ECCV 2006 tutorial on

Graph Cuts vs. Level Sets

part I

Basics of Graph Cuts

Yuri Boykov  Daniel Cremers  Vladimir Kolmogorov

University of Western Ontario  University of Bonn  University College London
Graph Cuts versus Level Sets

- Part I: Basics of *graph cuts*
- Part II: Basics of *level-sets*
- Part III: Connecting *graph cuts* and *level-sets*
- Part IV: Global vs. local optimization algorithms
Graph Cuts versus Level Sets

Part I: Basics of graph cuts

- Main idea for min-cut/max-flow methods and applications
- Implicit and explicit representation of boundaries
- Graph cut energy (different views)
  - submodularity, geometric functionals, posterior energy (MRF)
- Extensions to multi-label problems
  - convex and non-convex (robust) interactions, a-expansions
**1D Graph cut** $\iff$ **shortest path on a graph**

**Example:** find the shortest closed contour in a given domain of a graph.

Compute the shortest path $p \rightarrow p$ for a point $p$. Repeat for all points on the gray line. Then choose the optimal contour.

Graph Cuts approach: Compute the minimum cut that separates red region from blue region.
Graph cuts for optimal boundary detection
(simple example à la Boykov&Jolly, ICCV’01)

Minimum cost cut can be computed in polynomial time
(max-flow/min-cut algorithms)
Minimum $s$-$t$ cuts algorithms

- Augmenting paths [Ford & Fulkerson, 1962]
- Push-relabel [Goldberg-Tarjan, 1986]
“Augmenting Paths”

- Find a path from S to T along non-saturated edges
- Increase flow along this path until some edge saturates

A graph with two terminals
"Augmenting Paths"

- Find a path from S to T along non-saturated edges
- Increase flow along this path until some edge saturates
- Find next path...
- Increase flow...

A graph with two terminals
“Augmenting Paths”

- Find a path from $S$ to $T$ along non-saturated edges
- Increase flow along this path until some edge saturates

A graph with two terminals

Iterate until …
all paths from $S$ to $T$ have at least one saturated edge

MAX FLOW $\iff$ MIN CUT
Optimal boundary in 2D

“max-flow = min-cut”
Optimal boundary in 3D

3D bone segmentation (real time screen capture)
Graph cuts applied to multi-view reconstruction

Calibrated images of Lambertian scene

3D model of scene

CVPR’05 slides from Vogiatzis, Torr, Cippola
Graph cuts applied to multi-view reconstruction

\[ \rho(x) \]

photoconsistency

CVPR’05 slides from Vogiatzis, Torr, Cippola
Estimating photoconsistency in a narrow bad

- Occlusion

1. Get nearest point on outer surface
2. Use outer surface for occlusions
2. Discard occluded views

CVPR’05 slides from Vogiatzis, Torr, Cippola
Graph cuts applied to multi-view reconstruction

The cost of the cut integrates photoconsistency over the whole space

CVPR’05 slides from Vogiatzis, Torr, Cippola
Graph cuts applied to multi-view reconstruction

surface of good photoconsistency

visual hull (silhouettes)

CVPR’05 slides from Vogiatzis, Torr, Cippola
Graph cuts for video textures

Graph-cuts video textures
(Kwatra, Schodl, Essa, Bobick 2003)

3D generalization of “image-quilting” (Efros & Freeman, 2001)
Graph cuts for video textures

Graph-cuts video textures
(Kwatra, Schodl, Essa, Bobick 2003)

original short clip  synthetic infinite texture
Cuts on directed graphs

- Cost of a cut includes only edges from the source to the sink components
- Cut’s cost (on a directed graph) changes if terminals are swapped

Swapping terminals $s$ and $t$ is similar to switching surface orientation
Cuts on directed graphs

(a) directed graph  (b) image  (c) undir. result  (d) dir. result
Segmentation of elongated structures
Simple “shape priors”

[Funka-Lea et al.’06]  
“blob” prior

\[ w_{p \rightarrow q} \propto f(\nabla I) + c \cdot (1 - \cos(\alpha)) \]
Adding regional properties
(another segmentation example à la Boykov&Jolly’01)

regional bias example

suppose $I^s$ and $I^t$ are given “expected” intensities of object and background

$$D_p(s) = \text{const} - |I_p - I^s|$$

$$D_p(t) = \text{const} - |I_p - I^t|$$

NOTE: hard constrains are not required, in general.
Adding regional properties
(another segmentation example à la Boykov&Jolly’01)

