What Metrics Can Be Approximated by Geo-Cuts, or Global Optimization of Length/Area and Flux

Vladimir Kolmogorov
Microsoft Research
Cambridge, UK
vnk@microsoft.com

Yuri Boykov
University of Western Ontario
London, ON, Canada
yuri@csd.uwo.ca

Abstract

In [3] we showed that graph cuts can find hypersurfaces of globally minimal length (or area) under any Riemannian metric. Here we show that graph cuts on directed regular grids can approximate a significantly more general class of continuous non-symmetric metrics. Using submodularity condition [1, 11], we obtain a tight characterization of graph-representable metrics. Such "submodular" metrics have an elegant geometric interpretation via hypersurface functionals combining length/area and flux. Practically speaking, we extend "geo-cuts" algorithm [3] to a wider class of geometrically motivated hypersurface functionals and show how to globally optimize any combination of length/area and flux of a given vector field.

The concept of flux was recently introduced into computer vision by [13] but it was mainly studied within variational framework so far. We are first to show that flux can be integrated into graph cuts as well. Combining geometric concepts of flux and length/area within the global optimization framework of graph cuts allows principled discrete segmentation models and advances the state of the art for the graph cuts methods in vision. In particular, we address the "shrinking" problem of graph cuts, improve segmentation of long thin objects, and introduce useful shape constraints.

1. Introduction and Motivation

Graph cuts approach to problems in computer vision and medical image analysis offers numerically robust global optimization in N-dimensional settings. Discrete graph-cuts methods are easy to implement and their discrete framework is sufficiently flexible to include various forms of regional, boundary, or geometric constraints [2, 3]. Non-parametric implicit representation of hypersurfaces via graph cuts poses no restrictions on topological properties of segments. Besides, underlying max-flow/min-cut optimization algorithms run in seconds on 3D applications [4].

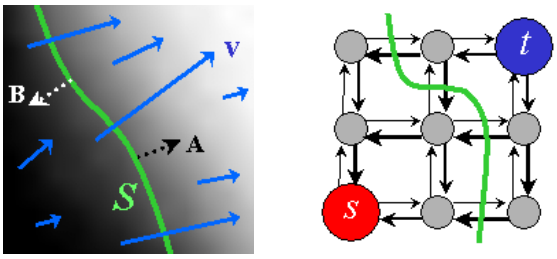
Until recently, one of the major criticism of graph cuts was based on discrete metrication artifacts. Such criticism was largely removed in [3] where we showed that metric properties of cuts can approximate properties of continuous space with any anisotropic Riemannian metric. The concept of geometric length/area under image-based Riemannian metric was extensively studied in segmentation within the framework of geodesic active contours, level-sets, and more recently graph cuts [3]. Introduction of geometrically motivated functionals into graph cuts reduces their metrication artifacts and leads to more principled segmentation models.

The main goal of our current work is to find a geometric interpretation for a general class of energy functionals that graph cuts can minimize in segmentation. This research was motivated by our understanding that minimization of Riemannian length/area as in [3] does not reach the full potential of graph cuts. For example, it is well known that a cost of an s/t cut on a directed graph depends on the cut's "orientation" (see Fig.1b). Indeed, the cut cost includes only directed edges from the s side of the cut to its t side. Swapping the terminals s and t can arbitrarily change the cost of the cut as it would consist of different weights for reverse edges. Normally, geometric (e.g. Riemannian) length/area of a hypersurface does not depend on any specific orientation assigned to it. Therefore, length/area do not describe all geometric properties of general cut metrics.

Flux optimization in image analysis: We found that the concept of length/area has to be combined with the concept of *flux* in order to get a complete geometric interpretation for the general case of cut metrics on directed graphs. Assuming any given field of vectors $\mathbf{v}$ (e.g. image gradients shown in Figure 1a), flux of that vector field through a given continuous hypersurface S is

$$flux(S) = \int_S \langle N, \mathbf{v} \rangle dS$$

where $\langle , \rangle$ is (Euclidean) dot product and N are unit normals



(a) oriented surface (b) s/t cut on directed graph

Figure 1. (a) Surface's orientation, A or B, determines the sign of flux. (b) Swapping the terminals, s and t, changes the cost of a cut.

to surface elements dS consistent with a given orientation. Figure 1a shows two possibles orientations that can be assigned to S. If vectors $\mathbf{v}$ are interpreted as a local speed in a stream of water then the absolute value of flux equals the volume of water passing through the hypersurface in a unit of time. Surface's orientation determines the sign of flux.

In fact, flux optimization has been previously considered in image analysis by a number of researchers. Image segmentation method in [8] uses minimum ratio cycle algorithm to find a contour with the largest ratio of (image gradient) flux and length. Global solution in [8] is limited to 2D contours. In a more general context, flux-based energy was proposed in [13] where the corresponding gradient flow equation was first derived. Practical effectiveness of flux within variational framework (level-sets) was demonstrated in medical image segmentation [13, 6] and in edge integration [9]. In particular, [13] showed that flux optimization helps to segment narrow elongated structures like vessels.

To the best of our knowledge, we are the first to demonstrate that flux can be integrated into the global optimization framework of graph cuts. We show that the geometric concept of flux allows more principled discrete graph-based models and helps to address some practical limitations of graph cuts methods in vision (e.g. "shrinking" problem).

