ELECTRICAL & COMPUTER
EngineeringKMUTNB

CPRE
010123210
IMAGE PROCESSING AND MACHINE VISION

Segmentation
Clustering

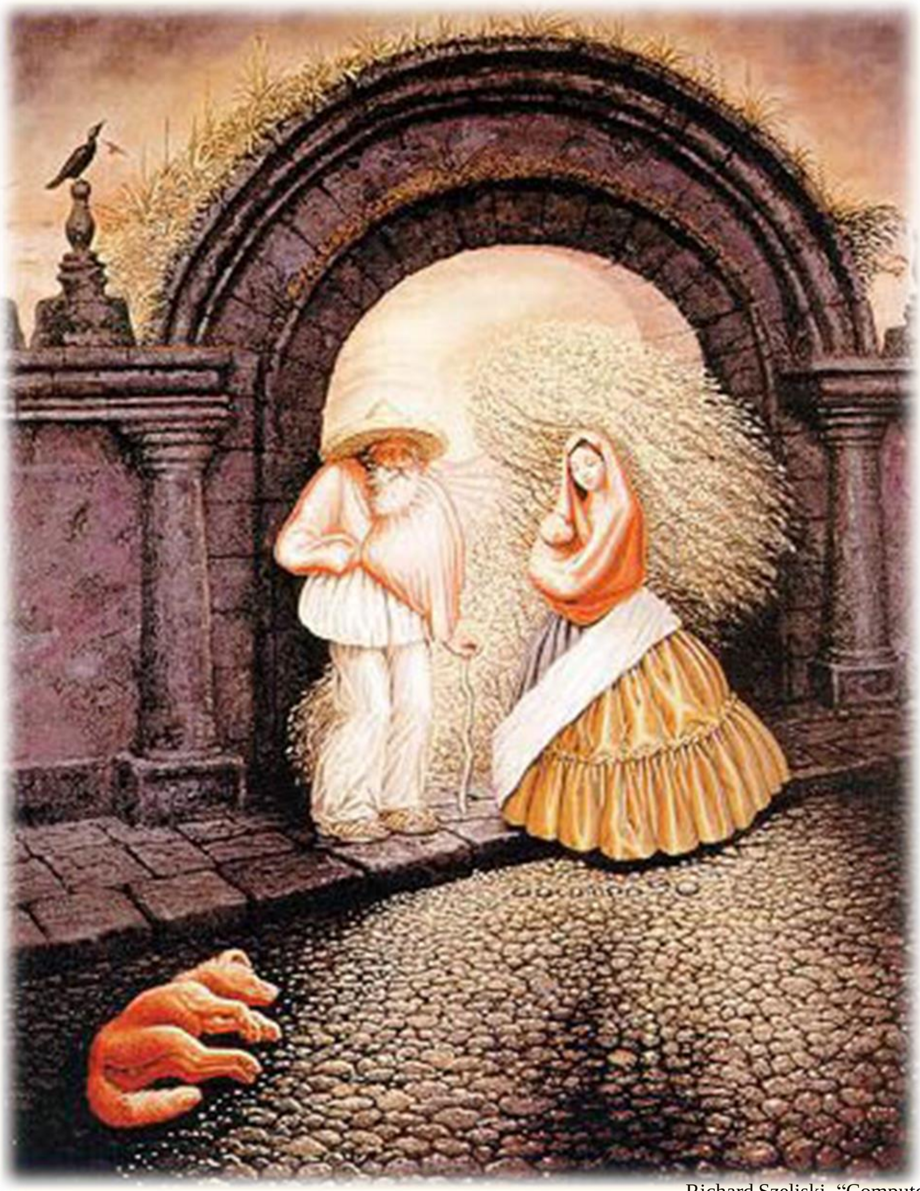
Presented By
Vera Sa-ing, Ph.D.

Section Outline

- Segmentation and Clustering
- Thresholding
- Agglomerative
- K-means
- Mean-shift
- Mean-shift Tracking

Segmentation and Clustering

Segmentation and Clustering



Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Segmentation and Clustering

Segmentation

- A task in computer vision
- Aims to **identify groups of pixels and image regions**

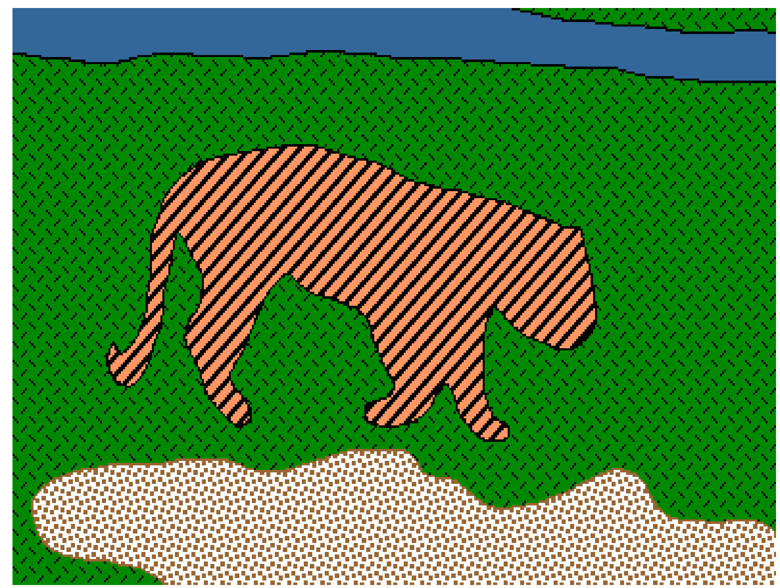
Pixels → Regions

Water

Grass

Tiger

Sand



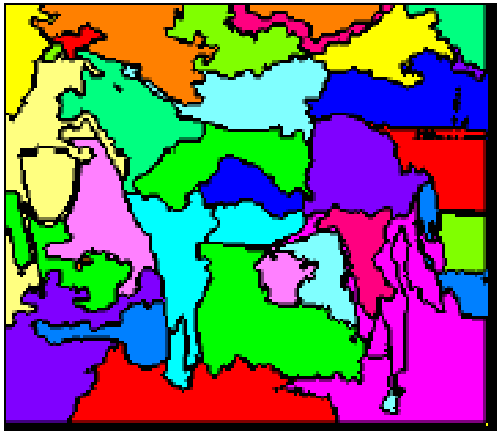
https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

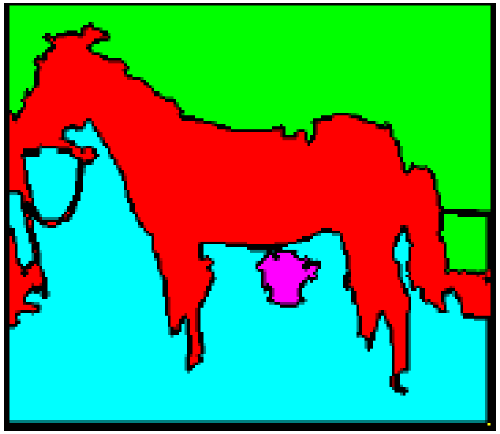
Segmentation and Clustering

Types of Segmentation

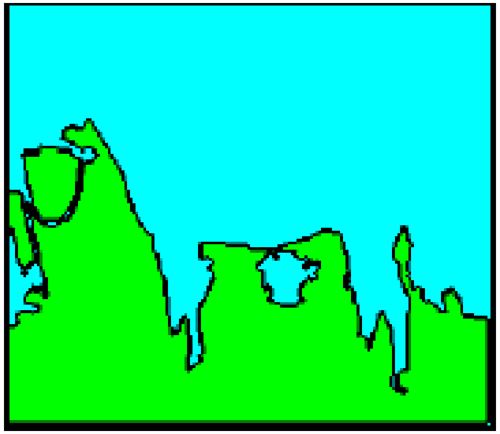
- How many separate coherent regions?



Over-Segmentation



Good-Segmentation



Under-Segmentation

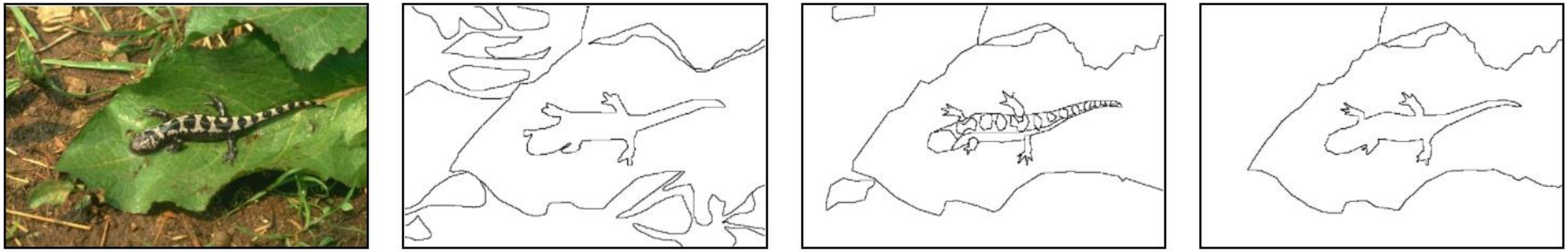
https://github.com/StanfordVL/CS131_notes

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Segmentation and Clustering

Types of Segmentation

- What is a good-segmentation?
- General idea: Compare to human segmentation
  - "**Ground Truth**"



Original Image Subject 1 Subject 2 Subject 3

- Image and ground-truth segment boundaries hand-drawn by three different human subjects

No objective definition of segmentation!

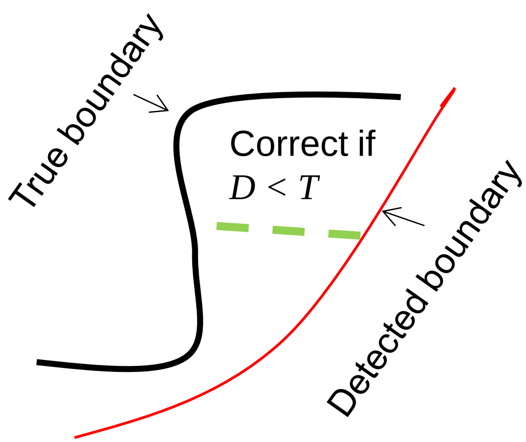
https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html
https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/papers/amfm_pami2010.pdf

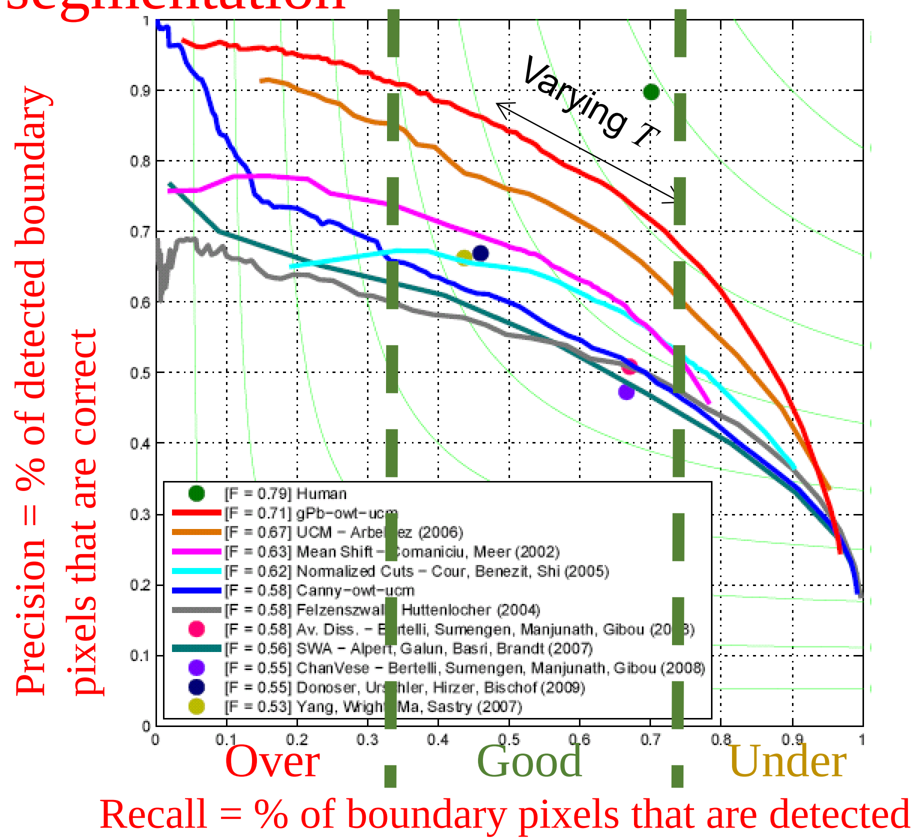
Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Segmentation and Clustering

Types of Segmentation

- What is a good-segmentation?
- General idea: Compare to human segmentation
  - "**Ground Truth**"
- General Evaluate:





https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html
https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/papers/amfm_pami2010.pdf

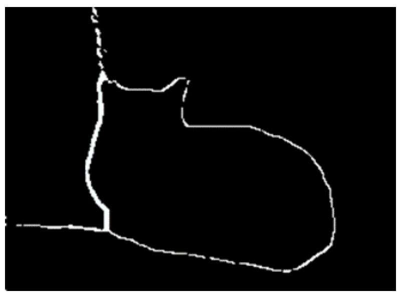
https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Segmentation and Clustering

Types of Segmentation

- What is a good-segmentation?
- General idea: Compare to human segmentation
  - "**Ground Truth**"
- General Evaluate:



$$OS(S,G) = \frac{|S \cap G|}{|S \cup G|}$$

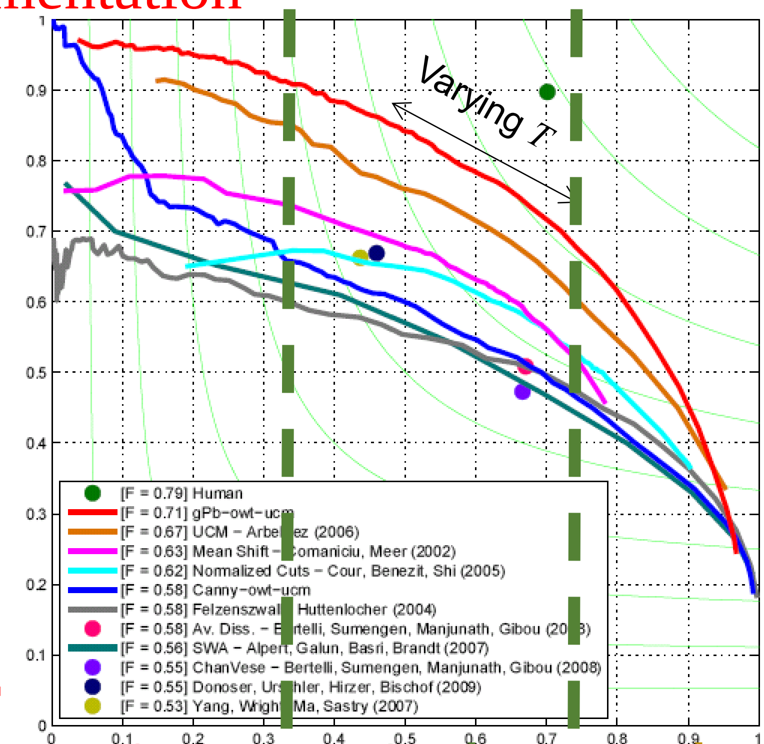
.825

Segment #1

.892

Segment #2

ion = % of detected boundary pixels that are correct



Over Good Under

Recall = % of boundary pixels that are detected

https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html
https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/papers/amfm_pami2010.pdf

https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Goals of Segmentation

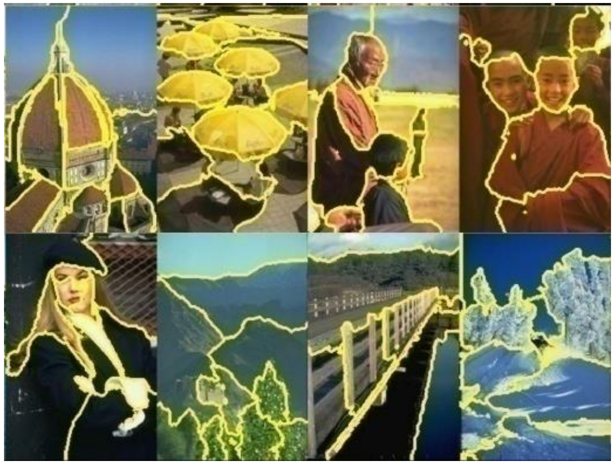
- Separate image into coherent "objects"

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Segmentation and Clustering

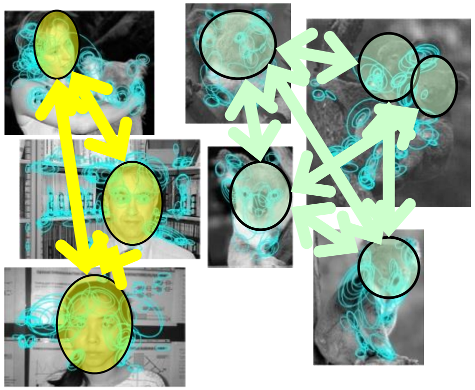
Clustering

- One way to think about "Segmentation"
- **Grouping similar data points**
- Represent them with a single token whatever needs to group
  - Such as, pixels, points, surface elements, etc.



Determining image regions

Figure-ground



Object-level grouping

What things should be grouped?

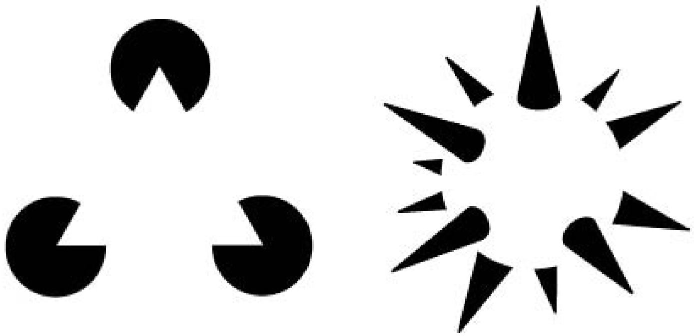
What cues indicate groups?

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

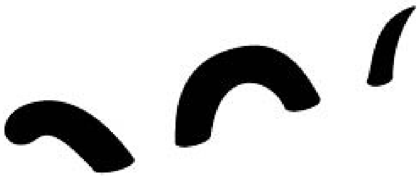
Segmentation and Clustering

Clustering from Gestalt Theory

- **Grouping is key to visual perception**
- Elements in a collection result have **relationships**
- Whole is greater than sum of its parts



Illusory contours



Familiar configuration

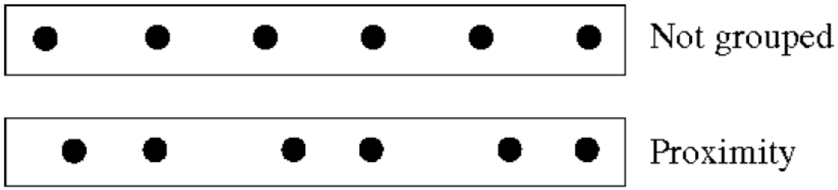
- Relationships among parts can yield new properties/features

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Segmentation and Clustering

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception

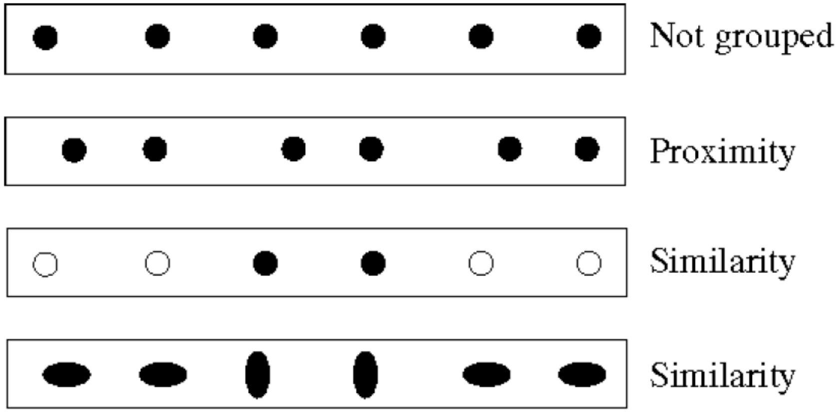


Proximity

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception



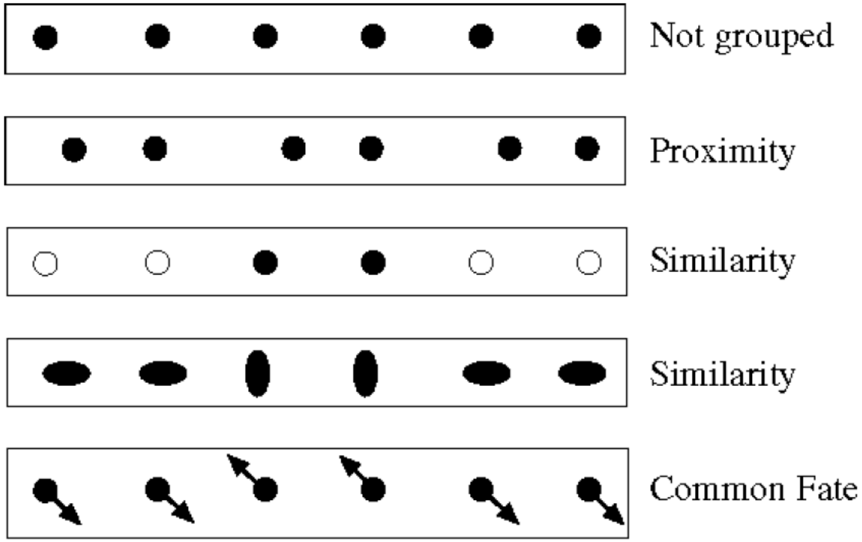
Similarity

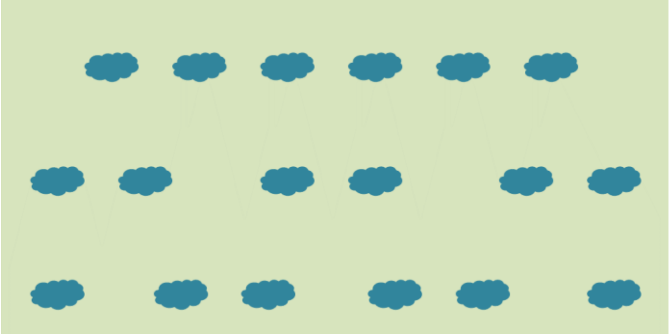
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Segmentation and Clustering