“expected” intensities of object and background $I^s$ and $I^t$ can be re-estimated

$D_p(s) = const - |I_p - I^s|$

$D_p(t) = const - |I_p - I^t|$

EM-style optimization of piece-wise constant Mumford-Shah model
Adding regional properties
(another example à la Boykov&Jolly, ICCV’01)
Adding regional properties
(another example à la Boykov&Jolly, ICCV’01)

More generally, regional bias can be based on any intensity models of object and background

\[ D_p(L_p) = -\ln \Pr(I_p \mid L_p) \]

given object and background intensity histograms
Iterative learning of regional models

- GMMRF cuts (Blake et al., ECCV04)
- Grab-cut (Rother et al., SIGGRAPH 04)

parametric regional model – Gaussian Mixture (GM)
designed to guarantee convergence
“Shrinking” bias

- Minimization of non-negative boundary costs (sum of non-negative n-links) gives regularization/smoothing of segmentation results
- Choosing n-link costs from local image gradients helps image-adaptive regularization
- May result in over-smoothing or “shrinking”
  - Typical for all surface regularization techniques
  - Graph cuts are no different from snakes or level-sets on that
“Shrinking” bias

Optimal cut depending on the size of the hard constrained region
“Shrinking” bias

Image intensities on one scan line

Shrinking (or under-segmentation) is allowed by “intensity gradient ramp”
“Shrinking” bias

Surface regularization approach to multi-view stereo

CVPR’05 slides from Vogiatzis, Torr, Cippola
Regional term can counter-act shrinking bias

All voxels in the center of the scene are connected to the object terminal

“Ballooning” force
- favouring bigger volumes that fill the visual hull


CVPR’05 slides from Vogiatzis, Torr, Cippola
“Shrinking” bias

CVPR’05 slides from Vogiatzis, Torr, Cippola
Uniform ballooning

CVPR’05 slides from Vogiatzis, Torr, Cippola
Regional term based on Laplacian zero-crossings (*flux*)

- Can be seen as intelligent ballooning
  - Vasilevsky and Sidiqqi, 2002
  - R. Kimmel and A. M. Bruckstein 2003

**Basic idea**

Image intensities on a scan line

Laplacian of intensities
Integrating Laplacian zero-crossings into graph cuts (Kolmogorov & Boykov’05)

The image is courtesy of David Fleet
University of Toronto
“Implicit” vs. “Explicit” graph cuts

- Most current graph cuts technique **implicitly** use surfaces represented via binary (interior/exterior) labeling of pixels
“Implicit” and “Explicit” graph cuts

- Except, a recent explicit surfaces representation method
  - Kirsanov and Gortler, 2004
“Explicit” Graph Cuts

- for multi-view reconstruction
  - Lempitsky et al., ECCV 2006
- Explicit surface patches allow local estimation of visibility when computing globally optimal solution

Compare with Vogiatzis et al., CVPR’05 approach for visibility
“Explicit” Graph Cuts

Regularization + uniform ballooning

some details are still over-smoothed

Lempitsky et al., ECCV 2006
“Explicit” Graph Cuts

Regularization + intelligent ballooning

Low noise and no shrinking

Boykov and Lempitsky, 2006
“Explicit” Graph Cuts

Kutulakos and Seitz, 2000
Three ways to look at energy of graph cuts

I: Binary submodular energy
II: Approximating continuous surface functionals
III: Posterior energy (MAP-MRF)
Simple example of energy

\[
E(L) = \sum_p -D_p(L_p) + \sum_{pq \in N} w_{pq} \cdot \delta(L_p \neq L_q)
\]

- **Regional term**
- **Boundary term**

**t-links** \( t \)-links

**n-links** \( n \)-links

**a cut**

**binary object segmentation**

- \( D_p(t) \)
- \( D_p(s) \)
Graph cuts for minimization of submodular binary energies

\[ E(L) = \sum_p E_p(L_p) + \sum_{pq \in N} E(L_p, L_q) \]

- **Regional term**
- **Boundary term**

\(L_p \in \{s, t\}\)

**Tight characterization of binary energies** that can be globally minimized by \(s-t\) graph cuts is known

\[ E(L) \text{ can be minimized by } s-t \text{ graph cuts} \iff E(s,s) + E(t,t) \leq E(s,t) + E(t,s) \]

