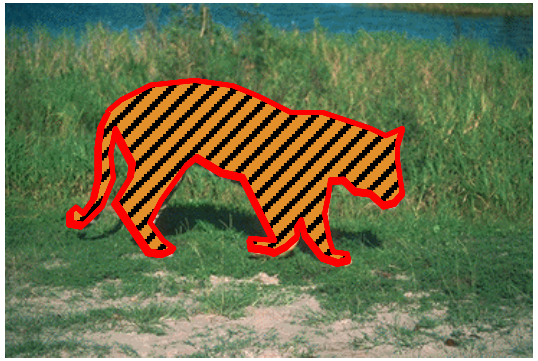
图像分割
Image Segmentation

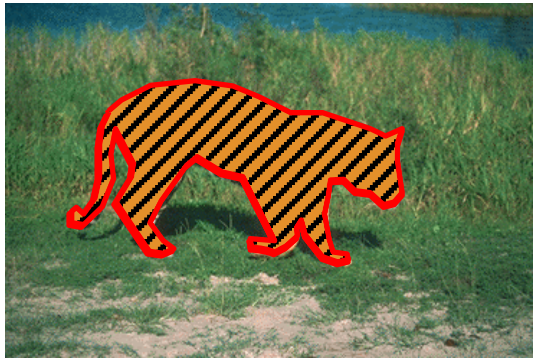
Ning Tang

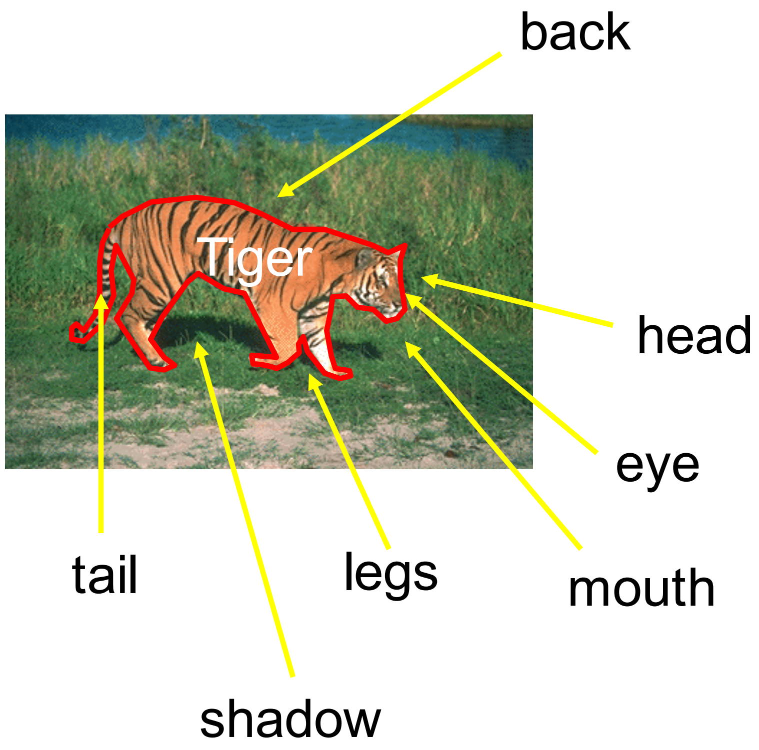
From Pixels to Perception

From Pixels to Perception

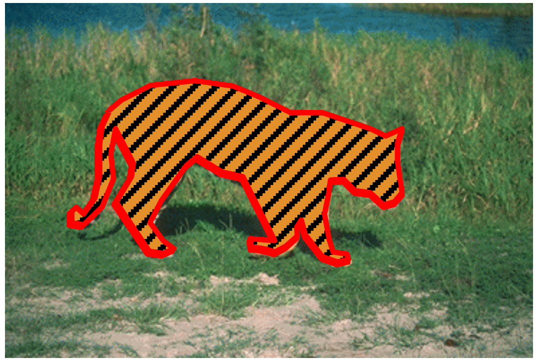


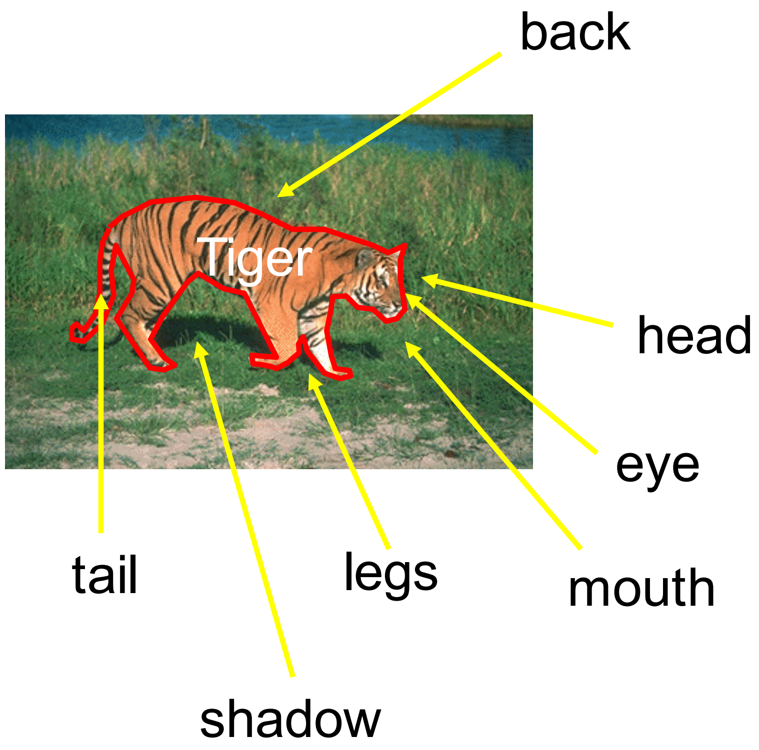
From Pixels to Perception

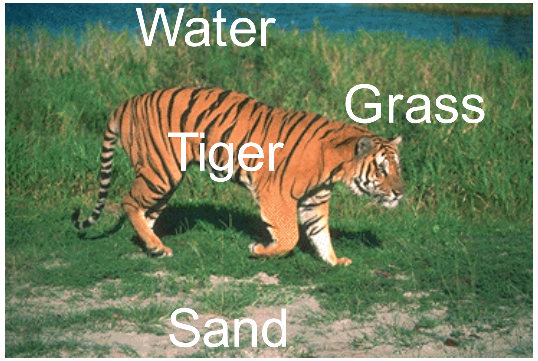




From Pixels to Perception







Visual Grouping

I stand at the window and see a house, trees, sky. Theoretically I might say there were 327 brightnesses and nuances of colour.

Do I have "327"?

No. I have sky, house, and trees.

---- Max Wertheimer, 1923

Human visual system

Computer Vision

Visual Grouping → **Image Segmentation**

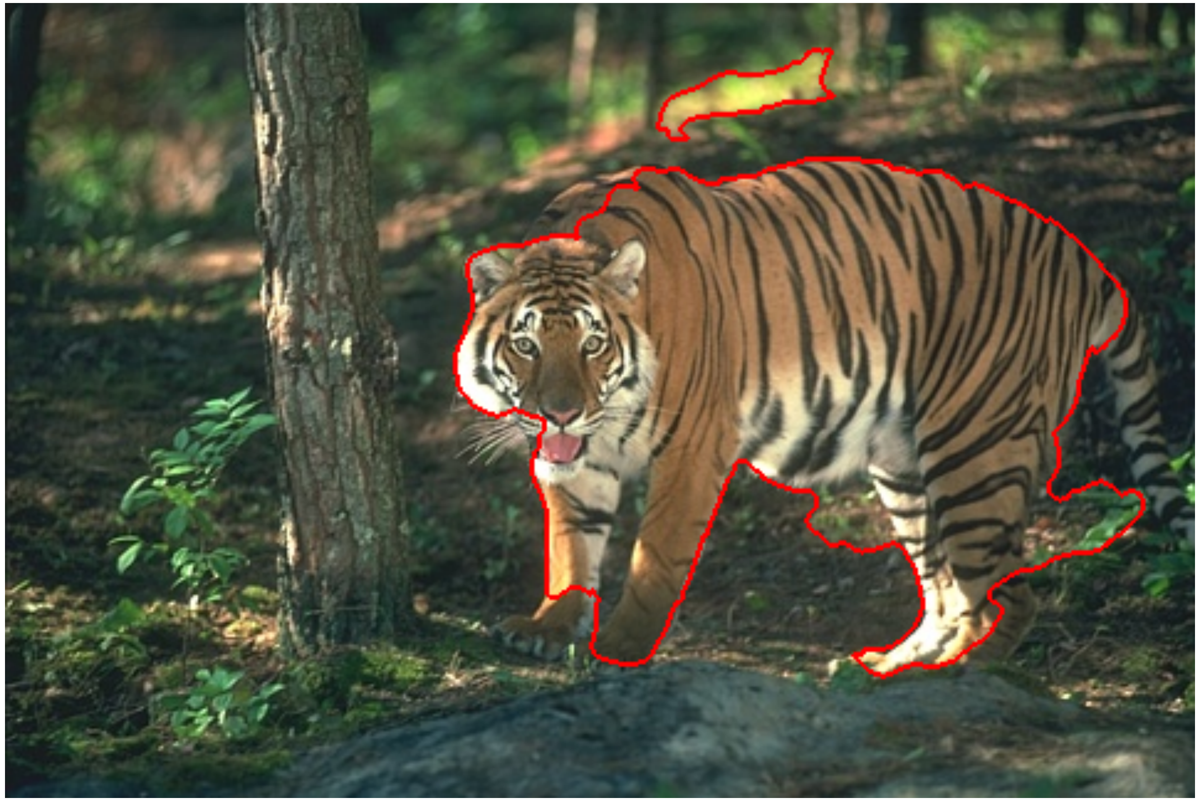
Image Segmentation

Image segmentation is the process of partitioning a image into multiple regions.

More precisely, image segmentation is the process of assigning a label to every pixel in an image so that pixels with the same label share certain characteristics (such as color, intensity, or texture).

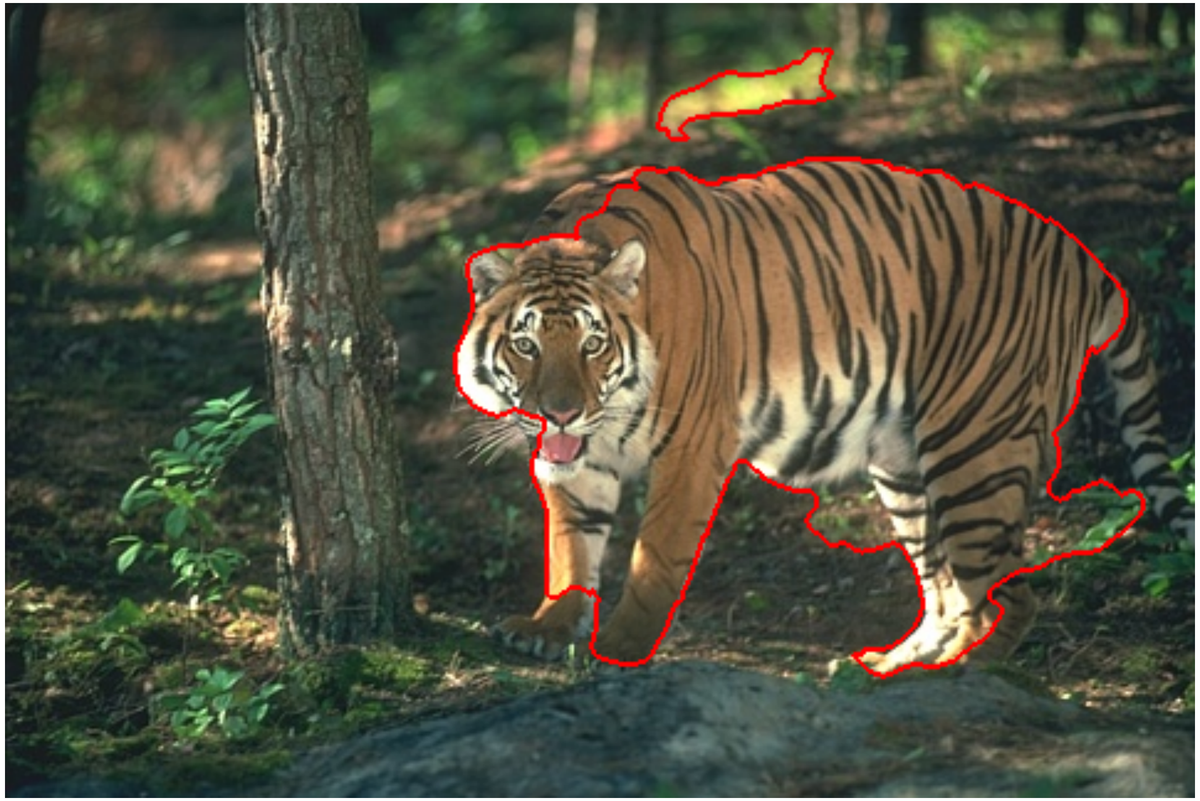
The Goal of Image Segmentation

Simplify the representation of an image into something that is more meaningful and easier to analyze.



The Goal of Image Segmentation

Simplify the representation of an image into something that is more meaningful and easier to analyze.



A pre-processing in Computer Vision

Image Segmentation Dataset

For image segmentation, we need :

original image

+

ground-truth

Compare real segmentation result with ground-truth

real segmentation result

≠

ground-truth

Manual annotation for getting ground-truth

original image

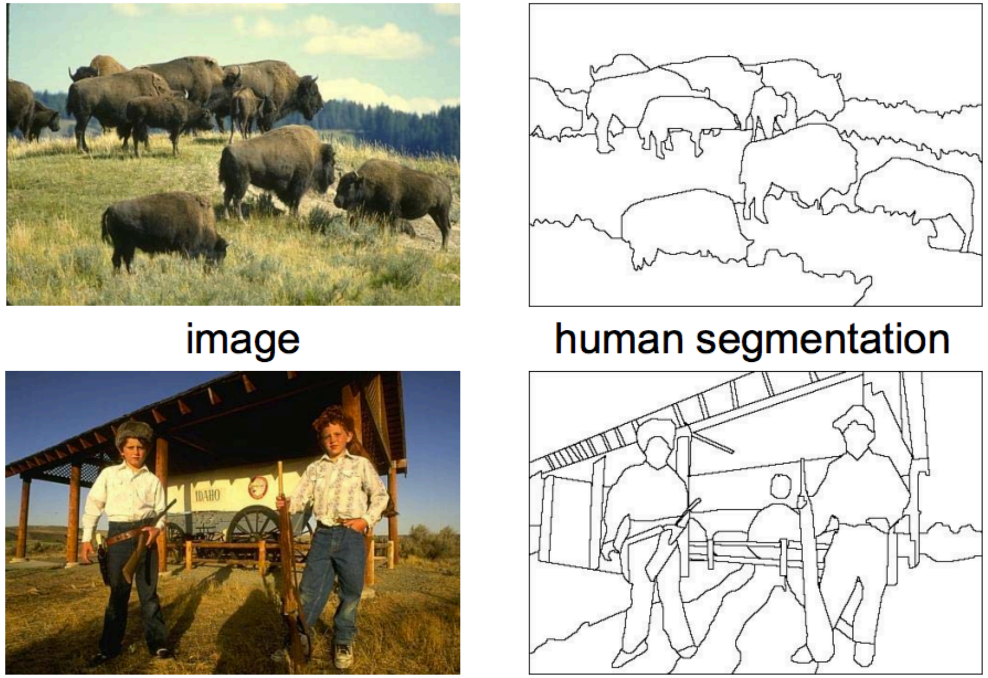
human labeling

ground-truth

Some Common Datasets in Image Segmentation

- **BSDS500：**

The dataset consists of 500 natural images (train, validation and test subsets) and ground-truth human annotations.

