A Probabilistic Method for Motion Pattern Segmentation

Daniel Weiler, Volker Willert, Julian Eggert, Edgar Körner

2007

Preprint:

This is an accepted article published in Proceedings of the 2007 International Joint Conference on Neural Networks (IJCNN). The final authenticated version is available online at: https://doi.org/[DOI not available]

A Probabilistic Method for Motion Pattern Segmentation

Daniel Weiler, Volker Willert, Julian Eggert, Edgar Körner

***Abstract*— In this paper we present an approach for probabilistic motion pattern segmentation. We combine level-set methods for image segmentation with motion estimations based on probability distribution functions (pdf's) calculated at each image position. To this end, we extend a region based level-set framework to exploit the motion pdf's. We then compare segmentation results of the pdf-based with those of optical-flow-based motion segmentation approaches. We found that the straightforward way of characterizing the segmented region by spatially averaging the motion measurement pdf's does not yield satisfactory results. However, describing the spatial characteristics of the motion pdf's with nonparametric density estimates enables to solve complex motion segmentation problems. In particular for situations with demanding motion patterns like partly overlapping objects and transparent motion, we show that the probabilistic approach yields better results. This confirms the idea that for motion processing it is beneficial to consistently retain the uncertainty and ambiguity of the measurement process right up to the final integration stage, instead of directly processing optical flow vectors.**

I. Introduction

In the field of image segmentation, two major approaches can be distinguished: *multi region segmentation* and *figure-background segregation*. While the former tries to group similar (by their image features **f**) and related (by their spatial properties like location, etc.) pixels of an image into separate regions, the latter attempts to find a few salient regions of an image considering them as a foreground "figure", labeling all the reminder without any further differentiation as background. In this paper we address the problem of figure-background segregation based on motion measurements, with a special focus on the segmentation of objects that are characterized by complex motion patterns that include transparency and partial overlaps with other objects.

The segmentation occurs by means of level-set methods [1], which separate all image pixels into two disjoint regions by favoring homogeneous image properties for pixels within the same region and distinct image properties for pixels belonging to different regions. The level-set formalism describes the region properties using an energy functional that implicitly contains the region description and that has to be minimized. The formulation of the energy functional dates back to e.g. Mumford and Shah [2] and to Zhu and Yuille [3]. Later on, the functionals were reformulated and minimized using the level-set framework by e.g. [4] and [5].

Daniel Weiler is with Darmstadt University of Technology, Institute of Automatic Control, Control Theory and Robotics Lab, Darmstadt D-64283, Germany (phone: +49-6151-16-6066; fax: +49-6151-16-2507; email: Daniel.Weiler@rtr.tu-darmstadt.de).

Volker Willert, Julian Eggert and Edgar Körner are with Honda Research Institute (Europe) GmbH, Offenbach D-63073, Germany (phone: +49-69-89011-750; fax: +49-69-89011-749; email: {Julian.Eggert, Volker.Willert, Edgar.Koerner}@honda-ri.de).

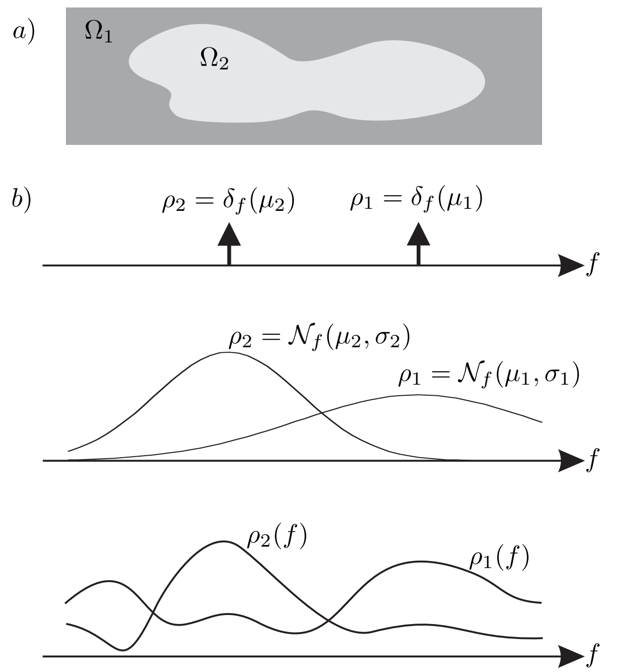

Fig. 1. a) Segmentation of an image into 2 disjoint regions Ω_1 and Ω_2 as used in level-set methods. b) Region descriptors of increasing complexity. Top row: Simple description via region averages, like the mean grey value. Middle: Description using parameterized feature distributions, like e.g. a Gaussian based on feature mean and variance. Bottom: Nonparameterized feature distributions as used in here for motion-based segmentation. f denotes the image feature.

