SIAM J. IMAGING SCIENCES
Vol. 5, No. 4, pp. 1333–1354
© 2012 Society for Industrial and Applied Mathematics

A Convex Max-Flow Approach to Distribution-Based Figure-Ground Separation*

Kumaradevan Punithakumar†, Jing Yuan‡, Ismail Ben Ayed§, Shuo Li§, and Yuri Boykov‡

Abstract. This study investigates a convex relaxation approach to figure-ground separation with a global distribution matching prior evaluated by the Bhattacharyya measure. The problem amounts to finding a region that most closely matches a known model distribution. It has been previously addressed by curve evolution, which leads to suboptimal and computationally intensive algorithms, or by graph cuts, which result in metrication errors. Solving a sequence of convex subproblems, the proposed relaxation is based on a novel bound of the Bhattacharyya measure which yields an algorithm robust to initial conditions. Furthermore, we propose a novel flow configuration that accounts for labeling-function variations, unlike existing configurations. This leads to a new max-flow formulation which is dual to the convex relaxed subproblems we obtained. We further prove that such a formulation yields exact and global solutions to the original, nonconvex subproblems. A comprehensive experimental evaluation on the Microsoft GrabCut database demonstrates that our approach yields improvements in optimality and accuracy over related recent methods.

Key words. image segmentation, convex optimization, max-flow min-cut, Bhattacharyya measure, distribution matching, iterative methods

AMS subject classifications. 68U10, 90C25

DOI. 10.1137/110850372

1. Introduction.

This study investigates the problem of partitioning an image into a figure and ground, so that the figure region is consistent with a known (a priori learned) distribution. The solution to this problem was recently shown to be very useful in the following computer vision tasks.

- *Cosegmentation of image pairs.* The problem consists of finding the same object (figure) in a pair of images and has attracted significant research attention in recent years [3, 29, 36, 37, 46, 51, 50]. It has been shown to be very useful in object recognition/image retrieval [19, 26, 46, 47], image editing [3], and summarization [22].
- *Interactive image segmentation.* Using minimal user inputs, e.g., bounding boxes or scribbles, interactive segmentation embeds clues on user intention, thereby limiting the space of possible solutions and dramatically easing the problem. It has been intensively investigated in recent years [9, 34, 40, 45, 49].
- *Segmentation with offline learning.* Segmenting groups of images that share similar properties occurs in important applications, e.g., in medical image analysis. In such scenarios,

*Received by the editors October 5, 2011; accepted for publication (in revised form) July 29, 2012; published electronically December 4, 2012.
http://www.siam.org/journals/siims/5-4/85037.html
†Corresponding author. Servier Virtual Cardiac Centre, Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton T6G 2B7, AB, Canada (punithak@ualberta.ca).
‡Department of Computer Science, University of Western Ontario, London N6A 5B7, ON, Canada (jyuan22@uwo.ca, yuri@csd.uwo.ca).
§GE Healthcare, London N6A 4V2, ON, Canada (ismail.benayed@ge.com, shuo.li@ge.com).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

offline learning of prior information from segmented training images is very useful [6, 18].

- *Tracking.* When tracking a target region within an image sequence, one can segment the current frame using information learned from the segmentation of a previous frame [5, 25, 55].

The problem amounts to optimizing a *global* measure of similarity/dissimilarity between distributions. Possible measures include the Bhattacharyya coefficient [4, 6, 5, 25, 40, 55], L_p norm of the difference between histograms [37, 46, 50], Wasserstein distance [38], and Kullback–Leibler divergence [2]. However, the Bhattacharyya coefficient has yielded outstanding performances recently [4, 40]. Furthermore, it has several desirable properties. First, it has a fixed (normalized) range which is convenient in many practical scenarios, e.g., in image retrieval and object recognition. Second, it is not affected by the choice of the bin width [1] (a proof of this property can be found in [1]). Third, it imposes a *metric* structure [20].

On the one hand, as several recent studies have shown [4, 6, 38, 35, 40, 46, 55, 25], global distribution-similarity measures can afford an accuracy unattainable with standard descriptions based on *pixelwise* information [9]. On the other hand, they generally lead to challenging optimization problems in both *continuous or variational* and *discrete or graph-based* settings.

In the continuous setting, the energy functional is defined over continuous variables and image domains [6, 25, 38, 55]. Active curve evolution driven by gradient-descent optimization has been the most prevalent and flexible choice in the literature for minimizing global functionals [2, 6, 25, 55]. However, such a choice suffers from local optima, slow convergence, and high sensitivities to initialization. Complex implementation in numerics is also a significant limitation of active curve methods.

In the discrete setting, the problem is stated as a label assignment where the input images are considered as graphs defined over the positional arrays [9, 11, 12]. In this respect, graph cut techniques have been of intense interest recently in computer vision because they can guarantee global optima in nearly real time [9]. Unfortunately, global energies have been generally avoided because they are not directly amenable to graph cut optimization. The recent works in [4, 37, 40, 46] are notable exceptions which use relaxations via bounds or approximations of the energy so as to befit graph cut optimization. These works led to significant improvements over active curves in regard to optimality, speed, and accuracy [4]. However, the well-known *grid bias (or metrication error)* is a limitation of such graph-based approaches [41]. Reducing metric artifacts can be done by increasing the number of neighboring graph nodes, but this may result in a heavy computation and memory load [32]. In practice, it is well known that graph cut algorithms can yield a state-of-the-art empirical performance in the context of 2-dimensional grids with 4-neighborhood systems. They use heuristics that efficiently handle sparse grids [11]. However, the efficiency of the algorithms may decrease when moving from 2- to 3-dimensional grids or when using denser (larger neighborhood) grids.

Continuous convex relaxation approaches share the advantages of both active curves and graph cuts and have recently attracted significant research attention in image partitioning problems. Recent notable studies have shown the potential of convex optimization in solving the classical piecewise constant Mumford–Shah segmentation model [28, 16, 53] and its multiphase (multiregion) variant [15, 33, 54], as well as a histogram-based segmentation using the Wasserstein distance [38]. Moreover, unlike graph cuts, the implementation of convex relaxation approaches can be parallelized [54] to reduce computational time. The recent study by

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

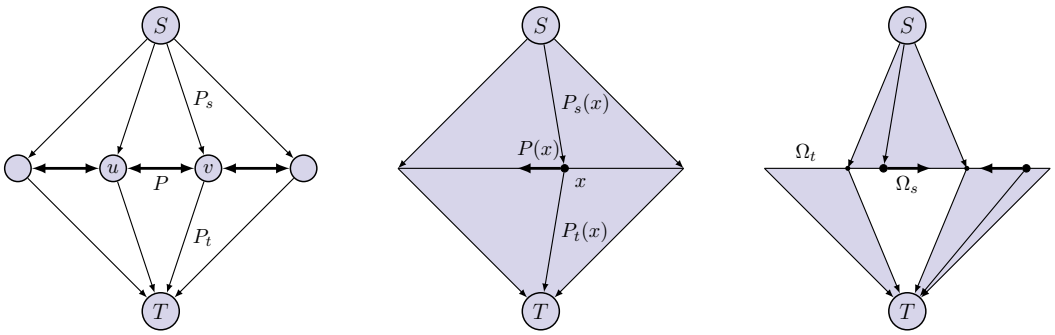


Figure 1. *Flow configurations. Left: spatially discrete (classical graph-based)* [9]. *Middle: spatially continuous* [52, 53]. *Right: the proposed spatially continuous configuration.*

Klodt et al. [31] compared discrete and continuous optimization techniques in computer vision. It showed that, in the case of 3-dimensional or higher-dimensional grids, convex relaxation approaches outperform graph cuts with regard to speed and accuracy.

This study investigates an iterative convex relaxation approach to figure-ground separation with a global distribution-matching prior evaluated by the Bhattacharyya measure. Here is a summary of the main contributions:

- Solving a sequence of convex subproblems, the proposed relaxation is based on a novel bound of the Bhattacharyya measure which can be viewed as a *generalization* of the bound we recently derived in our conference paper [4]. The bound in [4] assumes that, at each iteration, the new figure region is enclosed within the region obtained at the previous iteration (refer to the illustration in Figure 2(b)), an assumption which requires the initial region to include the target region. The proposed bound relaxes this assumption (refer to the illustration in Figure 2(a)), yielding an algorithm which is more robust to initial conditions (refer to the examples in Figure 3).
- We propose a novel flow configuration (a simple 1-dimensional example is depicted by the rightmost graph of Figure 1) which accounts for labeling-function variations, unlike existing configurations [52, 53] (see the middle graph of Figure 1). This leads to a new max-flow formulation which is dual to the convex relaxed subproblems we obtained. We further prove that such a formulation yields exact and global solutions to the original, nonconvex subproblems.
- A comprehensive experimental evaluation on the Microsoft GrabCut database (50 images with ground truth labelings) demonstrates that our approach yields improvements in optimality and accuracy over the Bhattacharyya measure graph cut (BMGC) method we recently presented in [4].

