we used a database containing different images of the object and we learn the relevance of the features on a set of positive and negative examples chosen by the user, as explained in Section 2.1. For color information in the tiger, zebra and the car objects, we used the CIE-L*a*b* space, as it is perceptually uniform. For the face, we used the components H and S of the HSV color space as they allow for a good separation between skin and non-skin pixels in an image [15]. For texture information, we used features derived from the correlogram matrix of each pixel neighborhood [1]. We recall that an element of the correlogram matrix, $C^{v,\phi}(c_i; c_j)$, gives the probability that, given a pixel \mathbf{x}_1 of color c_i , a pixel \mathbf{x}_2 at distance v and orientation ϕ from \mathbf{x}_1 is of color c_j [10]. We calculated the correlogram for 4 orientations $\phi \left\{ 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4} \right\}$, and 4 displacements $v \{1, 2, 3, 4\}$, and we derived from each correlogram the following features: Mean (ME), Variance (VR), Contrast (CT), Energy (ER), Entropy (ET), Homogeneity (HG) and Correlation (CR). We set the parameters (M, M') , as follows: (3, 8) for the red-car object, (3, 5) for the face object and (2, 5) for the tiger and zebra objects.

In Fig. 2, we show a sample of images where the feature weights are calculated for the above OOIs. The extraction of the positive and negative examples is performed as explained in Fig. 1 where we used 15 different images for each OOI. In Fig. 3, we show the value of the calculated feature weights for each object. Note that, for the texture features, we averaged the value of each calculated feature over the number of orientations ϕ and displacements v . We

remark that for the red-car object, the features I^* and a^* of color and the features VR and ME of texture are the most relevant. For the faces, the features H and S of color and the feature EN of texture are the most relevant. For the tigers, the most relevant features are the feature a^* of color and the features EN, ET and HG of texture. Finally, for the zebras, all the color features are relevant, while the features EL, VR and ME are the most relevant for texture.

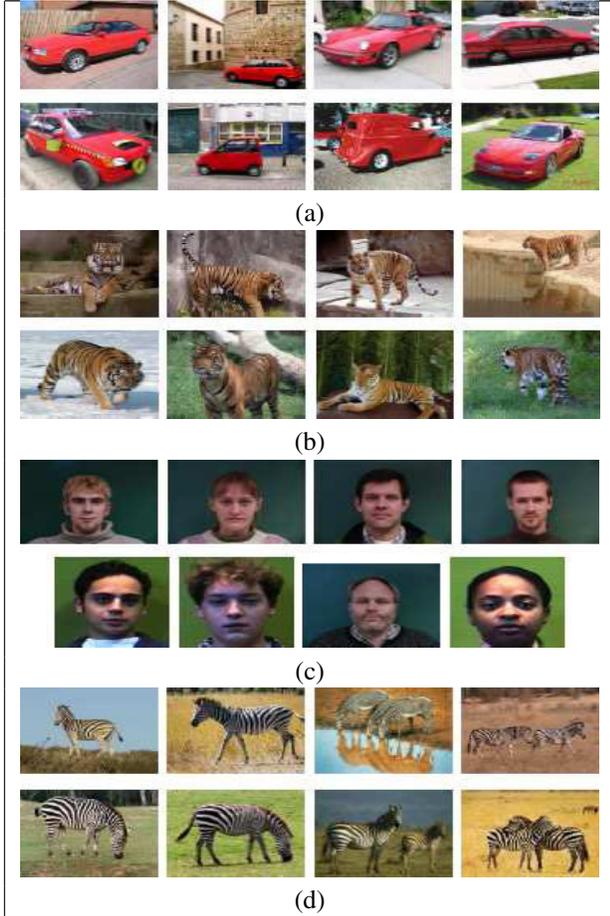


Figure 2. Sample images used for learning the segmentation of: (a) red-cars, (b) tigers, (c) faces and (d) zebras.