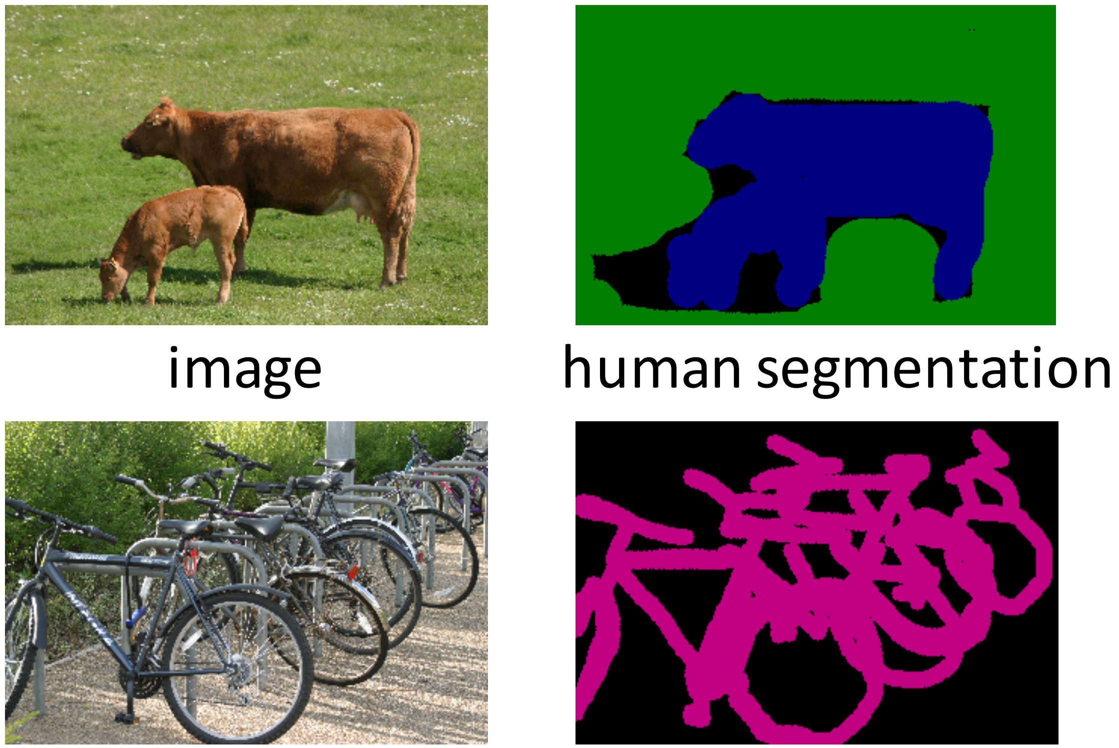


BSDS dataset:
https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html

Some Common Datasets in Image Segmentation

- **MSRC：**

The dataset consists of 591 natural images and ground-truth human annotations.



MSRC dataset:

https://www.microsoft.com/en-us/research/project/image-understanding/

Some Common Datasets in Image Segmentation

- **PASCAL VOC2012：**

The dataset consists of almost 10K natural images and ground-truth human annotations.

image

human segmentation

PASCAL dataset:
http://host.robots.ox.ac.uk/pascal/VOC/voc2012/

Image Segmentation Methods

- Threshold Based Segmentation
- Edge Based Segmentation
- Region Based Segmentation
- Active Contour Based Segmentation
- Graph Based Segmentation
- … …

Threshold Based Segmentation

- **The simplest image segmentation method**
- **Turn a gray-scale image $f(i,j)$ into a binary image $g(i,j)$**

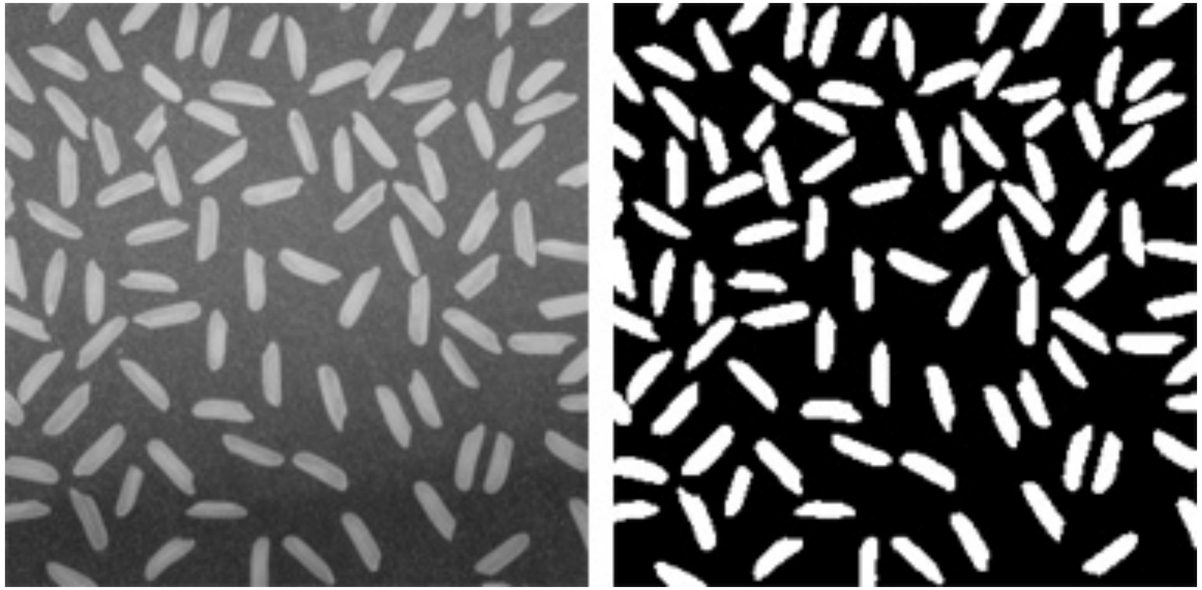
$$g(i,j) = \begin{cases} 1 & f(i,j) \geq T \\ 0 & f(i,j) < T \end{cases}$$

T : **threshold value**

The key of thresholding methods is to select a threshold value T

Otsu Thresholding

- Select the best threshold value to maximize the difference between background and target (maximize the inter-class variance)
- Global thresholding method



some images in BSDS500

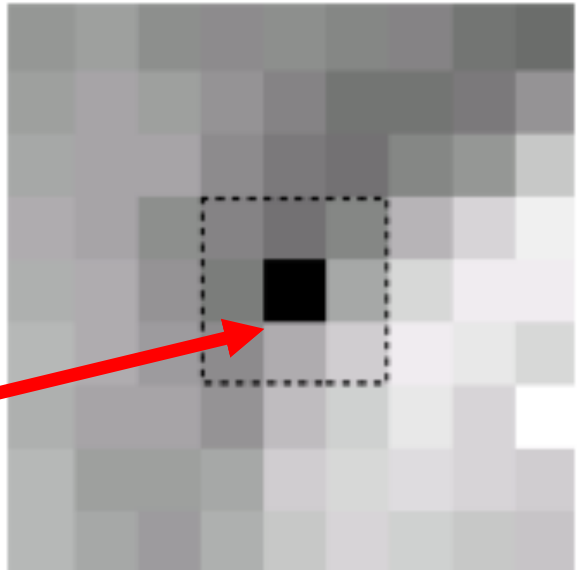
otsu

Adaptive Thresholding

- Finding the local threshold is to statistically examine the intensity values of the local neighborhood of each pixel

- Local thresholding method

a pixel and its 3×3 neighborhood



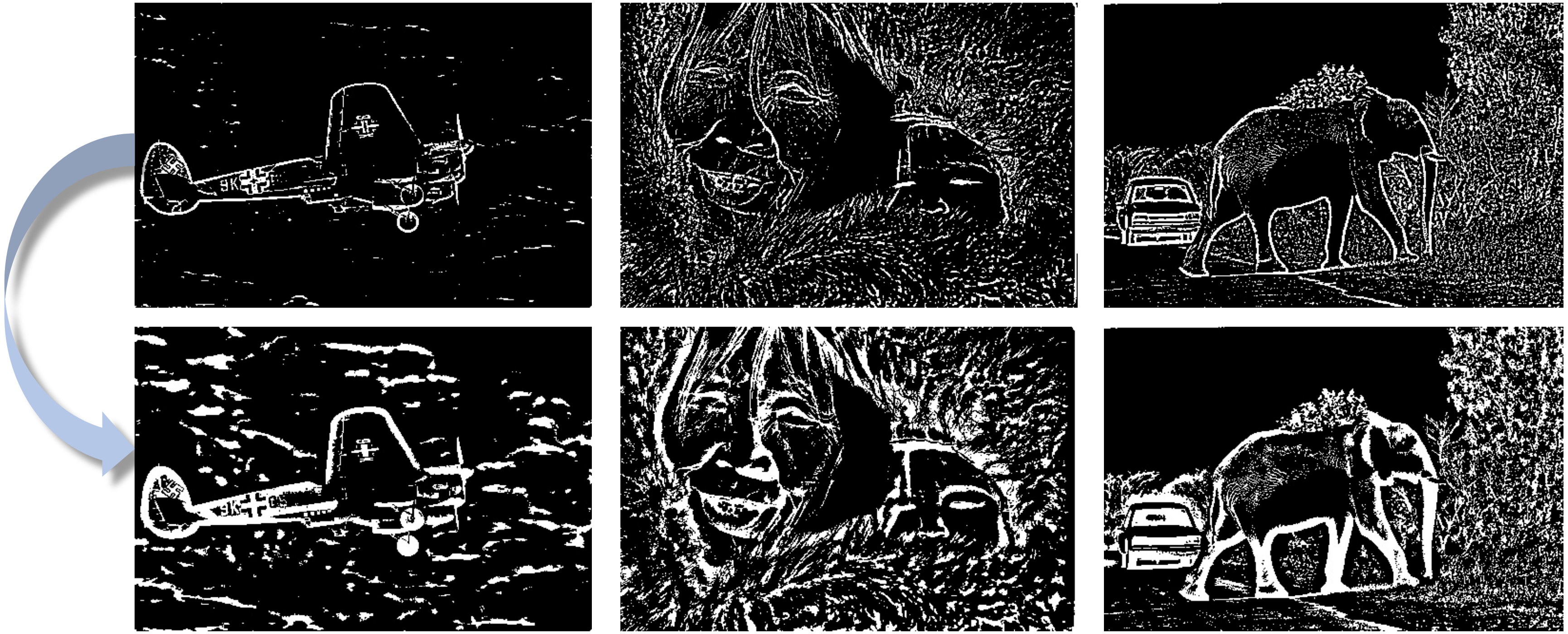
images in BSDS500

adaptive

more local details

images in BSDS500

adaptive



in larger neighborhood

more local details

About Thresholding Methods

Advantages:

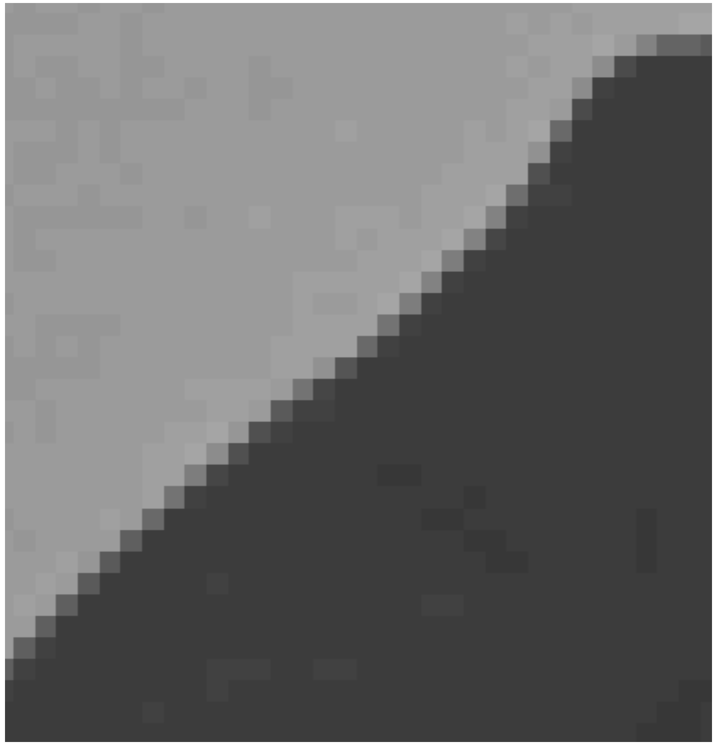
- Simplicity of calculation, fast

Disadvantages:

- Only consider gray-scale information
- Sensitive to noise

What is the edge?

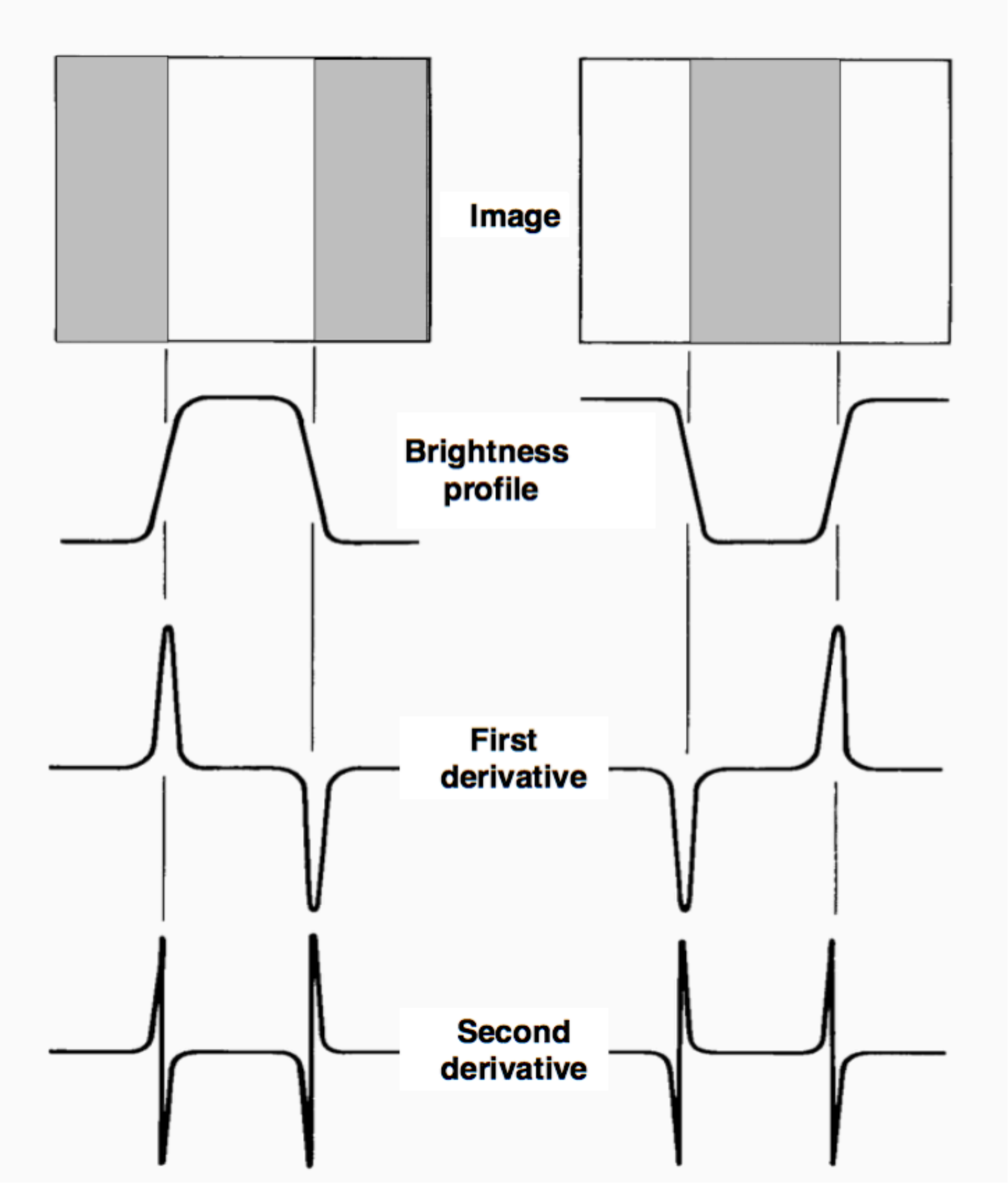
An edge is the boundary between two regions with distinct properties (color, intensity, or texture).