2. Preliminaries.

2.1. The basic max-flow and min-cut models. In this section, we review the dualities of max-flow and min-cut in both spatially discrete (graph) and continuous settings, based on the classical graph and flow configurations where flows stream from the source node s to each image node, and then from each image node to the sink (the leftmost and middle configurations in Figure 1 illustrate the discrete and continuous cases, respectively).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let $G = (\mathcal{V}, \mathcal{E})$ be a graph composed of a vertex set $\mathcal{V}$ and an edge set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. $\mathcal{V}$ includes all the nodes of the discretized image together with two additional terminals, a source s and a sink t. $\mathcal{E}$ includes all the links of the neighborhood system of image nodes, e.g., 4 or 8 connected, as well as the links between the terminal s or t and each image node $u \in \mathcal{V}\backslash\{s,t\}$ (see the leftmost configuration in Figure 1).

For such a graph, an s-t cut divides $\mathcal{V}$ into two disjoint subgraphs:

$$\mathcal{V} = \mathcal{V}_s \bigcup \mathcal{V}_t\,, \quad \mathcal{V}_s \cap \mathcal{V}_t = \emptyset$$

with $\mathcal{V}_s$ (respectively, $\mathcal{V}_t$) containing source s (respectively, sink t).

Let $C(e) \geq 0$ be the cost associated to each edge $e \in \mathcal{E}$. The energy of each s-t cut is the sum of all edge costs $C(e)$, where $e \in \mathcal{E}_{st} \subset \mathcal{E}$ links two nodes, one in $\mathcal{V}_s$ and the other in $\mathcal{V}_t$. On the one hand, the min-cut over $\mathcal{G}$ consists of finding a minimum-energy s-t cut, i.e.,

$$\min_{\mathcal{E}_{st} \subset \mathcal{E}} \sum_{e \in \mathcal{E}_{st}} C(e)\,. \tag{2.1}$$

On the other hand, one can regard each edge $e \in \mathcal{E}$ as a pipe and its edge cost $C(e)$ as the corresponding capacity. In this case, graph $\mathcal{G}$ gives a flow "network." Let $p(e)$ be the flow passing edge $e \in \mathcal{E}$. It is well known that the min-cut problem in (2.1) amounts to finding the maximal flow from s over the network [21, 24], i.e.,

$$\max_{p_s} \sum_{v \in \mathcal{V}\backslash\{s,t\}} p_s(v) \tag{2.2}$$

subject to the corresponding flow capacity constraints and flow conservation condition, i.e., the balance between incoming and outgoing flows at each node $v \in \mathcal{V}\backslash\{s,t\}$,

$$\left[\sum_{u \in \mathcal{N}(v)} p(e(u,v))\right] - p_s(v) + p_t(v) = 0, \tag{2.3}$$

where $\mathcal{N}(v)$ denotes the neighborhood set of v.

The recent studies by Yuan and coauthors [52, 53] proved that the equivalence between max-flow (2.2) and min-cut (2.1) holds in the spatially continuous context: given a continuous image domain Ω and two terminals s and t linked to each $x \in \Omega$, we have three types of flows (refer to the middle continuous configuration in Figure 1): the source flow $p_s(x)$ from s to x, the sink flow $p_t(x)$ from x to t, and the spatial flow $p(x)$ for each $x \in \Omega$. These three flow fields are constrained pointwise by properly defined capacities:

$$p_s(x) \leq C_s(x)\,, \quad p_t(x) \leq C_t(x)\,, \quad |p(x)| \leq C(x) \tag{2.4}$$

as well as by a flow conservation condition,

$$p_t(x) - p_s(x) + \operatorname{div} p(x) = 0 \quad \forall x \in \Omega, \tag{2.5}$$

where divergence $\operatorname{div} p(x)$ evaluates the excess of the spatial flow around x.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Finding the maximum total flow from s based on the above configuration yields the following continuous max-flow problem:

$$\max_{p_s,p_t,p} \int_\Omega p_s \, dx \qquad \text{subject to (s.t.) (2.4) and (2.5).} \tag{2.6}$$

Yuan and coauthors [52, 53] proved that the continuous max-flow formulation in (2.6) is equivalent to the continuous min-cut problem studied in [13, 16]:

$$\min_{u(x)\in[0,1]} \int_\Omega (1-u)C_s \, dx + \int_\Omega u C_t \, dx + \int_\Omega C \, |\nabla u| \, dx. \tag{2.7}$$

Moreover, a continuous max-flow algorithm can be derived from (2.6) to efficiently and reliably compute a solution of (2.7) via convex optimization [8, 44].

It is worth noting that the pioneering work of Strang [48] is one of the first attempts to investigate the key idea of formulating min-cut problems in a continuous optimization framework. However, Strang's max-flow formulation is different from the continuous min-cut formulation in (2.7) in the sense that it eventually results in a ratio min-cut.

2.2. The Bhattacharyya distribution matching energy. Let $I : \Omega \subset \mathbb{R}^2 \to \mathcal{Z} \subset \mathbb{R}^n$ be an image function which maps domain Ω to a space $\mathcal{Z}$ of a photometric variable such as a color vector. The problem we tackle in this study consists of finding a region $\mathcal{R} \subset \Omega$ whose distribution most closely matches a known reference distribution $\mathcal{M}$.

Let $\mathcal{P}$ be the nonparametric estimate of the distribution of I within $\mathcal{R}$:

$$\mathcal{P}(z) = \frac{\int_\mathcal{R} \mathcal{K}_z(x) \, dx}{|\mathcal{R}|} \qquad \forall z \in \mathcal{Z} \tag{2.8}$$

with $|\mathcal{R}|$ the area of region $\mathcal{R}$, i.e., $|\mathcal{R}| = \int_\mathcal{R} dx$. Typical choices of $\mathcal{K}(\cdot)$ include the Dirac function, which yields the histogram, or the Gaussian kernel [39]:

$$\mathcal{K}_z(x) = \frac{1}{(2\pi\sigma^2)^{(n/2)}} \exp\left(-\frac{\|z - I(x)\|^2}{2\sigma^2}\right), \tag{2.9}$$

where σ is the width of the kernel. Note that the Dirac function can be approximated by a Gaussian kernel with a very small value of σ.

Assume that the distribution of I within the target object, denoted $\mathcal{M}$, is given. Let $\mathcal{B}(\mathcal{P}, \mathcal{M})$ denote the Bhattacharyya measure which evaluates the similarity between $\mathcal{P}$ and $\mathcal{M}$:

$$\mathcal{B}(\mathcal{P}, \mathcal{M}) = \sum_{z\in\mathcal{Z}} \sqrt{\mathcal{P}(z)\mathcal{M}(z)}. \tag{2.10}$$

Adding a regularization term for the smooth region boundary, the problem amounts to the following optimization:

$$\min_\mathcal{R} \Big\{ E(\mathcal{R}) := -\mathcal{B}(\mathcal{P}, \mathcal{M}) + C \, |\partial\mathcal{R}| \Big\}, \tag{2.11}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where $|\partial\mathcal{R}|$ measures the boundary length of region $\mathcal{R}$ and C is a positive constant weighing the contribution of the regularization terms.

Let $u : \Omega \to \{0,1\}$ be the indicator function of $\mathcal{R} \subset \Omega$, i.e., $u(x) = 1$ when $x \in \mathcal{R}$ and $u(x) = 0$ otherwise. Using u, the distribution in (2.8) can be rewritten as

$$\mathcal{P}(z) = \frac{\int_\Omega \mathcal{K}_z(x)u(x)\,dx}{\int_\Omega u(x)\,dx} \qquad \forall z \in \mathcal{Z}, \tag{2.12}$$

and the Bhattacharyya similarity in (2.10) as

$$\mathcal{B}(u,\mathcal{M}) = \sum_{z\in\mathcal{Z}} \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x)u(x)\,dx}{\int_\Omega u(x)\,dx} \right)^{1/2}, \tag{2.13}$$

where $\mathcal{T}_{M,z}(x) = \mathcal{K}_z(x)\mathcal{M}(z)$.

The perimeter of region $\mathcal{R}$ can be evaluated by the total variation of indicator function u [17, 27]:

$$|\partial\mathcal{R}| = \int_\Omega C\,|\nabla u(x)|\,dx. \tag{2.14}$$

In view of (2.13) and (2.14), we can, therefore, reformulate (2.11) in terms of the binary-valued labeling function u:

$$\min_{u(x)\in\{0,1\}} \left\{ E(u) := -\sum_{z\in\mathcal{Z}} \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x)u(x)\,dx}{\int_\Omega u(x)\,dx} \right)^{1/2} + \int_\Omega C\,|\nabla u(x)|\,dx \right\}. \tag{2.15}$$

3. Global minimization of energy upper bound. Direct computation of (2.15) is a very challenging optimization problem, both in theory and numerics, due to the constraint of a pointwise binary-valued variable and the high nonconvexity of the energy. Equation (2.15) is nonconvex even when $u \in \{0,1\}$ is relaxed by $u \in [0,1]$, unlike the classical continuous min-cut problem in [16, 53].

