Segmentation of Microorganism in Complex Environments[1]

M. Kemmler, B. Fröhlich, E. Rodner, and J. Denzler

Chair for Computer Vision, Frendlich Schiller University of Jena

e-mail: michael.kemmler@uni-jena.de; bjoern.froehlich@uni-jena.de; erik.rodner@uni-jena.de; joachim.denzler@uni-jena.de

Abstract—In this paper, we tackle the problem of finding microorganisms in bright field microscopy images, which is an important and challenging step in various tasks, like classifying soil textures. Apart from bacteria or fungi, these images can contain impurities such as sand particles, which increase the difficulty of microbe detection. Following a semantic segmentation approach, where a label is inferred for each pixel, we achieve encouraging classification results on a database containing five different types of microbes. We review and evaluate multiple techniques including segment classification, conditional random field models, and level set approaches.

Keywords: segmentation of Microorganism, microbe detection.

DOI: 10.1134/S1054661813040056

1. INTRODUCTION

In many applications and research studies, analyzing microorganisms in images from bright field microscopes offers important insights. To derive quantitative results, biological researchers often put a lot of effort in labeling and counting bacteria in those images, which can be an exhausting task. This leads to only a small number of images used for evaluation and therefore to numbers which are doubtful from a statistical point of view.

Previous work on bacteria detection and cell segmentation has been done by Wu and Shah [1], who utilize a pixel-wise Conditional Random Field for binary cell segmentation making use of colored fluorescence images and an additional multi-spectral data source. In contrast, Gelas et al. [2] employ multiphase level set segmentation for 3D cell segmentation using a Gaussian shape prior. Their main assumption is an already existing, very accurate initial segmentation provided by an expert. Seeded watershed transform is used by [3] for segmenting and tracking cells. All of these previous works assume that input images do not include impurities, such as sand particles, which is often not the case for real-world images (Fig. 1). It should be noted that techniques such as fluorescence staining [4], which help to separate biotic and abiotic particles, can assist in this step. However, they cannot be applied in settings where the background itself fluoresces. Our study hence serves as a starting point for assessing bacteria identification methods in these difficult scenarios.

In particular, this work focuses on evaluating different state-of-the-art *semantic segmentation techniques* for the task of automatically detecting bacteria and other microbes. The benefit of semantic segmentation approaches is the availability of a label for each pixel, instead of a coarse labeling with a bounding box. Therefore, we analyze pixel-based level set segmentation and approaches based on segment classification and present how to incorporate application-specific prior knowledge such as shape information to distinguish cells from sand particles. The goal of this paper is to provide a comparison of several techniques to

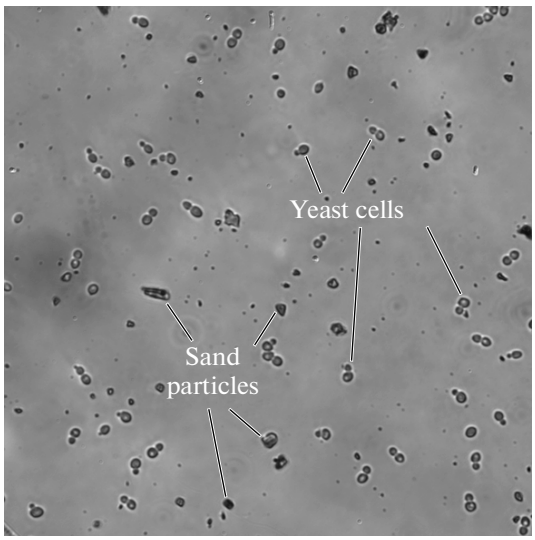

Fig. 1. Bright field microscopy image of yeast cells including sand particles and suffering from non–uniform illumination. As can be seen, differentiating between sand and cells is even challenging for human experts.

[1] The article is published in the original.