At the same time, we argue here that flux alone has a limited surface regularization effect. Consider the *divergence theorem* for differentiable vector fields

$$\oint_S \langle \mathbf{N}, \mathbf{v} \rangle \, dS = \int_V \operatorname{div} \mathbf{v} \;\; dV$$

where V is the region enclosed inside S. It follows that segmentation based on global maximization of flux is equivalent to thresholding all points where the field's divergence

$$\operatorname{div} \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}$$

is positive. In practice, this implies that flux may have to be combined with additional surface regularization criteria, e.g. geometric length of the segment boundary.

The divergence theorem also reveals a simple way for integrating flux optimization into graph cuts in case of differentiable vector field $\mathbf{v}$. If we can compute $\operatorname{div} \mathbf{v}$ then maximization of flux over surface S can be replaced by maximization of divergence over its interior. Divergence gives us a *regional* bias at each pixel which can be easily incorporated into graph cut framework using t-links [2].

One may argue that practical accuracy of estimating flux via divergence theorem may strongly depend on an exact filter used for estimating divergence. In practice, differentiability of $\mathbf{v}$ may be questionable as well. Later in this paper we avoid any use of the divergence theorem. A class of directed graph constructions for flux optimization is obtained in Section 5 as a simple consequence of our theories that do not explicitly rely on differentiability of $\mathbf{v}$. Not surprisingly, there is strong consistency with the ideas presented in the previous paragraph. We show that all our constructions can be (equivalently) converted to undirected graphs with certain t-links whose weights may be interpreted as specific finite difference schemes for divergence.

Summary of contributions: One of our major contributions is a complete geometric interpretation of discrete cut metrics on directed regular grids via standard continuous concepts of length/area and flux. We obtain a tight characterization of a class of continuous metrics that can be approximated by graph cuts. The result follows from submodularity of graph-representable binary energies [1, 11]. General cut metrics form a subclass of Finsler metrics[1] with asymmetric distance maps (see Fig.2c).

We show that graph-representable (*submodular*) metrics are in one-to-one correspondence with continuous hypersurface functionals combining geometric length with respect to any convex symmetric metric (e.g. Riemannian) and flux with respect to any given vector field. Our results are constructive. We explain how to build a regular grid where discrete cut metric approximates any such hypersurface functional. Minimum cut on this grid corresponds to a hypersurface with a globally optimal sum of length and flux. While global minimization of Riemannian length/area was covered in our earlier work [3], we provide an alternative method that applies to a slightly wider class of convex symmetric metrics. Incorporating flux optimization into graph cuts is one of the major practical contribution of this paper.

Technically speaking, this paper concentrates on theoretical characterization of the class of cut metrics for 2D grids. Nevertheless, all presented general concepts (e.g. Finsler metric, flux, length) extend to 3D and we conjecture that our main theoretical results can be proved in higher dimensions. All practical aspects of our work regarding global optimization of geometrically motivated functionals (length/area and flux) via graph cuts apply to the general N-D setting.

[1] For example, Finsler metric is used in Physics to describe the propagation of light in an inhomogeneous anisotropic medium.

Related work: Basic ideas introduced in [2, 3] work as a foundation for our work. We use implicit representation of continuous hypersurfaces via binary labeling of interior and exterior nodes on regular grids as in [3]. Presentation of the material may assume that the reader is familiar with the main ideas on how regional and boundary cues and hard constraints can be integrated within graph cuts [2].

A very recent work in [10] offers somewhat complementary approach where continuous hypersurfaces are explicitly represented by facets obtained by slicing space with a large number of lines/planes. Formed cells are used as irregularly connected graph nodes. This theoretically interesting construction allows global optimization of a fairly general continuous functionals. Explicit surface representation makes it easy to see that flux and geometric length/area can be encoded. Complete characterization of the corresponding metric properties is an interesting open question.

Organization of the paper: Section 2 is essential for the structure of the paper. It explains our basic mathematical approach and links general ideas and contributions outlined above with more detailed Sections 3 and 4 where necessary technical results are proved. Section 5 provides implementational details for building regular grids where discrete cut metrics approximate any specified "submodular" metric. This construction allows to compute globally optimal surfaces for any given functional combining geometric (e.g. Riemannian) length/area and flux with respect to any given vector field. Section 6 shows promising segmentation results combining image-based Riemannian length/area and flux-based functionals. We show that integrating flux into graph cuts helps to address the "shrinking" problem, to improve segmentation of long thin objects, and to introduce useful shape constraints.

2. Theoretical Framework in 2D

One of our theoretical goals is to characterize a class of continuous metrics that can be approximated by discrete cut metrics on regular grids. It turns out that such graph-representable ("submodular") metrics form a subclass of Finsler metrics. In general, Finsler metric is specified by a smooth non-negative Lagrangian function $L_p(\tau) \geq 0$ describing length of vector τ in a local neighborhood (tangent space) of point p. It is assumed that $L_p(t \cdot \tau) = t \cdot L_p(\tau)$ for any $t \geq 0$. For fixed p, hypersurface $L_p(\tau) = 1$ describes a local distance map visualizing metric properties of space near point p. Figure 2 shows some examples of local distance maps of Finsler metric, including Euclidean and Riemannian metrics as special cases.

