Image Segmentation: Mean Shift & Normalized Cut

M.-A. Augé & M. Karpe

École des Ponts ParisTech

30^{th} May 2018

Content

Goals of Image Segmentation (Lazebnik)

- Group together similar-looking pixels for efficiency of further processing
- Separate image into coherent "objects"

Content

Mean Shift

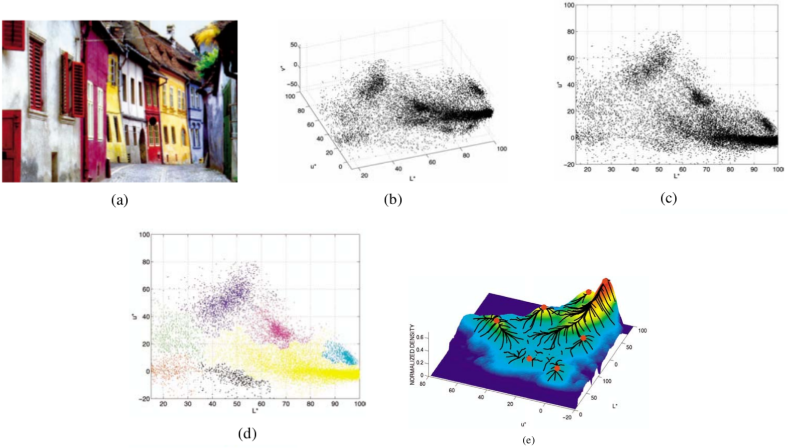

Figure – Illustration of mean shift image segmentation (Szeliski).

Illustration

Code available online : Link

Figure – Mean shift algorithm applied to drugs of different colors.

Density estimation

Let $(x_i)_{i \in \mathscr{I}}$ input samples, k kernel function, h : kernel width.

$$f(x) = \sum_{i \in \mathscr{I}} K(\mathbf{x} - \mathbf{x_i}) = \sum_{i \in \mathscr{I}} k\left(\frac{\|\mathbf{x} - \mathbf{x_i}\|^2}{h^2}\right)$$

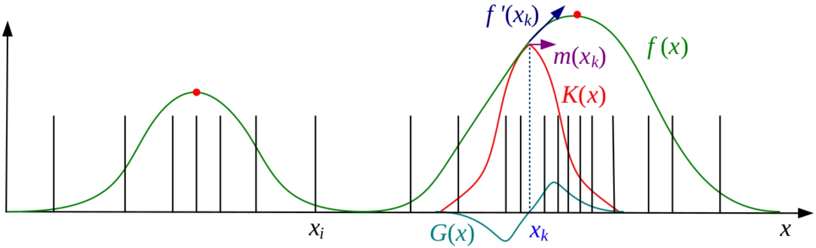

Figure – One-dimensional visualization of the kernel density estimate, its derivative, and a mean shift (Szeliski).

Computation of mean shift

- Gradient of $f(\mathbf{x})$:

$$\nabla f(\mathbf{x}) = \sum_{i \in \mathscr{I}} (\mathbf{x} - \mathbf{x_i}) G(\mathbf{x} - \mathbf{x_i}) = \sum_{i \in \mathscr{I}} (\mathbf{x} - \mathbf{x_i}) g\left(\frac{\|\mathbf{x} - \mathbf{x_i}\|^2}{h^2}\right)$$

$$g(r) = -k'(r)$$

- Re-write with *mean shift* vector :

$$\nabla f(\mathbf{x}) = \left(\sum_{i \in \mathscr{I}} G(\mathbf{x} - \mathbf{x_i})\right) \mathbf{m}(\mathbf{x})$$

$$\mathbf{m}(\mathbf{x}) = \frac{\sum_{i \in \mathscr{I}} \mathbf{x_i} G(\mathbf{x} - \mathbf{x_i})}{\sum_{i \in \mathscr{I}} G(\mathbf{x} - \mathbf{x_i})} - \mathbf{x}$$

Mean shift procedure

$$\mathbf{y}_{k+1} = \mathbf{y}_k + \mathbf{m}(\mathbf{y}_k) = \frac{\sum_{i \in \mathscr{I}} \mathbf{x_i} G(\mathbf{y_k} - \mathbf{x_i})}{\sum_{i \in \mathscr{I}} G(\mathbf{y_k} - \mathbf{x_i})}$$

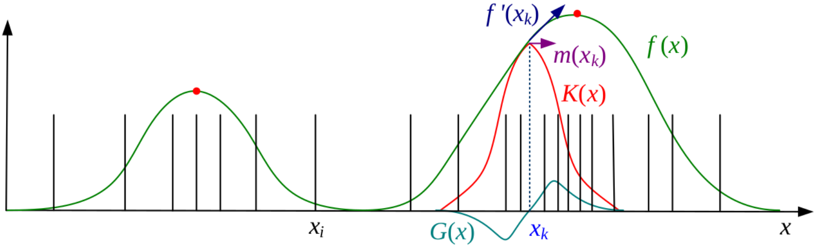

Figure – One-dimensional visualization of the kernel density estimate, its derivative, and a mean shift (Szeliski).

