Image Segmentation Using Minimum *st*-Cut

Tomáš Werner

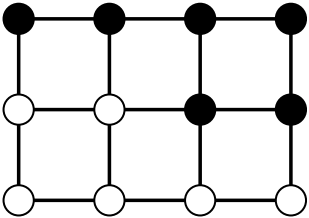
Center for Machine Perception
Czech Technical University
Prague, Czech Republic

Image Segmentation

Image segmentation: Label each pixel either as background or as foreground

Formalize this task as follows:

- Model the image as grid graph (V, E)
  - Pixels are nodes $v \in V$
  - Pairs of neighboring pixels are edges $vv' \in E$
- x_v = label of pixel v where $x_v \in \{0, 1\}$ (0 is background, 1 is foreground/object)



- f_v = intensity/color of pixel v; all intensities form vector $\mathbf{f} = (f_v \mid v \in V)$
- Segmentation: Compute the 'best' labeling $\mathbf{x} = (x_v \mid v \in V)$ from intensities $\mathbf{f}$

What is the 'Best' Labeling?

To be a good segmentation, labeling $\mathbf{x}$ must sastisfy two requirements:

1. Agreement with input data (independent for each pixel):
   - $p(0 \mid f_v)$ = probability that pixel with intensity f_v belongs to background
   - $p(1 \mid f_v)$ = probability that pixel with intensity f_v belongs to foreground
2. Contiguity of background and foreground (independent for each pixel pair):
   - $p(x_v, x_{v'}) = \begin{cases} a & \text{if } x_v = x_{v'} \\ b & \text{if } x_v \neq x_{v'} \end{cases}$ where $a > b$

What is the 'Best' Labeling?

To be a good segmentation, labeling **x** must sastisfy two requirements:

1. Agreement with input data (independent for each pixel):
   - $p(0 \mid f_v)$ = probability that pixel with intensity f_v belongs to background
   - $p(1 \mid f_v)$ = probability that pixel with intensity f_v belongs to foreground
2. Contiguity of background and foreground (independent for each pixel pair):
   - $p(x_v, x_{v'}) = \begin{cases} a & \text{if } x_v = x_{v'} \\ b & \text{if } x_v \neq x_{v'} \end{cases}$ where $a > b$

Maximum A-Posteriori (MAP) estimate

Find **x** maximising a-posteriori probability

$$p(\mathbf{x} \mid \mathbf{f}) = \frac{1}{Z(\mathbf{f})} \prod_{v \in V} p(x_v \mid f_v) \prod_{vv' \in E} p(x_v, x_{v'})$$

What is the 'Best' Labeling?

To be a good segmentation, labeling $\mathbf{x}$ must sastisfy two requirements:

1. Agreement with input data (independent for each pixel):
   - $p(0 \mid f_v)$ = probability that pixel with intensity f_v belongs to background
   - $p(1 \mid f_v)$ = probability that pixel with intensity f_v belongs to foreground
2. Contiguity of background and foreground (independent for each pixel pair):
   - $p(x_v, x_{v'}) = \begin{cases} a & \text{if } x_v = x_{v'} \\ b & \text{if } x_v \neq x_{v'} \end{cases}$ where $a > b$

Maximum A-Posteriori (MAP) estimate

Find $\mathbf{x}$ maximising a-posteriori probability

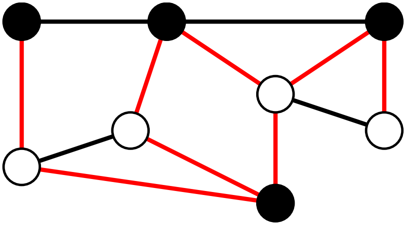
$$p(\mathbf{x} \mid \mathbf{f}) = \frac{1}{Z(\mathbf{f})} \prod_{v \in V} p(x_v \mid f_v) \prod_{vv' \in E} p(x_v, x_{v'})$$

