MAI4CAREU | Master programmes in Artificial Intelligence 4 Careers in Europe

Visual Computing Group

CYENS CENTRE OF EXCELLENCE

University of Cyprus – MSc Artificial Intelligence

MAI644 – COMPUTER VISION

Lecture 10: Visual Recognition – Segmentation

Melinos Averkiou

CYENS Centre of Excellence

University of Cyprus - Department of Computer Science

m.averkiou@cyens.org.cy

CYENS CENTRE OF EXCELLENCE

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Last time

- Linear least-squares
- RANSAC
- Panorama Stitching

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today’s Agenda

- Visual Recognition Tasks
- Introduction to segmentation and clustering
- Agglomerative clustering
- K-means clustering
- Mean-shift clustering
- Efficient Graph-based image segmentation

Reading: Forsyth Chapter 9

D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, TPAMI 2002

[material based on Niebles-Krishna]

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

- Visual Recognition Tasks
- Introduction to segmentation and clustering
- Agglomerative clustering
- K-means clustering
- Mean-shift clustering
- Efficient Graph-based image segmentation

Reading: Forsyth Chapter 9

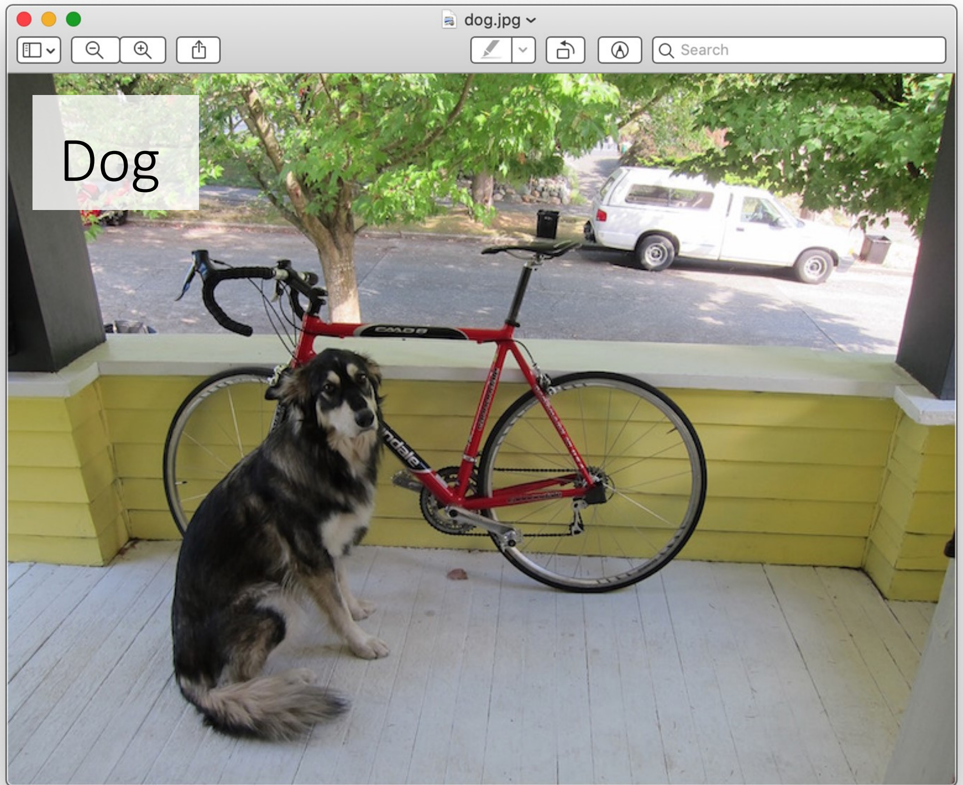
D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, TPAMI 2002

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Classification

- What is in the image?



