Video Image Segmentation with Graphical models

Jieyu Zhao

Outline

- Fundamentals
- Deterministic Methods
- Stochastic Methods
- Some Results
- Conclusions

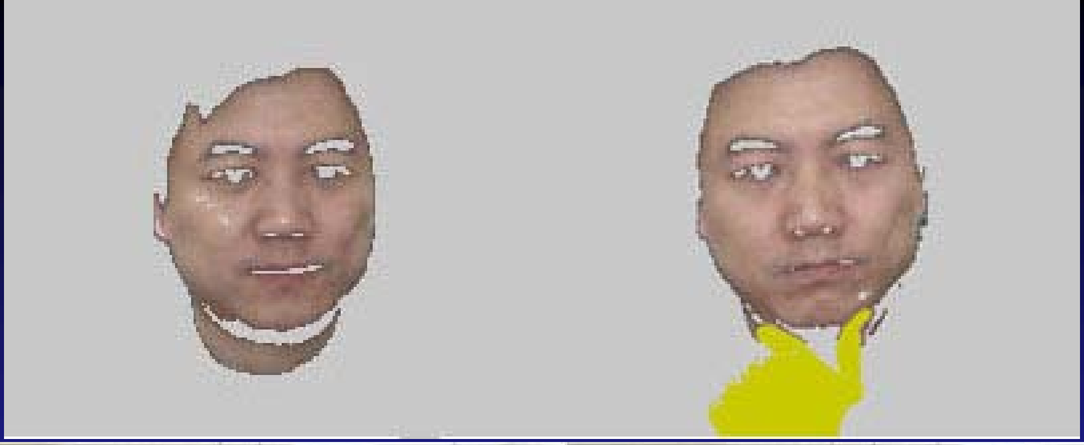
Video Image Segmentation

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


CNN
CNN



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

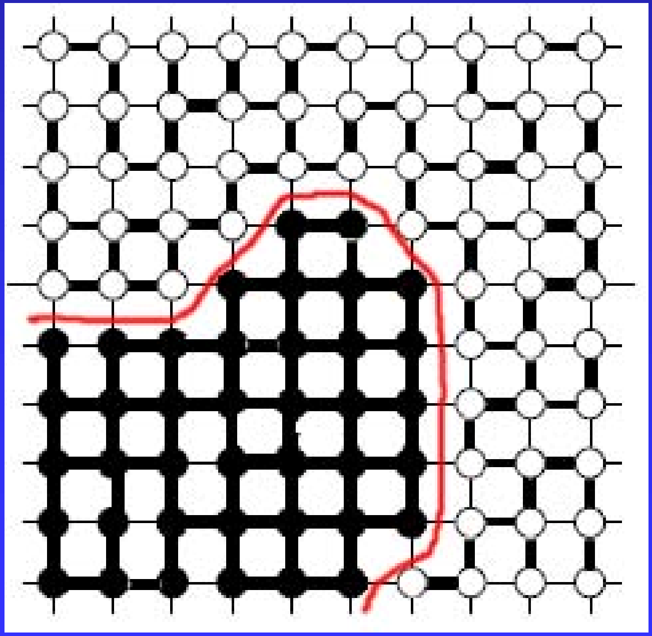
image represented with a weighted graph,

