

Continuous Maximal Flows and Wulff Shapes: Application to MRFs

Christopher Zach, Marc Niethammer, Jan-Michael Frahm
University of North Carolina
Chapel Hill, NC
{cmzach, mn, jmf}@cs.unc.edu

Abstract

Convex and continuous energy formulations for low level vision problems enable efficient search procedures for the corresponding globally optimal solutions. In this work we extend the well-established continuous, isotropic capacity-based maximal flow framework to the anisotropic setting. By using powerful results from convex analysis, a very simple and efficient minimization procedure is derived. Further, we show that many important properties carry over to the new anisotropic framework, e.g. globally optimal binary results can be achieved simply by thresholding the continuous solution. In addition, we unify the anisotropic continuous maximal flow approach with a recently proposed convex and continuous formulation for Markov random fields, thereby allowing more general smoothness priors to be incorporated. Dense stereo results are included to illustrate the capabilities of the proposed approach.

1. Introduction

Convex and continuous approaches to low level vision tasks are appealing for two reasons: (i) the convexity of the formulation ensures global optimality of the obtained solutions; and (ii) setting the problem in a continuous (i.e. non-combinatorial) domain often results in intrinsically data-parallel algorithms, that can be significantly accelerated e.g. by modern graphics processing units. Typically, the main challenge is to find a convex (and optionally continuous) formulation of an apparently non-convex problem. For instance, solving a pairwise Markov random field with multiple labels and linear or convex discontinuity costs requires an embedding of the original formulation into a higher dimensional space (see [6, 13, 11, 12] for combinatorial approaches and [20] for a continuous formulation).

Finding global optimizers for Markov random fields with pairwise and non-convex priors is generally NP-hard, and only approximation algorithms are known. Well-established approaches for MRF optimization in computer vision include belief propagation [26], graph cut meth-

ods [7, 15], message passing approaches [14, 16], and sequential fusion of labeling proposals [17, 25].

One major inspiration for this work is the continuous maximal flow framework proposed in [2], which provides globally optimal and efficient solutions to minimal surface problems for image segmentation (e.g. [9]) and stereo [21]. The original framework for continuous maximal flows uses isotropic, i.e. non-direction dependent, capacity constraints. The flow between neighboring nodes (as induced by the \mathbb{R}^N topology) is uniformly limited in all directions. The utilization of anisotropic capacities allows the flow to prefer certain, spatially varying directions.

It turns out, that the flow field with capacity constraints directly corresponds to the dual vector field employed in efficient minimization of total variation energies [10]. The spatially varying isotropic capacity constraint in [2] recurs as weighted total variation [8]. Anisotropic total variation for image denoising was theoretically analyzed in [19], and joint image smoothing and estimation of local anisotropy orientation was proposed in [3].

This work is outlined as follows: Section 2 introduces anisotropic capacity constraints and analyzes the relationship with continuous minimal surface and maximal flow formulations. An iterative optimization procedure maintaining primal and dual variables is presented as well. Section 3 uses strong duality results from non-smooth convex optimization to deduce the corresponding dual energy, that provides further theoretical insights, and is a useful indicator to stop the iterations. In Section 4 the results obtained in the previous sections are used to derive novel procedures for certain classes of Markov random fields, thus providing an extension and unification of [2] and [20]. Section 5 concludes this work.

2. Anisotropic Continuous Maximal Flows

This section presents an extension of continuous maximal flows with isotropic capacity constraints as proposed in [2] to anisotropic node capacities. Fortunately, the approach used in [2] to prove the correctness and stability of the underlying partial differential equations for

isotropic capacity constraints can be easily generalized to the anisotropic setting as shown in the next sections.

2.1. Wulff Shapes

The aim of this section is to introduce a technique allowing to rewrite non-differentiable terms like $\|\cdot\|$ as better manageable expressions: unconstrained optimization of certain non-differentiable objective functions can be formulated as nested optimization of a scalar product term subject to a feasibility constraint induced by the respective *Wulff shape* (e.g. [19]):

Definition 1 Let $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$ be a convex, positively 1-homogeneous function (i.e. $\phi(\lambda x) = \lambda\phi(x)$ for $\lambda > 0$). The Wulff shape W_ϕ is the set

$$W_\phi := \{y \in \mathbb{R}^N : \langle y, x \rangle \leq \phi(x) \forall x \in \mathbb{R}^N\}, \quad (1)$$

where we denote the inner product in \mathbb{R}^N by $\langle \cdot, \cdot \rangle$.