We propose a novel iterative convex relaxation solution to (2.15), which globally minimizes a sequence of upper bounds of $E(u)$, denoted $F(u,u^i)$, $i \geq 1$ (i is the iteration number):

$$u^{i+1} = \min_{u\in\{0,1\}} F(u,u^i), \quad i \geq 1 \quad \text{s.t.} \tag{3.1a}$$

$$E(u) \leq F(u,u^i), \quad i \geq 1, \tag{3.1b}$$

$$E(u) = F(u,u) \quad \forall u : \Omega \to \{0,1\}. \tag{3.1c}$$

We will give the expression of $F(u,u^i)$ in the next section. Let us first show that, in general, optimizing any energy upper bound that follows the constraints in (3.1) yields a monotonically decreasing sequence of the original energy E and, therefore, a minimum of E at convergence.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proposition 3.1. *Sequence $E(u^i)$ corresponding to the sequence of minima obtained in (3.1) is monotonically decreasing; i.e., $E(u^{i+1}) \leq E(u^i)$.*

Proof. Applying constraint (3.1c) to $u = u^i$ gives

$$E(u^i) = F(u^i, u^i). \tag{3.2}$$

Now, by the definition of minimum in (3.1a), we have

$$F(u^{i+1}, u^i) \leq F(u^i, u^i), \tag{3.3}$$

and combining (3.2) and (3.3) gives

$$F(u^{i+1}, u^i) \leq E(u^i). \tag{3.4}$$

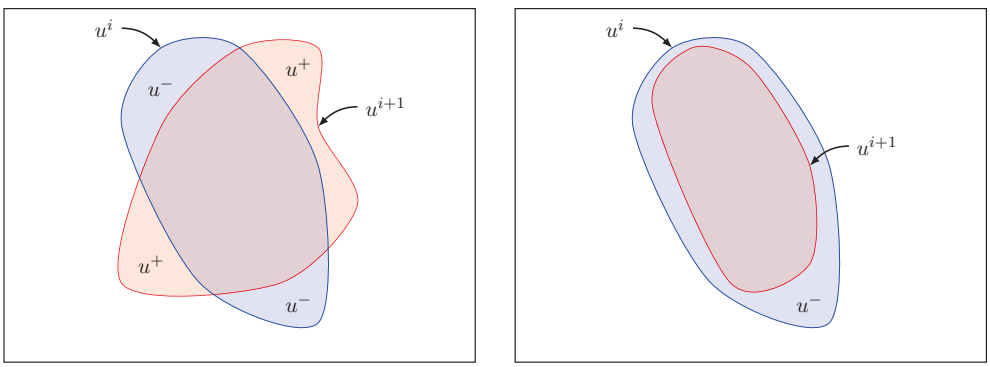
Also, applying constraint (3.1b) to $u = u^{i+1}$ gives

$$E(u^{i+1}) \leq F(u^{i+1}, u^i). \tag{3.5}$$

Finally, combining (3.4) and (3.5) proves Proposition 3.1. ∎

Proposition 3.2. *Sequence $E(u^i)$ is convergent.*

Proof. Because the Bhattacharyya measure is upper bounded by one, $E(u^i)$ is monotonically decreasing and lower bounded. Therefore $E(u^i)$ converges. ∎



(a) (b)

Figure 2. *Iterative bound optimization:* (a) *illustrates the proposed bound;* (b) *illustrates the bound in* [4]. *u^i denotes the labeling at iteration i.*

3.1. Energy upper bounds. In this section, we derive an energy upper bound $F(u, u^i)$, which is *convex with respect to labeling function* u. We will further show that $F(u, u^i)$ can be minimized *globally* and *exactly* over $u \in \{0, 1\}$ in section 4.

Let $u^i \in \{0, 1\}$ denote the binary labeling function, obtained at the previous ith step, as the starting point of the $(i + 1)$th iteration. At this stage, we are seeking an optimal labeling function $u \in \{0, 1\}$ which minimizes the energy upper bound. Let us express the difference between u and the given $u^i \in \{0, 1\}$ as a function of two new variables $u^+ \in \{0, 1\}$ and $u^- \in \{0, 1\}$ (refer to Figure 2(a)):

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

1. u^+ indicates the area where u^i is 0 and u becomes 1, i.e., area increase:

$$u^+(x) := \begin{cases} u(x), & \text{where } u^i(x) = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

2. u^- indicates the area where u^i is 1 and u becomes 0, i.e., area decrease:

$$u^-(x) := \begin{cases} 1 - u(x), & \text{where } u^i(x) = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

The Bhattacharyya measure (2.13) can then be rewritten in terms of u^+ and u^- as follows:

$$\mathcal{B}(u, \mathcal{M}) = \sum_{z \in \mathcal{Z}} \left(\frac{\int_\Omega (\mathcal{T}_{M,z} u^i + \mathcal{T}_{M,z} u^+ - \mathcal{T}_{M,z} u^-)\, dx}{\int_\Omega (u^i + u^+ - u^-)\, dx} \right)^{1/2}. \tag{3.8}$$

The following proposition describes the upper bound we propose.

Proposition 3.3. *Given a labeling $u^i \in \{0,1\}$, for any labeling function $u \in \{0,1\}$, we have the following upper bound which depends on u^+ and u^-:*

$$E(u) \leq -\mathcal{B}(u^i, \mathcal{M}) + J(u, u^i), \tag{3.9}$$

where

$$J(u, u^i) = \int_\Omega C_v^i u^-\, dx + \int_\Omega C_w^i u^+\, dx + \int_\Omega C\, |\nabla u|\, dx \tag{3.10}$$

and

$$C_v^i(x) = \sum_{z \in \mathcal{Z}} \frac{\mathcal{D}_{z,i} \mathcal{T}_{M,z}(x)}{\int_\Omega \mathcal{T}_{M,z}(x) u^i(x)\, dx}, \tag{3.11}$$

$$C_w^i(x) = \sum_{z \in \mathcal{Z}} \frac{\mathcal{D}_{z,i}}{2 \int_\Omega u^i(x)\, dx}, \tag{3.12}$$

$$\mathcal{D}_{z,i} = \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x) u^i(x)\, dx}{\int_\Omega u^i(x)\, dx} \right)^{1/2}. \tag{3.13}$$

Proof. We define

$$v^+ = \frac{\int_\Omega \mathcal{T}_{M,z}(x) u^+(x)\, dx}{\int_\Omega \mathcal{T}_{M,z}(x) u^i(x)\, dx}, \quad v^- = \frac{\int_\Omega \mathcal{T}_{M,z}(x) u^-(x)\, dx}{\int_\Omega \mathcal{T}_{M,z}(x) u^i(x)\, dx}$$

and

$$w^+ = \frac{\int_\Omega u^+(x)\, dx}{\int_\Omega u^i(x)\, dx}, \quad w^- = \frac{\int_\Omega u^-(x)\, dx}{\int_\Omega u^i(x)\, dx}.$$

Therefore, the Bhattacharyya measure (3.8) can be equally written as

$$\mathcal{B}(\mathcal{P}, \mathcal{M}) = \sum_{z \in \mathcal{Z}} \left\{ \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x) u^i(x)\, dx}{\int_\Omega u^i(x)\, dx} \right)^{1/2} \sqrt{\frac{1 + v^+ - v^-}{1 + w^+ - w^-}} \right\}. \tag{3.14}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Observing that all the variables v^+, v^-, w^+, and w^- are positive, we have

$$\mathcal{B}(u,\mathcal{M}) \geq \sum_{z\in\mathcal{Z}} \left\{ \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x)u^i(x)\,dx}{\int_\Omega u^i(x)\,dx} \right)^{1/2} \sqrt{\frac{1-v^-}{1+w^+}} \right\} \tag{3.15}$$

$$\geq \sum_{z\in\mathcal{Z}} \left\{ \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x)u^i(x)\,dx}{\int_\Omega u^i(x)\,dx} \right)^{1/2} (1-v^-)\left(1-\frac{1}{2}w^+\right) \right\} \tag{3.16}$$

$$\geq \sum_{z\in\mathcal{Z}} \left\{ \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x)u^i(x)\,dx}{\int_\Omega u^i(x)\,dx} \right)^{1/2} \left(1-v^--\frac{1}{2}w^+\right) \right\}. \tag{3.17}$$

The above result (3.16) can be seen by the fact that for any $v^- \in [0,1]$ and $w^+ \in \mathbb{R}^+$, we have

$$\sqrt{1-v^-} \geq 1-v^-, \quad 1/\sqrt{1+w^+} \geq 1-\frac{1}{2}w^+.$$

In view of (3.17), we have

$$\mathcal{B}(u,\mathcal{M}) \geq \mathcal{B}(u^i,\mathcal{M}) - \int_\Omega C_v^i(x)u^-(x)\,dx - \int_\Omega C_w^i(x)u^+(x)\,dx, \tag{3.18}$$

where

$$C_v^i(x) = \sum_{z\in\mathcal{Z}} \frac{\mathcal{D}_{z,n}\mathcal{T}_{M,z}(x)}{\int_\Omega \mathcal{T}_{M,z}(x)u^i(x)\,dx}, \quad C_w^i(x) = \sum_{z\in\mathcal{Z}} \frac{\mathcal{D}_{z,n}}{2\int_\Omega u^i(x)\,dx},$$

and

$$\mathcal{D}_{z,n} = \left(\frac{\int_\Omega \mathcal{T}_{M,z}(x)u^i(x)\,dx}{\int_\Omega u^i(x)\,dx} \right)^{1/2}.$$

