Video Image Segmentation with Graphical models

Jieyu Zhao
Outline

- Fundamentals
- Deterministic Methods
- Stochastic Methods
- Some Results
- Conclusions
Video Image Segmentation

**Goal:** to label the image regions with salient homogeneous properties, such as color, texture, motion or spatio-temporal structures.

The labeling algorithms based on graphical models become popular in recent years.

- Deterministic and stochastic
Deterministic Algorithms

- Belief Propagation, which infers marginal probabilities at the nodes of the graph by exchanging of messages initially designed on trees and later generalized

- Minimum Graph Cut, popular deterministic method maps the image segmentation task into a Max-Flow/Min Cut problem

- Other related approaches, such as normalized cut
Stochastic Algorithms

Mainly based on the Gibbs sampler, a Markov chain Monte Carlo algorithm

- Markov random field approaches
- random walk and diffusion approaches
- the Potts models, the Swendsen-Wang method.

Stochastic approaches are usually powerful but time-consuming
Representation

image represented with a weighted graph, vertices reflect the states of image pixels and weighted edges represent the relationship between pixels.

4-neighbour structure, weights represent the similarities.

Segmentation ~ Min Cut
Maximum flow / Minimum cut

“Max flow”: maximize the sum $\sum u f(u,t)$

“Min cut”: Delete the "best" set of edges to disconnect $t$ from $s$, with the smallest capacity
A weighted graph -- material flowing through the edges (railways, water pipelines)

Maximum flow: maximize the sum $\sum u f(u,t)$
A cut is a node partition \((S, T)\) such that \(s\) is in \(S\) and \(t\) is in \(T\).

\[
\text{capacity}(S, T) = \text{sum of weights of edges leaving } S.
\]
a min cut

Cut capacity = 28  ⇒  Flow value ≤ 28

Flow value = 28
**Max-flow min-cut theorem:** The value of the max flow is equal to the capacity of the min cut.

**Augmenting path theorem:** A flow $f$ is a max flow if and only if there are no augmenting paths.

The following are equivalent:

(i) $f$ is a max flow.
(ii) There is no augmenting path relative to $f$.
(iii) There exists a cut whose capacity equals the value of $f$. 
Augmenting path = path in residual graph.

- Increase flow along forward edges.
- Decrease flow along backward edges.

Flow $f(e)$.
Edge $e = v \rightarrow w$

original graph

residual graph

"Undo" flow sent.
$w \rightarrow v$
Image Segmentation Using Min Cut

- Calculating weighted graph
- Setting some seed points, automatically or interactively
- Max Flow Algorithm

Tends to have small and biased segmentation
Improved by the normalized cut:

\[ N_{\text{cut}}(A, B) = \frac{\text{cut}(A, B)}{\text{volume}(A)} + \frac{\text{cut}(A, B)}{\text{volume}(B)} \]
# History of Worst-Case Running Times

<table>
<thead>
<tr>
<th>Year</th>
<th>Discoverer</th>
<th>Method</th>
<th>Asymptotic Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1951</td>
<td>Dantzig</td>
<td>Simplex</td>
<td>$E^2 V U \uparrow$</td>
</tr>
<tr>
<td>1955</td>
<td>Ford, Fulkerson</td>
<td>Augmenting path</td>
<td>$E^2 V U \uparrow$</td>
</tr>
<tr>
<td>1970</td>
<td>Edmonds-Karp</td>
<td>Shortest path</td>
<td>$E^2 V$</td>
</tr>
<tr>
<td>1970</td>
<td>Edmonds-Karp</td>
<td>Max capacity</td>
<td>$E \log U \left( E + V \log V \right) \uparrow$</td>
</tr>
<tr>
<td>1970</td>
<td>Dinitz</td>
<td>Improved shortest path</td>
<td>$E V^2$</td>
</tr>
<tr>
<td>1972</td>
<td>Edmonds-Karp, Dinitz</td>
<td>Capacity scaling</td>
<td>$E^2 \log U \uparrow$</td>
</tr>
<tr>
<td>1973</td>
<td>Dinitz-Gabow</td>
<td>Improved capacity scaling</td>
<td>$E V \log U \uparrow$</td>
</tr>
<tr>
<td>1974</td>
<td>Karzanov</td>
<td>Preflow-push</td>
<td>$V^3$</td>
</tr>
<tr>
<td>1983</td>
<td>Sleator-Tarjan</td>
<td>Dynamic trees</td>
<td>$E V \log V$</td>
</tr>
<tr>
<td>1986</td>
<td>Goldberg-Tarjan</td>
<td>FIFO preflow-push</td>
<td>$E V \log \left( V^2 / E \right)$</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1997</td>
<td>Goldberg-Rao</td>
<td>Length function</td>
<td>$E^{3/2} \log \left( V^2 / E \right) \log U \uparrow$ ( EV^{2/3} \log \left( V^2 / E \right) \log U \uparrow$</td>
</tr>
</tbody>
</table>
Stochastic Algorithms

- Markov random field approaches
- Potts model, Swendsen-Wang method
- Random walk and diffusion approaches
Markov random fields

Positive:

\[ P(f) > 0, \forall f \in F \]

Markovian: state only depends on neighbors

\[ P(f_i \mid f_{S - \{i\}}) = P(f_i \mid f_{N_i}) \]

Homogenious: probability independent of positions of sites
Markov-Gibbs Equivalence

GRF -- global property (the Gibbs distribution)

MRF -- local property (the Markovianity)

The Hammersley-Clifford theorem, the equivalence of these two:

*F is an MRF on S with respect to N if and only if F is a GRF on S with respect to N.*
Gibbs distribution:

\[ P(f) = \frac{e^{-E(f)/T}}{\sum_{f \in F} e^{-E(f)/T}} \]

where \( E \) is the energy function, \( T \) is the temperature.

(a) maximization of the posterior probability in the Bayesian framework

\( \leftrightarrow \) (b) minimization of the posterior energy function of a MRF

\( \leftrightarrow \) (c) minimization of the energy in a stochastic recurrent network
Ising/Potts Models

Ising model has a choice of two possible spin states at each lattice point
Potts models have $q>2$ possible states:

$S_1, S_2, S_3, S_4, \ldots S_q$
Segmentation with Potts Models
Swendsen-Wang method

SW method speeds up the time-consuming process by flipping the color of all vertices in one or all clusters simultaneously.
My Work

- Add external fields for segmentation
- Working at low temperature or deterministically
- Noisy video image segmentation

Probability is given by:

\[
P_G(x \mid \beta, V) = W(\beta, V)^{-1} \exp\left(\sum_{i \in S} x_i^t V + \frac{1}{2} \beta \sum_{j \in N(i)} x_i^t x_j\right),
\]

EM algorithm developed to estimate the model parameters
Random Walk Methods

Labels:
L1, L2, L3

Weights: in [0,1]

\[ w_{ij} = \exp \left( -\beta (g_i - g_j)^2 \right) \]
Probability of reaching L1

Probability of reaching L2
Probability of reaching L3

Segmentation results
My Work

- Make it fast, local and limited steps
- Reduce noise while keeping edges
- Apply to facial feature extraction

the random walkers eliminate the noise and keep the mutually connected feature pixels from vagueness

like morphology filters but it does not need to define a structural element in advance
Conclusions and Future Work

- Graphical models are powerful and ideal for image segmentation
- Choice of the deterministic and stochastic algorithms, trade-off
- To make them more robust and develop some applications
Thank You!