Received March 3, 2012

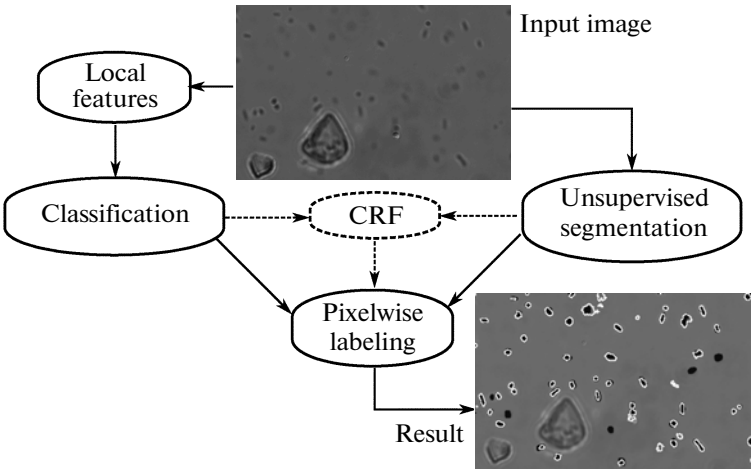

Fig. 2. Outline of our algorithm using segment classification and optional CRF optimization.

support their usage for automatic segmentation of microscopic images in life sciences.

Our experiments show that incorporating shape and appearance analysis, as well as modeling label dependencies between segments using Conditional Random Fields, substantially reduces the false positive rate of our detection scheme and gives more useful results compared to a level set segmentation approach. This is especially useful for input images that include a large number of outlier objects.

2. SEGMENT CLASSIFICATION

The task of inferring a label for each pixel in an image has been considered important in a lot of different tasks, such as scene understanding [5] and facade recognition [6]. In the following, we shortly outline a simplified variant of [6], which is subsequently used for microbe detection. A coarse outline is depicted in Fig. 2.

The first step is to obtain a segmentation of the image. This can be done fully unsupervised by employing some clustering algorithm such as mean shift [7]. This is followed by a pixel-wise classification on a predefined $M \times M$-grid using a classifier previously learned on a training set. For each segment, a label can then be inferred by averaging the classification results for all grid points falling inside of a segment. Since shape and appearance of microbes play a crucial role in distinguishing them from background, we include a post-processing step which takes these properties into account.

First of all, segments are labeled as background if they are too small (<5 pixels) or too big (>2000 pixels) in size. Secondly, shape and appearance features x_i are extracted from the remaining segments and used for a rating in a second classification stage. In our experiments, a probabilistic Parzen classifier was used for this step, since it allows for a fast leave-one-out estimation.

To obtain a hard decision, a segment i is classified as microbe if the probability $p(y_i = 1|x_i)$ of being microbe exceeds an automatically learned threshold. We used the minimal leave-one-out probability of all segments in the training image for this purpose. To take into account a possible over-segmentation of the image, neighboring segments are greedily clustered as to increase the segment likelihood of being microbe based on shape and appearance features.

3. INCORPORATING STRUCTURE WITH CONDITIONAL RANDOM FIELDS

In order to globally take into account interdependencies, the heuristic and greedy merging strategy from the last section is often insufficient. However, it is often sensible to deviate from the common independence assumption

$$p(y|\mathbf{X}) = \prod_{i=1}^{n} p(y_i|x_i) \quad (1)$$

and to include dependencies between segments. This can be done using graphical models where a graph $G = (V, E)$ is designed to capture certain dependencies between random variables. While the nodes $v \in V$ represent these random variables (segment labels), direct dependencies between two variables are described by edges $e \in E$. Conditional Random Fields (CRFs) are a famous subclass of graphical models which allow us to formulate dependencies between labels $\mathbf{y} = (y_1, ..., y_n)^T$ given some image features $\mathbf{X} = (x_1, ..., x_n)^T$.

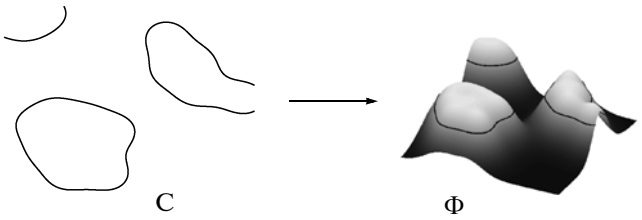

Fig. 3. Level set approach. Contours are described as those image coordinates which give rise to zero function values of Φ. Regions outside and inside the contours are then given as coordinates leading to positive and negative surface values, respectively.

It is known that for such an undirected dependency graph G, the joint probability of y factorizes w.r.t. maximal cliques c in G, i.e.,

$$p(y|\mathbf{X}) = Z^{-1}\prod_c \Psi_c(y_c|\mathbf{X}) \quad (2)$$

$$= \exp\left(\sum_c \log\Psi_c(y_c|\mathbf{X}) - \log Z\right). \quad (3)$$

One way to do inference with respect to labels y given observations $\mathbf{X}$, is to perform maximum a posteriori estimation. For sub-modular energies $\log p(y|X)$, this can be efficiently done using Graph-Cut [8]. A standard way is to specify the log-potential functions $\log\Psi_c(y_c|\mathbf{X})$ manually and we refer to this method as *Manual CRF* in our experiments.