Therefore, by (3.18), Proposition 3.3 can be proved. ∎

Proposition 3.3 leads us to the following conclusions:

- *Energy upper bound.* In view of (3.9), we have the energy upper bound

$$F(u,u^i) = -\mathcal{B}(u^i,\mathcal{M}) + J(u,u^i), \tag{3.19}$$

  where $\mathcal{B}(u^i,\mathcal{M})$ is a constant corresponding to the given u^i.
- *Continuous min-cut model.* Hence we can minimize $F(u,u^i)$ over $u \in \{0,1\}$ by

$$\min_{u\in\{0,1\}} J(u,u^i), \tag{3.20}$$

  where $J(u,u^i)$ is given in (3.10). The bound in (3.19) and its variable part $J(u,u^i)$ do not reference labeling variable u in the same way the standard min-cut problem in (2.7) does. $J(u,u^i)$ depends rather on labeling-change variables (u^+ and u^-). Therefore, it is not straightforward to optimize $J(u,u^i)$ with the standard flow configurations in the left and middle parts of Figure 1. The proposed flow configuration in the right part of Figure 1 is designed so as to account for u^+ and u^-. In the remainder of this paper, we will call (3.20) the *continuous min-cut model*.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

- *Cost of changing labels.* It is easy to see that $C_v^i(x)$ evaluates the cost of changing label $u(x)$ from $u^i(x) = 1$ to $u(x) = 0$; i.e., x is previously labeled as foreground, and cost $C_v^i(x)$ is paid when x is relabeled as background. $C_w^i(x)$ evaluates the cost of moving x from the background to the foreground.

4. Continuous max-flow approach. The new continuous min-cut model in (3.20) depends on labeling-function variations (u^+ and u^-), not on a single labeling function as with previous models [13, 16, 52, 53]. First, we examine a convex relaxation of (3.20). Then, we propose a novel flow configuration (refer to the rightmost graph of Figure 1) to solve such relaxation. This leads to a new max-flow formulation which is dual to the convex relaxed version of (3.20). Finally, we prove that our formulation yields a global and exact solution of (3.20).

4.1. Convex relaxed min-cut model. Let Ω_s and Ω_t be two disjoint domains given by the current labeling function $u^i \in \{0, 1\}$:

$$\Omega = \Omega_s \cup \Omega_t\,, \quad \Omega_s \cap \Omega_t = \emptyset, \tag{4.1}$$

where Ω_s is indicated by $u^i = 1$ and Ω_t by $u^i = 0$.

From the definitions of u^+ (3.6) and u^- (3.7), we reformulate (3.20) as

$$\min_{u(x)\in\{0,1\}} \int_{\Omega_t} C_v^i(1-u)\,dx + \int_{\Omega_s} C_w^i u\,dx + \int_{\Omega} C\,|\nabla u|\,dx \tag{4.2}$$

and propose a convex relaxation of (4.2) as follows:

$$\min_{u(x)\in[0,1]} \int_{\Omega_t} C_v^i(1-u)\,dx + \int_{\Omega_s} C_w^i u\,dx + \int_{\Omega} C\,|\nabla u|\,dx, \tag{4.3}$$

where the binary-valued constraint $u \in \{0, 1\}$ is relaxed to $u \in [0, 1]$. In the remainder of the paper, we will call (4.3) the *convex relaxed min-cut model.*

4.2. Continuous max-flow model. We propose a novel configuration of flows (a simple 1-dimensional example is depicted by the rightmost graph of Figure 1) which accounts for the previous partition defined by Ω_s and Ω_t, unlike the existing continuous flow setting [52, 53] (see the middle graph of Figure 1). Given Ω_s and Ω_t, we define three types of flows: the source flow $p_s(x)$ directed from s to every $x \in \Omega_s$, the sink flow $p_t(x)$ directed from every $x \in \Omega_t$ to t, and the spatial flow $p(x)$ given every $x \in \Omega$.

$p_s(x)$, $p_t(x)$, and $p(x)$ are constrained by the flow capacities

$$p_s(x) \le C_v^i(x) \quad \forall\, x \in \Omega_s\,; \tag{4.4}$$

$$p_t(x) \le C_w^i(x) \quad \forall\, x \in \Omega_t\,; \tag{4.5}$$

$$|p(x)| \le C \quad \forall\, x \in \Omega \tag{4.6}$$

and flow conservation conditions

$$-p_s(x) + \operatorname{div} p(x) = 0 \quad \forall\, x \in \Omega_s\,; \tag{4.7}$$

$$p_t(x) + \operatorname{div} p(x) = 0 \quad \forall\, x \in \Omega_t\,. \tag{4.8}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The continuous max-flow model can, therefore, be formulated by maximizing the total flow from the source s:

$$\max_{p_s,p_t,p} \int_{\Omega_s} p_s \, dx \qquad \text{s.t. (4.4)–(4.8).} \tag{4.9}$$

It is worth noting that the maximum in (4.9) is attained because it is a linear problem subject to linear constraints.

4.3. Equivalence between continuous max-flow (4.9) and continuous min-cut (4.3). Introduce the multiplier function $u(x) \in \mathbb{R}$ to the flow conservation constraints (4.7) at every $x \in \Omega_s$ and (4.8) at every $x \in \Omega_t$. We have the following *primal-dual* formulation equivalent to (4.9):

$$\min_u \max_{p_s,p_t,p} \int_{\Omega_s} p_s \, dx + \int_{\Omega_s} (\operatorname{div} p - p_s) u \, dx + \int_{\Omega_t} (\operatorname{div} p + p_t) u \, dx \qquad \text{s.t. (4.4)–(4.6),} \tag{4.10}$$

which can be rearranged as

$$\min_u \max_{p_s,p_t,p} \int_{\Omega_s} (1-u) p_s \, dx + \int_{\Omega_t} u \, p_t \, dx + \int_{\Omega} u \operatorname{div} p \, dx \qquad \text{s.t. (4.4)–(4.6).} \tag{4.11}$$

To build the equivalence between max-flow model (4.9) and min-cut formulation (4.3), we prove the following proposition.

Proposition 4.1. *The primal-dual model* (4.11) *amounts to*

$$\min_u \int_{\Omega_s} (1-u) C_v^i \, dx + \int_{\Omega_t} u C_w^i \, dx + \int_{\Omega} C \, |\nabla u| \, dx, \tag{4.12}$$

where

$$u(x) \le 1 \quad \forall \, x \in \Omega_s \,, \quad u(x) \ge 0 \quad \forall \, x \in \Omega_t \,.$$

Proof. The proof is based on the following variational facts: The constraints of flows are convex, and the energy function is linear in both the primal and dual functions $p_s(x)$, $p_t(x)$, $p(x)$, and $u(x)$, hence convex l.s.c. (lower semicontinuous) for fixed u and concave u.s.c. (upper semicontinuous) for fixed $p_s(x)$, $p_t(x)$, and $p(x)$ [44]. This implies the existence of at least one saddle point of the primal-dual model (4.10) or (4.11), i.e., strong duality. It follows that we can interchange the min and max order of (4.10) or (4.11). Here we first maximize the energy function of the primal-dual model over the three flow functions $p_s(x)$, $p_t(x)$, $p(x)$ and then minimize over $u(x)$.

Now we consider the optimization problem

$$f(v) = \sup_{w \le C} vw \,, \tag{4.13}$$

where v, w, and C are scalars. When $v < 0$, w can be negative infinity in order to maximize the value vw, i.e., $f(v) = +\infty$. To achieve a meaningful function $f(v)$ over v, we have $v \le 0$. It can also be easily seen that for $w \le C$,

$$\begin{cases} \text{if } v = 0 \,, & \text{then } w \le C \text{ maximizes the function and } f(v) = 0, \\ \text{if } v > 0 \,, & \text{then } w = C \text{ maximizes the function and } f(v) = vC. \end{cases} \tag{4.14}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In summary, we have

$$f(v) = vC \quad \text{and} \quad v \geq 0\,. \tag{4.15}$$

In view of (4.13) and (4.15), it follows that the maximization of (4.11) over p_s, constrained by (4.4), amounts to

$$\max_{p_s(x)\leq C_v^i(x)} \int_{\Omega_s} (1-u(x))\, p_s(x)\, dx \;=\; \int_{\Omega_s} (1-u(x))\, C_v^i(x)\, dx \quad \text{and} \quad u(x) \;\leq\; 1 \quad \text{for } x \in \Omega_s\,. \tag{4.16}$$

In the same manner, the maximization of (4.11) over p_t, constrained by (4.5), corresponds to

$$\max_{p_t(x)\leq C_w^i(x)} \int_{\Omega_t} u(x)\, p_t(x)\, dx \;=\; \int_{\Omega_t} u(x)\, C_w^i(x)\, dx \quad \text{and} \quad u(x) \;\geq\; 0 \quad \text{for } x \in \Omega_t\,. \tag{4.17}$$

It is also well known that [27]

$$\max_{|p(x)|\leq C(x)} \int_{\Omega} u\ \mathrm{div}\, p\, dx \;=\; \int_{\Omega} C\,|\nabla u|\ dx\,. \tag{4.18}$$

Observing (4.16), (4.17), and (4.18), Proposition 4.1 is proved. ∎

Observing Proposition 4.1, we further prove the following proposition.