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception





Common Fate

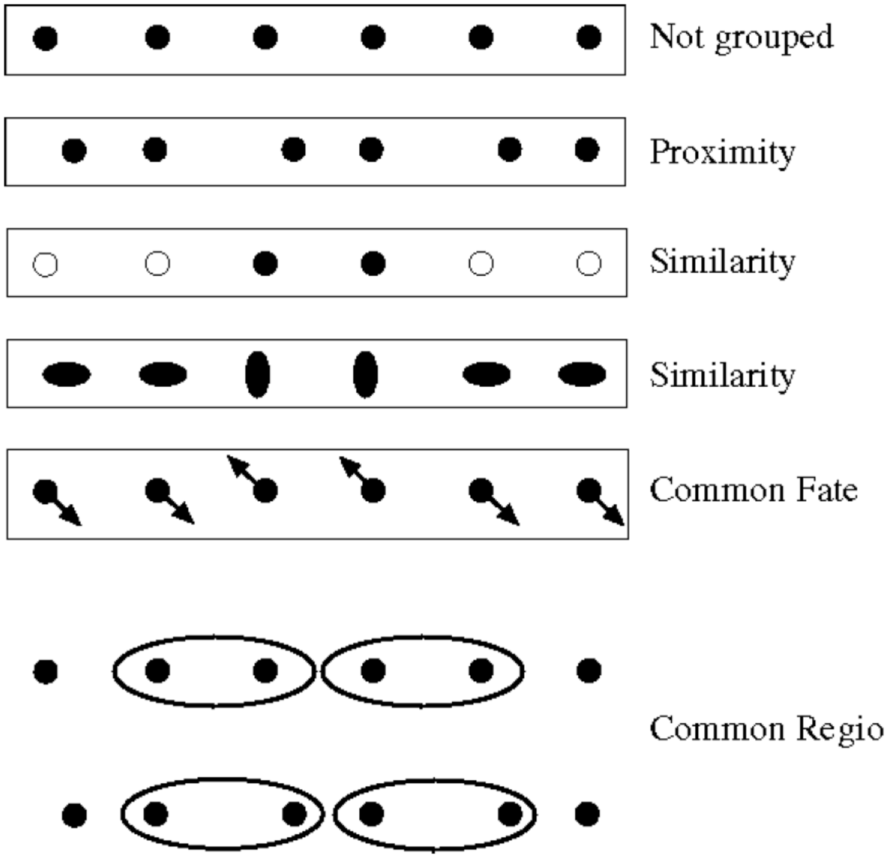
https://github.com/StanfordVL/CS131_notes

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Segmentation and Clustering

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception



Common Region

https://www.usertesting.com/blog/gestalt-principles#common

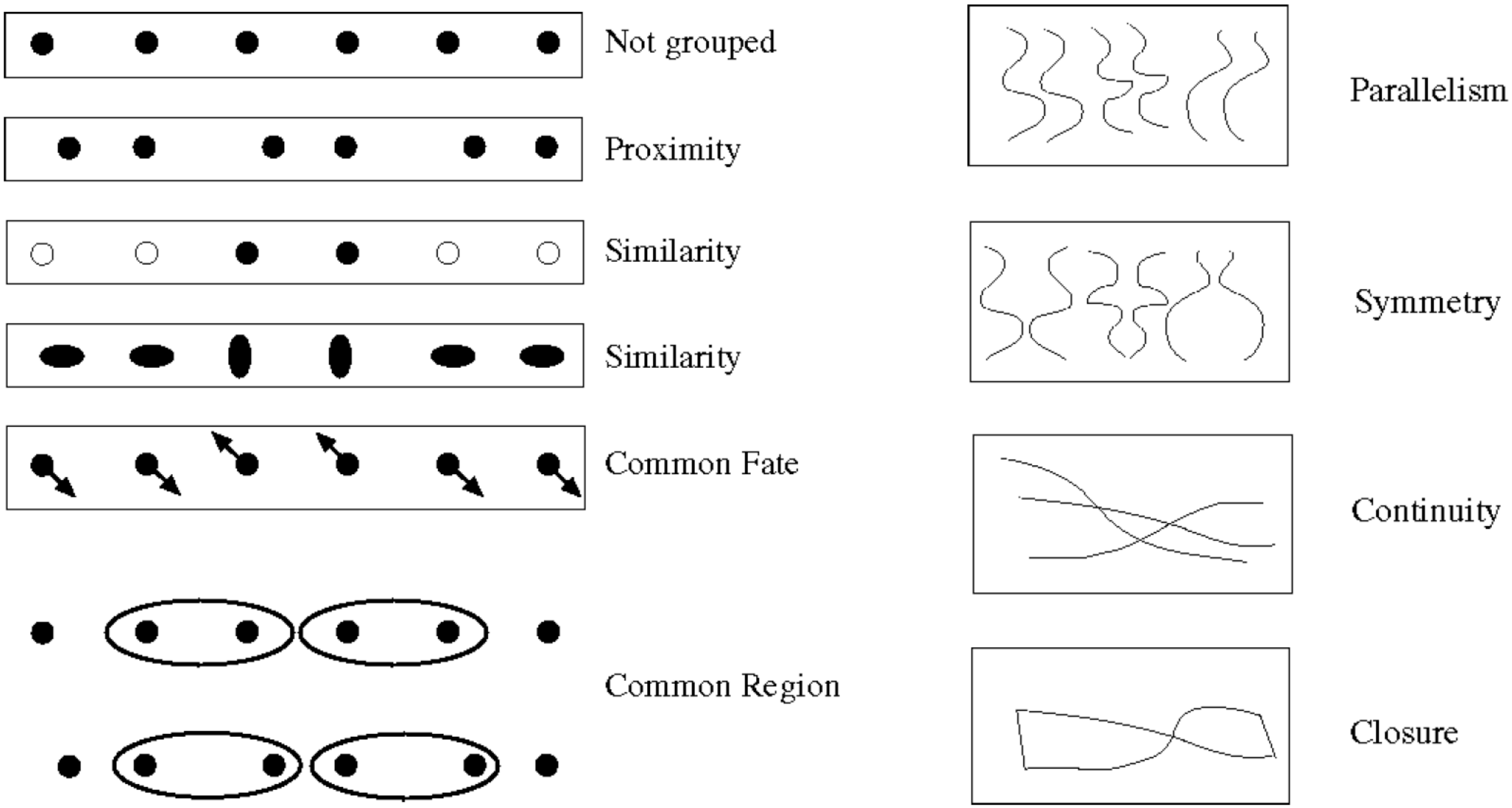
https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Segmentation and Clustering

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception



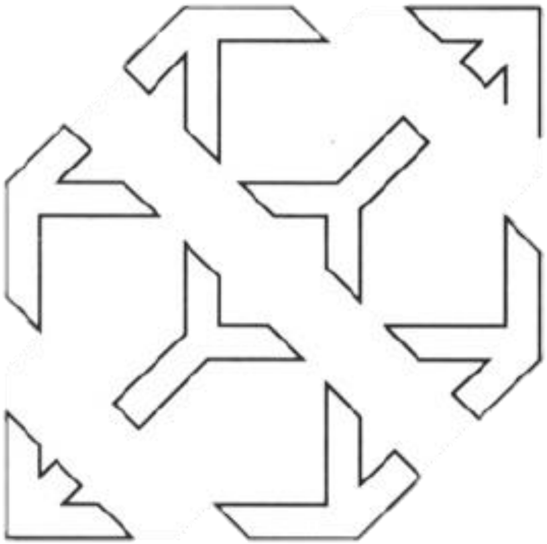
- Factors make intuitive sense, but difficult to translate into algorithms

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Segmentation and Clustering

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception



Is this one object?

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Segmentation and Clustering

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception

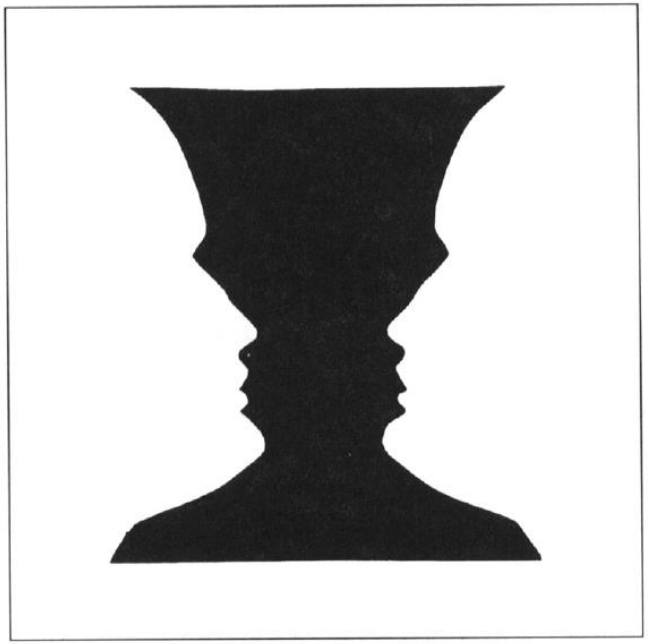
What numbers?

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Segmentation and Clustering

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception



Are two faces or one cup?

Figure and Ground Discrimination

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Clustering from Gestalt Theory

- Gestalt Factors: Grouping is key to visual perception

Did you see an animal?

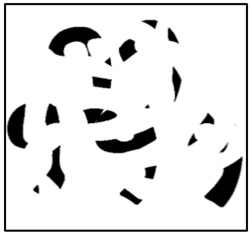
Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Segmentation and Clustering

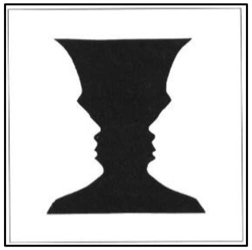
Clustering Objectives

- Look at large amounts of data
- Patch-based compression or denoising
- Represent a large continuous vector with the cluster number

- **Counting**
  - Histograms of texture, color, SIFT vectors



- **Segmentation**
  - Separate the image into different regions



- **Prediction**
  - Images in the same cluster may have the same labels

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

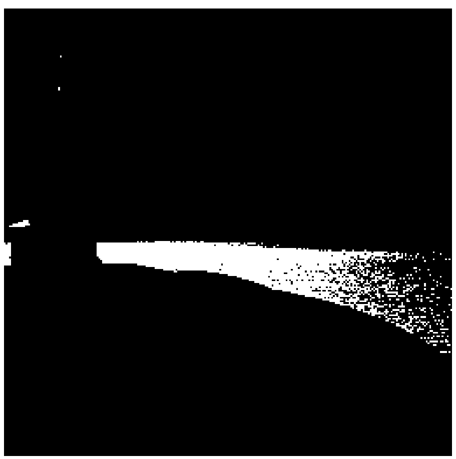
Thresholding

Thresholding

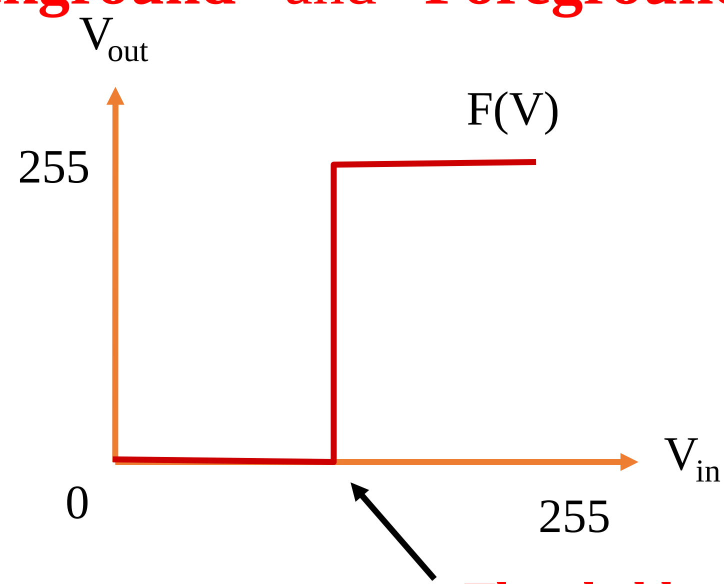
Thresholding Clustering

- Basic the segmentation algorithm to separate into two components
- Create a binary image from a grayscale image
  - Simplest way to segment objects from a background

Background and **Foreground**



Output
Selected Foreground



Threshold value
a constant gray value

Input

https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

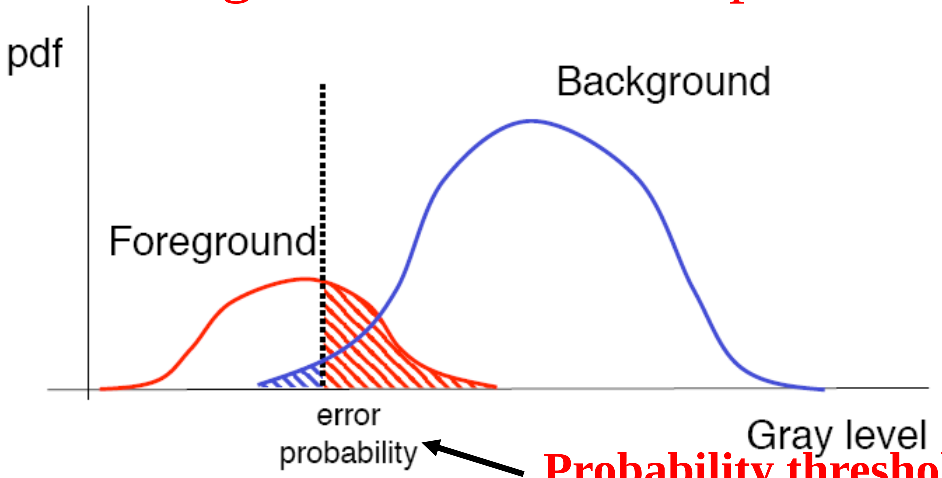
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Thresholding

Thresholding Clustering

- Basic the segmentation algorithm to separate into two components
- Create a binary image from a grayscale image
  - Simplest way to segment objects from a background

Example: Assuming there are two components in the image:



Probability threshold value T

- What is the probability of error given a threshold T?

https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

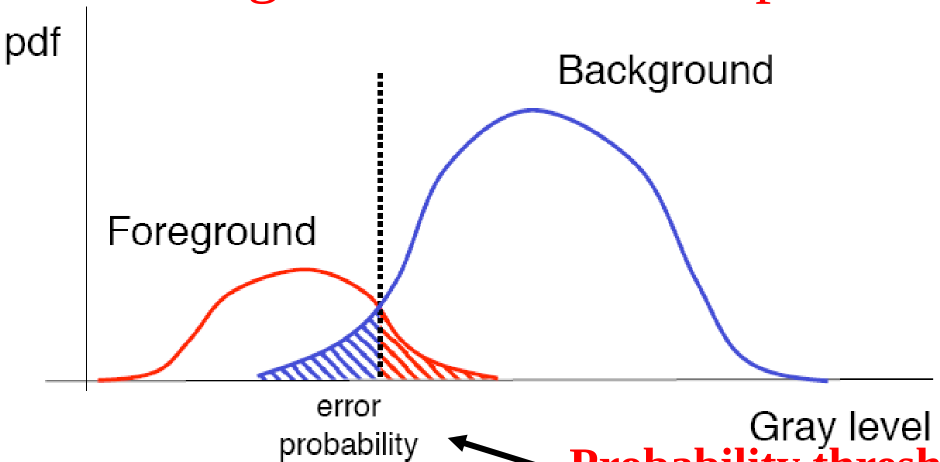
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Thresholding

Thresholding Clustering

- Basic the segmentation algorithm to separate into two components
- Create a binary image from a grayscale image
    - Simplest way to segment objects from a background

Example: Assuming there are two components in the image:



Probability threshold value T

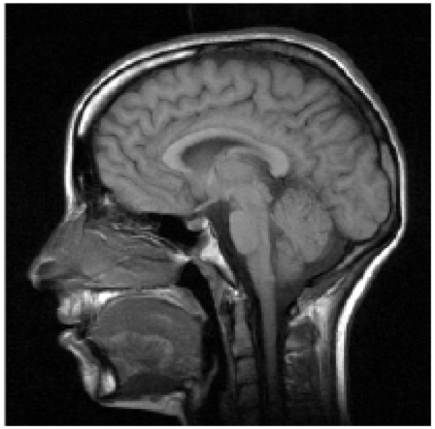
- Optimal T with respect to Maximum A Posteriori Estimation
- **Goal:** find T where $P(B \mid T) = P(F \mid T)$

Thresholding

Thresholding Clustering

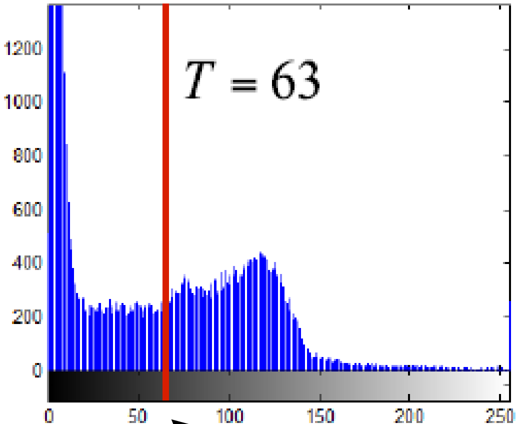
- Basic the segmentation algorithm to separate into two components
- Create a binary image from a grayscale image
  - Simplest way to segment objects from a background

Example: Assuming there are two components in the image:



Input

Histogram



Threshold value T

Output

Selected Foreground

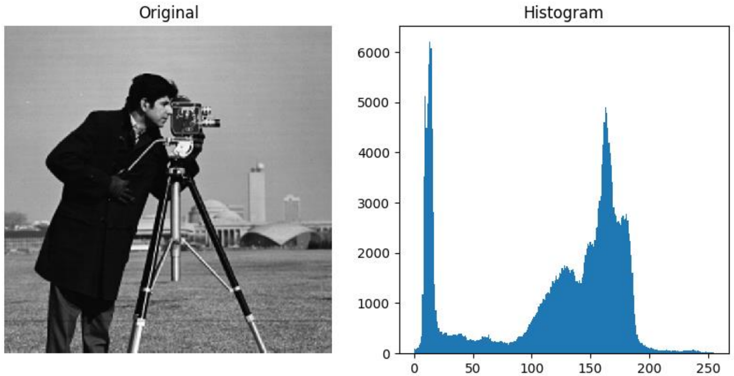
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

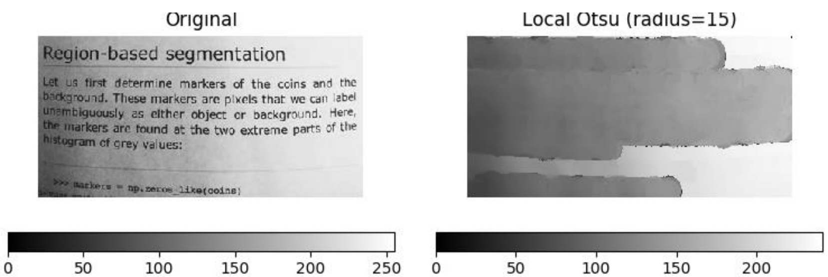
Thresholding

Thresholding Clustering

- Basic the segmentation algorithm to separate into two components
- Create a binary image from a grayscale image
  - Simplest way to segment objects from a background
- Implementation can be separated in two categories:
  - Histogram-based:



  - Local-based:



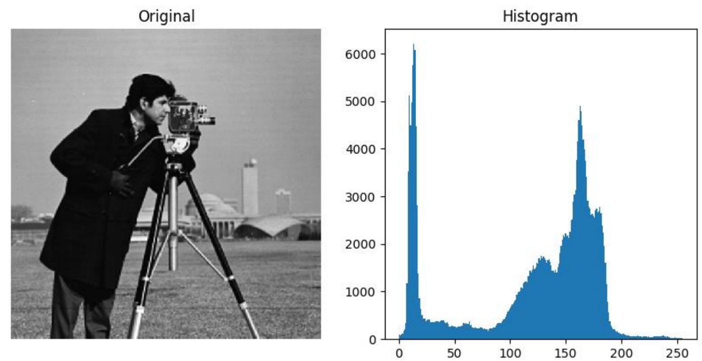
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

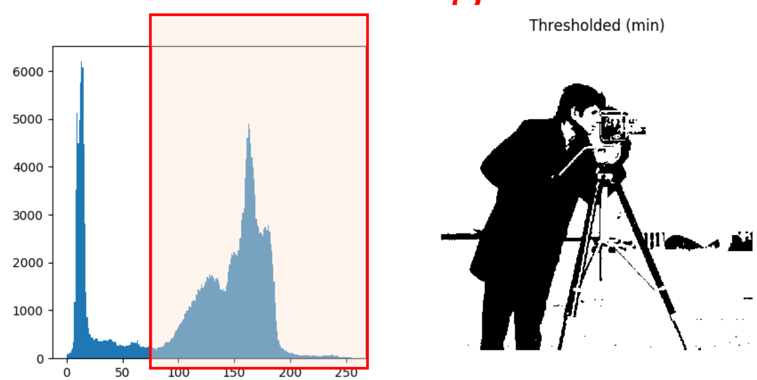
Thresholding

Thresholding Clustering

- Histogram-based:
  - Histogram of the pixels’ intensity is used



  - Certain assumptions are made on the properties of this histogram
    - Selecting the minimum histogram of the image



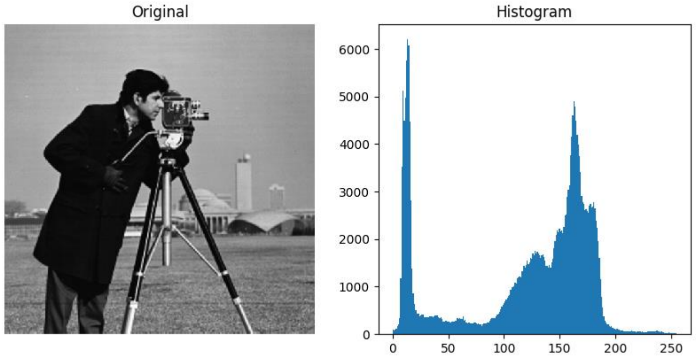
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

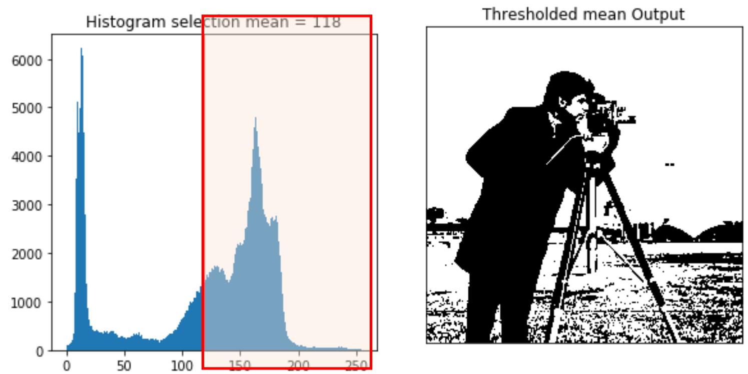
Thresholding

Thresholding Clustering

- Histogram-based:
  - Histogram of the pixels' intensity is used



  - Certain assumptions are made on the properties of this histogram
    - Selecting the mean histogram of the image



