

Fusion of a Multiple Hypotheses Color Model and Deformable Contours for Figure Ground Segmentation in Dynamic Environments

Francesc Moreno-Noguer, Alberto Sanfeliu
Institut de Robòtica i Informàtica Industrial
UPC-CSIC, Llorens Artigas 4-6
08028, Barcelona, Spain
fmoreno,asanfeliu@iri.upc.es

Dimitris Samaras
Computer Science Department
State University of New York at Stony Brook
NY 11794-4400, USA
samaras@cs.sunysb.edu

Abstract

In this paper we propose a new technique to perform figure-ground segmentation in image sequences of moving objects under varying illumination conditions. Unlike most of the algorithms that adapt color, the assumption of smooth change of the viewing conditions is no longer needed. To cope with this, in this work we introduce a technique that formulates multiple hypotheses about the next state of the color distribution (some of these hypotheses take into account small and gradual changes in the color model and others consider more abrupt and unexpected variations) and the hypothesis that generates the best object segmentation is used to remove noisy edges from the image. This simplifies considerably the final step of fitting a deformable contour to the object boundary, thus allowing a standard snake formulation to successfully track nonrigid contours. Reciprocally, the contour estimation is used to correct the color model. The integration of color and shape is done in a stage denominated 'sample concentration', that has been introduced as a final step to the well-known CONDENSATION algorithm.

1. Introduction

Color and deformable contours have been extensively used in computer vision applications, such as object detection and tracking tasks [1] [5]. Usually, these methods are based on a first step where the object is roughly (but robustly) located by the color module. This simplifies the subsequent step of accurately fitting the contour model to the rigidly or non-rigidly deformed object boundary. In environments with controlled lighting conditions and uncluttered background, color can be considered a reliable and invariant cue, but when dealing with real scenes with changing illumination and confusing backgrounds, the apparent color of the objects varies considerably over time,

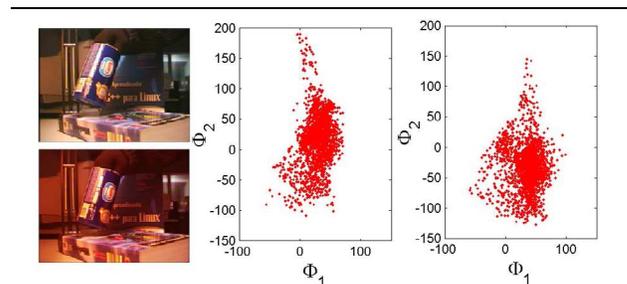


Figure 1. Abrupt change of illumination. Left: Two consecutive sequence frames. Light conditions have changed abruptly (from natural to red illuminant). Center and right: Corresponding Color distributions of the foreground (the can). Φ_1 and Φ_2 are the coordinates in a 2D colorspace.

and an important challenge for any figure-ground segmentation system to work in real unconstrained environments, is the ability to accommodate these changes (Fig. 1).

In the literature, the techniques that cope with change in color appearance can be divided in two groups. On the one side, there is a group of approaches that search for color constancy (e.g. [3]); but in practice, these methods work mostly in artificial and highly constrained environments. On the other hand, there are the techniques that generate a stochastic model of the color distribution, and adapt this model over time. In this sense, in [11], color is represented by a histogram that is adapted online, as the weighted function of previous histograms, and a predicted one. Yang and Lu [13], parameterize object color by a unique gaussian, the mean and covariance of which are estimated using a linear combination of the parameters in previous gaussians. Raja and McKenna [12] approximate color with a mixture of gaussians, and dynamically update it using a weighted sum

of previous estimates with estimates based on new data.

The drawback in all these approaches is that they assume that color varies slowly and that it can be predicted by a dynamic model based in only one hypothesis. However, this assumption does not suffice to cope with general scenes, where the dynamics of the color distribution might follow an unknown or unpredictable path.

In order to cope with these unconstrained environments, we propose a system with the following features, that are the main contributions of the paper:

- **Multihypotheses framework:** The use of a particle filter formulation to predict the color distribution in subsequent iterations, offers robustness to abrupt and unexpected changes in the color appearance of the object. In previous work [8] we have suggested a similar multihypotheses framework to track objects in which color could be approximated by an unimodal distribution, represented by a histogram. In the present work, we deal with multicolored objects, approximated by a mixture of gaussians (*MoG*). Note the difference between our work and all previous tracking approaches using a particle filter formulation (e.g. [2] [5] [10]), where multiple hypotheses are generated about the position of the object and not about the distribution of object color.
- **Integration of color and deformable contours in a particle filter framework:** The color estimation is used to generate a rough estimation about the object position and remove noisy edges from the image. This simplifies the stage of fitting a deformable contour to the object boundary, and even with a standard *snake* formulation, **nonrigid** objects can be accurately tracked in cluttered backgrounds with abrupt changes of illumination. The fusion of the multihypotheses color model and the deformable contour is done in a final stage that we have introduced to the well-known CONDENSATION algorithm.