In general, local distance maps $L_p(\tau) = 1$ should be convex but they do not have to be centrally symmetric. Indeed, Finsler metric allows $L_p(\tau) \neq L_p(-\tau)$. Unlike standard Riemannian geometry, a given path between A and B

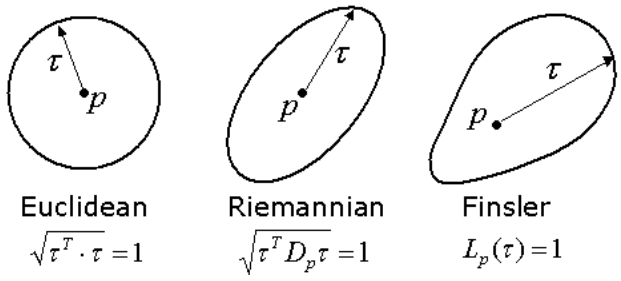


Figure 2. Local distance maps

may have two different Finsler lengths

$$\int_A^B L_s(\tau_s)ds, \qquad (\tau_s \text{ -path's tangent vector})$$

depending on whether the path is followed from A to B or the other way around. In other words, Finsler length of a contour depends on orientation assigned to it. This property is analogous to "directedness" of s/t cuts (Fig. 1b). It suggests that the class of Finsler metrics is sufficiently rich to describe geometric properties of cut metrics.

In Section 3.2 we derive a system of linear equations that relates edge weights on a regular grid to the corresponding cut metric's function L. Then, Section 4 characterizes all graph-representable metrics as functions L for which there are *submodular* edge weights solutions [1, 11].

Our general characterization of graph-representable (or *submodular*) metrics is based on breaking function

$$L_p(\tau) = L_p^+(\tau) + L_p^-(\tau)$$

into symmetric and antisymmetric parts (see Section 3.4). Section 4 shows that metric $L_p(\tau)$ is submodular if and only if distance maps of its symmetric part, $L_p^+(\tau) = 1$, are convex for all p and if its antisymmetric part is harmonic

$$L_p^-(\tau) = \langle \tau^\perp, \mathbf{v}_p \rangle$$

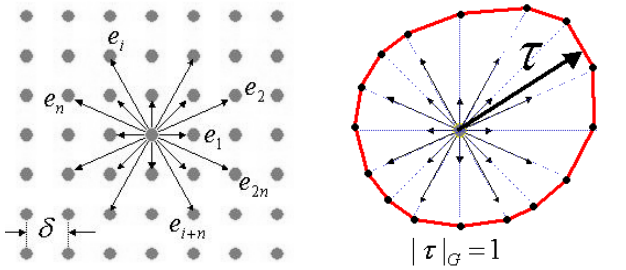
where symbol $\perp$ implies counter clockwise rotation by 90 degrees, and $\mathbf{v}_p$ is some vector defined at each location p.

It follows that Finsler length under general submodular metric can be represented by two terms

$$\int_A^B L_s(\tau_s)ds = \int_A^B L_s^+(\tau_s)ds + \int_A^B \langle \tau_s^\perp, \mathbf{v}_s \rangle ds$$

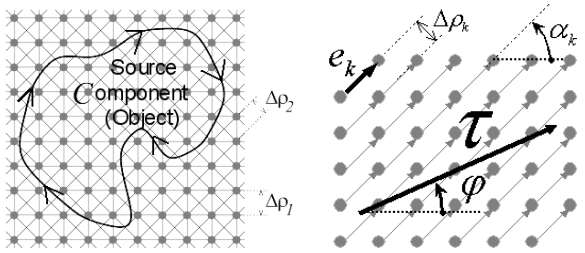
The first term (symmetric part of a cut metric) gives "standard" geometric length of the contour which is independent of its orientation[2]. The second term (antisymmetric part of a cut metric) is equal to *flux* of a given vector field $\{\mathbf{v}_p\}$ through the contour.

[2]Note that anisotropic Riemannian metric $L_p^+(\tau) = \sqrt{\tau^T D_p \tau}$ used in [3] is only one feasible example of symmetric part of a cut metric.



(a) Neighborhood system (b) Cut metric distance map

Figure 3. Edges on a regular 2D grid graph



(a) Contour Orientation (b) k-th family of edges

Figure 4.

3. Preliminaries

In this section we introduce our terminology for regular 2D grids and derive equations that form the basis for studying cut metrics in Section 4. We assume that all nodes are embedded in $\mathcal{R}^2$ in a regular grid-like fashion with cells of size δ. We also assume that all nodes have topologically identical neighborhood systems. One example of a neighborhood system is shown in Figure 3(a).

Neighborhood system of a regular grid $\mathcal{G}$ can be described by a set of directed edges $\mathcal{N}_\mathcal{G} = \{e_k : 1 \le k \le 2n\}$. We will assume that vectors e_k are enumerated in the increasing order of their angular orientation α_k so that $0 \le \alpha_1 < \alpha_2 < ... < \alpha_n < \pi$ and $\alpha_{i+n} = \alpha_i + \pi$ for $1 \le i \le n$. We also assume that e_k is the shortest length vector connecting two grid nodes in the given direction α_k.

Symbol $\tilde{w}$ will denote edge weights. In particular, $\tilde{w}_k(p)$ will be the weight of directed edge e_k at pixel p. Since our goal is graph-based approximation of continuous metrics as the grid gets arbitrarily fine $\delta \to 0$, we can assume that edge weights (and the corresponding cut metric) are homogeneous in small enough regions of the grid. In this case, $\tilde{w}_k(q) = \tilde{w}_k(p)$ for all nodes/pixels q sufficiently close to p and we will simply write $\tilde{w}_k$ to denote weights of all edges with orientation α_k in a small homogeneous region around some assumed pixel p.

