In B. Andres, B. Rosenhahn (Eds.): Pattern Recognition. Lecture Notes in Computer Science, Vol. 9796, pp. 117–128, Springer, Cham, 2016.

Robust Interactive Multi-label Segmentation with an Advanced Edge Detector

Sabine Müller[1,2], Peter Ochs[2], Joachim Weickert[2], and Norbert Graf[1]

[1] Department of Pediatric Oncology and Hematology, Saarland University Hospital, 66421 Homburg, Germany
[2] Mathematical Image Analysis Group, Saarland University, Campus E1.7, 66041 Saarbrücken, Germany

Abstract. Recent advances on convex relaxation methods allow for a flexible formulation of many interactive multi-label segmentation methods. The building blocks are a likelihood specified for each pixel and each label, and a penalty for the boundary length of each segment. While many sophisticated likelihood estimations based on various statistical measures have been investigated, the boundary length is usually measured in a metric induced by simple image gradients. We show that complementing these methods with recent advances of edge detectors yields an immense quality improvement. A remarkable feature of the proposed method is the ability to correct some erroneous labels, when computer generated initial labels are considered. This allows us to improve state-of-the-art methods for motion segmentation in videos by 5–10% with respect to the F-measure (Dice score).

1 Introduction

Image segmentation is a highly ambiguous task. By definition, the goal is a partitioning of an image into a finite number of meaningful regions. Unless there is a clear specification of the word "meaningful", the quality of a segmentation cannot be measured. Usually a "meaningful" segmentation refers to a separation of the visible objects in an image. However, even this task can be ambiguous. Consider for instance a fruit bowl with many apples. Should we consider each apple as an individual object, or the collection of apples in the fruit bowl as a single object?

In this paper, we argue that this distinction is important even when user input is provided. Assume the user seeds a label on a single apple. If only the seeded apple is to be segmented, the label can be propagated until the closest visual edge of the apple. Strong prior assumptions like the grouping of points that have a similar texture or colour can counteract the goal of segmenting a single apple. Similar structures are connected across edges and the label spreads

to other apples as well. This is beneficial when the collection of apples is to be segmented. Nevertheless, in any of the two cases edge information is essential for achieving the desired segmentation.

The goal of our paper is to complement a segmentation method that uses some user input with results from a sophisticated edge detector. The challenge for the segmentation method is to align the segments' boundaries with the "right" edges. Since the edge detector does not know what object the user is interested in, many edges are generated that are unimportant for the segmentation method. Moreover, often edges are not closed contours, hence selecting the right edges is not enough. The segmentation method needs to extend the proposed edges and close gaps. These challenges are by no means trivial. Additionally, the perception of important edges can vary significantly from image to image. While in a low contrast recording—for example taken at night—a weak change of colour is important, the same variation of colours is negligible in a high contrast recording. The proposed segmentation method deals with these challenges and shows a favourable performance on several semi-supervised segmentation benchmarks.

We apply our method also to video segmentation where many ambiguities are resolved by considering motion. Current state-of-the-art methods in motion segmentation in videos cluster a sparse set of point trajectories into similarly moving groups. Their output is an object-oriented sparse labelling of each video frame (more or less evenly distributed). We take these labellings as "user input" and turn the sparse labels into dense segmentations. A major challenge in this task is the ability to correct a moderate amount of these input labels. They are automatically generated labels and are often erroneous close to object boundaries. We significantly improve on the state-of-the-art method in this challenge.

Last but not least, we show that our method is also interesting for medical applications, where for example a clinician roughly marks a kidney tumour in a single slice of a volumetric MRI recording and receives a segmentation of the tumour in all slices.

2 Related Work

We propose a method for interactive segmentation. The "object" that is to be segmented in the image is seeded with labels. In the context of interactive segmentation these labels are often called scribbles and are set by a user. However, labels can also be generated automatically by an algorithmic method. In this paper, we consider both scenarios, which are fundamentally different. While user scribbles are usually assumed to be correct, i.e., the segmentation method needs to fill in labels between the given ones, automatically generated labels can be erroneous and need to be corrected partially.