Note: $$-\log p(\mathbf{x} \mid \mathbf{f}) = \underbrace{\sum_{v \in V} g(x_v \mid f_v) + \sum_{vv' \in E} g(x_v, x_{v'})}_{\text{image energy}} + \text{const}$$

where $g(x_v \mid f_v) = -\log p(\mathbf{x} \mid \mathbf{f})$ and $g(x_v, x_{v'}) = \begin{cases} 0 & \text{if } x_v = x_{v'} \\ c > 0 & \text{if } x_v \neq x_{v'} \end{cases}$

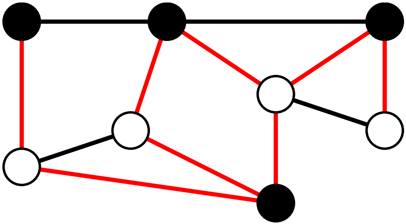
Minimum *st*-Cut

- Undirected graph (V, E) with nodes $v \in V$ and edges $vv' \in E \subseteq \binom{V}{2}$
- Every edge $vv' \in E$ has a non-negative weight $w_{vv'} \geq 0$
- Cut (S, T) is a partition of V into S and T such that $V = S \cup T$, $S \cap T = \emptyset$
- Weight of cut (S, T) is $\quad W(S, T) = \sum_{v \in S,\ v' \in T} w_{vv'}$



- Given special nodes s and t, an *st*-cut is a cut (S, T) such that $s \in S, t \in T$

- Undirected graph (V, E) with nodes $v \in V$ and edges $vv' \in E \subseteq \binom{V}{2}$
- Every edge $vv' \in E$ has a non-negative weight $w_{vv'} \geq 0$
- Cut (S, T) is a partition of V into S and T such that $V = S \cup T$, $S \cap T = \emptyset$
- Weight of cut (S, T) is $\quad W(S, T) = \sum_{v \in S,\ v' \in T} w_{vv'}$

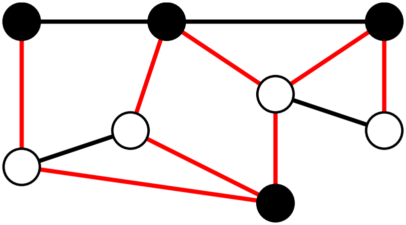


- Given special nodes s and t, an *st*-cut is a cut (S, T) such that $s \in S, t \in T$

Minimum *st*-cut problem

Find *st*-cut (S, T) that minimizes $W(S, T)$.

- Undirected graph (V, E) with nodes $v \in V$ and edges $vv' \in E \subseteq \binom{V}{2}$
- Every edge $vv' \in E$ has a non-negative weight $w_{vv'} \geq 0$
- Cut (S, T) is a partition of V into S and T such that $V = S \cup T$, $S \cap T = \emptyset$
- Weight of cut (S, T) is $\quad W(S, T) = \sum_{v \in S,\ v' \in T} w_{vv'}$



- Given special nodes s and t, an *st*-cut is a cut (S, T) such that $s \in S, t \in T$

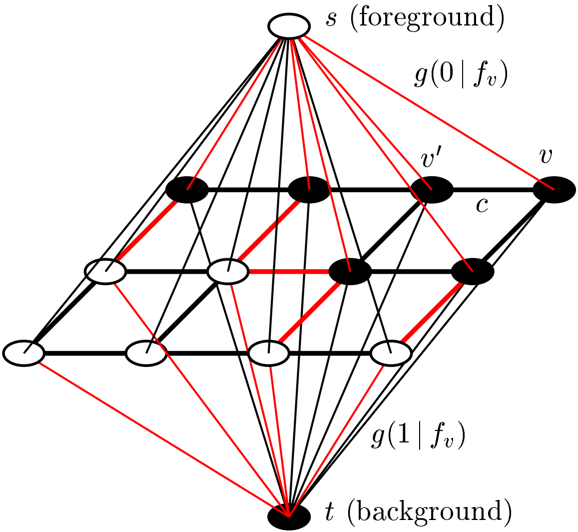
Minimum *st*-cut problem

Find *st*-cut (S, T) that minimizes $W(S, T)$.

There are fast algorithms for computing minimum *st*-cut in large sparse graphs! (They solve the related task, maximum flow.)