3.1 Local Distance Maps for Cut Metric

Approximation of Riemannian metric on graphs was studied in [3] using integral geometry. In contrast, we use linear algebra to study the general case of discrete cut metrics. The results in [3] can be obtained as a special case.

Our major starting point is a system of linear equations describing local distance maps of a discrete cut metric. As mentioned earlier, local properties of a cut metric will be studied by concentrating on sufficiently small homogeneous regions of the grid. Equations for local distance maps of a symmetric cut metric were first derived in [3] in order to visualize these maps for a given set of edge weights. In fact, our approach is to solve for edge weights directly from such distance map equations. Below, we generalize these equations to the case of directed edges e_k.

Consider vector τ of angular orientation φ in a given homogeneous region. According to Figure 4(b), the number of intersections between τ and a family of edges e_k is

$$n_k(\tau) = \frac{|\tau| \cdot |\sin(\alpha_k - \varphi)|}{\Delta\rho_k} = \frac{|\tau| \cdot |e_k| \cdot |\sin(\alpha_k - \varphi)|}{\delta^2}$$

where $\Delta\rho_k = \delta^2/|e_k|$ is the distance between two lines of edges e_k. We would like to estimate the cost of graph edges that τ severs assuming that it is a part of a hypersurface cutting the graph. Since the graph is directed, orientation of the hypersurface (cut) does matter. Throughout this paper we assign orientation to graph partitioning hypersurfaces (contours in 2D) as illustrated in Figure 4(a); we consider the *source* component of a cut to be an object segment and set the clockwise direction to be the orientation of the object boundary. Thus, the object (source component) is always "on the right" of vector τ when it is a part of an oriented cutting hypersurface. Then, τ severs edge e_k only if $\sin(\alpha_k - \varphi) > 0$. Weighted sum of intersections between τ and all families of severed directed edges gives the cut metric length of τ

$$|\tau|_\mathcal{G} = |\tau| \cdot \sum_{k=1}^{2n} w_k \cdot \sin(\alpha_k - \varphi)^+ \qquad (1)$$

where we use "normalized" weight $w_k = \frac{|e_k|}{\delta^2}\tilde{w}_k$. Equation (1) relates cut metric length $|\tau|_\mathcal{G}$ with graph's edge weights.

Figure 3(b) shows one example of a cut metric's local distance map which can be obtained as a solution of equation $|\tau|_\mathcal{G} = 1$. It is easy to check that this map is a polygon with exactly one vertex per edge e_k in the neighborhood system. Such vertices are aligned with the directions of edges. As shown later, such polygons must be convex.

3.2 From Metric to Edge Weights

We would like to find edge weights w_k such that cut metric length $|\tau|_\mathcal{G}$ approximates given Finsler metric $L(\tau)$.

This problem can be slightly simplified since both cut metric (1) and Finsler metric $L(\cdot)$ are homogeneous functions of degree one, i.e. $|t \cdot \tau|_{\mathcal{G}} = t \cdot |\tau|_{\mathcal{G}}$ and $L(t \cdot \tau) = t \cdot L(\tau)$ for any scalar $t > 0$. Introducing unit vectors $\mathbf{u}_\varphi = [\cos\varphi \ \sin\varphi]^T$ at angles φ, the problem is that their cut metric length $|\mathbf{u}_\varphi|_{\mathcal{G}}$ should approximate their given Finsler metric length

$$g(\varphi) := L(\mathbf{u}_\varphi)$$

In general, it is not always possible to find graph $\mathcal{G}$ such that equality $|\mathbf{u}_\varphi|_{\mathcal{G}} = g(\varphi)$ holds for all angles φ since the distance map of any cut metric is a polygon (Figure 3(b)). Instead, we will require that this equality holds for a given finite set of directions $\varphi \in [0, 2\pi]$. In particular, we choose angles $\varphi = \alpha_k$ corresponding to orientations of edges in the neighborhood system $\mathcal{N}_{\mathcal{G}} = \{e_k : 1 \le k \le 2n\}$. Then, (1) yields a system of linear equations for edge weights w_k:

$$\sum_{k=1}^{2n} w_k \cdot \sin(\alpha_k - \alpha_i)^+ = g_i \qquad (2)$$

where $g_i := g(\alpha_i) = L(\mathbf{u}_i)$ and $\mathbf{u}_i := \mathbf{u}_{\alpha_i} = \frac{e_i}{|e_i|}$. There are $2n$ equations ($1 \le i \le 2n$) and $2n$ variables w_k.

Section 4 characterizes graph-representable metrics g such that (2) yields feasible edge weights w_k. In the next subsection we explain what edge weights are "feasible".