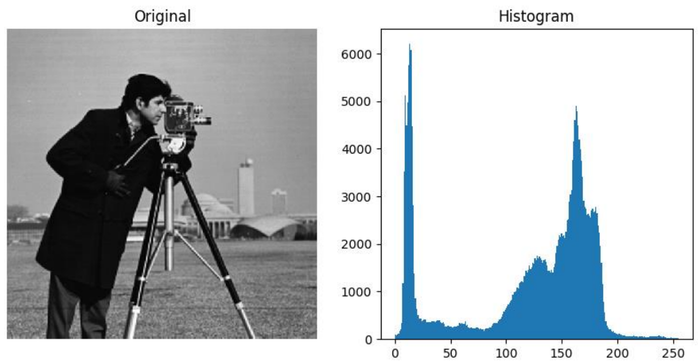
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

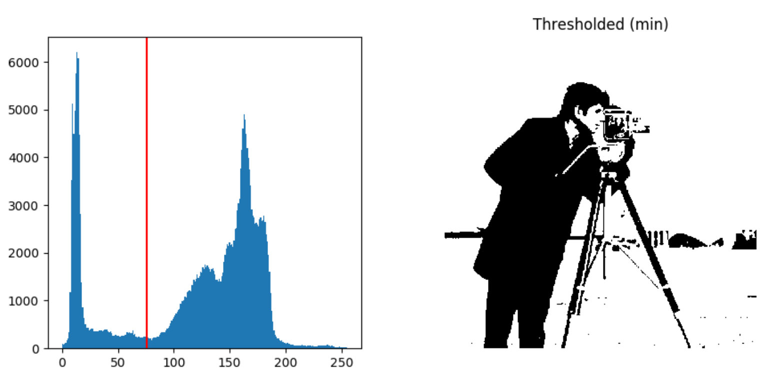
Thresholding

Thresholding Clustering

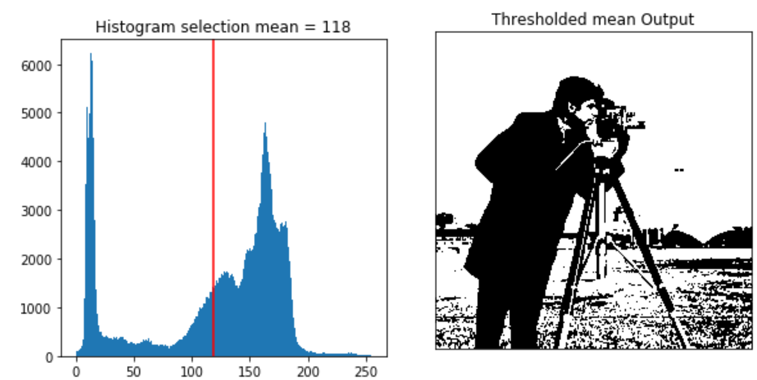
- Histogram-based:
  - Histogram of the pixels' intensity is used



Minimum Histogram



Mean Histogram



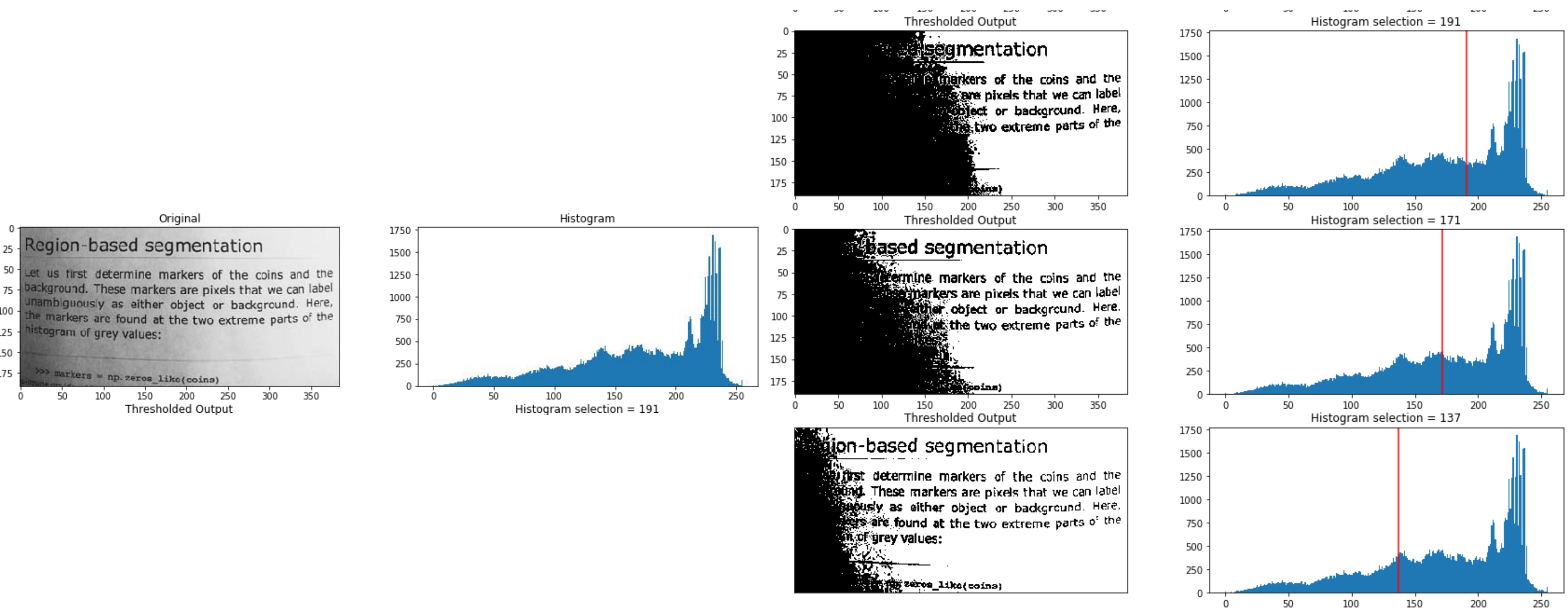
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Thresholding

Thresholding Clustering

- Histogram-based:
  - Histogram of the pixels’ intensity is used
  - Example: background subtraction



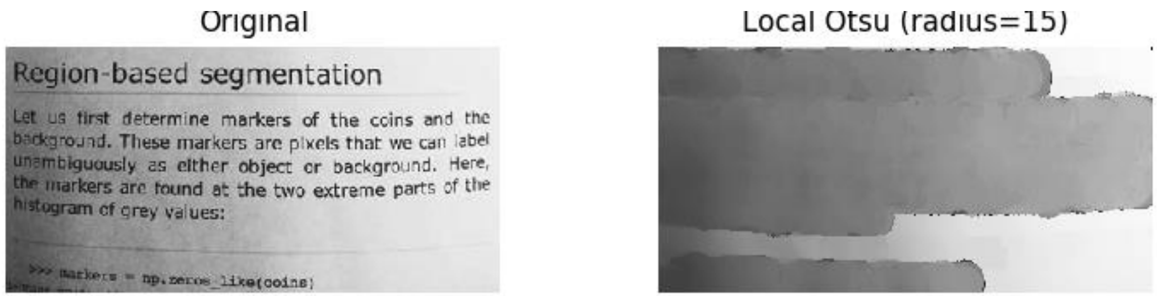
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

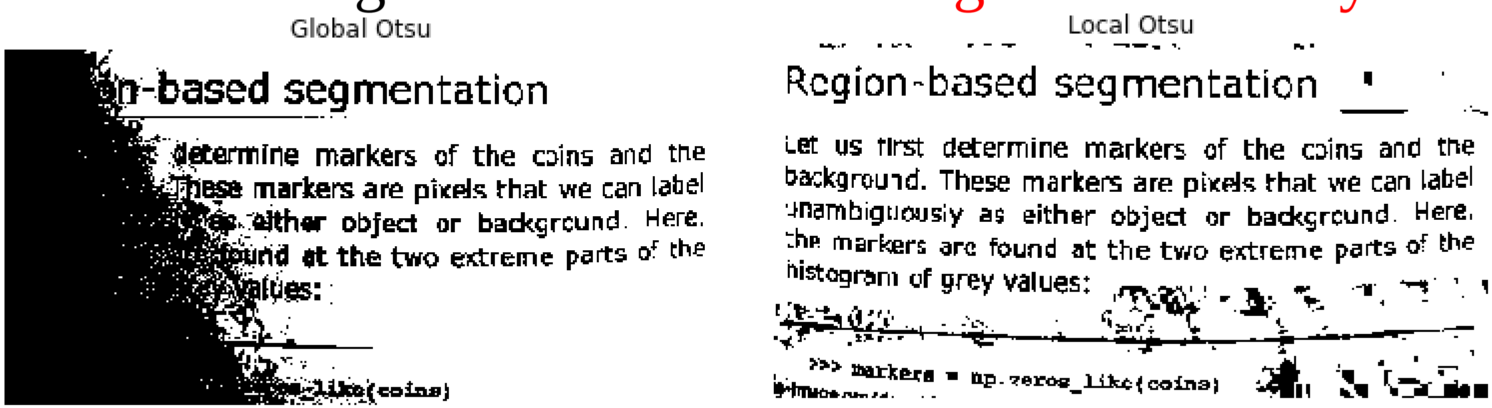
Thresholding

Thresholding Clustering

- Local-based:
  - To process a pixel, only the neighboring pixels are used



  - May produce better background subtraction results
    - If there is large variation in the background intensity



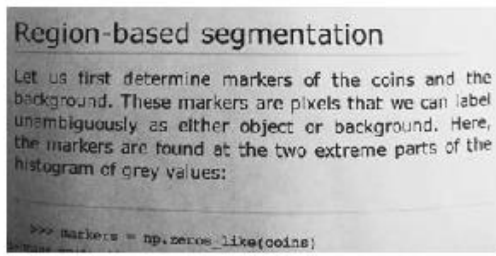
  - Often require more computation time

Thresholding

Thresholding Clustering

- Local-based:
  - Otsu’s threshold method can be applied locally
  - Calculate thresholds in regions with a characteristic size block_size surrounding each pixel (i.e. local neighborhoods)

Original

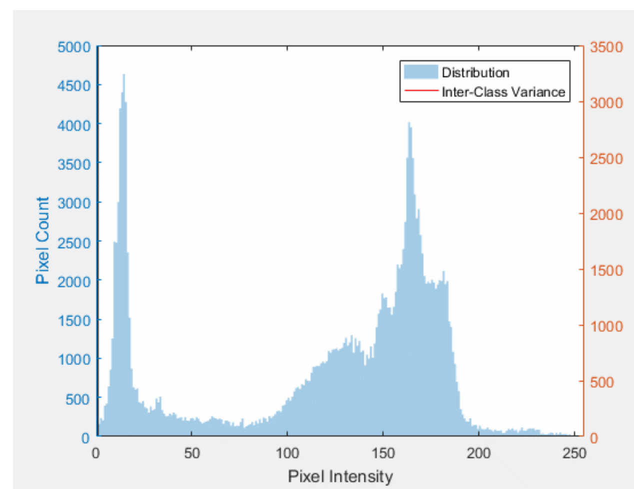

Region-based segmentation

Let us first determine markers of the coins and the background. These markers are pixels that we can label unambiguously as either object or background. Here, the markers are found at the two extreme parts of the histogram of grey values:

```
>>> markers = np.zeros_like(coins)
```

Local Otsu (radius=15)

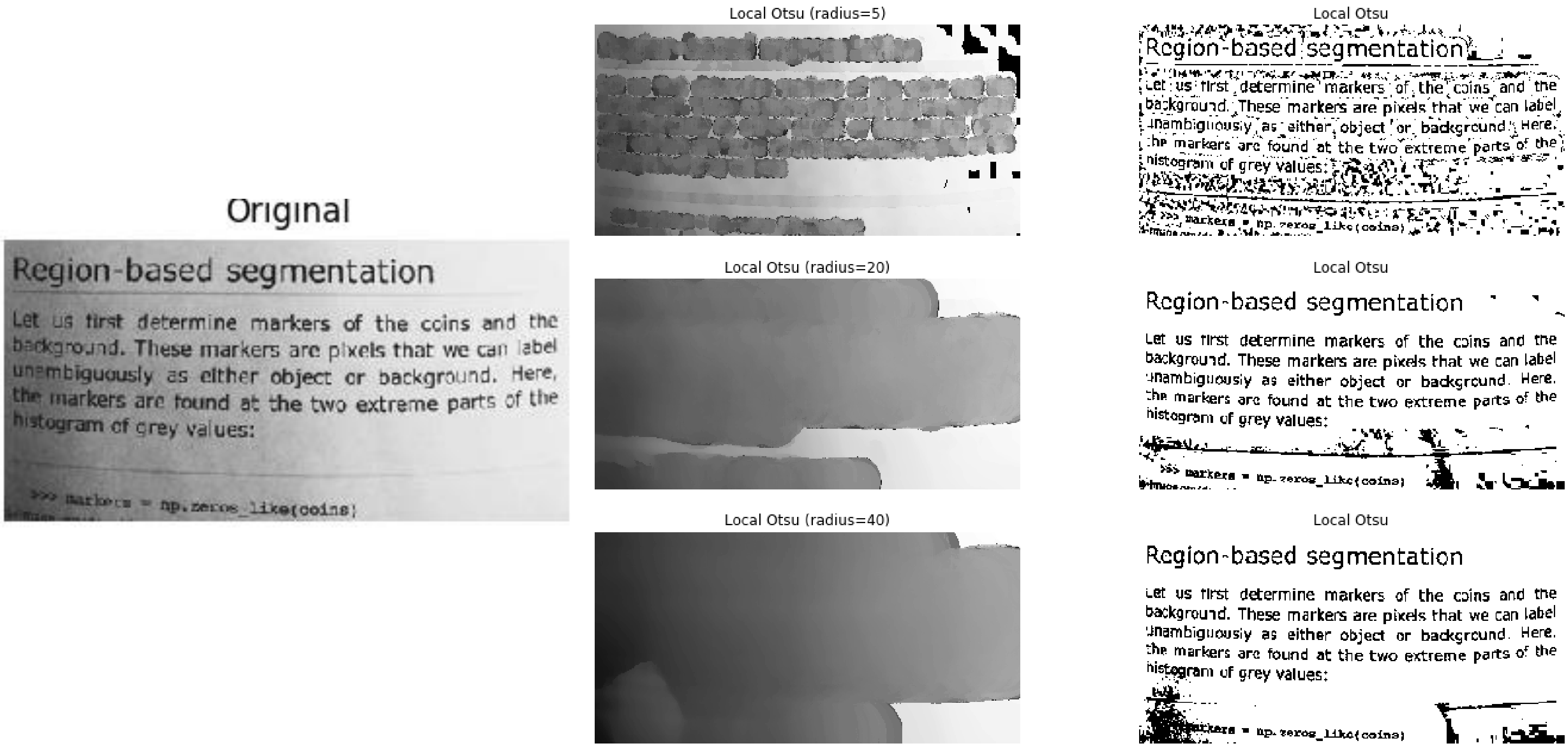
  - Determine the variance between two classes of pixels of the local neighborhood defined by a structuring element



Thresholding

Thresholding Clustering

- Local-based:
  - Otsu’s threshold method can be applied locally



https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Exercise

Thresholding

Thresholding Clustering Algorithm

Pros

- General, application-independent tool
- Model-free, does not assume any prior shape on data clusters
- Just a single parameter (threshold value)
- Fast computation

Cons

- Output depends on threshold value
- Does not Robust to outliers

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Agglomerative

Agglomerative

What is Similarity?

- Hard to define but… "We know it when we see it"
- The real meaning of similarity is a philosophical question
- We will take a more pragmatic approach

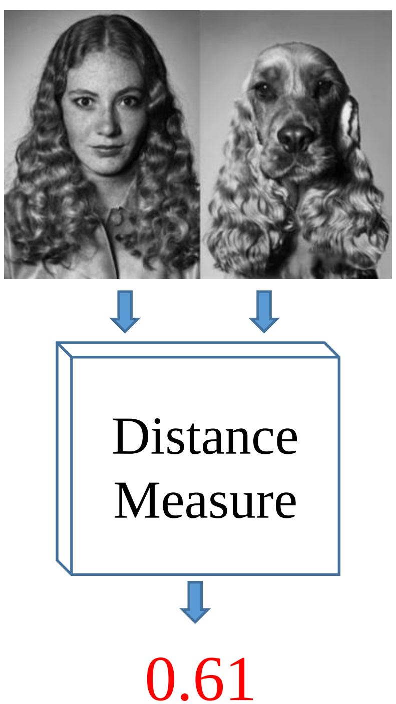
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

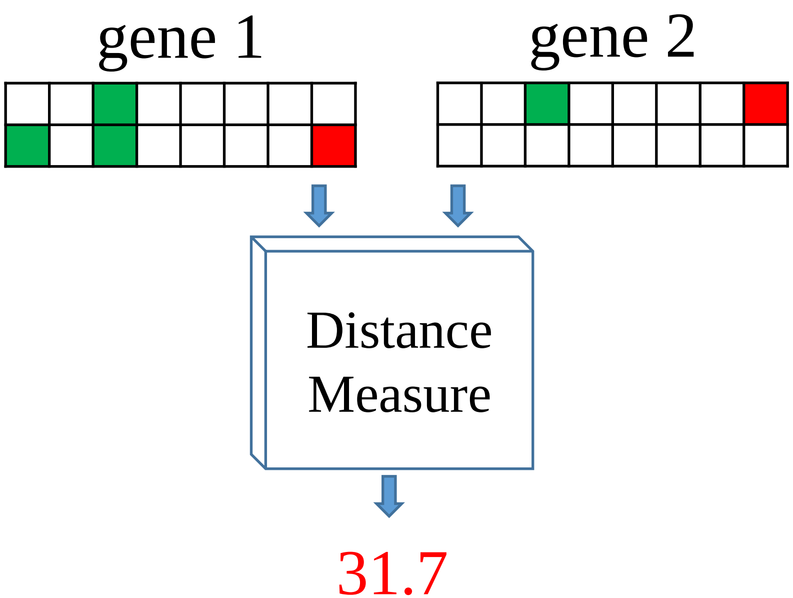
Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Agglomerative

Clustering defines distance measures

- Clustering is an unsupervised learning method
- Definition:
  - Let O_1 and O_2 be two objects from the universe of possible objects
  - The distance (dissimilarity) between O_1 and O_2 is a real number denoted by $\boldsymbol{D(O_1, O_2)}$





https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Agglomerative

Clustering defines distance measures

- Clustering is an unsupervised learning method
- Definition:
  - Let O_1 and O_2 be two objects from the universe of possible objects
  - The distance (dissimilarity) between O_1 and O_2 is a real number denoted by $\boldsymbol{D(O_1, O_2)}$

- The Euclidian distance:

$$\mathrm{D}(O_1, O_2) = {O_1}^T O_2 = \sqrt{\sum_i (O_{1,i} + O_{2,i})^2}$$

- The Cosine distance measure:

$$\mathrm{sim}(O_1, O_2) = \cos(\theta) = \frac{{O_1}^T O_2}{\sqrt{{O_1}^T O_1}\sqrt{{O_2}^T O_2}}$$

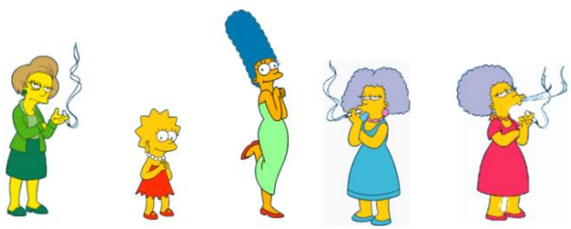
https://github.com/StanfordVL/CS131_notes and http://www.cs.cmu.edu/~tom/10601_fall2012/lectures.shtml

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Agglomerative

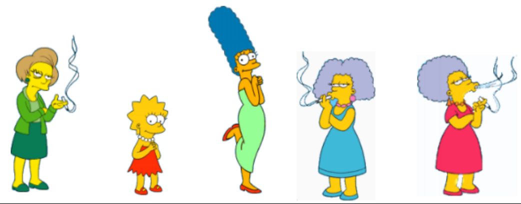
Hierarchical Clustering Algorithm

- Create a hierarchical decomposition of objects using some criterion



D() = 8

D() = 1



0	8	8	7	7
	0	2	4	4
		0	3	3
			0	1
				0

- Matrix represents distance between two elements
- Distance is usually assumed to be the inverse of similarity
- Low distance means high similarity

https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

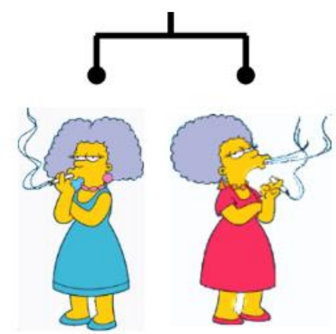
Agglomerative

Hierarchical Clustering Algorithm

- Bottom-Up (agglomerative):
  - Starting with each item in its own cluster
  - Find the best pair to merge into a new cluster
  - Repeat until all clusters are fused together

0	8	8	7	7
	0	2	4	4
		0	3	3
			0	1
				0

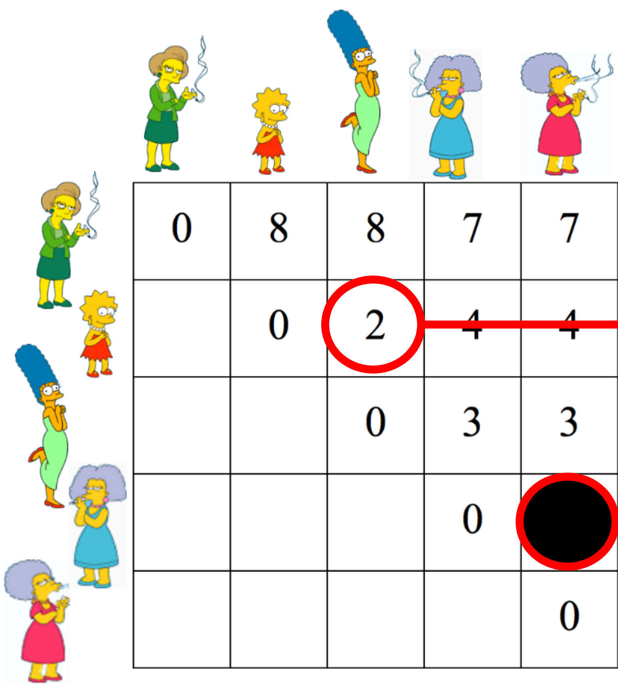
Choose
the best pair



Agglomerative

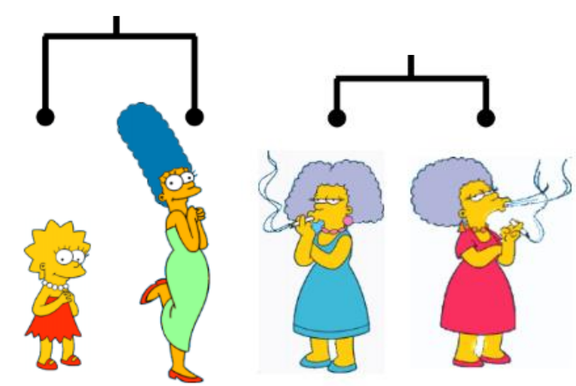
Hierarchical Clustering Algorithm

- Bottom-Up (agglomerative):
  - Starting with each item in its own cluster
  - Find the best pair to merge into a new cluster
  - Repeat until all clusters are fused together