A number of variations and extensions to the original region-based level-set methods were proposed in recent years. These went along with an increased refinement in the description of the region properties. Fig. 1 shows schematically a) a two-region separation as provided by standard level-set methods with b) different region descriptors. The early level-set formulations provided very reduced region characterizations, based e.g. on the mean values μ [4] of the entire outer and inner regions Ω_1 and Ω_2, as shown in the top row of Fig. 1b). Since this allows only a very poor representation of the properties of a spatially extended region, density functions ρ were introduced. These are commonly modeled by Gaussian approximations [6], see Fig. 1b) middle. In more complex cases, to be able to cope with multimodal distributions, nonparametric density estimates (see e.g. [7]) are used [5], see Fig. 1b), bottom row. All these region descriptors can of course be applied on scalar (e.g. pixel grey level) as well as vectorial image feature data (e.g. color, texture, motion vector gained from optical flow estimation, or even the combination of several cues, etc.).

Among all segmentation algorithms from computer vision, level-set methods provide perhaps the closest link with the biologically motivated, connectionist models as e.g. represented by [8]. Similar to neural models, level-set methods work on a grid of nodes located in image/retinotopic space, interpreting the grid as having local connectivity, and using local rules for the propagation of activity in the grid. Time is included explicitly into the model by a formulation of the dynamics of the nodes activity. Furthermore, the external influence from other sources (larger network effects, feedback from other areas, inclusion of prior knowledge) can be readily integrated on a node-per-node basis, which makes level-sets appealing for the integration into biologically motivated system frameworks.

In this paper we compare the representation of region characteristics by mean values and density functions for the special case of motion pattern segmentation and show the advantage of using nonparametric density functions which allow to represent multimodal distributions. Section II introduces the level-set method we used for image segmentation, its extension to vector-valued inputs and its coupling with motion probability distribution functions (pdf's). In section III we present the results. First, we show exemplary cases which use spatial averages of the motion pdf's but fail to find the right regions. We then extend the level-set formalism to make use of nonparametric density functions for the representation of region properties, and show how we can use this to distinguish a moving object in the presence of another moving object (with a different velocity) and a background, even in the case of object transparency. A short discussion finalizes the paper.

II. Object Segmentation Framework

A. Level-set based region segmentation

The object segmentation framework is based on a two-region level-set method [5], [9]. In a level-set framework, a level-set function $\phi \in \Omega \mapsto \mathbb{R}$ is used to divide the image plane Ω into two disjoint regions, Ω_1 and Ω_2, where $\phi(x) > 0$ if $x \in \Omega_1$ and $\phi(x) < 0$ if $x \in \Omega_2$. Here we adopt the convention that Ω_1 indicates the background and Ω_2 the segmented object. A functional of the level-set function ϕ can be formulated that incorporates the following constraints:

- The data within the two regions should be as similar as possible.
- The data between the two regions should be as dissimilar as possible.
- The length of the contour separating the two regions should be as short as possible.

This leads to the expression:

$$E(\phi) = \nu \int_\Omega |\nabla H(\phi)| dx - \sum_{i=1}^{2} \int_\Omega \chi_i(\phi) \log p_i \, dx \qquad (1)$$

with the Heaviside function $H(\phi)$ and $\chi_1 = H(\phi)$ and $\chi_2 = 1 - H(\phi)$. The first term favors small region boundary contours, whereas the second term contains assignment probabilities $p_1(x)$ and $p_2(x)$ that a pixel at position x belongs to the inner and outer regions Ω_1 and Ω_2, respectively, favoring a unique region assignment.