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

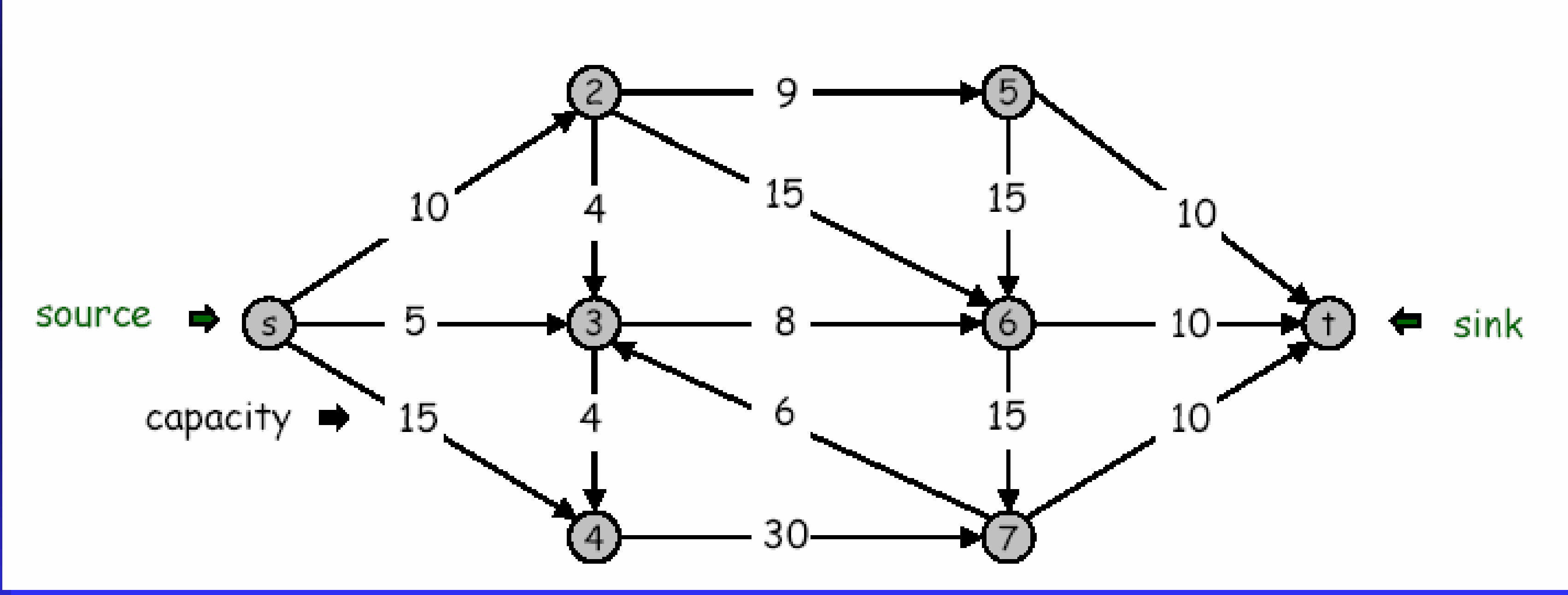


Maximum flow / Minimum cut

“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

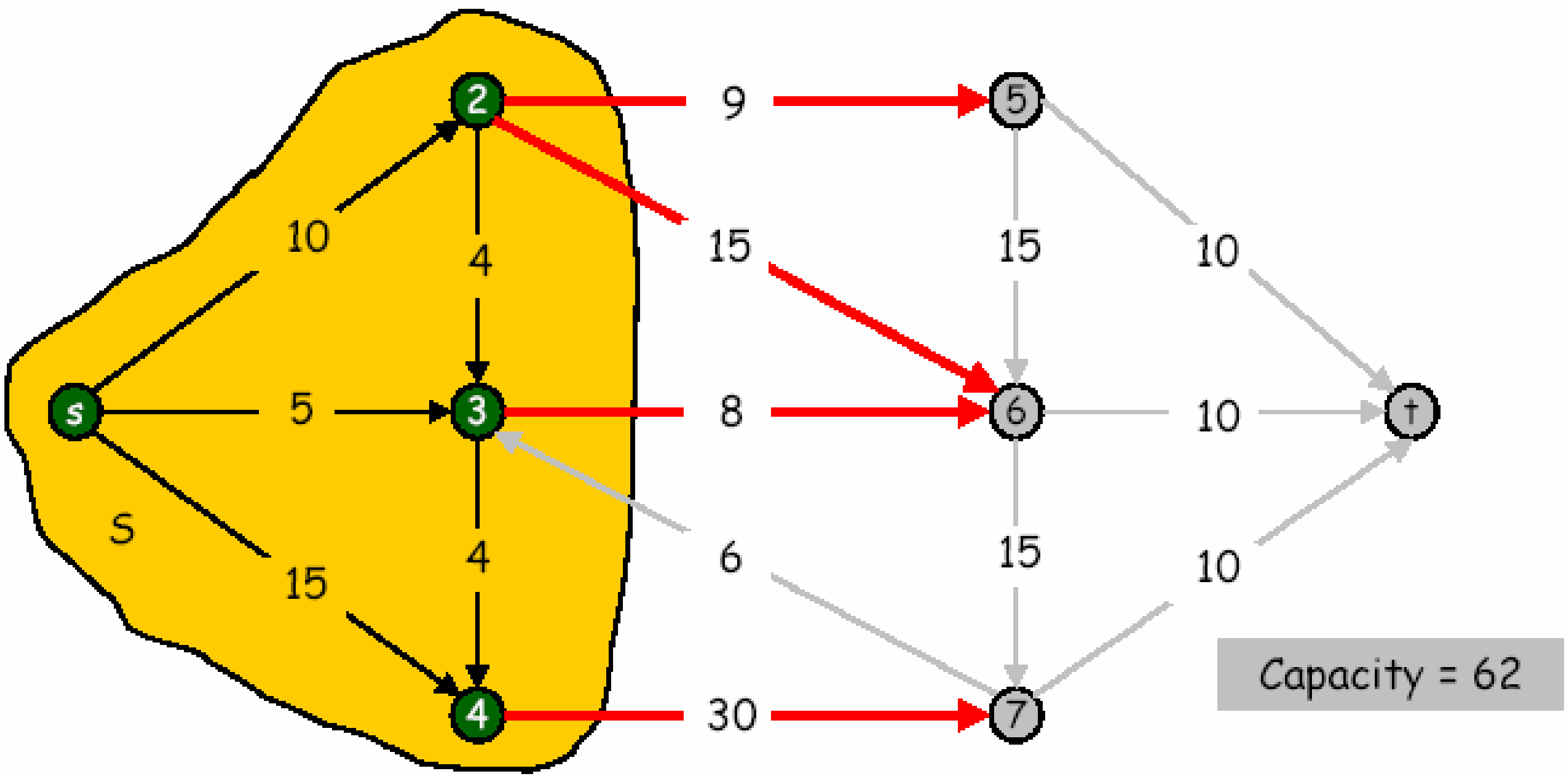
A weighted graph -- material flowing through the edges (railways, water pipelines)



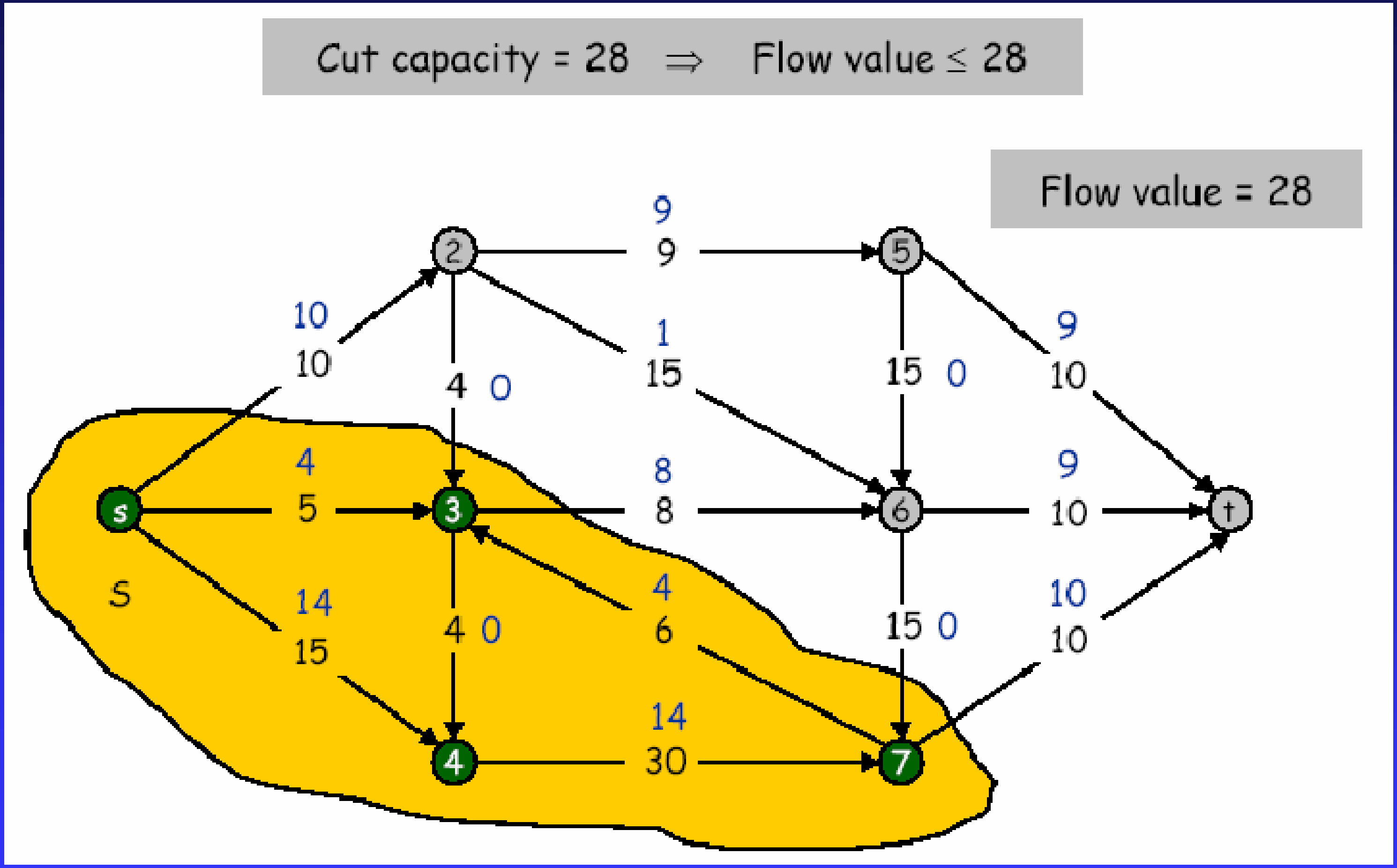
Maximum flow: maximize the sum $\sum_u f(u,t)$

A cut is a node partition (S, T) such that s is in S and t is in T.

capacity(S, T) = sum of weights of edges leaving S.