0	8	8	7	7
	0	2	4	4
		0	3	3
			0	
				0

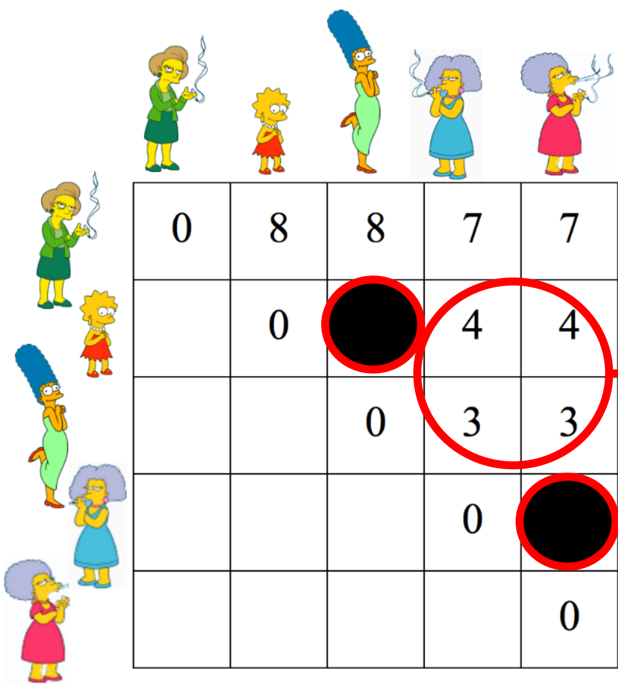
Choose
the best pair



Agglomerative

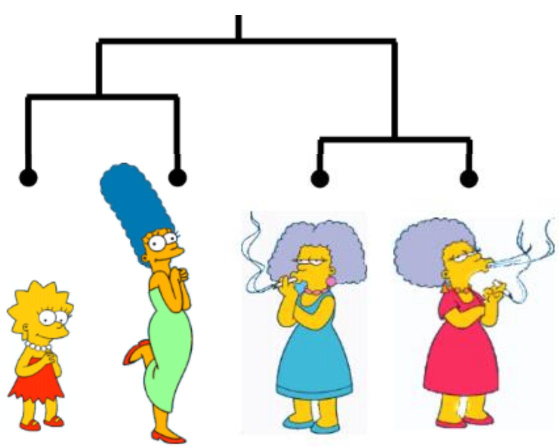
Hierarchical Clustering Algorithm

- Bottom-Up (agglomerative):
  - Starting with each item in its own cluster
  - Find the best pair to merge into a new cluster
  - Repeat until all clusters are fused together



0	8	8	7	7
	0		4	4
		0	3	3
			0	
				0

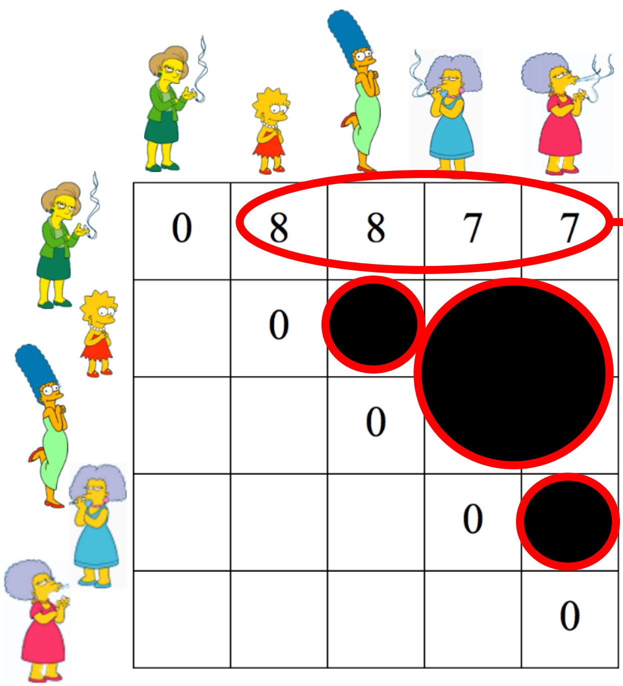
Choose
the best pair



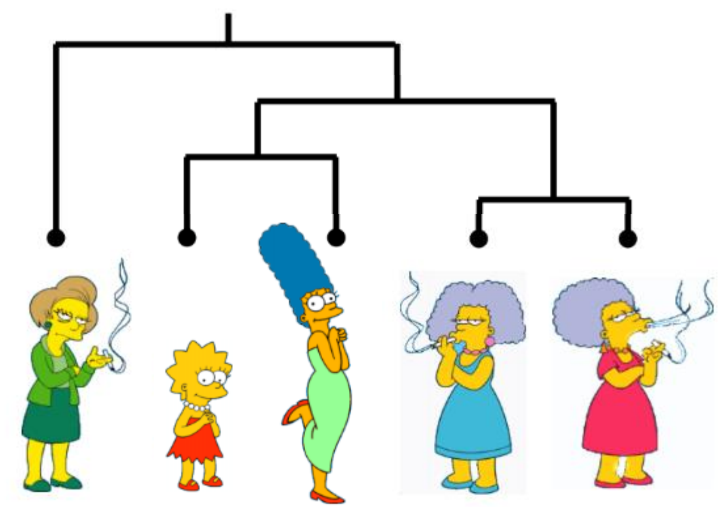
Agglomerative

Hierarchical Clustering Algorithm

- Bottom-Up (agglomerative):
  - Starting with each item in its own cluster
  - Find the best pair to merge into a new cluster
  - Repeat until all clusters are fused together



Choose the best pair



https://github.com/StanfordVL/CS131_notes

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

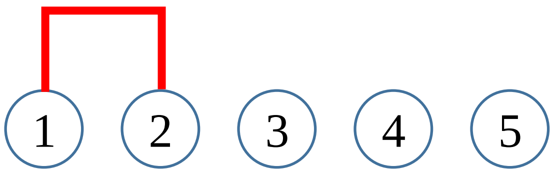
Agglomerative

Agglomerative Clustering Algorithm

- Cluster distance: **Single Link**
  - Distance of two closest members in each class

$$\begin{array}{cc} & \begin{array}{ccccc}1&2&3&4&5\end{array} \\ \begin{array}{c}1\\2\\3\\4\\5\end{array} & \begin{bmatrix}0&&&&\\2&0&&&\\6&3&0&&\\10&9&7&0&\\9&8&5&4&0\end{bmatrix}\end{array}$$

Choose
the best pair



https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Agglomerative

Agglomerative Clustering Algorithm

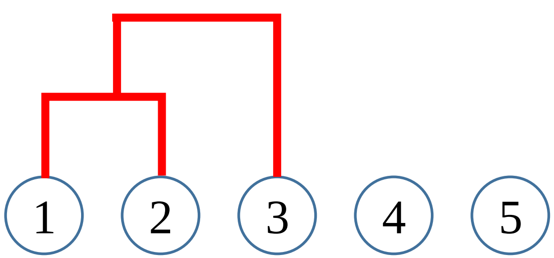
- Cluster distance: Single Link
  - Distance of two closest members in each class

$$
\begin{array}{c|ccccc}
 & 1 & 2 & 3 & 4 & 5 \\
\hline
1 & 0 & & & & \\
2 & 2 & 0 & & & \\
3 & 6 & 3 & 0 & & \\
4 & 10 & 9 & 7 & 0 & \\
5 & 9 & 8 & 5 & 4 & 0
\end{array}
\Rightarrow
\begin{array}{c|cccc}
 & (1,2) & 3 & 4 & 5 \\
\hline
(1,2) & 0 & & & \\
3 & 3 & 0 & & \\
4 & 9 & 7 & 0 & \\
5 & 8 & 5 & 4 & 0
\end{array}
$$

$$d_{(1,2),3} = \min\{d_{1,3}, d_{2,3}\} = \min\{6,3\} = 3$$

$$d_{(1,2),4} = \min\{d_{1,4}, d_{2,4}\} = \min\{10,9\} = 9$$

$$d_{(1,2),5} = \min\{d_{1,5}, d_{2,5}\} = \min\{9,8\} = 8$$



https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Agglomerative

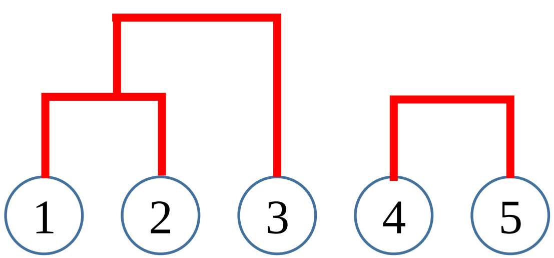
Agglomerative Clustering Algorithm

- Cluster distance: Single Link
  - Distance of two closest members in each class

$$
\begin{array}{c|ccccc}
 & 1 & 2 & 3 & 4 & 5 \\
\hline
1 & 0 & & & & \\
2 & 2 & 0 & & & \\
3 & 6 & 3 & 0 & & \\
4 & 10 & 9 & 7 & 0 & \\
5 & 9 & 8 & 5 & 4 & 0
\end{array}
\Rightarrow
\begin{array}{c|cccc}
 & (1,2) & 3 & 4 & 5 \\
\hline
(1,2) & 0 & & & \\
3 & 3 & 0 & & \\
4 & 9 & 7 & 0 & \\
5 & 8 & 5 & 4 & 0
\end{array}
\Rightarrow
\begin{array}{c|ccc}
 & (1,2,3) & 4 & 5 \\
\hline
(1,2,3) & 0 & & \\
4 & 7 & 0 & \\
5 & 5 & 4 & 0
\end{array}
$$

$$d_{(1,2,3),4} = \min\{d_{(1,2),4}, d_{3,4}\} = \min\{9,7\} = 7$$

$$d_{(1,2,3),5} = \min\{d_{(1,2),5}, d_{3,5}\} = \min\{8,5\} = 5$$



https://github.com/StanfordVL/CS131_notes

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

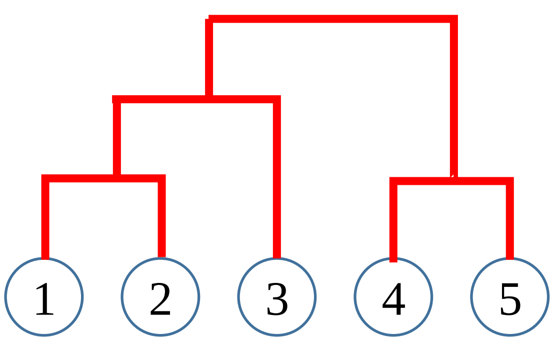
Agglomerative

Agglomerative Clustering Algorithm

- Cluster distance: Single Link
  - Distance of two closest members in each class

$$
\begin{array}{c|ccccc}
 & 1 & 2 & 3 & 4 & 5 \\
\hline
1 & 0 & & & & \\
2 & 2 & 0 & & & \\
3 & 6 & 3 & 0 & & \\
4 & 10 & 9 & 7 & 0 & \\
5 & 9 & 8 & 5 & 4 & 0
\end{array}
\Rightarrow
\begin{array}{c|cccc}
 & (1,2) & 3 & 4 & 5 \\
\hline
(1,2) & 0 & & & \\
3 & 3 & 0 & & \\
4 & 9 & 7 & 0 & \\
5 & 8 & 5 & 4 & 0
\end{array}
\Rightarrow
\begin{array}{c|ccc}
 & (1,2,3) & 4 & 5 \\
\hline
(1,2,3) & 0 & & \\
4 & 7 & 0 & \\
5 & 5 & 4 & 0
\end{array}
$$

$$
d_{(1,2,3),(4,5)} = \min\{d_{(1,2,3),4}, d_{(1,2,3),5}\} = 5
$$



https://github.com/StanfordVL/CS131_notes

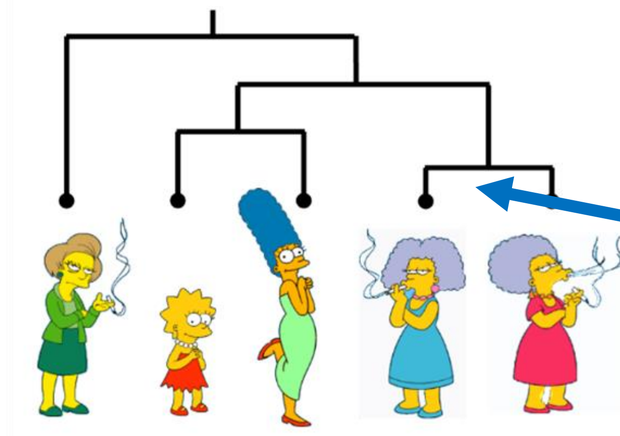
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Agglomerative

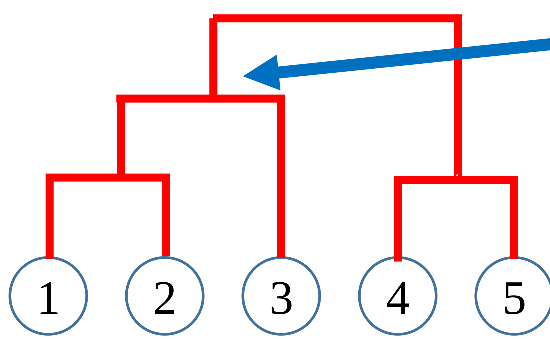
Agglomerative Clustering Algorithm

0	8	8	7	7
	0	2	4	4
		0	3	3
			0	1
				0

$$
\begin{array}{c|ccccc}
 & 1 & 2 & 3 & 4 & 5 \\
\hline
1 & 0 & & & & \\
2 & 2 & 0 & & & \\
3 & 6 & 3 & 0 & & \\
4 & 10 & 9 & 7 & 0 & \\
5 & 9 & 8 & 5 & 4 & 0
\end{array}
$$

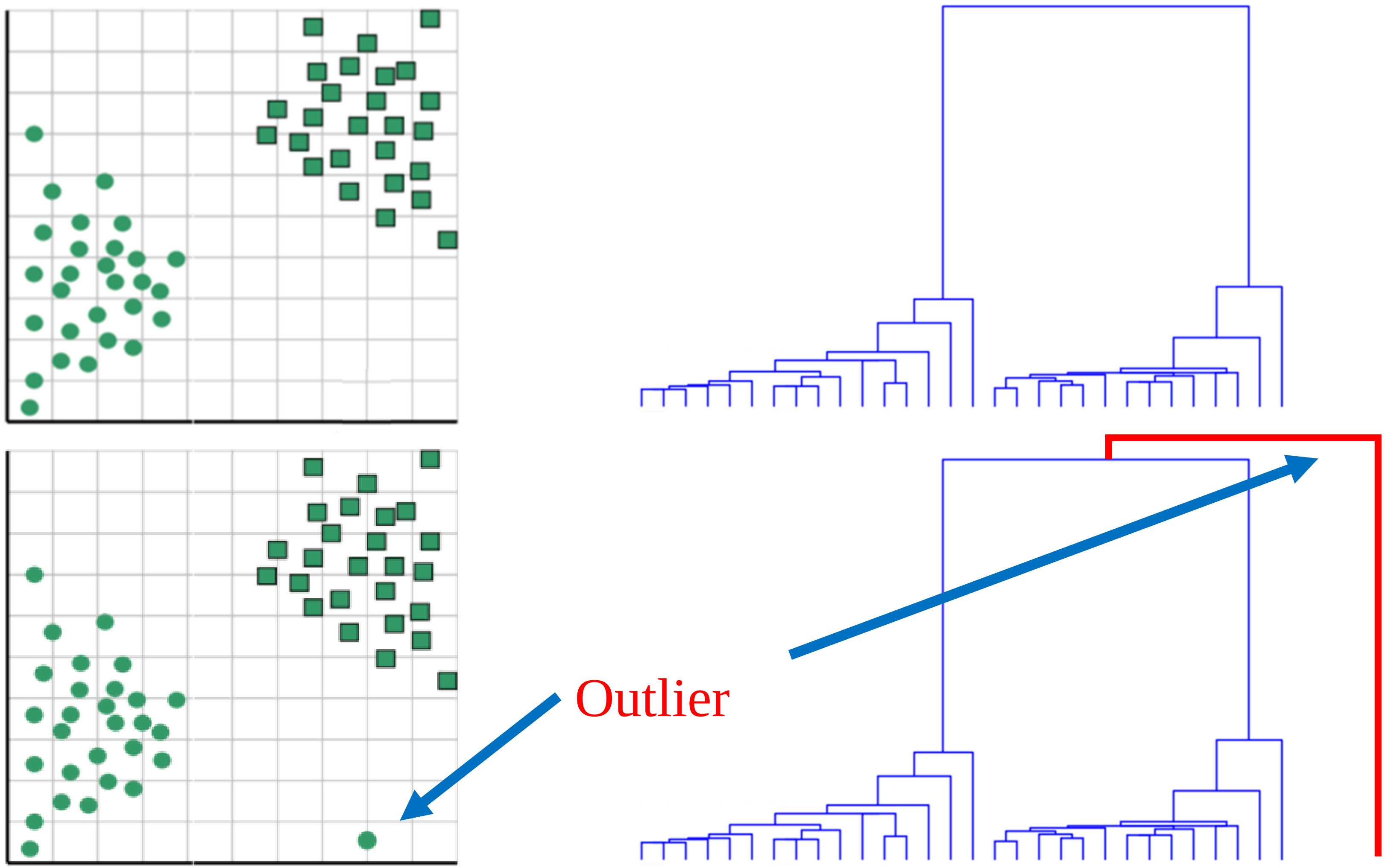


Height represents distance between objects / clusters



Agglomerative

Agglomerative Clustering Algorithm



https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Agglomerative

Agglomerative Clustering Algorithm

Good

- Simple to implement, widespread application
- Clusters have adaptive shapes
- Provides a hierarchy of clusters
- No need to specify number of clusters in advance

Bad

- May have imbalanced clusters
- Still have to choose number of clusters or threshold
- Does not scale well and more runtime of O(n3)
- Can get stuck at a local optima

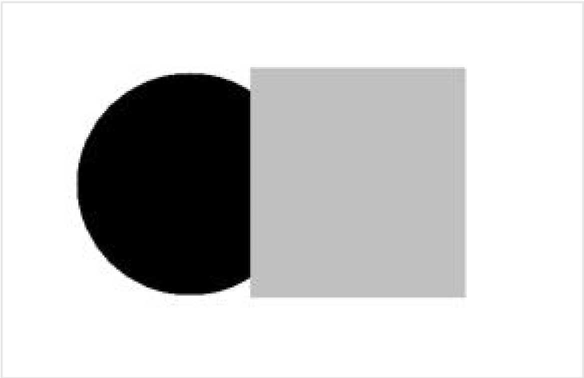
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Exercise

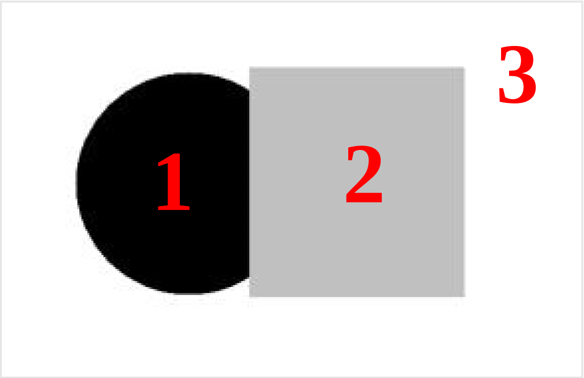
K-means

K-means

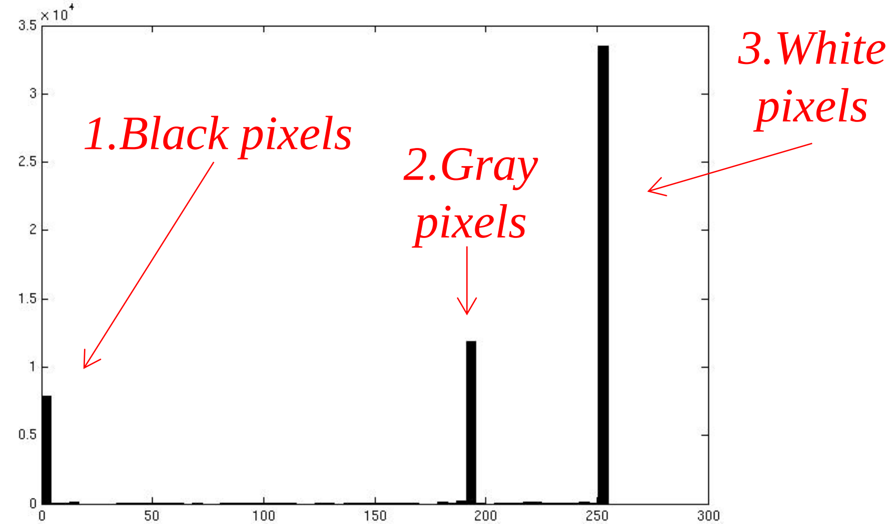
Image Segmentation



Input image



Labels of Input image



Intensity of Input image

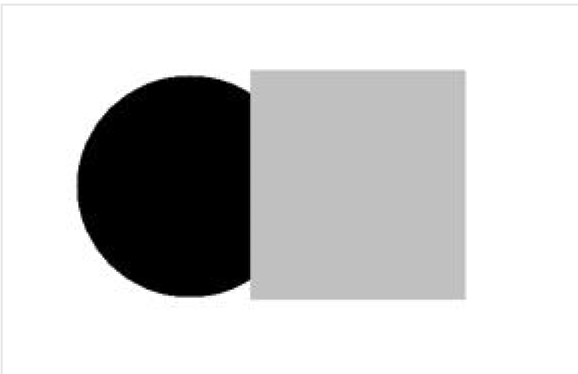
- These intensities define the three groups

What if the image isn't quite so simple?

