In Pattern Recognition, Springer LNCS 3175, C.-E. Rasmussen, H. Bülthoff, M. Giese, and B. Schölkopf (Eds.), pp. 415-423, Tübingen, Germany, Aug. 2004

Level Set Based Image Segmentation with Multiple Regions ⋆

Thomas Brox and Joachim Weickert

Mathematical Image Analysis Group, Faculty of Mathematics and Computer Science,
Saarland University, Building 27, 66041 Saarbrücken, Germany
{brox,weickert}@mia.uni-saarland.de
www.mia.uni-saarland.de

Abstract. We address the difficulty of image segmentation methods based on the popular level set framework to handle an arbitrary number of regions. While in the literature some level set techniques are available that can at least deal with a fixed amount of regions greater than two, there is very few work on how to optimise the segmentation also with regard to the number of regions. Based on a variational model, we propose a minimisation strategy that robustly optimises the energy in a level set framework, including the number of regions. Our evaluation shows that very good segmentations are found even in difficult situations.

1 Introduction

Image segmentation has a long tradition as one of the fundamental problems in computer vision. Relatively early, the problem has been formalised by Mumford and Shah as the minimisation of an energy functional that penalises deviations from smoothness within regions and the length of their boundaries [13]. Later, Zhu and Yuille found out that this formulation is closely related to the *minimum description length* criterion and the *maximum a-posteriori* criterion [22]. They presented a new energy functional that unified many of the existing approaches on image segmentation. It can be interpreted as the joint minimisation of the boundary length (as in the Mumford-Shah functional) and the Bayes error in the regions' interior. This is based on the fact that segmentation is actually a clustering problem with a neighbourhood constraint. Since penalising the Bayes error is optimal from the statistical point of view, the variational formulation of Zhu-Yuille describes the segmentation problem very accurately.
However, a tricky issue on image segmentation is the representation of regions and their boundaries. Although there exist neat energy functionals like the one of Mumford-Shah or that of Zhu-Yuille, it is not easy to minimise them in practice. A very nice tool to deal with this problem appeared with the introduction of level sets [8, 14]. One application to image segmentation has been the active contour model [3, 4, 10], which is completely edge based, and therefore a rather local approach to image segmentation. Level set based segmentation that takes the region information into account has been proposed later in [15] and [5]. Using level sets for image segmentation has many advantages. First of all, level sets yield a nice representation of regions and their boundaries

⋆ We gratefully acknowledge partial funding by the *Deutsche Forschungsgemeinschaft (DFG)* and many interesting discussions with Mikaël Rousson from INRIA Sophia-Antipolis.

on the pixel grid without the need of complex data structures. This considerably simplifies optimisation, as variational methods and standard numerics can be employed. Furthermore, level sets can describe topological changes in the segmentation, i.e. parts of a region can split and merge. Finally, the possibility to describe the image segmentation problem with a variational model increases the flexibility of the model and allows to employ, for instance, additional features [1], shape knowledge [11, 7], or joint motion estimation and segmentation [6].

The main problem of the level set representation lies in the fact that a level set function is restricted to the separation of two regions. As soon as more than two regions are considered, the level set idea looses parts of its attractiveness. This is why only a few papers focus on level set based segmentation in the case of more than two regions. In [21], a level set function is assigned to each region. This framework has been adapted to classification in [18]. In another approach, the bi-modal case is extended to tri-modal segmentation [20]. Both techniques, however, assume an initially fixed number of regions. This assumption is omitted in [16] where the number of regions is estimated in a preliminary stage by means of a Gaussian mixture estimate of the image histogram. This way, the number of mixture coefficients determines the number of regions. However, this kind of estimation is only loosely connected to the energy functional that is minimised. A considerably different approach is proposed in [19]. Here, the level set functions are used in such a way that N regions are represented by only $\log_2 N$ level set functions. Unfortunately, this will result in empty regions, if less than N regions are present in the image. These empty regions have undefined statistics, though the statistics still appear in the evolution equations.

Altogether, the prominence of level set based segmentation is yet lost as soon as more than two regions come into play, and other segmentation methods based for instance on algebraic multigrid [9] often perform better. The purpose of this paper is to solve the remaining problem of the level set framework while saving its advantages.