**Non-submodular cases** can be addressed with some optimality guarantees, e.g. \(QPBO\) algorithm reviewed in Boros and Hummer, 2002, nice slides by Kolmogorov at CVPR and Oxford-Brooks 2005

more in PART IV
Graph cuts for minimization of continuous surface functionals

\[ E(C) = \int_C g(\cdot) \, ds + \int_C \left\langle \tilde{N}, \tilde{\nabla}_x \right\rangle \, ds + \int_{\Omega(C)} f(x) \, dp \]

- **Geometric length**
  - any convex, symmetric metric \( g \)
  - e.g. Riemannian

- **Flux**
  - any vector field \( \mathbf{v} \)

- **Regional bias**
  - any scalar function \( f \)

- **Tight characterization of energies of binary cuts** \( C \) as functionals of continuous surfaces

[Boykov&Kolmogorov, ICCV 2003]
[Kolmogorov&Boykov, ICCV 2005]

more in PART III
Graph cuts for minimization of posterior energy

- Greig at. al. [IJRSS, 1989]
  - Posterior energy (MRF, Ising model)

\[
E(L) = \sum_p -\ln \Pr(D_p | L_p) + \sum_{pq \in N} V_{pq}(L_p, L_q)
\]

Likelihood (data term) \quad Spatial prior (regularization) \quad L_p \in \{s, t\}

Example: binary image restoration
Graph cuts algorithms can minimize multi-label energies as well.
Multi-scan-line stereo
with $s$-$t$ graph cuts (Roy& Cox’98)
Multi-scan-line stereo with $s$-$t$ graph cuts (Roy&Cox’98)
s-t graph-cuts for multi-label energy minimization

- Modification of construction by Roy&Cox 1998

\[
E(L) = \sum_p -D_p(L_p) + \sum_{pq \in N} V(L_p, L_q) \quad L_p \in R^1
\]

Linear interactions

“Convex” interactions
Pixel interactions $V$:

“convex” vs. “discontinuity-preserving”

"Convex" Interactions $V$

Robust "discontinuity preserving" Interactions $V$

"linear" model

$V(dL)$

$dL = L_p - L_q$

$Potts model$

$V(dL)$

$dL = L_p - L_q$
Pixel interactions: “convex” vs. “discontinuity-preserving”
Robust interactions

- NP-hard problem (3 or more labels)
  - two labels can be solved via $s$-$t$ cuts (Greig et al., 1989)

- $\alpha$-expansion approximation algorithm
  (Boykov, Veksler, Zabih 1998, 2001)
  - guaranteed approximation quality (Veksler, thesis 2001)
    - within a factor of 2 from the global minima (Potts model)
  - applies to a wide class of energies with robust interactions
    - Potts model (BVZ 1989)
    - “Metric” interactions (BVZ 2001)
    - “Submodular” interactions (e.g. Boros and Hummer, 2002, KZ 2004)
a-expansion algorithm

1. Start with any initial solution
2. For each label “a” in any (e.g. random) order
   1. Compute optimal a-expansion move (s-t graph cuts)
   2. Decline the move if there is no energy decrease
3. Stop when no expansion move would decrease energy
**a-expansion move**

Basic idea: break multi-way cut computation into a **sequence of binary s-t cuts**
a-expansion moves

In each a-expansion a given label “a” grabs space from other labels.

For each move we choose expansion that gives the largest decrease in the energy: **binary optimization problem**
\(a\)-expansions:
examples of \textit{metric} interactions

\[ V(\alpha, \beta) \]

\textbf{Potts} \( V \)
\( \alpha - \beta \)

\textbf{“noisy shaded diamond”}

\textbf{Truncated “linear”} \( V \)
Multi-way graph cuts

Multi-object Extraction

Obvious generalization of binary object extraction technique (Boykov, Jolly, Funkalea 2004)
Multi-way graph cuts

Stereo/Motion with slanted surfaces
(Birchfield & Tomasi 1999)

Labels = parameterized surfaces

EM based: E step = compute surface boundaries
M step = re-estimate surface parameters
Multi-way graph cuts

stereo vision

original pair of “stereo” images

depth map

KZ 2002
Multi-way graph cuts

Graph-cut textures
(Kwatra, Schodl, Essa, Bobick 2003)

similar to “image-quilting” (Efros & Freeman, 2001)
Multi-way graph cuts

Graph-cut textures
(Kwatra, Schodl, Essa, Bobick 2003)
a-expansions vs. simulated annealing

- annealing: 9 hours, 24.7% err
- a-expansions (BVZ 89,01): 90 seconds, 5.8% err

Graph showing the smoothness energy over time for both methods. The x-axis represents time in seconds, ranging from 1 to 100,000, and the y-axis represents smoothness energy, ranging from 0 to 100,000.