REGION TRACKING WITH NARROW PERCEPTION OF BACKGROUND

Julien Mille

Université de Lyon, CNRS
Université Lyon 1, LIRIS, UMR5205
F-69622, France
julien.mille@liris.cnrs.fr

Jean-Loïc Rose

Université de Lyon, CNRS
Université Lyon 2, LIRIS, UMR5205
F-69676, France
jean-loic.rose@liris.cnrs.fr

ABSTRACT

We address the problem of object tracking within image sequences through region-based energy minimization. A common underlying assumption in region tracking is that color statistics can be confidently estimated in a global manner over object and background regions. This can be a drawback for tracking in real scenes with cluttered backgrounds, where statistical color data is highly scattered, preventing the estimation of reliable color statistics for object/background discrimination. To overcome this limitation, we propose an approach based on a *narrow perception* of background, which concentrates on the vicinity of tracked objects and thus extract more consistent statistical data for region separation. The benefits of our approach are demonstrated using two different statistical color models.

Index Terms— Region tracking, energy minimization, color probability density, narrow approach

1. INTRODUCTION

A large variety of variational methods have been proposed and applied to object tracking or segmentation within image sequences. In this context, region-based energy minimization approaches [1, 2] appear generally to be powerful. In [3], the region tracking method is formulated as a Bayesian estimation problem. Local feature distributions of both the object and background regions were used for tracking. Brox and Cremers [4] define an extended version of the Mumford-Shah functional considering local estimation of region statistics. These approaches consider the tracking functional over the entire image domain. This can be a drawback for tracking in real scenes, especially for the background region, which may be cluttered and contain many objects. In such case, statistical color data is highly scattered, so that background distribution may not be confident. To overcome this problem, limiting the spatial range of the energy within a narrow domain may be considered.

There have been several methods in the literature which are related to our work. Local modeling of statistical data within narrow regions was recently addressed for object segmentation in still images (see for instance [5]). This kind of approach enables to relax assumptions made by global region-based methods such as the Chan-Vese model [1], as they basically consist in replacing image homogeneity terms over entire regions with combinations of piecewise homogeneity terms over local subregions. To some extent, we adapt this philosophy to the tracking problem and introduce a relaxed version of the minimization problem.

We propose an approach based on a *narrow perception* of background, which concentrates on the close neighborhood of tracked objetcs to extract statistical color data. It allows to obtain consistent indicators for separation between background and object regions. We provide two possible ways of implementation. The first one is based on kernel estimation of global color Probability Density Functions and the second one relies on a local matching approach and motion prior. Energy minimization is performed thanks to the recent variational region growing approach developed in [6]. Experiments compare our "narrow perception" energy against classical energy and show the efficiency of our model.

2. BAYESIAN INFERENCE FOR REGION TRACKING

Let us consider an input color image sequence where a single object is being tracked at current time t. Given current image frame $\boldsymbol{I}_t$ defined over $\mathcal{D} \subset \mathbb{R}^2$ and a partition into object and backgound regions R_t and $\overline{R}_t$, object tracking consists in determining next region R_{t+1}. We first rely on the *Maximum A Posteriori* (MAP) framework introduced by Mansouri [3], who uses Baye's theorem and assumes conditional independence between image pixels:

$$\begin{aligned} R_{t+1}^* &= \underset{R_{t+1}}{\operatorname{argmax}}\, \mathrm{p}(R_{t+1}|\boldsymbol{I}_t, \boldsymbol{I}_{t+1}, R_t) \\ &= \underset{R_{t+1}}{\operatorname{argmax}} \prod_{\boldsymbol{x}\in\mathcal{D}} \mathrm{p}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t, R_{t+1}) \mathrm{p}(R_{t+1}|\boldsymbol{I}_t, R_t) \end{aligned} \tag{1}$$