We show a way how to minimise the energy of Zhu-Yuille by means of level sets. This includes also the minimisation with regard to the number of regions. As the objective function can be assumed to have plenty of local minima, we employ multi-scale ideas and a divide-and-conquer strategy. The most precarious part of the segmentation, namely the determination of the number of regions as well as the initialisation of the level set functions, is based on the very robust two-region segmentation which splits a domain into two parts in a way that is optimal according to the energy (Section 2). The multi-phase level set evolution has then just to adapt the regions in the global scope with more than two regions present (Section 3). With this minimisation strategy the level set framework can be fully exploited, what leads to excellent segmentation results. This will be demonstrated in some experiments in Section 4.

2 Two-region segmentation

Contrary to the general segmentation problem, two-region segmentation by means of a level set framework is well understood. Consider the Bayes error, i.e. the probability of misclassified pixels

$$L = 1 - \int_{\Omega_1} p_1 P_1 \, dx - \int_{\Omega_2} p_2 P_2 \, dx \tag{1}$$

with the probability densities $p_1 = p(x|\Omega_1)$ and $p_2 = p(x|\Omega_2)$ of the regions Ω_1 and Ω_2, and under the side conditions $\Omega = \Omega_1 \cup \Omega_2$ and $\Omega_1 \cap \Omega_2 = \emptyset$, i.e. the regions cover the whole image domain Ω and do not overlap. The a-priori probabilities of both regions are equal, so $P_1 = P_2 = 0.5$. Moreover, instead of minimising the Bayes error directly, it is beneficial from the numerical point of view to work on the logarithms. Together with a penalty on the length of the boundary Γ, weighted by the parameter ν, this leads to the energy functional

$$E(\Omega_1, \Omega_2, p_1, p_2) = -\int_{\Omega_1} \log p_1 \, dx - \int_{\Omega_2} \log p_2 \, dx + \nu \int_{\Gamma} ds. \qquad (2)$$

For minimising this energy, now a level set function is introduced. Let $\Phi : \Omega \to \mathbb{R}$ be the level set function with $\Phi(x) > 0$ if $x \in \Omega_1$ and $\Phi(x) < 0$ if $x \in \Omega_2$. The zero-level line of Φ is the searched boundary between the two regions. We also introduce the regularised Heaviside function $H(s)$ with $\lim_{s\to-\infty} H(s) = 0$, $\lim_{s\to\infty} H(s) = 1$, and $H(0) = 0.5$. This allows to rewrite Eq. 2 as

$$E(\Phi, p_1, p_2) = -\int_{\Omega} H(\Phi) \log p_1 + (1 - H(\Phi)) \log p_2 - \nu |\nabla H(\Phi)| \, dx. \qquad (3)$$

The minimisation with respect to the regions can now be performed according to the gradient descent equation

$$\partial_t \Phi = H'(\Phi) \left(\log \frac{p_1}{p_2} + \nu \, \mathrm{div} \left(\frac{\nabla \Phi}{|\nabla \Phi|} \right) \right) \qquad (4)$$

where $H'(s)$ is the derivative of $H(s)$ with respect to its argument. Note that the side conditions are automatically satisfied due to the level set representation.

However, the probability densities p_1 and p_2 still have to be estimated. This is done according to the *expectation-maximisation principle*. Having the level set function initialised with some partitioning, the probability densities can be computed by a nonparametric Parzen density estimate using the smoothed histogram of the regions. Then the new densities are used for the level set evolution, leading to a further update of the probability densities, and so on. This iterative process converges to the next local minimum, so the initialisation matters.

In order to attenuate this dependency on the initialisation, two measures are recommendable. Firstly, the initialisation should be far from a possible segmentation of the image, as this enforces the search for a minimum in a more global scope. We always use an initialisation with many small rectangles scattered across the image domain.

The second measure is the application of a coarse-to-fine strategy. Starting with a downsampled image, there are less local minima, so the segmentation is more robust. The resulting segmentation can then be used as initialisation for a finer scale, until the original optimisation problem is solved.

Under the assumption of exactly two regions in the image, this framework works very well. For some nice results obtained with this method we refer to [17, 1]. The only remaining problem is the fact, that the assumption of exactly two regions in an image is mostly not true.

3 Multiple Region Segmentation

For the before-mentioned reasons, the generalised version of the segmentation problem with an arbitrary number of regions N will now be considered. The general model is described by the energy of Zhu-Yuille [22]

$$E(\Omega_i, p_i, N) = \sum_{i=1}^{N} \left(-\int_{\Omega_i} \log p_i \, dx + \frac{\nu}{2} \int_{\Gamma_i} ds + \lambda \right). \tag{5}$$

The additional term of this energy functional penalises the number of regions with the parameter λ. Now also the number of regions is a free variable that has to be optimised. Moreover, this variable is discrete and the increased number of regions is very sensitive to different initialisations. Furthermore, the nice splitting into two regions by a single level set function as described in the last section is not applicable anymore.