3.3 Submodular Edge Weights

It is well-known that graph cut/max flow algorithm can be applied only if edge weights in a graph are non-negative. This places some restriction on metrics g that graph cuts can handle. At the first glance, it may seem that we need equation 2 to have a non-negative solution w. However, we argue that in some cases it is possible to use graph cuts even if some components w_k are negative. Indeed, consider edge e_k connecting pixels p and q ($1 \le k \le n$). We pay some penalty only if p and q are given different segmentation labels x_p and x_q. This penalty is $\tilde{w}_k$ if $x_p = 0$, $x_q = 1$, and $\tilde{w}_{k+n}$ if $x_p = 1$, $x_q = 0$. Therefore, edge (p, q) adds the following pairwise term $V_{pq}(x_p, x_q)$ to the functional that we minimize:

$V_{pq}(0,0)$	$V_{pq}(0,1)$		0	$\tilde{w}_k$
$V_{pq}(1,0)$	$V_{pq}(1,1)$	$=$	$\tilde{w}_{k+n}$	0

It is known (see e.g. [1, 11]) that term V_{pq} can be handled by a broad class of graph constructions if and only if this term is *submodular*, i.e. the sum of diagonal elements is the same or smaller than the sum of off-diagonal elements[3]. This motivates the following definition.

[3]If one of the weights is negative (for example $\tilde{w}_k$), but $\tilde{w}_k + \tilde{w}_{k+n} \ge 0$, then the graph is constructed as follows [11]: we add edges $(source \to p)$, $(q \to p)$, $(p \to sink)$ with non-negative weights so that their combined effect is the same as that of edges $(p \to q)$ and $(q \to p)$ with weights $\tilde{w}_k$ and $\tilde{w}_{k+n}$, respectively.

Definition 3.1. *Edge weight vector* $\{w_k\}$ *is called* submodular *if* $w_k + w_{k+n} \ge 0$ *for any index* $k \in [1, n]$.

Thus, the question what metrics can be approximated via geo-cuts reduces to the following central questions:

- For which g equation (2) has a solution?
- For which g equation (2) has a submodular solution?

An answer to these questions is the main technical contribution of the paper. Our solution in presented in Section 4.

3.4 Symmetric and Antisymmetric Parts

To characterize cut metrics and to solve for edge weights we decompose vectors $\{g_i\}$, $\{w_k\}$ into symmetric and antisymmetric parts. Below we explain basic properties of such decompositions and introduce some notation that is used throughout the paper.

We will work with $2n \times 1$ vectors $f = \{f_k\}$ whose elements correspond to fixed discrete angles $\alpha_1, \ldots, \alpha_{2n}$. It is convenient to extend such vectors and angles periodically, i.e. define $f_{i+2nk} = f_i$, $\alpha_{i+2nk} = \alpha_i$ for $1 \le i \le 2n$, $k \in \mathbb{Z}$.

We say that f is symmetric if $f_i = f_{i+n}$, and antisymmetric if $f_i = -f_{i+n}$. Note that any periodic vector f can be split uniquely into symmetric and antisymmetric parts $f = f^+ + f^-$ where f^+ and f^- are determined as follows:

$$f_i^+ = \tfrac{1}{2}[f_i + f_{i+n}]$$
$$f_i^- = \tfrac{1}{2}[f_i - f_{i+n}]$$

The submodularity condition for a periodic edge weight vector $\{w_k\}$ can then be reformulated as follows:

Lemma 3.2. *Vector* $\{w_k\}$ *is submodular iff* $w_k^+ \ge 0$ *for any* k.

4 Characterization of Submodular Metrics

Two theorems below provide complete characterization of cut metrics from discrete formulation (2) by answering two fundamental questions posted in Section 3.3. Although vector $g = \{g_i\}$ was introduced under the assumption that it corresponds to some Finsler metric, in this section we remove this restriction and consider arbitrary vectors g.

Theorem 4.1. *Equation (2) has a solution if and only if antisymmetric part* $g^- = \{g_i^-\}$ *is harmonic, i.e. there exists vector* $\mathbf{v}$ *such that* $g_i^- = \langle \mathbf{u}_i^\perp, \mathbf{v} \rangle$ *for any* i.

Proof. Equation (2) can be rewritten as

$$Kw = g \qquad (3)$$

where K is $2n \times 2n$ matrix with entries

$$K_{ij} = \sin(\alpha_j - \alpha_i)^+$$

It can be seen that K maps symmetric (antisymmetric) vector w to symmetric (antisymmetric) vector $g = Kw$. Therefore, system (3) is equivalent to two independent equations (subject to constraints that w^+ is symmetric and w^- is antisymmetric):

$$Kw^+ = g^+ \tag{4a}$$

$$Kw^- = g^- \tag{4b}$$

Consider $2n \times 2n$ matrix D with entries

$$D_{ii} = -\frac{\sin(\alpha_{i+1} - \alpha_{i-1})}{\sin(\alpha_{i+1} - \alpha_i)\sin(\alpha_i - \alpha_{i-1})}$$

$$D_{ij} = \frac{1}{|\sin(\alpha_j - \alpha_i)|} \quad \text{if } j - i = \pm 1 \mod 2n$$

$$D_{ij} = 0 \quad \text{for all other entries}$$

It can be seen that

$$DK = KD = \begin{bmatrix} I_n & I_n \\ I_n & I_n \end{bmatrix}$$

where I_n is $n \times n$ identity matrix. Therefore, multiplying equation (4a) by D on the left we obtain

$$w^+ = \frac{1}{2} Dg^+ \tag{5}$$

Conversely, substituting w^+ into equation (4a) we obtain an identity. Thus, equation (4a) has unique symmetric solution given by formula (5).