Most of the available literature deals with the former scenario. A successful idea is the estimation of statistical features from given scribbles in order to define likelihoods for each label, e.g. mean value [9], colour/intensity histograms [4,13,23,18], or texture [1,25,24,19]. As a regulariser usually the boundary length in the image metric is minimised. In [28] the boundary length is penalised in the

metric induced by the image gradients, which aligns the segmentation boundaries to image edges. In [29] the boundary length is measured in a non-local metric, which achieves good results for small details, but suffers from expensive non-local computations.

Of course, the image metric can be induced by more sophisticated edge indicators such as the traditional ones [6,22], which are based on colour and intensities, or more recent ones that include also texture information. We propose to induce the image metric by one of the state-of-the-art trained edge detectors by Dollár and Zitnick [11,12]. Besides having better information about edges in the image, much fewer edges are detected compared to the simple gradient magnitude measure. This has also a positive effect on the computation time, as the propagation of labels is hampered less by unimportant structures.

A flexible framework for multi-label segmentation is the formulation as a minimal partitioning problem [15,17], which can be solved via convex relaxation methods [7]. This framework, which is sometimes referred to as Potts multi-label segmentation model, is very flexible, since all the above data likelihoods can be incorporated. The boundary length is represented using the total variation of the label indicator functions, and is easily adjusted to a modification of the image metric. This convex relaxation framework is used for example in [28,18,19,20,29] for the task of interactive segmentation. In [27] it has been extended to a non-metric prior, in [26] to generalised ordering constraints, which constraints labels appearing adjacent to each other in a certain direction, in [10] to RGB-D data, and in [3] to the context of semantic segmentation. All these models can be easily complemented in the way we propose by a sophisticated edge detector.

The second scenario where automatically generated labels must be corrected has been considered in [20,21], which is also formulated in the convex relaxation framework mentioned above. We propose a generalisation of their method to an arbitrary edge indicator function, and this extension has a strong impact on practical results. On the FBMS-59 benchmark [21] we improve their method by about 5–10%. For recent works on interactive segmentation on medical images, we refer to [14,2] and the references therein.

3 Method

In this section, we discuss a generic multi-label segmentation model. Based on this model, we recapitulate the interactive segmentation method proposed in [18] and our improvements: We take erroneous user scribbles into account and introduce a new regularisation term with an advanced edge detector.

3.1 A Multi-label Segmentation Model

Following Chambolle et al. [7], we consider a minimal partitioning problem of the rectangular image domain $\Omega \subset \mathbb{R}^2$ into $\Omega_1, \ldots, \Omega_n \subset \mathbb{R}^2$ non-overlapping

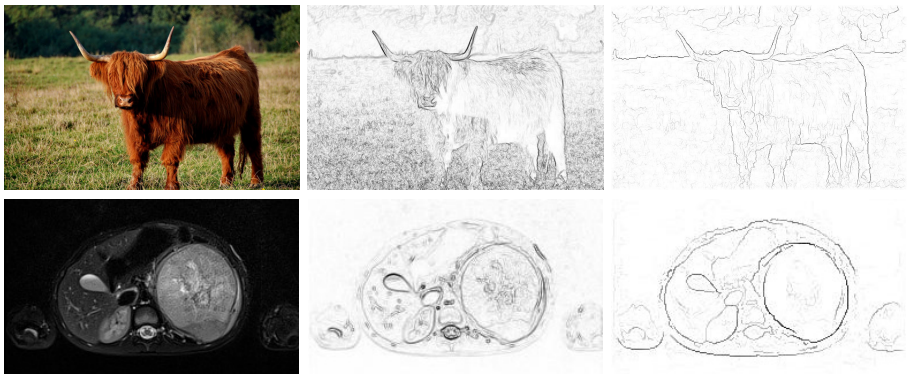

Fig. 1. Exemplary results for edge detection. From left to right: Input image. Gradient magnitude image. Result of structured edge detector. Edge maps are inverted and gamma corrected for visualisation. The structured edge detector shows more object edges and less clutter for unimportant structures.