[Redmon]

Tagging

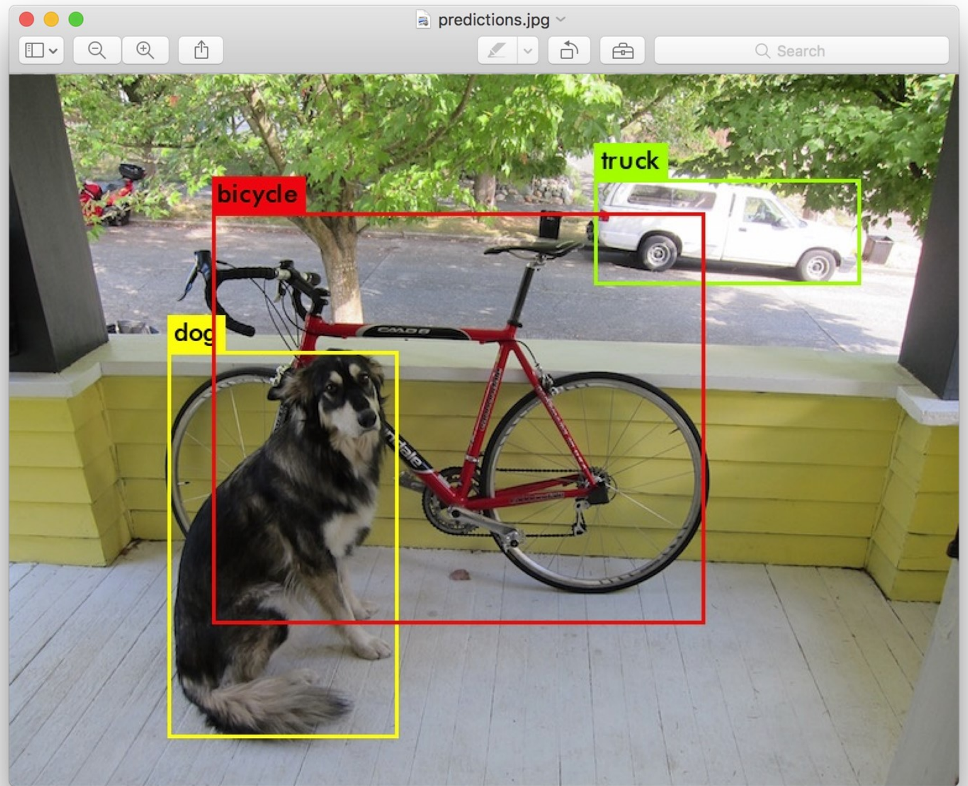
- What are ALL the things in the image?

[Redmon]

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

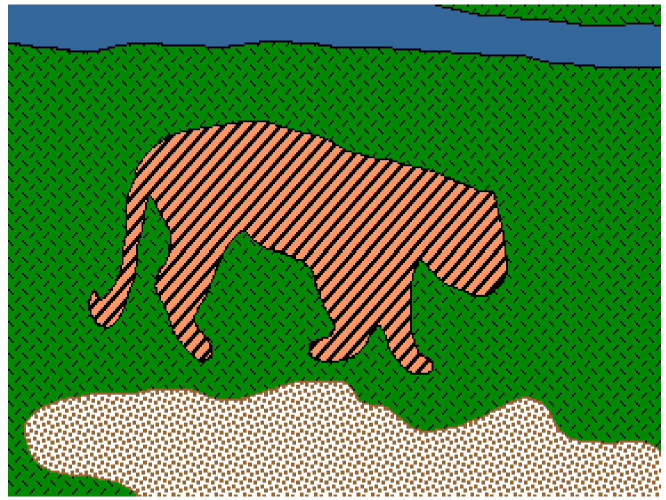
Detection

- What are ALL the things in the image?
- Where are they?