Consider the second equation. We will periodically extend matrix K as $K_{i+2nk,j+2nl} = K_{ij}$. Then the LHS can be rewritten as

$$[Kw^-]_i = \sum_{j=1}^{n} K_{ij} w_j^- + \sum_{j=n+1}^{2n} K_{ij} w_j^- =$$

$$= \sum_{j=1}^{n} [K_{ij} - K_{i,j+n}] w_j^- = \sum_{j=1}^{n} \langle \mathbf{u}_i{}^\perp, \mathbf{u}_j \rangle w_j^- = \langle \mathbf{u}_i{}^\perp, \mathbf{v} \rangle$$

where we used bilinear form $\langle a^\perp, b \rangle = a_x b_y - a_y b_x$ and

$$\mathbf{v} := \sum\nolimits_{j=1}^{n} \mathbf{u}_j w_j^- \tag{6}$$

Therefore, the antisymmetric part g^- must be harmonic if equation (4b) has an antisymmetric solution. Conversely, if g^- is harmonic ($g_i^- = \langle \mathbf{u}_i{}^\perp, \mathbf{v} \rangle$) then it is easy to construct solution w^- of equation (4b): we just need to choose antisymmetric vector w^- which satisfies equality (6). (If $n = 2$, then w^- is unique, otherwise there are many solutions). □

Theorem 4.2. *Suppose that antisymmetric part of $g = \{g_i\}$ is harmonic, i.e. $g_i^- = \langle \mathbf{u}_i{}^\perp, \mathbf{v} \rangle$ for some vector $\mathbf{v}$. Then*

(1) Equation (2) has a submodular solution w if and only if

$$Dg^+ \geq \mathbf{0} \tag{7}$$

(2) Condition (7) implies that vector g^+ is non-negative.

(3) If vector g^+ is strictly positive, then condition (7) is equivalent to convexity of the distance map for g^+.

Proof. Part (1) is a direct consequence theorem 4.1. Indeed, submodularity is equivalent to the condition $w^+ \geq \mathbf{0}$, which by formula (5) is equivalent to condition (7).

Let us prove part (2). We will do it by assuming that $c = \min_i g_i^+ < 0$ and deriving a contradiction.

Without loss of generality we can assume that this minimum is achieved at $i = 0$. Thus, we have $g_0^+ = g_n^+ = c$. Consider vector

$$h_i = \frac{g_i^+ - c/2}{\sin(\alpha_i - \alpha_0)}, \quad 0 < i < n$$

Let us define $k = \arg\max_{0<i<n} h_i$ and $A = h_k$. Now consider vector H defined as follows:

$$H_i = g_i^+ - c/2 - A\sin(\alpha_i - \alpha_0), \quad i \in \mathbb{Z}$$

We have $H_k = 0$ and $H_i \leq 0$ for any $0 \leq i \leq n$. This implies that $[DH]_k \leq 0$. Therefore,

$$[Dg^+]_k = [DH]_k + [D_{k,k-1} + D_{kk} + D_{k,k+1}] \cdot c/2 < 0$$

We get a contradiction, which proves part (2).

We now assume that g^+ is strictly positive. Then we can consider the distance map of g^+, i.e. the polygon with vertices $A_1, \ldots, A_{2n}$ where A_i is vertex $(\alpha_i, 1/g_i^+)$ in polar coordinates: $\overrightarrow{OA_i} = \frac{1}{g_i^+}\mathbf{u}_i$. We denote $r_i = |\overrightarrow{OA_i}| = \frac{1}{g_i^+}$.

Let us prove that condition $w^+ = \frac{1}{2}Dg^+ \geq \mathbf{0}$ is equivalent to the condition that the distance map is convex. For each $k \in \mathbb{Z}$ we can write

$$w_k^+ = \frac{1}{2}[D_{k,k-1}g_{k-1}^+ + D_{kk}g_k^+ + D_{k,k+1}g_{k+1}^+] =$$

$$= c \cdot [\ \tfrac{1}{2}\ r_k r_{k+1}\sin(\alpha_{k+1} - \alpha_k)$$
$$-\tfrac{1}{2}\ r_{k-1}r_{k+1}\sin(\alpha_{k+1} - \alpha_{k-1})$$
$$+\tfrac{1}{2}\ r_{k-1}r_k\sin(\alpha_k - \alpha_{k-1})\ ]$$

where

$$c = \frac{g_{k-1}^+ g_k^+ g_{k+1}^+}{\sin(\alpha_k - \alpha_{k-1})\sin(\alpha_{k+1} - \alpha_k)} > 0$$

The three terms inside parentheses are $Area(OA_{k-1}A_k)$, $-Area(OA_{k-1}A_{k+1})$, $Area(OA_kA_{k+1})$, respectively (see Fig. 5). Therefore, their sum equals $Area(A_{k-1}A_kA_{k+1})$ which is taken with positive sign if A_k and O are at different sides of line $A_{k-1}A_{k+1}$, and with negative sign otherwise. This proves the desired result. □

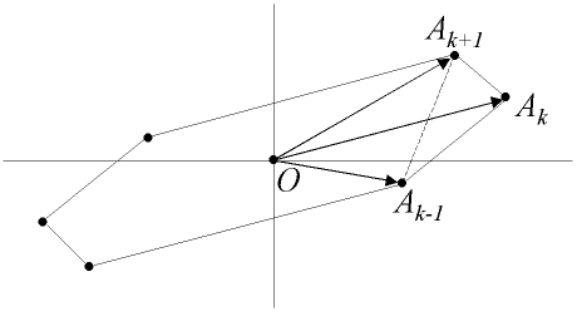


Figure 5. Illustration for the proof that $Dg^+ \geq 0$ is equivalent to convexity of the distance map for strictly positive vector g^+.

