

TV-Based Multi-Label Image Segmentation with Label Cost Prior

Jing Yuan
 cn.yuanjing@gmail.com
 Yuri Boykov
 yboykov@gmail.com

Department of Computer Science
 University of Western Ontario
 London, Ontario, Canada, N6A 5B7

Abstract

This paper studies image segmentation based on the minimum description length (MDL) functional combining spatial regularization with a penalty for the number of distinct segments, a.k.a. *label cost prior*. Continuous MDL-based segmentation functionals were introduced in [19]. We propose a convex relaxation approach for optimizing MDL criterion that leads to a globally optimal solution.

As common in recent continuous convex formulations [6, 5], we use the total-variation functional to encode spatial regularity of segmentation boundaries. To the best of our knowledge, we are the first to demonstrate that the *label cost prior* can be also addressed within a continuous convex framework. The second-order cone programming algorithm is applied to tackle such nonsmooth convex energy functional. The experiments validate the proposed approach and theoretical results.

1 Introduction

The minimum description length principle (MDL) is an important information-theoretic concept. It states that any regularity in a given set of data can be used to compress the data, i.e. to describe it using fewer symbols than needed to describe the data literally [19]. Zhu and Yuille [6] proposed to segment images based on the continuous formulation of MDL principle, which boils down to the minimization of the following energy function:

$$\min_{\Omega_i} \sum_{i=1}^n \left\{ \int_{\Omega_i} \rho(\ell_i, x) dx + \lambda \int_{\partial\Omega_i} ds \right\} + \gamma M, \quad (1)$$

where Ω_i , $i = 1, \dots, n$, are homogeneous segments corresponding to n models/labels ℓ_i , $M = \#\{1 \leq i \leq n \mid \Omega_i \neq \emptyset\}$ is the number of nonempty segments, and data fidelity function $\rho(\ell_i, x) = -\log P(I_x \mid \ell_i)$ is a negative log-likelihood for model ℓ_i at pixel x . The second term in (1) describes the total perimeter of segments and favours spatially regular segments with minimum length boundary. Constants λ and γ describe relative weights of spatial regularity and label cost prior, correspondingly. Zhu and Yuille applied a local searching method, namely region competition, to approximate the highly nonconvex optimization problem (1). Their method converges to a local minimum.

In this work, we use standard total-variation (TV) based approach to representing segment boundaries, which allows to rewrite equation (1) as follows:

$$\min_{u_i(x) \in \{0,1\}} \sum_{i=1}^n \left\{ \int_{\Omega} u_i(x) \rho(l_i, x) dx + \lambda \int_{\Omega} |\nabla u_i| dx \right\} + \gamma M, \quad \text{s.t.} \quad \sum_{i=1}^n u_i(x) = 1, \quad \forall x \in \Omega \quad (2)$$

where u_i are indicator functions over x such that $\Omega_i = \{x \in \Omega | u_i(x) = 1\}$ and M is the number of appearing models $M = \#\{1 \leq i \leq n | u_i \neq 0\}$. We propose to solve (2) by relaxing the integer constraints $u_i(x) \in \{0, 1\}$ and reformulating the label cost term as a convex functional as follows:

$$\min_{u(x) \in S, x \in \Omega} \sum_{i=1}^n \left\{ \int_{\Omega} u_i(x) \rho(l_i, x) dx + \lambda \int_{\Omega} |\nabla u_i(x)| dx + \gamma \max_{x \in \Omega} u_i(x) \right\}. \quad (3)$$

where S is the pixelwise simplex constraint for the labeling functions. The third term in (3) is the infinity norm of $u_i(x)$ providing a convex analogue for the label cost term in (1) and (2). Energy function (3) is convex and can be globally optimized.

Without the label cost term, thresholding the solution of the convex relaxation u_i gives the global minimum for energy (2) in case of $n = 2$ [30]. The case of more than 2 labels is still debated, but many practical experiments in imaging suggest that thresholding gives a good approximation of the global minima for the original binary valued optimization problem. Our experiments provide empirical evidence that thresholding works well also in our case with label costs.