4. PARAMETER ESTIMATION FOR LOG-LINEAR CRFS

Manual specification of a CRF is a tricky task and can lead to models that do not generalize well on unseen images. An alternative is to linearly decompose the log-potential functions in a parametric label term $\theta(y_c)$ and a feature term $f_c(\mathbf{X})$, i.e.,

$$\log\Psi_c(y_c|\mathbf{X}) = \theta(y_c)^T f_c(\mathbf{X}). \quad (4)$$

By doing so, parameters $\theta(y_c)$ can be learned in these log-linear CRFs by maximum likelihood estimation. Since the log-partition function $\log Z$ depends on those newly introduced parameters, the exact procedure is intractable and one has to resort to approximate inference techniques such as loopy belief propagation (LBP). Learned parameters can then be used for inference, where again Graph-Cut can be employed. However, instead of using MAP estimation given an image $\mathbf{X}_*$, we infer the marginal probabilities $p(y_i^* = 1|\mathbf{X}_*)$ of region i belonging to a microbe, using LBP. A hard decision can then be derived by thresholding those marginals. In this work, the threshold is empirically set to the 10th percentile of all marginal probabilities inferred for the training segments which corresponds to an assumed outlier ratio of 10%.

5. LEVEL SET SEGMENTATION

Active contours are curves that evolve to minimize a predefined energy functional, e.g., classical snakes [9] are moved through the image in order to fit strong edges such as clear object boundaries and at the same time obeying a smoothness property.

Level sets as first proposed by Osher and Sethian [10] can be understood as an extension to active contours which can directly handle complex morphology changes such as splitting and merging of object boundaries. The idea is to evolve a two-dimensional surface $\Phi : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ being defined over the whole (continuous) image domain Ω instead of just a parameterized one-dimensional contour C (see Fig. 3). This framework also enables us to easily include region information which is important for a large variety of segmentation tasks.

One of the most well-known region-based energy functionals is the piecewise-constant Mumford–Shah functional [11, 12], also known as the Chan–Vese functional. In the binary case, it is given as

$$E_{CV} = \left(\sum_{i=1}^{2}\lambda_i\int_{\Omega_i} l(\mathbf{I}(x), \theta_i)dx\right) + \nu L(C), \quad (5)$$

where $\Omega = \Omega_1 \cup \Omega_2 \cup C$ and $l(\mathbf{I}(x), \theta_i) = |\mathbf{I}(x) - c_i|^2$ is the quadratic loss between pixel intensity and constant $\theta_i = c_i$ which is a very coarse intensities model for region Ω_i. The parameter ν controls the trade-off between the intensity model and curve length regularization term $L(C)$.

A more general model uses Gaussian intensity distributions for each region, where the first term of E_{CV} can be replaced by $l(\mathbf{I}(x), \theta_i) = -\log p(\mathbf{I}(x)|\mu_i, \sigma_i)$ with $\theta_i = (\mu_i, \sigma_i)^T$ being mean and standard deviation of the assumed Gaussian density over intensities in respective regions [13]. To end up with an energy term whose components all act on the whole image domain Ω, the Heaviside function

$$H(z) = \begin{cases} 1\text{: } z > 0, \\ 0\text{: otherwise} \end{cases} \quad (6)$$

can be utilized. Since regions inside and outside contours are represented as those coordinates that give rise to negative and positive values of a 2D surface Φ, the Mumford–Shah functional based on some intensity loss $l(\mathbf{I}(x), \theta_i)$ can be formulated as

$$\begin{aligned} E_{CV} = \int_\Omega [&\lambda_1 H(\Phi(x))l(\mathbf{I}(x), \theta_1) \\ &+ \lambda_2(1 - H(\Phi(x)))l(\mathbf{I}(x), \theta_2) \\ &+ \alpha|\nabla H(\Phi(x))|]dx. \end{aligned} \quad (7)$$

When several objects in the image are represented by a single region, using a standard contour length