5. Summary of Graph Constructions

Theorems 4.1 and 4.2 gave a precise characterization of *submodular* metrics, i.e. vectors $g = \{g_i\}$, for which equation (2) has a submodular solution. Here we summarize how to compute edge weights for any g at pixel p assuming that it is submodular. We discuss graph implementations of symmetric and antisymmetric parts of $g = g^+ + g^-$ separately because they correspond to conceptually different geometric functionals (length/area and flux) and because edges for each part can be combined (added) on a graph.

Length in 2D: Symmetric part g^+ of a general submodular metric should be convex. This part of the metric corresponds to cut's geometric length. To implement any given g^+ we compute vector w^+ according to formula (5) and set the weight of edge e_k originating from p as $\tilde{w}_k^+ = \frac{\delta^2}{|e_k|} w_k^+$ for $1 \leq k \leq 2n$.

Area in 3D: The general result for convex symmetric part g^+ of an arbitrary submodular metric applies to 2D grids only. In order to integrate geometric area into graph cuts in case of 3D grids, it is possible to use n-link construction given in our earlier work [3] assuming that the symmetric part of the metric is Riemannian. We conjecture that it might be possible to construct 3D grids that integrate area with respect to any convex symmetric metric. Theories in Sections 2, 3, and 4 would have to be generalized to 3D.

Flux in ND: Antisymmetric part $g_i^- = \langle \mathbf{u}_i{}^\perp, \mathbf{v} \rangle$ of a general submodular metric at pixel p is defined by a given vector $\mathbf{v}$. This part of the metric corresponds to flux of vector field $\mathbf{v}$ over a cut. Our construction allows optimization of flux with respect to any vector field and applies to grids of arbitrary dimensions.

Assume that N-D grid's neighborhood system is given by vectors $e_1, \ldots, e_{2n}$ such that $e_{k+n} = -e_k$ for $1 \leq k \leq n$. Equation (6) suggests that vector $\mathbf{v}$ should be decomposed into a sum of vectors parallel to edges e_k and find coefficients $w_1^-, \ldots, w_n^-$ such that

$$\sum_{k=1}^{n} w_k^- \mathbf{u}_k = \mathbf{v} \tag{8}$$

We set $w_{k+n}^- := -w_k^-$ for $1 \leq k \leq n$ and compute unnormalized weights $\tilde{w}_k^- = \frac{\delta^d}{|e_k|} w_k^-$ where d is space dimension.

For each edge e_k connecting pixel p to its neighbor q_k we should add directed n-links $(p \rightarrow q_k)$ and $(q_k \rightarrow p)$ with weights $\tilde{w}_k^-$ and $\tilde{w}_{k+n}^-$, respectively. Note that one of these edge weights is negative. Normally, the graph may already have some non-negative weights assigned to its n-links based on symmetric part g^+ of the metric. If the combination of weights from g^+ and some negative weights for g^- is positive at each n-link then we have a valid graph construction. This directed graph construction will work for some combinations of g^+ and g^- but it will put restrictions on the magnitude of vector $\mathbf{v}$ defining g^-.

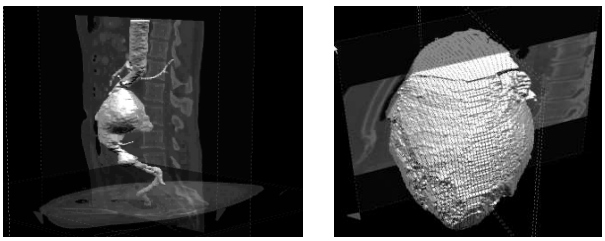
In case of an arbitrary vector field, we can use construction in [11] that allows to implement any submodular n-links. We have $\tilde{w}_k^- + \tilde{w}_{k+n}^- = 0$, therefore n-links $(p \rightarrow q_k)$ and $(q_k \rightarrow p)$ above are submodular for any vector fields. The construction in [11] uses *t-links*, i.e. edges to the terminals (source s and sink t). If $\tilde{w}_k^- < 0$ then we add edges $(s \rightarrow p)$ and $(q_k \rightarrow t)$ with weight $-\frac{1}{2}\tilde{w}_k^-$, otherwise we add edges $(p \rightarrow t)$ and $(s \rightarrow q_k)$ with weight $\frac{1}{2}\tilde{w}_k^-$. (Coefficient $\frac{1}{2}$ is necessary in order to avoid double counting - both pixels p and q_k contribute to the costs).

The implementation of flux via t-links seems consistent with the divergence theorem approach explained in Section 1 even though here we used completely different principles. In particular, we obtain specific weights for t-links based on a decomposition of vector $\mathbf{v}$ in equation (8) and not from a direct computation of divergence of the vector field at pixel p. There are many decompositions of $\mathbf{v}$ and any specific one may lead to different t-links whose weights can be interpreted as a specific estimate of divergence.