Probability $\mathrm{p}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t, R_{t+1})$ is the likelihood of observing a particular color at space-time location $(\boldsymbol{x}, t+1)$ given current image and both current and next object regions. This term will represent our assumptions about color constancy over time, whereas prior probability $\mathrm{p}(R_{t+1}|\boldsymbol{I}_t, R_t)$ models available prior knowledge about object shape and/or motion. A tractable expression is obtained by making the reasonable assumption that the likelihood of observing $\boldsymbol{I}_{t+1}(\boldsymbol{x})$

depends only on $\boldsymbol{I}_t$, R_t and the region which $\boldsymbol{x}$ will belong to at time $t+1$. Moreover, object and background have their distinct color likelihoods in next frame, conditioned on their respective current configurations, leading to a piecewise definition of the likelihood function:

$$\mathrm{p}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t, R_{t+1}) = \begin{cases} \mathrm{p}^{\text{in}}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t) & \text{if } \boldsymbol{x} \in R_{t+1} \\ \mathrm{p}^{\text{out}}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t) & \text{if } \boldsymbol{x} \in \overline{R}_{t+1} \end{cases} \quad (2)$$

With a view to simplicity, we shorten $\mathrm{p}^{\text{in}}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t)$ to $\mathrm{p}^{\text{in}}_{t+1}(\boldsymbol{x})$ and similarly for p^{out} in the remainder of the paper. The MAP estimation of object region R_{t+1} is turned into minimization of energy $E[R_{t+1}]$, taken as the negative log of posterior probability (1):

$$E[R_{t+1}] = -\int_{R_{t+1}} \log \mathrm{p}^{\text{in}}_{t+1}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \int_{\overline{R}_{t+1}} \log \mathrm{p}^{\text{out}}_{t+1}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \log \mathrm{p}(R_{t+1}|\boldsymbol{I}_t, R_t) \quad (3)$$

One may note that if we removed temporal consistency between successive images and object states, both in the likelihood and the prior, the problem would boil down to a two-region segmentation of image $\boldsymbol{I}_{t+1}$ with respect to color distribution, regardless of previous image and object configuration. Such an energy would assume that object and background could be discriminated from each other relying only on their respective color distributions (notable examples include the region competition [7] or the information theory-based approach of [8]).

3. NARROW PERCEPTION OF BACKGROUND

It is common to turn the minimization of energy (3) into its corresponding curve evolution problem. Suppose that the boundary ∂R is described by closed curve Γ parameterized by arc-length s. Calculus of variations with respect to Γ gives the following gradient flow (see for instance [3]):

$$\frac{\partial \Gamma(s)}{\partial \tau} = [\log \mathrm{p}^{\text{out}}(\Gamma(s)) - \log \mathrm{p}^{\text{in}}(\Gamma(s))]\mathbf{n}(s) + ...$$

where time index $t+1$ is dropped for simplicity, τ is the algorithmic time and $\mathbf{n}$ is the unit inward normal. Regardless of curve implementation, which may rely on parametric contours or level-sets, the curve will locally expand if color $\boldsymbol{I}_{t+1}(\Gamma(s))$ matches inner statistical features more than outer ones, and shrink in the opposite case. Estimating these statistical features over entire regions can be a drawback for tracking in real scenes, especially for the background region distribution, which may be cluttered and contain many objects. In such case, statistical color data is highly scattered, so that p^{out} may not be confident.

To overcome this limitation and obtain reliable background image data, we head towards a background model based on "narrow perception". To some extent, we adapt the philosophy of local modeling approaches [5, 4] to our tracking problem and introduce a relaxed version of the minimization problem (3). Instead of considering statistical knowledge over the entire background, we limit ourselves to the outer neighborhood around R, i.e. the following narrow band[1] of width w: $L = \{\boldsymbol{x} \in \overline{R} \mid \min_{\boldsymbol{y}\in R} \|\boldsymbol{x} - \boldsymbol{y}\| \leq w\}$ We consider that background color statistics are relevant only within L, and thus ignore available knowledge about color appearance in the "far" background $B = \overline{R} \backslash L$. Extending this principle to multiple object tracking, each object would have its own local perception of surrounding background, as shown in fig. 1. The outer neighborhood and the far background have distinct color likelihoods, so $\mathrm{p}^{\text{out}}_{t+1}(\boldsymbol{x})$ is rewritten piecewisely, depending on the membership of $\boldsymbol{x}$:

Fig. 1. Background perceived by target region (outlined with red curve) is limited within narrow band (outlined with blue curve). Other moving people do not intervene in the background representation of target

$$\mathrm{p}^{\text{out}}_{t+1}(\boldsymbol{x}) = \begin{cases} \mathrm{p}^{L}_{t+1}(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in L_{t+1} \\ \mathrm{p}^{B}_{t+1}(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in B_{t+1} \end{cases} \quad (4)$$