The Wulff shape W_ϕ is a convex, bounded and closed set. Further $0 \in W_\phi$. Given a Wulff shape W_ϕ the generating function $\phi(\cdot)$ can be recovered by

$$\phi(x) = \max_{y \in W_\phi} \langle y, x \rangle. \quad (2)$$

For consistency with [2] and with the dual energy formulation (Section 3), we will use Eq. 2 in a slightly modified form with negated dual variables, i.e.

$$\phi(x) = \max_{-y \in W_\phi} \langle -y, x \rangle = \max_{y \in -W_\phi} \langle -y, x \rangle, \quad (3)$$

where we introduce the negated Wulff shape,

$$-W_\phi = \{-y : y \in W_\phi\}.$$

An important set of functions ϕ satisfying the convexity and positive 1-homogeneity constraint are norms. The Wulff shape for the Euclidean norm, $\|\cdot\|_2$ is the unit ball with respect to $\|\cdot\|_2$. In general, the Wulff shape for the l^p norm is the unit ball of the dual norm, $\|\cdot\|_q$, with $1/p + 1/q = 1$.

One useful observation about Wulff shapes in the context of this work is the geometry of the Wulff shape for linear combinations of convex and positively 1-homogeneous functions. It is easy to see that the following holds:

Observation 1 Let ϕ and ψ be two convex, positively 1-homogeneous functions, and $k \in \mathbb{R} > 0$. Then, the following relations hold:

$$W_{\phi+\psi} = W_\phi \oplus W_\psi, \quad (4)$$

where $A \oplus B$ denotes the Minkowski sum of two sets A and B . Further,

$$W_{k\phi} = \frac{1}{k}W_\phi = \left\{ \frac{y}{k} : y \in W_\phi \right\} \quad (5)$$

This observation allows us to derive the geometry of Wulff shapes for positive linear combinations of given functions from their respective individual Wulff shapes in a straightforward manner.

In [19] it is shown, that we have $\phi(x) = \langle y, x \rangle$ for non-zero x and $y \in W_\phi$ if and only if

$$y \in \partial W_\phi \text{ and } x \in N_{W_\phi}(y), \quad (6)$$

where ∂W_ϕ denotes the boundary of the Wulff shape and $N_{W_\phi}(y)$ is the set of outward pointing normals at $y \in \partial W_\phi$. Note that this observation is equivalent to the Fenchel-Young equation (see Section 3), but expressed directly in geometric terms.

2.2. Minimal Surfaces and Continuous Maximal Flows

Appleton and Talbot [2] formulated the task of computing a globally minimal surface separating a known source location from the sink location as a continuous maximal flow problem. Instead of directly optimizing the shape of the separating surface ∂A , the indicator function of the respective region A enclosing the source is determined.

In this section we derive the main result of [2] for anisotropic capacity constraints. Let $\Omega \subset \mathbb{R}^N$ be the domain of interest, and $S \subset \Omega$ and $T \subset \Omega$ the source and sink regions, respectively. Further, let $\phi_{\mathbf{x}}$ be a family of convex and positively 1-homogeneous functions for every $\mathbf{x} \in \Omega$. Then the task is to compute a binary valued function $u : \Omega \rightarrow \{0, 1\}$, that is the minimizer of

$$E(u) = \int_{\Omega} \phi_{\mathbf{x}}(\nabla u) \, d\mathbf{x}, \quad (7)$$

such that $u(\mathbf{x}) = 1$ for all $\mathbf{x} \in S$ and $u(\mathbf{x}) = 0$ for all $\mathbf{x} \in T$. Note that we have replaced the usually employed weighted Euclidean norm by the general, spatially varying weighting function $\phi_{\mathbf{x}}$. In the following we will drop the explicit dependence of $\phi_{\mathbf{x}}$ on \mathbf{x} and will only use ϕ . Since ϕ is convex and positively 1-homogeneous, we can substitute $\phi(\nabla u)$ by $\max_{y \in W_\phi} \langle -y, \nabla u \rangle$ (recall Eq. 3) and obtain:

$$E(u) = \int_{\Omega} \max_{-y \in W_\phi} \langle -y, \nabla u \rangle \, d\mathbf{x}. \quad (8)$$

We will collect all *dual* variables $y \in -W_\phi$ into a vector field $\mathbf{p} : \Omega \rightarrow -W_\phi$. Hence, we arrive at the following optimization problem:

$$\min_u \max_{\mathbf{p}} E(u, \mathbf{p}) = \min_u \max_{\mathbf{p}} \int_{\Omega} \langle -\mathbf{p}, \nabla u \rangle \, d\mathbf{x} \quad (9)$$

together with the source/sink conditions on u and the generalized capacity constraints on \mathbf{p} ,

$$-\mathbf{p}(\mathbf{x}) \in W_{\phi_{\mathbf{x}}} \quad \forall \mathbf{x} \in \Omega. \quad (10)$$

strained optimization problems. The primal energy Eq. 7 is generally non-smooth and has constraints for the source and sink regions. Thus, we employ the main result on strong duality from convex optimization to obtain the corresponding dual energy (e.g. [5]).

