

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Optimal solutions for semantic image decomposition

Daniel Cremers

Department of Computer Science, TU Munich, Germany

ARTICLE INFO

Keywords:

Optimization
Efficient algorithms
Convexity
Semantic labeling

ABSTRACT

Bridging the gap between low-level and high-level image analysis has been a central challenge in computer vision throughout the last decades. In this article I will point out a number of recent developments in low-level image analysis which open up new possibilities to bring together concepts of high-level and low-level vision. The key observation is that numerous multi-label optimization problems can nowadays be efficiently solved in a near-optimal manner, using either graph-theoretic algorithms or convex relaxation techniques. Moreover, higher-level semantic knowledge can be learned and imposed on the basis of such multi-label formulations.

© 2012 Elsevier B.V. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Combining low-level vision...

Starting in the 1980s, researchers have tackled the image segmentation problem by means of energy minimization approaches [1,12]. While early approaches were generally not convex and respective algorithms would only compute locally optimal solutions, in recent years researchers have developed algorithms to compute optimal or near optimal solutions for respective energies using either graph-theoretic approaches [2,7] or convex relaxation techniques [3,4,10,15]. The underlying energies typically take into account local color information and aim at grouping regions of coherent color information, possibly enhanced with interactive user input indicating the rough location of objects of interest. By now, respective methods allow to separate objects of interest in rather challenging images, despite similar colors of object and background and strong variation of the illumination — see Fig. 1.

2. ...with semantic knowledge

Somewhat independent from the above developments in low-level image analysis, researchers have developed algorithms for high-level image analysis which allow to detect and recognize objects in images and even allow to perform an entire semantic scene analysis. Rather than modeling the color variations on a pixel-level they compute histograms of sparse features which are then related to respective features of previously observed objects [5]. Respective methods exhibit excellent performance on challenging high-level tasks. Yet the choice of features is generally somewhat heuristic and computed solutions typically do not come with a notion of statistical

optimality with respect to the original image data, nor do they provide a per-pixel semantic decomposition of images.

In contrast, the framework of multi-label optimization allows to perform semantic image parsing on a per-pixel level with higher-level knowledge. Fig. 2 shows a recent example where the multi-label optimization process was enhanced with a statistical prior on label co-occurrence [9]. Fig. 3 shows an example of multi-label segmentation with label ordering constraints [6,11,14]. In my view, the fusion of low-level and high-level aspects of visual processing on the basis of efficient multi-label optimization methods bears great potential for future research in computer vision.

References

- [1] A. Blake, A. Zisserman, Visual Reconstruction, MIT Press, 1987.
- [2] Y. Boykov, O. Veksler, R. Zabih, Markov random fields with efficient approximations, In: Proc. IEEE Conf. on Comp. Vision Patt. Recog. (CVPR'98), Santa Barbara, California, 1998, pp. 648–655.
- [3] A. Chambolle, D. Cremers, T. Pock, A convex approach for computing minimal partitions, In: Technical report TR-2008-05, Dept. of Computer Science, University of Bonn, Bonn, Germany, November 2008.
- [4] T. Chan, S. Esedoğlu, M. Nikolova, Algorithms for finding global minimizers of image segmentation and denoising models, SIAM J. Appl. Math. 66 (5) (2006) 1632–1648.
- [5] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, In: Int. Conf. on Computer Vision and Pattern Recognition (CVPR), 2005, pp. 886–893.
- [6] P.F. Felzenszwalb, O. Veksler, Tiered scene labeling with dynamic programming, In: Int. Conf. on Computer Vision and Pattern Recognition (CVPR), 2010, pp. 3097–3104.
- [7] D.M. Greig, B.T. Porteous, A.H. Seheult, Exact maximum *a posteriori* estimation for binary images, J. Roy. Stat. Soc. Ser. B 51 (2) (1989) 271–279.
- [8] M. Klodt, D. Cremers, A convex framework for image segmentation with moment constraints, In: IEEE Int. Conf. on Computer Vision (ICCV), 2011.
- [9] L. Ladicky, C. Russell, P. Kohli, P.H. Torr, Graph cut based inference with co-occurrence statistics, In: Europ. Conf. on Computer Vision, 2008.
- [10] J. Lellmann, J. Kappes, J. Yuan, F. Becker, C. Schnörr, Convex multi-class image labeling by simplex-constrained total variation, In: Tech. rep., IPA, HCL, Dept. Of Mathematics and Computer Science, Univ. Heidelberg, October 2008.