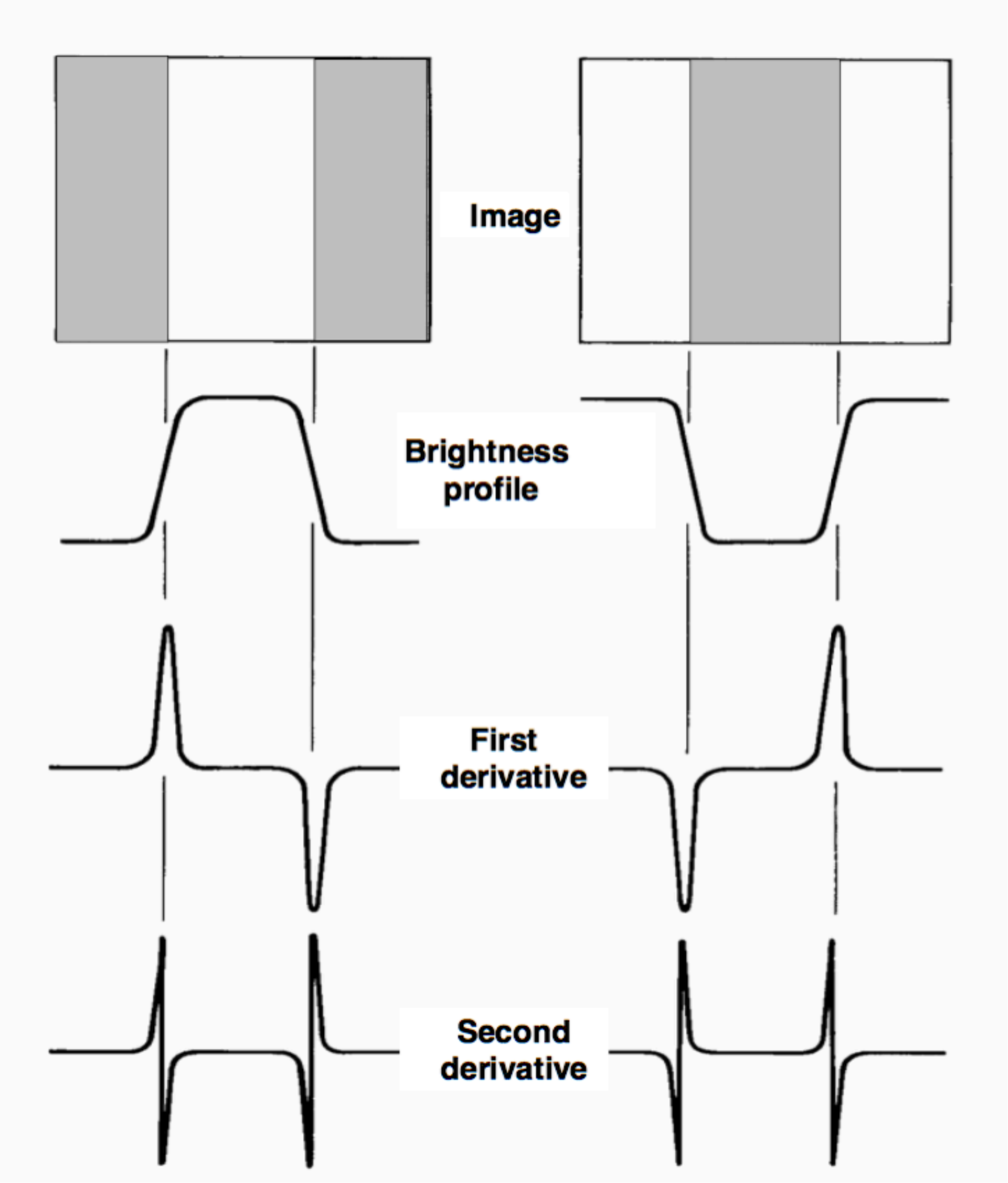
Edge Based Segmentation

- Edges are important features in a image to separate regions
- Based on discontinuity
- Edge detection is the principal approaches to image segmentation

What is edge detection?

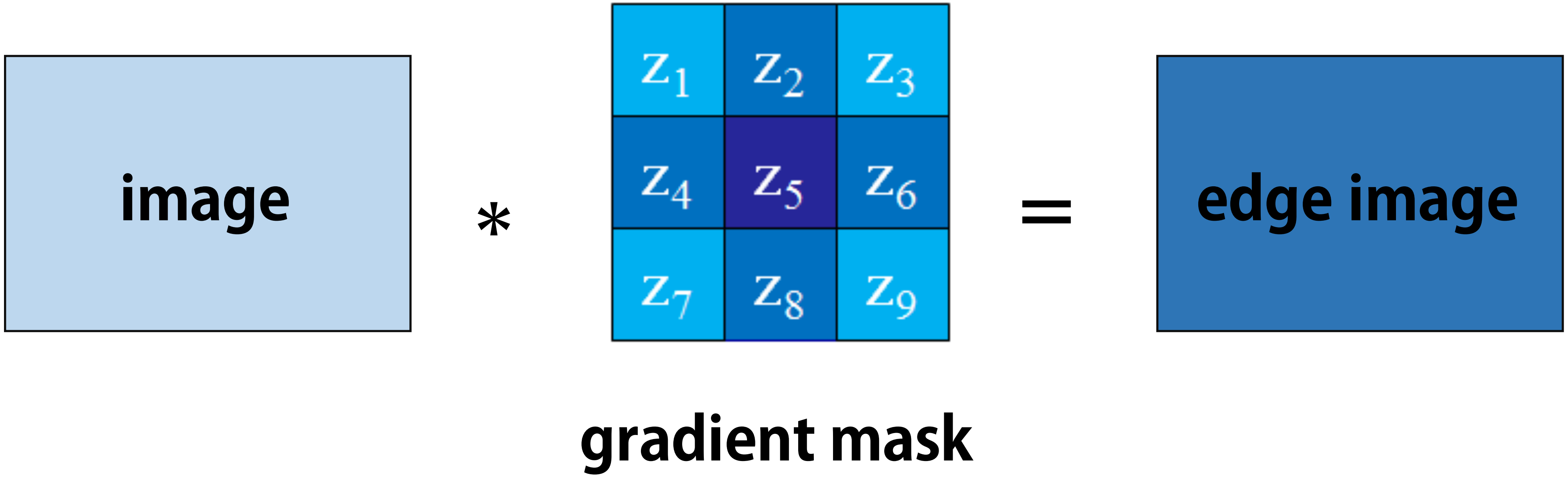


What is edge detection?



Edge detection by means of gradient operators

How gradient operators work?



Common Gradient Masks

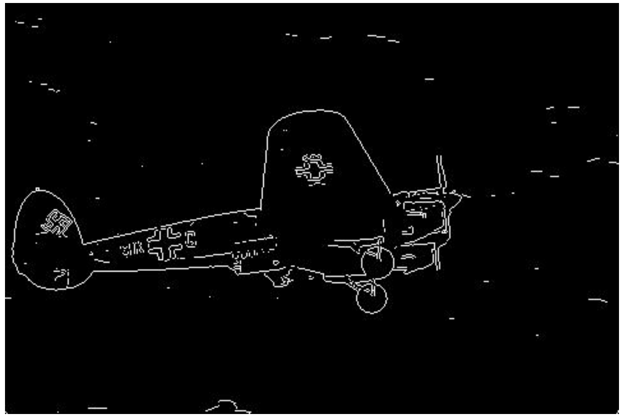
First order gradient masks：

- **Sobel**
- **Roberts**
- **Prewitt**
- **… …**

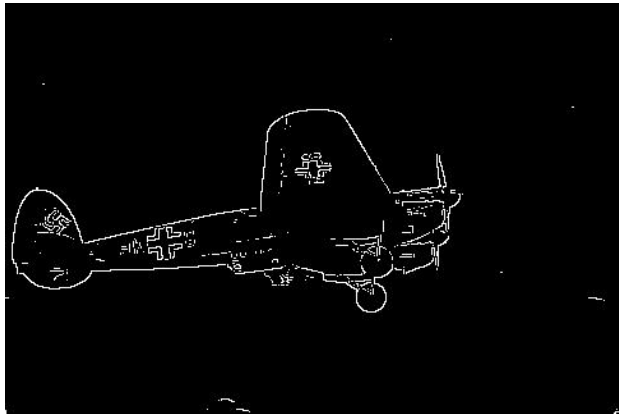
Second order gradient masks：

- **Canny**
- **LoG（Laplacian-Gauss）**
- **… …**

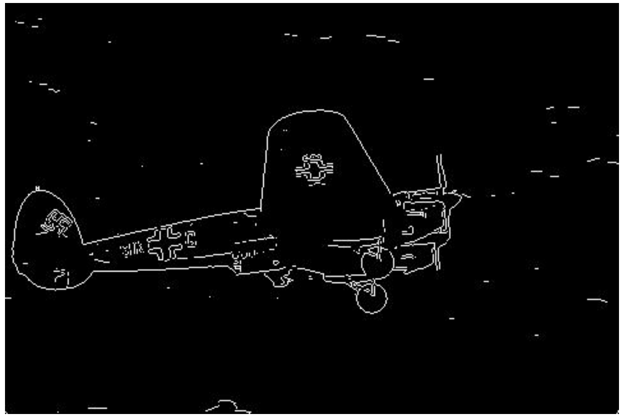
Sobel



Roberts

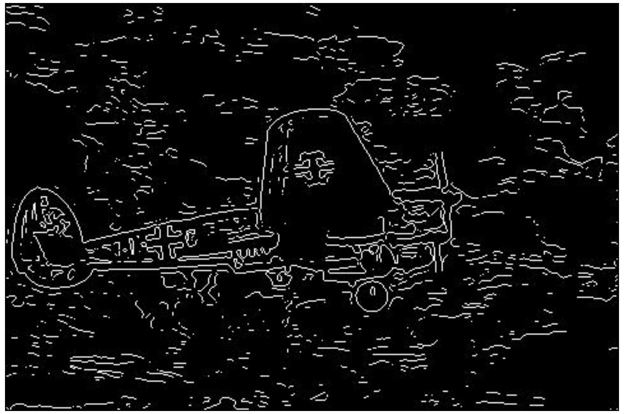


Prewitt



image

LoG



Canny

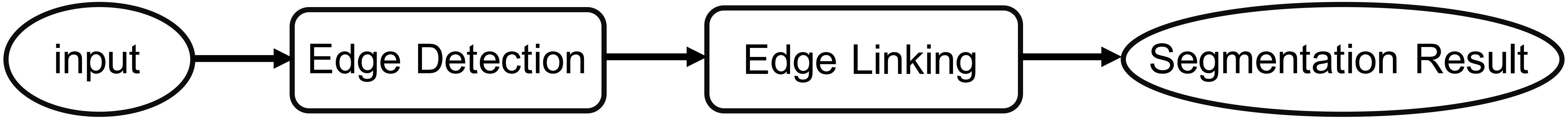
The Problem of Edge Detection

- **Noise and background affect the accurate of edge detection**

Edge presence in locations where there is no border, but no edge presence where a real border exists.

- **Edges are the sign of lack of continuity, and ending.**

The edges produced by edge detection can not be used directly, because edges are discontinuous.



From Edge to Region

The simplest way is some Morphological operations to fill the edge

About Edge Based Segmentation Methods

Advantages:

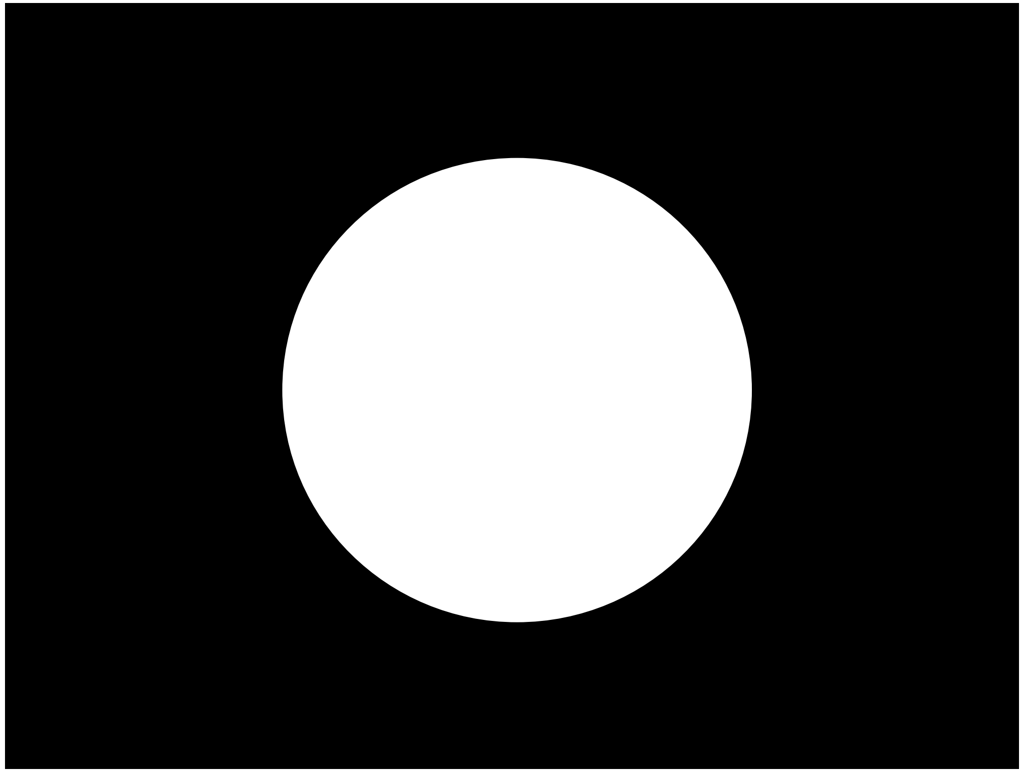
- Simplicity of calculation, fast
- Works well in images with good contrast between object and background

Disadvantages:

- Does not work well on images with smooth transitions and low contrast
- Sensitive to noise

What is a region?

A group of connected pixels with similar properties

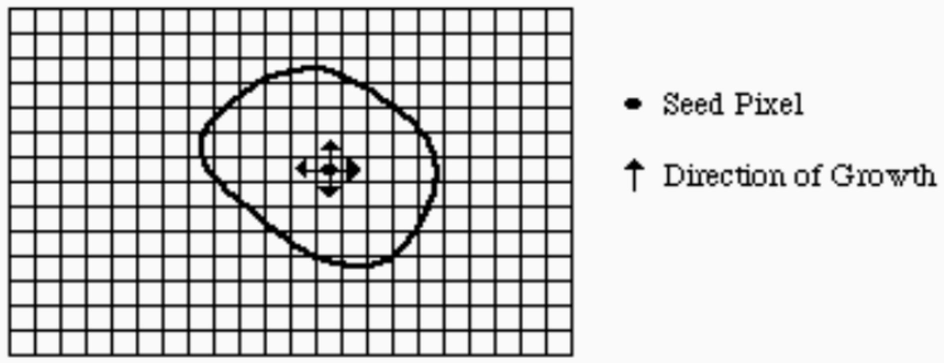


Region Based Segmentation

- Region is an important concept because regions may correspond to objects in a scene
- Based on similarity
- Those pixels that correspond to an object are grouped together and marked
- Region based segmentation is a technique that allows us to determine the regions directly

Region Growing Segmentation

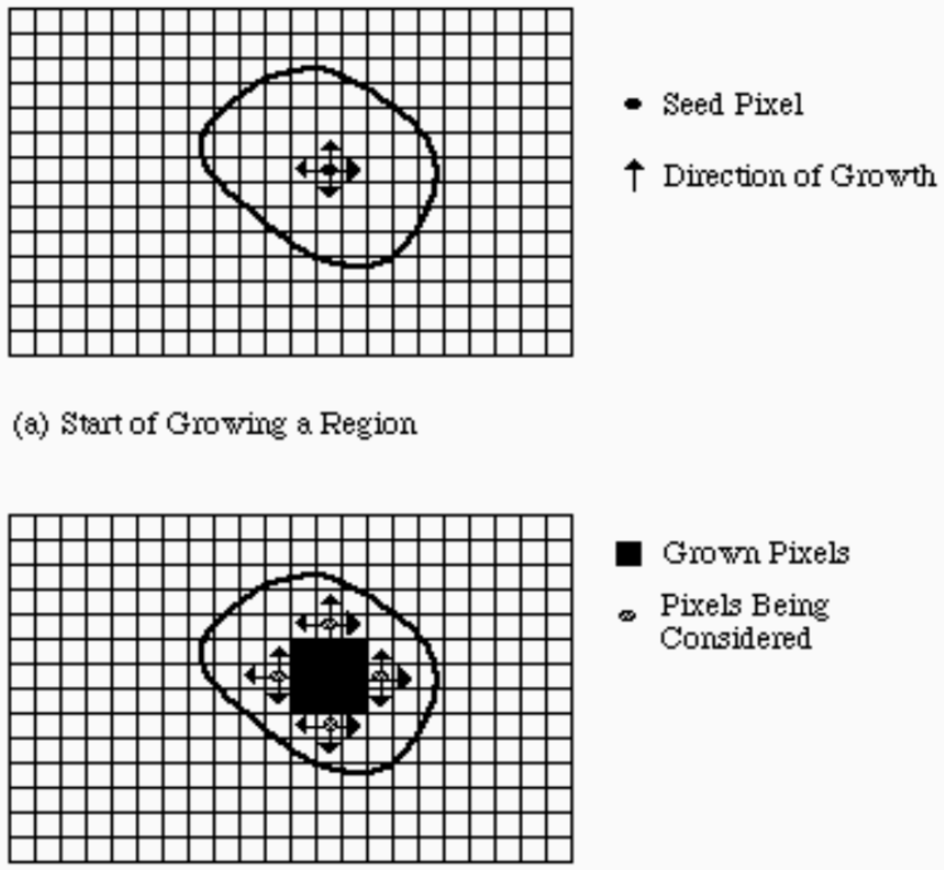
- Start from some seed points to represent the property we want, then grow the region
- A bottom up method