Proposition 4.2. *The optimum u^* of the minimization problem in (4.12) verifies $u^* \in [0,1]$ for almost everywhere in Ω. Therefore, we have the equivalence*

$$(4.9) \iff (4.10) \quad \text{or} \quad (4.11) \iff (4.3)\,.$$

Proof. Let u^* be the minimum of (4.12). Due to the convexity of (4.12), u^* is simply accepted as a global minimum. Now we prove that $0 \leq u^*(x) \leq 1$ for all $x \in \Omega$.

Through Proposition 4.1, we have

$$u^*(x) \;\leq\; 1 \quad \text{for } x \in \Omega_s\,, \quad u^*(x) \;\geq\; 0 \quad \text{for } x \in \Omega_t\,.$$

Now we prove $u^*(x) \geq 0$ for $x \in \Omega_s$. If $u^*(x) < 0$ at $\tilde{\Omega} \subset \Omega_s$, then we define the function $u'(x)$ which just sets the value $u^*(x)$ to be nonnegative; i.e.,

$$u'(x) \;=\; \begin{cases} 0 & \text{for } x \in \tilde{\Omega}, \\ u^*(x) & \text{for } x \in \Omega\backslash\tilde{\Omega}. \end{cases}$$

Obviously $u'(x) \leq u^*(x)$ for $x \in \Omega_s$; then we have

$$\int_{\Omega_s} (1-u^*(x))\, C_v^i(x)\, dx \;=\; \int_{\Omega_s} (1-u'(x))\, C_v^i(x)\, dx \;+\; \int_{\tilde{\Omega}} (0-u^*(x))\, C_v^i(x)\, dx\,.$$

Observing that $C_v^i(x) > 0$ by (3.11) and $u^*(x) < 0$ at $\tilde{\Omega} \subset \Omega_s$, we have

$$\int_{\Omega_s} (1-u'(x))\, C_v^i(x)\, dx \;<\; \int_{\Omega_s} (1-u^*(x))\, C_v^i(x)\, dx. \tag{4.19}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

On the other hand, by the coarea formula of the total variation term,

$$\int_\Omega C\,|\nabla u|\;dx \;=\; \int_{-\infty}^{+\infty} L_\gamma(u)\,d\gamma\,,$$

where $L_\gamma(u)$ is the weighted length of the γ-upper level set of $u(x)$. It follows that

$$\int_\Omega C\,\left|\nabla u'\right|\;dx \;<\; \int_\Omega C\,|\nabla u^*|\;dx\,, \tag{4.20}$$

because for $u'(x)$, all the γ-upper level sets of u', when $\gamma < 0$, are set to vanish.

Through (4.19), (4.20), and $u'(x) = u^*(x)$ for $x \in \Omega_t$, we have

$$\begin{aligned}&\left\{\int_{\Omega_s}(1-u'(x))\,C_v^i(x)\,dx \;+\; \int_{\Omega_t} u'(x)\,C_w^i(x)\,dx \;+\; \int_\Omega C\,\left|\nabla u'\right|\;dx\right\}\\ <\;&\left\{\int_{\Omega_s}(1-u^*(x))\,C_v^i(x)\,dx \;+\; \int_{\Omega_t} u^*(x)\,C_w^i(x)\,dx \;+\; \int_\Omega C\,|\nabla u^*|\;dx\right\}.\end{aligned}$$

This is in contradiction to the fact that $u^*(x)$ is the global optimum of (4.12).

Likewise, we can prove that $u^*(x) \le 1$ for $x \in \Omega_t$. We therefore have $u^*(x) \in [0,1]$ for $x \in \Omega$.

The equivalence between the continuous max-flow model (4.9), the primal-dual model (4.10) or (4.11), and the dual convex min-cut model (4.3) can be established consequently. ∎

4.4. Exactness of the convex relaxed min-cut model in (4.3). In this section, we show that the binary-valued optimization problem in (4.2) can be solved globally and exactly by its convex relaxation in (4.3).

Clearly, (4.3) is convex. Let $u^*(x)$ be its global optimum. For any $\gamma \in [0,1]$, we define the γ-upper level set $u_\gamma^*(x)$ for all $x \in \Omega$ by

$$u_\gamma^*(x) \;=\; \begin{cases} 1 & \text{when } u^*(x) > \gamma,\\ 0 & \text{when } u^*(x) \le \gamma.\end{cases} \tag{4.21}$$

Moreover, we have the following proposition.

Proposition 4.3. *When $u^*(x)$ gives one global optimum of (4.3), its thresholding $u_\gamma^*(x) \in \{0,1\}$ by (4.21), for any $\gamma \in [0,1]$, solves the binary-valued optimization problem in (4.2) globally and exactly.*

The proof of Proposition 4.3 follows the ideas of [52, 53], is discussed in [14, 16], and is given below.

Proof. There exists $(p_s^*, p_t^*, p^*; u^*(x))$ as the optimal primal-dual pair of the primal-dual model (4.10) or (4.11); see the proof of Proposition 4.1. Clearly, p_s^*, p_t^*, p^* optimize the continuous max-flow problem (4.9) and give the globally maximal max-flow energy

$$E_{mf}^* \;=\; \int_{\Omega_s} p_s^*(x)\,dx\,.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$u^*(x)$ optimizes the dual convex min-cut formulation (4.3) and gives the globally minimal energy of the convex min-cut model (4.3), i.e.,

$$E^*_{mc} = \int_{\Omega_s} (1 - u^*(x))\, C^i_v(x)\, dx + \int_{\Omega_t} u^*(x)\, C^i_w(x)\, dx + \int_{\Omega} C\, |\nabla u^*|\, dx\,.$$

Let $E_{pd}(p_s, p_t, p; u)$ be the primal-dual energy of (4.10) over the minimax of p_s, p_t, p, and u, i.e.,

$$E_{pd}(p_s, p_t, p; u) = \int_{\Omega_s} (1 - u(x))\, p_s(x)\, dx + \int_{\Omega_t} u\, p_t(x)\, dx + \int_{\Omega} u(x)\ \mathrm{div}\, p(x)\, dx\,.$$

Observing the proof of Proposition 4.1 (the interchangeable min-max order), we have

(4.22)
$$E^*_{mf} = \max_{p_s,p_t,p} \min_{u} E_{pd}(p_s, p_t, p; u) = E_{pd}(p^*_s, p^*_t, p^*; u^*) = \min_{u} \max_{p_s,p_t,p} E_{pd}(p_s, p_t, p; u) = E^*_{mc}\,.$$

Given the γ-level set $u^*_\gamma(x)$ of $u^*(x)$ for any $\gamma \in (0, 1]$, now we prove that $(p^*_s, p^*_t, p^*; u^*_\gamma)$ is also an optimal primal-dual pair. This can be validated by the following facts:

- When the flow function $p^*(x)$ maximizes $\int_\Omega u\, \mathrm{div}\, p^*\, dx$, for any γ-upper level set of $u(x)$, i.e., $u_\gamma(x)$, $\gamma \in (u_{\min}, u_{\max}]$, $p^*(x)$ also maximizes $\int_\Omega u_\gamma\, \mathrm{div}\, p^*\, dx$. This is another representation of the coarea theorem. Then, it follows that $p^*(x)$ also maximizes integral $\int_\Omega u^*_\gamma\, \mathrm{div}\, p^*\, dx$.
- Let $\tilde{\Omega} \subset \Omega_s$ be the area where $u^*(x) < \gamma$ for $x \in \tilde{\Omega}$. In Ω_s, p^*_s maximizes the integral $\int_{\Omega_s} (1 - u^*)\, p_s\, dx$ over the flow capacity constraint (4.4) for p_s. Observing (4.14) and $1 - u^*(x) > 0$ for $x \in \tilde{\Omega}$, we have $p^*_s(x) = C^i_v(x)$ for $x \in \tilde{\Omega}$. In addition, for $u^*_\gamma(x)$, we have
$$1 - u^*_\gamma(x) = 1 \ \text{ for } \ x \in \tilde{\Omega}\,, \quad 1 - u^*_\gamma(x) = 0 \ \text{ for } \ x \in \Omega_s \backslash \tilde{\Omega}\,.$$
Therefore, by (4.14), $p^*_s(x)$ maximizes the integral $\int_{\Omega_s} (1 - u^*_\gamma)\, p_s\, dx$ over the flow capacity constraint (4.4).
- Let $\tilde{\Omega} \subset \Omega_t$ be the area where $u^*(x) \geq \gamma$ for $x \in \tilde{\Omega}$. In Ω_t, p^*_t maximizes the integral $\int_{\Omega_t} u^*\, p_t\, dx$ over the flow capacity constraint (4.5) for p_t. Observing (4.14) and $u^*(x) > 0$ for $x \in \tilde{\Omega}$, we have $p^*_t(x) = C^i_w(x)$ for $x \in \tilde{\Omega}$. In addition, for $u^*_\gamma(x)$, we have
$$u^*_\gamma(x) = 1 \ \text{ for } \ x \in \tilde{\Omega}\,, \quad u^*_\gamma(x) = 0 \ \text{ for } \ x \in \Omega_t \backslash \tilde{\Omega}\,.$$
Therefore, by (4.14), $p^*_t(x)$ maximizes the integral $\int_{\Omega_t} u^*_\gamma\, p_t\, dx$ over the flow capacity constraint (4.4).