3.1. Fenchel Conjugates and Duality

The notion of dual programs in convex optimization is heavily based on subgradients and conjugates:

Definition 2 Let $f : \mathbb{R}^N \rightarrow \mathbb{R}$ be a convex and continuous function. y is a subgradient at x_0 if

$$\forall x : f(x_0) + \langle y, x - x_0 \rangle \leq f(x)$$

The subdifferential $\partial f(x)$ is the set of subgradients of f at x , i.e. $\partial f(x) = \{y : y \text{ is a subgradient of } f \text{ at } x\}$.

Intuitively, for functions $f : \mathbb{R} \rightarrow \mathbb{R}$, subgradients of f at x_0 are slopes of lines containing $(x_0, f(x_0))$ such that the graph f is not below the line.

Definition 3 Let $f : \mathbb{R}^N \rightarrow \mathbb{R}$ be a convex and semi-continuous function. The Fenchel conjugate of f , denoted by $f^* : \mathbb{R}^N \rightarrow \mathbb{R}$, is defined by

$$f^*(z) = \sup_y [\langle z, y \rangle - f(y)]. \quad (16)$$

$f^{**} = f$ if and only if f is convex and semi-continuous. From the definition of f^* , we obtain

$$f(y) \geq \langle z, y \rangle - f^*(z) \quad \forall y, z \in \mathbb{R}^N, \quad (17)$$

i.e. the graph of f is above the hyperplane $[y, \langle z, y \rangle - f^*(z)]$ (Fenchel-Young inequality). Equality holds for given y, z if and only if z is a subgradient of f (or equivalently, y is a subgradient of f^* , *Fenchel-Young equation*).

Let $A \in \mathbb{R}^{M \times N}$ be a matrix (or a linear operator in general). Consider the following convex optimization problem for convex and continuous f and g :

$$p = \inf_{y \in \mathbb{R}^N} [f(y) + g(Ay)]. \quad (18)$$

The corresponding dual program is [5]:

$$d = \sup_{z \in \mathbb{R}^M} [-f^*(A^T z) - g^*(-z)]. \quad (19)$$

Fenchel's duality theorem states, that strong duality holds (under some technical condition), i.e. $p = d$. Further, since the primal program is a minimization problem and the dual program is maximized,

$$f(y) + g(Ay) \geq -f^*(A^T z) - g^*(-z) \quad (20)$$

is always true for any $y \in \mathbb{R}^N$ and $z \in \mathbb{R}^M$. This inequality allows to provide an upper bound on the optimality gap (i.e. the distance to the true optimal energy) in iterative algorithms maintaining primal and dual variables.

3.2. Dual Energy for Continuous Maximal Flows

Now we consider again the basic energy to minimize (recall Eq. 7)

$$E(u) = \int_{\Omega} \phi_{\mathbf{x}}(\nabla u(\mathbf{x})) dx.$$

Without loss of generality, we assume Ω to be 3-dimensional, since our applications are in such a setting. Let A denote the gradient operator, i.e. $A(u) = \nabla u$. In the finite setting, A will be represented by a sparse $3|\Omega| \times |\Omega|$ matrix with ± 1 values at the appropriate positions.

The function g represents the primal energy,

$$g(A \circ u) := \int_{\Omega} \phi_{\mathbf{x}}((A \circ u)(\mathbf{x})) dx. \quad (21)$$

$(A \circ u)(\mathbf{x})$ denotes $\nabla u(\mathbf{x})$. Since A is a linear operator and $\phi_{\mathbf{x}}$ is convex, g is convex with respect to u .

The second function in the primal program, f , encodes the restrictions on u . The value attained by $u(\mathbf{x})$ should be 1 for source locations $\mathbf{x} \in S$, and 0 for sink regions, $u(\mathbf{x}) = 0$ for $\mathbf{x} \in T$. W.l.o.g. we can assume $u(\mathbf{x}) \in [0, 1]$ for all $\mathbf{x} \in \Omega$. Without this assumption it turns out that the dual energy is unbounded ($-\infty$), if $\text{div } \mathbf{p}(\mathbf{x}) \neq 0$ for some $\mathbf{x} \in \Omega$ not in the source or the sink region (see below). We choose f as

$$f(u) = \begin{cases} 0 & \text{if } u \text{ complies with the constraints above} \\ \infty & \text{otherwise} \end{cases} \quad (22)$$

We denote the space of functions satisfying the constraints (equivalently, with finite $f(u)$) by \mathcal{U} .