regions. The generic variational problem is

$$\begin{aligned} \min_{\Omega_1,\ldots,\Omega_n \subset \Omega} \; & \frac{1}{2} \sum_{i=1}^{n} \mathrm{Per}(\Omega_i; \Omega) + \sum_{i=1}^{n} \int_{\Omega_i} h_i(\mathbf{x}) \, d\mathbf{x} \,, \\ \text{s.t.} \quad & \Omega = \bigcup_{i=1}^{n} \Omega_i, \quad \Omega_i \cap \Omega_j = \emptyset, \quad \forall \; i \neq j \end{aligned} \tag{1}$$

where $h_i \colon \mathbb{R} \to \mathbb{R}_+$ are potential functions reflecting the cost for each pixel being assigned to a certain label $i = 1, \ldots, n$, and $\mathrm{Per}(\Omega_i; \Omega)$ denotes the perimeter of region Ω_i inside Ω. A weighting parameter $\lambda > 0$ is not required, as it can be absorbed in the functions h_i. In order to improve the alignment of region boundaries with image edges the perimeter is usually measured in a metric that is induced by the underlying image $f \colon \Omega \to \mathbb{R}^d$. A common choice is a weighting with the image gradient with $\exp(-\gamma |\nabla f(\mathbf{x})|)$, where ∇f denotes the Jacobian of f and $|\nabla f|$ is its Frobenius norm. This reduces the measure of the boundary length where the image gradient magnitude is high. As the middle column of Fig. 1 demonstrates, this choice is suboptimal when we seek for segmentations of objects. The image gradient magnitude shows clutter, i.e., it is high for unimportant edges. In this paper, we propose to use a sophisticated edge detector instead. We built on the fast structured edge detector [11,12], which in contrast to traditional edge detectors, incorporates texture, colour, and brightness. The right column of Fig. 1 shows that this state-of-the-art edge detector is well-suited to identify also texture edges and illusory contours. Let $\mathcal{E} \colon \Omega \to \mathbb{R}$ be the output of this edge detector, i.e., $\mathcal{E}(\mathbf{x})$ is large on edges and low otherwise. We weight the perimeter of region boundaries with the function

$$g(\mathbf{x}) = \exp(\mathcal{E}(\mathbf{x})^{\beta} / \, \bar{\mathcal{E}}) \,, \tag{2}$$

where $\bar{\mathcal{E}} := \frac{2}{\Omega} \int_\Omega |\mathcal{E}(\mathbf{x})| \, d\mathbf{x}$ and β is a positive parameter.

Let us now turn our attention to the potential functions in the second term of (1). Any method that proposes a new way to estimate these potential functions may be combined with the perimeter regularisation discussed above. In the following, we recap the model proposed by Nieuwenhuis et al. [18].

Spatially varying colour distributions. Assume the user provides a (measurable) set of scribbles $\mathcal{S}_i \subset \Omega$ for each label i. Nieuwenhuis and Cremers [18] suggest to define the potential function $h_i(\mathbf{x})$ in (1) as the negative logarithm of the linearly to $[0, 1]$-scaled function $\tilde{h}(\mathbf{x})$ of [3]

$$\frac{1}{|\mathcal{S}_i|} \int_{\mathcal{S}_i} k_{\rho_i(\mathbf{x})}(\mathbf{x} - \mathbf{y}) \; k_\sigma(f(\mathbf{x}) - f(\mathbf{y})) \, d\mathbf{y} \,, \tag{3}$$

where $|\mathcal{S}_i|$ denotes the area that is occupied by ith label, k_σ and k_{ρ_i} are Gaussians with standard deviation σ in colour space and adaptive standard deviation $\rho_i(\mathbf{x}) = \alpha \inf_{\mathbf{y} \in \mathcal{S}_i} |\mathbf{x} - \mathbf{y}|$ in the spatial domain, respectively. The idea of this spatially adaptive standard deviation is to reduce the influence of the colour distribution from scribbles that are far away. This influence is reduced proportionally to distance from $\mathbf{x}$ to the closest scribble location.