In the following sections, a detailed description of the method is given. In Section 2 the object color model and initialization step are presented. The dynamic model for generating multiple hypotheses of the (object and background) color distributions is depicted in Section 3. Section 4 deals with the global and local deformable model fitting process. In Section 5, the complete tracking algorithm and model adaptation is explained in detail, and results and conclusions are presented in Sections 6 and 7, respectively.

2. Color Model

The selection of the colorspace is an important initial issue for any color-based figure ground segmentation system. The typical selection criterion is based on the invari-

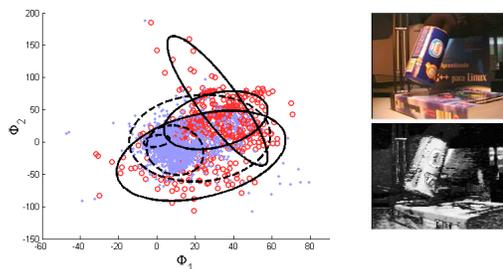


Figure 2. Gaussian mixture components of \mathcal{O} (the can) and \mathcal{B} . Left image: Filled dots and continuous lines are \mathcal{O} data points (in Fisher colorspace) and gaussian components, respectively. Empty dots and dashed lines are \mathcal{B} data and gaussians. Lower right image: $p(\mathcal{O}|\mathbf{x})$, where brighter points correspond to more likely pixels.

ance of the color representation to illumination changes, and according to this idea, color is usually represented by two components of the *rgb*, *HSV* or *CIE xyz* (lower case letters represent normalized components) colorspace or by linear combinations of the *RGB* components. However, these representations are not robust enough to cope with abrupt illumination changes. In the experiments presented in this paper, we use a colorspace (*Fisher colorspace*) that maximizes the separability of the foreground and background classes [9].

In order to represent the color distribution of a monochrome object, color histograms have been demonstrated to be an effective technique (e.g.[8]). However, when the object to be modeled contains regions with different colors, the number of pixels representing each color can be relatively low and a histogram representation may not suffice. In this case, a better approach is to use the *MoG* model, that expresses the conditional probability for a pixel \mathbf{x} belonging to a multi-colored object \mathcal{O} as a sum of M_o gaussian components: $p(\mathbf{x}|\mathcal{O}) = \sum_{j=1}^{M_o} p(\mathbf{x}|j) P(j)$. Similarly, the background color will be represented by a mixture of M_b gaussians.

Given the classes foreground (\mathcal{O}) and background (\mathcal{B}), the a posteriori probability that a pixel \mathbf{x} belongs to object \mathcal{O} is computed using the Bayes rule:

$$p(\mathcal{O}|\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{O}) P(\mathcal{O})}{p(\mathbf{x}|\mathcal{O}) P(\mathcal{O}) + p(\mathbf{x}|\mathcal{B}) P(\mathcal{B})} \quad (1)$$

where $P(\mathcal{O})$, $P(\mathcal{B})$ represent the a priori probabilities of \mathcal{O} and \mathcal{B} , respectively. These values are approximated to the expected relative area values of each class (see Fig. 2).

Similar to the problem of selecting the number of bins

in histogram models, using *MoG* conceals the challenge of choosing the number of gaussian components that better adjust the data. We initialize this, with the modified *EM* algorithm proposed in [4], that is based on a *Minimum Message Length* criterion and iteratively fits and annihilates an initially large number of components (introduced by the user).

Once we have learnt the initial configurations of the *MoG* for \mathcal{O} and \mathcal{B} , they are parameterized by:

$$\mathcal{G}_\varepsilon = [\mathbf{p}_\varepsilon, \mu_\varepsilon, \lambda_\varepsilon, \theta_\varepsilon] \quad (2)$$

where $\varepsilon = \{\mathcal{O}, \mathcal{B}\}$, \mathbf{p}_ε contains the priors for each gaussian component, μ_ε the centroids, λ_ε the eigenvalues of the principal directions and θ_ε the angles between the principal directions with the horizontal. $\mathcal{G} = \{\mathcal{G}_\mathcal{O}, \mathcal{G}_\mathcal{B}\}$ will be the state vector representing the color model.