The dual variable is a vectorial function (vector field) $\mathbf{p} : \Omega \rightarrow \mathbb{R}^3$. Since $g(\nabla u)$ is a sum (integral) of $\phi_{\mathbf{x}}$ acting on independent components of ∇u , we have for the conjugate function g^* :

$$g^*(-\mathbf{p}) = \int_{\Omega} \phi_{\mathbf{x}}^*(-\mathbf{p}(\mathbf{x})) dx. \quad (23)$$

Since all $\phi_{\mathbf{x}}$ share the same properties, we drop the index and derive the Fenchel conjugate for any convex, positively 1-homogeneous function $\phi(y)$. First, note that the Wulff shape W_{ϕ} is nothing else than the subdifferential of ϕ at 0, i.e. $W_{\phi} = \partial\phi(0)$:

$$\begin{aligned} z \in \partial\phi(0) &\Leftrightarrow \forall y : \phi(0) + \langle z, y - 0 \rangle \leq \phi(y) \\ &\Leftrightarrow \forall y : \langle z, y \rangle \leq \phi(y) \Leftrightarrow z \in W_{\phi}, \end{aligned}$$

where we made use of $\phi(0 \cdot y) = 0 \cdot \phi(y) = 0$. Hence, we obtain for every $z \in W_{\phi}$ (by the Fenchel-Young equation)

$$\phi^*(z) = \langle z, 0 \rangle - \phi(0) = 0. \quad (24)$$

If $z \notin W_{\phi}$, then there exists a y such that $\langle z, y \rangle > \phi(y)$. Further, $\langle z, \lambda y \rangle > \phi(\lambda y) = \lambda\phi(y)$ for any $\lambda > 0$, and

therefore $\phi^*(z) = \sup_y [\langle z, y \rangle - \phi(y)]$ is unbounded. In summary, we obtain

$$\phi^*(z) = \begin{cases} 0 & \text{if } z \in W_\phi \\ \infty & \text{otherwise,} \end{cases} \quad (25)$$

i.e. the conjugate of ϕ is the indicator function of the respective Wulff shape.

The conjugate of f , f^* , can be computed directly. The argument of f^* is a mapping from Ω to negated divergences, $q : \Omega \rightarrow \mathbb{R}$ with

$$q(\mathbf{x}) = (A^* \circ \mathbf{p})(\mathbf{x}) = -\text{div } \mathbf{p}(\mathbf{x}),$$

where A^* is the adjoint operator to A (recall that $\nabla^* = -\text{div}$). Further, let $\overset{\circ}{\Omega}$ denote the domain without the source and sink sets, i.e. $\overset{\circ}{\Omega} = \Omega \setminus (S \cup T)$, then

$$f^*(q) = \sup_u [\langle q, u \rangle - f(u)] \quad (26)$$

$$= \max_{u \in \mathcal{U}} [\langle q, u \rangle] \quad (27)$$

$$= \max_{u \in \mathcal{U}} \left[\int_{\overset{\circ}{\Omega}} qu \, d\mathbf{x} + \int_S qu \, d\mathbf{x} + \int_T qu \, d\mathbf{x} \right] \quad (28)$$

$$= \int_{\overset{\circ}{\Omega}} \max(0, q) \, d\mathbf{x} + \int_S q \, d\mathbf{x}, \quad (29)$$

since u is fixed to 1 for $\mathbf{x} \in S$, $u(\mathbf{x}) = 0$ for $\mathbf{x} \in T$, and between 0 and 1 for locations neither at the source nor the sink.

In summary, the primal energy to evaluate is given in Eq. 7, and the dual energy (recall Eq. 19) is given by

$$E^*(\mathbf{p}) = \int_S \text{div } \mathbf{p} \, d\mathbf{x} + \int_{\overset{\circ}{\Omega}} -\max(0, -\text{div } \mathbf{p}) \, d\mathbf{x} \quad (30)$$

$$= \int_S \text{div } \mathbf{p} \, d\mathbf{x} + \int_{\overset{\circ}{\Omega}} \min(0, \text{div } \mathbf{p}) \, d\mathbf{x} \quad (31)$$

subject to $\mathbf{p}(\mathbf{x}) \in -W_{\phi_{\mathbf{x}}}$. The two terms in the dual energy have direct interpretations: the first term,

$$\int_S \text{div } \mathbf{p} \, d\mathbf{x} \quad (32)$$

measures the total outgoing flow from the source, while the second term, $\int_{\overset{\circ}{\Omega}} \min(0, \text{div } \mathbf{p}) \, d\mathbf{x}$, penalizes additional

sinks in the interior of Ω . Thus, it corresponds directly to the well-known min-cut/max-flow theorem.

If we do not add a bounds constraint on u , here $u(\mathbf{x}) \in [0, 1]$, then it is easy to see that a strict penalizer on the

divergence outside S and T is obtained. The dual energy contains just the source term,

$$E^*(\mathbf{p}) = \int_S \text{div } \mathbf{p} \, d\mathbf{x} \quad (33)$$

subject to $\mathbf{p} \in -W_{\phi_{\mathbf{x}}}$ and $\text{div } \mathbf{p} = 0$ in $\overset{\circ}{\Omega}$. For \mathbf{p} not satisfying these constraints the dual energy is $-\infty$. Algorithms as the gradient descent updates in Eq. 12 achieve $\text{div}(\mathbf{p}) = 0$ only in the limit, hence the strict penalty on $\text{div}(\mathbf{p}) \neq 0$ is not beneficial. Nevertheless, Eq. 33 is useful for analysis after convergence.