Hence, we have

$$\begin{aligned} \max_{p_s,p_t,p} E_{pd}(p_s, p_t, p; u^*_\gamma) &= E_{pd}(p^*_s, p^*_t, p^*; u^*_\gamma) \\ &= \int_{\Omega_s} p^*_s(x)\, dx + \int_{\Omega_s} (\mathrm{div}\, p^* - p^*_s) u^*_\gamma\, dx + \int_{\Omega_t} (\mathrm{div}\, p^* + p^*_t) u^*_\gamma\, dx \\ &= \int_{\Omega_s} p^*_s(x)\, dx = E^*_{mf}, \end{aligned} \tag{4.23}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

where the last term comes from the flow conservation conditions (4.7) and (4.8).

On the other hand, we also have

$$E_{pd}(p_s^*, p_t^*, p^*; u_\gamma^*) = \int_{\Omega_s} (1 - u_\gamma^*(x))\, C_v^i(x)\, dx + \int_{\Omega_t} u_\gamma^*(x)\, C_w^i(x)\, dx + \int_\Omega C \left|\nabla u_\gamma^*\right| dx\,, \tag{4.24}$$

following the same procedure as in the proof for Proposition 4.1.

Combining (4.22), (4.23), and (4.24), we have

$$\begin{aligned} &\int_{\Omega_s} (1 - u_\gamma^*(x))\, C_v^i(x)\, dx + \int_{\Omega_t} u_\gamma^*(x)\, C_w^i(x)\, dx + \int_\Omega C \left|\nabla u_\gamma^*\right| dx \\ &= \int_{\Omega_s} (1 - u^*(x))\, C_v^i(x)\, dx + \int_{\Omega_t} u^*(x)\, C_w^i(x)\, dx + \int_\Omega C \left|\nabla u_\gamma^*\right| dx \\ &= E_{mc}^*\,. \end{aligned}$$

Observing that E_{mc}^* gives the global minimal energy of the convex min-cut (4.3), the relaxation of (4.3), we conclude that $u_\gamma^*(x) \in \{0, 1\}$ solves (4.2) globally and exactly. ∎

5. Continuous max-flow based algorithm. The following summarizes the proposed procedure:

1. Start with an arbitrary initial labeling u^1 $(i = 1)$;
2. For the ith outer iteration, set the flow configuration based on labeling function u^i. Compute C_v^i and C_w^i by (3.11) and (3.12);
3. Compute $u^{i+1}(x) \in \{0, 1\}$ as the global and exact optimum in (4.2) via the convex relaxation in (4.3) and the corresponding max-flow model in (4.9). The max-flow algorithm is detailed in Algorithm 1;
4. Let $i = i + 1$ and repeat steps 2 and 3 until convergence.

ALGORITHM 1. Continuous max-flow based algorithm.

- $i = 1$: Initialize flows p_s^1, p_t^1, p^1 and labeling $u^1 \in \{0, 1\}$. Then start $(i+1)$th iteration;
- Fix p^i and u and maximize in a closed form $L_c(p_s, p_t, p, u)$ over p_s and p_t:

$$(p_s^{i+1}, p_t^{i+1}) := \arg\max_{p_s, p_t} L_c(p_s, p_t, p^i, u)\,;$$

- Fix p_s^{i+1}, p_t^{i+1}, and u and maximize iteratively $L_c(p_s, p_t, p, u)$ over p via the simple projection-descent step:

$$p^{i+1} := \arg\max_{p} L_c(p_s^{i+1}, p_t^{i+1}, p, u)\,;$$

- Update labeling function u by

$$u(x) = \begin{cases} u(x) + c\,(\operatorname{div} p^{i+1}(x) - p_s^{i+1}(x)) & \forall x \in \Omega_s, \\ u(x) + c\,(\operatorname{div} p^{i+1}(x) + p_t^{i+1}(x)) & \forall x \in \Omega_t. \end{cases}$$

Repeat the previous two steps until convergence.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Note that the energy in (4.10) is the Lagrangian function of (4.9), where labeling u is the multiplier to flow conservation conditions (4.7) and (4.8). In this regard, we define the following augmented Lagrangian function:

$$\begin{aligned} L_c(p_s, p_t, p, u) = &\int_{\Omega_s} p_s \, dx + \int_{\Omega_s} (\operatorname{div} p - p_s) u \, dx \\ &+ \int_{\Omega_t} (\operatorname{div} p + p_t) u \, dx - \frac{c}{2} \|\operatorname{div} p - p_s\|^2_{\Omega_s} \\ &- \frac{c}{2} \|\operatorname{div} p + p_t\|^2_{\Omega_t}, \end{aligned} \tag{5.1}$$

where $c > 0$. Finally, we derive the continuous max-flow algorithm from the classical augmented Lagrangian method [42, 43], whose convergence properties can be found in [7, 8]. The solution is sought by optimization over the flows and labeling simultaneously.

6. Experiments.

We applied the proposed method to the 50 images of the Microsoft GrabCut segmentation database [46], which also contains ground truth segmentations of the foreground objects. The results were compared with similar experiments on the same data using several other algorithms. We measured the optimality by computing (1) the Bhattacharyya measure at convergence and (2) the difference in histogram w.r.t. ground truth, i.e., the L_1 norm of the difference between the obtained region and ground truth histograms. We also evaluated the accuracy of the algorithms by computing the number of misclassified pixels in comparison to the ground truth. For the proposed algorithm, the parameters for all 50 images were selected as follows: $C = 0.3$, error threshold (convergence criteria) $= 1 \times 10^{-2}$, and $c = 0.05$. The density is estimated from the red/green/blue (RGB) components with a 3-dimensional histogram consisting of $256 \times 256 \times 256$ bins. Following the experiment in [46], we estimate the model from the ground truth for all compared methods.

6.1. Quantitative evaluations.

Table 1 reports average number of misclassified pixels (error), the norm of the average difference between the obtained histogram and the model (L_1-norm), average computational time (time), the Bhattacharyya measure at convergence ($\mathcal{B}$), and the number of KDE (kernel density estimation) evaluations per image (# KDEs) for the following:

- the proposed method;
- the bound in [4] in conjunction with the standard convex max-flow framework in [52, 53];
- the bound in [4] optimized via discrete graph cuts;

Table 1

Quantitative accuracy and optimality evaluations.

Methods	L_1-norm	$\mathcal{B}$	# KDEs	Time	Error
Proposed flow config. + proposed bound	373.2	0.9991 ± 0.0009	13	18.8	383
Flow config. in [52, 53] + bound in [4]	737.5	0.9943 ± 0.0059	3	3.2	750
Graph cuts + bound in [4]	501.1	0.9984 ± 0.0016	4	3.5	520
Log-likelihood model in [10, 45]	1418.2	0.9903 ± 0.0071	1	0.2	1426
Active curve optimization	–	0.9791 ± 0.01	1737	631	–