Reduced problem with N regions. In order to cope with all these additional difficulties, the complexity of the problem is first reduced by setting N fixed and assuming that a reasonable initialisation of the regions is available. In this case it is possible to introduce again a level set based energy functional with a set of level set functions Φ_i, each representing one region as $\Phi_i(x) > 0$ if and only if $x \in \Omega_i$.

$$E(\Phi_i, p_i) = \sum_{i=1}^{N} \left(-\int_{\Omega} H(\Phi_i) \log p_i - \frac{\nu}{2} |\nabla H(\Phi_i)| \, dx \right) \tag{6}$$

Note that, in contrast to the two-region case, this formulation does not implicitely respect the side condition of disjoint regions anymore. Minimising the energy according to the expectation-maximisation principle and the following evolution equations

$$\partial_t \Phi_i = H'(\Phi_i) \left(\log p_i - \max_{j \neq i,\, H(\Phi_j) > 0} \left(\log p_j \right) + \frac{\nu}{2} \operatorname{div} \left(\frac{\nabla \Phi_i}{|\nabla \Phi_i|} \right) \right). \tag{7}$$

ensures the adherence to the side conditions at least for the statistical part, since the maximum a-posteriori criterion ensures that a pixel is assigned uniquely to the region with the maximum a-posteriori probability. The smoothness assumption, however, can result in slight overlapping of regions close to their boundaries, like in all existing level set based methods dealing with an arbitrary number of regions, beside [19]. If this is not wanted in the final result, the pixels of such overlapping areas can be assigned to the region, where the level set function attains its maximum value.

So up to this point we can handle the following two cases:

- A domain of the image can be split into two parts by the two-region segmentation framework.
- A set of regions can evolve, minimising the energy in Eq. 5, if the number of regions is fixed and reasonable initialisations for the regions are available.

Solving the general problem. By means of these two special cases, also the general problem according to the model in Eq. 5 can be solved. Starting with the whole image domain Ω being a single region, the two-region segmentation can be applied in order to find the best splitting of the domain. If the energy decreases by the splitting, this results in two regions. On these regions, again the two-region splitting can be applied, and so on, until the energy does not decrease by further splits anymore. With this proceeding, not only the optimum number of regions is determined, but also suitable initialisations for the regions. Of course, the resulting partitioning is not optimal yet, as for the two-region splitting, possibilities of a region to evolve have been ignored. However, as the region number and the initialisation are known, the energy can now be minimised in the global scope by applying the evolution of Eq. 7, adapting the regions to the new situation where they have more competitors.

This procedure is applied in a multi-scale setting. Starting the procedure as described on the coarsest scale, with every refinement step on the next finer scale, it is checked whether any further splitting or merging decreases the energy before the evolution according to Eq. 7 is applied. So for each scale the optimum N is updated, as well as the region boundaries and the region statistics.

Though a global optimum still cannot be guaranteed[1], this kind of minimisation avoids quite reliably to be trapped by far-away local minima, as it applies both a coarse-to-fine strategy and the divide-and-conquer principle. The two-region splitting completely ignores the cluttering rest of the image. This consistently addresses the problems of optimising the discrete variable N and of not knowing good initialisations for the regions.

4 Results

We evaluated this scheme with a couple of artificial and real-world images. In order to handle texture and colour images, the features were computed and incorporated as described in [1]. We also used the local scale measure proposed in [2] as additional texture feature.

As Fig. 1 reveals, the method works fine for the artificial texture images. The optimum number of regions has been detected. The same holds for the test image depicted in Fig. 2, which is often used in the literature, e.g. in [9]. Often much more difficult, are real world images. Comparing, however, the segmentation result of the penguin image in Fig. 4 to the result in [12] shows that our method is competitive to other well-known methods. While in [12] 6 regions have been detected, the 3 regions found by our method are more reasonable. Our level set framework also compares favourably to the algebraic multigrid method in [9], as can be observed by means of the difficult squirrel image in Fig. 5a. Also Fig. 5b and Fig. 6 show an almost perfect segmentation.

It should be noted that all parameters that appear in the method have been set to fixed values, so all results shown here have been achieved with the *same* parameters. This is important, as of course it is much easier to obtain good segmentation results, if the parameters are tuned for each specific image. However, we think that this contradicts somehow the task of *unsupervised* segmentation.

The algorithm is reasonably fast. The 169×250 koala image took 22.5 seconds on an Athlon XP 1800+ including feature computation.