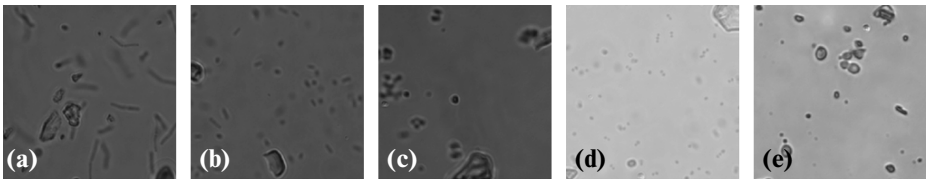

Fig. 4. Examples for usual microscopy images of *Bacillus subtilis* (a), *Escherichia coli* (b), *Micrococcus luteus* (c), *Staphylococcus epidermidis* (d), and *Yeast* (e). For better visibility, only a cropped region of the original images is displayed.

penalty $L(C)$ as in (7) might not have the desired effect. An alternative strategy is to include a geodesic active contour regularizer which encourages object boundaries to coincide with strong edges [14]:

$$E_{GAC} = \int_{\Omega} g(x)|\nabla H(\Phi(x))|dx, \tag{8}$$

where $g = [1 + |\nabla G_\sigma * \mathbf{I}|^2]^{-1}$ and G_σ being a Gaussian smoothing filter with standard deviation σ. Using the piecewise constant or Gaussian intensity-based Mumford–Shah functional with some loss function l and geodesic active contour penalty, the curve evolution equation following the principle of variational steepest descent can be written as

$$\frac{\partial \Phi(x)}{\partial t} = \delta(\Phi(x))\Big[\alpha \mathrm{div}\Big(g(x)\frac{\nabla \Phi(x)}{|\nabla \Phi(x)|}\Big) \tag{9}$$

$$-\lambda_1 l(\mathbf{I}(x), \theta_1) + \lambda_2 l(\mathbf{I}(x), \theta_2)\Big], \tag{10}$$

with $\mathrm{div}(\cdot)$ denoting the divergence operator and $d(z) = \frac{\partial}{\partial z} H(z)$ being the Dirac-delta function.[2] Please note that the update of parameters θ_i and the evolution of Φ are interleaved. See [12, 13] for a more detailed algorithmic treatment.

For stabilizing the evolution process, the surface Φ is often explicitly converted to a signed distance function. Here, we follow the approach of [15] and incorporate an additional term which penalizes surfaces which are far away from being a signed distance function. We refer the interested reader to [12, 15] for more theoretical insights of the used double-well nonlinear diffusion penalty and further implementation details.

As initial surface Φ^0 1% of the pixels most likely belonging to a microbe are assigned negative values $\Phi(x) = -2$, the remaining 99% pixels are initialized with $\Phi(x) = +2$. The likelihood of being microbe is here estimated by a simple normal distribution over gray values learned on the training image.

[2] The Heaviside function $H(z)$ is in fact not differentiable at point $z = 0$. However, it can be shown that its *distributional* derivative is equal to the Dirac-delta function $\delta(z)$, which is equal to one at $z = 0$ and zero elsewhere.

6. EXPERIMENTS

In the following, methods presented in previous sections are utilized for pixel-wise microbe detection and evaluated on real-world images.

Performance Assessment

We use a database including 5 different microbe species with 40 up to 470 microbes per class for training and testing. Examples from the different microbe classes are displayed in Fig. 4.

While the same parameters are used for all microbe categories, recognition accuracy is measured within those categories. As performance measures, average recognition rate as well as the overlap between inferred microbe labeling and ground-truth microbe labeling are used. In our setting, a single image is used for training and another one for testing. To arrive at the final results, all measures are averaged over all possible combinations of training-testing pairs within each category.

Implementation Details

For semantic segmentation as presented in the second section, pixel-wise classification is done using a decision tree based Gaussian process classifier [16]. The resulting pseudo-likelihood on the segment level is also used as unary potential of the manual CRF, where pairwise potentials are proportional to the average edge strength between two neighboring segments. For the log-linear CRF, gray values, thresholded intensity values, and an integrity measure are used as input features $\mathbf{X}$. Inference was done with the UGM package of Mark Schmidt [17].