In our opinion, a major drawback of this model is the assumption that all labels are correct. Formally, $h_i(\mathbf{x})$ must be set to $+\infty$ for $\mathbf{x} \in \mathcal{S}_i$, since (3) does not make sense for $\rho_i(\mathbf{x}) = 0$. Therefore, we propose to use potential functions that allow the segmentation method to correct possibly wrong scribbles/labels—this issue arises for instance when scribbles are provided automatically by a computer. This is achieved by setting for scribble positions $\mathbf{x} \in \mathcal{S}_j$ the function values $\tilde{h}_i(\mathbf{x}) = 1 - \zeta$ if $i = j$ and $\tilde{h}_i(\mathbf{x}) = \zeta/(n-1)$ otherwise, where $1 - \zeta$ is the assumed probability for the scribble being correct.

The variational minimization problem (1) can be solved efficiently after a convex relaxation [7]. The regions $\Omega_1, \ldots, \Omega_n$, their non-overlapping and covering criterion can be easily represented by label indicator functions, whose ranges get relaxed to $[0, 1]$. In this representation, the perimeter of the regions is measured by the weighted total variation. Using the dual definition of the total variation makes the application of the efficient primal–dual algorithm in [8] straightforward.

4 Experimental Evaluation

We evaluated our approach on the GRAZ benchmark [24], the FBMS-59 [21] data sets, and MRI data sets from patients with kidney tumours. We show results in terms of the metrics suggested in [18,19,21,24] : *precision* and *recall*

[3] Instead of integrating over the set of scribbles they sum over all scribbled pixels.

are computed as

$$P_{\hat{\Omega}_i,\Omega_i} := \frac{|\hat{\Omega}_i \cap \Omega_i|}{|\Omega_i|}\,, \qquad R_{\hat{\Omega}_i,\Omega_i} := \frac{|\hat{\Omega}_i \cap \Omega_i|}{|\hat{\Omega}_i|}\,, \tag{4}$$

where $\hat{\Omega}_i$ is the ground truth labelling, and Ω_i the result of our algorithm for label i. The harmonic mean of precision and recall, the *F-measure* or *Dice-score* relates the area of a cluster to its overlap with the ground truth:

$$F_{\hat{\Omega}_i,\Omega_i} = \frac{2P_{\hat{\Omega}_i,\Omega_i} R_{\hat{\Omega}_i,\Omega_i}}{P_{\hat{\Omega}_i,\Omega_i} + R_{\hat{\Omega}_i,\Omega_i}}\,. \tag{5}$$

Finally, the average F-measure determines the overall segmentation accuracy.

4.1 Multi-label Segmentation of Colour Images

The GRAZ benchmark dataset [24] consists of 262 seed-ground-truth pairs from 158 natural images for interactive multilabel segmentation. We use the following manually tuned parameters for experiments with space-variant colour distributions: $\alpha = 15$, $\beta = 2$, $\sigma = 3$, and $\zeta = 0.05$.
We compare our results in Tab. 1 to the original approach by Nieuwenhuis et al. [18], as well as their advanced method that incorporates texture information [19]. The results indicate that using an advanced edge detector textural information in the data term can be neglected. Fig. 2 shows an exemplary result from our evaluation. The information from an advanced edge detection is sufficient for segmentations of high quality.

Table 1. Comparison to spatially variant approaches by Nieuwenhuis et al [18,19].

Method	Dim	Dice Score
Nieuwenhuis/Cremers, spatially constant [18]	3	0.89
Nieuwenhuis/Cremers, space-variant [18]	5	0.92
Nieuwenhuis/Cremers, space-variant [18]	13	0.93
Nieuwenhuis et al., space-variant + texture [19]	13	0.94
Our approach, spatially constant, no colour	5	0.77
Our approach, spatially constant, colour	5	0.90
Our approach, space-variant, no colour	5	0.80
Our approach, space-variant, colour	5	0.93
Our approach, space-variant, colour	13	0.94

Decreasing the diameter of user scribbles leads to slightly worse approximations of the colour distributions. However, our model can compensate this lack of information since textural and colour information are also included in the edge detector. In fact, the texture based model [19] seems to rely on a large diameter.