a min cut


Cut capacity = 28 ⇒ Flow value ≤ 28
Flow value = 28
s
2
3
4
5
6
7
t
S
9
9
10
10
4 0
1
15
15 0
9
10
4
5
8
8
9
10
14
15
4 0
4
6
15 0
10
10
14
30

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:

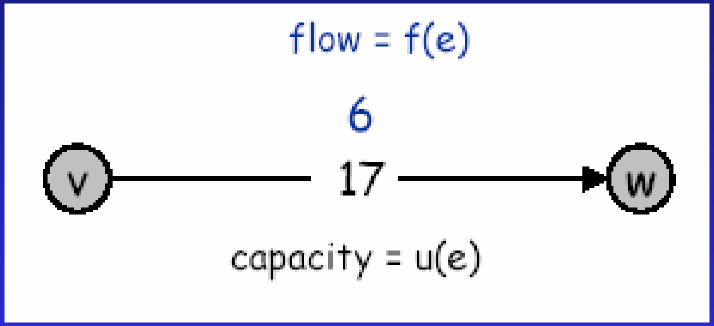
(i) f is a max flow.

(ii) There is no augmenting path relative to f.

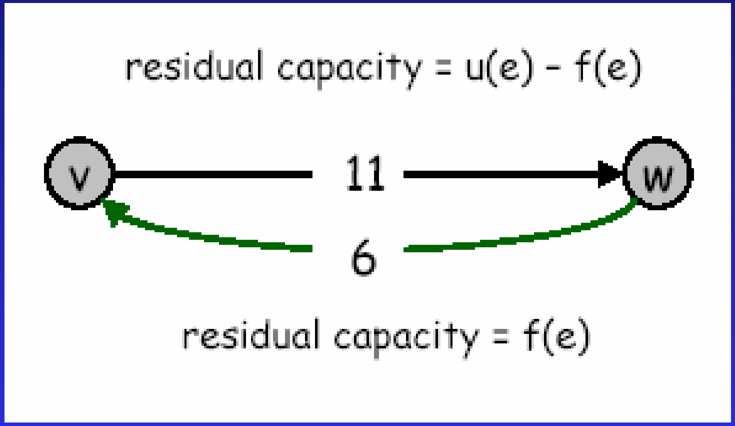
(iii) There exists a cut whose capacity equals the value of f.

Augmenting path = path in residual graph.