Discussion of Results

Final results are displayed in Fig. 5. It becomes apparent that there is no method outperforming all other ones for all measures and categories. However, the fraction of false positives can be clearly reduced when dependencies between segments are taken into account. This can be also seen in Fig. 6, where example results are visualized for the microbe class *E. coli*. While the CRF-based approach labels less segments wrongly as bacteria, both level set approaches lead to many false positives. Since for the latter, predicted

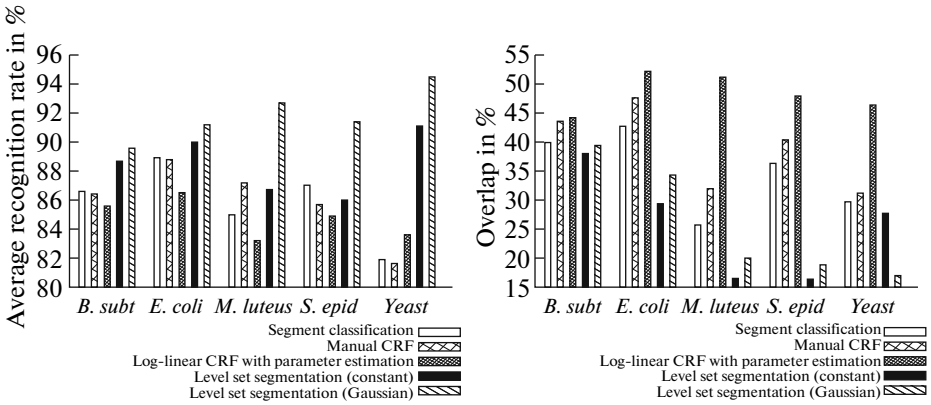

Fig. 5. Average recognition rate and overlap ratio achieved by all methods for each the bacteria types.

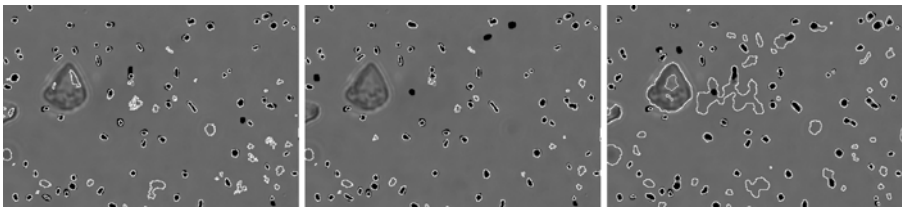

Fig. 6. Segmentation result using (left) intensity-based segment classification [5] with post-processing; (center) employing Graph-Gut based dependencies between segments; (right) and using a binary Mumford–Shah level set segmentation with Gaussian intensity distribution and geodesic active contour penalty. Ground-truth regions are filled with black color and the automatic results are shown by white contours.

microbe patches often include both classes, a post-processing step is often counterproductive. By comparing the two types of level set methods used in this work, we observed that Gaussian intensity models generally lead to more accurate results than constant intensity models.

The overall analysis hence suggests that segment-based models are to be preferred for detecting microorganisms in complex media. Especially CRFs seem to be suitable for tasks, which require a low number of false positives.

7. CONCLUSIONS AND FURTHER WORK

This paper focused on an in-depth analysis and evaluation of several computer vision and machine learning techniques applied to the task of pixel-wise microbe detection. We compared several semantic segmentation techniques such as segment-based classification using Conditional Random Fields and region-based level set segmentation. These techniques show a good performance on our challenging dataset when learned from labeled data. However, their usage in life sciences is still not common and established. In our opinion, the results of this paper can serve as a good guideline to select suitable algorithms for microbe segmentation which allow for learning from specific datasets.

Future research will focus on studying active learning methods to reduce the manual effort needed to label training images. Another interesting topic would be a comparison of different approximation methods for parameter estimation in Conditional Random Fields. Furthermore, a large-scale public dataset and common evaluation strategies would support application–driven research in this area.

ACKNOWLEDGMENTS

This work was partially funded by the TMBWK ProExzellenz project “MikroPlex” (PE113-1) and by the TMBWK ProExzellenz project “Graduate School on Image Processing and Image Interpretation.” We also would like to thank Prof. Erika Kothe for acquiring the microscopy images.

REFERENCES

1. X. Wu and S. K. Shah, “A bottom-up and top-down model for cell segmentation using multispectral data,” in *Proc. IEEE Int. Conf. on Biomedical Imaging* (Rotterdam, 2010), pp. 592–595.

2. A. Gelas, K. Mosaliganti, A. Gouaillard, L. Souhait, R. Noche, N. Obholzer, and S. G. Megason, “Variational level-set with gaussian shape model for cell seg-

mentation," in *Proc. IEEE Int. Conf. on Image Processing* (Cairo, 2009), pp. 1089–1092.