Minimization of this functional with respect to the level-set function ϕ using gradient descent leads to

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \log \frac{p_1}{p_2} \right] \qquad (2)$$

A region descriptor $\rho_i(\mathbf{f})$ that depends on the image feature vector $\mathbf{f}$ serves to describe the characteristic properties of the outer vs. the inner regions. The assignment probabilities $p_i(x)$ for each image position are calculated based on an image feature vector via $p_i(x) := \rho_i(\mathbf{f}(x))$. For standard images, there may be only a single feature vector component like the pixel grey values. The case with several image features can be covered by assuming independent contributions from each feature vector channel f_j using assignment probabilities $p_1 = \prod_j p_{1j}$ and $p_2 = \prod_j p_{2j}$. In many cases, the p_{ij}'s are modeled by unimodal Gaussian region descriptor distributions so that $p_{ij}(x) = \mathcal{N}_{f_j}(\mu_{ij}, \sigma_{ij})$ [6], with mean μ_{ij} and variance σ_{ij}. Furthermore, μ_{ij} and σ_{ij} may act as locally calculated parameters that depend on the pixel position x. Remark that if we assume a single μ_{ij} and σ_{ij} for the entire region, Eq. 1 reduces to the standard Mumford-Shah functional as used in [4].

There are also approaches where the distributions are approximated in a multimodal way [5] e.g. by Gaussian mixture models or nonparametric Parzen density estimates [7]. For motion-based segmentation as presented in this paper, we propose to use nonparametric region descriptor functions, i.e., representing them extensively in a grid-based way. To this end, we calculate for each feature channel j inside the region i a normalized histogram $\mathbf{h}_{ij}$,

$$h_{ijk} = \frac{\int_\Omega \chi_i(\phi) \hat{h}_{ijk}(x) dx}{\int_\Omega \chi_i(\phi) dx}$$

with

$$\hat{h}_{ijk}(x) = H(f_j(x) - b_k) - H(f_j(x) - b_{k+1})$$

with bins indexed by k and borders b_k of the histogram bin intervals. [1] Here, the histogram $\mathbf{h}_{ij}$ takes the role of the feature-dependent region descriptor $\rho_i(\mathbf{f})$. The region assignment probability is then calculated by

$$p_{ij}(x) = \hat{\mathbf{h}}_{ij}(x) \, \mathbf{h}_{ij} := \sum_k \hat{h}_{ijk}(x) h_{ijk}$$

i.e., by extracting the histogram entry of $\mathbf{h}_{ij}$ that corresponds to the bin index indicated by $f_j(x)$. In this way, both the region descriptor function as well as the computation of the region assignment become computationally inexpensive, since they amount to calculating and extracting single entries from normalized histograms.

[1] Smoothed versions of the histogram can be gained by convolving $K_h * \mathbf{h}_{ij}$, but in our applications smoothing this did not change the results substantially.

B. Probabilistic motion information

In level-set-based figure-background segregation using motion information our aim is to find the most discriminative motion representation that enables a separation of the figure from the background, e.g., that most precisely describes the motion pattern of an object. In the ideal case this includes not only the local pixel movements but also the consideration of spatial motion coherence and the existence of multiple motion hypotheses. Spatial relations are usually incorporated to constrain the object movement to some limited class, most often by assuming affine motion [10]. In this paper we take spatial relations into account but only at the segmentation level where we consider the degree of spatial occurrence of coexisting pixel-motion hypotheses to discriminate object movements from the rest.

The characteristic motion pattern of an object in an image sequence $\mathbf{I}^{1:t}$ at time t is given by the optical flow $\mathbf{V}$ within the region that shapes the object. The optical flow $\mathbf{V} = \{\mathbf{v}_x\}$ is the set of the velocity vectors $\mathbf{v}_x$ of all pixels at every location x in the image $\mathbf{I}$, meaning that the movement of each pixel is represented with one velocity hypothesis. This representation neglects the fact that in most cases the pixel movement cannot be unambiguously detected because of different kinds of motion-specific correspondence problems, like e.g. the well-known aperture problem [11], and noisy data the measurement is based on. Especially for the case of transparent moving objects that overlap or partly occlude each other several motion hypotheses are needed to fully describe the image movement within the overlapping regions.