The effectiveness of the proposed convex label cost prior of (3) can be shown by a simple example (see Fig. 1): the method without considering the label cost term fails to discover the reasonable result, i.e. valid segment boundaries and proper number of segments. In contrast, the proposed approach based on (3), by adding both smoothness and label cost terms, captures the optimal segmentation result with reasonable boundaries and number of segments.

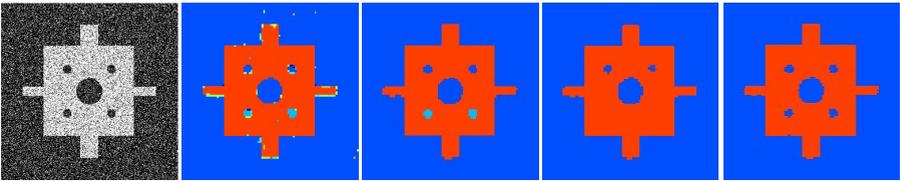


Figure 1: **From left to right:** 1st figure shows the given image to be labeled by 11 labels, associated to gray-scales $\{0, 0.1, 0.2, \dots, 1\}$. The 2nd, 3rd and 4th figures show the segmenting results without the proposed label cost prior, the total-variation penalty parameter is set to $\lambda = 0.07, 0.1, 0.2$ respectively. Obviously, both the number and edges of appearing segments are biased, such that increasing the penalty of the total-variation term helps to reduce the number of appearing labels but fails to recover the true segment edges! The 5th figure shows the labeling result computed by the proposed approach (3) with the label cost prior, where $\lambda = 0.07$ and $\gamma = 10$. The result gives much reasonable segments with 2 appearing labels and the correct segmentation boundaries as well.

1.1 Contributions

We generalize our main contributions as follows: we study the MDL based image segmentation and propose its novel convex relaxation model, which can be applied to solve the nonconvex MDL segmentation problem in an efficient and global way. More specially, we model the penalty of appearance number, i.e. the label cost prior, by the convex infinity norm instead. This is new. We propose the second order programming based algorithm to build up an effective scheme for the proposed convex MDL based image segmentation.

2 Previous Works

2.1 MDL Approaches

For a detailed review of MDL, we refer to [27]. In fact, such requirements of model reduction has been considered in model fitting problems for a long history, *e.g.* image segmentation [39], motion segmentation [26, 37] etc. The introduction to label costs amounts the well-studied uncapacitated location problem (UFL) [24]. In [8, 23], MDL principle was applied in the evolution of level sets to assist merges. Recently, an α -expansion combined with MDL was first suggested in Hoiem et al [20], with a specific application in object recognition; Delong *et al.* [16] independently extended α -expansion to handle similar but more general energies.

2.2 MDL Based Image Segmentation

In this paper, we consider the MDL principle to segment an image. The image to be segmented can often be represented by parametric models associated with labels $l_i, i = 1, \dots, n$, *e.g.* the piecewise smooth image function of the Mumford-Shah model [29] and of the weak membrane model by Blake and Zisserman [9], or the simplest piecewise constant image function introduced by Chan and Vese [13] where labels are directly related to gray values. To achieve a reasonable image segmentation, two issues are often emphasized: the first one concerns how well each image pixel fits the assigned image model represented by parameters associated to labels $l_i, i = 1, \dots, n$; the second one takes care of the spatial coherence of assigned segments, *e.g.* the Potts model with the minimal boundary length [2, 6, 31, 34], random walks [17, 18] or shortest-path [9, 15]. Applying the optimal length criterion, these two considerations can be adopted in the total-variation based energy function:

$$\min_{u_i(x) \in \{0,1\}} \sum_{i=1}^n \int_{\Omega} \{u_i(x)\rho(l_i, x) + \lambda |\nabla u_i|\} dx, \quad \text{s.t.} \quad \sum_{i=1}^n u_i(x) = 1, \quad \forall x \in \Omega, \quad (4)$$

where the first term evaluates the cost of model assignment at each image pixel and the sum of total variations of the labeling functions $u_i(x), i = 1, \dots, n$, measures the total perimeter.