The bounds constraint $u(\mathbf{x}) \in [0, 1]$ is redundant for the primal energy, since $u(\mathbf{x}) \in [0, 1]$ is always true for minima of Eq. 7 subject to the source and sink constraints. In the dual setting this range constraints has the consequence, that any capacity-constrained flow \mathbf{p} is feasible. Any additional flow generated by \mathbf{p} at the source S , but vanishing in the interior is subtracted from the total flow in Eq. 31, hence the overall dual energy remains the same, and bounded and unbounded formulations are equivalent. In practice, we observed substantially faster convergence if u is clamped to $[0, 1]$ after each update step, and a meaningful value for the duality gap is thus obtained.

3.3. Equivalence of the Dual Variables

It remains to show, that the dual vector field \mathbf{p} introduced in Section 2 (Eq. 9) in fact corresponds to the dual vectorial function \mathbf{p} introduced in the dual energy formulation (Eq. 31). This is easy to see: after convergence of the primal-dual method, the primal-dual energy is

$$\begin{aligned} E(u, \mathbf{p}) &= \int_{\Omega} \langle -\mathbf{p}, \nabla u \rangle \, d\mathbf{x} = \int_{\Omega} \text{div } \mathbf{p} \cdot u \, d\mathbf{x} \\ &= \int_{\overset{\circ}{\Omega}} \text{div } \mathbf{p} \cdot u \, d\mathbf{x} + \int_S \text{div } \mathbf{p} \cdot u \, d\mathbf{x} + \int_T \text{div } \mathbf{p} \cdot u \, d\mathbf{x} \\ &= \int_S \text{div } \mathbf{p} \, d\mathbf{x}, \end{aligned}$$

since u is one in S and zero in T , and $\text{div } \mathbf{p} = 0$ in $\overset{\circ}{\Omega}$.

Consequently, the primal-dual energy in terms of \mathbf{p} after convergence is identical to the total flow emerging from the source, i.e. equal to the dual energy. Thus, \mathbf{p} as maintained in the primal-dual formulation is equivalent to the argument of the pure dual one.

4. Application to Markov Random Fields

In this section we address the task of globally optimizing a Markov random field energy, where every pixel in the image domain can attain a label from a discrete and totally ordered set of labels \mathcal{L} . Without loss of generality, we assume

that the labels are represented by non-negative integers, i.e. $\mathcal{L} = \{0, \dots, L - 1\}$. Similar to [22, 21, 2, 20], we formulate the label assignment problem with linearly ordered label sets as a segmentation task in higher dimensions.

4.1. From Markov Random Fields to Anisotropic Continuous Maximal Flows

We denote the (usually rectangular) image domain by \mathcal{I} and particular pixels in bold font, e.g. $\mathbf{x} \in \mathcal{I}$. We will restrict ourselves to the case of two-dimensional image domains. A labeling function $\Lambda : \mathcal{I} \rightarrow \mathcal{L}$, $\mathbf{x} \mapsto \Lambda(\mathbf{x})$ maps pixels to labels. The task is to find a labeling function that minimizes an energy functional comprised of a label cost and a regularization term, i.e. to find the minimizer of

$$E(\Lambda) = \int_{\mathcal{I}} (c(\mathbf{x}, \Lambda(\mathbf{x})) + V(\nabla\Lambda, \nabla^2\Lambda, \dots)) d\mathbf{x}, \quad (34)$$

where $c(\mathbf{x}, \Lambda(\mathbf{x}))$ denotes the cost of selecting label $\Lambda(\mathbf{x})$ at pixel \mathbf{x} and $V(\cdot)$ is the regularization term. Even for convex regularizers $V(\cdot)$ this optimization problem is generally difficult to solve exactly, since the data costs are typically highly non-convex. By embedding the labeling assignment into higher dimensions, and by appropriately restating the energy in Eq. 34, a convex formulation can be obtained. Let Ω be the product space of \mathcal{I} and \mathcal{L} , i.e. $\Omega = \mathcal{I} \times \mathcal{L}$.

In the following, we introduce the level function u ,

$$u(\mathbf{x}, l) = \begin{cases} 1 & \text{if } \Lambda(\mathbf{x}) < l \\ 0 & \text{otherwise.} \end{cases} \quad (35)$$

Since one label must be assigned to every pixel, we require that $u(\mathbf{x}, L) = 1$ for all $\mathbf{x} \in \mathcal{I}$. Further, $u(\mathbf{x}, 0) = 0$ by construction. In particular, the gradients of these function in spatial direction ($\nabla_{\mathbf{x}}$) and in the label direction (∇_l) play an important role, i.e.

$$\nabla_{\mathbf{x}}u = \begin{pmatrix} \partial u / \partial x \\ \partial u / \partial y \end{pmatrix} \quad \nabla_l u = \partial u / \partial l.$$

Finally, ∇u specifies the full gradient of u in spatial and label directions, $\nabla u = (\partial u / \partial x, \partial u / \partial y, \partial u / \partial l)^T$.