[Redmon]

Segmentation



Today’s Agenda

- Visual Recognition Tasks
- Introduction to segmentation and clustering
- Agglomerative clustering
- K-means clustering
- Mean-shift clustering
- Efficient Graph-based image segmentation

Reading: Forsyth Chapter 9

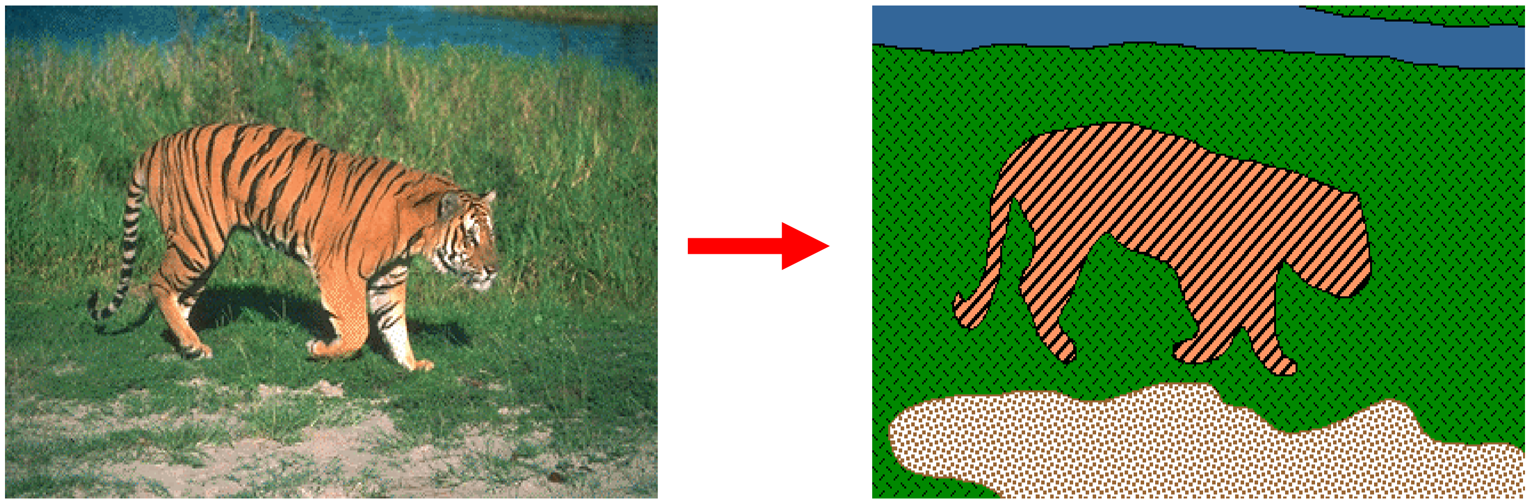
D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, TPAMI 2002

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Image Segmentation

Goal: identify groups of pixels that go together



Slide credit: Steve Seitz, Kristen Grauman

The Goals of Segmentation

- Separate image into coherent “objects”

Image | Human segmentation

Slide credit: Svetlana Lazebnik

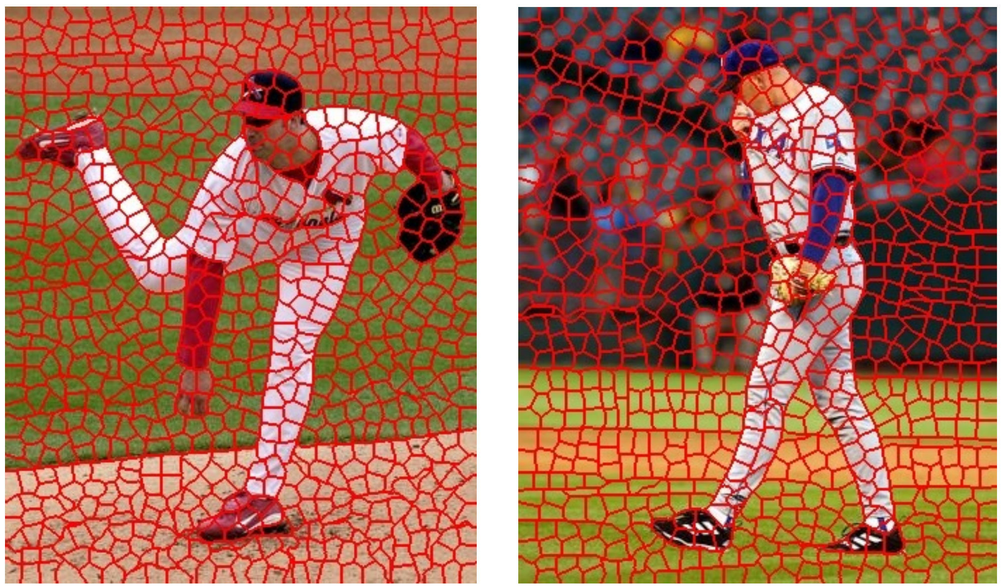
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