The third principle of image segmentation is motivated by the so-called Akaike information criterion (AIC) [10] under the perspective of information theory. One may also use other criterion like BIC [8, 33]. Such AIC principle leads to the minimization of the energy

$$E = -2\log L + 2C \quad (5)$$

where $\log L$ gives the model fitting cost and C measures the complexity of assigned models. In most cases, C is considered to be proportional to the total number of present segments,

i.e. one seek an segmentation result with the least labels. In contrast to vast works of image segmentation based on the two optimality considerations, relatively few of works were contributed to this rule [8, 23, 24, 59].

The combination of the three principles motivates the energy function for image segmentation (2) considered in this paper. Obviously, increasing the number M of 'active' segments results in an extra cost proportional to the related penalty parameter $\gamma > 0$, which gives the label cost prior. The global optimum of the energy function (2) seeks the optimal image labeling functions $u_i(x) \in \{0, 1\}$, $i = 1, \dots, n$, which concerns both the minimal length and minmal number of appearing labels. This result possesses optimalities of both geometry and model simplicity.

2.3 Convex TV-Based Optimization Approaches

Image segmentation and labeling subject to the minimal total-perimeter has been intensively studied, e.g. [6, 11, 12, 50, 55] etc. Current studies [10, 11, 25, 51] of total-variation based segmentation approaches proposed to solve (4) in a convex relaxation way. This is in contrast to the traditional level set and active contour based methods [56] which are formulated in a highly nonconvex way. By the relaxation of such binary constraints to $u_i(x) \in [0, 1]$, $i = 1, \dots, n$, the nonconvex image labeling model (4) boils down to the convex minimal image partition problem with multiple labels:

$$\min_{u(x) \in S} \sum_{i=1}^n \left\{ \langle u_i, d_i \rangle + \lambda \int_{\Omega} |\nabla u_i(x)| dx \right\} \quad (6)$$

where $u(x) = (u_1(x), \dots, u_n(x))^T$ and S denotes the pixelwise simplex constraint, i.e.

$$S := \{(u_1(x), \dots, u_n(x))^T \mid \sum_{i=1}^n u_i(x) = 1, \quad u_i(x) \geq 0; \quad \forall x \in \Omega\}.$$

In [58], Zach *et al.* proposed an alternating optimization method to solve (6) approximately. In [25], Lellmann *et al.* applied a Douglas-Rachford splitting approach to (6) with a variant of the total-variation term. Both proposed methods involve two substeps within one outer loop, where one substep is for exploring the pointwise simplex constraint $u(x) \in S$ and the other one is for tackling the total-variation term. In [10, 51], Pock *et al.* introduced a variant implementation of the constraint $u(x) \in S$, i.e. a tighter relaxation based on the multi-layered configuration, and this gives a more complex constraint on the concerning dual variable p to avoid multiple counting.

In contrast to the works of [10, 25, 51, 58] which tried to tackle the labeling function u of (6) in a direct way, Bae *et al.* [1] proposed to solve the convex relaxed Potts problem (6) based on its equivalent dual formulation. The highly nonsmooth dual formulation can then be efficiently approximated by a smooth convex energy function.

3 A Convex Relaxation Approach

Directly solving the proposed optimization problem (2) is very hard due to its nonconvexity and nonlinearity: the labeling functions $u_i(x)$, $i = 1, \dots, n$, are binary constrained and the number M is even unknown. We propose a convex relaxation approach to (2) in this paper, like [1, 25, 51] etc.