We assume, that the regularization energy in Eq. 34 can be written in terms of the gradient of the level function, i.e. $V(\cdot)$ can be formulated as

$$\int_{\Omega} \psi_{\mathbf{x}, l}(\nabla u) d\mathbf{x} dl, \quad (36)$$

where $\psi_{\mathbf{x}, l}$ is a family of convex and 1-positively homogeneous functions, that shapes the regularization term. The same technique as proposed in [20] can be applied to rewrite the data term:

$$\int_{\mathcal{I}} c(\mathbf{x}, \Lambda(\mathbf{x})) d\mathbf{x} = \int_{\mathcal{L}} \int_{\mathcal{I}} c(\mathbf{x}, l) |\nabla_l u(\mathbf{x}, l)| d\mathbf{x} dl \quad (37)$$

$$= \int_{\Omega} c |\nabla_l u| d\mathbf{x} dl, \quad (38)$$

where $|\nabla_l u|$ ensures that costs stay positive even for general choices of u , that attain any real value and are not monotonically increasing in their label argument.

By combining the data fidelity and regularization terms, we can rewrite the original energy Eq. 34 purely in terms of u (where we omit the explicit dependence of \mathbf{x} and l),

$$E(u) = \int_{\Omega} [\psi(\nabla u) + c |\nabla_l u|] d\mathbf{x} dl. \quad (39)$$

Since $\psi_{\mathbf{x}, l}$ and $c|\cdot|$ are convex and 1-positively homogeneous, their sum $\phi_{\mathbf{x}, l}(\cdot) := \psi_{\mathbf{x}, l}(\cdot) + c_{\mathbf{x}, l}|\cdot|$ shares this property. Finally, we can rewrite Eq. 39 as

$$E(u) = \int_{\Omega} \phi(\nabla u) d\mathbf{x} dl \quad (40)$$

Since $\phi_{\mathbf{x}, l}$ is again a family of positively 1-homogeneous functions, the anisotropic continuous maximal flow approach described in Section 2 can be applied. Specifically, the solution u after convergence is essentially binary and monotone in the label direction, thus the optimal labeling function Λ can be easily recovered from u .

4.2. Choices for ψ

In [20] a different approach is taken to analyze the particular choice $\psi_{\mathbf{x}}(\nabla u) = \|\nabla_{\mathbf{x}}u\|_2$ (resulting in a special case of the updates in Eq. 12). By the co-area formula it can be shown, that this choice for $\psi_{\mathbf{x}}$ corresponds to the total variation regularization of the underlying labeling function Λ ,

$$\int_{\mathcal{I}} \|\nabla\Lambda\| d\mathbf{x} = \int_{\Omega} \|\nabla_{\mathbf{x}}u\|_2 d\mathbf{x} dl. \quad (41)$$

As such, using $\psi_{\mathbf{x}}$ for the smoothness term results in the preference of piecewise constant assigned labels. The corresponding Wulff shape $W_{\phi_{\mathbf{x}, l}}$ of $\phi_{\mathbf{x}, l} = \|\nabla_{\mathbf{x}}u\|_2 + \alpha c |\nabla_l u|$ is a cylinder with its height proportional to the label cost $c(\mathbf{x}, l)$. Since the regularization term is isotropic in the image domain, i.e. the smoothness term does not depend on the actual image content, we denote this particular regularization as homogeneous total variation. Note that it is easy to “squeeze” this cylinder depending on the strength and orientation of edges in the reference image in the stereo pair (Figure 2). In order to use only relevant image edges, the structure tensor of the denoised reference image can be utilized to shape the resulting Wulff shape. We use the phase field approach [1] for the Mumford-Shah functional [18] to obtain piecewise smooth images. The positive impact on the resulting depth maps (with the sampling insensitive Birchfield-Tomasi matching cost [4]) is illustrated in Figure 3.

The total variation regularization on the depth map is rather inappropriate if the labels directly correspond to metric depth values in 3D space instead of pure image disparities. A more suitable choice is $\psi = \|\nabla u\|$, if we assume

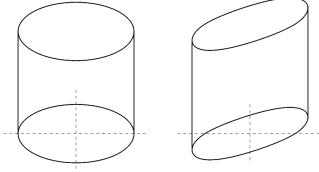
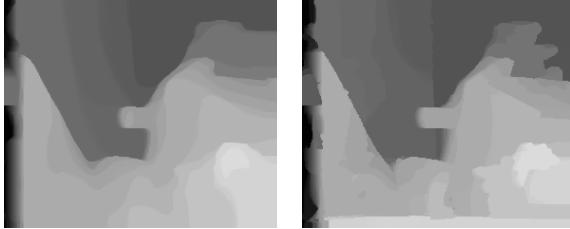


Figure 2. The Wulff shape for homogeneous (left) and edge driven (right) total variation regularizers.