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

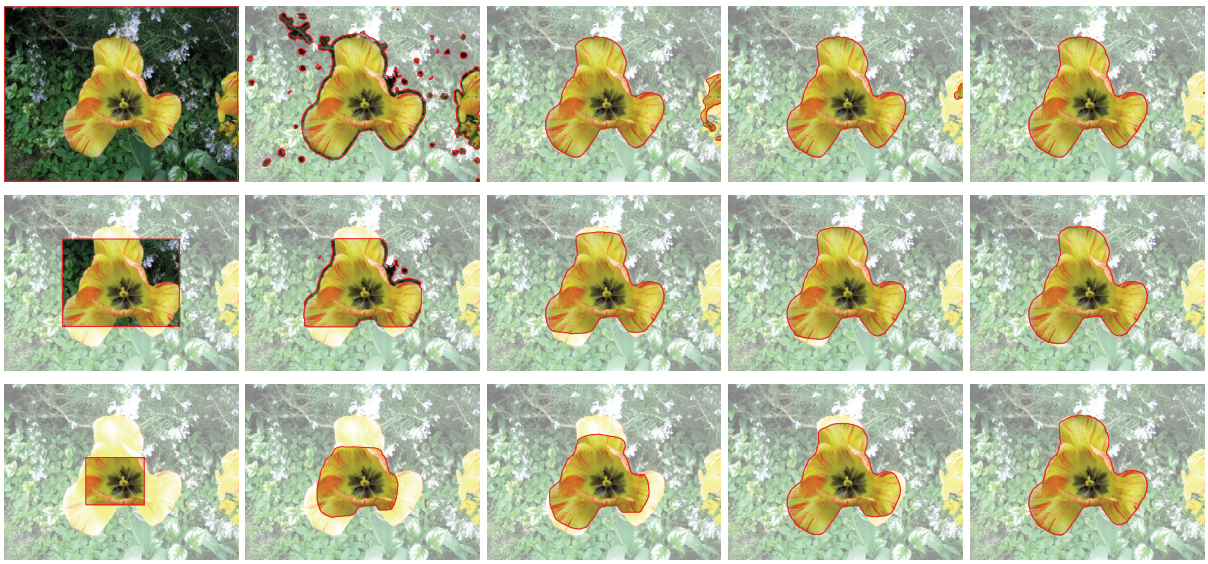


Figure 3. *Evolution of the solution of the proposed algorithm w.r.t. different initial labelings. We initialized the figure region with* (1) *the whole image domain (1st row);* (2) *a part of the figure and ground (2nd row); and* (3) *a part of the figure (3rd row). The algorithm converged to the same final solution regardless of initialization. The density is estimated from the RGB components with a 3-dimensional histogram consisting of* $256 \times 256 \times 256$ *bins.*

- the log-likelihood model in [10, 45] optimized via discrete graph cuts; and
- active curve optimization [6, 25].

The proposed method yielded the lowest error, highest Bhattacharyya measure, and lowest L_1-norm of histogram difference, which indicates improvements in accuracy and optimality. For the proposed method, the average time for computing the KDEs and the Bhattacharyya distances is 4.5 seconds, which corresponds to approximately 25% of the overall computational time (18.8 seconds).

Note that the figure and ground prior distributions required by the likelihood-based model [10, 45] were learned from the ground truth and, therefore, were known in advance. Even though this comparison is not in favor of our method (as it does not use the background-model information), our experiments indicate that the proposed algorithm can bring an improvement on the order of 10^3 pixels.

6.2. Robustness w.r.t. initial conditions. The examples in Figure 3 illustrate the robustness of the proposed bound w.r.t. different initial labelings. We initialized the figure region with (1) the whole image domain (1st row); (2) a part of the figure and ground (2nd row); and (3) a part of the figure (3rd row). Our algorithm converged to the same final solution. The typical examples in Figure 4 use various initializations that, by definition, preclude the application of the bound in [4]. The figure demonstrates how the proposed bound can deal with these different types of initializations, unlike the bound in [4], which requires the initial region to include the target region.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

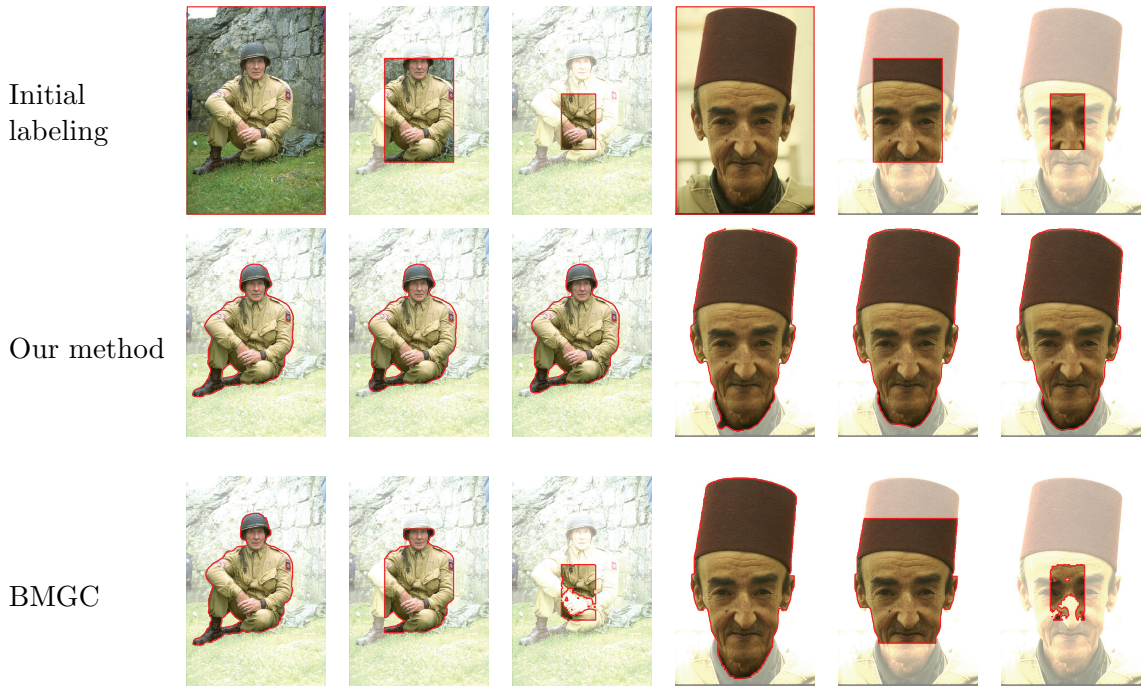


Figure 4. *Robustness of the proposed algorithm w.r.t. initializations and comparison to BMGC* [4]. *We initialized the figure region with* (1) *the whole image domain;* (2) *a part of the figure and ground; and* (3) *a part of the figure (*1*st row).* $C(x) = 1.5$ *and* $c = 2$. *The results demonstrate that the proposed bound (*2*nd row) is more robust than the bound in* [4] *(*3*rd row), which requires the initial region to include the target region. The density is estimated from the RGB components with a* 3*-dimensional histogram consisting of* $256 \times 256 \times 256$ *bins.*

6.3. Segmentation examples. Figure 5 depicts some representative examples of segmentations with the proposed method along with the ground truth segmentations from the GrabCut database. The pixel error, the initial Bhattacharyya measure, the Bhattacharyya measure corresponding to the final segmentation, total computational time, and image size are given for each example. The results with difficult examples undergoing strong camouflage demonstrate the effectiveness of the proposed method.

6.4. The effect of kernel width σ. Figure 6 depicts the effect of kernel width σ on the results. This typical example shows that the histogram representation, which can be viewed as a KDE with a very small σ, yields the best performance. The higher σ is, the further the result is from the ground truth. This is expected in camouflage examples, where the figure and ground regions have some similarities in their color distributions. In such cases, a high value of σ affects the ability of the algorithm in distinguishing fine color differences.

It is worth noting that bandwidth selection for density estimation has been intensively studied in statistics [30, 23], and there are several ways of computing the optimal value of parameter σ. Generally, the optimal kernel width minimizes a given error between the original density and its estimate and depends on the size and statistics of the data [30]. Note that more accurate estimates of the optimal kernel width can be computed via iterative procedures [30]. However, embedding such procedures into segmentation leads to a heavy computation load.

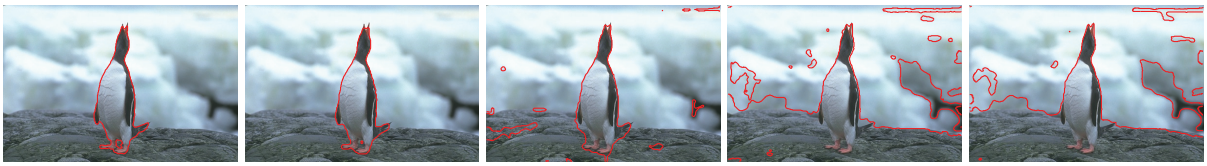
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Ground truth				
Segmentation				
Error	0.02%	0.03%	0.06%	0.15%
$\mathcal{B}_{initial}$	0.3726	0.4486	0.3636	0.4089
$\mathcal{B}_{final}$	0.9994	0.9998	0.9991	0.9958
Run time	7.4 secs	13.0 secs	24.6 secs	11.0 secs
Image size	321×481	450×640	600×450	481×321

Figure 5. *Segmentation results of the proposed method from sample images from the GrabCut database. 1st row: the image and ground truth segmentation boundary (green curve). 2nd row: foreground segmentation obtained with the proposed method. The density is estimated from the RGB components with a 3-dimensional histogram consisting of* $256 \times 256 \times 256$ *bins.*



(a) Histogram (b) $\sigma = 1$ (c) $\sigma = 2$ (d) $\sigma = 3$ (e) $\sigma = 4$

Figure 6. *The effect of kernel width* σ*.*

REFERENCES

[1] F. Aherne, N. Thacker, and P. Rockett, *The Bhattacharyya metric as an absolute similarity measure for frequency coded data*, Kybernetika (Prague), 34 (1998), pp. 363–368.