3.1 Convex Formulation with Label Cost Prior

Given n labels $\{l_1, \dots, l_n\}$, we introduce the auxilliary indicating function $y_i \in \{0, 1\}$, $i = 1, \dots, n$, which indicates if the label l_i appears in the segmentation result: $y_i = 1$ when l_i appears in the final segments and $y_i = 0$ otherwise. Therefore, we have

$$\sum_{i=1}^n y_i = M. \tag{7}$$

Clearly, when the maximum of the labeling function $u_k(x) \in \{0, 1\}$, $1 \leq k \leq n$, on the whole image domain Ω is 1, the label l_k presents in the image segmentation result. Hence the variable y_i is related to the labeling function $u_i(x) \in \{0, 1\}$, $i = 1, \dots, n$, such that

$$\max_{x \in \Omega} u_i(x) = y_i, \quad i = 1, \dots, n. \tag{8}$$

In view of (7) and (8), we can reformulate (2) by

$$\min_{u_i(x) \in \{0,1\}} \sum_{i=1}^n \left\{ \int_{\Omega} u_i(x) \rho(l_i, x) dx + \lambda \int_{\Omega} |\nabla u_i| dx + \gamma y_i \right\} \tag{9}$$

$$\text{s.t.} \quad \sum_{i=1}^n u_i(x) = 1, \quad u_i(x) \leq y_i, \quad \forall x \in \Omega; \tag{10}$$

or equivalently

$$\min_{u_i(x) \in \{0,1\}} \sum_{i=1}^n \left\{ \int_{\Omega} u_i(x) \rho(l_i, x) dx + \lambda \int_{\Omega} |\nabla u_i| dx + \gamma \max_{x \in \Omega} u_i(x) \right\} \tag{11}$$

$$\text{s.t.} \quad \sum_{i=1}^n u_i(x) = 1, \quad \forall x \in \Omega.$$

We relax the binary constraint of the labeling functions $u_i(x) \in \{0, 1\}$, $i = 1 \dots n$, to $u(x) := (u_1(x), \dots, u_n(x))^T \in S$ where S denotes the pointwise simplex constraint of the vector in \mathbb{R}^n , i.e.

$$S := \{(u_1(x), \dots, u_n(x))^T \mid \sum_{i=1}^n u_i(x) = 1, u_i(x) \geq 0\}. \tag{12}$$

Therefore, the nonconvex optimization problem (11) is written as the continuous convex optimization problem (3), which can be solved globally. Hence (3) gives rise to the convex relaxed approach to the proposed image segmentation (2). Its third term penalizes the infinity norm of each labeling function $u_i(x)$, $i = 1 \dots n$, which corresponds to the label cost prior.

In the same manner, (9) can be also formulated as:

$$\min_{u(x)} \sum_{i=1}^n \left\{ \int_{\Omega} u_i(x) \rho(l_i, x) dx + \lambda \int_{\Omega} |\nabla u_i| dx + \gamma y_i \right\} \tag{13}$$

$$\text{s.t.} \quad \sum_{i=1}^n u_i(x) = 1, \quad 0 \leq u_i(x) \leq y_i, \quad \forall x \in \Omega, \quad i = 1, \dots, n. \tag{14}$$

3.2 Equivalent Convex Formulations

It is well-known that the total-variation functions can be equivalently written as [\[14\]](#), [\[28\]](#)

$$\lambda \int_{\Omega} |\nabla u_i(x)| dx = \max_{p_i(x)} \langle \operatorname{div} p_i, u_i \rangle, \quad i = 1, \dots, n \quad (15)$$

where the functions $p_i(x)$, $i = 1 \dots n$, are constrained by the convex set C_p^λ

$$C_p^\lambda := \{p(x) \mid |p(x)| \leq \lambda, \forall x \in \Omega; \quad p_n = 0\}. \quad (16)$$

For the infinity norm of the labeling function $u_i(x) \in [0, 1]$, $i = 1 \dots n$, we have

$$\gamma \max_{x \in \Omega} u_i(x) = \max_{v_i(x)} \langle v_i, u_i \rangle, \quad i = 1, \dots, n \quad (17)$$

where the functions $v_i(x)$, $i = 1 \dots n$, are constrained by the convex set C_v^γ

$$C_v^\gamma := \{v(x) \mid \int_{\Omega} v(x) = \gamma; \quad v(x) \geq 0, \forall x \in \Omega\}. \quad (18)$$