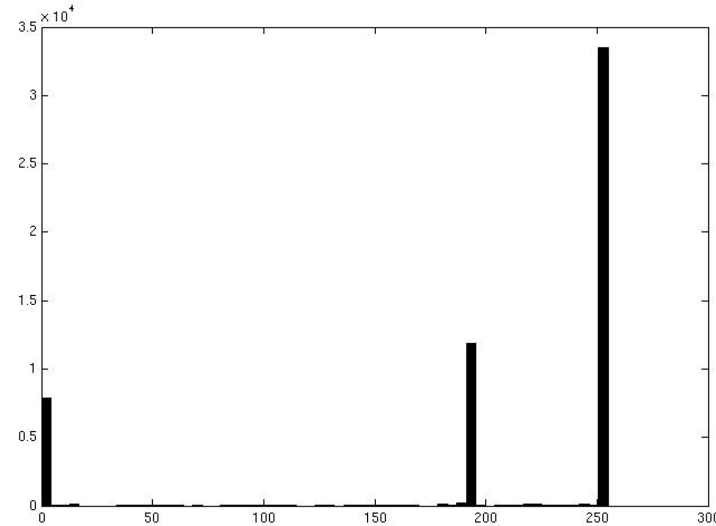
Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

K-means

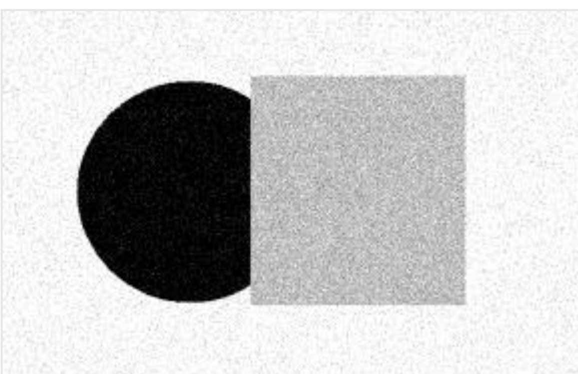
Image Segmentation



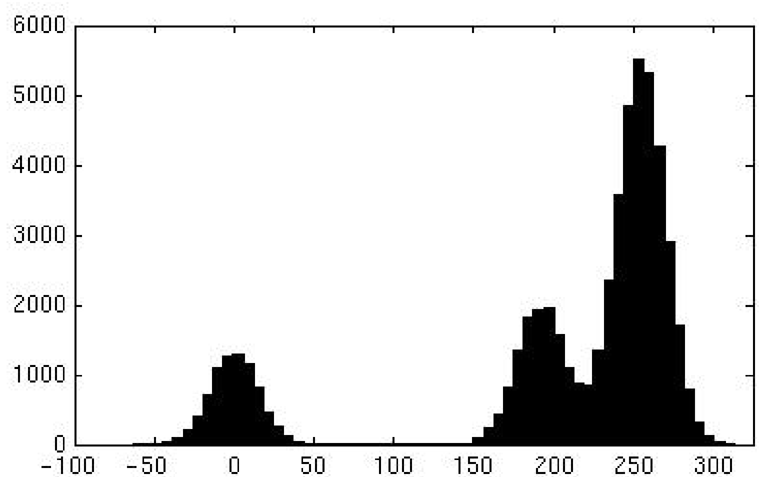
Input image



Intensity of Input image

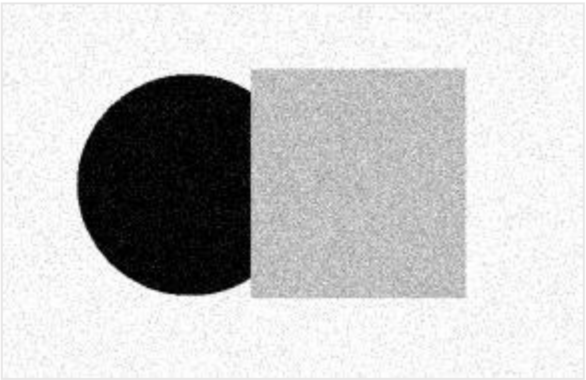


Input image
with Noise

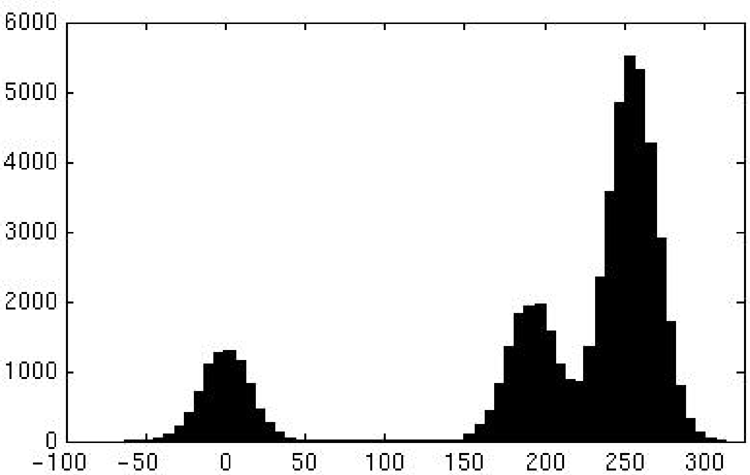


K-means

Image Segmentation



Input image with Noise



Intensity of Noise image

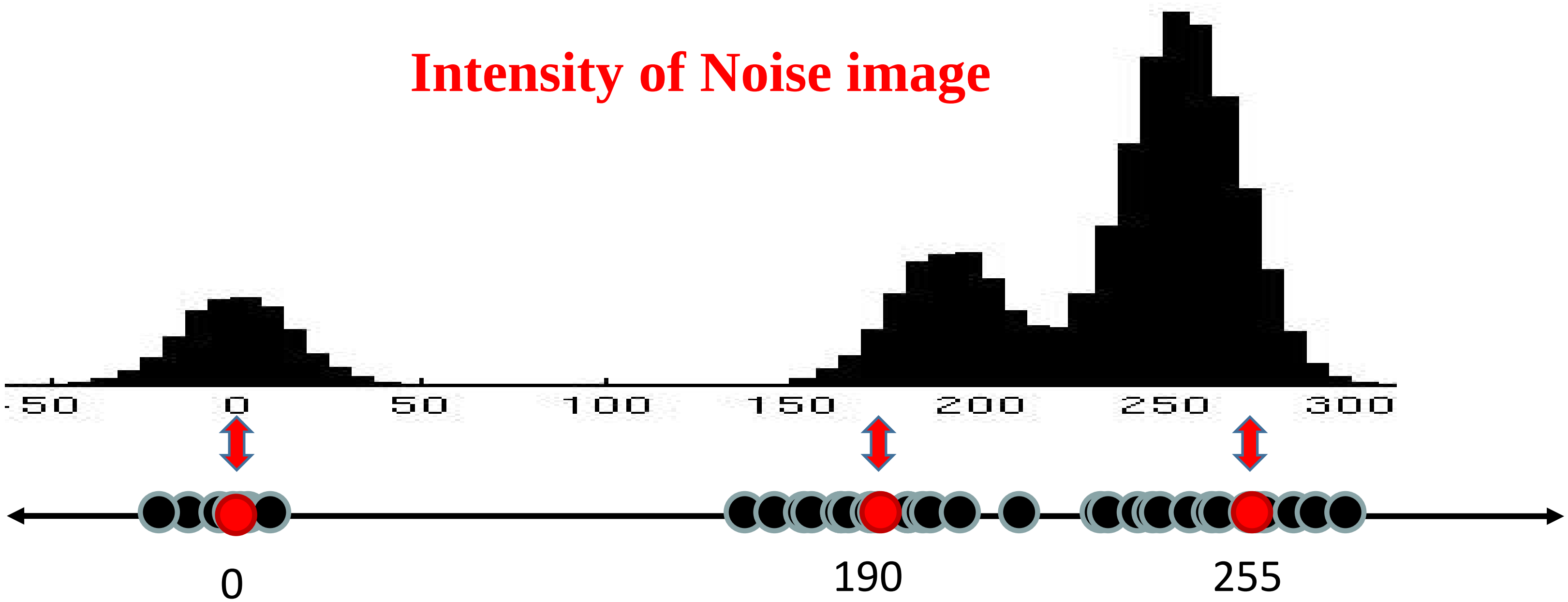
How many groups to determine these intensities?

Need to use a cluster algorithm

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

K-means

Image Segmentation

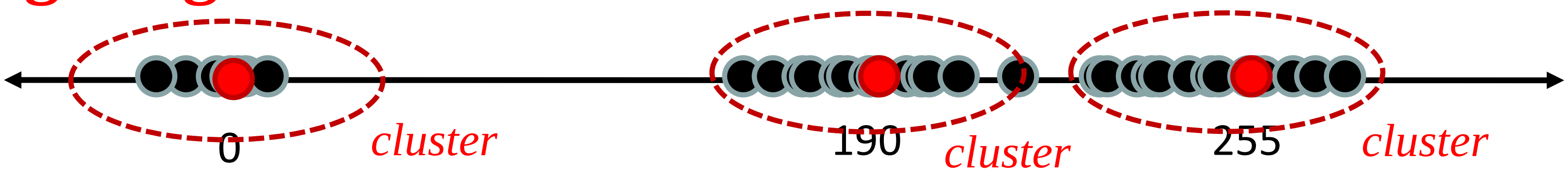


- **Goal**: choose three "centers" intensities and label every pixel according to which is nearest of these centers

Are these the best cluster center?

K-means

Image Segmentation



- **Goal**: choose three "centers" intensities and label every pixel according to which is nearest of these centers
- Best cluster centers:
  - Minimize Sum of Square Distance (SSD)

$$SSD = \sum_{cluster\ i} \sum_{x \in cluster\ i} (x - c_i)^2$$

  - Between all points x in cluster and their nearest cluster center c_i

- **Goal**: cluster to minimize variance in data given clusters
  - by using the K-mean clustering algorithm

https://github.com/StanfordVL/CS131_notes

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

K-means

Image Segmentation by K-means Clustering

- Consists of three main steps

1. **Initialization**
   - Set the number of clusters, K
   - Initial "means" (centroids) are generated at random
2. **Assignment**
   - Clusters are created by observation with the nearest centroid
3. **Update**
   - Centroid of the clusters becomes the new mean
   - Repeat the assignment step until any same reassignment

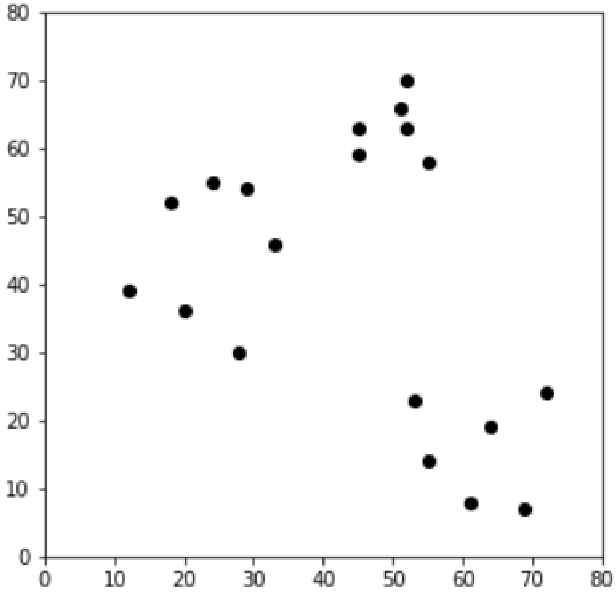
K-means

Image Segmentation by K-means Clustering

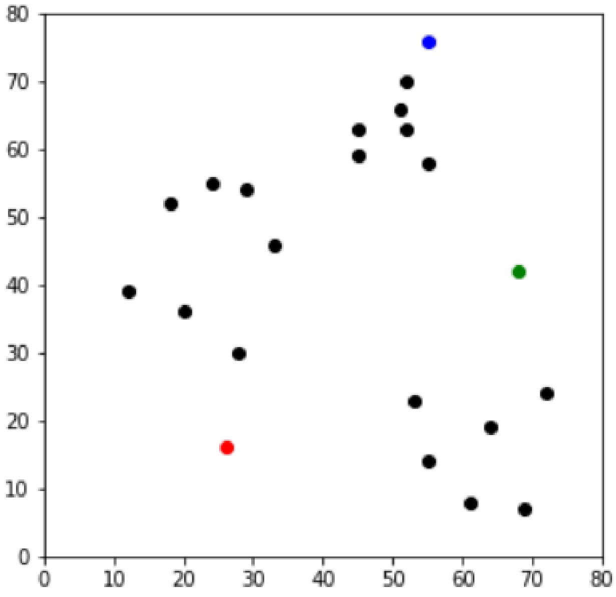
- Consists of three main steps

1. Initialization

- Set the number of clusters, K
- Initial "means" (centroids) are generated at random



Set K =3



http://benalexkeen.com/k-means-clustering-in-python/

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

K-means

Image Segmentation by K-means Clustering

- Consists of three main steps

2. Assignment

- Clusters are created by observation with the nearest centroid

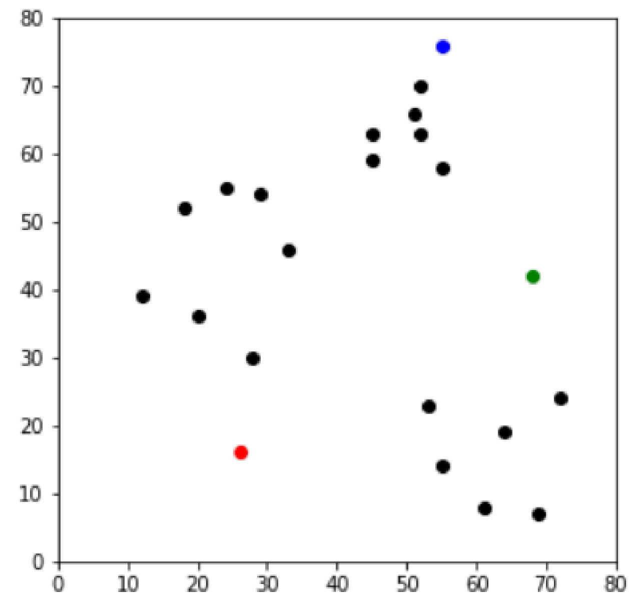
$$\delta^t = \underset{\delta}{argmin} \frac{1}{N} \sum_{j}^{N} \sum_{i}^{K} \delta_{ij}^{t-1} (c_i^{t-1} - x_j)^2$$

$\delta^t = Assignment\ set\ of\ each\ point$

$x_j = Each\ point$

$c_i = Cluster\ center$

$t = Iteration$



http://benalexkeen.com/k-means-clustering-in-python/

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

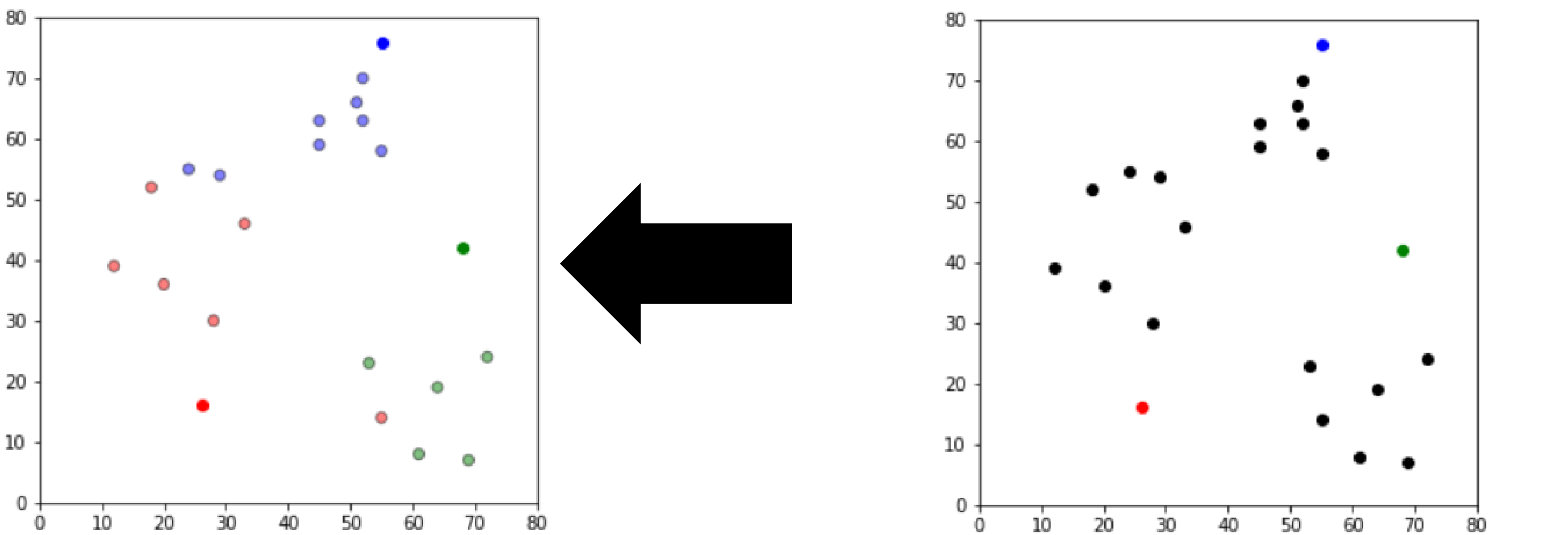
K-means

Image Segmentation by K-means Clustering

- Consists of three main steps

2. **Assignment**
   - Clusters are created by observation with the nearest centroid

$$\delta^t = \underset{\delta}{argmin} \frac{1}{N} \sum^{N} \sum^{K} \delta_{ij}^{t-1} (c_i^{t-1} - x_j)^2$$



http://benalexkeen.com/k-means-clustering-in-python/

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

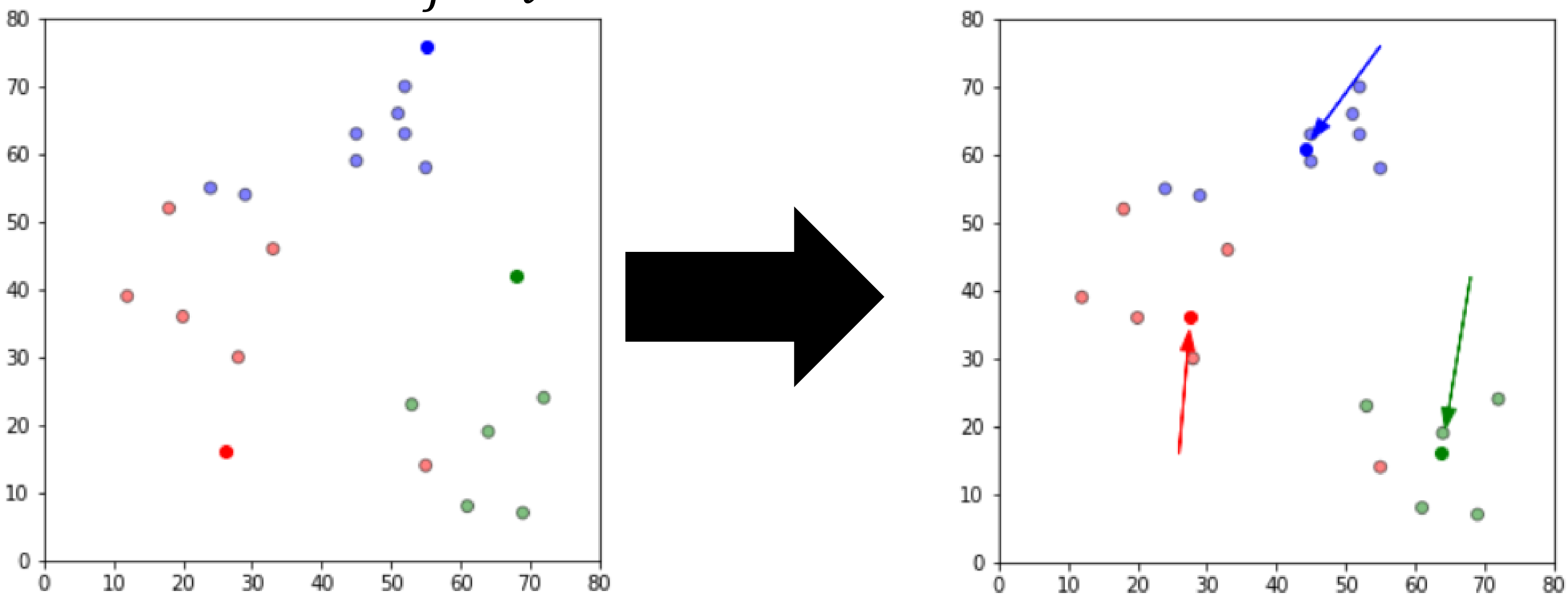
K-means

Image Segmentation by K-means Clustering

- Consists of three main steps

3. **Update**
   - Centroid of the clusters becomes the new mean

$$c^t = \underset{c}{argmin}\, \frac{1}{N} \sum_{j}^{N} \sum_{i}^{K} \delta_{ij}^{t} (c_i^{t-1} - x_j)^2$$



http://benalexkeen.com/k-means-clustering-in-python/

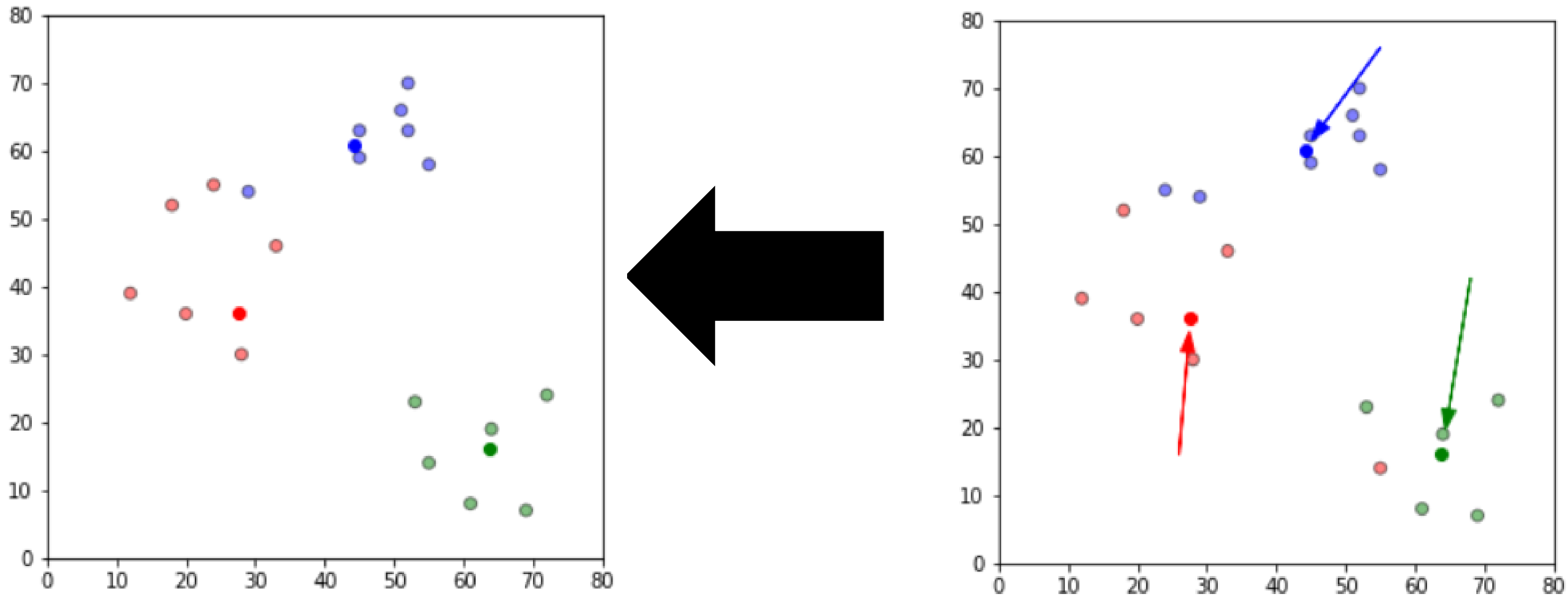
Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