As has been suggested and discussed by several authors [12], [13], [14], velocity pdf's are well suited to handle several kinds of motion ambiguities. Following these ideas we model the uncertainty for the optical flow $\mathbf{V}$ like the following:

$$P(\mathbf{V}|Y^t) = \prod_x P(\mathbf{v}_x|Y^t) \text{ with } Y^t = \{\mathbf{I}^t, \mathbf{I}^{t-1}\}\,, \quad (3)$$

where the probability for the optical flow $P(\mathbf{V}|Y^t)$ is composed of locally independent velocity pdf's $P(\mathbf{v}_x|Y^t)$ for every image location x. $P(\mathbf{v}_x|Y^t)$ can be calculated using several standard methods, for details refer e.g. to [13], [14]. These pdf's fully describe the motion estimations available for each position x, taking along (un)certainties and serving as a basis for the motion segmentation in the sections that follow.

C. Coupling probabilistic motion estimation with object segmentation

For the coupling of motion estimation and region based image segmentation methods, two major approaches can be found. A first approach treats motion estimation and image segmentation as a single, combined problem, which can be solved by minimizing a common functional that estimates the motion and segments the image simultaneously [10], [15]. A second approach treats motion estimation and image segmentation as two independent processes. In the second case, the motion estimation acts only as a preprocessing step for the segmentation [16].

The first approach is able to incorporate spatial considerations provided by the segmentation process into the motion estimation, therefore reducing ambiguities that occur during the motion measurement. To the contrary, in the second approach the segmentation stage has to rely on a sufficiently good motion estimation. As motivated in the previous section, optical flow estimations deteriorate under conditions where ambiguous solutions exist. Therefore, in this paper we use the full motion probability distribution function $P(\mathbf{v}_x|Y^t)$ as the input to the level-set-based segmentation stage, since this allows to carry along the uncertainties and resolve them later implicitly during the spatial integration of the segmentation process.

The straightforward approach for vector-based level-sets would be to consider the pdf from motion preprocessing as a feature vector $\mathbf{f}(x)$ at each image position. We can do this by taking into account a discrete set of possible velocity vectors $\mathbf{v}^j$, so that

$$f_j(x) = P(\mathbf{v}_x = \mathbf{v}^j|Y^t) \quad (4)$$

i.e., the j-th feature channel corresponds to the locally measured probabilities at the j-th velocity.

Let us consider that we now take the approach of a unimodal Gaussian region descriptor as introduced in Fig. 1b), middle and in section II-A. This is then equivalent to representing the level-set region using a "mean" motion pdf, gained by simple spatial averaging of the pixelwise motion pdf's over the region. Region assignment then occurs by checking if the locally measured motion pdf is similar to the mean motion pdf of a region. An alternative is to use the nonparametric region descriptor functions as introduced in section II-A. These two ways of coupling the motion pdf's to the level-set-based segmentation will be examined in the next section.

III. Main Results

Contrary to the velocity pdf's introduced in sections II-B, II-C, standard approaches for motion-based pattern segmentation that use level-set methods usually work directly on the two-dimensional optical flow vectors $\mathbf{v}_x$ [6], [9] instead of the velocity probabilities $P(\mathbf{v}_x|Y^t)$. These vectors $\mathbf{v}_x$ are then coupled with the level-set segmentation framework by considering the two vector components as segmentation features. For cases of nontransparent, translational motion and homogeneous velocity fields the segmentation methods describing the regions with mean values or on the basis of a normal distribution as discussed in the introduction (see Fig. 1b) top and middle row) produce reasonable results.

But cases of motion with high level of ambiguities lead to large errors for methods based on optical flow field estimation, since they are not able to represent more than one velocity at a specific image position. Therefore, for transparent motion, nontranslational motion (e.g. rotation) or in the presence of objects with different velocities, these methods are likely to fail. Nevertheless, using both motion pdf's and region descriptors that are able to capture the information provided by the pdf's we may be able to circumvent some of the problems of complex motion patterns.

In the following section, we use the framework introduced in section II that makes use of motion pdf's and histogram-based level-sets and show how we can do motion-based object segmentation for the case of transparency.