Minimizing Image Energy Using Minimum *st*-Cut

$$\min_{\mathbf{x}} \left[\sum_{v \in V} g(x_v \mid f_v) + \sum_{vv' \in E} g(x_v, x_{v'}) \right] \quad \text{where} \quad g(x_v, x_{v'}) = \begin{cases} 0 & \text{if } x_v = x_{v'} \\ c > 0 & \text{if } x_v \neq x_{v'} \end{cases}$$



Example

Let the data terms be Gaussian distributions with the same variance:

$$p(x_v \mid f_v) = \mathrm{const} \cdot \exp \frac{-[f_v - \mu(x_v)]^2}{2\sigma^2}$$

where $\mu(0)$, $\mu(1)$ are expected gray levels of background/foreground. Thus

$$g(x_v \mid f_v) = -\log p(x_v \mid f_v) = [f_v - \mu(x_v)]^2 + \mathrm{const}$$

Example

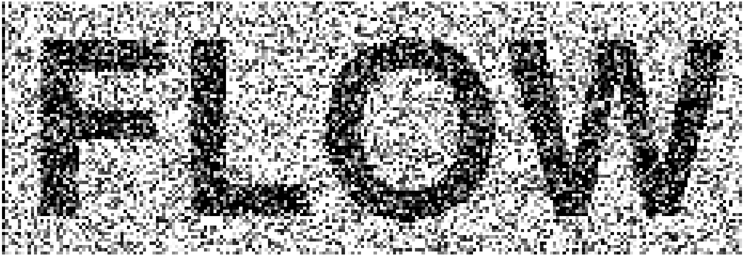
Let the data terms be Gaussian distributions with the same variance:

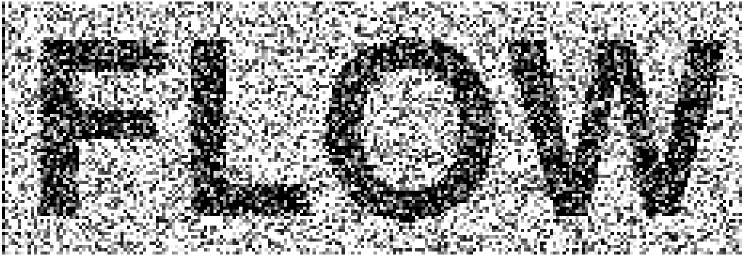
$$p(x_v \mid f_v) = \mathrm{const} \cdot \exp \frac{-[f_v - \mu(x_v)]^2}{2\sigma^2}$$

where $\mu(0)$, $\mu(1)$ are expected gray levels of background/foreground. Thus

$$g(x_v \mid f_v) = -\log p(x_v \mid f_v) = [f_v - \mu(x_v)]^2 + \mathrm{const}$$

For this image, we know that $\mu(0) = 1$ and $\mu(1) = 0$:

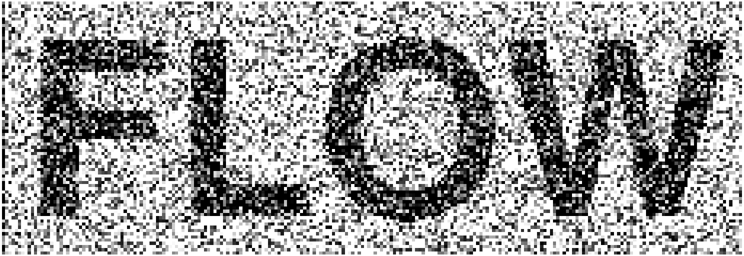




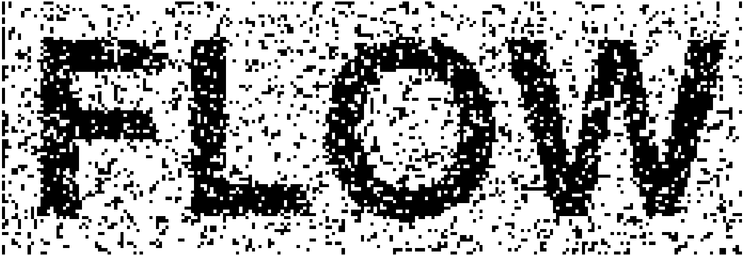
input image

segmentation, $c = 0$ (pure thresholding)