The Goals of Segmentation

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

“superpixels”



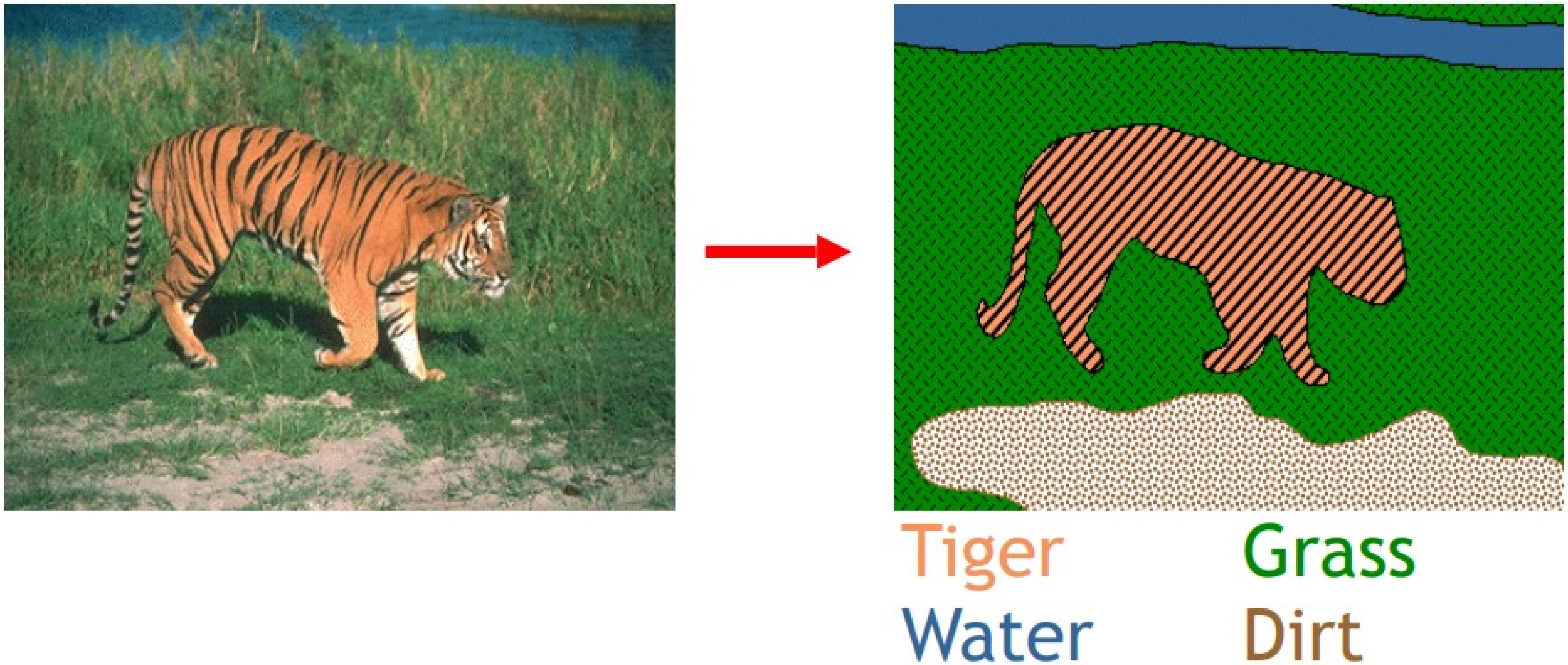
X. Ren and J. Malik. [Learning a classification model for segmentation.] ICCV 2003.