To demonstrate the performance of incorporating the feature relevance, we run our segmentation model on a set of images with and without the feature relevance (to eliminate the features relevance, the weights $\alpha_i, \forall i$, are assigned the value 1). The database where the test is performed contains 40 images cars, 100 images of faces, 50 images of tigers and 50 images of zebras. In each segmentation, the object contour was initialized using a rectangle with a center aligned with the image center. The width and height of

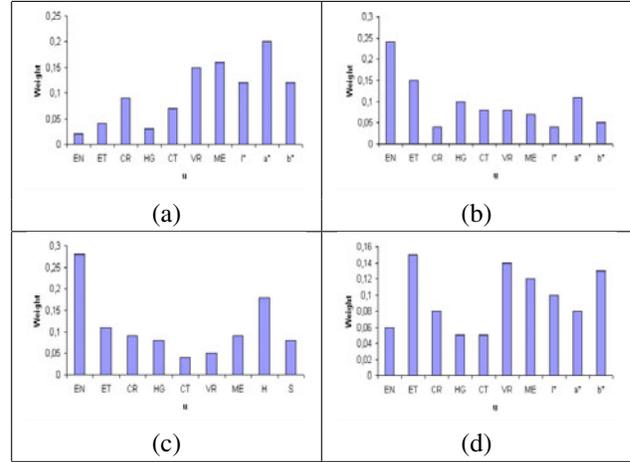


Figure 3. Values of the color and texture feature weights for the: (a) red-cars, (b) tigers, (c) faces and (d) zebras.

the rectangle in each segmented image are $\frac{1}{3}$ of the width and height of the image, respectively. In Fig. (4), (5), (6) and (7) we show, respectively, the segmentation of a sample of red-cars, faces, tigers and zebras, where the performance of the using features relevance was the most significant. The first row of each figure shows the contour initialization. The second and third rows show the final segmentation of the images using and without using the relevance of the features in the segmentation. To evaluate how far the segmentation deviate from the ground truth, we measure the *Kullback – Leibler* (KL) distance between the mixture models resulting from the segmentation of the OOI and the learnt mixture models using the positive and negative examples (considered as as the ground-truth). Recall that the KL distance between two distributions f and g is given by: $KL(f, g) = \int_z f(z) \log(f(z)/g(z)) dz$, where z is the variable of the distributions.

We can note that, for all the segmented OOI, when using the feature relevance the segmented object statistics are much more closer to the ground truth than without using the feature relevance. Intuitively, this may be explained by the effect of weighting the features, in narrowing the space of mixture models in which the segmentation of the OOI may converge. In other words, having the segmentation guided most importantly by the relevant features, i.e., which have the best discrimination between the positive ad negative examples in the learning phase, allows us to drive the segmentation toward objects that are "similar" to the positive examples and "dissimilar" to the negative examples.

Finally in Fig. 8, we show some examples where the proposed model failed in capturing the OOI. In the first two examples, the segmentation of the tiger failed because no

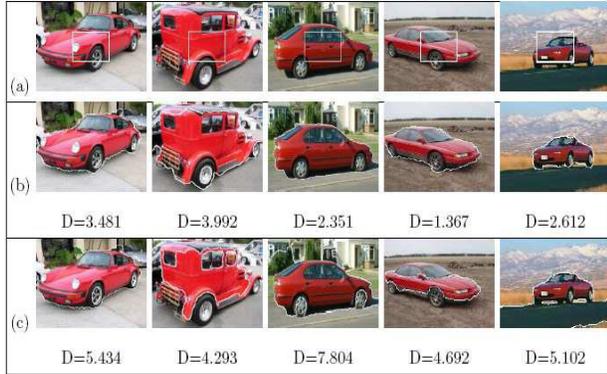


Figure 4. Extraction of the red-car object in a set of images: (a) shows the contour initialization, (b) shows the segmentation using the feature relevance, (c) shows the segmentation without using the feature relevance.

feature was discriminative enough to separate the OOI from the background. In the third example, the OOI was missed in two parts. The bottom part of image is highly illuminated and the zebra feet were erroneously segmented to the background. The top part is covered by a thick shadow produced by the trees behind the zebra. Having some light beams passing through the trees created a texture in the shadow area similar to the one of the zebra, which distracted the contour to include pixels from the background to the interior of the object. In the fourth example, as the beard has the color and texture as the hair, the contour has missed the beard area of the face. Finally, the contour erroneously included the tinted hair into the face object. Indeed, the color and texture of the hair are very similar to be discriminated using the above color and texture features. We believe that using more features, such as prior shape information for OOI, should mitigate the above problems.