3. A. Pinidiyaarachchi and C. Wöhlby, "Seeded watersheds for combined segmentation and tracking of cells," in *Proc. Int. Conf. on Image Analysis and Processing* (Genoa, 2005), pp. 336–343.
4. M. Krause, P. Rösch, B. Radt, and J. Popp, "Localizing and identifying living bacteria in an abiotic environment by a combination of raman and fluorescence microscopy," Anal. Chem. **80**, 8568–8575 (2008).
5. G. Csurka and F. Perronnin, "A simple high performance approach to semantic segmentation," in *Proc. British Machine Vision Conf.* (Leeds, 2008), pp. 213–222.
6. B. Fröhlich, E. Rodner, and J. Denzler, "A fast approach for pixelwise labeling of facade images," in *Proc. Int. Conf. on Pattern Recognition* (San Francisco, 2010), Vols. 7, 8, pp. 3029–3032.
7. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Anal. Mach. Intell. **24** (5), 603–619 (2002).
8. Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," IEEE Trans. Pattern Anal. Mach. Intell. **23**, 1222–1239 (2001).
9. M. Kass, A. Witkin, and D. Terzopoulus, "Snakes: active contour models," Int. J. Comput. Vision **1** (4), 321–331 (1987).
10. S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations," J. Comput. Phys. **79**, 12–49 (1988).
11. D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," Commun. Pure Appl. Math. **42**, pp. 557–685 (1989).
12. T. F. Chan and L. A. Vese, "Active contours without edges," IEEE Trans. Image Processing **10**, 266–277 (2001).
13. D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion, and shape," Int. J. Comput. Vision **72**, 195–215 (2007).
14. V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," Int. J. Comput. Vision **22**, 61–79 (1997). http://dl.acm.org/citation.cfm?id=250488.250495
15. C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," IEEE Trans. Image Processing **19**, 3243–3254 (2010).
16. B. Fröhlich, E. Rodner, M. Kemmler, and J. Denzler, "Efficient Gaussian process classification using random decision forests," Pattern Recogn. Image Anal. **21**, 184–187 (2011). doi: 10.1134/S1054661811020337.
17. M. Schmidt, UGM: A matlab toolbox for probabilistic undirected graphical models (2007). http://www.di.ens.fr/mschmidt/Software/UGM.html

Michael Kemmler, born January 31, 1983, received the Diploma degree in Computer Science with honors in 2009 from the Friedlich Schiller University of Jena, Germany. As a PhD student of the Jena Graduate School for Microbial Communication, he is currently pursuing his studies under supervision of Joachim Denzler at the Chair for Computer Vision, University of Jena. His research interests are in the area of machine learning, object recognition and bioinformatics, including kernel methods, visual image and scene classification as well as bacterial classification.

Björn Fröhlich, born September 10, 1984, earned the degree "Diplom-Informatiker" from the Friedrich Schiller University of Jena in the year 2009. He is currently a holder of a scholarship in the Graduate School on Image Processing and Image Interpretation from the Free State of Thuringia (Germany) and a PhD student at the chair of Computer Vision, Institute of Computer Science, Friedrich Schiller University in Jena. Research interests are focused on object recognition and image segmentation.

Erik Rodner, born May 22, 1983 earned the Diploma degree in Computer Science with honours in 2007 from the University of Jena, Germany. He pursued and received his Ph. D. in 2011 with summa cum laude for his work on learning with few examples, which was done under supervision of Joachim Denzler at the computer vision research group of the University of Jena. Erik is currently continuing his research as a postdoctoral researcher. His research interests include kernel methods, visual object discovery, rare animals, and scene understanding.

Joachim Denzler, born April 16, 1967, earned the degrees "Diplom–Informatiker", "Dr.-Ing.," and "Habilitation" from the University of Erlangen in the years 1992, 1997, and 2003, respectively. Currently, he holds a position of full professor for computer science and is head of the Chair for Computer Vision, Faculty of Mathematics and Informatics, Friedrich–Schiller-University of Jena. His research interests comprise active computer vision, object recognition and tracking, 3D reconstruction, and plenoptic modeling, as well as computer vision for autonomous systems. He is author and coauthor of over 90 journal papers and technical articles. He is a member of the IEEE, IEEE computer society, DAGM, and GI. For his work on object tracking, plenoptic modeling, and active object recognition and state estimation, he was awarded the DAGM best aper awards in 1996, 1999, and 2001, respectively.