Slide credit: Svetlana Lazebnik

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Types of Segmentation

- Semantic segmentation: Assign labels



Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Types of Segmentation

- Semantic segmentation: Assign labels

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Types of Segmentation

- Instance segmentation: Assign labels per object

http://www.youtube.com/watch?v=OOT3UIXZztE

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

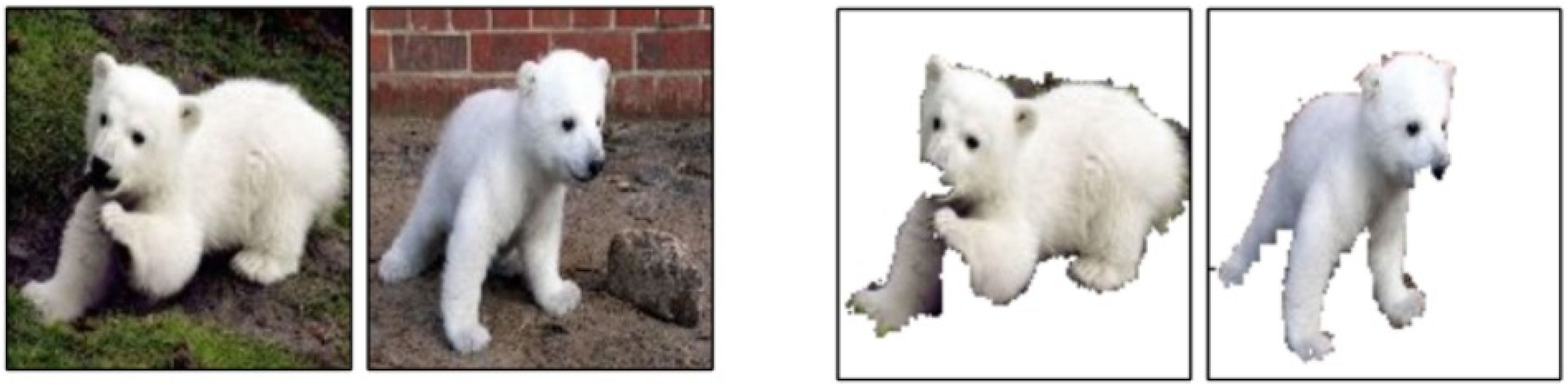
Types of Segmentation

- Foreground / background segmentation

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Types of Segmentation

- Co-segmentation: Segment common object in multiple images



Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

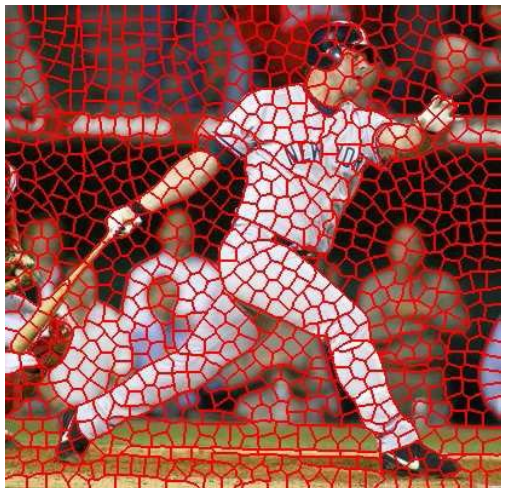
Application: as a result

GrabCut: Rother et al. 2004

Application: for efficiency – e.g., speed up recognition

[Felzenszwalb and Huttenlocher 2004]