A. Segmentation on velocity distributions

In Fig. 2 the first example is shown, with two objects moving on a static background [2]. The upper object is moving from left to right and the lower object moves in the opposite direction. The background pattern as well as the patterns for the two objects are generated by creating random gray values with a uniform distribution. Thus no separation of objects and background is possible without motion of the objects because other features, like e.g. colour or texture, of figures and background do not provide any information that can contribute to the segmentation process.

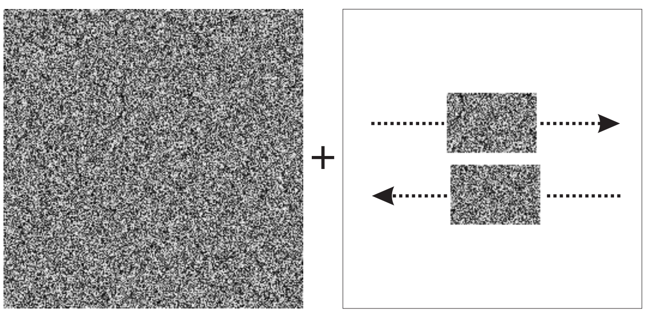

Fig. 2. Assembly of the first artificial motion sequence. The sequence is composed by a background pattern (left) superimposed with two patterns, simulating moving objects (right), that translate from left to right and vice versa.

The aim would be that only the object which is favored by the initialization of the level-set contour is cut out as figure and the object moving in the opposite direction should be assigned as part of the background.

Applying the commonly used level-set segmentation method for vector valued data proposed by [9] to the motion pdf's obtained as described in II-B leads to the results depicted in Fig. 3 middle. The method from [9] describes the regions by their mean values, as illustrated in Fig. 1b) top. In our case of motion pdf's the mean value of the features within a region corresponds to the spatially averaged pdf's within that region. The Mumford-Shah level-set functional then segments an object for which the distance of the distribution for each position inside the object to the mean distribution inside the object gets minimized (and analogous for the outside region). Since we are already working with velocity distributions one might assume that the method should be able to distinguish the two objects from another, but this is not the case. Fig. 3 middle shows that not only the object, focused by the initial level-set contour (see Fig. 3 left) is segmented, but also the second object, which moves in the opposite direction.

[2]Our method also produces competitive segmentation results on real-world examples, but in this paper we concentrate on artificially generated examples for the sake of the argument and because it is easier to show transparency effects.

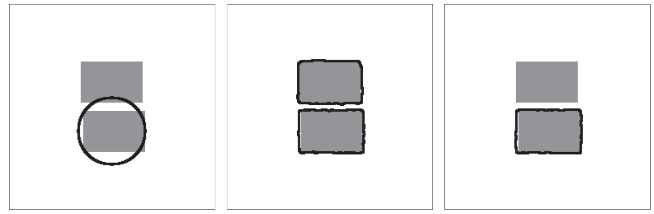

Fig. 3. Level-set contours (black) delimiting the two regions Ω_1 and Ω_2 superimposed with the flow-field's ground truth data (gray), depicting the real object position. Left: Initial level-set contour for the gradient descent of Eq. 2. Middle: Final level-set contour gained with a segmentation method with mean values of features as region descriptors. Right: final level-set contour achieved with the proposed segmentation method with a region descriptor exploiting a nonparametric spatial distribution of the feature vector.

For the case of distributions modelled by their mean values only (see Fig. 1b) top) the last term of Eq. 2, $\log(p_1/p_2)$, leads to an equation that compares the Euclidean distances from $\mathbf{f}$ to the region descriptor parameters μ_1 and μ_2 for the outside and inside region, respectively. For a specific image position x_0 this leads to

$$||\mu_2 - \mathbf{f}(x_0)||^2 - ||\mu_1 - \mathbf{f}(x_0)||^2$$

which can be rewritten as

$$\sum_j \left((\mu_{2j} - f_j(x_0))^2 - (\mu_{1j} - f_j(x_0))^2 \right) = \sum_j d_j$$

with $d_j = d_j(f_j(x_0), \mu_1, \mu_2)$ giving an evaluation for each feature channel $f_j(x_0)$ concerning its assignment to the inside or outside region, i.e., the sign of d_j describes the assignment to the corresponding region and the value its discrimination power.

