Interactive Graph Cuts

for Optimal Boundary & Region Segmentation of Objects in N-D Images

Yuri Y. Boykov and Marie-Pierre Jolly
From [3]
Vision problems as image labeling

- depth (stereo)
- object index (segmentation)
- original intensity (image restoration)
Labeling problems can be cast in terms of energy minimization

\[ E(L) = \sum_{p \in P} D_p(L_p) + \sum_{p,q \in N} V_{p,q}(L_p, L_q) \]

- \( L : P \rightarrow L_p \) \hspace{1cm} Labeling of pixels
- \( D_p : L_p \rightarrow \mathbb{R} \) \hspace{1cm} Penalty for pixel labeling
- \( V_{p,q} : P \times P \rightarrow \mathbb{R} \) Interaction between neighboring pixels. Smoothing term.
Energy minimization can be solved with graph cuts
• Energy function and graph construction
• Min-cut of graph minimizes energy
• Summar max-flow/min-cut algorithms
• Rest of bone segmentation example
Some Notation

\[ P = \text{set of pixels } p \]

\[ N = \text{set of unordered pairs } (p, q) \text{ of neighbors in } P \]

\[ L = (L_1, \ldots, L_p, \ldots, L_{|P|}) \text{ Binary vector representing a binary segmentation} \]

\[ G = (V, E) \text{ graph with nodes, } V, \text{ and edges, } E \]

\[ B = \text{set of user defined background pixels} \]

\[ O = \text{set of user defined object pixels} \]
Labeling problems can be cast in terms of energy minimization

\[ E(L) = \sum_{p \in P} D_p(L_p) + \sum_{p, q \in N} V_{p, q}(L_p, L_q) \]

\( L : P \rightarrow L_p \)  
Labeling of pixels

\( D_p : L_p \rightarrow \mathbb{R} \)  
Penalty for pixel labeling

\( V_{p, q} : P \times P \rightarrow \mathbb{R} \)  
Interaction between neighboring pixels.
Smoothing term.
Their Energy Function

\[ E(L) = \lambda \cdot \sum_{p \in P} R_p(L_p) + \sum_{(p,q) \in N} B(p, q) \cdot \delta(L_p, L_q) \]

\( D_p(L_p) \) becomes regional term \( R_p(L_p) \)

\( V_{p,q} \) becomes boundary term \( B(p, q) \cdot \delta(L_p, L_q) \)
Regional term

\[ E(L) = \lambda \cdot \sum_{p \in P} R_p(L_p) + \sum_{(p,q) \in N} B(p, q) \cdot \delta(L_p, L_q) \]

Penalize pixel label based on local properties

Negative log-likelihood of intensity

\[ R_p(obj) = -\ln Pr(I_p|\mathcal{O}) \]

\[ R_p(bkq) = -\ln Pr(I_p|\mathcal{B}) \]
Boundary term: penalize dissimilar neighbors

\[ E(L) = \lambda \cdot \sum_{p \in P} R_p(L_p) + \sum_{(p,q) \in N} B(p, q) \cdot \delta(L_p, L_q) \]

\[ B(p, q) \propto e^{-\frac{(I_p - I_q)^2}{2\sigma^2}} \cdot \frac{1}{d_{\text{ist}}(p, q)} \]
Boundary term: penalize dissimilar neighbors

\[ E(L) = \lambda \cdot \sum_{p \in P} R_p(L_p) + \sum_{(p,q) \in N} B(p, q) \cdot |L_p - L_q| \]

\[ B(p, q) \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\text{dist}(p, q)} \]
Graph construction: cost of n-links

c(p, q) = B(p, q) if (p, q) ∈ N

B(p, q) ∝ \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{dist(p, q)}
Graph construction: cost of t-link \((p, S)\)

If \(p \in \mathcal{O}\) then \(c(p, q) = K\)

\[
K = 1 + \max_{p \in P} \sum_{q: (p, q) \in N} B(p, q)
\]
Graph construction: cost of t-link \((p, S)\)

If \(p \notin O \cup B\) then \(c(p, q) = \lambda \cdot R_p(bkg)\)
Graph construction: cost of t-link \((p, S)\)

If \(p \in \mathcal{B}\) then \(c(p, q) = 0\)
Claim: min-cut of graph minimizes energy

- Min-cut on G is a feasible cut
- Each feasible cut has a unique binary segmentation
- Segmentation associated with min-cut that satisfies user defined constraints minimizes the energy function
Summary of max-flow/min-cut algorithms

• Augmenting paths (Ford and Fulkerson)
• Push-relabel (Goldberg and Tarjan)
• Their implementation (see [2])
From [3]
Citations