(a) Start of Growing a Region

Region Growing Segmentation

- Start from some seed points to represent the property we want, then grow the region
- A bottom up method

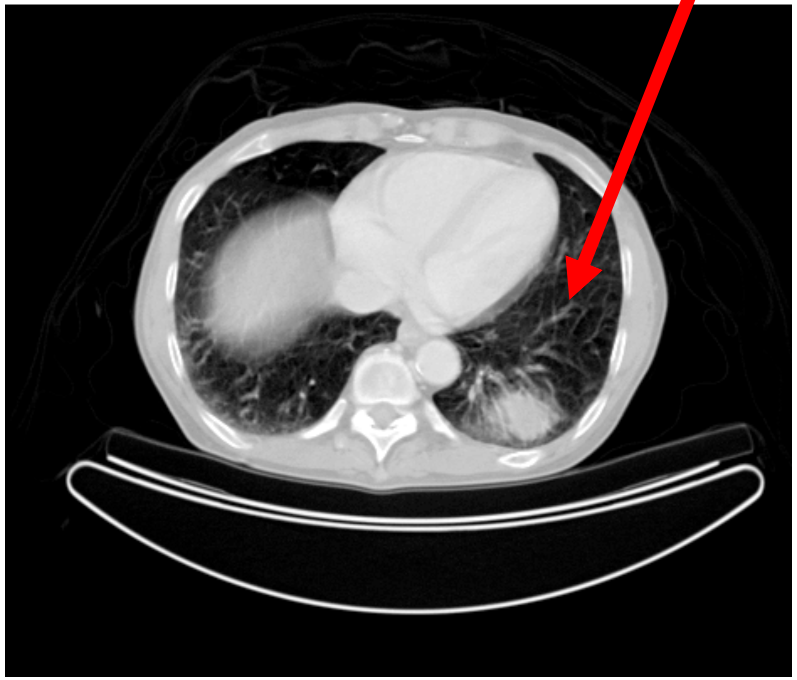


(a) Start of Growing a Region

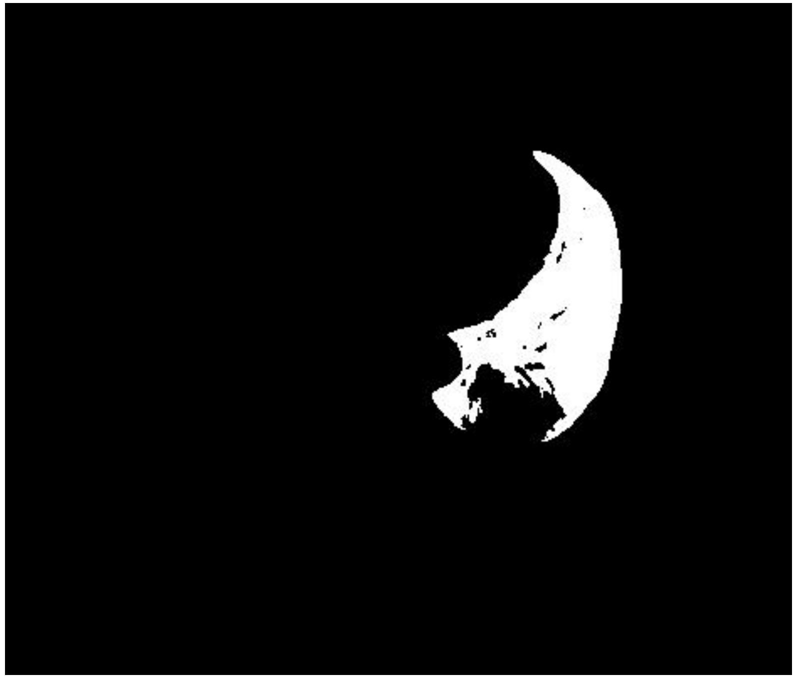
(b) Growing Process After a Few Iterations

An Example of Region Growing

set a seed around here



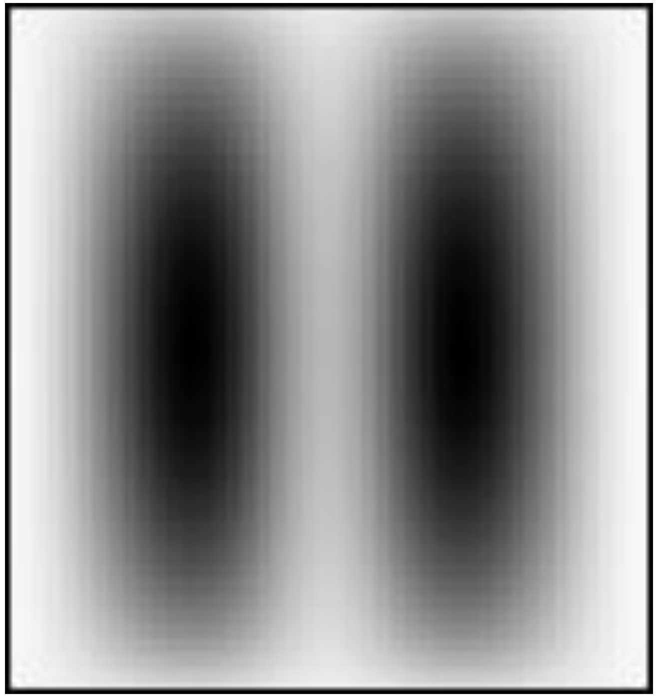
image

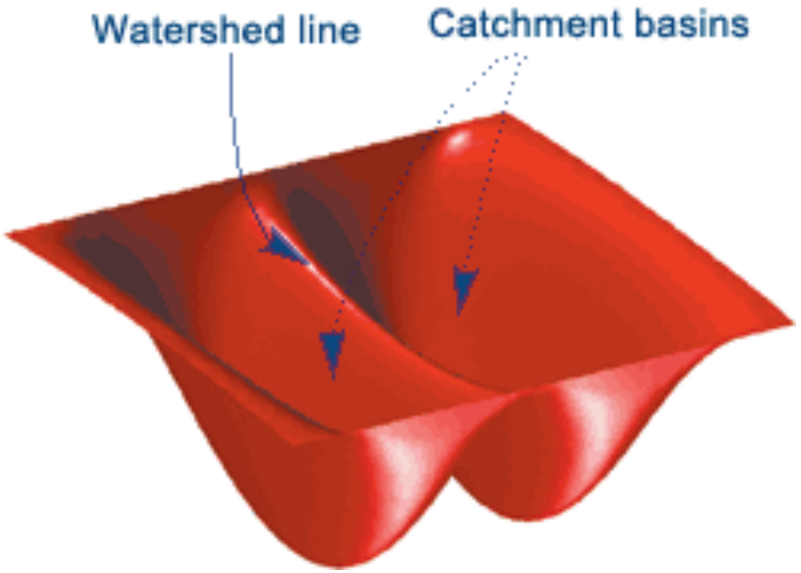


result

Watershed Segmentation

- An image is considered as a 3D topographic surface with valleys and mountains

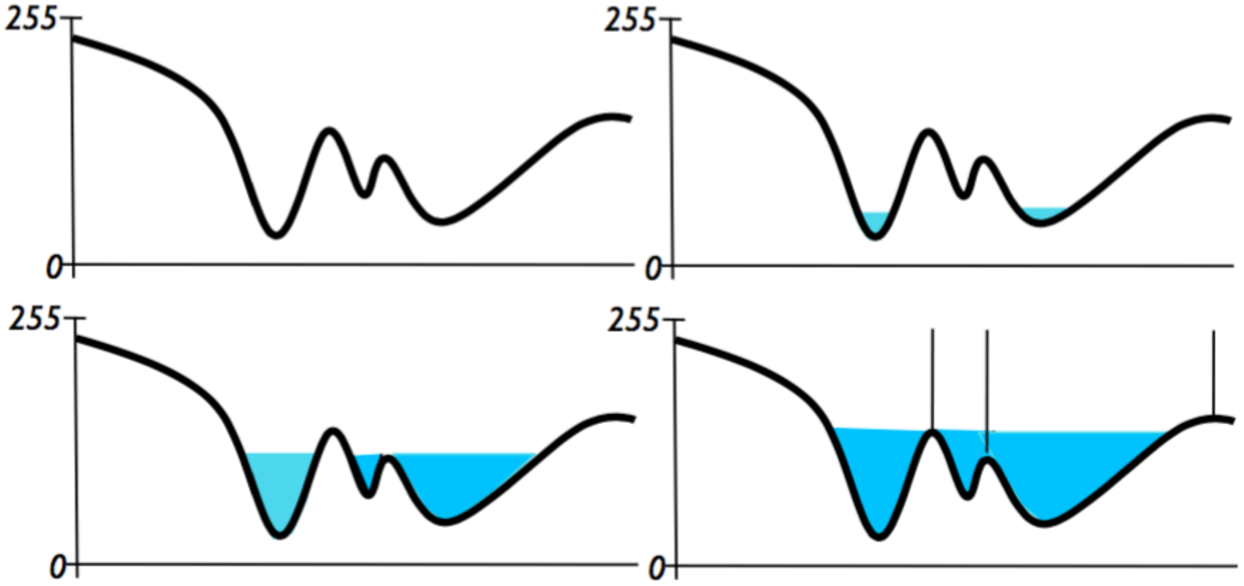


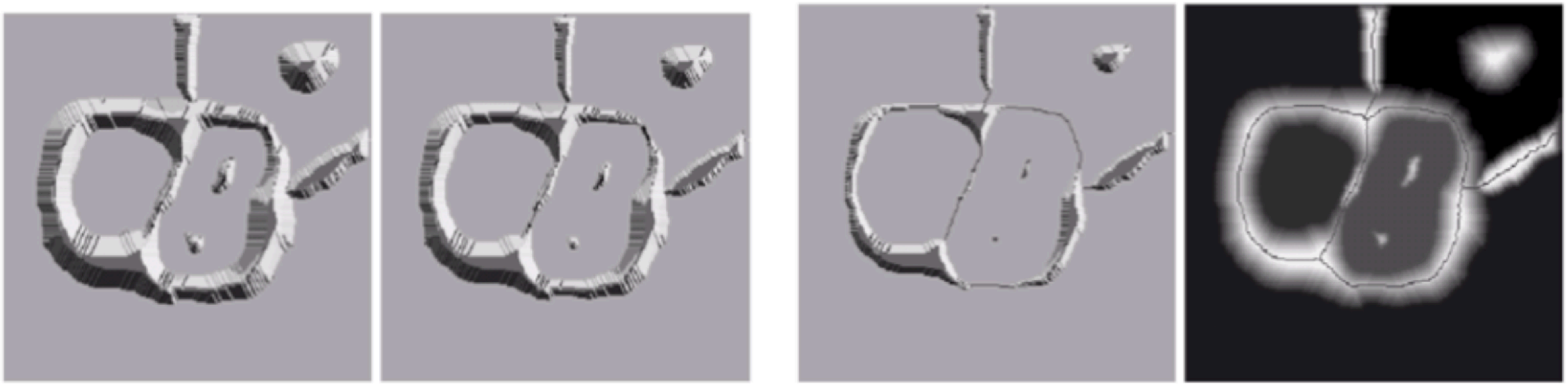


Watershed Segmentation

- The goal of watershed segmentation:
To find the watershed lines

- Watershed lines are remained by flooding water





Watershed Segmentation

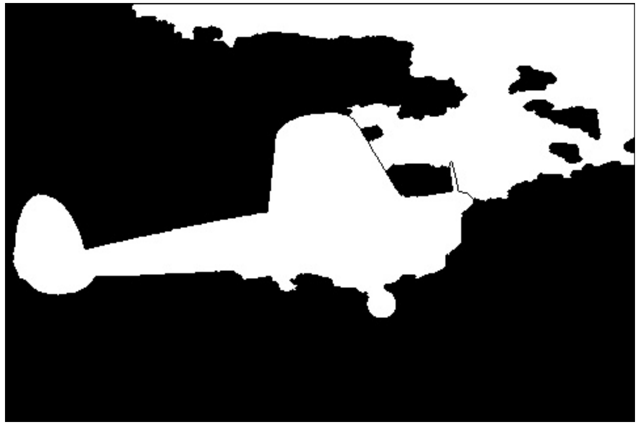
The difficulty of watershed segmentation is determining the masks as flooding water

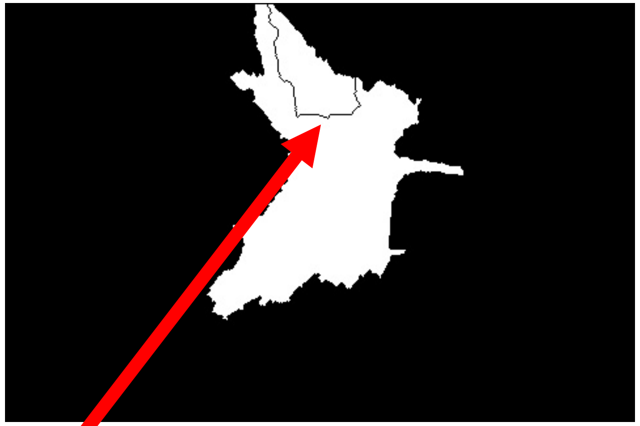
image

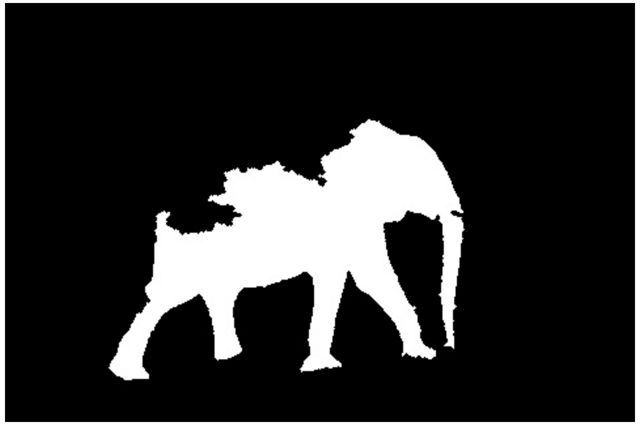
watershed

image

watershed



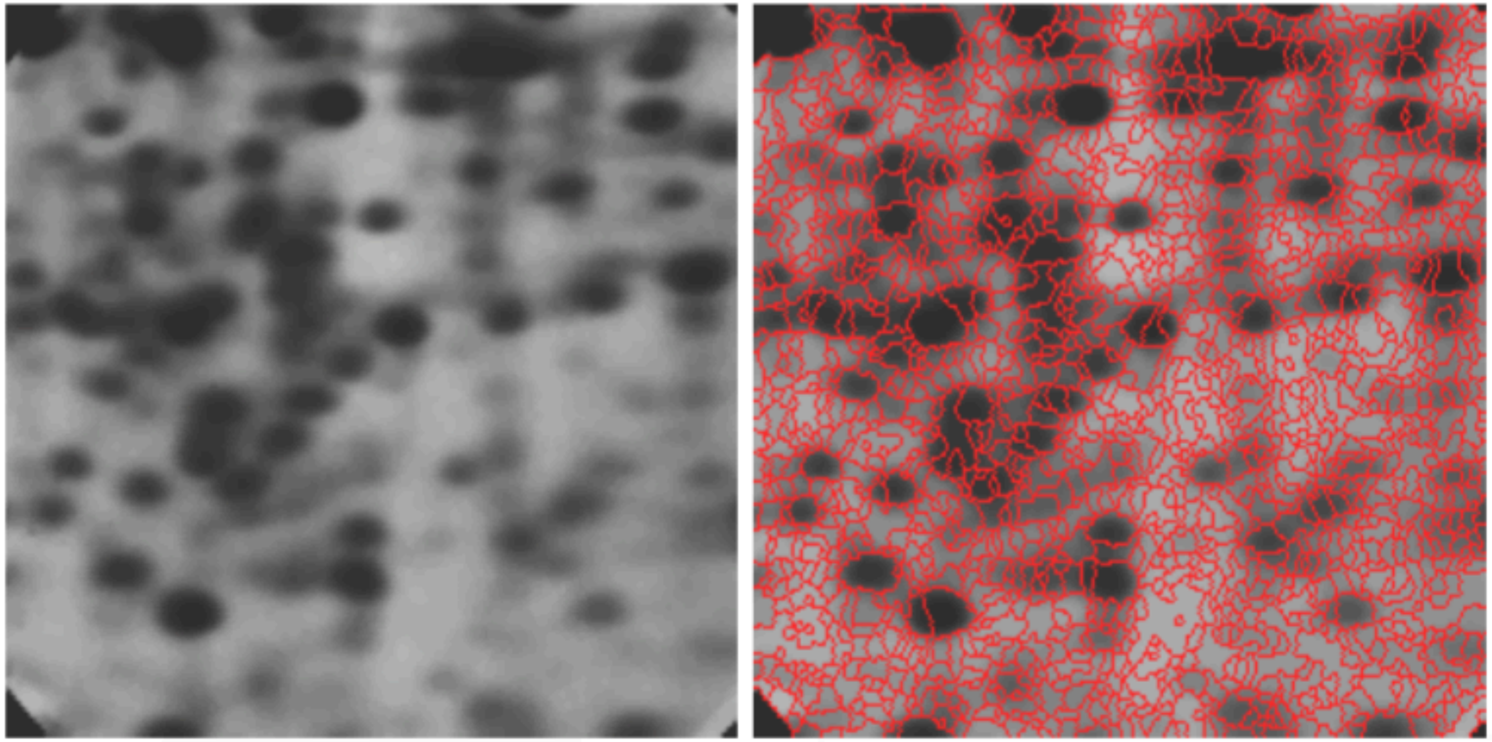




over-segmentation

Over-segmentation

A division into too many segments, as when attempting to recognize parts of an image.