3.2 Application to image object tracking in videos

In the past, video object tracking using appearance models has been substantially investigated [7, 21]. Using object histogram, for example, can achieve efficient tracking through partial occlusion and pose variation [7]. However, the tracking success or failure depends primarily on how distinguishable the object is from its surroundings. Indeed, as the target object moves, for several reasons (e.g., shadows, illumination changes) the appearance of both the object and background may change.

The contribution of the present paper for tracking is threefold. First, we use an extensive set of features that

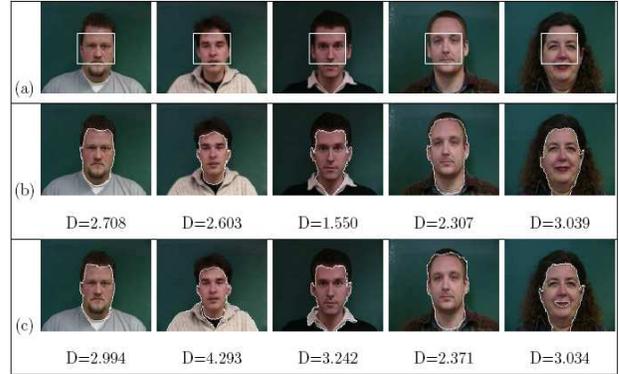


Figure 5. Extraction of the face object in a set of images: (a) shows the contour initialization, (b) shows the segmentation using the feature relevance, (c) shows the segmentation without using the feature relevance.

include color and texture features for tracking. Thus, our model is capable of selecting the best features for tracking when the surroundings of the object changes from a texture to non-texture area and vice versa. Second, the feature weights for tracking are determined using our probabilistic model that selects through time the features which maximize the *inner likelihood* of the target object and the background and minimize the *outer likelihood*, according to eq. (7). Finally, using level set formalism for tracking allows us for tracking the boundaries of non-rigid objects and handling automatically topology changes for the objects.

In Fig. 10, we show an example of tracking using the feature relevance. The sequence is composed of 102 frames, where the immediate neighborhood of the target object changes constantly color and texture. To determine the feature weights to be used for tracking in the current frame, we extract the positives and negative examples using the object location resulting from the tracking in the $C = 4$ last frames. The extraction of the set of positive and negative examples in each frame is performed as follows. Suppose that the bounding rectangle B of the object contour has a height H and a width W in the frame as shown in Fig. 9, where the object contour is represented by the red line. The surrounding of the object is defined as the part of the background delimited by the rectangle B' having the same center as B and having the height $H+2L$ and the width $W+2L$, where L is chosen sufficiently high (60 pixels in our experiments). Given the objects and their surroundings for the last C frames, the positive and negative examples are extracted from the objects and the surroundings, respectively. From these examples, we calculate the mixture models \mathcal{M}_i and \mathcal{M}'_i for the feature u_i , and the feature weights using

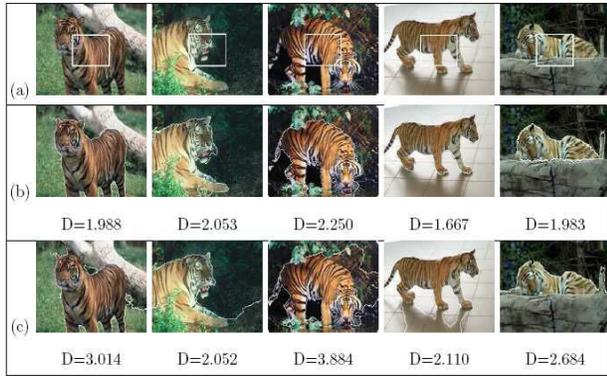


Figure 6. Extraction of the tiger object in a set of images: (a) shows the contour initialization, (b) shows the segmentation using the feature relevance, (c) shows the segmentation without using the feature relevance.