K-means

Image Segmentation by K-means Clustering

- Consists of three main steps

3. **Update**
   - Centroid of the clusters becomes the new mean
   - Repeat the assignment step until any same reassignment



http://benalexkeen.com/k-means-clustering-in-python/

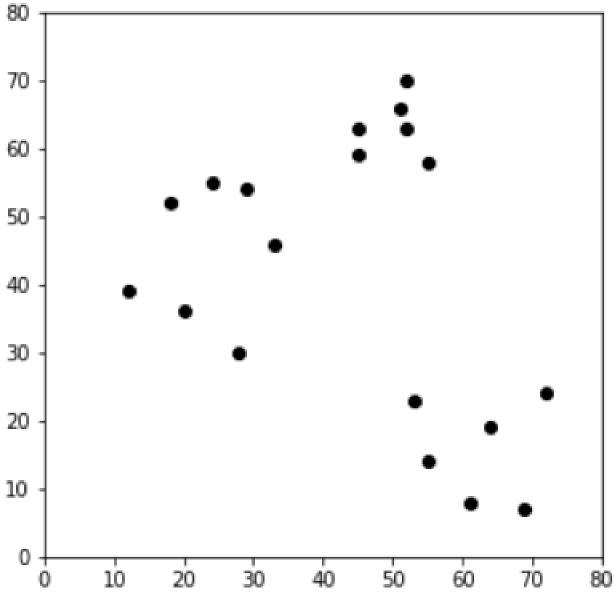
Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

K-means

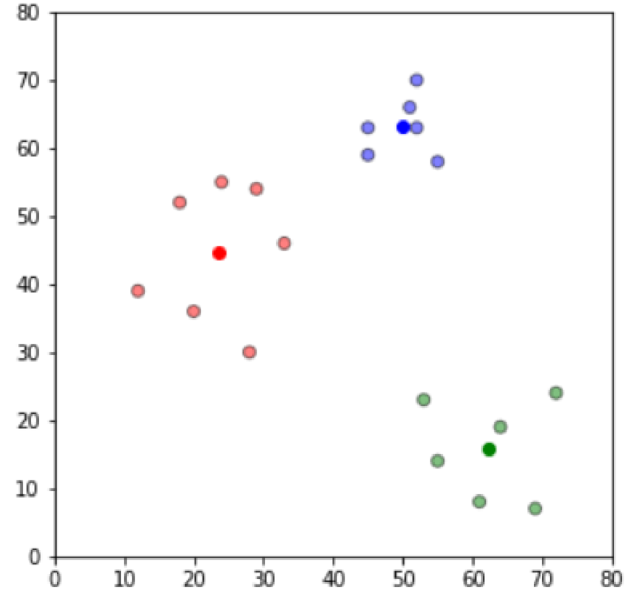
Image Segmentation by K-means Clustering

- Consists of three main steps

2. **Assignment**
3. **Update**
    - Repeat the assignment step until any same reassignment



Set K =3

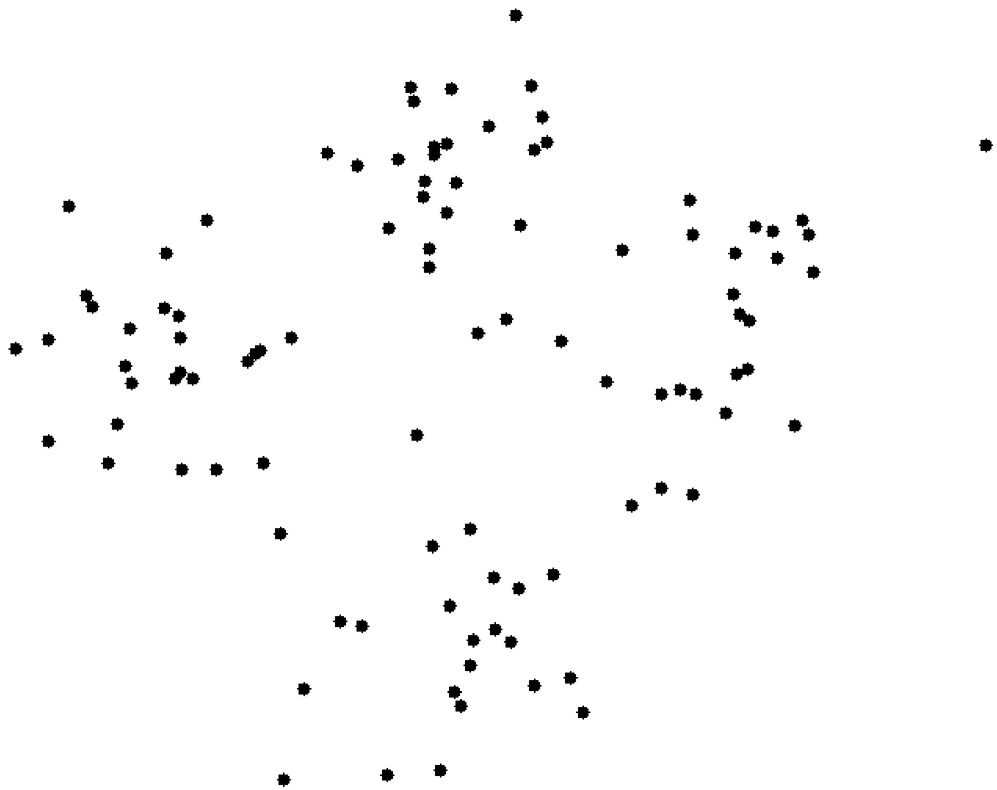


K-means

Image Segmentation by K-means Clustering

- Example

Set K =4



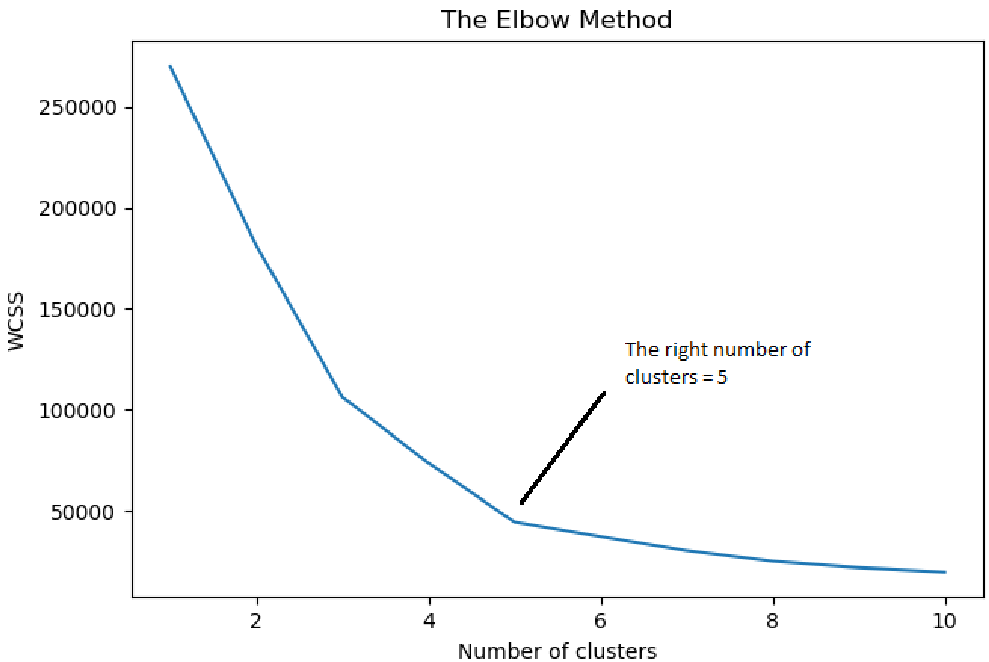
K-means

Image Segmentation by K-means Clustering

How many numbers do well the data clusters?

- **Elbow method**
- Calculate total within-cluster sum of square (WCSS) is minimized
- Total WCSS measures the compactness of the clustering
- Want it to be as small as possible

$$wcss = \sum_{j=1}^{N} (x_j - c_j)^2$$

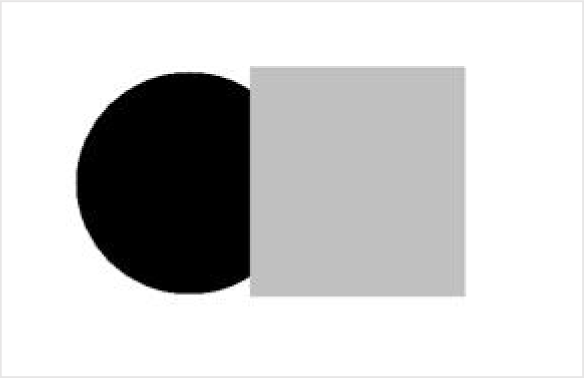


https://towardsdatascience.com/introduction-to-image-segmentation-with-k-means-clustering-83fd0a9e2fc3

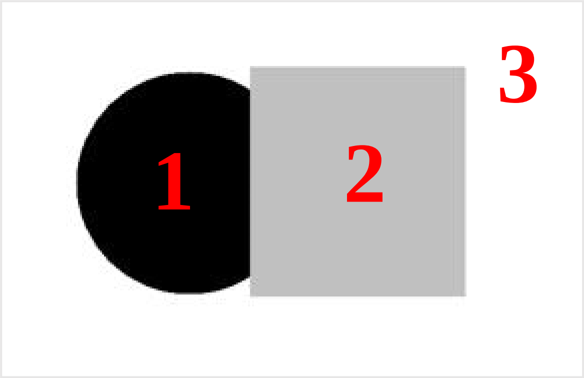
Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

K-means

Image Segmentation by K-means Clustering



Input image



Three Objects

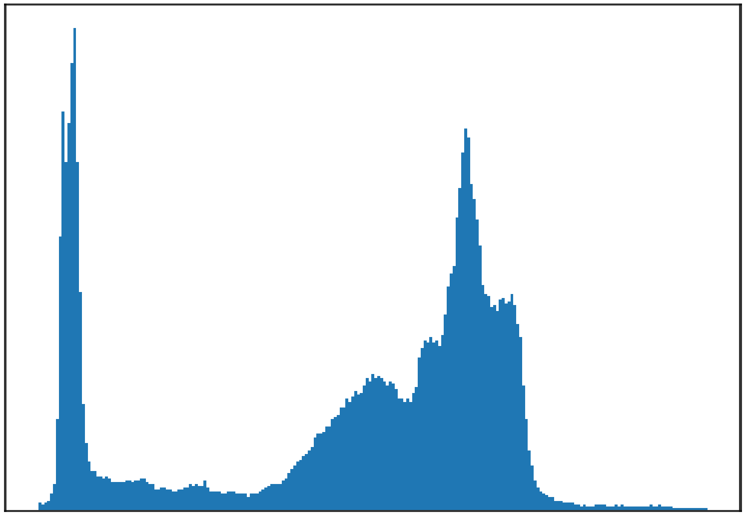
Two Objects

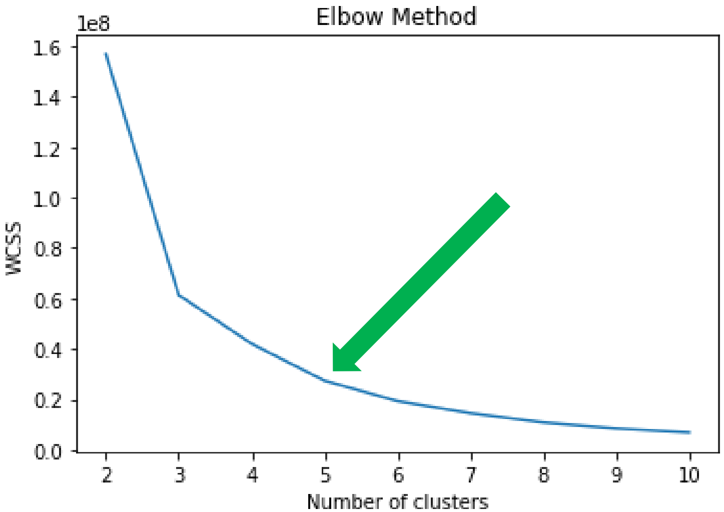
https://github.com/StanfordVL/CS131_notes

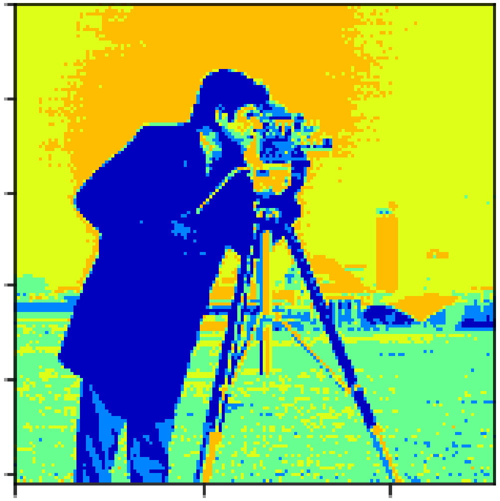
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

K-means

Image Segmentation by K-means Clustering







https://github.com/StanfordVL/CS131_notes

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Exercise

K-means

K-means Clustering Algorithm

Pros

- Finds cluster centers that minimize conditional variance
- Simple, Fast, and Easy to implement
- Unsupervised clustering

Cons

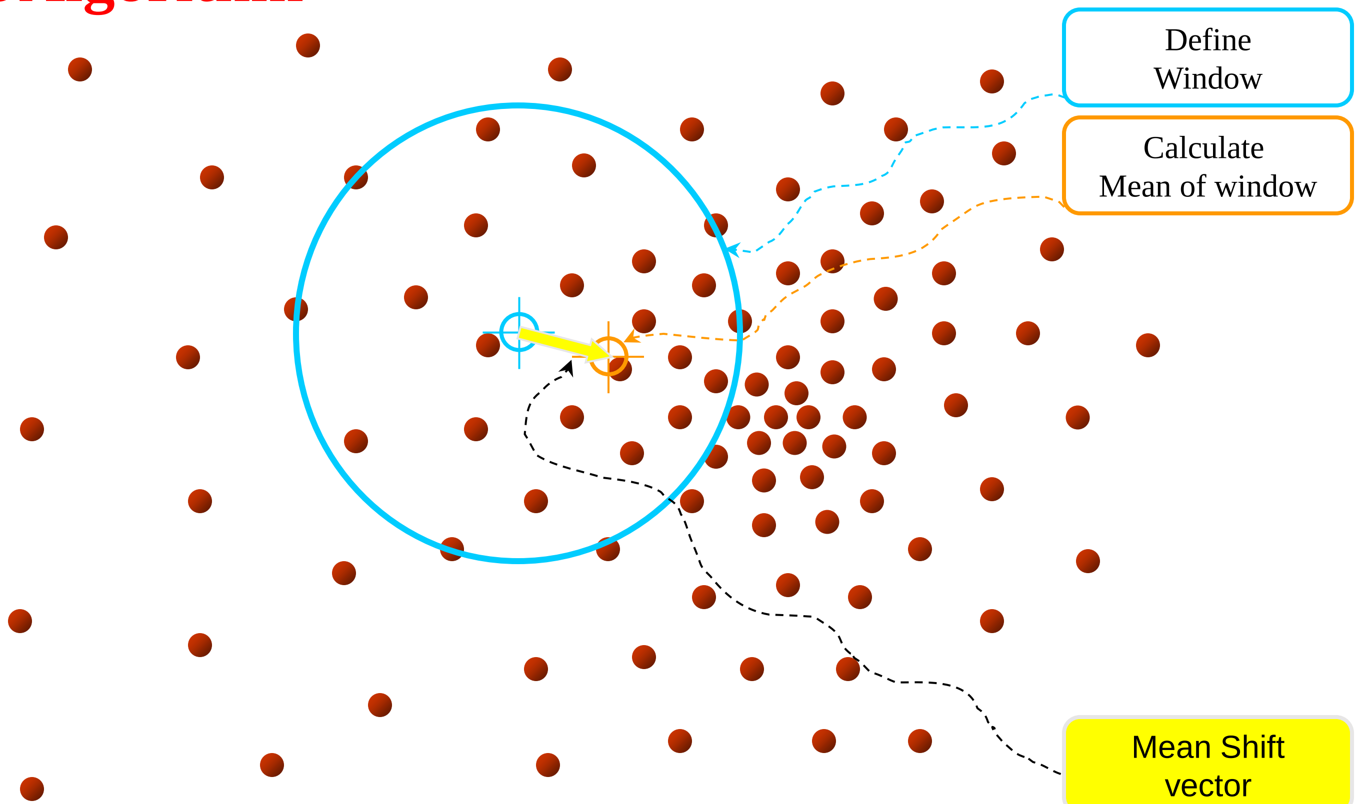
- Need to set the number of K clustering
- Sensitive to outliers
- All clusters have the same parameters