3. Dynamic Color Model

Let $\mathcal{X}_{\varepsilon,t-1} = [\mathbf{x}_{1,t-1}, \dots, \mathbf{x}_{N_\varepsilon,t-1}]^T$, be the vectors containing the set of points (in a particular colorspace) belonging to the classes \mathcal{O} and \mathcal{B} , at time $t-1$. One of the stages of the tracking algorithm, consists of propagating the components $\mathcal{G}_{\varepsilon,t-1}$ of the state vector to $\tilde{\mathcal{G}}_{\varepsilon,t}$, given a specific dynamical model and the image at time t , denoted as \mathbf{Z}_t . Instead of applying directly the dynamic model to $\mathcal{G}_{\varepsilon,t-1}$, we apply it to the distribution $\mathcal{X}_{\varepsilon,t-1}$, to obtain the estimation $\tilde{\mathcal{X}}_{\varepsilon,t}$, that will be used later to reestimate $\tilde{\mathcal{G}}_{\varepsilon,t}$. With this aim we define the following *affine random* dynamic model:

$$\tilde{\mathcal{X}}_{\varepsilon,t} = \mathcal{A}_\varepsilon \mathcal{X}_{\varepsilon,t-1} + \mathbf{v}_\varepsilon \quad (3)$$

In the case of representing color distributions in a 2-dimensional colorspace, matrix \mathcal{A}_ε and translation vector \mathbf{v}_ε are written as:

$$\mathcal{A}_\varepsilon = \begin{bmatrix} 1 + a_{\varepsilon,11} & a_{\varepsilon,12} \\ a_{\varepsilon,21} & 1 + a_{\varepsilon,22} \end{bmatrix} \quad \mathbf{v}_\varepsilon = \begin{bmatrix} v_{\varepsilon,1} \\ v_{\varepsilon,2} \end{bmatrix}$$

Variables $a_{\varepsilon,ij}$ and $v_{\varepsilon,i}$ are approximated by normal random distributions, $a_{\varepsilon,ij} \sim \mathcal{N}(0, \sigma_{a_{\varepsilon,ij}})$, $v_{\varepsilon,i} \sim \mathcal{N}(\mu_{v_{\varepsilon,i}}, \sigma_{v_{\varepsilon,i}})$, where parameters $\sigma_{a_{\varepsilon,ij}}$ and $\sigma_{v_{\varepsilon,i}}$ are learned a priori by a least-squares procedure, from a training sequence of a still object under changing illumination. The parameter $\mu_{v_{\varepsilon,i}}$ accounts for the expected displacement between the class distributions in $t-1$ and t , and is approximated by the translation vector between the centroids of the sets $\mathcal{X}_{\varepsilon,t-1}$ and \mathcal{X}_t . Note that the vector $\mathcal{X}_t = [\mathcal{X}_{\mathcal{O},t}, \mathcal{X}_{\mathcal{B},t}]^T$ representing the color distribution of the whole image \mathbf{Z}_t is known, but the subsets $\mathcal{X}_{\mathcal{O},t}$ and $\mathcal{X}_{\mathcal{B},t}$ are unknown.

Using the *EM* algorithm initialized on $\mathcal{G}_{\varepsilon,t-1}$, a new mixture of Gaussians $\tilde{\mathcal{G}}_{\varepsilon,t}$ is fitted to each predicted distribution $\tilde{\mathcal{X}}_{\varepsilon,t}$, and used to compute the a posteriori probability maps for the foreground class, following eq. 1. In Fig. 3 we show several hypotheses (with the corresponding $p(\mathcal{O}|\mathbf{x})$

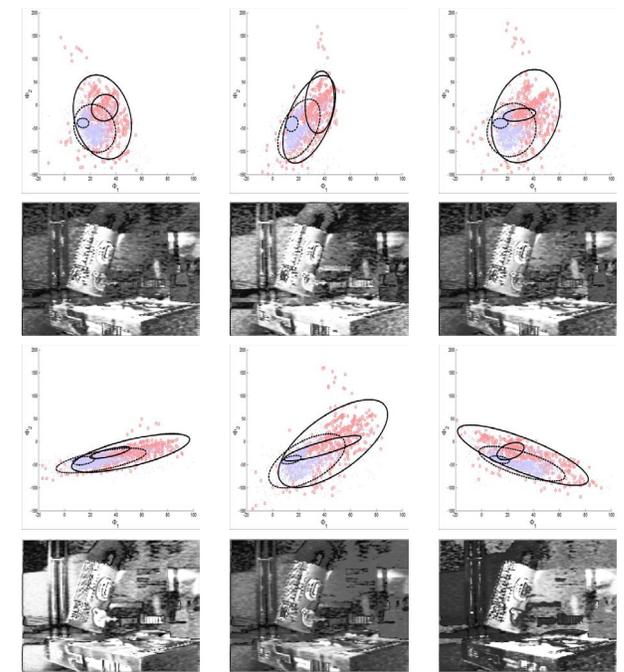


Figure 3. Several hypotheses and their respective $p(\mathcal{O}|\mathbf{x})$ map, corresponding to the abrupt illumination transition presented in Fig. 1

maps) used to estimate the abrupt change of illumination that exists in the pair of images of Fig. 3. Observe that some of the hypotheses are able to provide a ‘good’ foreground/background discrimination.