Trivially, $\mathrm{p}^{\text{out}}_{t+1}(\boldsymbol{x}) = 0$ if $\boldsymbol{x} \in R_{t+1}$. Color likelihood in the far background is intentionally ignored by making all colors equiprobable in B_{t+1}, independently from previous configuration $(\boldsymbol{I}_t, R_t)$:

$$\mathrm{p}^{B}_{t+1}(\boldsymbol{x}) = \mathrm{p}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t, \boldsymbol{x} \in B_{t+1}) = \frac{1}{|\mathcal{C}|}$$

where $\mathcal{C}$ is the subset of admissible colors in the chosen colorimetric space. The outer term of energy (3) is split with respect to definition (4), which gives:

$$\int_{\overline{R}_{t+1}} \log \mathrm{p}^{\text{out}}_{t+1}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{L_{t+1}} \log \mathrm{p}^{L}_{t+1}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - |B_{t+1}| \log |\mathcal{C}|$$

The energy to be minimized with respect to candidate object is finally:

$$E_{\text{LB}}[R_{t+1}] = -\int_{R_{t+1}} \log \mathrm{p}^{\text{in}}_{t+1}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \int_{L_{t+1}} \log \mathrm{p}^{L}_{t+1}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + |B_{t+1}| \log |\mathcal{C}| - \log \mathrm{p}(R_{t+1}|\boldsymbol{I}_t, R_t) \quad (5)$$

Functionals over regions are most often optimized by gradient descent applied on a level set-based reformulation, either of the Euler-Lagrange equation or of the energy itself. Instead of doing so, we minimize energy (5) with the recent variational region growing approach [6], which we embed in a greedy evolution scheme. In addition to its purely algorithmic benefits - direct evolution of a set of pixels instead of a real-valued level set function, no need for time step parameter, *ad hoc* stopping criterion - it advantageously avoids to perform calculus of variations. The greedy minimization of energy E_{LB} is summarized in algorithm 1 and holds for both implementations presented in sections 4.1 and 4.2.

[1] This is different from the so-called narrow band technique employed for level set-based segmentation

Algorithm 1 Basic greedy algorithm to minimize E_{LB}

$k := 0$; Estimate $\boldsymbol{d}^*$ if local matching is used
$R_{t+1}^{(0)} := \mathcal{T}_{\boldsymbol{d}^*}(R_t)$
repeat
  Find best candidate pixel:
  $$\boldsymbol{x}^* := \underset{\boldsymbol{x}\in R_{t+1}^{(k)}\cup\partial R_{t+1}^{(k)}}{\text{argmin}} \begin{cases} E_{\text{LB}}[R_{t+1}^{(k)}\backslash\{\boldsymbol{x}\}] & \text{if } \boldsymbol{x}\in R_{t+1}^{(k)} \\ E_{\text{LB}}[R_{t+1}^{(k)}\cup\{\boldsymbol{x}\}] & \text{if } \boldsymbol{x}\notin R_{t+1}^{(k)} \end{cases}$$
  Add or remove $\boldsymbol{x}^*$ if E_{LB} decreases:
  if $\boldsymbol{x}^*\in R_{t+1}^{(k)}$ and $E_{\text{LB}}[R_{t+1}^{(k)}\backslash\{\boldsymbol{x}^*\}] < E_{\text{LB}}[R_{t+1}^{(k)}]$
    $R_{t+1}^{(k+1)} := R_{t+1}^{(k)}\backslash\{\boldsymbol{x}^*\}$
  else if $\boldsymbol{x}^*\notin R_{t+1}^{(k)}$ and $E_{\text{LB}}[R_{t+1}^{(k)}\cup\{\boldsymbol{x}^*\}]<E_{\text{LB}}[R_{t+1}^{(k)}]$
    $R_{t+1}^{(k+1)} := R_{t+1}^{(k)}\cup\{\boldsymbol{x}^*\}$
  else $R_{t+1}^{(k+1)} := R_{t+1}^{(k)}$
  endif
  $k := k+1$
until $R_{t+1}^{(k)} = R_{t+1}^{(k-1)}$