About Region Based Segmentation Methods

Advantages:

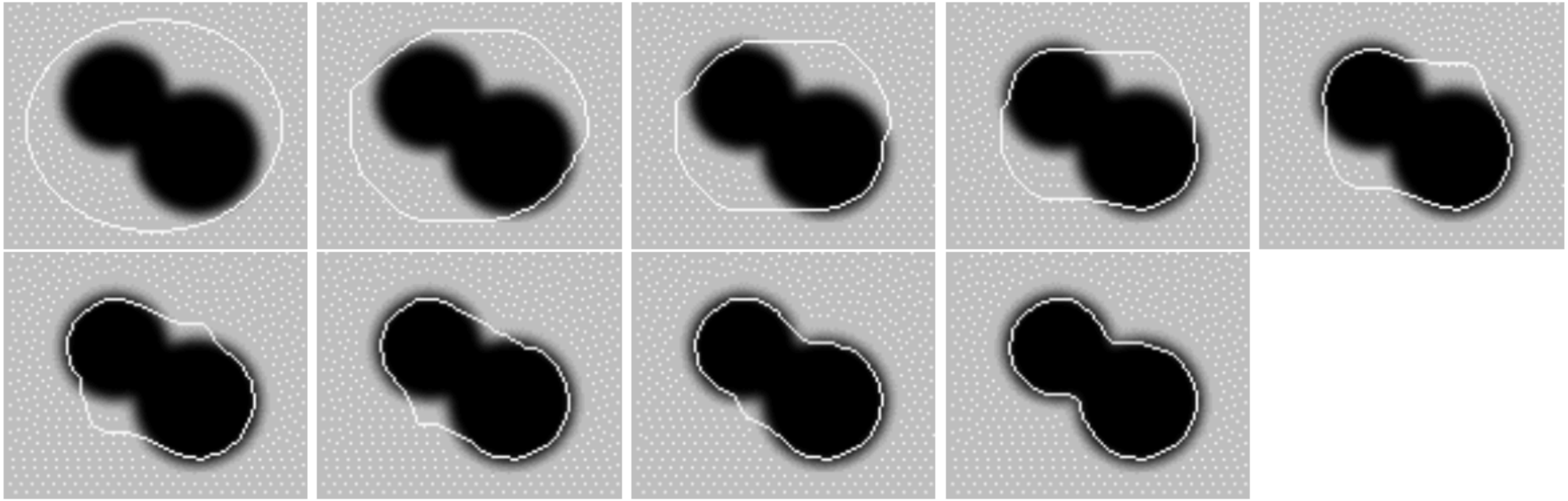
- Simplicity of calculation, fast

Disadvantages:

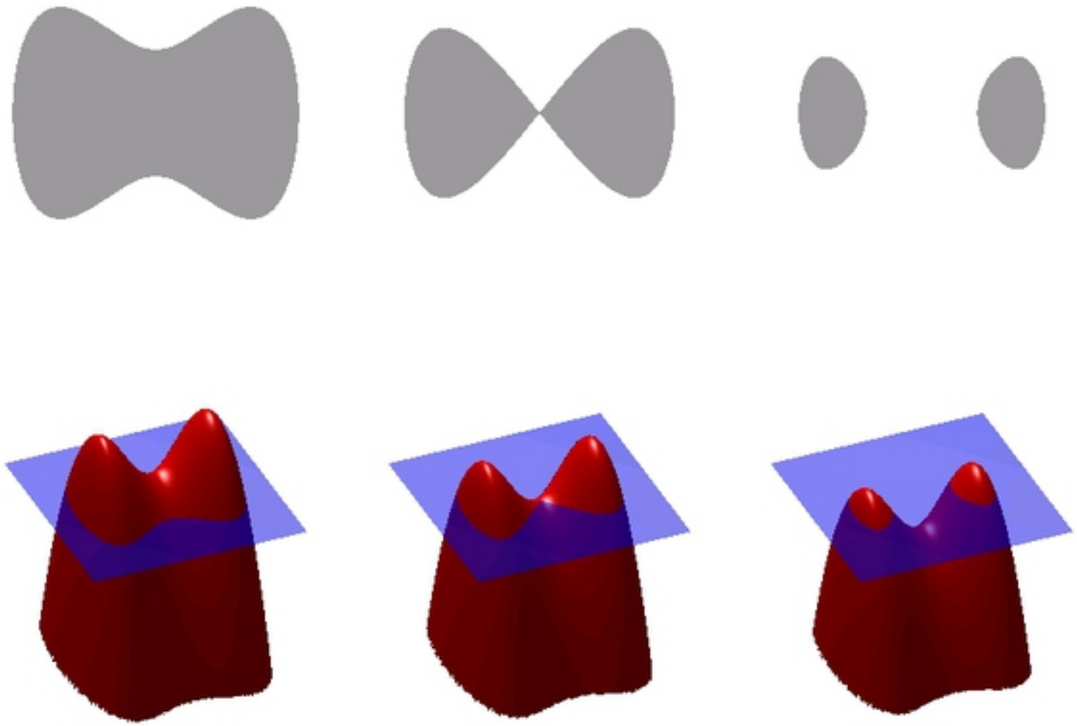
- These two region based segmentation methods need human marks
- Sensitive to noise
- Watershed segmentation usually causes over-segmentation phenomenon

Active Contour Based Segmentation

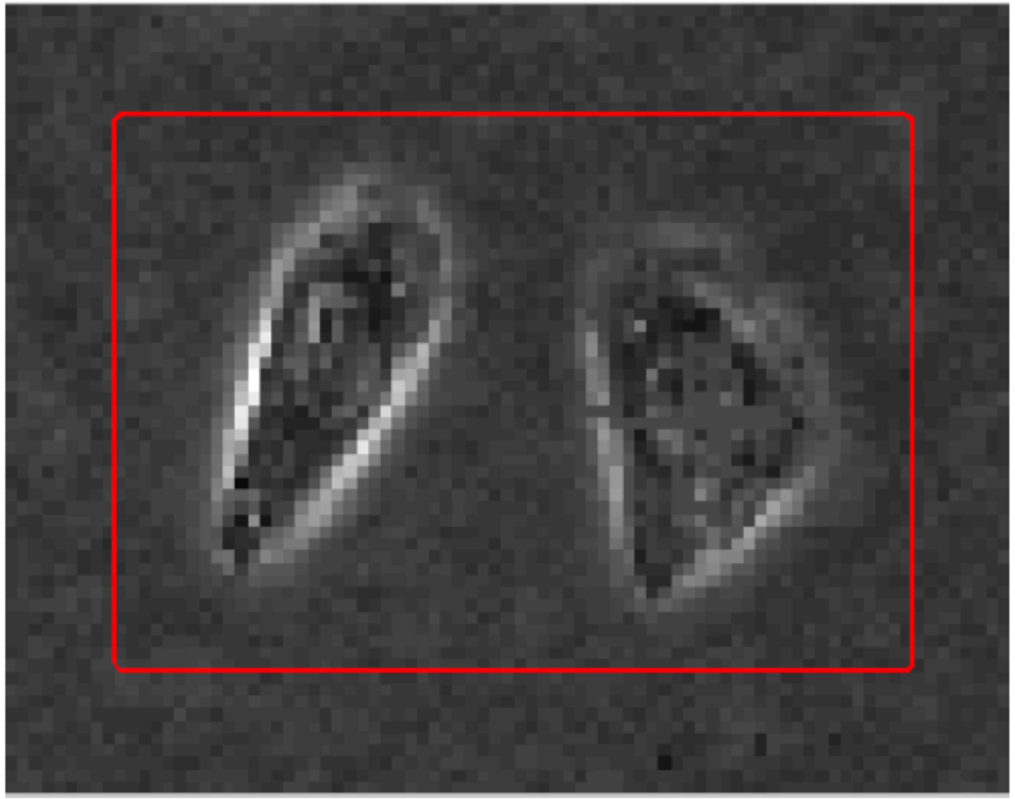
- **Start from a contour**
- **Snake, Level set**



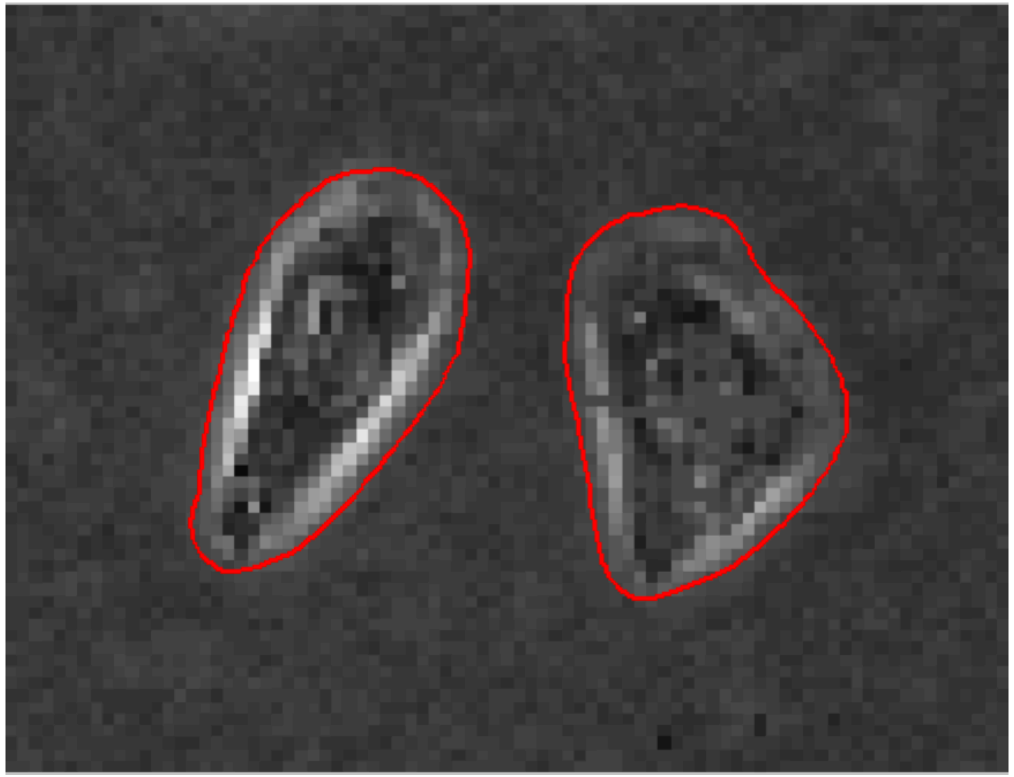
Level Set Segmentation



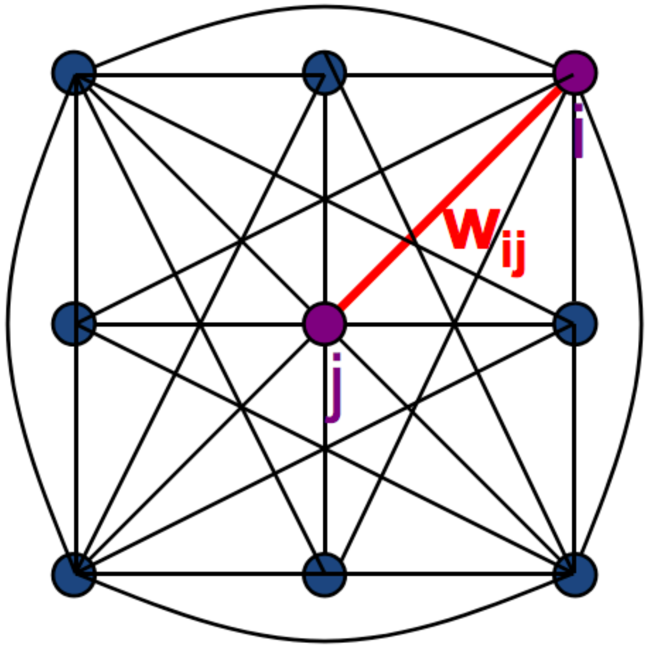
An Example of Level Set Segmentation



An Example of Level Set Segmentation



Graph Based Segmentation

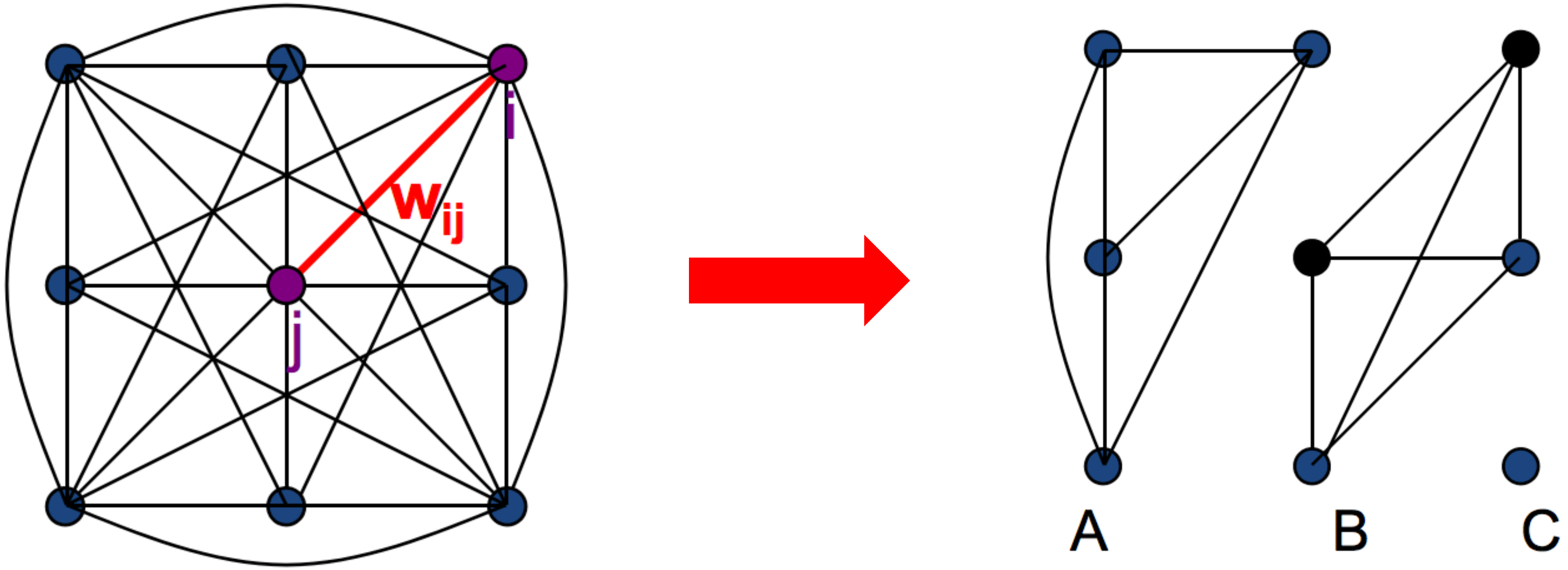


Images as Graphs : G = {V,E}

V (Vertex): graph nodes
E (Edge): edges connection nodes

Image = { pixels }
Pixel similarity

Break Graph into segments



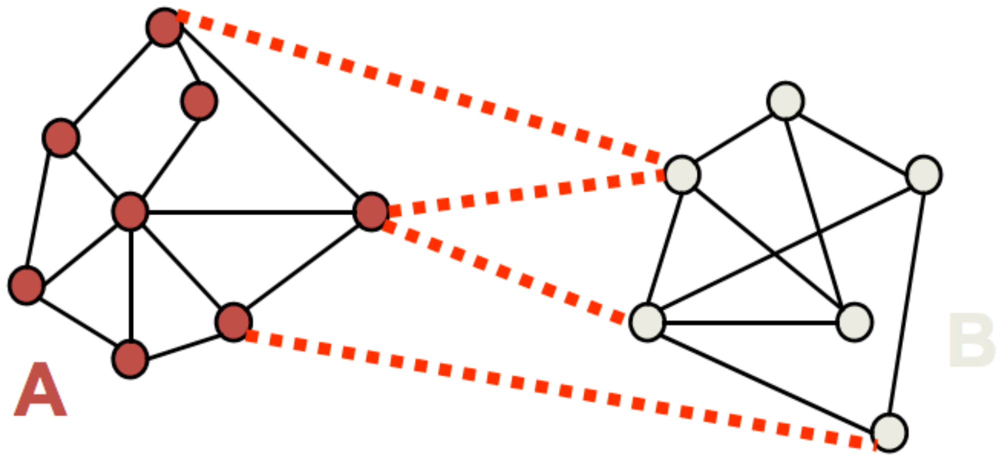
Delete links that cross between segments
Easiest to break links that have low cost (low similarity)