4. Global and Local Deformable Model Fitting

As color segmentation usually only gives a rough estimation about the object location, we use a deformable model ([2], [7]) to fit its boundary and obtain a more precise information about its position. This process is highly simplified by using the data that is estimated by the color model (Section 3) in order to preprocess the contour image and to remove those noisy edges that might disturb the deformable model fitting process. This simplification allows us to obtain good tracking results in rigid and nonrigid objects, even using the simple and classical snake algorithm [6]. In the boundary adjustment process, first a global fit of an affine contour is performed, that basically deals with object translation and orientation (rigid motion), followed by local deformations that consider nonrigid motions. The following are some details of these processes:

Let the contour of the object be parameterized by a curve $\mathbf{r}(s) = [u(s), v(s)]$, $s \in [0, 1]$, that moves through the image. In the traditional *snake* formulation [6], the problem of snake fitting can be viewed as a force balance equation:

$$\mathbf{F}_{int}(\mathbf{r}(s)) + \mathbf{F}_{ext}(\mathbf{r}(s)) = 0 \quad (4)$$

where $\mathbf{F}_{int}(\mathbf{r}(s)) = \alpha \frac{\partial^2 \mathbf{r}(s)}{\partial s^2} + \beta \frac{\partial^4 \mathbf{r}(s)}{\partial s^4}$ are the internal forces that control the bending and stretching of the snake (α and β are the elasticity and rigidity parameters, respectively). $\mathbf{F}_{ext}(\mathbf{r}(s))$ are the external forces that pull the curve towards the edge image features. In the literature, there exist several definitions for this external function. In particular, we use the *Gradient Vector Flow* (GVM) external force proposed in [14], because it has a larger capture range and better convergence performance in boundary concavities than other methods.

Eq. 4 is solved by making the snake also a function of time, i.e., $\mathbf{r}(s, t)$ (we will write \mathbf{r}_t) and iterating over the following expression:

$$\frac{\mathbf{r}_t - \mathbf{r}_{t-1}}{\Delta t} = \alpha \frac{\partial^2 \mathbf{r}_{t-1}}{\partial s^2} + \beta \frac{\partial^4 \mathbf{r}_{t-1}}{\partial s^4} + \mathbf{F}_{ext}(\mathbf{r}_{t-1}) \quad (5)$$

When the solution stabilizes ($\mathbf{r}_{t-1} = \mathbf{r}_t$), eq. 4 is satisfied.

For the numerical implementation we approximate the derivatives with finite differences, and discretize the curve $\mathbf{r}(s, t)$ with N_P points, so that the previous gradient descent method can be rewritten as:

$$\mathcal{R}_t = (I - \Delta t \mathcal{Q})^{-1} (\mathcal{R}_{t-1} + \Delta t \mathbf{F}_{ext}(\mathcal{R}_{t-1})) \quad (6)$$

where $\mathcal{R} = [(u_1, v_1, 1), \dots, (u_{N_P}, v_{N_P}, 1)]^T$ contains the homogeneous coordinates of the N_P discretized points of the snake, \mathcal{Q} is a $N_P \times N_P$ pentadiagonal matrix including the α and β parameters, and I is the N_P -identity matrix.

Iterating over eq. 6 the snake is locally fitted to the image edges, only governed by the internal and external forces. But previous to this stage we perform a global deformation in order to find the suitable translation and orientation of the snake. For this fitting, the next additional constraint of affine deformation is introduced to eq. 6:

$$\mathcal{R}_t = \mathcal{R}_{t-1} \mathcal{H}_A = \mathcal{R}_{t-1} \begin{bmatrix} a_{11} & a_{12} & v_1 \\ a_{21} & a_{22} & v_2 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

Combining equations 6 and 7, we obtain the following iterative procedure for the affine snake deformation:

1. $\mathcal{H}_A = (\mathcal{J}^T \mathcal{J})^{-1} \mathcal{J}^T (\mathcal{R}_{t-1} + \mathbf{F}_{ext}(\mathcal{R}_{t-1}))$
where $\mathcal{J} = \mathcal{R}_{t-1} - \Delta t S \mathcal{R}_{t-1}$
2. Normalize \mathcal{H}_A using the component $\mathcal{H}_A(3, 3)$
Set $\mathcal{H}_A(3, 1) = \mathcal{H}_A(3, 2) = 0$

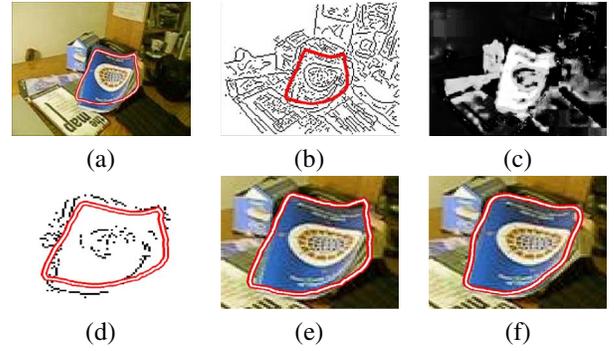


Figure 4. Global and local fitting procedures: (a) Original image and contour from previous iteration. (b) Edge features image. The process of contour fitting in such image is quite difficult because of the noisy edges. (c) Foreground a posteriori probability map obtained using the color module. (d) Refined edge image. (e) Contour fitted after global deformations. (f) Contour fitted after local deformations.

$$3. \mathcal{R}_t = \mathcal{R}_{t-1} \mathcal{H}_A$$

Steps 1 – 3 are iterated until the convergence of R_t and R_{t-1} . In Fig.4 we show the results of the global and local fitting in a nonrigid movement.

5. Tracking Algorithm

In this section, we will use the tools described previously to explain in detail the whole method implemented to track rigid and nonrigid objects in a cluttered and changing illumination environment. The basic steps of the tracking algorithm follow the procedure of the particle filters, but we introduce a modification to the classical CONDENSATION algorithm (similar to the ICONDENSATION technique [5]), and in order to ‘direct’ the search for the next iteration we add a final stage that concentrates the future hypotheses on those areas of the state-space containing more information about $p(\mathcal{O}|\mathbf{x})$. Moreover, in this final stage we fuse object color and shape information to obtain precise results about object pose. Next, we present the steps of our algorithm:

- **pdf of the color point set:** At time t , a set of N samples $\mathcal{S}_{t-1}^{(n)}$ ($n = 1, \dots, N$) with the same structure as \mathcal{G} , parameterizing N color distributions, are available from the previous iteration. Each sample has an associated weight $\pi_{t-1}^{(n)}$ and a classification $\mathcal{X}_{t-1}^{(n)} = [\mathcal{X}_{\mathcal{O}, t-1}^{(n)}, \mathcal{X}_{\mathcal{B}, t-1}^{(n)}]^T$ of the image colorpoints in the foreground and background sets. The whole set

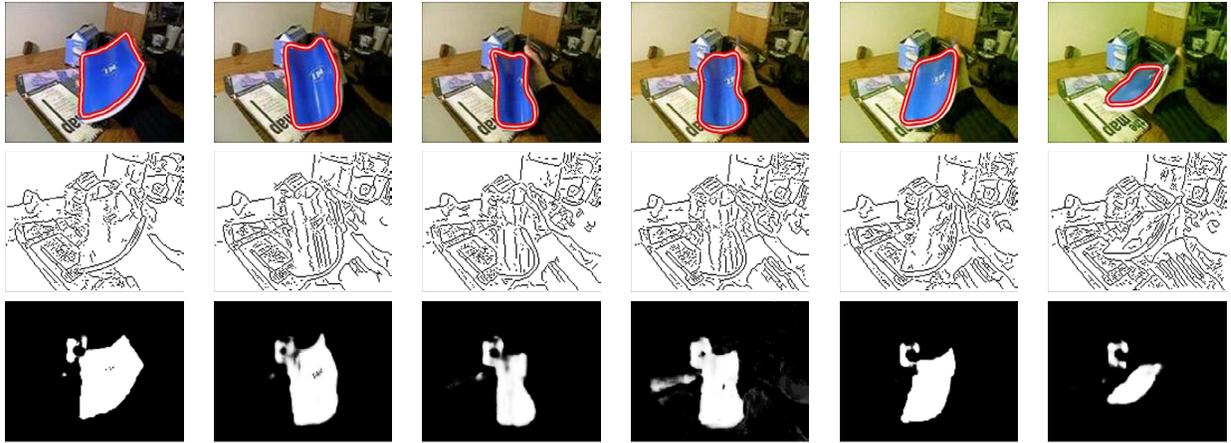


Figure 5. Tracking results of a non-rigid object in a sequence with smooth lighting changes. First row: Tracking results. The tracked contour is superimposed to the original images. Second row: Edge map. The task of fitting a deformable model to the contour of the object is extremely difficult because of the presence of noisy edges. Third row: Foreground a posteriori map obtained using the proposed multihypotheses color model. This map provides a rough estimation of the object position and remove most of the noisy edges, so that the deformable contour fitting procedure is highly simplified.