(a) $\alpha = 3$, w/o tensor (b) $\alpha = 3$, with tensor

Figure 3. Depth maps with and without edge driven regularization. (a) is based on homogeneous total variation regularization, and (b) utilizes edge driven regularization. Observe that in (a) the low data fidelity weight yields to substantial loss of detail in the depth map, which is clearly better preserved in (b).

that labels represent equidistantly sampled depth values. This particular choice penalizes the 3D surface area and yields smoother 3D models largely suppressing staircasing artifacts. This specific choice for regularization complies with smoothing the depth map based on an induced surface metric [24]. It can be readily verified that the respective Wulff shape is a capsule-like shape (i.e. a cylinder with half-spheres attached to its base and top face). Figure 4(a–b) visually compares the 3D meshes for the “Dino ring” dataset [23] obtained using total variation regularization $\|\nabla_{\mathbf{x}}u\|$ (Fig. 4(a)) and 3D regularization $\|\nabla u\|$ (Fig. 4(b)). Figure 4(c) and (d) display the 3D models obtained for a brick wall, for which a laser scanned reference range image is available. The RMS for (c) is 5.1cm, and 3.2cm for (d) with respect to the known range image.

The underlying update equations Eq. 12 are very suitable to be accelerated by a modern GPU. Our current CUDA-based implementation executed on a Geforce 8800 Ultra is able to achieve two frames per second for 320×240 images and 32 disparity levels (aiming for a 2% duality gap at maximum).

5. Conclusion

This work analyzes the relationship between continuous maximal flows, Wulff shapes, convex analysis and Markov random fields with convex and homogeneous pairwise priors. It is shown that strong results from isotropic continuous maximal flow carry over to flows with anisotropic capacity constraints. The underlying theory yields extensions to a

recently proposed continuous and convex formulation for Markov random fields with total variation-based smoothness priors. The numerical simplicity and the data-parallel nature of the minimization method allows an efficient implementation on highly parallel devices like modern graphics processing units.

Future work will address the extension of this work to more general classes of MRFs by investigating other linear operators than ∇ , e.g. using $\|\nabla_{\mathbf{x}}\nabla_l u\|$ in the regularization term corresponds to the LP-relaxation of MRFs with a Potts discontinuity prior.

A. The Potential u is Essentially Binary

This section shows, that thresholding of a not necessarily binary primal solution u^* of the anisotropic geodesic energy Eq. 7 yields to a equally globally optimal solution. Let u^* and \mathbf{p}^* be a pair of primal and dual globally optimal solutions for the continuous maximal flow energy Eq. 9. Recall that the optimal dual energy is the total flow leaving the source S (Eq. 33):

$$v^* := E^*(p^*) = \int_S \operatorname{div} \mathbf{p} \, d\mathbf{x}, \quad (42)$$

and that v^* is equal to the primal energy $E(u^*) = \int \phi(\nabla u^*)$. The thresholded result of u^* , u_θ , is given by

$$u_\theta(\mathbf{x}) = \begin{cases} 1 & \text{if } u(\mathbf{x}) \geq \theta \\ 0 & \text{otherwise.} \end{cases}$$

Further, define $A_\theta := \{\mathbf{x} \in \Omega : u^*(\mathbf{x}) \geq \theta\}$. Note that for $\mathbf{x} \in \Omega$ with $\nabla u_\theta(\mathbf{x}) \neq 0$ (or equivalently $\mathbf{x} \in \partial A_\theta$) we have $\nabla u^*(\mathbf{x}) \neq 0$ by construction. Even a stronger relation holds: $\nabla u_\theta(\mathbf{x}) = k(\mathbf{x})\nabla u^*(\mathbf{x})$ for a positive scalar $k(\mathbf{x})$, since both gradients point in the same direction. Further, the capacity constraint is active at $\mathbf{x} \in A_\theta$: $-\mathbf{p}^*(\mathbf{x}) \in \partial W_{\phi_{\mathbf{x}}}$. Thus, for all $\mathbf{x} \in A_\theta$ (omitting the explicit dependence on \mathbf{x})

$$\begin{aligned} \langle -\mathbf{p}^*, \nabla u_\theta \rangle &= \langle -\mathbf{p}^*, k\nabla u^* \rangle \\ &= k \langle -\mathbf{p}^*, \nabla u^* \rangle = k \phi(\nabla u^*) \\ &= \phi(k\nabla u^*) = \phi(\nabla u_\theta). \end{aligned}$$

Overall we have:

$$\begin{aligned} \int_\Omega \phi(\nabla u_\theta) \, d\mathbf{x} &= \int_\Omega \langle -p^*, \nabla u_\theta \rangle \, d\mathbf{x} \\ &= \int_\Omega \operatorname{div} p^* \cdot u_\theta \, d\mathbf{x} = \int_S \operatorname{div} p^* = v^*, \end{aligned}$$

since $\operatorname{div} p^* = 0$ in $\overset{\circ}{\Omega}$ and u_θ is 1 in S and 0 in T . This means, that the binary function u_θ has the same energy as u^* and is likewise optimal.