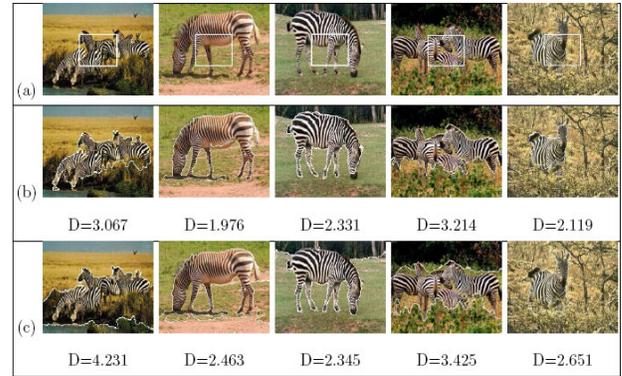


Figure 7. Extraction of the zebra object in a set of images: (a) shows the contour initialization, (b) shows the segmentation using the feature relevance, (c) shows the segmentation without using the feature relevance.

eq. (7). The tracking is then performed inside the rectangle B' by deforming the contour from the last position of the object using eq. (11).

In Fig. 10, we show the result of tracking by, respectively, using and without using the feature relevance. Indeed, using the feature relevance allowed to overcome the distraction of the contour and having a successfully tracking for the object, whereas, without the use of feature relevance, we had to re-initialize the contour in the frames where the contour has been distracted (see Fig. 10).

4 Conclusions

In this paper, we proposed to incorporate the relevance of low-level features for the segmentation of OOI. We can summarize the main contributions in the present work as follows. First, we proposed a model for the calculation of the feature relevance. This is performed through a set of positive and negative examples where the feature weights are determined by optimizing an objective function. Second, we incorporated the features relevance in segmentation and tracking of OOI using active contours. Several examples showed the advantage of using the feature relevance for enhancing the accuracy of object segmentation.

5 Acknowledgements

The completion of this research was made possible thanks to the Natural Sciences and Engineering Council of Canada (NSERC) and Bell Canada's support through its Bell University Laboratories R&D programs.



Figure 8. Examples of failure of the extraction of the OOI using the feature relevance.

References

- [1] M.S. Allili and D. Ziou. Automatic Color-Texture Image Segmentation by Using Active Contours. *IEEE Int'l Workshop on Intelligent Computing in Pattern Analysis and Synthesis*, LNCS 4153, 495-504, 2006.
- [2] M.S. Allili and D. Ziou. A Robust Video Object Tracking by Using Active Contours. *IEEE Conf. on Computer Vision and Pattern Recognition Workshop*, p. 135, 2006.
- [3] E. Borenstein, E. Sharon and S. Ullman. Combining Top-Down and Bottom-Up Segmentation. *IEEE Conf. on Computer Vision and Pattern Recognition Workshop*, P. 46, 2004.
- [4] C. Carson, S. Belongie, H. Greenspan and J. Malik. Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(8):1026-1038, 2002.
- [5] T. Chan and L. Vese. Active Contours Without Edges. *IEEE Trans. on Image Processing*, 10(7):266-277, 2001.
- [6] H-T. Chen, T-L. Liu and C-S. Fuh. Probabilistic Tracking with Adaptive Feature Selection. *IEEE Int'l Conf. on Pattern Recognition*, II-736-739, 2004.

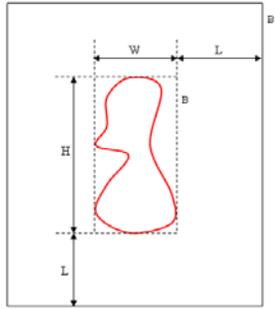


Figure 9. Illustration of the object surroundings defined using the bounding object rectangle B and the rectangle B'.