---

[1] This will only be the case, if the simplified objective function at the coarsest scale is unimodal and the global optimum of each next finer scale is the optimum closest to the global optimum at the respective coarser scale.

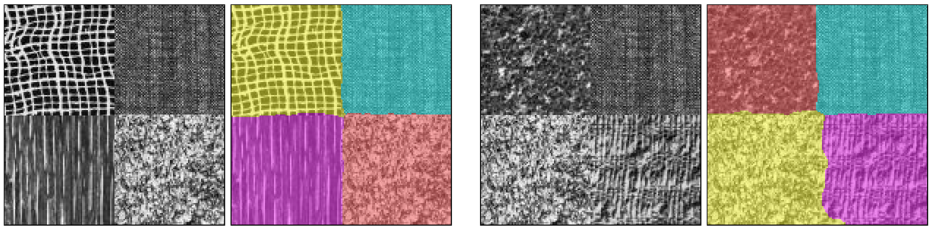

Fig. 1. Segmentation of two artificial texture images: In both cases 4 regions were detected.

5 Summary

In this paper we proposed a level set based minimisation scheme for the variational segmentation model of Zhu-Yuille. While the popular level set framework has so far only been used for two-region segmentation or segmentation with a fixed number of regions, we described a way how to optimise the result also regarding the number of regions. Moreover, the divide-and-conquer principle provides good initialisations, so the method is less sensitive to local minima than comparable methods. All advantages of the level set framework are preserved, while its main problem has been solved. The performance of the variational model and its minimisation strategy has been demonstrated in several experiments. It compares favourably to existing approaches from the literature.

References

1. T. Brox, M. Rousson, R. Deriche, and J. Weickert. Unsupervised segmentation incorporating colour, texture, and motion. In N. Petkov and M. A. Westenberg, editors, *Computer Analysis of Images and Patterns*, volume 2756 of *Lecture Notes in Computer Science*, pages 353–360. Springer, Berlin, 2003.
2. T. Brox and J. Weickert. A TV flow based local scale measure for texture discrimination. In T. Pajdla and J. Matas, editors, *Proc. 8th European Conference on Computer Vision*, Lecture Notes in Computer Science. Springer, Berlin, May 2004. To appear.
3. V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66:1–31, 1993.
4. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proc. Fifth International Conference on Computer Vision*, pages 694–699, Cambridge, MA, June 1995. IEEE Computer Society Press.
5. T. Chan and L. Vese. An active contour model without edges. In M. Nielsen, P. Johansen, O. F. Olsen, and J. Weickert, editors, *Scale-Space Theories in Computer Vision*, volume 1682 of *Lecture Notes in Computer Science*, pages 141–151. Springer, 1999.
6. D. Cremers. A multiphase levelset framework for variational motion segmentation. In L. D. Griffin and M. Lillholm, editors, *Scale Space Methods in Computer Vision*, volume 2695 of *Lecture Notes in Computer Science*, pages 599–614. Springer, Berlin, June 2003.
7. D. Cremers, N. Sochen, and C. Schnörr. Towards recognition-based variational segmentation using shape priors and dynamic labeling. In L. D. Griffin and M. Lillholm, editors, *Scale Space Methods in Computer Vision*, volume 2695 of *Lecture Notes in Computer Science*, pages 388–400. Springer, Berlin, June 2003.
8. A. Dervieux and F. Thomasset. A finite element method for the simulation of Rayleigh–Taylor instability. In R. Rautman, editor, *Approximation Methods for Navier–Stokes Problems*, volume 771 of *Lecture Notes in Mathematics*, pages 145–158. Springer, Berlin, 1979.

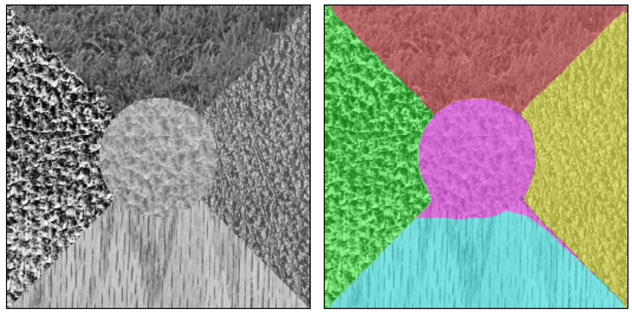

Fig. 2. Segmentation of a texture image: 5 regions have been detected.