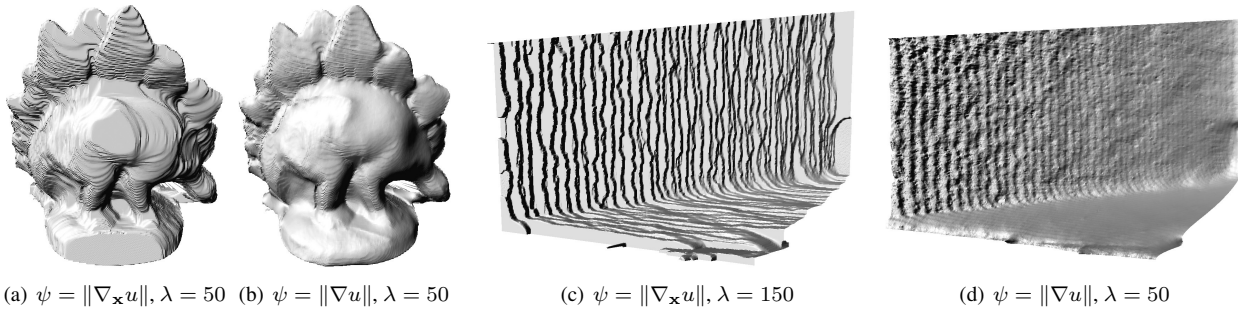


Figure 4. 3D meshed obtained for $\psi = \|\nabla_x u\|$ (a, c) and $\psi = \|\nabla u\|$ (b, d). Clearly, $\|\nabla_x u\|$ favors piece-wise constant results. In (d) the strong matching term causes a visible step structure especially in the foreground. Using a lower value of λ in (c) removes the floor (bottom part of the mesh).

References

- [1] L. Ambrosio and V. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ -convergence. *Comm. Pure Appl. Math.*, 43:999–1036, 1990.
- [2] B. Appleton and H. Talbot. Globally minimal surfaces by continuous maximal flows. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(1):106–118, 2006.
- [3] B. Berkels, M. Burger, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. In *Vision, Modeling, and Visualization Proceedings*, pages 293–300, 2006.
- [4] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(4):401–406, 1998.
- [5] J. Borwein and A. S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*. Springer, 2000.
- [6] Y. Boykov, O. Veksler, and R. Zabih. Markov random fields with efficient approximations. In *Proc. CVPR*, pages 648–655, 1998.
- [7] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, 2001.
- [8] X. Bresson, S. Esedoglu, P. Vanderghenst, J. Thiran, and S. Osher. Fast Global Minimization of the Active Contour/Snake Model. *J. Math. Imag. Vision*, 2007.
- [9] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *IJCV*, 22(1):61–79, 1997.
- [10] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imag. Vision*, 20(1–2):89–97, 2004.
- [11] A. Chambolle. Total variation minimization and a class of binary MRF models. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 136–152, 2006.
- [12] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part II: Levelable functions, convex priors and non-convex cases. *J. Math. Imag. Vision*, 26(3):277–291, 2006.
- [13] H. Ishikawa. Exact optimization for Markov random fields with convex priors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(10):1333–1336, 2003.
- [14] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(10):1568–1583, 2006.
- [15] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2):147–159, 2004.
- [16] N. Komodakis, N. Paragios, and G. Tziritas. MRF optimization via dual decomposition: Message-passing revisited. In *Proc. ICCV*, 2007.
- [17] V. Lempitsky, S. Roth, and C. Rother. FusionFlow: Discrete-continuous optimization for optical flow estimation. In *Proc. CVPR*, 2008.
- [18] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
- [19] S. Osher and S. Esedoglu. Decomposition of images by the anisotropic Rudin-Osher-Fatemi model. *Comm. Pure Appl. Math.*, 57:1609–1626, 2004.
- [20] T. Pock, T. Schoenemann, D. Cremers, and H. Bischof. A convex formulation of continuous multi-label problems. In *Proc. ECCV*, 2008.
- [21] S. Roy. Stereo without epipolar lines: A maximum flow formulation. *IJCV*, 34(2/3):147–161, 1999.
- [22] S. Roy and I. J. Cox. A maximum flow formulation of the N-camera stereo correspondence problem. In *Proc. ICCV*, pages 492–499, 1998.
- [23] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proc. CVPR*, pages 519–526, 2006.
- [24] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Trans. in Image Processing*, 7:310–318.
- [25] W. Trobin, T. Pock, D. Cremers, and H. Bischof. Continuous energy minimization via repeated binary fusion. In *Proc. ECCV*, 2008.
- [26] J. Yedidia, W. T. Freeman, and Y. Weiss. Understanding belief propagation and its generalizations. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2001.