Indeed, consider one specific example of decomposing vector $\mathbf{v}$ along grid edges using a natural choice of n-links oriented along the main axes. In 2D, for example, we have $\mathbf{v} = v_x \mathbf{u}_x + v_y \mathbf{u}_y$ where $\mathbf{u}_x$ and $\mathbf{u}_y$ are unit vectors in X and Y direction. It is easy to check that this choice of decomposition (8) leads to t-link weights

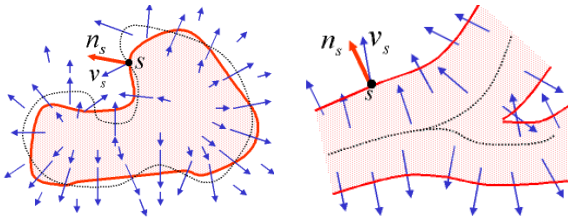
$$t_p = \frac{\delta^2}{2} \left[(v_x^{right} - v_x^{left}) + (v_y^{up} - v_y^{down}) \right] \tag{9}$$

where v_x^{right} is the X-component of vector $\mathbf{v}$ taken at the right neighbor of pixel p and all other terms are defined similarly. We add edge $(s \rightarrow p)$ with weight $-t_p$ if $t_p < 0$, or edge $(p \rightarrow t)$ with weight t_p otherwise. Obviously, equation (9) is a particular finite difference scheme for computing divergence of vector field $\mathbf{v}$.



(a) Vessels and aneurism (b) Heart extraction

Figure 6. 3D experimental results



(a) From prior shape(s) (b) From a center line

Figure 7. Shape constraints using flux

6. Experimental Results

In general, flux is a good counter-balance to length/area. Flux causes stretching while length/area causes shrinking. Specific applications may need to balance them correctly.

Figure 6(a) demonstrates segmentation of a long vessel with an aneurism. We use graph-representable surface functional based on a weighted combination of image gradients flux and image-based Riemannian area. Note that functionals using only area may cut short any elongated structures due to "shrinking": a short expensive boundary may (aggregatively) cost less than a very long cheap one. The divergence theorem explains why flux helps to avoid this problem in the context of global optimization via graph cuts. Divergence of image gradients is positive right inside the object boundary and negative right outside of it. Therefore, we have regional t-links all along the vessel boundary helping to avoid shrinking. Flux sensitivity to orientation also helps the vessel's segment to avoid sticking to the wrong boundary even though the gradients on a spine near by are stronger than those on the vessel. Presence of length/area in the functional helps to add regularity.

We also discuss a novel way of using flux for a shape constraint. Models using distance maps for shape priors are popular [12, 7, 5]. Our shape constraint for N-D image segmentation uses flux of distance map gradients. Two examples of such vector fields are shown in Figure 7. The gradients are for a signed distance map from a given prior shape (dotted line) or from a center line. The magnitude of vectors could be reduced with a distance from a shape. A vector field could be also obtained (learned) as a sum of normals of training shapes that pass through point p. We propose the the following submodular "shape functional"

$$\oint_S \left(L^+(\tau_s) - \langle \mathbf{N}, \mathbf{v} \rangle\right) dS$$

where the second term encourages the surface normals to align with the largest vectors in the field and the first term attempts to regularize via (e.g. Euclidean) length. The minima of this functional (zero) is achieved exactly at the prior shape and it is non-negative elsewhere. Example in Figure 6(b) shows the whole heart extracted from a 3D volume using shape functional above. The vector field was obtained from a center point provided. Without the shape prior, segmentation of the heart as a whole is problematic as the 3D data of heart contains many internal structures with high image gradients. Note that our shape constraint have a reasonable degree of scale invariance.

References

[1] E. Boros and P. L. Hammer. Pseudo-boolean optimization. *Discrete Appl. Math.*, 123(1-3):155–225, Nov. 2002.

[2] Y. Boykov and M.-P. Jolly. *Interactive graph cuts* for optimal boundary & region segmentation of objects in N-D images. In *International Conference on Computer Vision*, volume I, pages 105–112, July 2001.

[3] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *International Conference on Computer Vision*, volume I, pages 26–33, 2003.

[4] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, September 2004.

[5] D. Cremers and C. Schnorr. Statistical shape knowledge in variational motion segmentation. *Image and Vision Computing*, 21(1):77–86, Jan 2003.

[6] M. Descoteaux, D. L. Collins, and K. Siddiqi. Geometric flows for segmenting vasculature in mri: Theory and validation. In *MICCAI (1)*, pages 500–507, 2004.

[7] P. Dimitrov, J. N. Damon, and K. Siddiqi. Flux invariants for shape. In *CVPR*, 2003.

[8] I. H. Jermyn and H. Ishikawa. Globally optimal regions and boundaries as minimum ratio weight cycles. *PAMI*, 23(10):1075–1088, October 2001.

[9] R. Kimmel and A. M. Bruckstein. Regularized Laplacian zero crossings as optimal edge integrators. *International Journal of Computer Vision*, 53(3):225–243, 2003.

[10] D. Kirsanov and S. J. Gortler. A discrete global minimization algorithm for continuous variational problems. *Harvard Computer Science Technical Report*, TR-14-04, May 2004.

[11] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(2):147–159, Feb. 2004.

[12] M. E. Leventon. *Statistical Models in Medical Image Analysis*. PhD thesis, MIT, Boston, MA, 2000.

[13] A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *PAMI*, 24(12):1565–1578, December 2002.