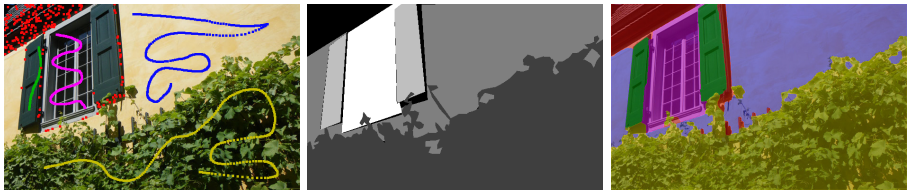

Fig. 2. Exemplary segmentation result exclusively based on colour variation and the structured edge detector. From left to right: Scribble image, dim = 5. Ground truth labelling. Our result.

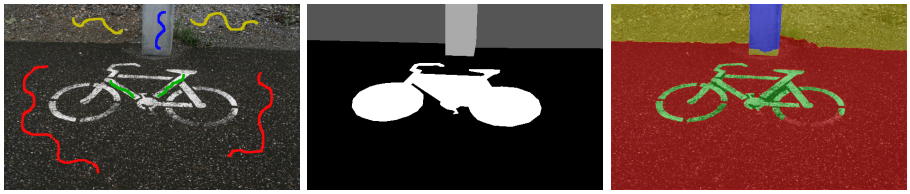

Fig. 3. Limitations of GRAZ benchmark. From left to right: Scribble image. Ground truth label image. Our result. It is clearly visible, that the quality of image labels is limited and the segmentation outcome can not reflect the user intention. Dice score: 0.86

The creators of the GRAZ benchmark made an important contribution towards evaluation of interactive multi-label segmentation, but current state-of-the-art segmentation reached the limit of these data sets. Small details are very important, and not all ground truth labellings are optimal, see Fig. 3. Overall, the results on this data set are very close to the optimum. Therefore it might not be desirable to further improve the Dice score on this data set.

4.2 Video Segmentation

The FBMS-59 [21] data set, an extended version of [5], contains 29 video sequences for training and 30 video sequences for testing. We used minimum cost multicuts [16] as well as [21] to generate and automatically label point trajectories. Ideally, they provide sparse, and temporally consistent labels for each frame in a video. In contrast to interactive or supervised segmentation, trajectory labels are single pixels spread over the whole image domain and tend to be erroneous. Video segmentation is a challenging task. Even the considered state-of-the-art methods [16,21] provide erroneous trajectories. Therefore, methods that turn the sparse labels into dense segmentations, such as our method, must be able to correct some of the wrong labels. We incorporate this prior knowledge by increasing the uncertainty parameter ζ. The manually tuned parameters we used for all video frames are $\alpha = 30$, $\beta = 2$, $\sigma = 3$, and $\zeta = 0.2$.

We compare our approach to the recent approach [21] that densifies the sparse labels from the trajectories in each frame based on image gradients. We

Table 2. Results on the Video Segmentation Benchmark FBMS-59.

Method	Density	Dice	Precision	Recall	F≥ 75%
Training set					
Ochs et al., MoSegDense [21]	100%	0.69	0.84	0.59	15/65
Ochs et al., MoSegSparse [21]	0.87%	0.72	0.85	0.62	17/65
Our approach, SPT-C, NC (SC [21])	100%	0.79	0.83	0.75	18/65
Our approach, SPT-V, C (SC [21])	100%	0.81	0.84	0.78	20/65
Keuper, MCe sparse, prior 0.5 (MT [16])	0.86%	0.79	**0.87**	0.73	31/65
Our approach, SPT-V, C (MT [16])	100%	**0.82**	0.85	**0.80**	24/65
Test set					
Ochs et al., MoSegDense [21]	100%	0.66	0.78	0.57	17/69
Ochs et al., MoSegSparse [21]	0.92%	0.69	0.80	0.61	24/65
Our approach, SPT-C, NC (SC [21])	100%	0.71	0.75	0.68	18/65
Our approach, SPT-V, C (SC [21])	100%	0.74	0.76	**0.72**	21/69
Keuper, MCe sparse, prior 0.5 (MT [16])	0.87%	**0.76**	**0.88**	0.68	25/69
Our approach, SPT-V, C (MT [16])	100%	0.75	0.81	0.71	23/69