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Mean-shift

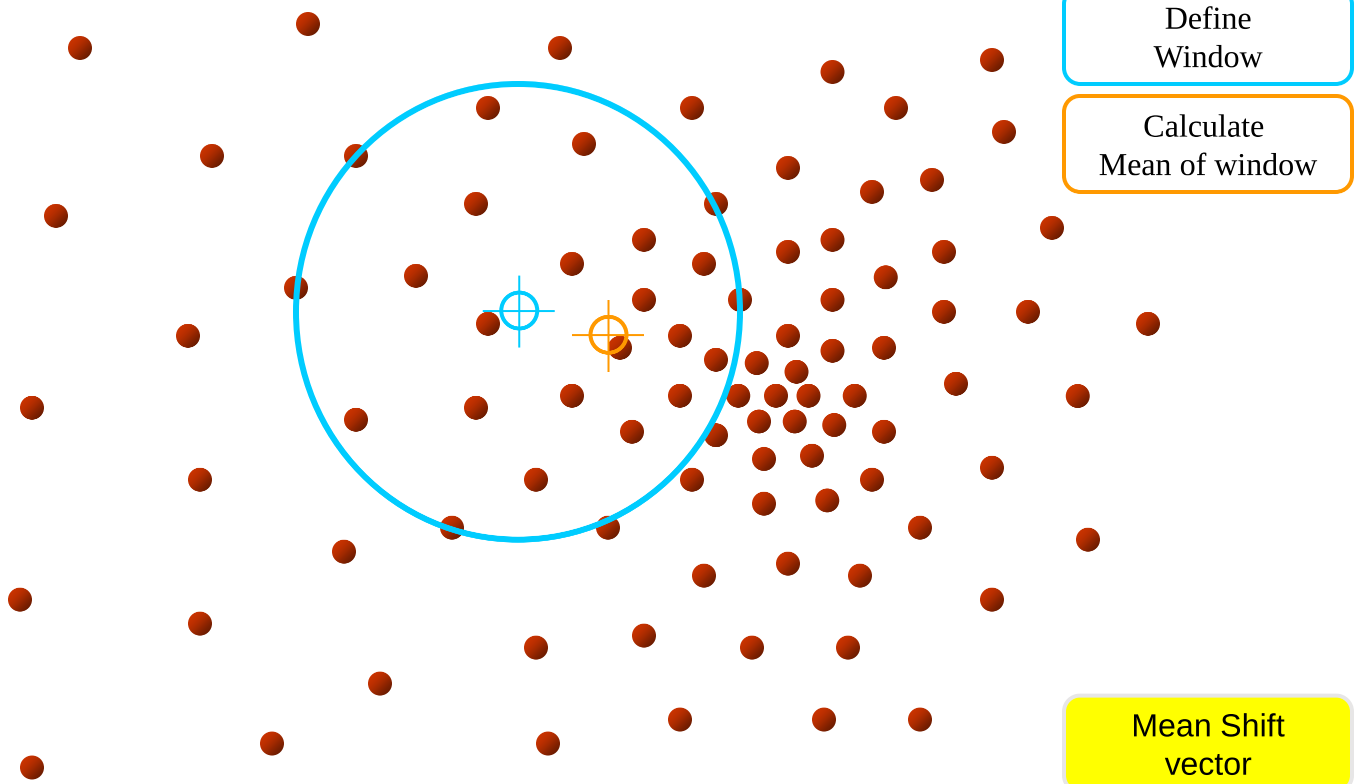
Mean-shift

Mean-Shift Algorithm



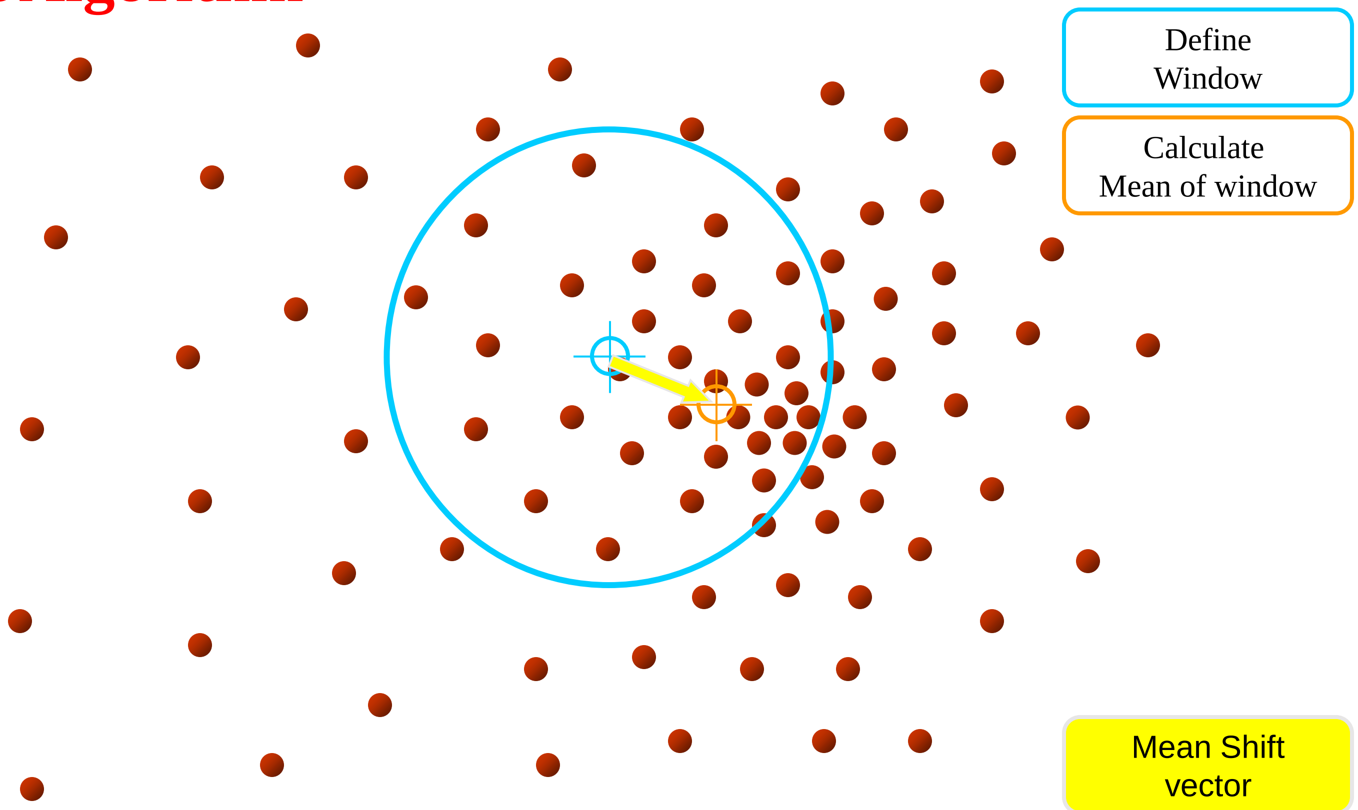
Mean-shift

Mean-Shift Algorithm



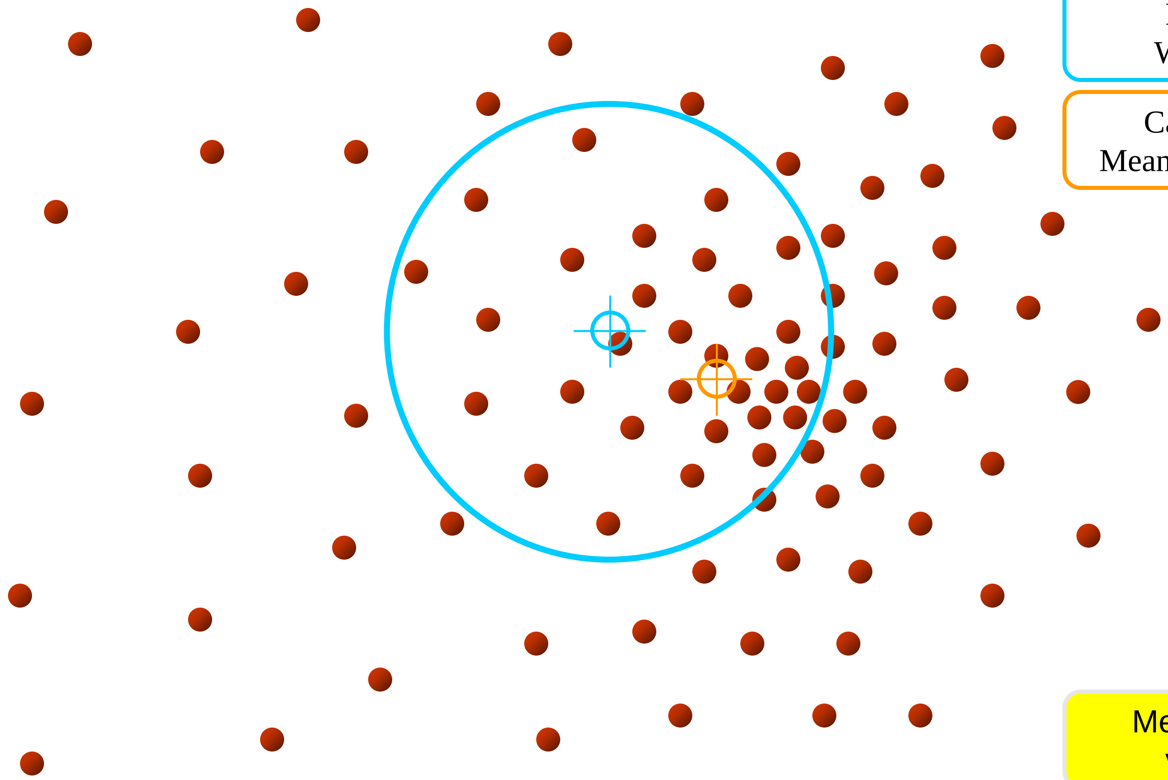
Mean-shift

Mean-Shift Algorithm



Mean-shift

Mean-Shift Algorithm

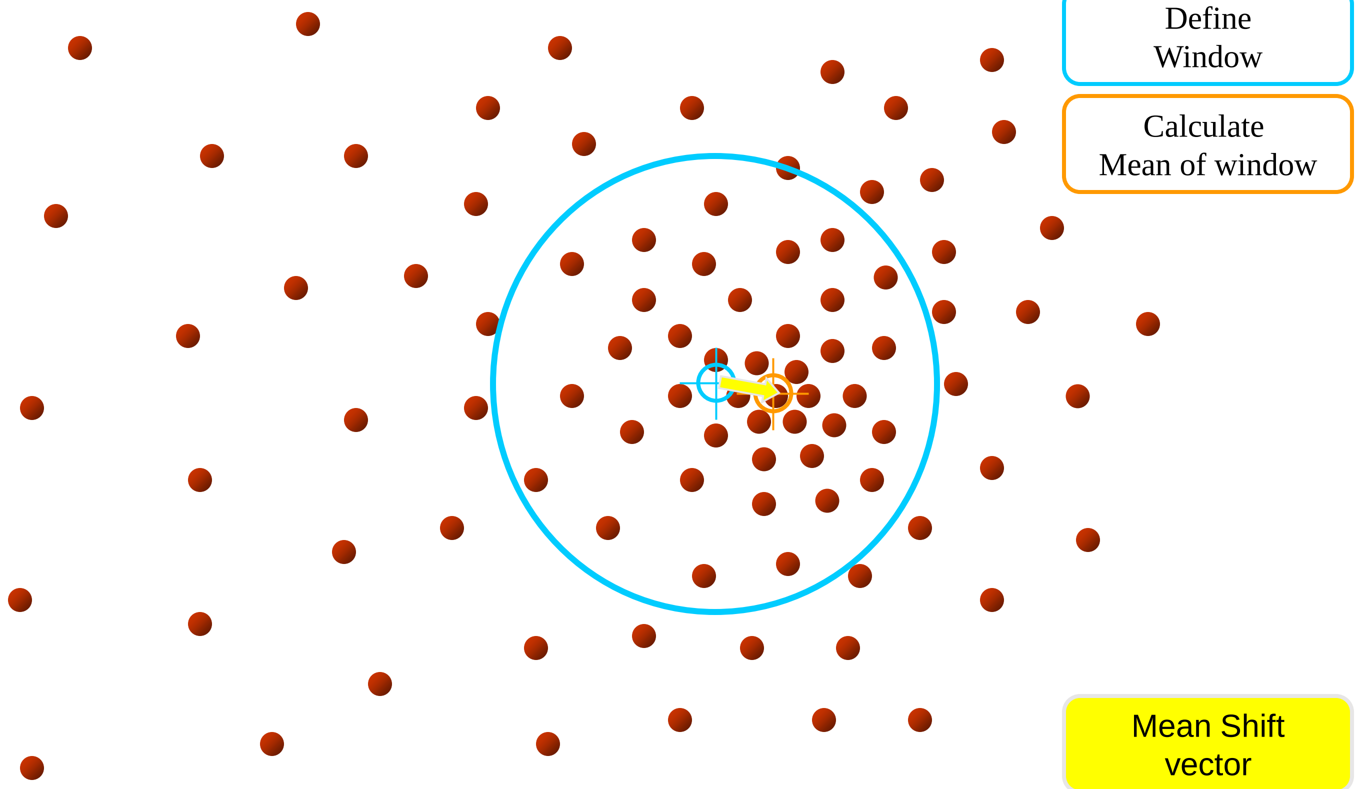


Define Window

Calculate Mean of window

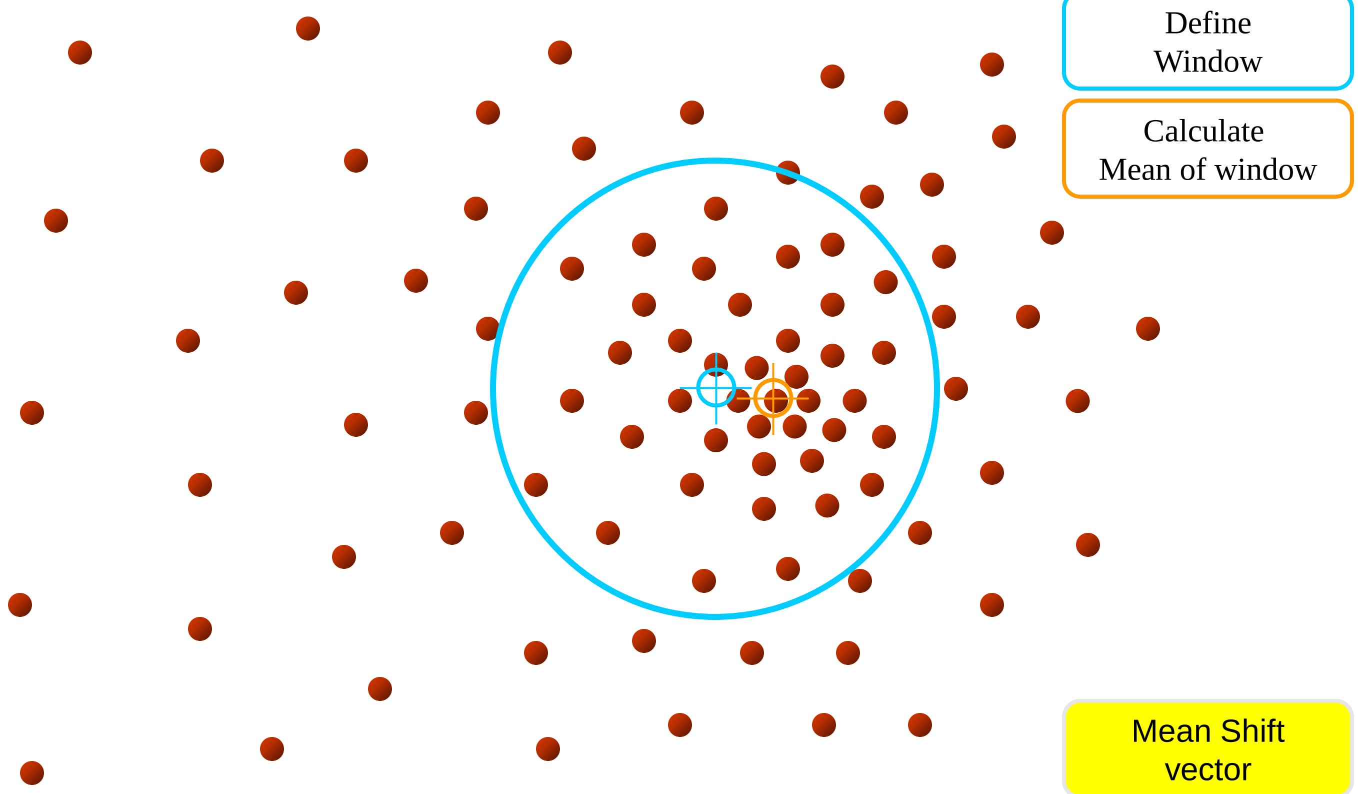
Mean Shift vector

Mean-Shift Algorithm



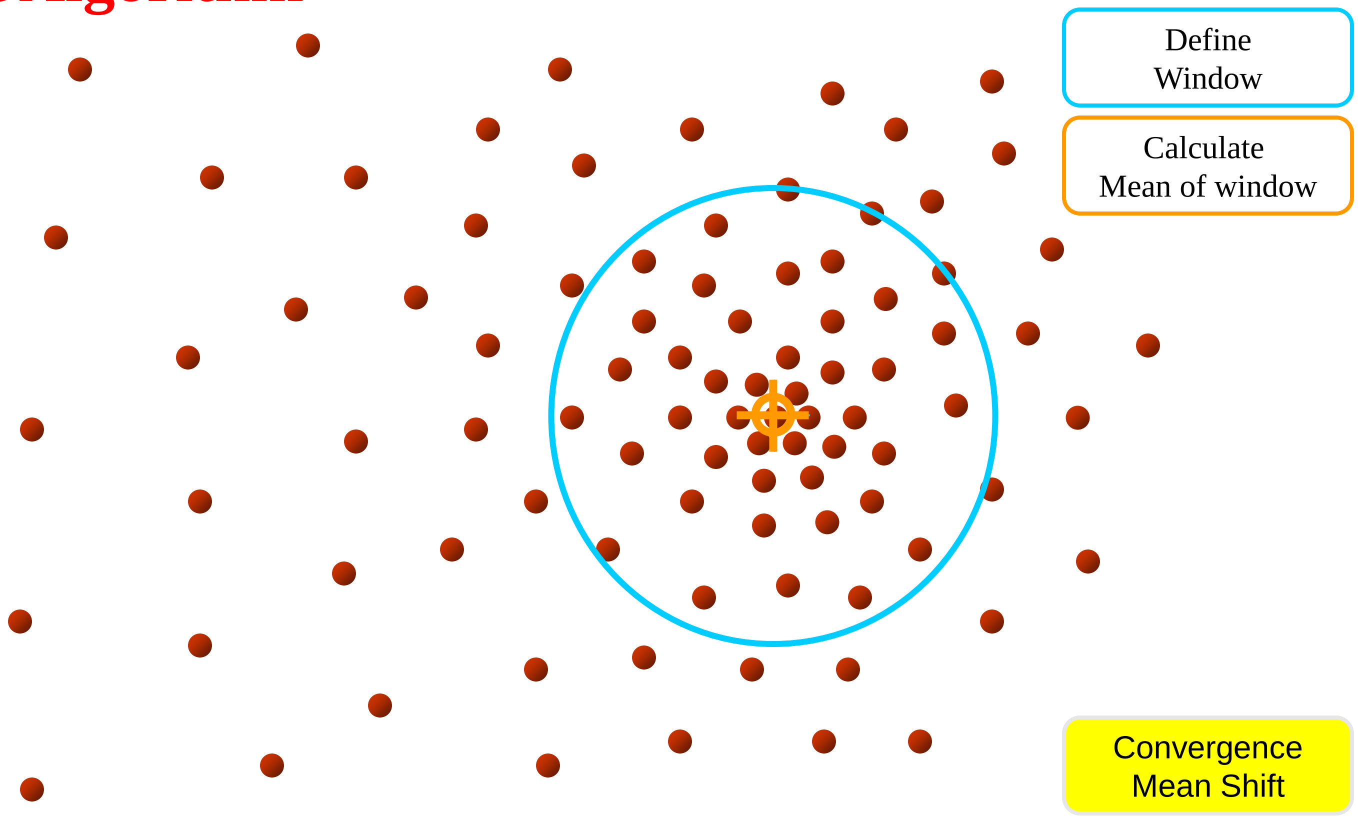
Mean-shift

Mean-Shift Algorithm



Mean-shift

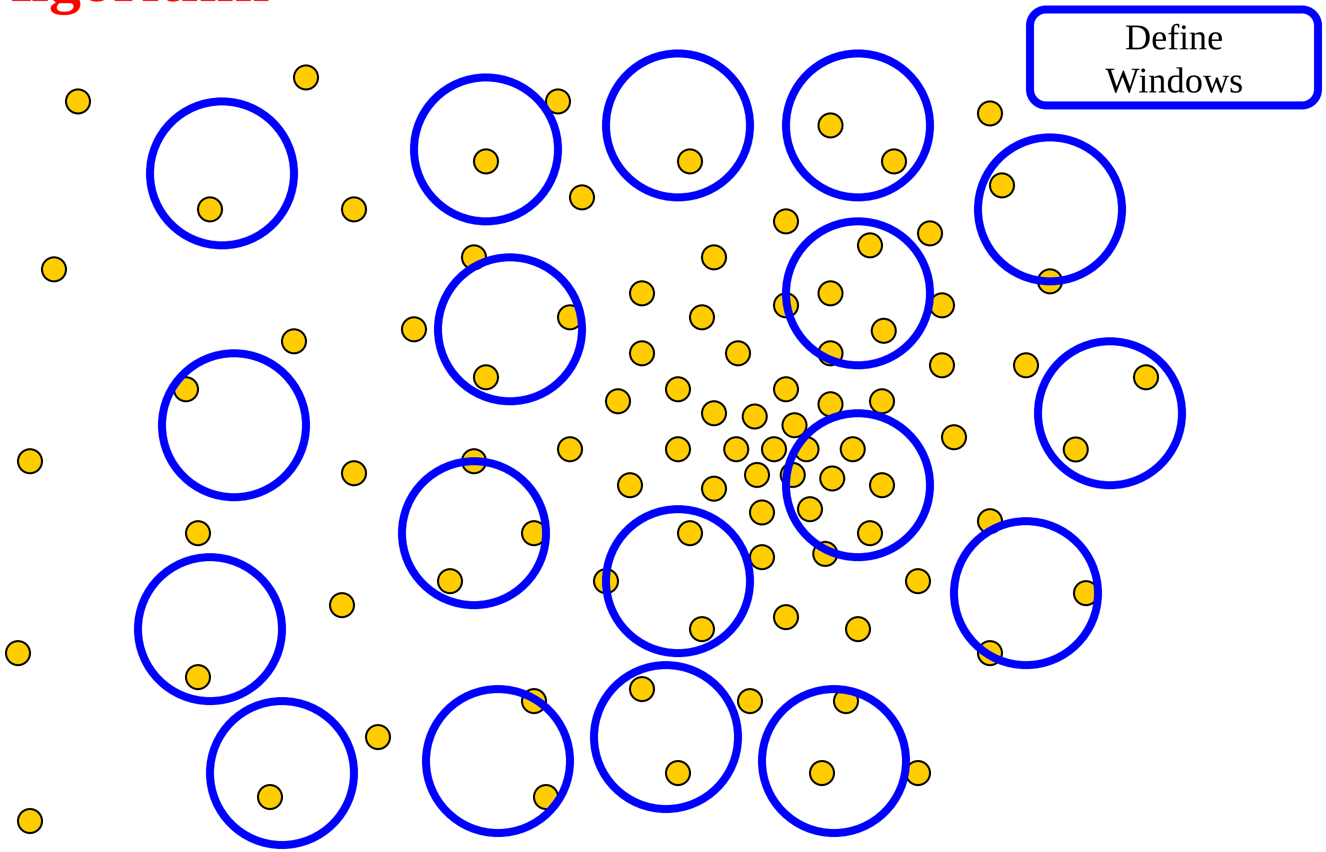
Mean-Shift Algorithm



Mean-shift

Mean-Shift Algorithm

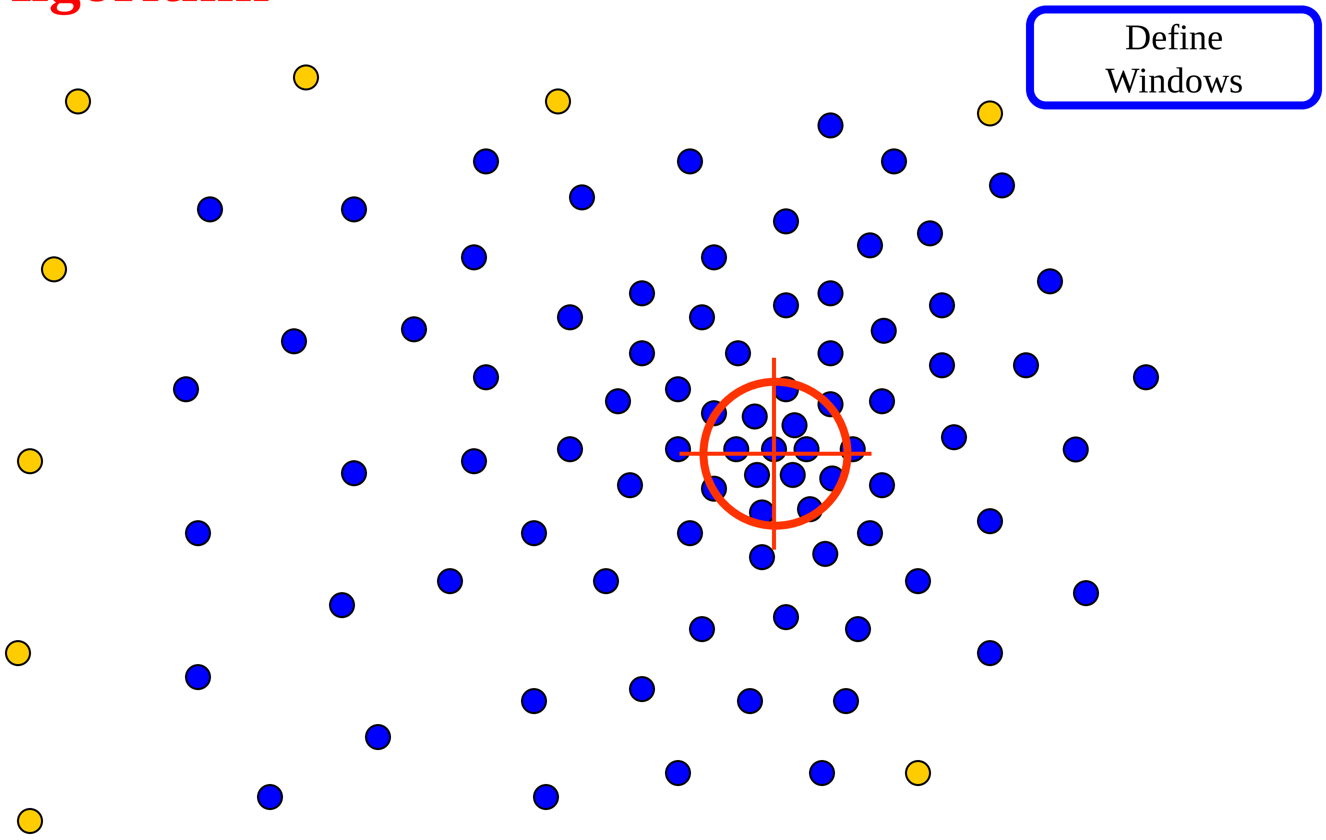
Run the parallel procedure



Mean-shift

Mean-Shift Algorithm

Run the parallel procedure



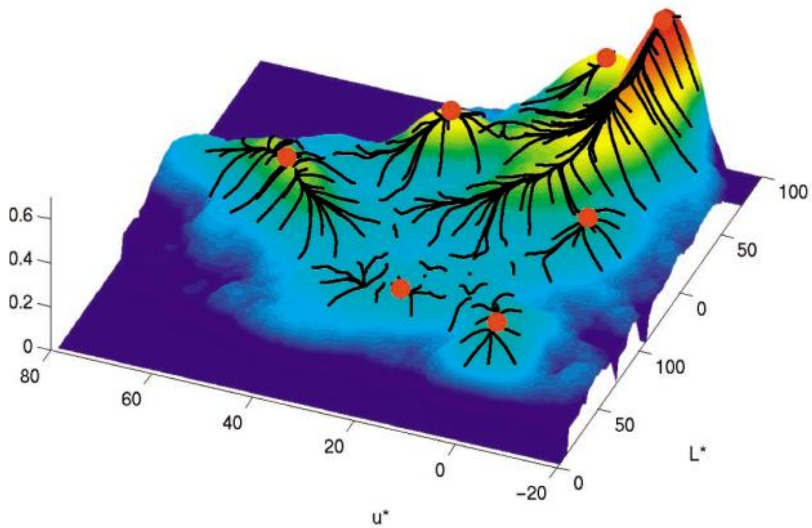
The blue data points were traversed by the windows towards the mode.