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



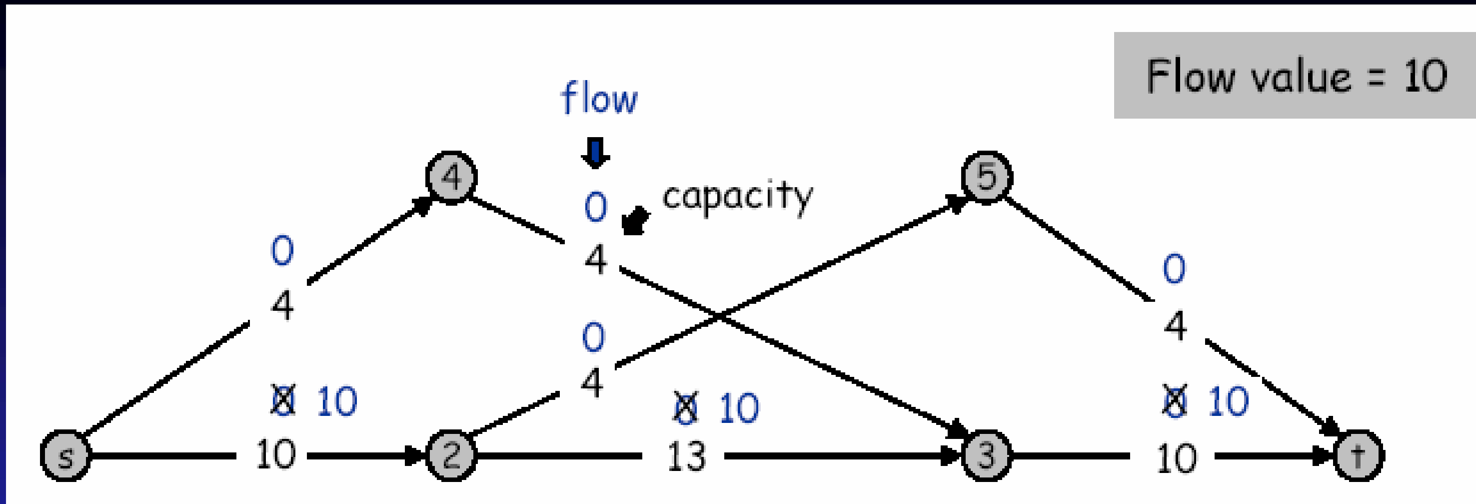
original graph

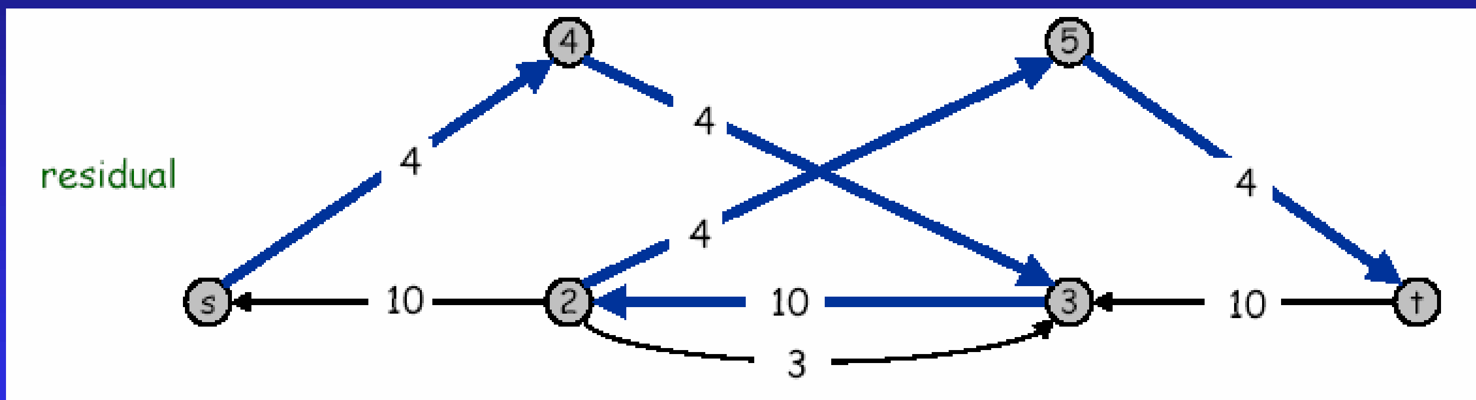


residual graph

Flow f(e).
Edge e = v→w

"Undo" flow sent.
w→v


Flow value = 10
flow
capacity
s
2
3
4
5
t
0
4
0 10
10
13


residual
s
2
3
4
5
t
4
10
3

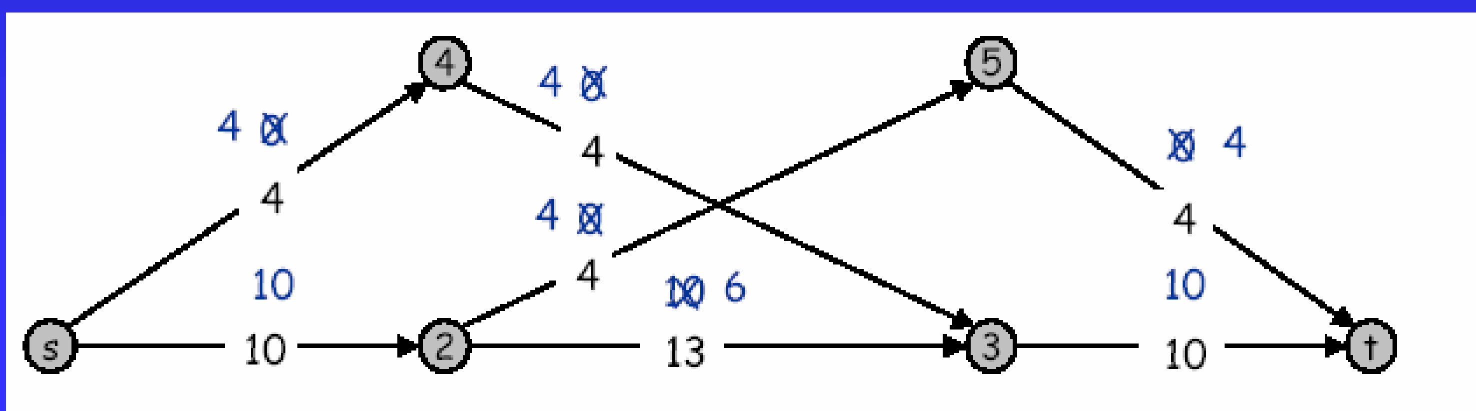

s
2
3
4
5
t
4 0
4
10 6
10
13
0 4

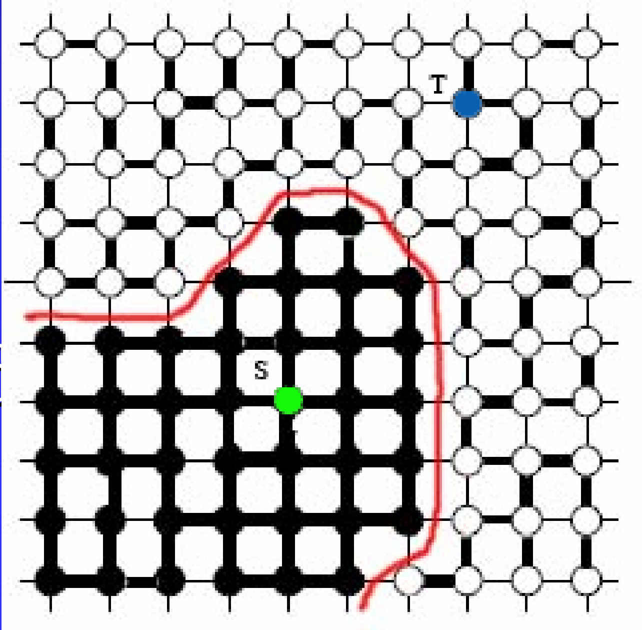
Image Segmentation Using Min Cut

- 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:

$$Ncut(A,B) = \frac{cut(A,B)}{volume(A)} + \frac{cut(A,B)}{volume(B)}$$



History of Worst-Case Running Times

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

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.

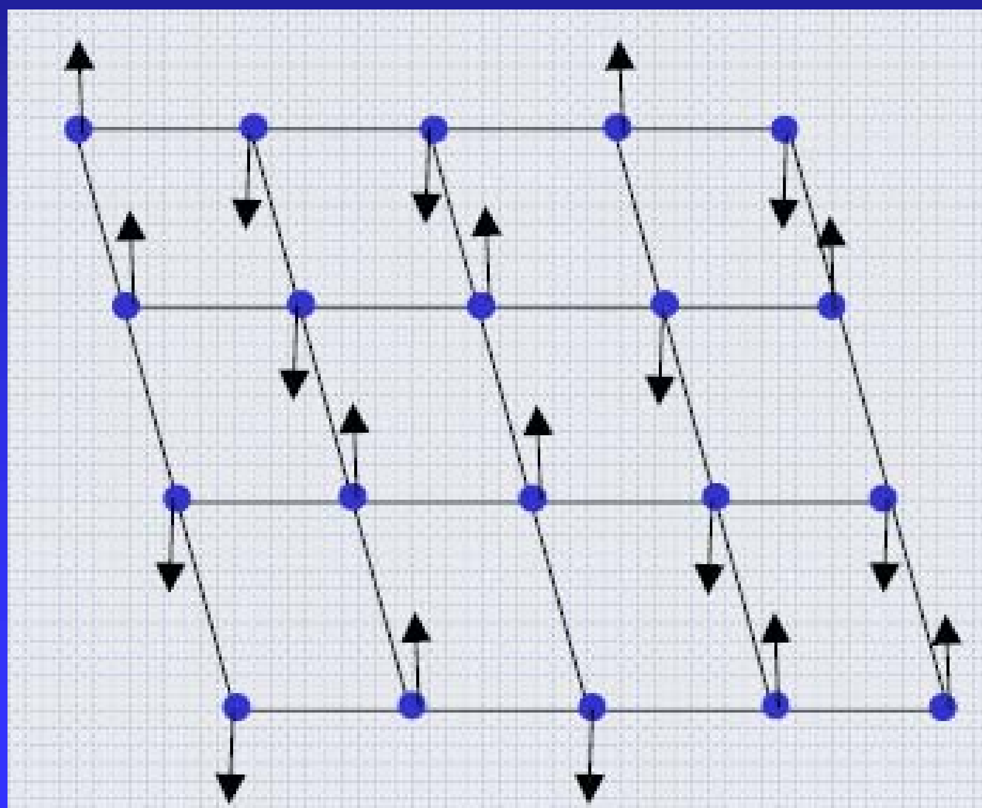
(a) maximization of the posterior probability in the Bayesian framework

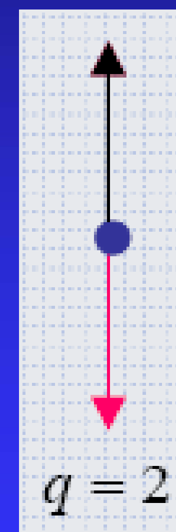
←→ **(b)** minimization of the posterior energy function of a MRF