Observe [\(15\)](#) and [\(17\)](#), [\(3\)](#) can, therefore, be equivalently formulated as

$$\begin{aligned} \min_{u(x) \in \mathcal{S}, x \in \Omega} \max_{p, v} \sum_{i=1}^n \langle \rho(l_i, x) + \operatorname{div} p_i + v_i, u_i \rangle, \\ \text{s.t. } p_i(x) \in C_p^\lambda, \quad v_i(x) \in C_v^\gamma; \quad i = 1, \dots, n. \end{aligned} \quad (19)$$

3.2.1 Equivalent Dual Formulation

We consider the same technique as [\[9\]](#):

Given any vector $a = (a_1, \dots, a_n)^\top$, it is easy to see that

$$\min_i a_i = \min_{(q_1, \dots, q_n)^\top \in \mathcal{S}} \sum_{i=1}^n q_i a_i.$$

It follows that the optimization [\(19\)](#) over $(u_1(x), \dots, u_n(x)) \in \mathcal{S}$, $\forall x \in \Omega$, gives rise to an identical formulation of [\(19\)](#) as

$$\begin{aligned} \max_{p, v} \int_{\Omega} \min_i (\rho(l_i, x) + \operatorname{div} p_i(x) + v_i(x)) dx, \\ \text{s.t. } p_i(x) \in C_p^\lambda, \quad v_i(x) \in C_v^\gamma; \quad i = 1, \dots, n. \end{aligned} \quad (20)$$

We call [\(20\)](#) as the *equivalent dual model* of [\(3\)](#) in this paper.

4 Implementation and Algorithm

In this paper, we apply the second order cone programming (SOCP) [\[32\]](#) to solve the highly nonsmooth convex optimization problem [\(3\)](#) and [\(13\)](#). SOCP provides an effective way to model and solve such complex convex optimization problems globally and accurately.

4.1 Discretization

Based on the mimetic finite difference method [21, 22], we develop the cell-wise discretization scheme: each image pixel corresponds to a square cell with the equal size which is assumed to be 1; a scalar function $u(x)$ is mapped to the discrete scalar space H_V , i.e. defined at the center of each cell and $u(\alpha, \beta)$ is given at the cell $C_{\alpha, \beta}$; for each cell, there are four adjacent cells: top, bottom, right and left, which is similar to a 4-connected graph construction; hence, the gradient of the image function $u(\alpha, \beta)$ is given in the discrete vector space H_S [21]; in addition, the inner products are defined for the linear spaces H_V and H_S as [21, 22].

Based on this discretization scheme, the total variation term is evaluated over each cell $C_{\alpha, \beta}$, i.e.

$$\text{TV}_d(u) := \sum_{C_{\alpha, \beta}} \sqrt{((u^L - u^C)^2 + (u^R - u^C)^2 + (u^T - u^C)^2 + (u^B - u^C)^2) / 2} \quad (21)$$

where at the cell $C_{\alpha, \beta}$, $u^C = u(\alpha, \beta)$ and $u^{B, T, R, L}$ are the values defined at the four adjacent cells of $C_{\alpha, \beta}$.

4.2 Second Order Cone Programming

We formulate (3) in the dual form of SOCP, i.e.

$$\begin{aligned} \max_x \quad & b^T x \\ \text{s.t.} \quad & c - A^T x \in \mathbf{K} \end{aligned} \quad (22)$$

where \mathbf{K} denotes the Lorentz cone.

The constraint $u(x) \in S$ of (3) amounts to the formulation (12) $\forall x \in \Omega$, which can be easily represented by SOCP as the simple positive orthant constraint along with a linear equality constraint. The convex label cost term can also be implemented based on the introduction of auxiliary variables y_i of (13), along with the linear inequalities: $u_i(x) \leq y_i, \quad \forall x \in \Omega, \quad i = 1, \dots, n$.