4. ESTIMATION OF PROBABILITY FUNCTIONS

Until now, our probabilistic model has been described in a general manner, which is suitable for any likelihood functions $\text{p}_{t+1}^{\text{in}}$ and p_{t+1}^{L} as well as prior probability $\text{p}(R_{t+1}|\boldsymbol{I}_t, R_t)$. The choice of likelihoods depends on the assumptions made about temporal consistency of color, whereas prior probability depends on constraints shape and motion of the tracked object. We actually provide two possible examples of implementation. The first one is based on kernel estimates of global color Probability Density Functions (PDFs) and a simple non-temporal smoothness prior, while the second one relies on a local matching approach and motion prior.

4.1. Global kernel-based estimation

The kernel-based estimation of PDFs is global in the extent that a single distribution is used to describe color statistics in an entire region. In image segmentation, this principle leads for instance to the maximization of histogram entropy [8] or discrepancy between object and background histograms [9]. In our tracking application, likelihood functions are estimated as follows:

$$\text{p}_{t+1}^{\text{in}}(\boldsymbol{x}) = \frac{1}{|R_t|}\int_{R_t} K_\sigma(\boldsymbol{I}_{t+1}(\boldsymbol{x}) - \boldsymbol{I}_t(\boldsymbol{y}))\text{d}\boldsymbol{y}$$

and similarly for $\text{p}_{t+1}^{L}(\boldsymbol{x})$ over band L_t. Kernel K_σ is a zero-mean isotropic Gaussian with standard deviation σ. Estimating color PDFs in this way may be simply thought of as computing "smoothed" normalized color histograms within regions. To some extent, this instantiation is a "time-consistent" counterpart of the histogram-based segmentation model of [8], since pixels are assigned to object or background regarding the statistics to which they best match. We consider that no prior knowledge regarding shape or motion is available. It is thus relevant to consider the length of object boundary as a regularizer:

$$-\log \text{p}(R_{t+1}|\boldsymbol{I}_t, R_t) = \omega\,|\partial R_{t+1}|$$

where ω controls the significance of the smoothness term.

4.2. Local matching

In many cases, global modeling of color statistics is not sufficient to guarentee discrimination between object and background pixels. It is then useful to consider a local matching approach, relying on the fact that the non-rigid transformation from $\boldsymbol{I}_t$ to $\boldsymbol{I}_{t+1}$ can be expressed with motion field $\boldsymbol{v}$ and an additive white Gaussian noise $\boldsymbol{b}$:

$$\boldsymbol{I}_{t+1}(\boldsymbol{x}+\boldsymbol{v}(\boldsymbol{x})) = \boldsymbol{I}_t(\boldsymbol{x}) + \boldsymbol{b}(\boldsymbol{x}) \tag{6}$$

where $\boldsymbol{b}\sim\mathcal{N}(\boldsymbol{0};\sigma_{\boldsymbol{b}}^2)$. Local matching often requires motion field $\boldsymbol{v}$ to be estimated. In our region tracking framework, it is possible to rely on a simpler global motion descriptor instead of a dense motion field. A simple prior consists in assuming that the transformation of R_t into R_{t+1} is made up of translation $\boldsymbol{d}^*$ and a non-translational component (possibly including rotation, scaling and non-rigid deformation). There is no such prior on background motion, but it is reasonable to describe the variability of motion vectors with Gaussian distributions for both object and background:

$$\boldsymbol{v}(\boldsymbol{x}) \sim \begin{cases} \mathcal{N}(\boldsymbol{d}^*;\sigma_{\text{in}}^2) & \text{if } \boldsymbol{x}\in R_{t+1} \\ \mathcal{N}(\boldsymbol{0};\sigma_{\text{out}}^2) & \text{otherwise} \end{cases} \tag{7}$$