Kernels

- Epanechnikov kernel : $k_E(r) = \max(0, 1 - r)$
- Gaussian (normal) kernel : $k_N(r) = \exp\left(-\frac{r}{2}\right)$

$$K(\mathbf{x_i}) = k\left(\frac{\|\mathbf{x_r}\|^2}{h_r^2}\right) k\left(\frac{\|\mathbf{x_s}\|^2}{h_s^2}\right)$$

- $\mathbf{x_s} = (x, y)$: spatial coordinates
- $\mathbf{x_r}$: color value
- h_s (resp. h_r) : spatial (resp. range) bandwith.

Illustration

Figure – Original picture (left) and mean shifted picture (right) of Bienvenüe (top) and La Source (bottom) at ENPC.

Illustration


Figure – Original picture (left) and mean shifted picture (right) of Imagine researcher at Ecole des Ponts ParisTech (Aubry).

Pros and cons

Pros :

- Few assumptions
- Few parameters
- Robust to outliers

Cons :

- Choose of window size
- Computationally expensive
- Not adapted for high-dimensional features

Content

Graph to Image (Lazebnik)

- Node for every pixel
- Edge between every pair of pixels (or every pair of "sufficiently close" pixels)
- Each edge is weighted by the similarity of the two nodes

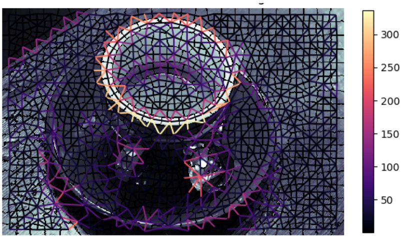

Figure – Region Adjacency Graph (RAG) obtained with `skimage`

Region Adjacency Graph (RAG)

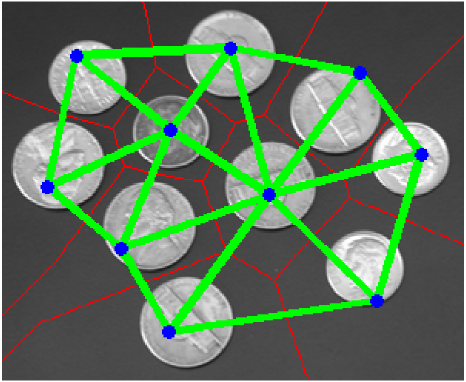

Figure – RAG on coins
https://fr.mathworks.com/matlabcentral/fileexchange/16938-region-adjacency-graph--rag-

Min Cut

$$\min_{A} cut(A, \bar{A})$$

where

$$cut(A, B) = \sum_{w \in A, v \in B} w(u, v)$$

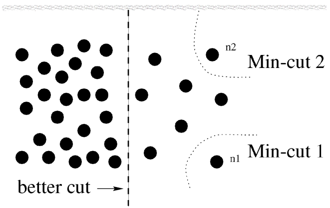

Figure – Min Cut where results are bad
(*Normalized Cuts and Image Segmentation*, J. Shi, J. Malik)

Normalized Cut

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

where

$$assoc(A,V) = \sum_{w \in A, t \in V} w(u,t)$$

Normalized Cut to Spectral Clustering

$$(D - W)x = \lambda Dx$$

where

- $N = |V|$
- $W \in \mathbb{R}^{N \times N}$: weight matrix
- $D \in \mathbb{R}^{N \times N}$: diagonal matrix with $D(i,i) = \sum_j w(i,j)$

Signs of the second eigenvector x decide on the cut :

$$i \in A \text{ iff } x(i) > 0$$

Our code

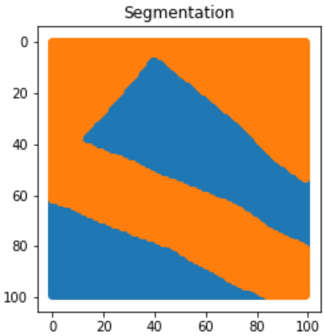

Figure – First results for a black and white picture

Scikit-image


Figure – Original picture (left) and normalized cut picture with scikit-image (right) of La Source at ENPC.

Pros and Cons (Lazebnik)

Pros :

- Generic framework, can be used with many different features and graphs formulations

Cons :

- High storage requirement and time complexity
- Bias towards partitioning into equal segments

Content

Conclusion (Szeliski)

- Mean-shift technique tries to find clusters of similar pixels using mode finding
- Normalized cuts technique examines the affinities between nearby pixels and tries to separate groups that are connected by weak affinities

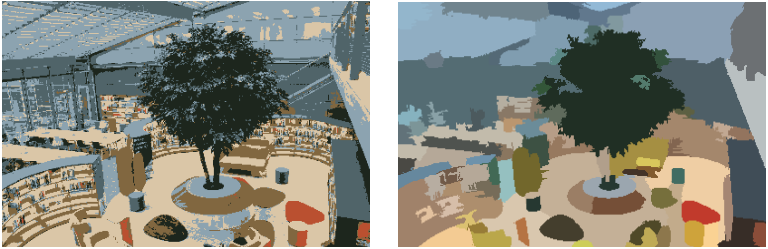

Figure – Mean shifted picture (left) and normalized cut picture (right) of La Source (ENPC).

References I

Richard Szeliski
Computer Vision : Algorithms and Applications
Springer, 3rd September 2010
http://szeliski.org/Book/

Jianbo Shi, Jitendra Malik
Normalized Cuts and Image Segmentation
University of Pennsylvania, 1st August 2000
https://repository.upenn.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1101&context=cis_papers

References II

Svetlana Lazebnik
Image Segmentation
University of North Carolina, Spring 2009
http://cs.unc.edu/~lazebnik/spring09/lec24_segmentation.pdf