represents an approximation to $p(\mathcal{G}_{t-1}|\mathcal{Z}_{t-1})$ where $\mathcal{Z}_{t-1} = \{\mathbf{Z}_0, \dots, \mathbf{Z}_{t-1}\}$ is the history of the images. The algorithm aims to construct a new sample set $\{\mathcal{S}_t^{(n)}, \pi_t^{(n)}\}$ to estimate $p(\mathcal{G}_t|\mathcal{Z}_t)$.

- **Sampling from $p(\mathcal{G}_{t-1}|\mathcal{Z}_{t-1})$:** A sampling with replacement is performed N times on the set $\{\mathcal{S}_{t-1}^{(n)}\}$, where each element has probability $\pi_{t-1}^{(n)}$ of being chosen. This will give us a set $\{\mathcal{S}'_{t-1}^{(n)}\}$.
- **Probabilistic propagation of the samples:** Each sample $\mathcal{S}'_{t-1}^{(n)}$ is propagated to $\tilde{\mathcal{S}}_t^{(n)}$, using the dynamic model explained in Section 3.
- **Measure and Weight:** Each element $\tilde{\mathcal{S}}_t^{(n)}$ has to be weighted according to some measured features. Based on the propagated *MoG* samples $\tilde{\mathcal{S}}_t^{(n)}$ we compute $p(\mathcal{O}|\mathbf{x})$ for the whole image using the Bayes rule (eq. 1). With this probability map, we assign the following weight to each sample:

$$\pi_t^{(n)} = \frac{\sum_{\mathbf{x} \in W} p(\mathcal{O}|\mathbf{x})}{N_w} - \frac{\sum_{\mathbf{x} \notin W} p(\mathcal{O}|\mathbf{x})}{\overline{N_w}} \quad (8)$$

where W is the interest region around the previous object position (where we predict the object will be), and $N_w, \overline{N_w}$ are the number of image pixels in and out of this interest region.

- **Sample Concentration:** In the last stage of our algorithm, we concentrate the samples around the local

maxima, so that in the following iteration the hypotheses are formulated around these more likely regions of the state space. In our case, this is absolutely necessary because the state vector \mathcal{G} has high dimensionality, and if we let the samples move freely, uniquely governed by the dynamic model, the number of hypotheses needed to find the samples representing a correct color configuration, is extremely high.

The concentration is performed by taking the sample with maximum weight, $\pi_t^* = \max\{\pi_t^{(1)}, \dots, \pi_t^{(n)}\}$ and based on the a posteriori map generated by this sample, the object of interest is accurately segmented from the image using the deformable model fitting procedure explained in section 4. The different steps of this stage, can be summarized as follows:

1. Using morphologic operations on the probability map image a coarse approximation of object shape is obtained that allows us to eliminate noisy edges from the original image (Fig. 4b,c,d).
2. The contour of the object in the previous iteration, is used as initialization of an affine snake, that is adjusted (only by affine deformations) to the image of refined edges (Fig. 4e) in order to solve the global deformation. Next, to cope with nonrigid deformations the process is repeated with a non-affine snake (Fig. 4f).