[2] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, *Image segmentation using active contours: Calculus of variations or shape gradients?*, SIAM J. Appl. Math., 63 (2003), pp. 2128–2154.

[3] D. Batra, A. Kowdle, D. Parikh, J. Luo, and T. Chen, *iCoseg: Interactive co-segmentation with intelligent scribble guidance*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2010.

[4] I. Ben Ayed, H.-M. Chen, K. Punithakumar, I. Ross, and S. Li, *Graph cut segmentation with a global constraint: Recovering region distribution via a bound of the Bhattacharyya measure*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2010.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[5] I. Ben Ayed, Shuo Li, and I. Ross, *Tracking distributions with an overlap prior*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[6] I. Ben Ayed, S. Li, and I. Ross, *A statistical overlap prior for variational image segmentation*, Int. J. Comput. Vision, 85 (2009), pp. 115–132.
[7] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Comput. Sci. Appl. Math., Academic Press, New York, 1982.
[8] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Nashua, NH, 1999.
[9] Y. Boykov and G. Funka Lea, *Graph cuts and efficient N-D image segmentation*, Int. J. Comput. Vision, 70 (2006), pp. 109–131.
[10] Y. Boykov and M.-P. Jolly, *Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images*, in Proceedings of the Eighth IEEE International Conference on Computer Vision, Vol. 1, 2001, pp. 105–112.
[11] Y. Boykov and V. Kolmogorov, *An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision*, IEEE Trans. Pattern Anal. Mach. Intell., 26 (2004), pp. 1124–1137.
[12] Y. Boykov, O. Veksler, and R. Zabih, *Fast approximate energy minimization via graph cuts*, IEEE Trans. Pattern Anal. Mach. Intell., 23 (2001), pp. 1222–1239.
[13] X. Bresson, S. Esedoglu, P. Vandergheynst, J. P. Thiran, and S. Osher, *Fast global minimization of the active contour/snake model*, J. Math. Imaging Vision, 28 (2007), pp. 151–167.
[14] A. Chambolle, *Total variation minimization and a class of binary MRF models*, in Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Comput. Sci. 3757, A. Rangarajan et al., eds., Springer-Verlag, Berlin, Heidelberg, 2005, pp. 136–152.
[15] A. Chambolle, D. Cremers, and T. Pock, *A Convex Approach for Computing Minimal Partitions*, Technical report TR-2008-05, Department of Computer Science, University of Bonn, Bonn, Germany, 2008.
[16] T. F. Chan, S. Esedoḡlu, and M. Nikolova, *Algorithms for finding global minimizers of image segmentation and denoising models*, SIAM J. Appl. Math., 66 (2006), pp. 1632–1648.
[17] T. F. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, Philadelphia, 2005.
[18] S. Chen and R. J. Radke, *Level set segmentation with both shape and intensity priors*, in Proceedings of the 12th IEEE International Conference on Computer Vision, 2009.
[19] M. Cho, Y. Min Shin, and K. Mu Lee, *Co-recognition of image pairs by data-driven Monte Carlo image exploration*, in Proceedings of the Tenth European Conference on Computer Vision, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 144–157.
[20] D. Comaniciu, V. Ramesh, and P. Meer, *Kernel-based object tracking*, IEEE Trans. Pattern Anal. Mach. Intell., 25 (2003), pp. 564–575.
[21] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 2nd ed., MIT Press, Cambridge, MA, 2001.
[22] J. Cui, Q. Yang, F. Wen, Q. Wu, C. Zhang, L. V. Gool, and X. Tang, *Transductive object cutout*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[23] P. Deheuvels, *Estimation non paramétrique de la densité par histogrammes généralisés*, Rev. Statist. Appl., 25 (1977), pp. 5–42.
[24] L. R. Ford and D. R. Fulkerson, *Maximal flow through a network*, Canad. J. Math., 8 (1956), pp. 399–404.
[25] D. Freedman and T. Zhang, *Active contours for tracking distributions*, IEEE Trans. Image Process., 13 (2004), pp. 518–526.
[26] A. C. Gallagher and T. Chen, *Clothing cosegmentation for recognizing people*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[27] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Australian National University, Canberra, Australia, 1977.
[28] T. Goldstein, X. Bresson, and S. Osher, *Geometric applications of the split Bregman method: Segmentation and surface reconstruction*, J. Sci. Comput., 45 (2010), pp. 272–293.
[29] D. S. Hochbaum and V. Singh, *An efficient algorithm for co-segmentation*, in Proceedings of the 12th IEEE International Conference on Computer Vision, 2009.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[30] M. C. Jones, J. S. Marron, and S. J. Sheather, *A brief survey of bandwidth selection for density estimation*, J. Amer. Statist. Assoc., 91 (1996), pp. 401–407.
[31] M. Klodt, T. Schoenemann, K. Kolev, M. Schikora, and D. Cremers, *An experimental comparison of discrete and continuous shape optimization methods*, in Proceedings of the Tenth European Conference on Computer Vision, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 332–345.
[32] V. Kolmogorov and Y. Boykov, *What metrics can be approximated by geo-cuts, or global optimization of length/area and flux*, in Proceedings of the Tenth IEEE International Conference on Computer Vision, 2005.
[33] J. Lellmann, J. Kappes, J. Yuan, F. Becker, and C. Schnörr, *Convex multi-class image labeling by simplex-constrained total variation*, in Proceedings of the Second International Conference on Scale Space and Variational Methods in Computer Vision, Springer-Verlag, Berlin, Heidelberg, 2009, pp. 150–162.
[34] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp, *Image segmentation with a bounding box prior*, in Proceedings of the 12th IEEE International Conference on Computer Vision, 2009, pp. 277–284.
[35] O. Michailovich, Y. Rathi, and A. Tannenbaum, *Image segmentation using active contours driven by the Bhattacharyya gradient flow*, IEEE Trans. Image Process., 16 (2007), pp. 2787–2801.
[36] L. Mukherjee and V. Singh, *Scale invariant cosegmentation for image groups*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011.
[37] L. Mukherjee, V. Singh, and C. R. Dyer, *Half-integrality based algorithms for cosegmentation of images*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2009.
[38] K. Ni, X. Bresson, T. F. Chan, and S. Esedoglu, *Local histogram based segmentation using the Wasserstein distance*, Int. J. Comput. Vision, 84 (2009), pp. 97–111.
[39] E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Statist., 33 (1962), pp. 1065–1076.
[40] V.-Q. Pham, K. Takahashi, and T. Naemura, *Foreground-background segmentation using iterated distribution matching*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011.
[41] T. Pock, A. Chambolle, D. Cremers, and H. Bischof, *A convex relaxation approach for computing minimal partitions*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2009.
[42] R. T. Rockafellar, *The multiplier method of Hestenes and Powell applied to convex programming*, J. Optim. Theory Appl., 12 (1973), pp. 555–562.
[43] R. T. Rockafellar, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Oper. Res., 1 (1976), pp. 97–116.
[44] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin, 1998.
[45] C. Rother, V. Kolmogorov, and A. Blake, *"GrabCut": Interactive foreground extraction using iterated graph cuts*, ACM Trans. Graph., 23 (2004), pp. 309–314.
[46] C. Rother, T. Minka, A. Blake, and V. Kolmogorov, *Cosegmentation of image pairs by histogram matching—incorporating a global constraint into MRFs*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006.
[47] B. C. Russell, W. T. Freeman, A. A. Efros, J. Sivic, and A. Zisserman, *Using multiple segmentations to discover objects and their extent in image collections*, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, pp. 1605–1614.
[48] G. Strang, *Maximal flow through a domain*, Math. Programming, 26 (1983), pp. 123–143.
[49] S. Vicente, V. Kolmogorov, and C. Rother, *Joint optimization of segmentation and appearance models*, in Proceedings of the 12th IEEE International Conference on Computer Vision, 2009.
[50] S. Vicente, V. Kolmogorov, and C. Rother, *Cosegmentation revisited: Models and optimization*, in Proceedings of the 11th European Conference on Computer Vision, Springer-Verlag, Berlin, Heidelberg, 2010.
[51] S. Vicente, C. Rother, and V. Kolmogorov, *Object cosegmentation*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 2217–2224.
[52] J. Yuan, E. Bae, and X.-C. Tai, *A study on continuous max-flow and min-cut approaches*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2010.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[53] J. Yuan, E. Bae, X.-C. Tai, and Y. Boykov, *A Study on Continuous Max-Flow and Min-Cut Approaches. Part* I: *Binary Labeling*, Technical report CAM 10-61, UCLA, Los Angeles, CA, 2010.

[54] J. Yuan, E. Bae, X.-C. Tai, and Y. Boykov, *A continuous max-flow approach to Potts model*, in Proceedings of the 11th European Conference on Computer Vision, Springer-Verlag, Berlin, Heidelberg, 2010.

[55] T. Zhang and D. Freedman, *Improving performance of distribution tracking through background mismatch*, IEEE Trans. Pattern Anal. Mach. Intell., 27 (2005), pp. 282–287.

Downloaded 02/13/13 to 129.100.249.50. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.