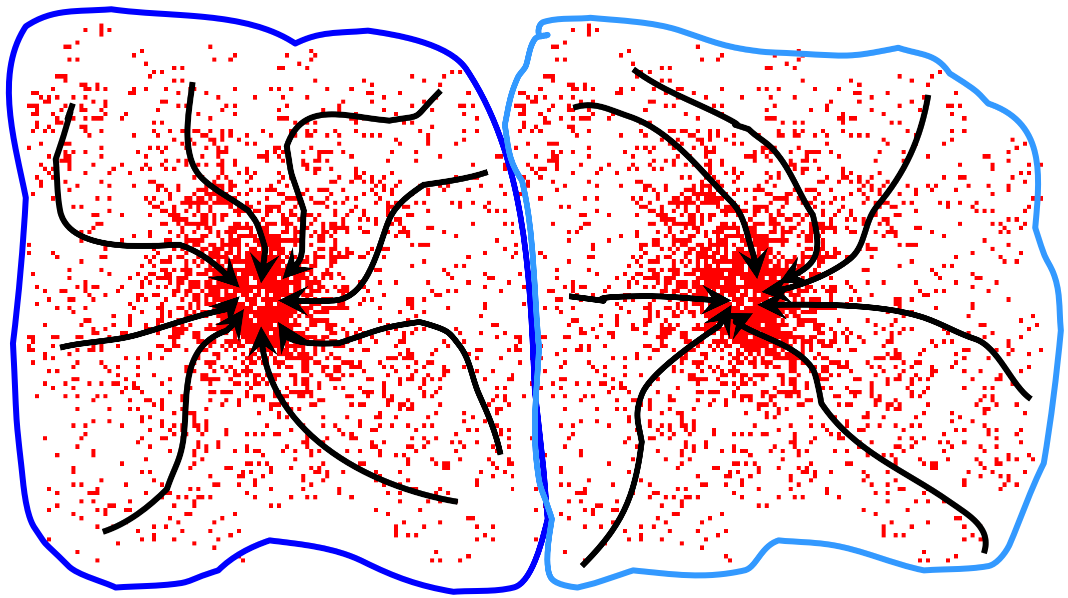
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Mean-shift

Mean-Shift Clustering

- Cluster: all data points in the attraction basin of a mode
- Attraction basin: the region of all trajectories lead to the same mode

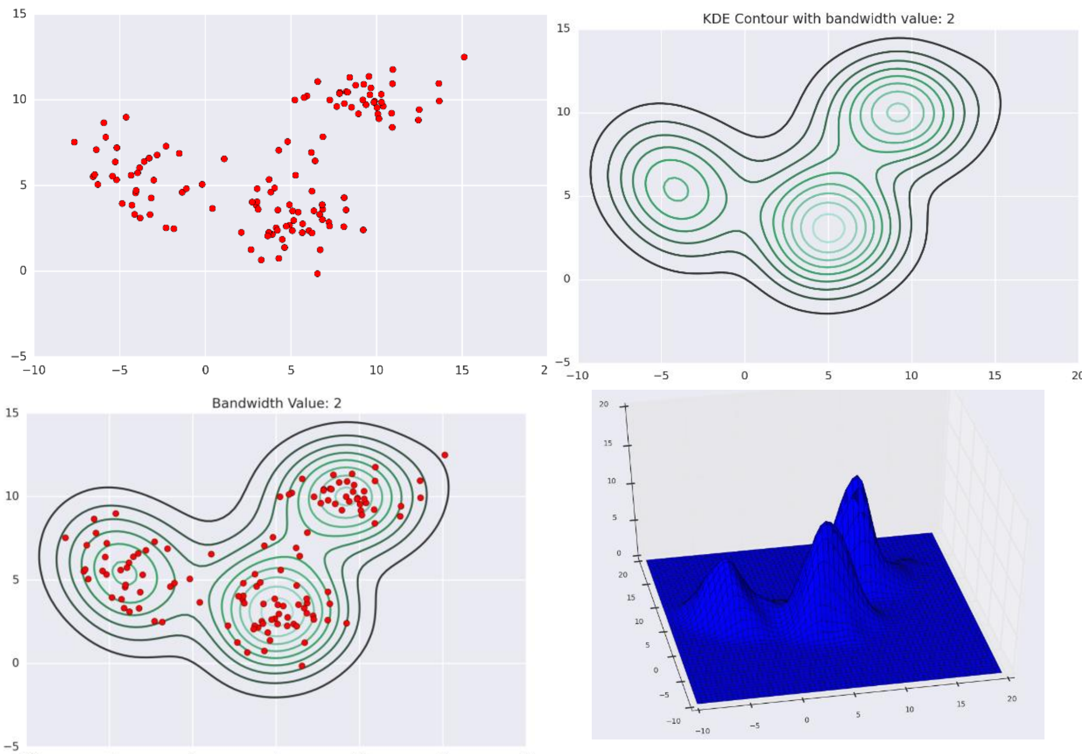




Mean-shift

Image Clustering/Segmentation by Mean-Shift

- Find features (color, gradients, texture, etc)
- Initialize windows at individual pixel locations
- Perform mean shift for each window until convergence
- Merge windows that end up near the same "peak" or mode



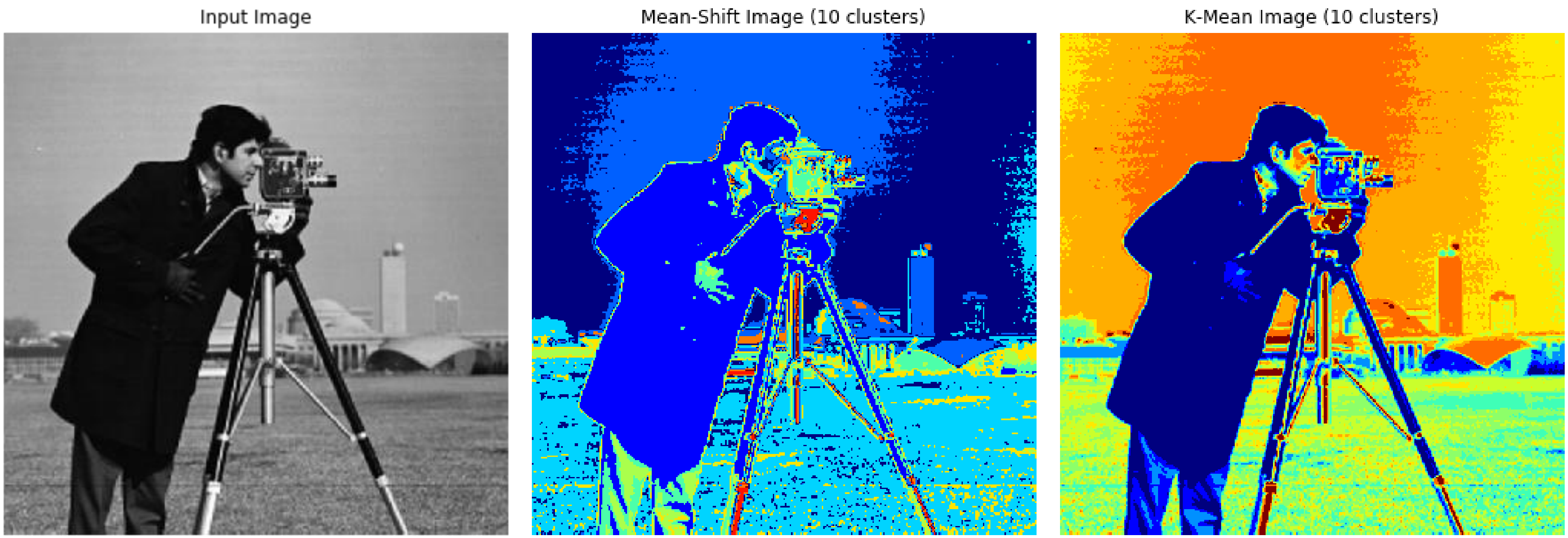
http://vision.stanford.edu/teaching/cs131_fall1314_nope/lectures/lecture13_kmeans_cs131.pdf

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Mean-shift

Image Clustering/Segmentation by Mean-Shift

- Find features (color, gradients, texture, etc)
- Merge windows that end up near the same "peak" or mode



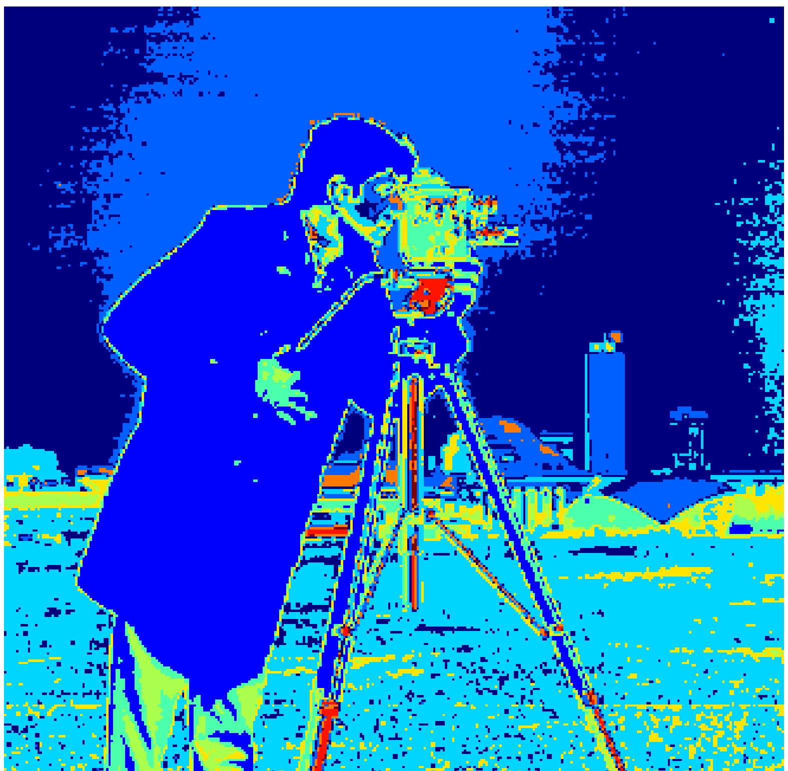
Mean-shift

Image Clustering by Mean-Shift vs K-Mean

- Find features (color, gradients, texture, etc)
- Merge windows that end up near the same "peak" or mode

Input Image

Mean-Shift Image (10 clusters)



K-mean (5 clusters)

http://vision.stanford.edu/teaching/cs131_fall1314_nope/lectures/lecture13_kmeans_cs131.pdf

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Exercise

Mean-shift

Mean-Shift Clustering Algorithm

Pros

- General, application-independent tool
- Model-free, does not assume any prior shape on data clusters
- Just a single parameter (window size h)
- Finds variable number of modes
- Robust to outliers

Cons

- Output depends on window size
- Window size (bandwidth) selection is not trivial(little value)
- Computationally expensive (~2s/image)
- Does not scale well with dimension of feature space

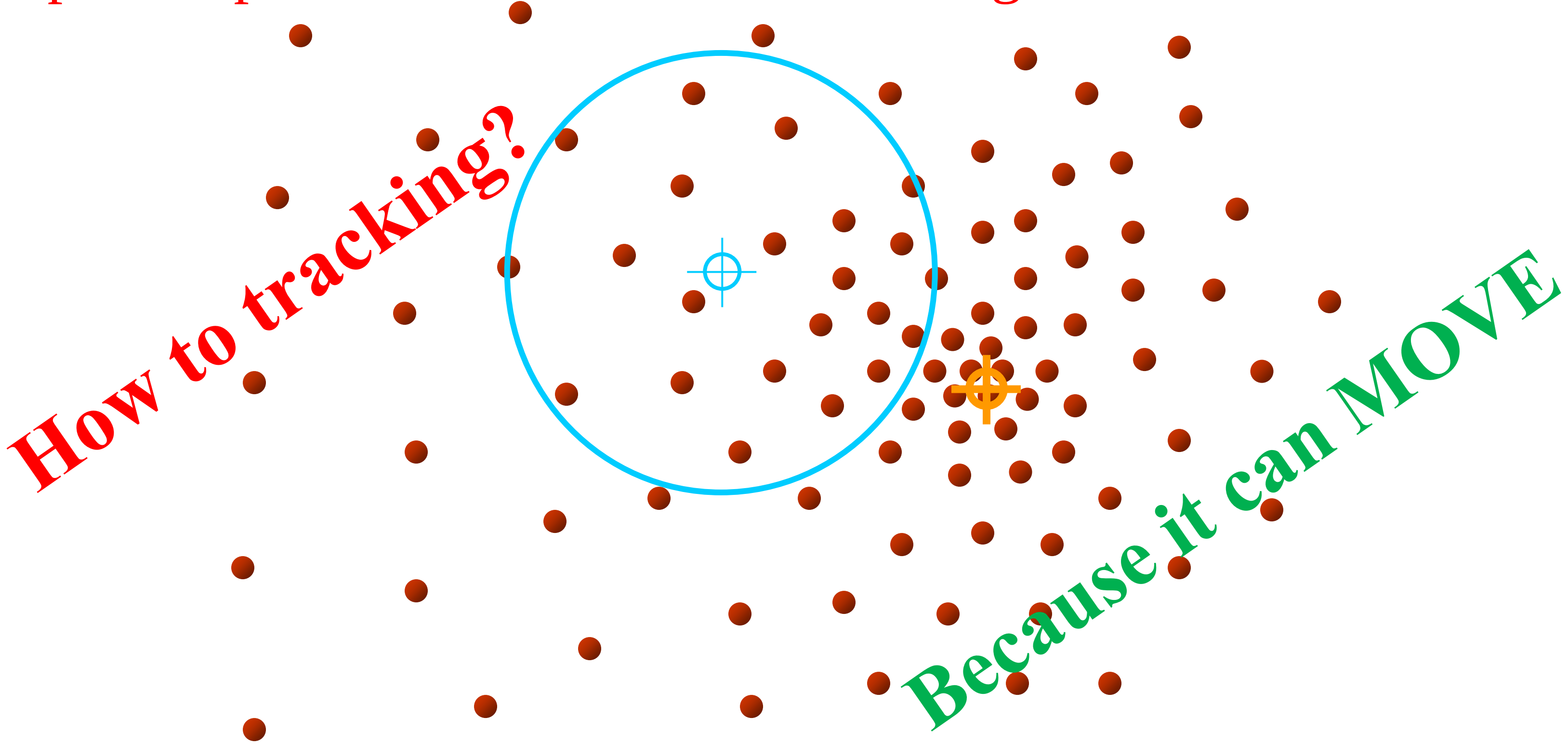
Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Mean-shift Tracking

Mean-shift Tracking

Mean-Shift Algorithm

1. Define a window (W) and random the initial data point
2. Calculate the mean for all the points in the window
3. **Move the center of the window** to the location of the mean
4. Repeat steps 2 and 3 until there is convergence



https://www.bogotobogo.com/python/OpenCV_Python/python_opencv3_mean_shift_tracking_segmentation.php

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

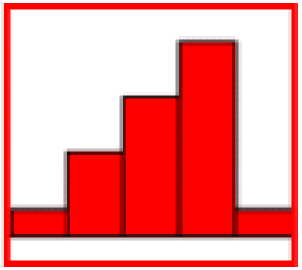
Mean-shift Tracking

Mean-Shift Tracking

1. Set the initial Tracking Area consists of just one and the object centroid

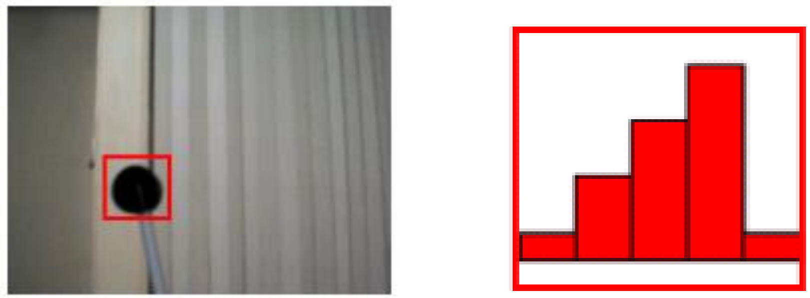
Frame 1 – Target Initialisation

2. From the initial area, the spatial coordinates of a pixel *x* and sample weight *w(x)* define the histogram belongs to an object model



Mean-Shift Tracking

Frame 1 – Target Initialisation

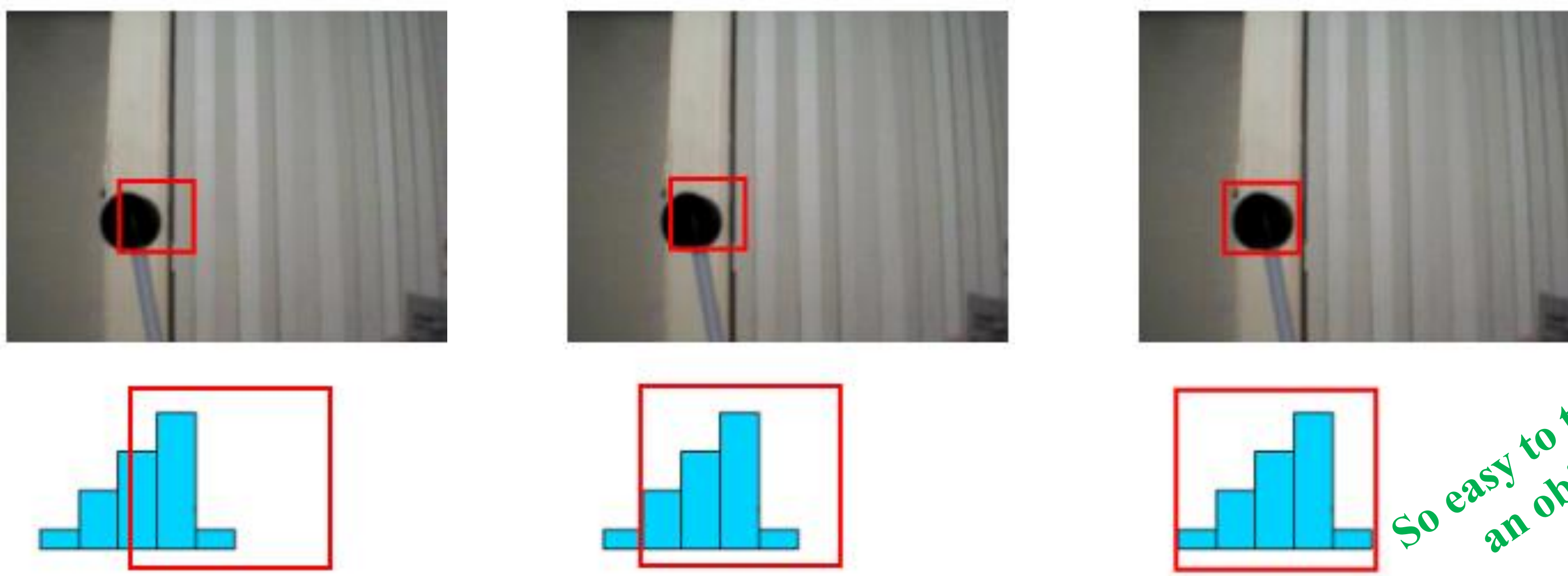


3. Mean-shifts seek the mode of the kernel density computed with these weights

Frame 2 – Target Mode Finding

Frame 2 – Meanshift Iteration 1

Frame 2 – Meanshift Iteration 2



So easy to track an object

https://www.codeproject.com/Articles/36288/Computer-Vision-Application-with-C-Part-III

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Mean-shift Tracking

Example: Mean-Shift Tracking

https://www.codeproject.com/Articles/36288/Computer-Vision-A

"Computer Vision: Algorithms and Applications," Springer, 2010

Mean-shift Tracking

Object Tracking for Laparoscopic Surgery Using the Adaptive Mean-Shift Kalman Algorithm

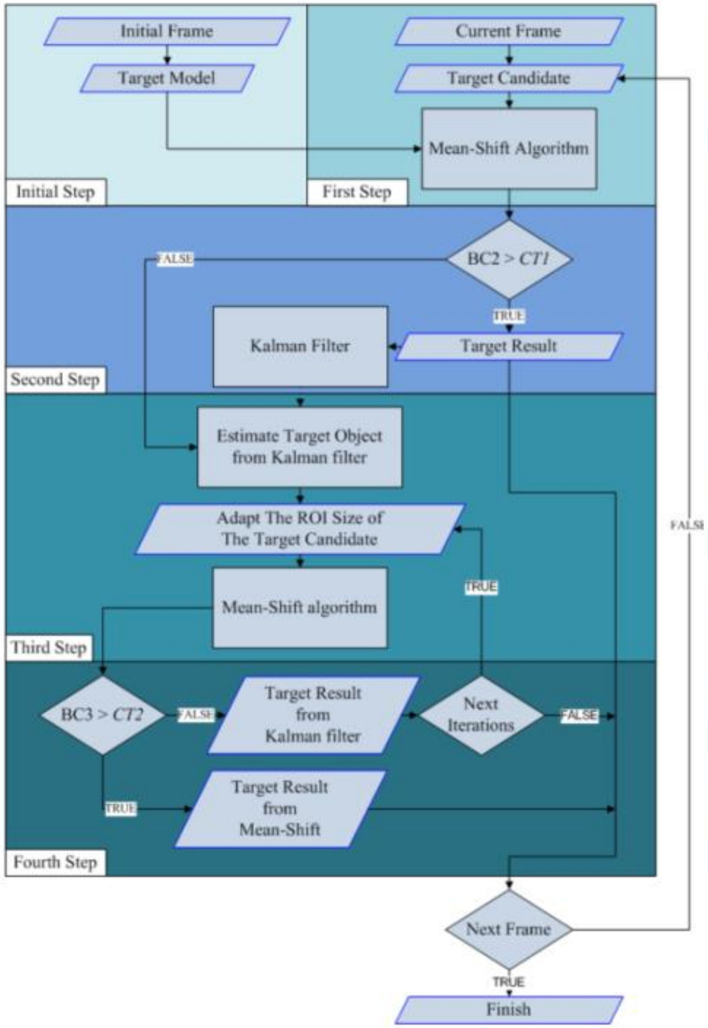


Fig.2. The overall process of the adaptive Mean-Shift Kalman tracking

Fig. 10. Sample results in Real Laparoscopic Surgery Experiment #2 using the proposed algorithm

https://www.ijml.org/papers/66-A341.pdf

Richard Szeliski, “Computer Vision: Algorithms and Applications,” Springer, 2010

Mean-shift Tracking

Mean-Shift Tracking Algorithm

Pros

- Suitable for real data analysis
- Does not assume any prior shape (such as elliptical) on data clusters
- Only 1 parameter to choose

Cons

- Window size (bandwidth) selection is not trivial(little value)
- Computationally expensive (~2s/image)
- Does not scale well with dimension of feature space
- Inappropriate window size can cause modes to be merged so need to use adaptive window size

Richard Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2010

Homework

ELECTRICAL & COMPUTER
Engineering KMUTNB

Thank You
for
Your Attention

Vera Sa-ing, Ph.D.