As it can be seen in Fig. 3 middle, the method is not able to segment only one of the moving objects, i.e. to discriminate between the two moving objects, even if they exhibit motion in contrary directions. The reason for that becomes clear when evaluating the d_j's in Fig. 4 bottom. The motion estimation pdf is represented by 25 supporting points, that span a discrete velocity vector space of 5×5 velocities, where the origin, the zero-velocity, is located in the middle. The features 11, 13, and 14 are of particular interest, as they correspond to the velocity of the upper object, the background and the lower object, respectively. All other features correspond to velocity measurements with only minor impact on the segmentation result. Evaluating the d_j's in Fig. 4 bottom, feature 13 (background) holds the most discriminative power, which is even larger than the joint discriminative power of features 11 (velocity of the object the image location x_0 belongs to) and 14 (velocity of the other, i.e., the favored object). Thus the background, which in fact is the largest "object" in the scene dominates the solution and supresses the differences between the two objects.

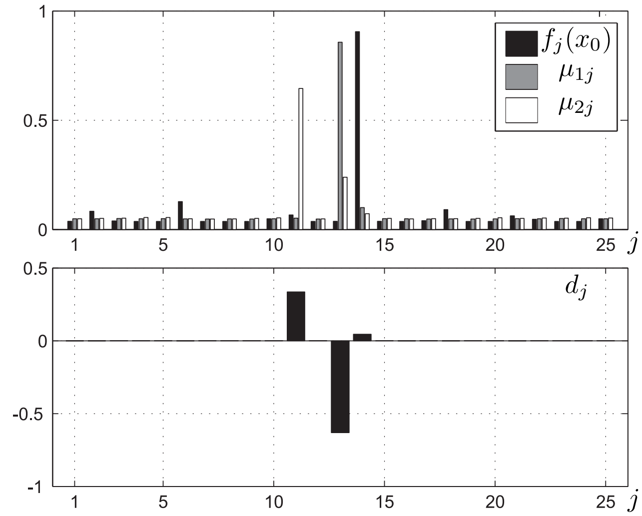

Fig. 4. Analysis of the level-set method using mean values as region descriptors for a specific image position x_0 that belongs to the upper object in the first test sequence, which is wrongly segmented. Top: Region descriptors, i.e. mean values of the motion pdf's for the regions Ω_1 (gray) and Ω_2 (white) and the motion pdf at the image position x_0 (black). Bottom: Contributions $d_j(f_j(x_0), \mu_{1j}, \mu_{2j})$ for each element of the feature vector to the overall gradient descent equation. Values taken at the beginning of the level-set relaxation. See text for the detailed explanation why this leads to a wrong solution.

To overcome these problems we use nonparameterized distribution based region descriptors (see Fig. 1b) bottom). In our case with motion estimation pdf's as features (see Eq. 4) this leads to a two dimensional distribution over discrete velocities and the values of velocity probabilities $P(\mathbf{v}_x|Y^t)$ (see Fig. 5 top). Applying the proposed method to the motion sequence from Fig. 2 succeeds in completing the segmentation of the object selected by the initial level-set contour (see Fig. 3 left) not only from the background but also from the second moving object (see Fig. 3 right).

For the case of region descriptors based on a nonparameterized distribution (see. Fig. 1b) bottom) the last term of Eq. 2, $\log(p_1/p_2)$ leads to

$$\sum_j (\log p_{1j} - \log p_{2j}) = \sum_j D_j$$

where $D_j = D_j(f_j(x_0), \rho_1, \rho_2)$ gives an evaluation for each feature $f_j(x_0)$ concerning its assignment to the inside or outside region as for the previous case.

Evaluating again the D_j's (see Fig. 5 bottom right) clarifies the advantage of using a multimodal distribution of the features instead of the mean feature vector as region descriptors. Looking at features 11, 13, and 14 (velocity of upper object, background and lower object, respectively) and their contributions to the overall gradient descent Eq. 2 shows that the discriminative power of the opposed velocities of the two objects exceeds the discriminative power of the background.

The reason for the higher discriminative power is that with the two-dimensional region descriptors we are better able to assign pixel features to single peaks of a multimodal distribution.