stick to point trajectories generated by [21] for the sake of comparability. Tab. 2 states our benchmark results for two variants of our model: In the first version, we include exclusively information about label position and rely on information already contained in our edge detector (SPT-C, NC). This version coincides with [21] except for the edge detection. In the second variant, we include all available information for segmentation and use spatially variant colour distributions (SPT-V, C). Though the edge detector should already contain all relevant color and texture information in the image, favouring slightly color similarity improves the results. This is due to a suboptimal performance of the edge detector. We clearly outperform [21] on all error metrics.
Moreover, we complement the sparse labels of the state-of-the-art in motion segmentation from Keuper et al. [16] with our approach. The ability to correct also erroneous labels (see Fig. 4) allows us to even improve the results in the dense segmentation. Note that the difference in performance on test and training set is due to different challenges and the (still limited) number of video sequences rather than over-fitting.

4.3 Volumetric MRI Data

Additionally, we evaluate the performance of our method on MRI data of unilateral kidney tumours and rely on image data of 12 nephroblastoma-patients before and after chemotherapy. This encompasses in total 24 pre- and post-chemotherapeutic monomodal volumetric images of T_2 modality (inslice-sampling ranges from 0.4mm to 0.6mm, across-slice sampling from 1mm to 7mm). The (pre- and post-chemotherapy) ground truth of the tumour's outline was defined by manual expert segmentation of two human expert raters. A clinician, familiar

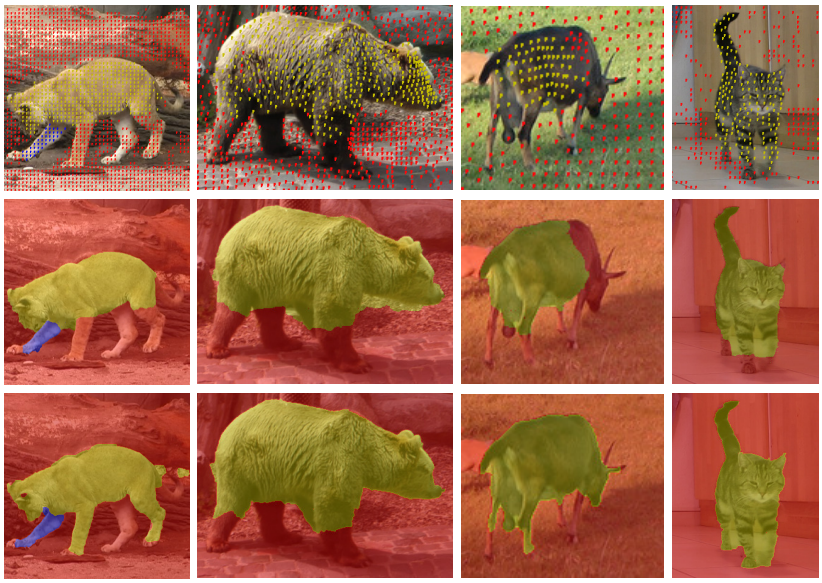

Fig. 4. Exemplary results on FMBS-59 [21] data sets. From top to bottom: Labels from SC-point trajectories [21]. Segmentation results of Ochs et al., MoSegDense [21]. Our segmentation result, no colour, spatially constant. Trajectory labels are enhanced for visualisation. Our method is able to correct a significant amount of labels.

with tumours, draw user scribbles in a single depth slice for each data set. We use the following parameters for our experiments on MRI data: $\alpha = 15$, $\beta = 2$, $\sigma = 5$, and $\zeta = 0.05$. We compare the results of our approach for several kind

Table 3. Evaluation for unilateral nephroblastoma patients.

Method	Dim	Dice	Precision	Recall
Our approach, SPT-C, no intensity	5	0.76	0.79	0.74
Our approach, SPT-C, intensity	5	0.77	0.80	0.75
Our approach, SPT-V, no intensity	5	0.79	0.83	0.75
Our approach, SPT-V, intensity	5	0.90	0.92	0.89
Our approach, SPT-V, intensity	13	0.91	0.92	0.91

of information in Tab. 3. Our method does not perform well when no intensity information is included as only two labels are provided - tumour and not tumour. Incorporating intensity information dramatically improves the result, but similar to our results on the GRAZ benchmark in Tab. 1, few user scribbles are sufficient for highly accurate segmentation results.