←→ **(c)** minimization of the energy in a stochastic recurrent network

Ising/Potts Models

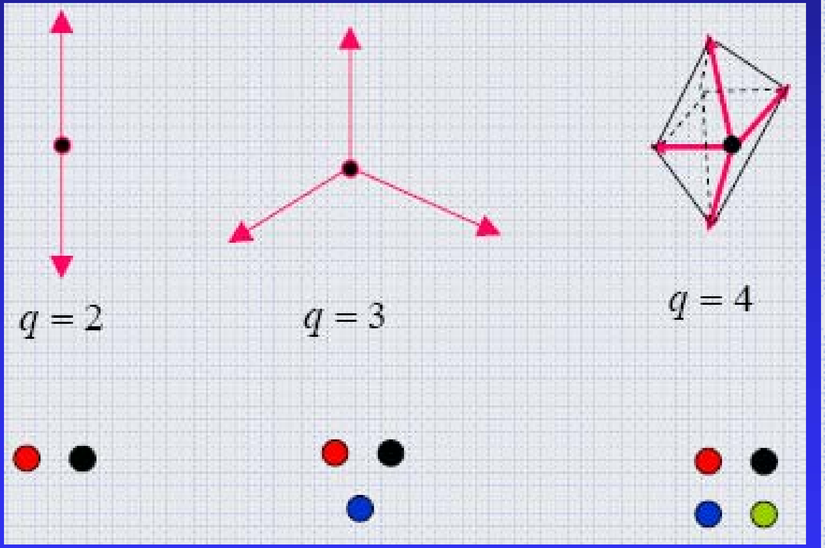
Ising model has a choice of two possible spin states at each lattice point

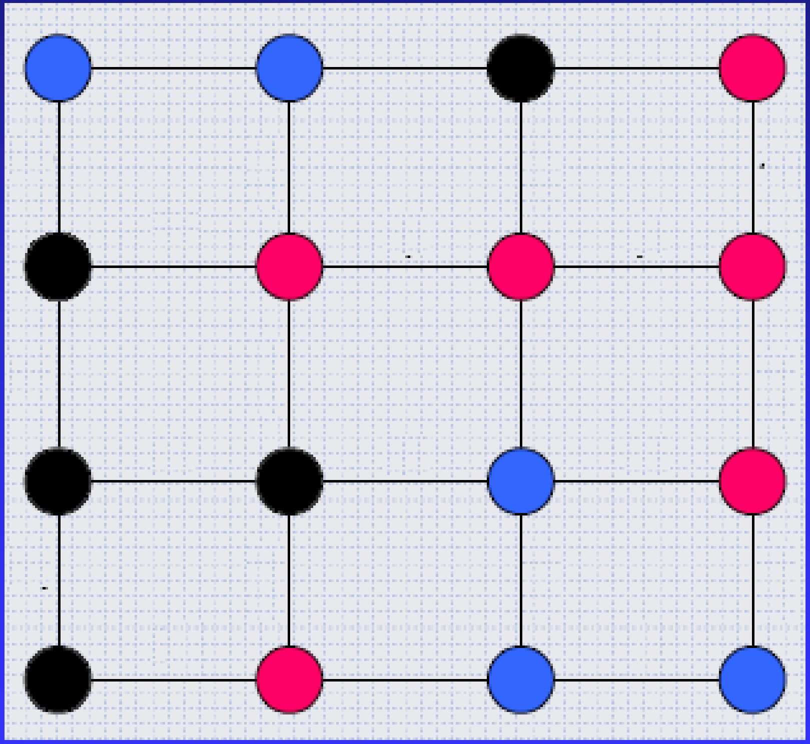




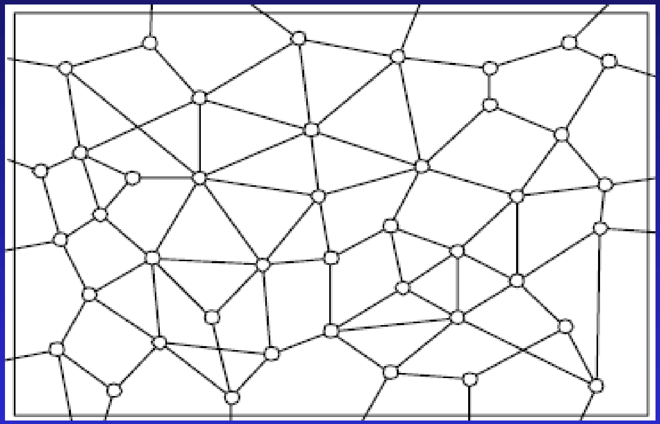
Potts models have q>2 possible states:

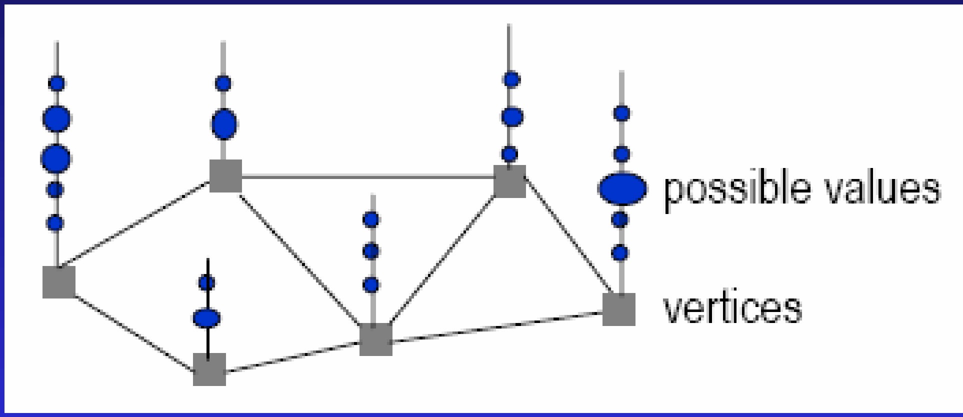
S1, S2, S3, S4, … Sq

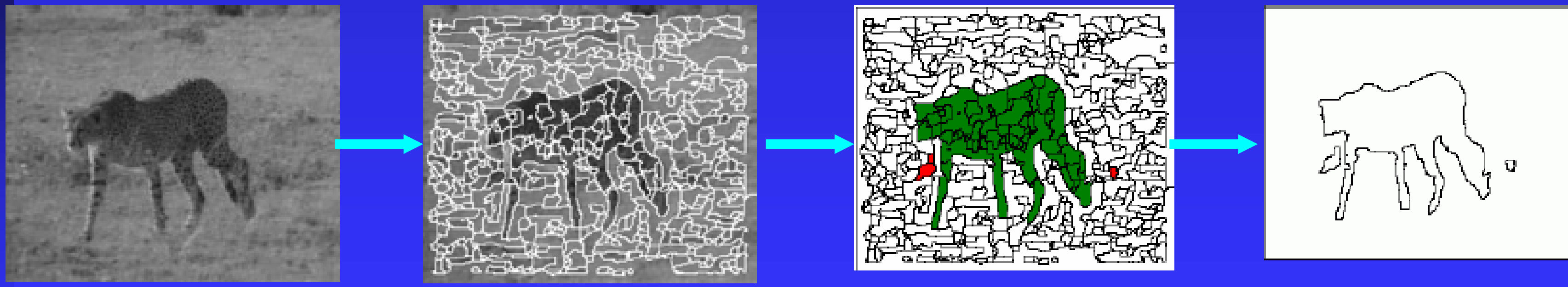




Segmentation with Potts Models

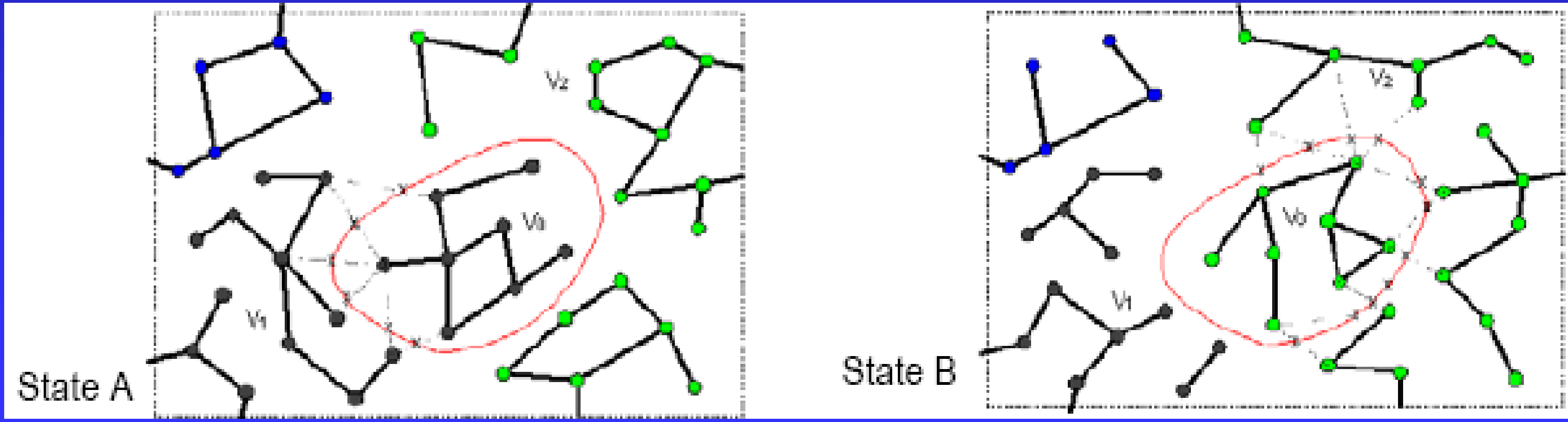






Swendsen-Wang method

- 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:

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

EM algorithm developed to estimate the model parameters

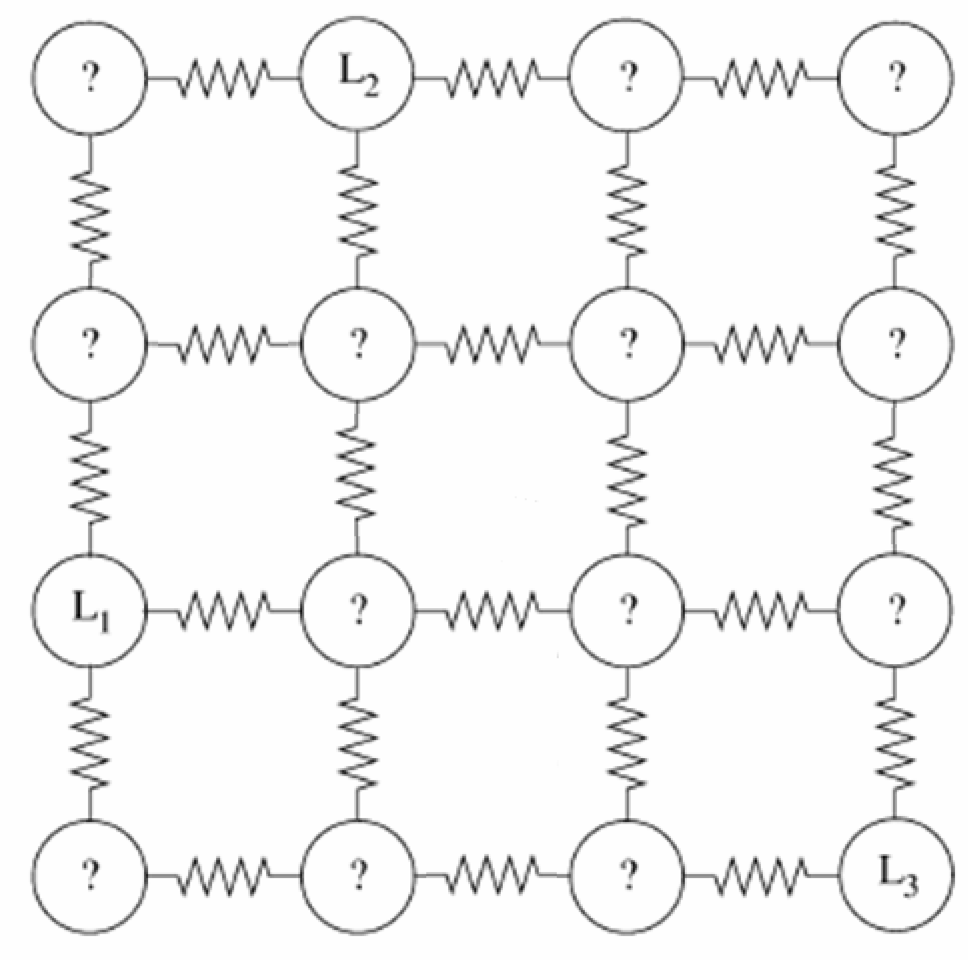
Random Walk Methods

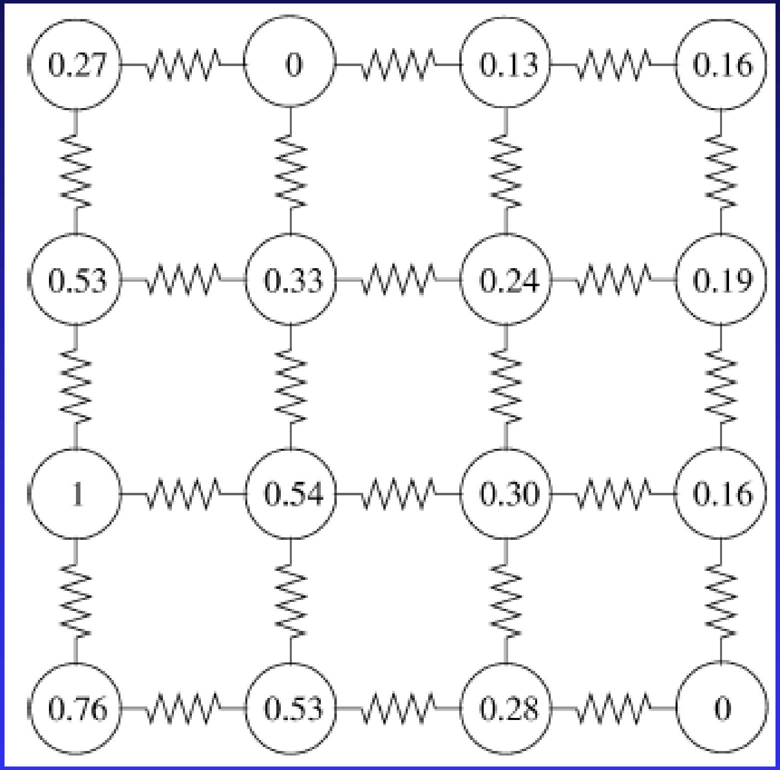
Labels:

L1, L2, L3

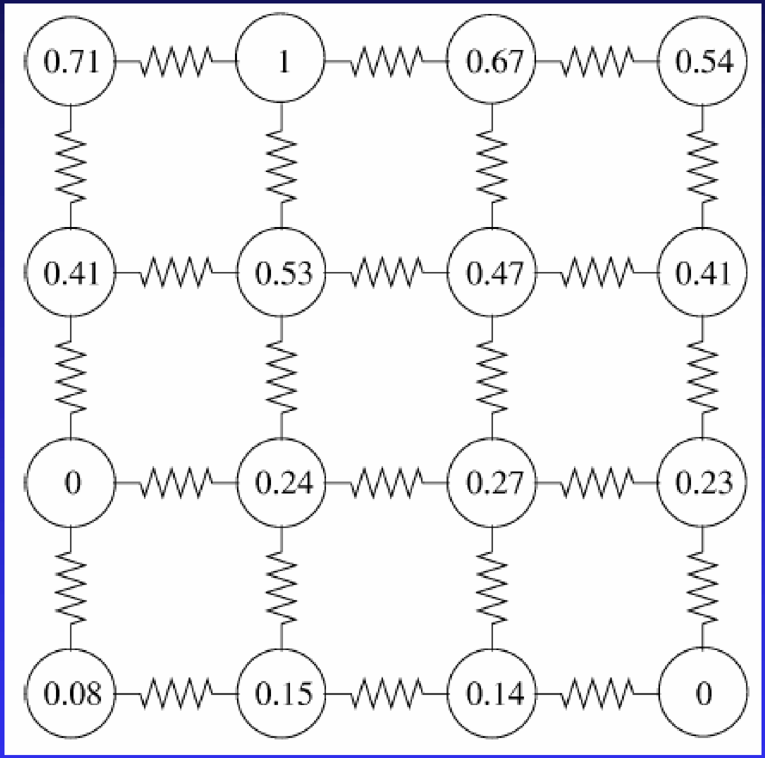
Weights: in [0,1]