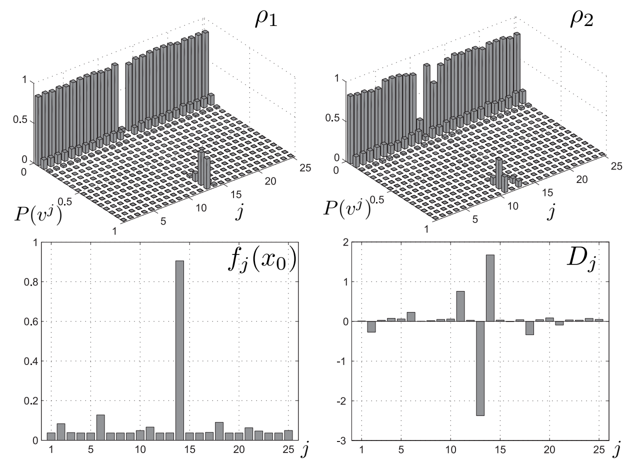

Fig. 5. Analysis of the level-set method using nonparameterized distribution based region descriptors for a specific image position x_0 that belongs to the upper object in the first test sequence, which now gets segmented as expected. Top row: Two-dimensional region descriptor for region Ω_1 (left) and Ω_2 (right). Bottom: The velocity distribution at a specific image position x_0 (left) and the contributions $D_j(f_j(x_0), \rho_1, \rho_2)$ for each element of the feature vector to the overall gradient descent from Eq. 2. Values again taken from the beginning of the level-set relaxation.

B. Segmentation on velocity distributions with transparent motion

To show the advantages of the proposed method a second example was created. This time we incorporated transparent motion, which leads to multimodal distributions. Again, two objects are moving on a static background. The objects and the background are as in the first example created by uniformly distributed random gray values. In the artificial sequence one object moves from top to bottom and the other one from left to right, overlapping each other (see Fig. 6). The transparency of the objects is achieved by a zero-mean additive superposition of the objects and the background, leading again to an image where the objects can be best identified by motion cue.

For the artificial image sequence the level-set function was initialized with a "signed distance" circle, covering half

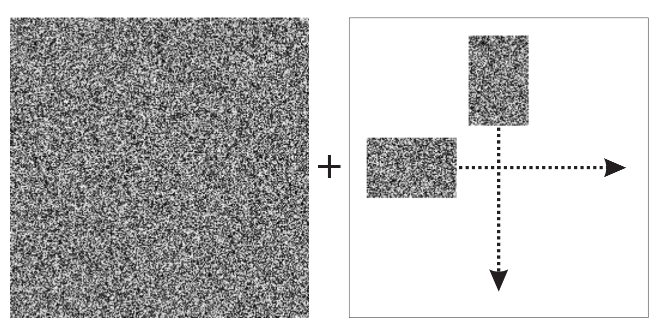

Fig. 6. Assembly of the second artificial motion sequence. The sequence is composed by a background pattern (left) superimposed with two transparent patterns, simulating transparent moving objects translating from left to right and from top to bottom (right), overlapping each other.

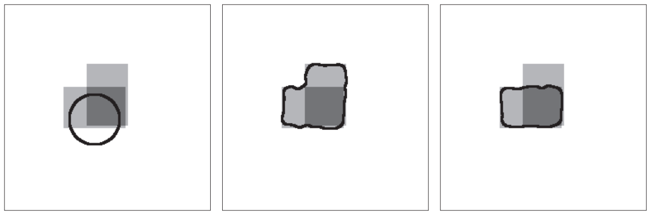

Fig. 7. Final level-set contour for region based level-set methods working on transparent motion. Left: Initial level-set contour for the gradient descent of Eq. 2. Middle: Assuming unimodally distributed data. Right: Using the pdf of the motion estimates instead of the velocity field and assuming multimodally distributed data.

the object to segment and half the background. In Fig. 7 we show the segmentation result for the same methods as compared in the previous section. Fig. 7 middle shows the result obtained when assuming unimodally distributed data, where the method is not able to distinguish the two objects from each other and only a separation of the moving objects from the background is achieved. On Fig. 7 right one can see that the method that assumes multimodally distributed data and uses the pdf of the motion estimates instead of the velocity field, is able to distinguish the favored object both from the background and the other object, even under the condition of transparent overlap.