The total-variation term of (21) can be implemented in the form of Lorentz cones, where

$$(t_i(x); u_i^L(x) - u_i^C(x), u_i^R(x) - u_i^C(x), u_i^T(x) - u_i^C(x), u_i^B(x) - u_i^C(x)) \in \mathbf{K}, \quad \forall x \in \Omega, \quad i = 1, \dots, n,$$

i.e.

$$t_i(x) \geq \sqrt{((u_i^L(x) - u_i^C(x))^2 + (u_i^R(x) - u_i^C(x))^2 + (u_i^T(x) - u_i^C(x))^2 + (u_i^B(x) - u_i^C(x))^2) / 2}.$$

Through these formulations, we encode (3) by Sedumi [82].

5 Experiments

Fig. 2 shows the experiment to validate the theoretical results and the effectiveness of the convex label cost prior. The experiment tries to assigne the given image f with 11 labels $l_1 \dots l_{11}$, which are the gray-scale values. The data term is measured by

$$d_i(x) = |f(x) - l_i|, \quad i = 1, \dots, 11; \quad \forall x \in \Omega.$$

The computed label information is colorized by matlab, from blue to red corresponds to different label values from 0 to 1. As shown in Fig. 2, the total appearing number of labels can be evidently reduced by the proposed label cost prior, without taking large length penalty which corresponds to the large value of λ and causes the distortion of label regions. The results given at the first row are computed by vanishing the label cost prior, i.e. $\gamma = 0$; the results given at the second row are computed by vanishing the total-variation cost, i.e. $\lambda = 0$ (only the label cost is considered); the results given at the third row are computed by considering both the total-variation cost and the label cost. As shown by the experiments, the total number of appearing labels decreases but the region edges are distorted at the same time, when the total-variation penalty λ increases. Only applying the label cost works effectively to reduce the total number of labels, but loses spatial regularities. When both the spatial smoothness term and label cost term work, the total number of appearing labels decreases and the edges concerning present labeled regions are kept well simultaneously!

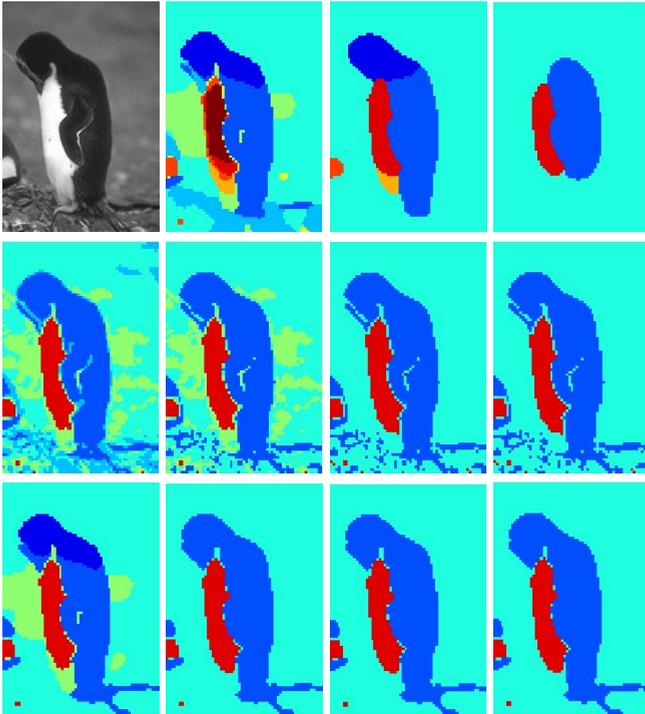


Figure 2: **At the first row (from left to right):** 1st figure shows the given image to be labeled by 11 labels $\{0, 0.1, \dots, 1\}$. 2nd - 4th figures show the labeling results without the proposed label cost prior, where the total-variation penalty parameter $\lambda = 0.05, 0.25, 0.65$ respectively and the result is colorized by matlab; different color is associated to different color. **At the second row (from left to right):** the figures show the result computed by the proposed approach only regularized by the label cost prior, i.e. without the total-variation term, where $\lambda = 0$ and the label-cost parameter $\gamma = 25, 43, 50, 100$. **At the third row (from left to right):** the figures show the result computed by the proposed MDL approach (3) with the label cost prior, where λ is fixed to be 0.05 and the label-cost parameter $\gamma = 10, 25, 50, 100$ respectively.