Cuts in a Graph

Link Cut

- set of links whose removal makes a graph disconnected
- cost of a cut:

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



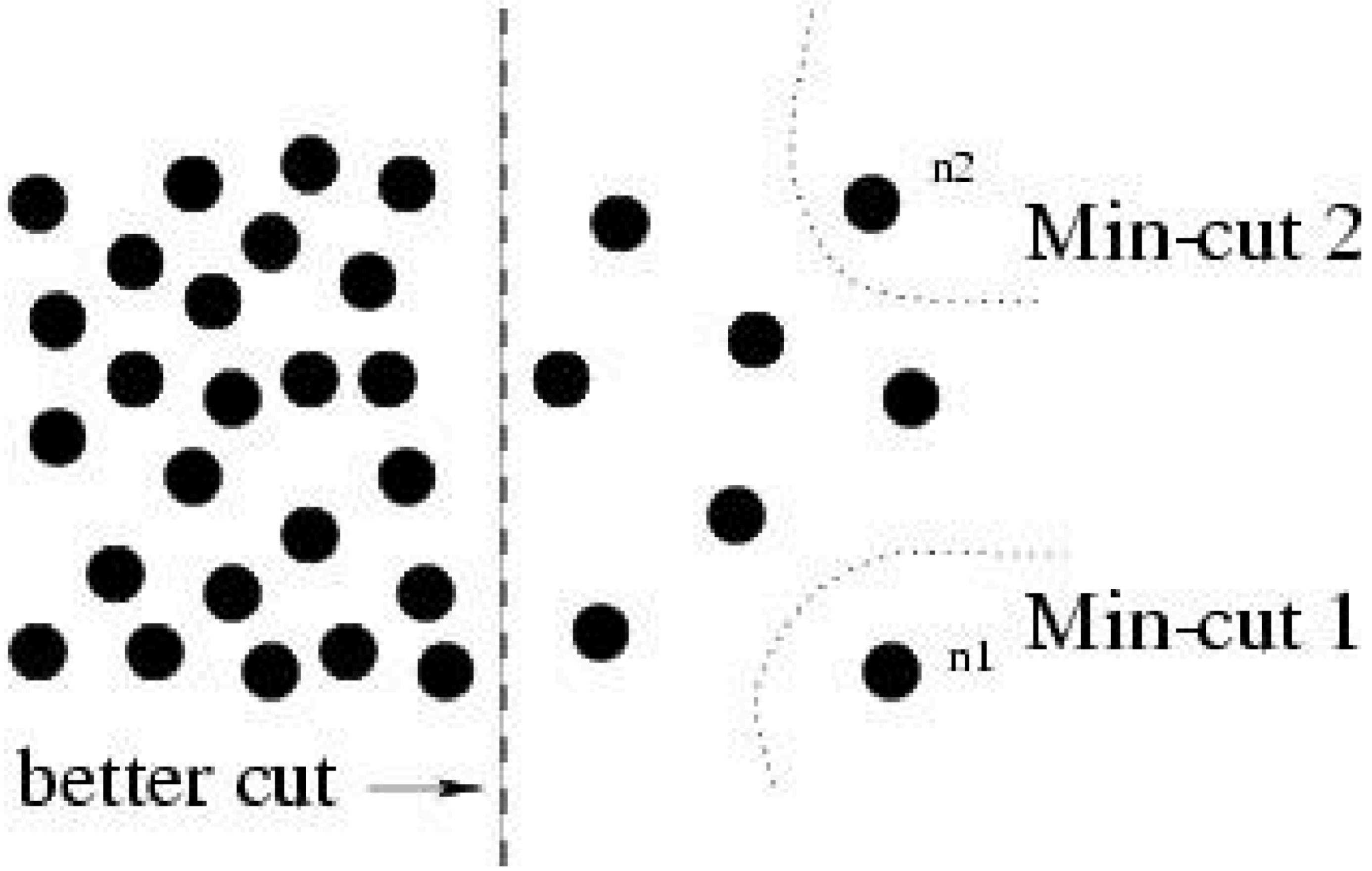
One idea: find minimum Cut (Min Cut)

Two Interactive Min Cut Based Segmentation Methods

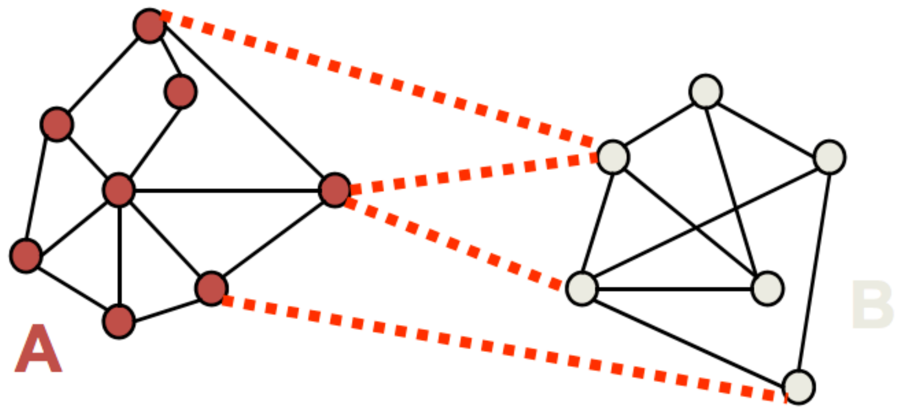
- **Graph Cuts**

- **Grab Cuts**

But Min Cut is not always the best Cut



Normalized Cuts



- A cut penalizes large segments
- Fix by normalizing for size of segments

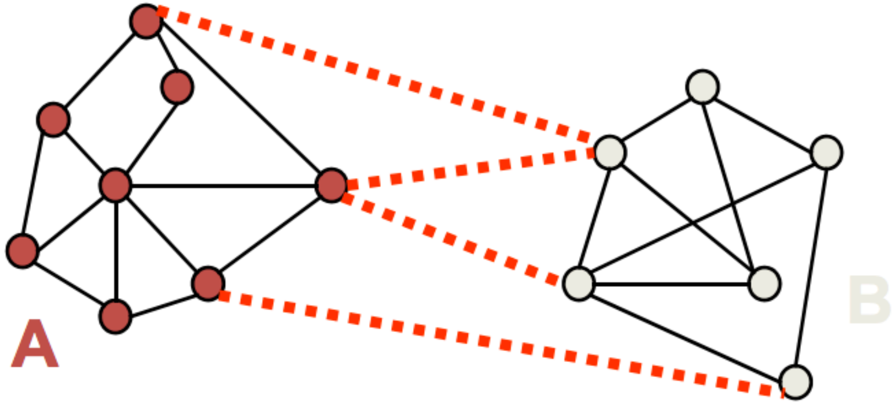
$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)}$$

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

- assoc(A,V) = sum of costs of all edges that touch A

"Normalized Cuts and Image Segmentation", JianBo Shi and Jitendra Malik, PAMI, 2000

Normalized Cuts



- A cut penalizes large segments
- Fix by normalizing for size of segments

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

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

$$minNcut(A,B) \text{ is NP} - \text{Hard}$$

- assoc(A,V) = sum of costs of all edges that touch A

"Normalized Cuts and Image Segmentation", JianBo Shi and Jitendra Malik. PAMI, 2000

Solution

$$\boxed{minNcut(A,B)}$$

approximation

Solve for eigenvectors with the smallest eigenvalues:

$$\boxed{(D-W)y=\lambda Dy}$$

$$D=\begin{bmatrix}\sum_j w(1,j) & \cdots & 0\\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sum_j w(N,j)\end{bmatrix}$$

$$W=\begin{bmatrix}w(1,1) & \cdots & w(1,N)\\ \vdots & \ddots & \vdots \\ w(N,1) & \cdots & w(N,N)\end{bmatrix}$$

Solution

$$(\boldsymbol{D} - \boldsymbol{W})\boldsymbol{y} = \boldsymbol{\lambda D y}$$

- **Use the eigenvector with the second smallest eigenvalue to bipartition the graph**

- **Recursively repartition the segmented parts if necessary**

Some Normalized Cuts Results

need to set
the number of
segments (K)

Some Normalized Cuts Results

need to set
the number of
segments (K)

here, K=10

About Normalized Cuts

Advantages:

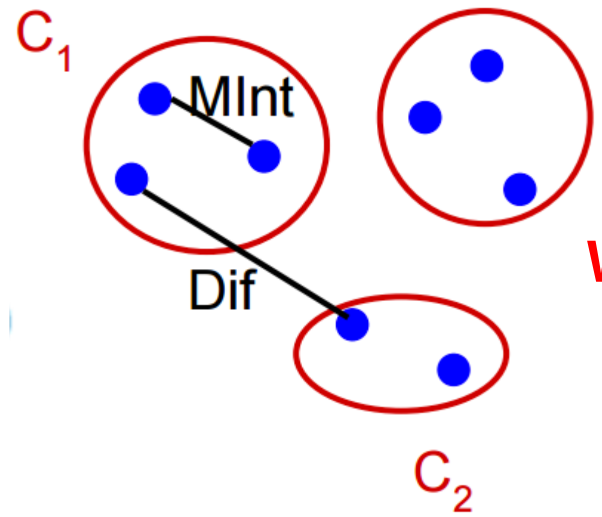
- Generic framework, can be used with many different features and affinity formulations

Disadvantages:

- Need to chose number of segments
- High storage requirement and time complexity

Efficient Graph-Based Image Segmentation

Predicate D determines whether there is a boundary for segmentation



$$D(C_1, C_2) = \begin{cases} \text{true} & \text{if} Dif(C_1, C_2) > MInt(C_1, C_2) \\ \text{false} & \text{otherwise} \end{cases}$$

Where

$Dif(C_1, C_2)$ **is the difference between two components**

$MInt(C_1, C_2)$ **is the minimum internal different in the components** C_1**and** C_2

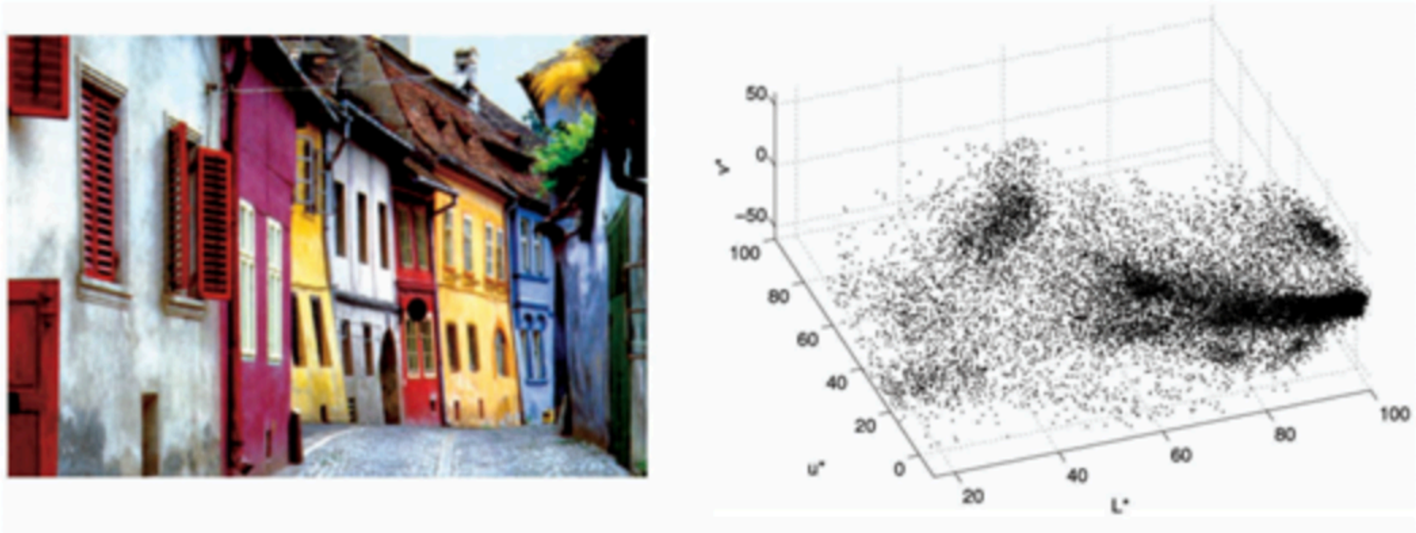
"Efficient Graph-Based Image Segmentation", P. Felzenszwalb and D. Huttenlocher, IJCV, 2004

Efficient Graph-Based Image Segmentation Results

Need to tune parameters by hand

Mean Shift

- **A feature space analysis method**

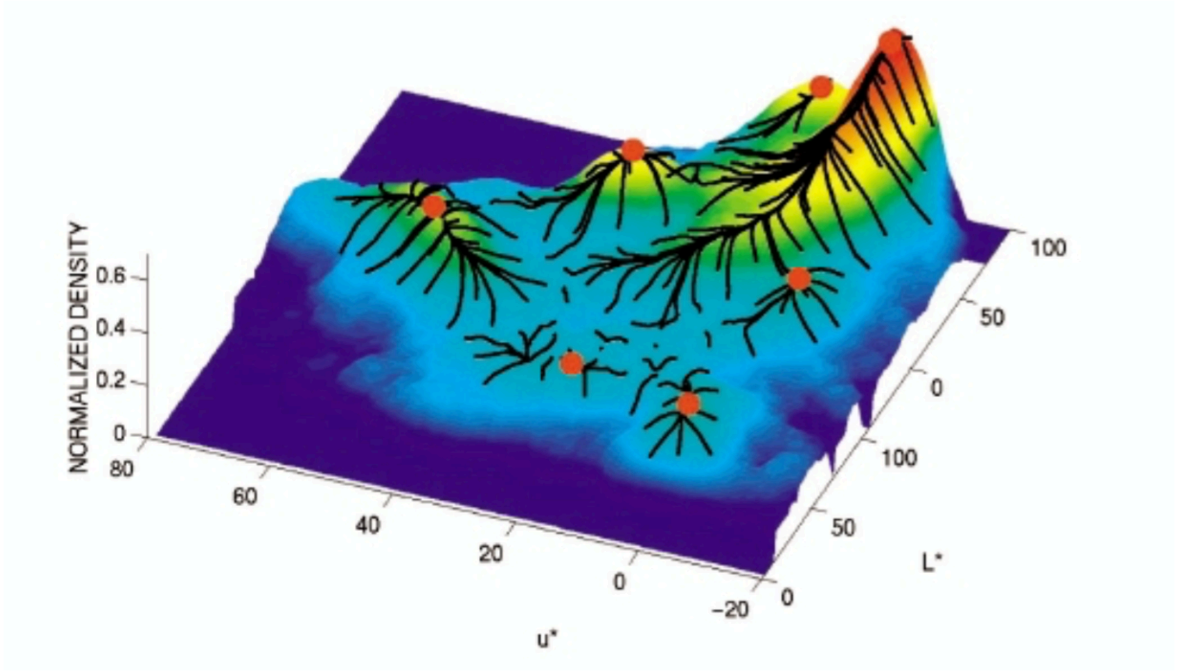


- **Built on probabilistic intuitions**

"Mean shift: a robust approach toward feature space analysis", D. Comaniciu and P. Meer, PAMI, 2002