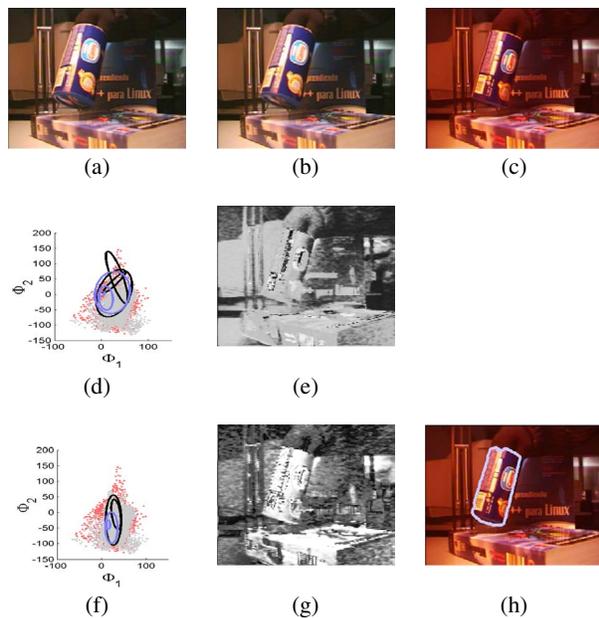


Figure 6. Performance comparison of a smooth prediction color dynamic model and the multihypotheses one, for an abrupt change in illumination and rigid motion object. (a)(b)(c) Frames $t - 2$, $t - 1$ and t are three consecutive images of the sequence. Note the abrupt change in illuminant between frames $t - 1$ and t . (d) Ellipses correspond to the *MoG* predicted with a smooth color dynamic model. There are also shown the real distributions of points in colorspace. (e) $p(\mathcal{O}|\mathbf{x})$ map obtained with the smooth model. There is no good discrimination between fore and background. (f) *MoG* of the best sample using the multihypotheses color dynamic model. (g) $p(\mathcal{O}|\mathbf{x})$ map obtained with this color model. There is good fore/back discrimination. (h) Tracking results obtained after using $p(\mathcal{O}|\mathbf{x})$ to eliminate false edges from image and fitting a deformable contour.

3. Once the boundary of the object has been accurately detected, the color estimations are refined. Inner image pixels are separated from outer pixels and the vector $\mathcal{X}_t^* = [\mathcal{X}_{\mathcal{O},t}^*, \mathcal{X}_{\mathcal{B},t}^*]^T$ is generated. Mixtures of gaussians are fitted to these color distributions (using the EM algorithm), giving a state vector \mathcal{S}_t^* , around which samples $\{\tilde{\mathcal{S}}_t^{(n)}\}$ are ‘concentrated’ with the equation $\mathcal{S}_t^{(n)} = (1 - a)\tilde{\mathcal{S}}_t^{(n)} + a\mathcal{S}_t^*$, where the parameter a governs the level of concentra-

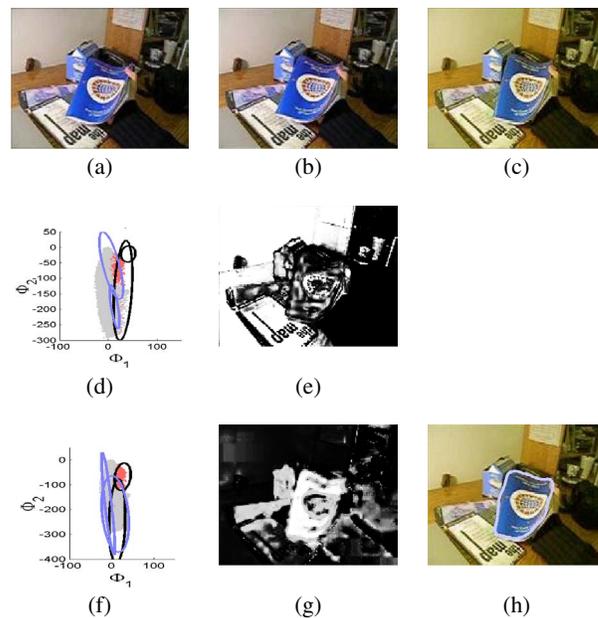


Figure 7. Performance comparison of a smooth prediction color dynamic model and the multihypotheses one, for an abrupt change in illumination and nonrigid motion object. See Fig. 6 for interpretation of results.

tion. Similarly, weights $\{\tilde{\pi}_t^{(n)}\}$ and distributions $\{\tilde{\mathcal{X}}_t^{(n)}\}$ associated to these samples, are concentrated around π_t^* and \mathcal{X}_t^* .

6. Experimental Results

In this section, two different experimental results are presented in order to illustrate the robustness of our system to several changing conditions of the environment. As the algorithm has been implemented in an interpretative language (MATLAB), speed issues are not discussed, focusing on the effectiveness of the method, instead.

In the first experiment, we track the boundary of a bending book (nonrigid motion) in a video sequence where the lighting conditions change smoothly from natural lighting to yellow lighting. In this case, as the displacement of the color distribution in color space was relatively small, we have used ‘only’ 5 hypotheses. Fig. 5 shows some frames of the sequence with the obtained results, the corresponding edge images and the a posteriori probability maps of the foreground (the book). The sequence of edge images contains a lot of noisy boundaries that pose difficulties for the tracking process and for the adjustment of a deformable

model to the edges of the object. However, the integration with color information gives a first estimation of the object position, that allows us to eliminate many false edges and relax the complexity of the deformable model fitting procedure.