We expect background motion not be oriented towards a particular direction and to exhibit a relatively high variability, at least greater than motions within a single object. Thus, in practice, we set $\sigma_{\text{out}}^2 > \sigma_{\text{in}}^2$ (conversely, σ_{out}^2 may be set close to 0 in case of a nearly static background). The prior term can be expressed using translational component $\boldsymbol{d}^*$:

$$-\log \text{p}(R_{t+1}|\boldsymbol{I}_t, R_t) = \omega_1\,|\partial R_{t+1}| + \omega_2 D(R_{t+1}, \mathcal{T}_{\boldsymbol{d}^*}(R_t))$$

where $\mathcal{T}_{\boldsymbol{d}}$ denotes transformation with respect to translation $\boldsymbol{d}$. The oriented measure $D(A,B)$ is the total distance from points in A to closest points in B,

$$D(A,B) = \int_A \inf_{\boldsymbol{b}\in B} \|\boldsymbol{a}-\boldsymbol{b}\|^2\,\text{d}\boldsymbol{a}$$

This constrains R_{t+1} to be in the vicinity of a displaced counterpart of R_t. Color likelihoods are inferred from definitions in eqs (6) and (7). At a given location $\boldsymbol{x}$, probability $\text{p}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t, R_{t+1})$ is expanded by marginalizing $\text{p}(\boldsymbol{I}_{t+1}(\boldsymbol{x})|\boldsymbol{I}_t, R_t, R_{t+1}, \boldsymbol{y})$, integrating over all transformations relating $(\boldsymbol{x}, t{+}1)$ to possible predecessors $(\boldsymbol{y}, t)$, such that $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{v}(\boldsymbol{y})$. As in [3], we make the major simplifying assumption that the marginal probability is concentrated on a small set of transformations. Neglecting any normalizing constant and denominators of Gaussian probabilities, which are independent from $\boldsymbol{x}$ and $\boldsymbol{y}$, we write:

$$\text{p}_{t+1}^{\text{in}}(\boldsymbol{x}) \approx \sup_{\boldsymbol{y}\in R_t}\Big\{\exp\left(-\|\boldsymbol{I}_{t+1}(\boldsymbol{x}) - \boldsymbol{I}_t(\boldsymbol{y})\|^2\,(2\sigma_{\boldsymbol{b}}^2)^{-1}\right) \cdot \exp\left(-\|(\boldsymbol{x}-\boldsymbol{y}) - \boldsymbol{d}^*\|^2\,(2\sigma_{\text{in}}^2)^{-1}\right)\Big\}$$

$$\text{p}_{t+1}^{L}(\boldsymbol{x}) \approx \sup_{\boldsymbol{y}\in R_t}\Big\{\exp\left(-\|\boldsymbol{I}_{t+1}(\boldsymbol{x}) - \boldsymbol{I}_t(\boldsymbol{y})\|^2\,(2\sigma_{\boldsymbol{b}}^2)^{-1}\right) \cdot \exp\left(-\|\boldsymbol{x}-\boldsymbol{y}\|^2\,(2\sigma_{\text{out}}^2)^{-1}\right)\Big\}$$

Fig. 2. Global histogram-based probability estimation applied on the *tenniswoman* sequence: "entire background" energy minimization reveals some inaccuracies (blue, top row) and narrow perception improves segmentation (red, bottom row)

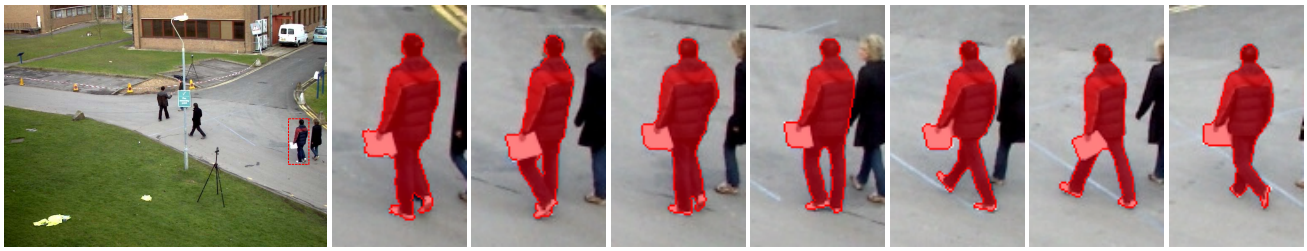

Fig. 3. Local matching-based probability estimation applied on a PETS 2009 crowd sequence (with narrow perception)