$$w_{ij} = \exp\left(-\beta(g_i - g_j)^2\right)$$

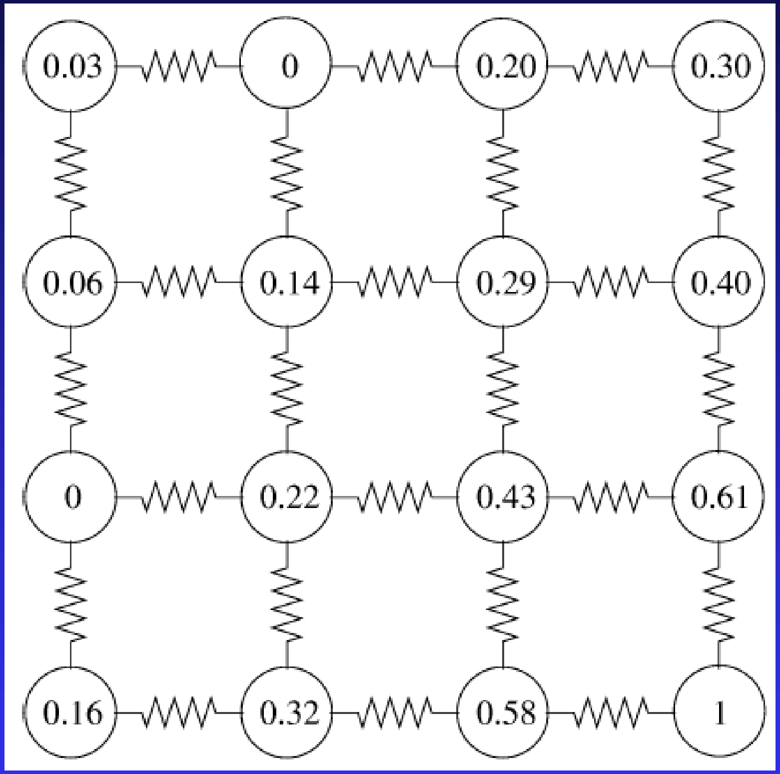




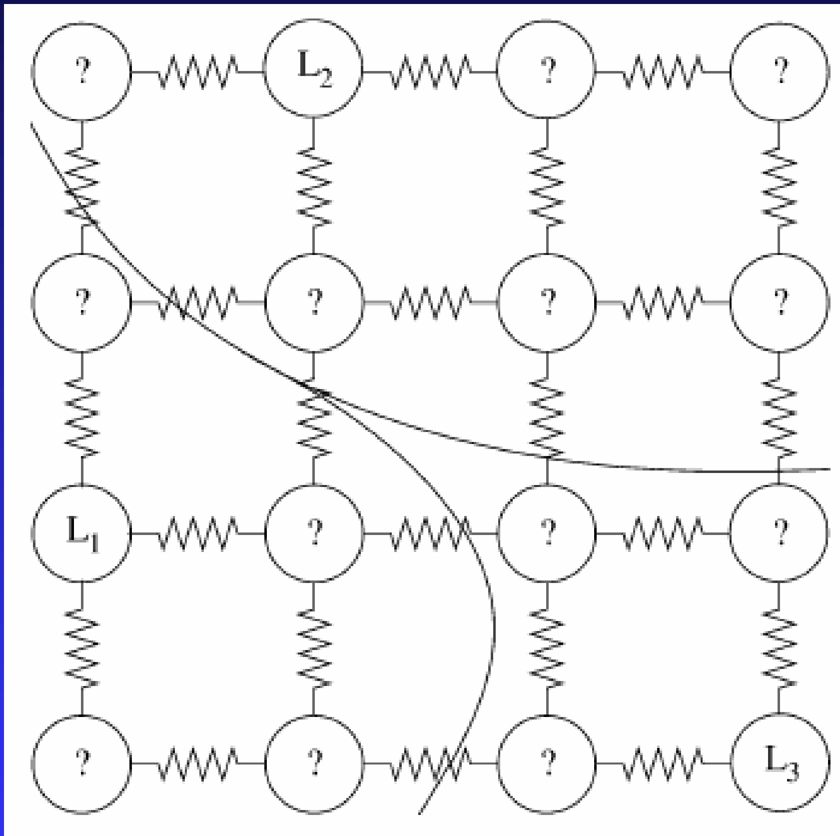
Probability of reaching L1



Probability of reaching L2



Probability of reaching L3



Segmentation results

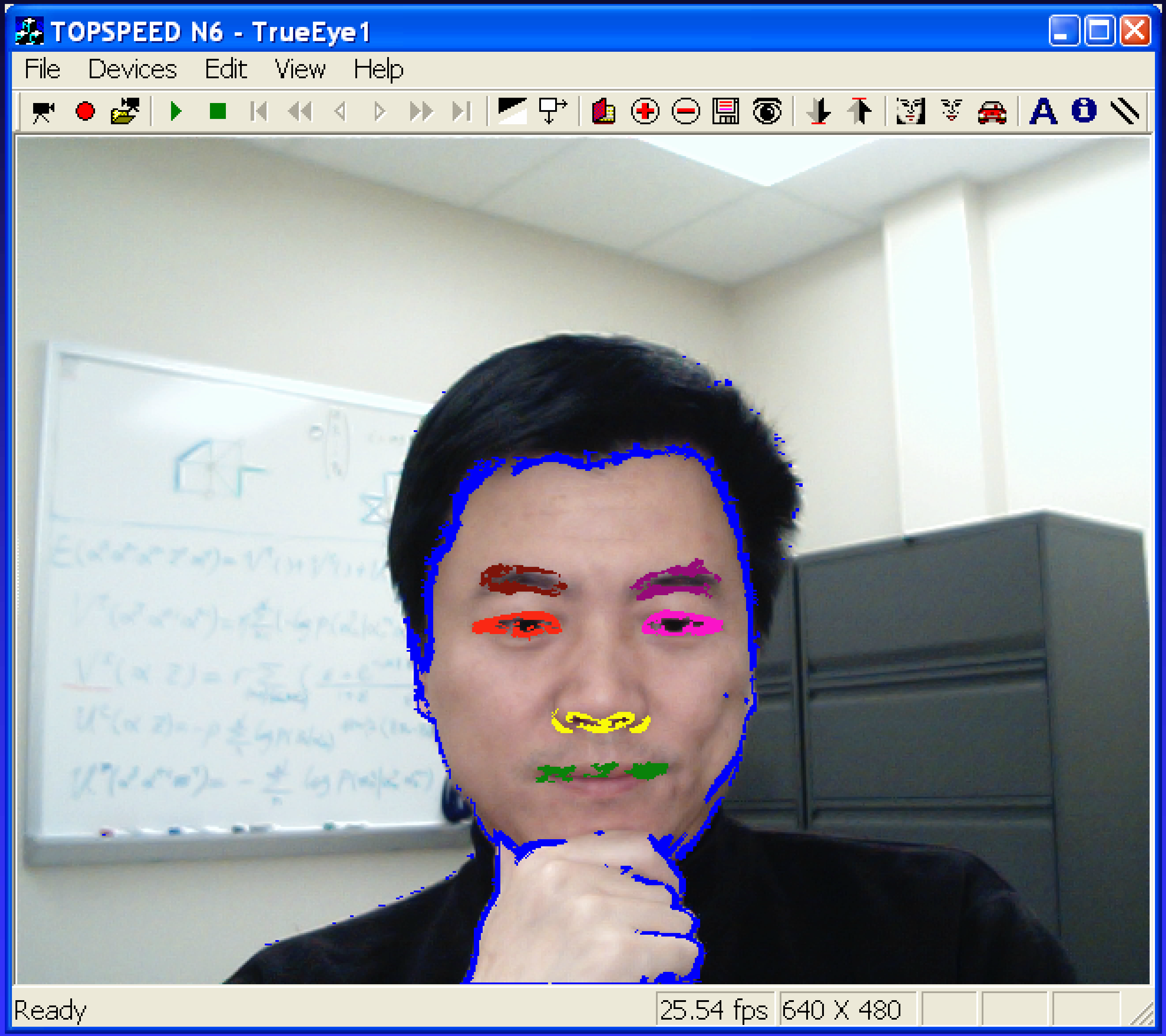
My Work

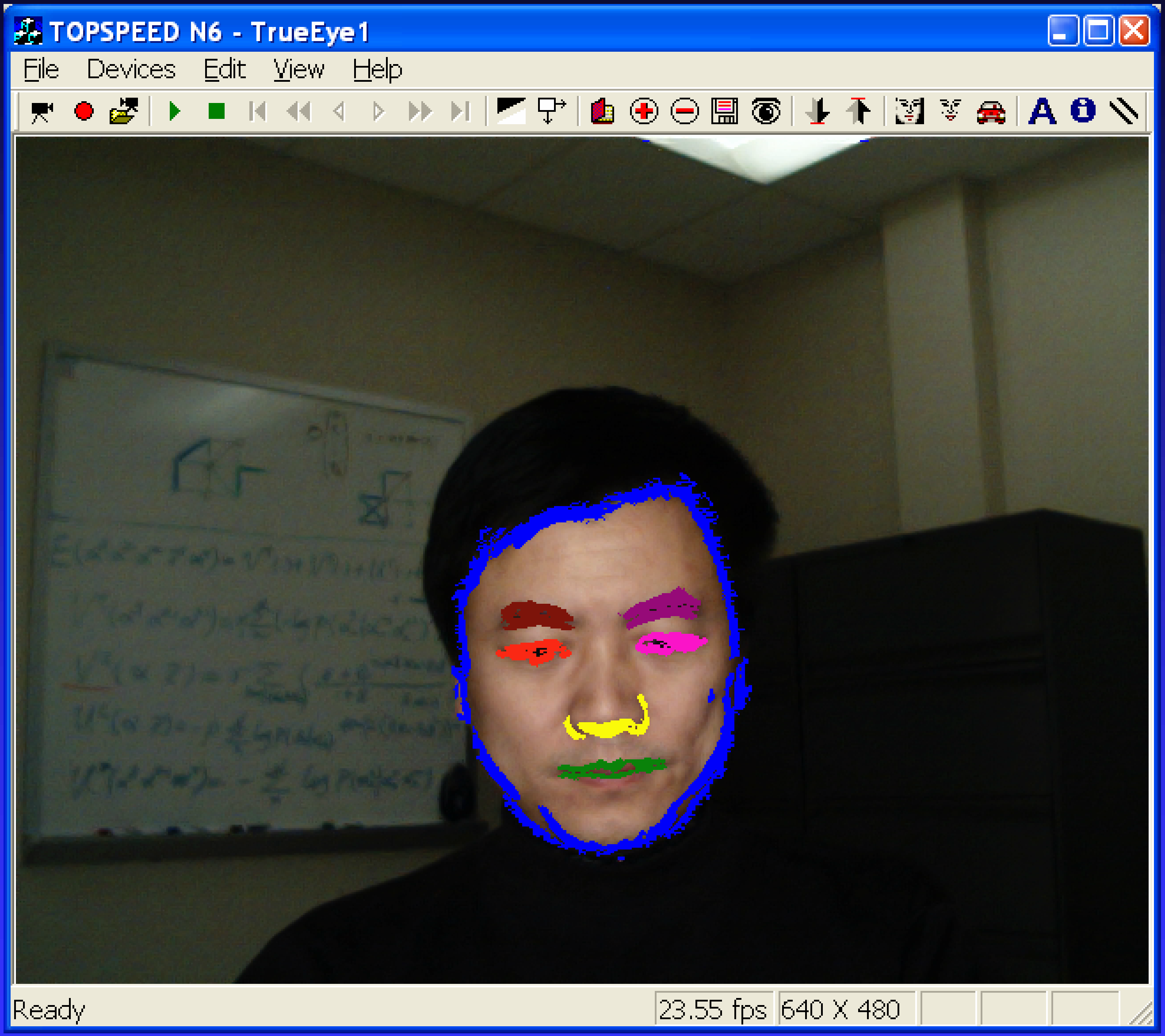
- Make it fast, local and limited steps
- Reduce noise while keeping edges
- Apply to facial feature extraction

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


TOPSPEED N6 - TrueEye1
File Devices Edit View Help
Ready
21.71 fps
640 X 480


TOPSPEED N6 - TrueEye1
File Devices Edit View Help
Ready
25.54 fps
640 X 480


TOPSPEED N6 - TrueEye1
File
Devices
Edit
View
Help
Ready
23.55 fps
640 X 480

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 !