IV. Conclusions

We have presented an approach for motion pattern segmentation as a two-stage model that makes use of velocity probability distribution functions as a first step and then incorporates the motion pdfs into a level-set-based segmentation framework that uses nonparametric density estimates as region descriptors.

Contrary to standard motion segmentation approaches which directly use velocity *vectors* $\mathbf{v}_x$, the incorporation of the *full motion pdfs* allows to handle situations which require the representation of motion uncertainty resp. ambiguity. With artificial test sequences we have shown that for these cases the nonparametric density estimation for the description of the segmented region becomes essential. We also obtained good results for segmentation problems from real-world scenes, nevertheless the question to what extent the pdf and histogram based processing of motion information provides an advantage in real-world conditions still has to be evaluated.

In all examples, a single two-phase level-set method was used for *each* object, which allows for the separation of a single object from a background. For the simultaneously segmentation of several objects, multiphase level-sets are often used. This was not an option here, since they do not allow an overlapping of the objects as needed for the transparent case.

Further work will include the evaluation of the method for real-world scenes and the integration of additional segmentation cues.

References

[1] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, vol. 79, pp. 12–49, 1988.

[2] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math*, vol. 42, pp. 577–685, 1989.

[3] S. C. Zhu and A. L. Yuille, "Region competition: Unifying snakes, region growing, and bayes/MDL for multiband image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 9, pp. 884–900, 1996.

[4] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[5] J. Kim, J. W. Fisher, A. J. Yezzi, M. Çetin, and A. S. Willsky, "Nonparametric methods for image segmentation using information theory and curve evolution," in *International Conference on Image Processing, Rochester, New York*, vol. 3, Sept. 2002, pp. 797–800.

[6] M. Rousson and R. Deriche, "A variational framework for active and adaptative segmentation of vector valued images," *IEEE Workshop on Motion and Video Computing, Orlando, Florida*, Dec. 2002.

[7] E. Parzen, "On the estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, 1962.

[8] Grossberg, Stephen, Hong, and Simon, "A neural model of surface perception: Lightness, anchoring, and filling-in," *Spatial Vision*, vol. 19, no. 2-4, pp. 263–321, 2006.

[9] T. Chan, B. Sandberg, and L. Vese, "Active contours without edges for vector-valued images," *J. Visual Communication Image Representation*, vol. 11, no. 2, pp. 130–141, 2000.

[10] T. Brox, A. Bruhn, and J. Weickert, "Variational motion segmentation with level sets," in *European Conference on Computer Vision (ECCV)*, ser. LNCS, A. Leonardis, H. Bischof, and A. Pinz, Eds., vol. 3951. Graz, Austria: Springer, May 2006, pp. 471–483.

[11] S. S. Beauchemin and J. L. Barron, "The computation of optical flow," *ACM Comp. Surv.*, vol. 27, no. 3, pp. 433–467, 1995.

[12] E. P. Simoncelli, E. H. Adelson, and D. J. Heeger, "Probability distributions of optical flow," in *Proc Conf on Computer Vision and Pattern Recognition*. Mauii, Hawaii: IEEE Computer Vision and Pattern Recognition (CVPR), 1991, pp. 310–315.

[13] V. Willert, J. Eggert, J. Adamy, and E. Körner, "Non-gaussian velocity distributions integrated over space, time, and scales," in *IEEE Trans. Syst., Man, Cybern. B*, vol. 36, no. 3, June 2006, pp. 482–493.

[14] Y. Weiss and D. Fleet, "Velocity likelihoods in biological and machine vision," in *Probabilistic Models of the Brain: Perception and Neural Function*. MIT Press, 2002, pp. 77–96.

[15] D. Cremers, "A multiphase level set framework for motion segmentation," in *4th Int. Conf. on Scale Space Theories in Computer Vision*, L. D. Grififn and M. Lillholm, Eds., vol. 2695. LNCS c Springer, June 2003, pp. 599–614.

[16] T. Brox, M. Rousson, R. Deriche, and J. Weickert, "Unsupervised segmentation incorporating colour, texture, and motion," *Computer Analysis of Images and Patterns*, vol. 2756, pp. 353–360, 2003.