Whereas in the first experiment we demonstrate the need for integration of the different vision modules, in the second experiment we demonstrate the need for a multihypotheses model to face abrupt changes in the illuminant. For this experiment, we have computed the prediction of the color distribution using 20 hypotheses. In Fig. 6 we compare the results obtained using a smooth color dynamic model and our multihypotheses model, for a rigid object moving in an environment in which the lighting changes abruptly. The MoG for frame t predicted by the smooth model, is based on a weighting function $\mathcal{G}_t = (1 - a)\mathcal{G}_{t-2} + a\mathcal{G}_{t-1}$, where \mathcal{G} is the parameterization of the color distribution and a is the mixing factor. Results prove the inability of the smooth color model to predict the change (the a posteriori probability map of the foreground region does not discriminate between fore and background, Fig. 6e) whereas a good result is obtained with the method proposed in the paper (where simple morphologic operations over the a posteriori probability map, allow obtain a good estimation of the object position, Fig. 6f). In Fig. 7 we show similar results for the contour tracking of a nonrigid object under an abrupt change of illumination.

7. Conclusions

Most of the methods that adapt color are based on the assumption of smooth change on the color model, so that the predicted color of the target is computed based on a weighting function of previous color distributions. In this paper we have presented a method where this constraint is no longer needed, and the dynamic model is based on the formulation of multiple hypotheses about the next state of the target color distribution. The best of these hypotheses is used to make a rough estimation of the object position, and eliminate false and noisy edges, so that the task of fitting a deformable contour to the object boundary is considerably simplified. Reciprocally, this boundary is used to refine the color estimation. We have used this method to obtain a precise figure-ground segmentation in rigid and nonrigid objects, moving in an environment with abrupt light changes (where smooth dynamic color models fail). We plan to continue this work by integrating other cues such as texture and optical flow techniques to improve the robustness of the method. We also plan to apply our multihypotheses framework into tracking of objects in 3D.

Acknowledgments

This work was supported by CICYT projects DPI2001-2223 and DPI2000-1352-C02-01, and by a fellowship from the Spanish Ministry of Science and Technology.

References

- [1] S.Birchfield, Elliptical Head Tracking Using Intensity Gradients and Color Histograms. *Proc. CVPR*, pp.232-237,1998.
- [2] A.Blake, M.Isard, Active contours. *Springer*,1998.
- [3] G.D.Finlayson, B.V.Funt, K.Barnard. Color Constancy under Varying Illumination. *Proc. ICCV*, pp.720-725, 1995.
- [4] M.A.T.Figueiredo, A.K.Jain. Unsupervised Learning of Finite Mixture Models. *Trans. PAMI*, Vol.24, num.3, pp.381-396, 2002.
- [5] M.Isard, A.Blake Icondensaton: Unifying Low-Level and High-Level Tracking in a Stochastic Framework. *Proc. ECCV*, Vol.1, pp.893-908, 1998.
- [6] M.Kass, A.Witkin, D.Terzopoulos Snakes: Active contour models. *IJCV*, Vol.1, pp.321-331, 1987.
- [7] T.McInerney, D.Terzopoulos Deformable Models in Medical Image Analysis: A Survey *Medical Image Analysis*, Vol.1(2), pp.91-108, 1996.
- [8] F.Moreno-Noguer, J.Andrade-Cetto, A.Sanfeliu. Fusion of Color and Shape for Object Tracking under Varying Illumination. *Proc.IBPRIA, LNCS 2652, Springer*, pp.580-588, 2003.
- [9] F.Moreno-Noguer, A.Sanfeliu. Integration of Shape and a Multihypothesis Fisher Color Model for Figure-Ground Segmentation in Non-Stationary Environments. *To appear in Proc. ICPR04*, 2004.
- [10] K.Nummiaro, E.Koller-Meier, L.Van Gool An Adaptive Color-Based Particle Filter *IVC*, Vol.2(1), pp.99-110, 2003.
- [11] L.Sigal, S.Sclaroff, V.Athitsos. Estimation and Prediction of Evolving Color Distributions for Skin. Segmentation under Varying Illumination. *Proc. CVPR*, Vol.2, pp.152-159, 2000.
- [12] Y.Raja, S.McKenna, S.Gong. Colour Model Selection and Adaption in Dynamic Scenes. *Proc. ECCV*, Vol.1, pp.460-475, 2000.
- [13] J.Yang, W.Lu, A.Waibel. Skin-Color Modeling and Adaption. *Proc. ACCV*, Vol.2, pp.687-694, 1998.
- [14] C.Xu, J.L.Prince. Snakes, Shapes, and Gradient Vector Flow. *Trans. IP*, Vol.7, num.3, pp.359-369, 1998.