- [7] D. Comaniciu, V. Ramesh and P. Meer. Kernel-based Object Tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(5):564- 577, 2003.
- [8] J.H. Elder, A. Krupnik and L.A. Johnston. Contour Grouping with Prior Models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(6):661-674, 2003.
- [9] D.A. Forsyth, J. Malik, M. M. Fleck, H. Greenspan, T. Leung, S. Belongie, C. Carson and C. Bregler. Finding pictures of objects in large collections of images. *2nd Int'l Workshop on Object Representation in Computer Vision*, 1996.
- [10] J. Huang and S. R. Kumar and M. Mitra and W.-J. Zhu and R. Zabih. Spatial Color Indexing and Applications. *Int'l J. of Computer Vision*, 35(3):245-268, 1999.
- [11] R.T. Collins, Y. Liu. Online Selection of Discriminative Tracking Features. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(10):1631-1643, 2005.
- [12] M.L. Kherfi and D. Ziou. Relevance Feedback for CBIR: A New Approach Based on Probabilistic Feature Weighting with Positive and Negative Examples. *IEEE Trans. on Image Processing*, 15(4):1017-1030, 2006.
- [13] B. Ko and H. Byun. FRIP: A Region-Based Image Retrieval Tool Using Automatic Image Segmentation and Stepwise Boolean AND Matching. *IEEE Trans. on Multimedia*, 7(1):105-113, 2005.
- [14] M. Law, M. Figueiredo and A.K. Jain. Simultaneous Feature Selection and Clustering Using a Mixture Model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(9):1154-1166, 2004.
- [15] S.Z. Li and A.K. Jain. Handbook of Face Recognition. Springer, 2004.
- [16] N. Paragios and R. Deriche. Geodesic Active Regions and Level Set Methods for Supervised Texture Segmentation. *IEEE Int'l Conf. on Computer Vision*, 46(3):223-247, 2002.
- [17] S. Osher and J. Sethian. Fronts Propagating With Curvature-Dependant Speed: Algorithms Based on Hamilton-Jacobi Formulations. *J. of Computational Physics*, 79(1):12-49, 1988.

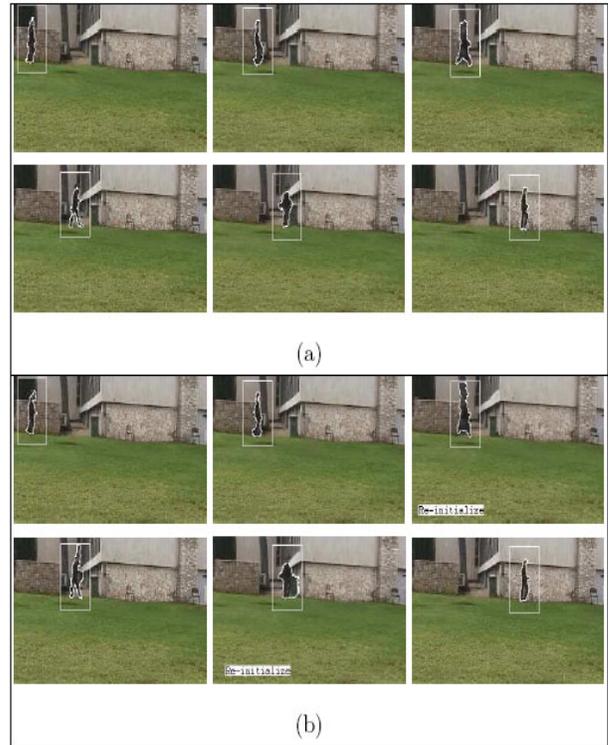


Figure 10. The two rows of (a) show the tracking result using the feature relevance for frames 20, 32, 34, 40, 46 and 64, from left to right and top to bottom. The two rows of (b) show for the same frames the tracking result without using the feature relevance.

- [18] V. Roth and T. Lange. Adaptive Feature Selection in Image Segmentation. *Pattern Recognition- 26th DAGM Symposium*, 9-17, 2004.
- [19] Y. Rui. Efficient Indexing Browsing and Retrieval of Image/Video Content. *PhD thesis of Electrical Engineering in the Graduate College of the University of Illinois at Urbana-Champaign*, 1999.
- [20] T. Wang, Y. Rui and J-G Sun. Constraint Based Region Matching for Image Retrieval. *Int'l J. of Computer Vision*, 56(1-2):37-45, 2004.
- [21] T. Zhang and D. Freedman. Improving Performance of Distribution Tracking through Background Mismatch. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(2):282-287, 2005.