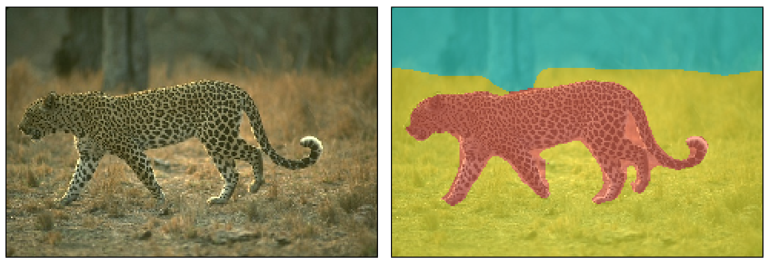

Fig. 3. Segmentation of a leopard image (colour): 3 regions have been detected.

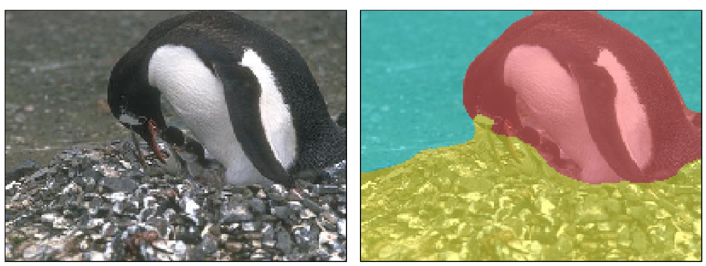

Fig. 4. Segmentation of a penguin image (colour): 3 regions have been detected.

Fig. 5. LEFT: (a) Segmentation of a squirrel image: 2 regions have been detected.
RIGHT: (b) Segmentation of a koala image (colour): 4 regions have been detected.

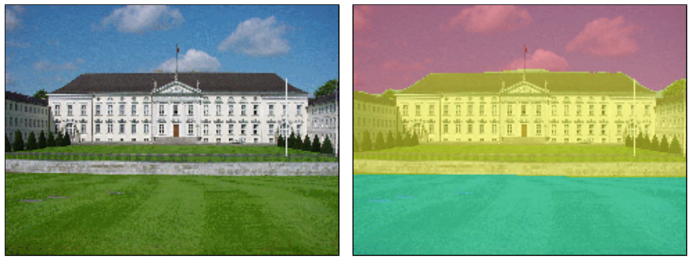

Fig. 6. Segmentation of a castle image (colour): 3 regions have been detected.

9. M. Galun, E. Sharon, R. Basri, and A. Brandt. Texture segmentation by multiscale aggregation of filter responses and shape elements. In *Proc. IEEE International Conference on Computer Vision*, pages 716–723, Nice, France, Oct. 2003.
10. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. In *Proc. Fifth International Conference on Computer Vision*, pages 810–815, Cambridge, MA, June 1995. IEEE Computer Society Press.
11. M. E. Leventon, W. E. L. Grimson, and O. Faugeras. Statistical shape infuence in geodesic active contours. In *Proc. 2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 316–323, Hilton Head, SC, June 2000. IEEE Computer Society Press.
12. J. Malik, S. Belongie, J. Shi, and T. K. Leung. Textons, contours and regions: cue integration in image segmentation. In *Proc. IEEE International Conference on Computer Vision*, pages 918–925. Corfu, Greece, Sept. 1999.
13. D. Mumford and J. Shah. Boundary detection by minimizing functionals, I. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 22–26, San Francisco, CA, June 1985. IEEE Computer Society Press.
14. S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations. *Journal of Computational Physics*, 79:12–49, 1988.
15. N. Paragios and R. Deriche. A PDE-based level-set approach for detection and tracking of moving objects. In *Proc. Sixth International Conference on Computer Vision*, pages 1139–1145, Bombay, India, Jan. 1998.
16. N. Paragios and R. Deriche. Coupled geodesic active regions for image segmentation: A level set approach. In *Proc. Sixth European Conference on Computer Vision*, number 2, pages 224–240, Dublin, Ireland, 2000.
17. M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. In *Proc. 2003 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, volume 2, pages 699–704, Madison, WI, June 2003.
18. C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia. A level set model for image classification. *International Journal of Computer Vision*, 40(3):187–197, 2000.
19. L. Vese and T. Chan. A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, 50(3):271–293, Dec. 2002.
20. A. Yezzi, A. Tsai, and A. Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *Proc. 7th International Conference on Computer Vision*, volume 2, pages 898–903, Kerkyra, Greece, Sept. 1999.
21. H. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *Journal of Computational Physics*, 127:179–195, 1996.
22. S. Zhu and A. Yuille. Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(9):884–900, Sept. 1996.