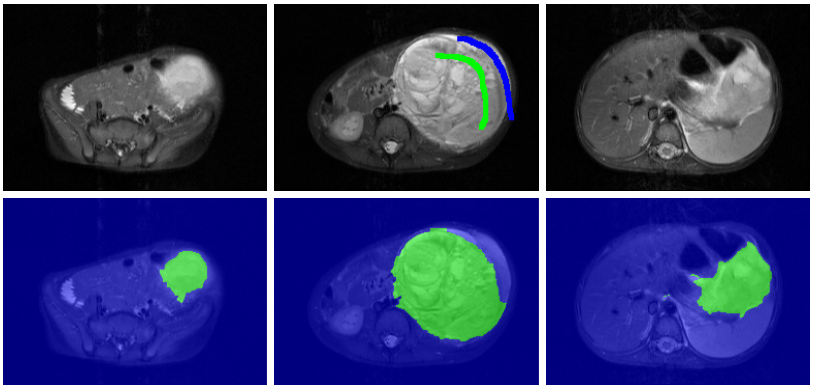

Fig. 5. Example results for 3D medical image segmentation. From left to right: Slice 2, slice 7, slice 19. User scribbles are set for one slice. Cost volume computation, as well as optimisation are conducted in 3D.

4.4 Runtimes

We conducted our experiments on a desktop PC with 3.4GHz Intel Core i7 CPU, 16GB RAM, and a NVidia GeForce GTX 970 graphics card with 4GB RAM and 1664 CUDA cores. The edge maps [12] for all benchmarks were pre-computed with an average computation time of 0.9 s. The average computation time for an image of size 625×391 was 1.08 s. This is comparable to [19,18] and [24].

5 Summary

We proposed a robust algorithm for interactive multi-label segmentation based on a sophisticated edge detector that includes texture, colour, and brightness. As interaction, we considered two scenarios: The labels are set by a user and are always correct, and the labels are computed by an algorithm and are likely to be erroneous. Especially the second setting is rarely considered. However, a common methodology for motion segmentation in videos is to cluster a sparse set of point trajectories. These trajectories can be considered as the (erroneous) user scribbles in each frame and can be made dense by our approach. For this problem, we improve on state-of-the-art by about 5–10% in the Dice score. In our ongoing research, we are also considering extensions of our approach that allow uncertainty quantification labellings of other computer vision tasks.

Acknowledgements. This work was partially funded by the European Union's FP7 under the project Computational Horizons in Cancer (grant agreement No 600841).

References

1. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(5), 898–916 (2011)
2. Bauer, S., Porz, N., Meier, R., Pica, A., Slotboom, J., Wiest, R., Reyes, M.: Interactive segmentation of MR images from brain tumor patients. In: Proc. 11th International Symposium on Biomedical Imaging. pp. 862–865. Beijing, China (2014)
3. Bergbauer, J., Nieuwenhuis, C., Souiai, M., Cremers, D.: Proximity priors for variational semantic segmentation and recognition. In: ICCV Workshop on Graphical Models for Scene Understanding. pp. 15–21. Sydney, Australia (2013)
4. Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: Proc. 2001 IEEE International Conference on Computer Vision. pp. 105–112. Vancouver, Canada (2001)
5. Brox, T., Malik, J.: Object segmentation by long term analysis of point trajectories. In: Kostas Daniilidis, Petros Maragos, N.P. (ed.) Computer Vision–ECCV 2010, Lecture Notes in Computer Science, vol. 6315, pp. 282–295. Springer, Heraklion, Greece (2010)
6. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8, 679–698 (1986)
7. Chambolle, A., Cremers, D., Pock, T.: A convex approach to minimal partitions. SIAM Journal on Applied Mathematics 5(4), 1113–1158 (2012)
8. Chambolle, A., Pock, T.: A first-order primal–dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision 40(1), 120–145 (May 2011)
9. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Transactions on Image Processing 10(2), 266–277 (Feb 2001)
10. Diebold, J., Demmel, N., Hazrba, C., Möller, M., Cremers, D.: Interactive multi-label segmentation of RGB-D images. In: Aujol, J.F., Nikolova, M., Papadakis, N. (eds.) Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, vol. 9087, pp. 294–306. Springer, Berlin (2015)
11. Dollár, P., Zitnick, C.L.: Structured forests for fast edge detection. In: Proc. 2013 IEEE International Conference on Computer Vision. pp. 1841–1848. Washington, DC, USA (2013)
12. Dollár, P., Zitnick, C.L.: Fast edge detection using structured forests. IEEE Transactions on Pattern Analysis and Machine Intelligence 37(8), 1558–1570 (2015)
13. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. International Journal of Computer Vision 59(2), 167–181 (2004)
14. Feng, C., Zhao, D., Huang, M.: Segmentation of stroke lesions in multi-spectral MR images using bias correction embedded FCM and three phase level set. MICCAI Ischemic Stroke Lesion Segmentation p. 3 (2015)
15. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721–741 (1984)
16. Keuper, M., Andres, B., Brox, T.: Motion trajectory segmentation via minimum cost multicuts. In: Proc. 2015 IEEE International Conference on Computer Vision. pp. 3271–3279. Santiago, Chile (2015)
17. Mumford, D., Shah, J.: Optimal approximation of piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics 42, 577–685 (1989)