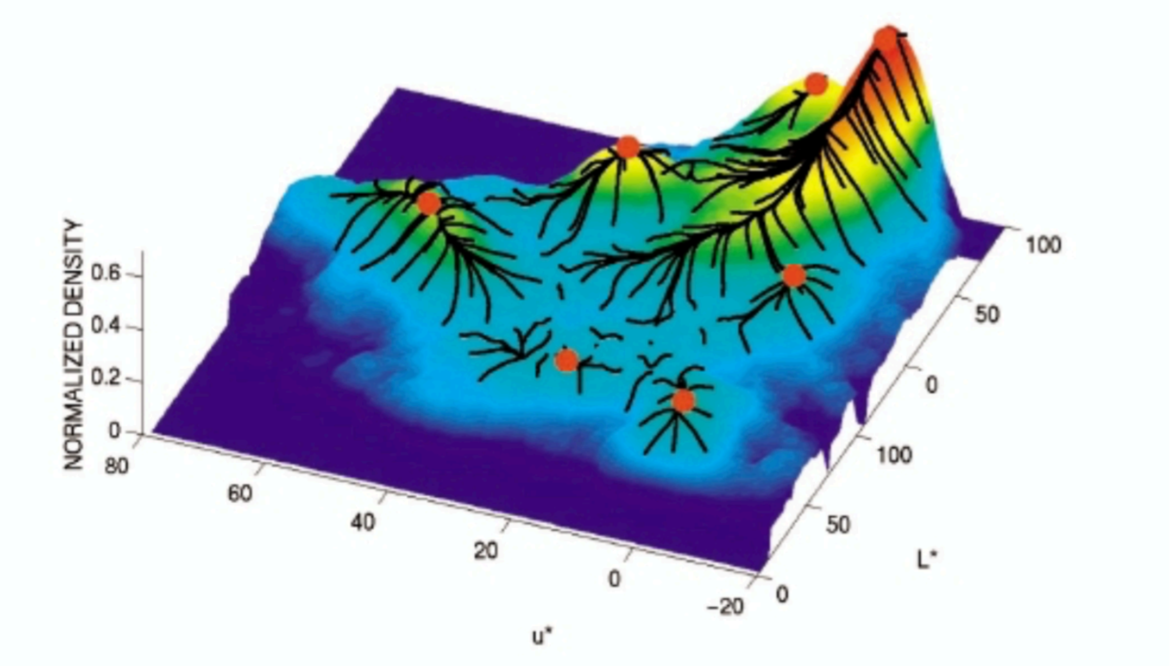
Non-Parametric Methods

- **Basic Idea: Use the data to define the distribution**
- **Kernel-density estimates**



What is Mean-Shift?

- The density will have peaks (also called modes)
- If we started at point and did gradient-ascent, we would end up at one of the modes
- Cluster based on which mode each point belongs to

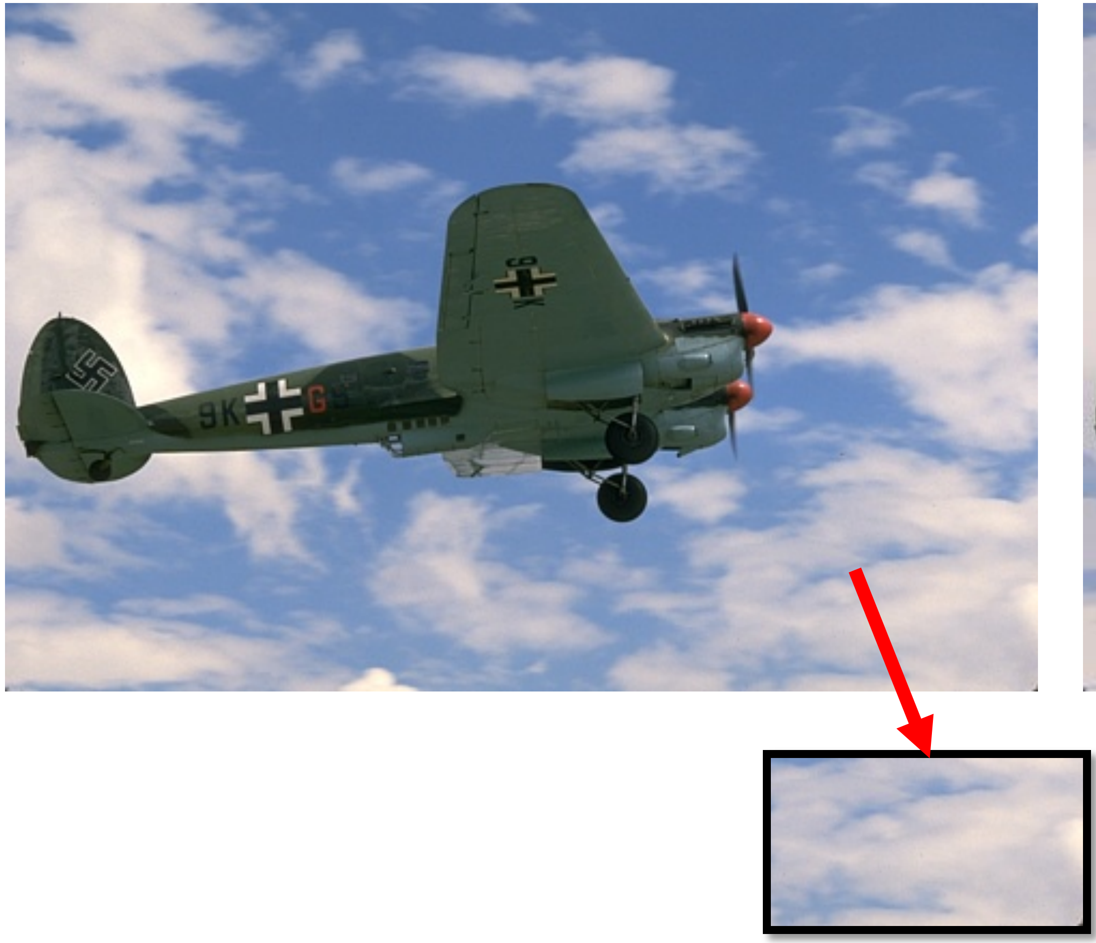


Gradient Ascent?

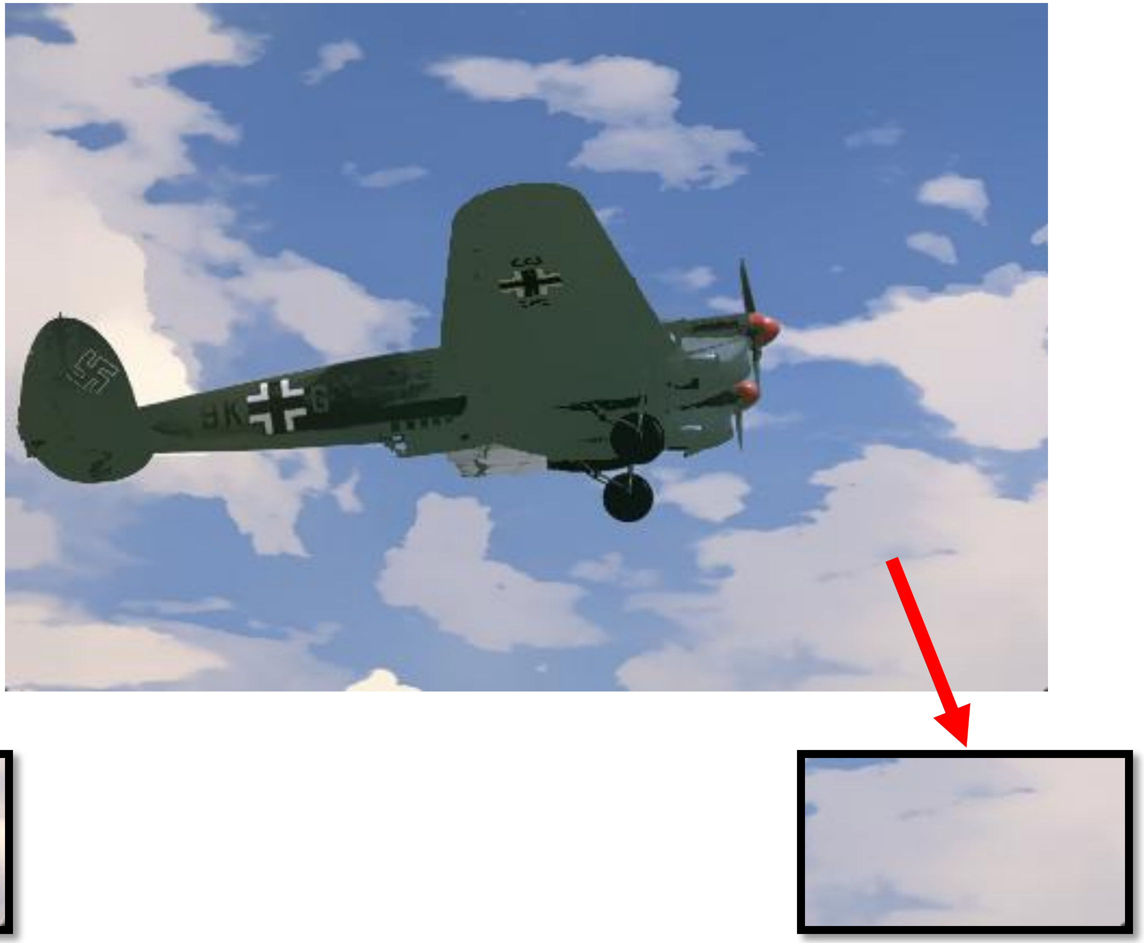
- Actually, no
- A set of iterative steps can be taken that will monotonically converge to a mode
  - No worries about step sizes
  - This is an adaptive gradient ascent

Mean Shift Result

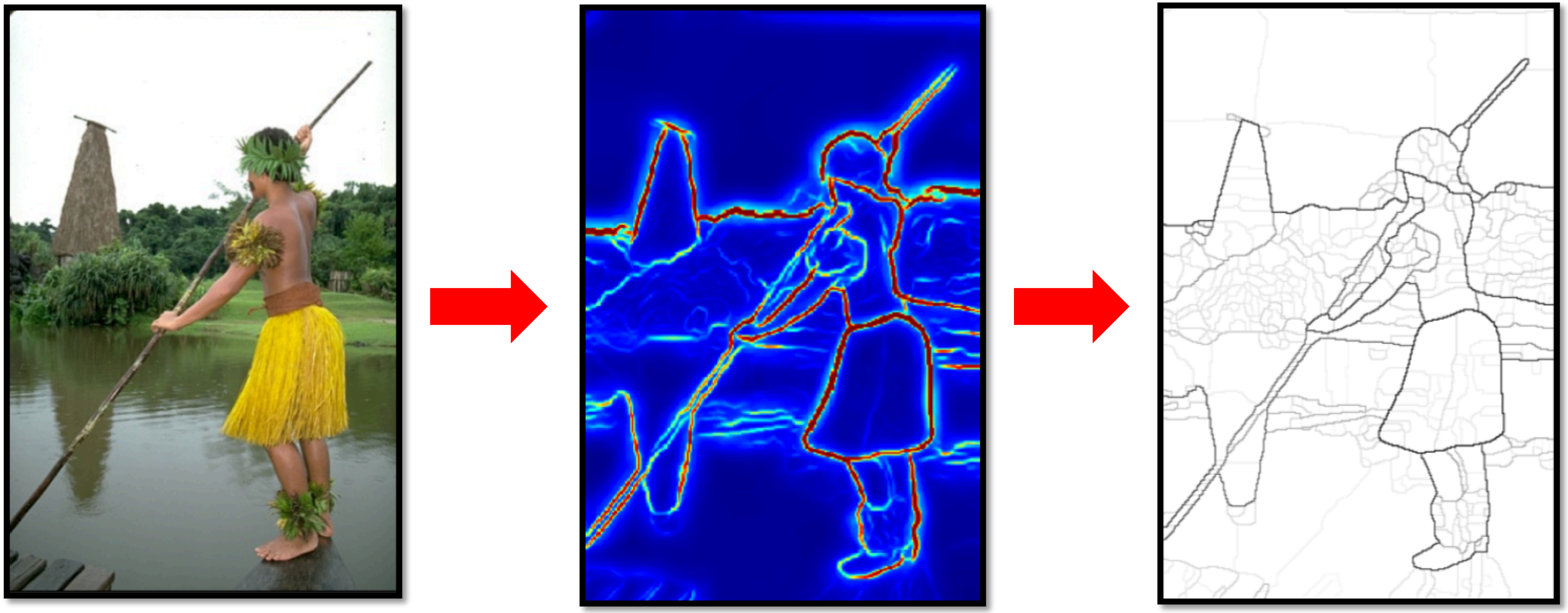
image



result



gPb-owt-ucm



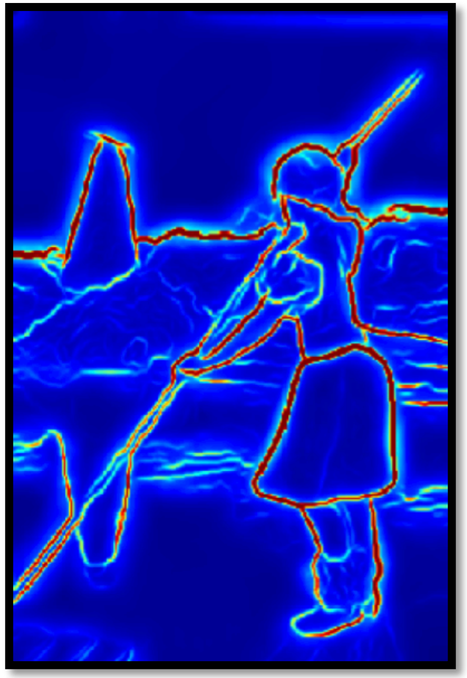
Image

Contour

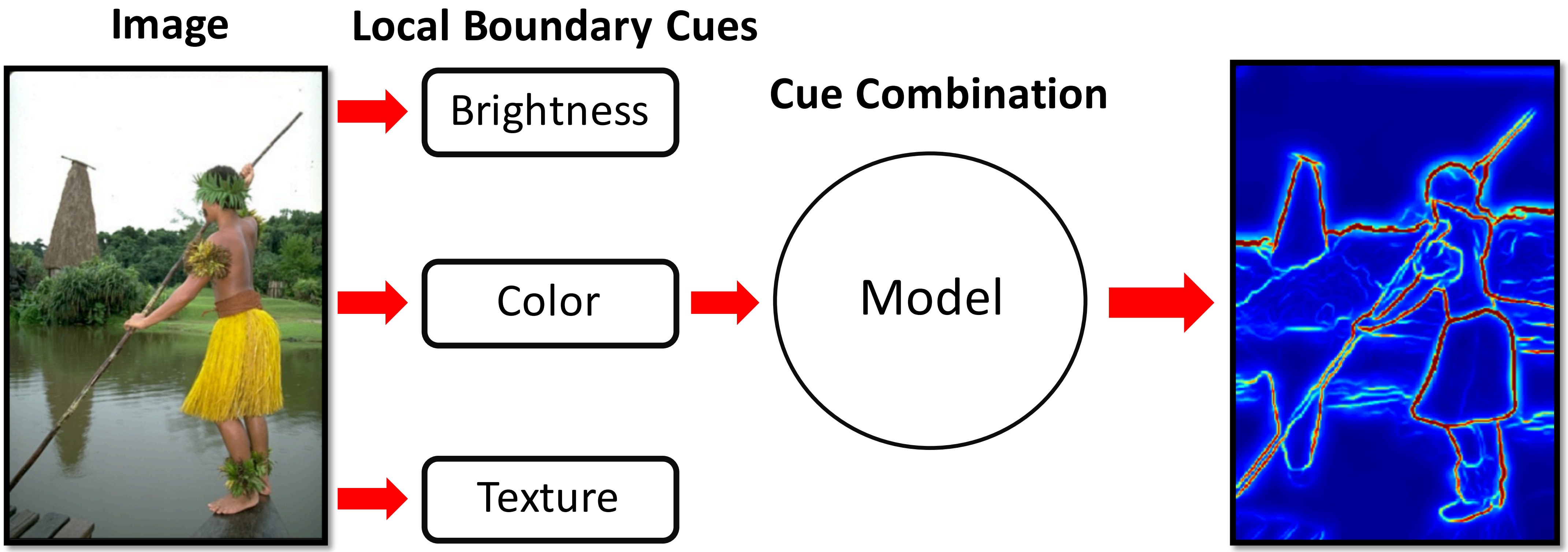
Segmentation

Contour Detection

- Learn local boundary cues
- Global framework to capture closure, continuity
- Local cues and global cues combination