The images of experiments, shown in Fig. 3, come from the *Berkeley segmentation dataset and benchmark*. All the experiments take 11 labels associated to gray values. These experiments further confirm the effectiveness of the label cost prior.

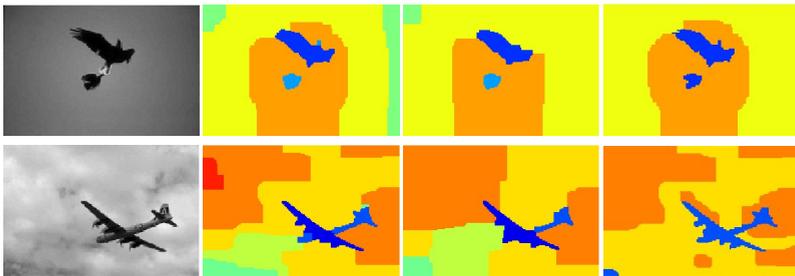


Figure 3: **At the first row (from left to right):** the input image $f(x)$, the computed label image when $\lambda = 0.1$ without label cost, the computed label image when $\lambda = 0.2$ without label cost, the result label image when $\lambda = 0.06$ with the label cost $\gamma = 50$. **At the second row (from left to right):** the input image $f(x)$, the computed label image when $\lambda = 0.1$ without label cost, the computed label image when $\lambda = 0.2$ without label cost, the result label image when $\lambda = 0.06$ with the label cost $\gamma = 100$.

6 Conclusions

We study the MDL based image segmentation in this work. More specially, we propose the convex relaxation approach to the considered nonconvex MDL formulation (2) so as to achieve its globally optimal approximation. In this regard, we suggest the new convex label cost function and implement the SOCP-based algorithm. Experiments validate the effectiveness and great advantages of such convex label cost term in image segmentation. However, a fast algorithm to the convex MDL based image segmentation (3) is still open, due to the highly nonsmoothness of such energy function. The authors will devote to this direction in future researches.

References

- [1] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Trans. Automatic Control*, AC-19:716–723, 1974. ISSN 0018-9286. System identification and time-series analysis.
- [2] E. Bae, J. Yuan, and X.C. Tai. Convex relaxation for multipartitioning problems using a dual approach. Technical report CAM09-75, UCLA, CAM, September 2009.
- [3] Xue Bai and Guillermo Sapiro. Geodesic matting: A framework for fast interactive image and video segmentation and matting. *International Journal of Computer Vision*, 82(2):113–132, 2009.
- [4] Andrew Blake and Andrew Zisserman. *Visual Reconstruction*. MIT Press, 1987. ISBN 0-262-02271-0.

-
- [5] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on PAMI*, 26: 359–374, 2001.
- [6] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on PAMI*, 23:1222 – 1239, 2001.
- [7] X. Bresson, S. Esedoglu, P. Vandergheynst, J.P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, 2007.
- [8] T. Brox and J. Weickert. Level set segmentation with multiple regions. *IEEE Transactions on Image Processing*, 15(10):3213–3218, October 2006.
- [9] Kenneth P. Burnham and David R. Anderson. Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods Research*, 33(2):261–304, November 2004. ISSN 0049-1241.
- [10] A. Chambolle, D. Cremers, and T. Pock. A convex approach for computing minimal partitions. Technical report TR-2008-05, Dept. of Computer Science, University of Bonn, Bonn, Germany, November 2008.
- [11] Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1):89–97, January 2004.
- [12] T. Chan and L.A. Vese. *Active contours without edges*. *IEEE Image Proc.*, 10, pp. 266–277, 2001.
- [13] Tony E Chan and Luminita A. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. In *IEEE WORKSHOP ON VARIATIONAL AND LEVEL SET METHODS*, page 161. IEEE Computer Society, 2001.
- [14] Gerard Cornuejols, Marshall L. Fisher, and George L. Nemhauser. Location of bank accounts to optimize float: An analytic study of exact and approximate algorithms. *Management Science*, 23(8):789–810, 1977.
- [15] Antonio Criminisi, Toby Sharp, and Andrew Blake. Geos: Geodesic image segmentation. In *ECCV (1)*, pages 99–112, 2008.
- [16] Andrew DeLong, Anton Osokin, Hossam Isack, and Yuri Boykov. Fast approximate energy minimization with label costs. In *CVPR*, 2010.
- [17] Leo Grady. Multilabel random walker image segmentation using prior models. In *CVPR*, pages 763–770, 2005.
- [18] Leo Grady. Random walks for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(11):1768–1783, 2006.
- [19] Peter D. Gruenwald. *The Minimum Description Length Principle*, volume 1 of *MIT Press Books*. The MIT Press, October 2007.
- [20] D. Hoiem, C. Rother, and J. Winn. 3D LayoutCRF for Multi-View Object Class Recognition and Segmentation. In *CVPR*, 2007.