* This paper has been recommended for acceptance by special issue Opinions Editor Sinisa Todorovic.

E-mail address: cremers@tum.de.

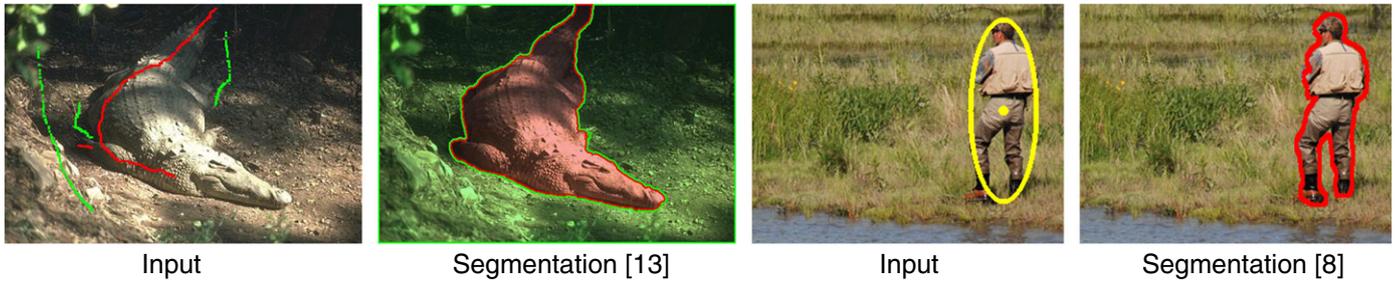


Fig. 1. Interactive segmentations obtained using space-varying color models (left) and low-order moment constraints (right) [8,13].



Fig. 2. Semantic segmentation obtained using label co-occurrence statistics.

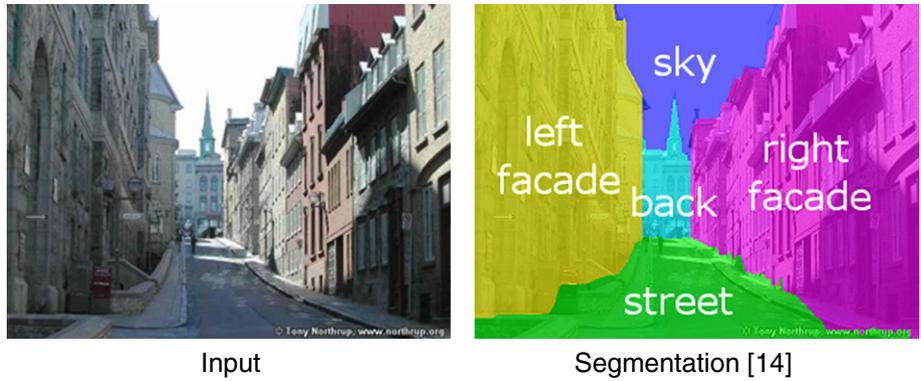


Fig. 3. Semantic segmentation obtained using ordering constraints.

[11] X. Liu, O. Veksler, J. Samarabandu, Order-preserving moves for graph-cut-based optimization, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (2010) 1182–1196.
 [12] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Commun. Pure. Appl. Math.* 42 (1989) 577–685.
 [13] C. Nieuwenhuis, D. Cremers, Spatially varying color distributions for interactive multiregion segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.*, to appear.
 [14] E. Strelakovski, D. Cremers, Generalized ordering constraints for multilabel optimization, In: *IEEE Int. Conf. on Computer Vision (ICCV)*, 2011.
 [15] C. Zach, D. Gallup, J.-M. Frahm, M. Niethammer, Fast global labeling for real-time stereo using multiple plane sweeps, In: *Workshop on Vision, Modeling and Visualization*, October 2008.