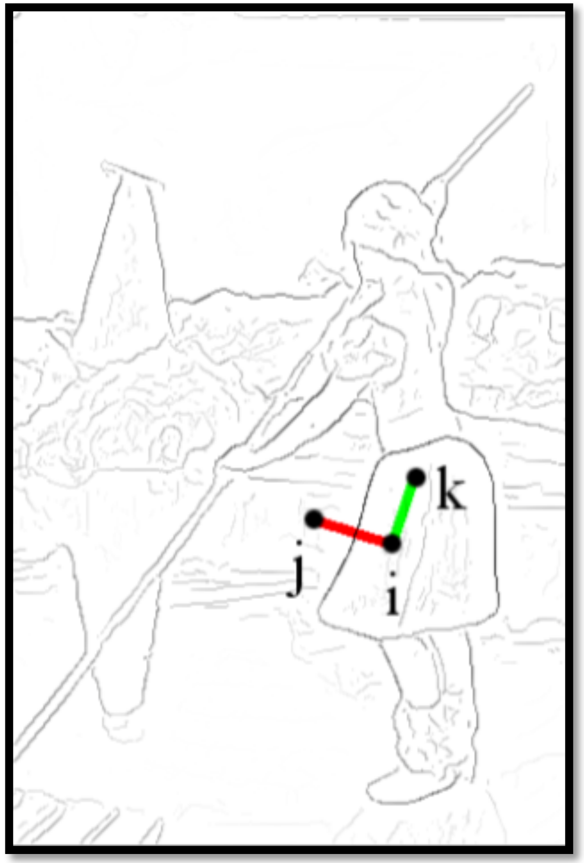


Learn Local Boundary Cues



$$mPb(x, y, \theta) = \sum_{s} \sum_{i} \alpha_{i,s} G_{i,\sigma(i,s)}(x, y, \theta)$$

Global framework



Build a weighted graph G=(V,E) from image

$$W_{ij} = \exp\left(-\max_{p \in \overline{ij}}\{mPb(p)\}/\rho\right)$$

$$sPb(x, y, \theta) = \sum_{k=1}^{n} \frac{1}{\sqrt{\lambda_k}} \cdot \nabla_\theta \mathbf{v_k}(x, y)$$

Local cues and global cues combination

$$gPb(x,y,\theta) = \sum_s \sum_i \underbrace{\beta_{i,s} G_{i,\sigma(i,s)}(x,y,\theta)}_{\textbf{Local Cues}} + \underbrace{\gamma \cdot sPb(x,y,\theta)}_{\textbf{Global Cues}}$$

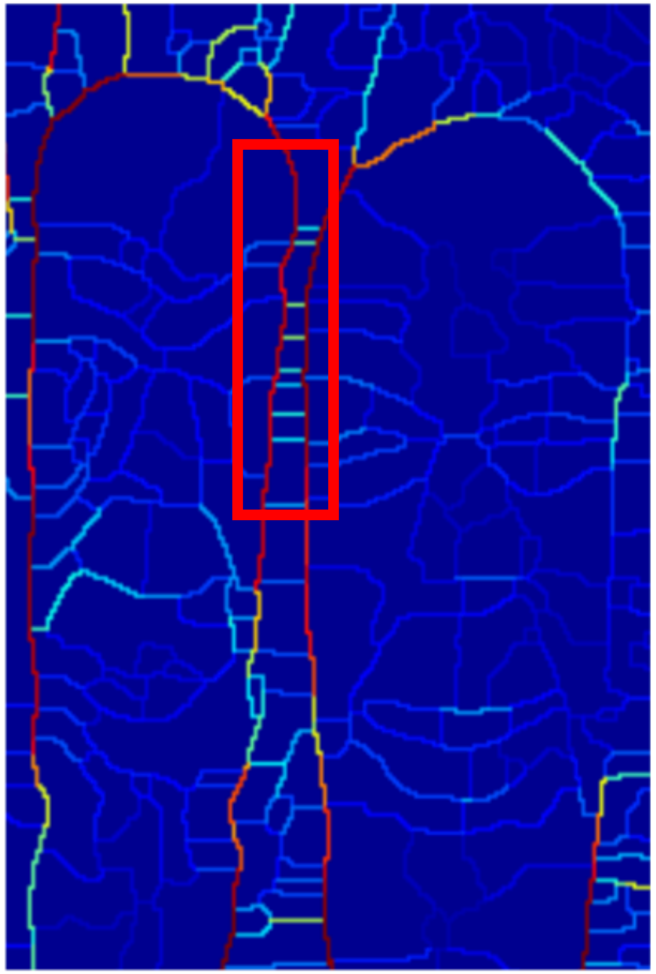
Hierarchical Segmentation

- Oriented Watershed Transform
- Ultrametric Contour Map

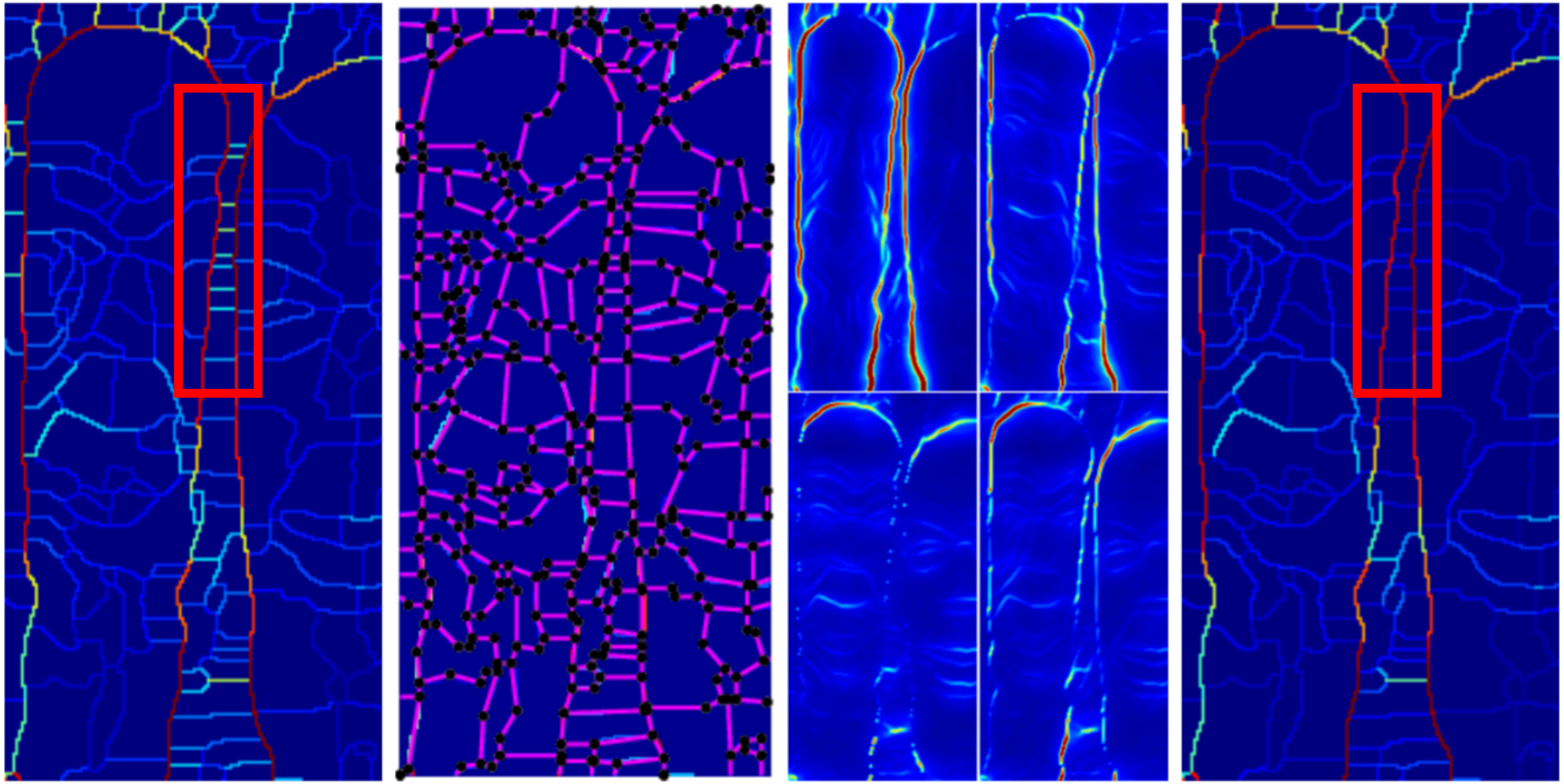
Watershed Transform

Image

Watershed

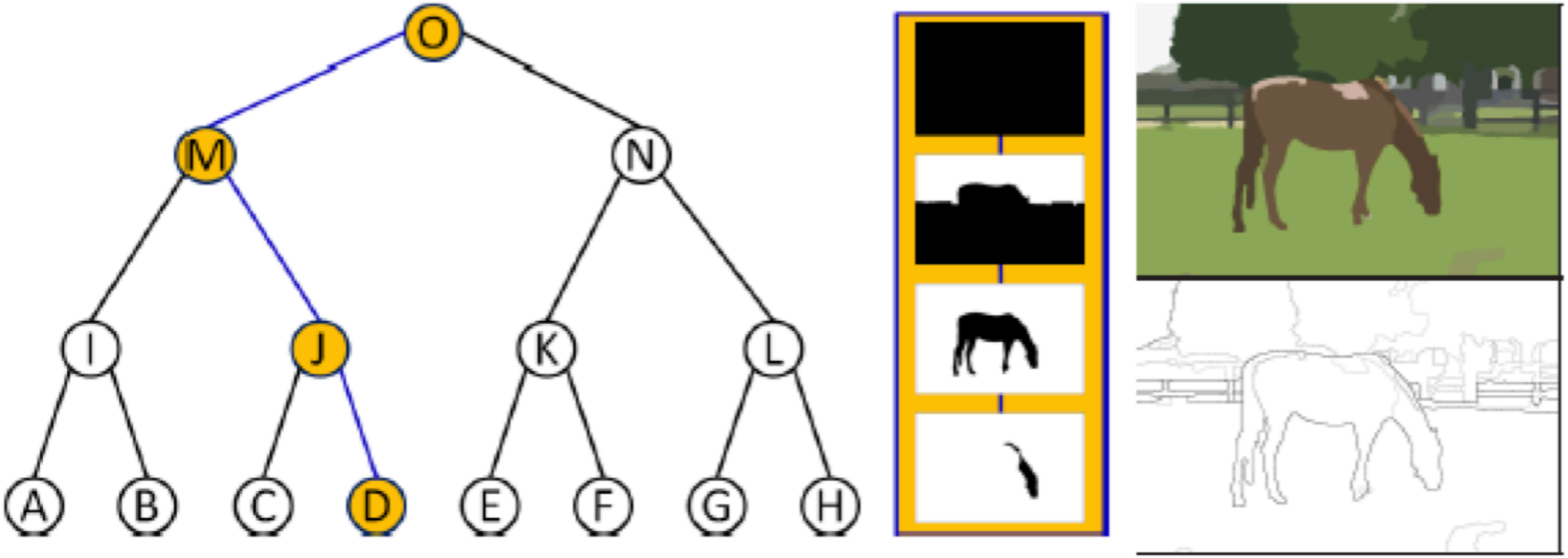


Oriented Watershed Transform

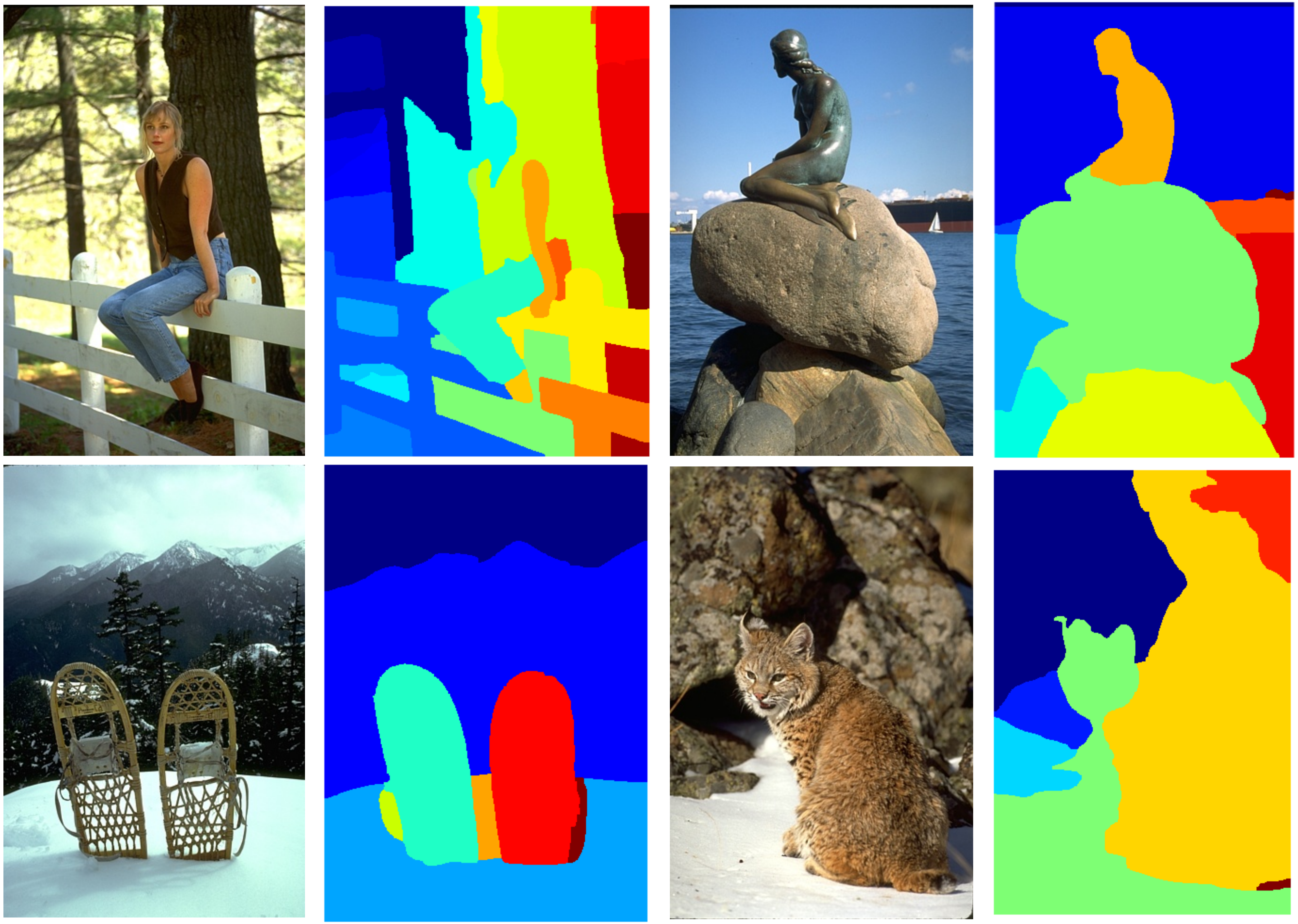


Ultrametric Contour Map

Iterative Merging



Some gPb-owt-ucm Results



Thanks