18. Nieuwenhuis, C., Cremers, D.: Spatially varying color distributions for interactive multilabel segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 35(5), 1234–1247 (2013)
19. Nieuwenhuis, C., Hawe, S., Kleinsteuber, M., Cremers, D.: Co-sparse textural similarity for interactive segmentation. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) Computer Vision–ECCV 2014, Lecture Notes in Computer Science, vol. 8694, pp. 285–301. Springer, Cham, Switzerland (2014)
20. Ochs, P., Brox, T.: Object segmentation in video: A hierarchical variational approach for turning point trajectories into dense regions. In: Proc. 2011 IEEE International Conference on Computer Vision. pp. 1583–1590. Barcelona, Spain (2011)
21. Ochs, P., Malik, J., Brox, T.: Segmentation of moving objects by long term video analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 36(6), 1187–1200 (2014)
22. Perona, P., Malik, J.: Scale space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 629–639 (1990)
23. Rother, C., Kolmogorov, V., Blake, A.: "GrabCut": interactive foreground extraction using iterated graph cuts. In: Proc. SIGGRAPH 2004. pp. 309–314. New York, NY, USA (Aug 2004)
24. Santner, J., Pock, T., Bischof, H.: Interactive multi-label segmentation. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) Computer Vision – ACCV 2010, Lecture Notes in Computer Science, vol. 6492, pp. 397–410. Springer, Queenstown, New Zealand (2010)
25. Santner, J., Unger, M., Pock, T., Leistner, C., Saffari, A., Bischof, H.: Interactive texture segmentation using random forests and total variation. In: Proc. 2009 British Machine Vision Conference. pp. 66.1–66.12. London, UK (2009)
26. Strekalovskiy, E., Cremers, D.: Generalized ordering constraints for multilabel optimization. In: Proc. 2011 IEEE International Conference on Computer Vision. pp. 2619–2626. Barcelona, Spain (2011)
27. Strekalovskiy, E., Nieuwenhuis, C., Cremers, D.: Nonmetric priors for continuous multilabel optimization. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) Computer Vision - ECCV 2012, pp. 208–221. Lecture Notes in Computer Science, Springer, Berlin (2012)
28. Unger, M., Pock, T., Trobin, W., Cremers, D., Bischof, H.: TVSeg - interactive total variation based image segmentation. In: Proc. 2008 British Machine Vision Conference. pp. 40.1–40.10. Leeds, UK (2008)
29. Werlberger, M., Unger, M., Pock, T., Bischof, H.: Efficient minimization of the non-local potts model. In: Bruckstein, A.M., ter Haar Romeny, B.M., Bronstein, A.M., Bronstein, M.M. (eds.) Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, vol. 6667, pp. 314–325. Springer, Berlin (2011)