Hence, at tested location $\boldsymbol{x}$, we seek for the most probable predecessor $\boldsymbol{y}$ that $\boldsymbol{x}$ would have in case it belonged to R_{t+1} or L_{t+1}, respectively. We actually estimate the negative log-likelihoods as follows:

$$-\log \mathrm{p}^{\mathrm{in}}_{t+1}(\boldsymbol{x}) \approx \inf_{\boldsymbol{y}\in R_t} \left\{ \|\boldsymbol{I}_{t+1}(\boldsymbol{x}) - \boldsymbol{I}_t(\boldsymbol{y})\|^2 + \lambda^{\mathrm{in}} \|(\boldsymbol{x}-\boldsymbol{y}) - \boldsymbol{d}^*\|^2 \right\}$$

$$-\log \mathrm{p}^{L}_{t+1}(\boldsymbol{x}) \approx \inf_{\boldsymbol{y}\in \overline{R}_t} \left\{ \|\boldsymbol{I}_{t+1}(\boldsymbol{x}) - \boldsymbol{I}_t(\boldsymbol{y})\|^2 + \lambda^{\mathrm{out}} \|\boldsymbol{x}-\boldsymbol{y}\|^2 \right\}$$

with $\lambda^{\mathrm{in}} = \sigma_{\boldsymbol{b}}^2/\sigma_{\mathrm{in}}^2$ and $\lambda^{\mathrm{out}} = \sigma_{\boldsymbol{b}}^2/\sigma_{\mathrm{out}}^2$.

5. EXPERIMENTS AND DISCUSSION

We give an overview of a few tracking results to demonstrate the benefits of our narrow background energy (5). Firstly, we report a test with the histogram-based probability estimation. For comparison purpose, we also apply this probability estimation on the classical "entire background" energy of eq. (3). Color likelihoods are estimated using $64 \times 64 \times 64$-bin histograms, corresponding to downsampled RGB color space subsequently smoothed at scale $\sigma = 0.75$. Average processing time for a single frame is $0.8s$ for a C++ implementation running on an 2.4GHz Intel Core2 Duo. As shown in the *tenniswoman* sequence depicted in fig. 2, the classical approach may lead to several inaccuracies, whereas considering a narrow perception of background significantly improves fitting to actual target boundaries. Note that the propagation to neighboring parts with similar colors (from leg to hand, for instance) is inherent to the probability model.

The local matching approach was applied to pedestrian tracking within static camera configuration, on videos taken from the PETS 2009 benchmark database[2], as shown in fig. 3. Average processing time for a single frame is $2.5s$. Local color statistics over object and narrow band surrounding background, combined with the motion prior, manage well to recover target boundaries. Notice that of the proximity of the non-tracked pedestrian does not disrupt the tracking of the target, whether it is partially included into the narrow band of the target or not.

[2] http://www.cvg.rdg.ac.uk/PETS2009

6. REFERENCES

[1] T. Chan and L. Vese, "Active contours without edges," *IEEE TIP*, vol. 10, no. 2, pp. 266–277, 2001.

[2] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for motion estimation and tracking," *CVIU*, vol. 97, no. 3, pp. 259–282, 2005.

[3] A.R. Mansouri, "Region tracking via level set PDEs without motion computation," *IEEE TPAMI*, vol. 24, no. 7, pp. 947–961, 2002.

[4] T. Brox and D. Cremers, "On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional," *IJCV*, vol. 84, no. 2, pp. 184–193, 2009.

[5] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE TIP*, vol. 17, no. 11, pp. 2029–2039, 2008.

[6] J-L. Rose, C. Revol-Muller, T. Grenier, and C. Odet, "Unifying variational approach and region growing segmentation," in *EUSIPCO*, Aalborg, Denmark, 2010.

[7] S. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, Bayes/MDL for multiband image segmentation," *IEEE TPAMI*, vol. 18, no. 9, pp. 884–900, 1996.

[8] J. Kim, J.W. Fisher, A. Yezzi, M. Çetin, and A.S. Willsky, "A nonparametric statistical method for image segmentation using information theory and curve evolution," *IEEE TIP*, vol. 14, no. 10, pp. 1486–1502, 2005.

[9] O. Michailovich, Y. Rathi, and A. Tannenbaum, "Image segmentation using active contours driven by the Bhattacharyya gradient flow," *IEEE TIP*, vol. 16, no. 11, pp. 2787–2801, 2007.