- [21] J. M. Hyman and M. J. Shashkov. Natural discretizations for the divergence, gradient, and curl on logically rectangular grids. *Comput. Math. Appl.*, 33(4):81–104, 1997. ISSN 0898-1221.
- [22] James M. Hyman, Mikhail Shashkov, A If, Only If, A Curl, and B Curl. Adjoint operators for the natural discretizations of the divergence, gradient, and curl on logically rectangular grids. *Appl. Numer. Math.*, 25:413–442, 1997.
- [23] Timor Kadir and Michael Brady. Unsupervised non-parametric region segmentation using level sets. In *ICCV*, pages 1267–1274, 2003.
- [24] Yvan G. Leclerc. Constructing simple stable descriptions for image partitioning. *International Journal of Computer Vision*, 3(1):73–102, 1989.
- [25] J. Lellmann, J. Kappes, J. Yuan, F. Becker, and C. Schnörr. Convex multi-class image labeling by simplex-constrained total variation. Technical report, HCI, IWR, Uni. Heidelberg, IWR, Uni. Heidelberg, November 2008.
- [26] Hongdong Li. Two-view motion segmentation from linear programming relaxation. In *CVPR*, 2007.
- [27] David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [28] Yves Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001. ISBN 0-8218-2920-3. The fifteenth Dean Jacqueline B. Lewis memorial lectures.
- [29] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
- [30] Mila Nikolova, Selim Esedoglu, and Tony F. Chan. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal on Applied Mathematics*, 66(5):1632–1648, 2006.
- [31] T. Pock, A. Chambolle, H. Bischof, and D. Cremers. A convex relaxation approach for computing minimal partitions. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, Florida, 2009.
- [32] Jos F. Sturm. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones (updated for version 1.05), 1998.
- [33] P. H. S. Torr. Geometric motion segmentation and model selection. *Phil. Trans. Royal Society of London A*, 356:1321–1340, 1998.
- [34] Markus Unger, Thomas Pock, Werner Trobin, Daniel Cremers, and Horst Bischof. H.: Tvseg - interactive total variation based image segmentation. In *BMVC*, 2008.
- [35] Olga Veksler. Image segmentation by nested cuts. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 339–344, 2000.

- [36] L. A. Vese and T. F. Chan. A new multiphase level set framework for image segmentation via the mumford and shah model. *International Journal of Computer Vision*, 50: 271–293, 2002.
- [37] René Vidal, Roberto Tron, and Richard Hartley. Multiframe motion segmentation with missing data using powerfactorization and gpca. *International Journal of Computer Vision*, 79(1):85–105, 2008.
- [38] C. Zach, D. Gallup, J.-M. Frahm, and M. Niethammer. Fast global labeling for real-time stereo using multiple plane sweeps. In *Vision, Modeling and Visualization Workshop (VMV)*, 2008.
- [39] Song Chun Zhu and Alan Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multi-band image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:884–900, 1996.