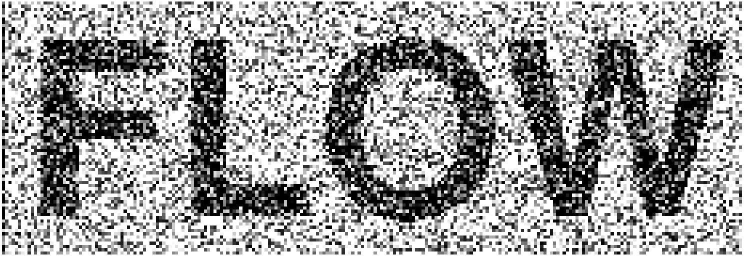
Example



input image



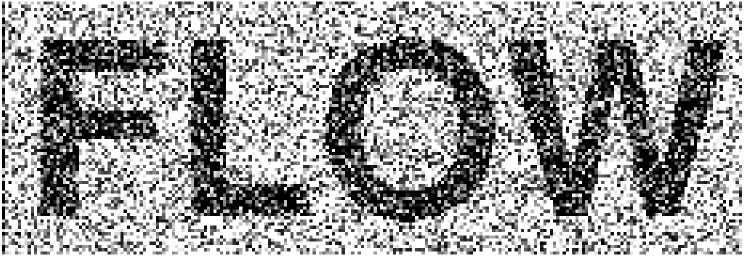
segmentation, $c = 20$



input image

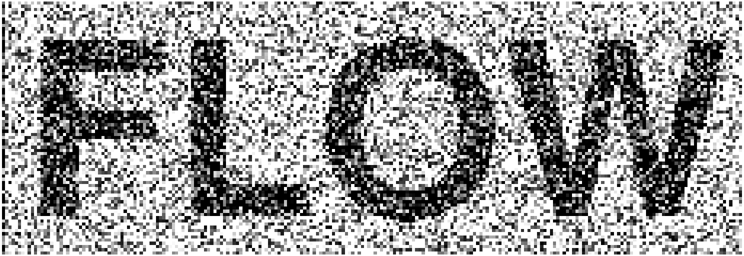
segmentation, $c = 30$

Example



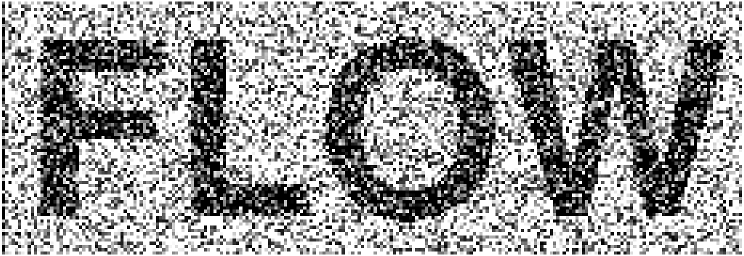
input image

segmentation, $c = 40$



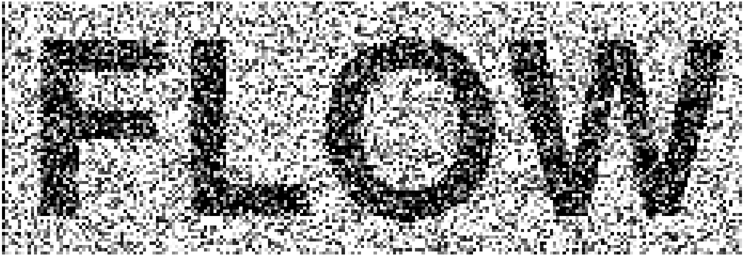
input image

segmentation, $c = 50$



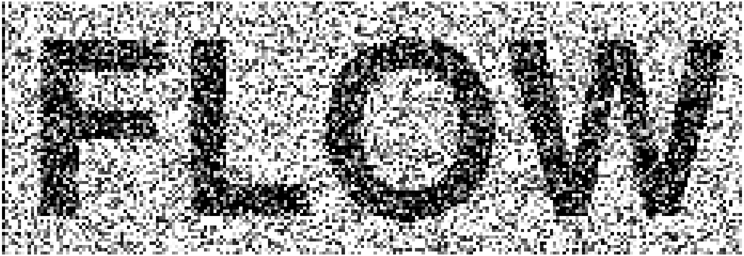
input image

segmentation, $c = 60$



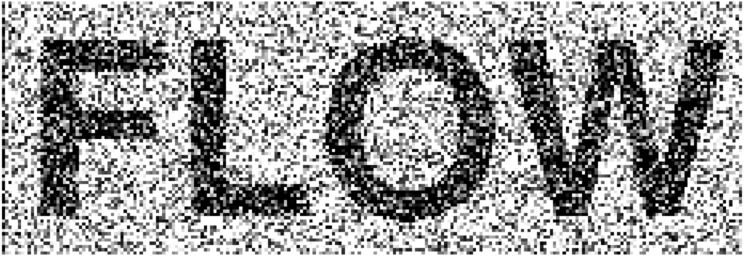
input image

segmentation, $c = 62$



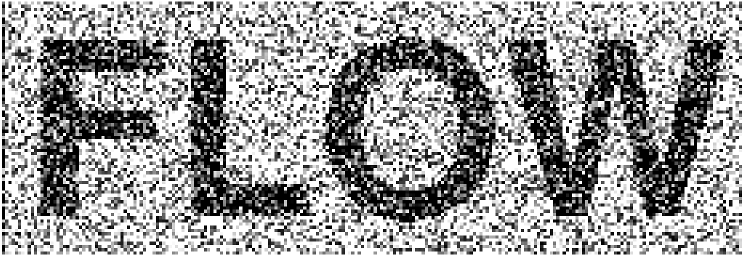
input image

segmentation, $c = 64$



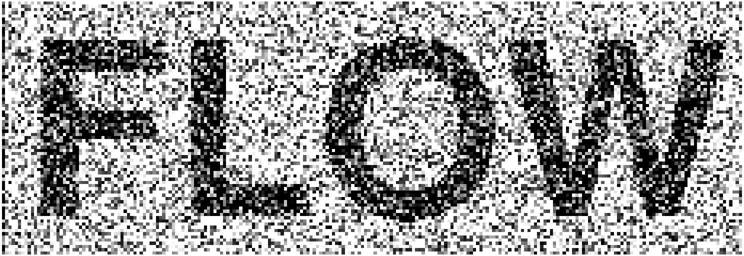
input image

segmentation, $c = 65$



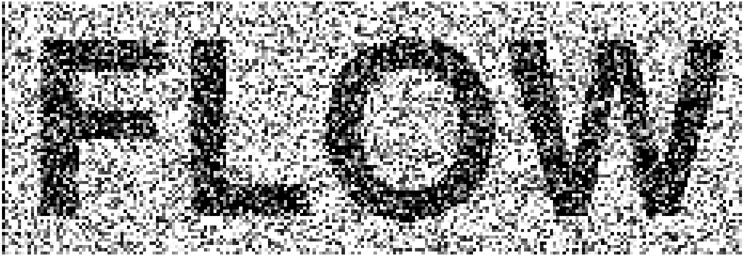
input image

segmentation, $c = 66$



input image

segmentation, $c = 67$



input image

segmentation, $c = 68$

Unknown Parameters of Back-/Foreground Model

- Often, statistical model of foreground/background is a family of distributions $p(x_v|f_v,\theta)$ parameterized by unknown θ.

  Note: In above example, $\theta = (\mu(0), \mu(1))$.
- We want to minimize a-posteriori probability simultaneously over $\mathbf{x}$ and θ:

Maximum A-Posteriori (MAP) estimate

Find $\mathbf{x}$ and θ maximising

$$p(\mathbf{x}\,|\,\mathbf{f},\theta) = \frac{1}{Z(\mathbf{f},\theta)} \prod_{v\in V} p(x_v\,|\,f_v,\theta) \prod_{vv'\in E} p(x_v,x_{v'})$$

A hard problem. A suboptimal solution found by alternating maximisation:

- Fix $\mathbf{x}$ and minimise over θ.
- Fix θ and minimise over $\mathbf{x}$.

Repeat until convergence.