[Hoiem et al. 2005, Mori 2005]



[Shi and Malik 2001]

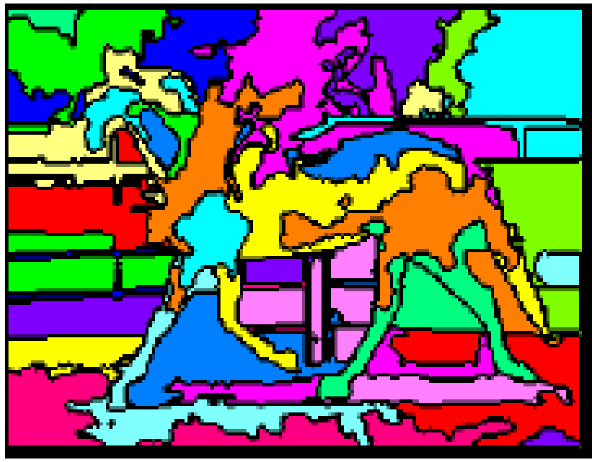
Slide: Derek Hoiem

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

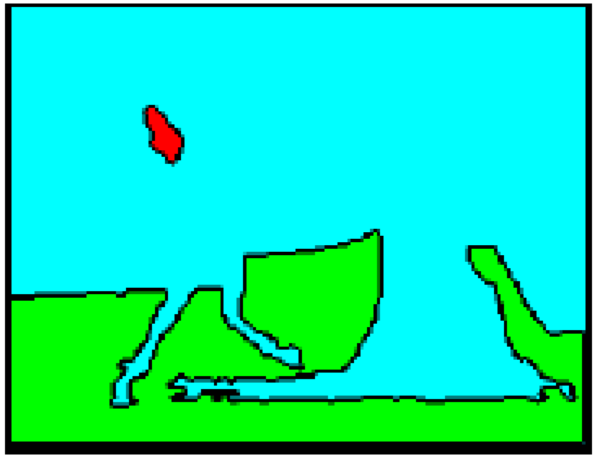
Application: better classification

Angelova and Zhu, 2013

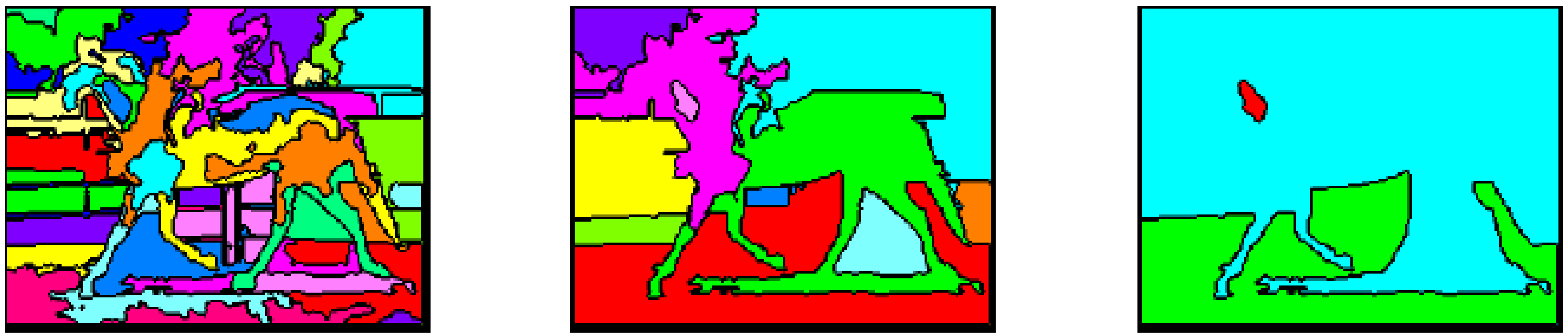
Over/under segmentation



Oversegmentation



Undersegmentation



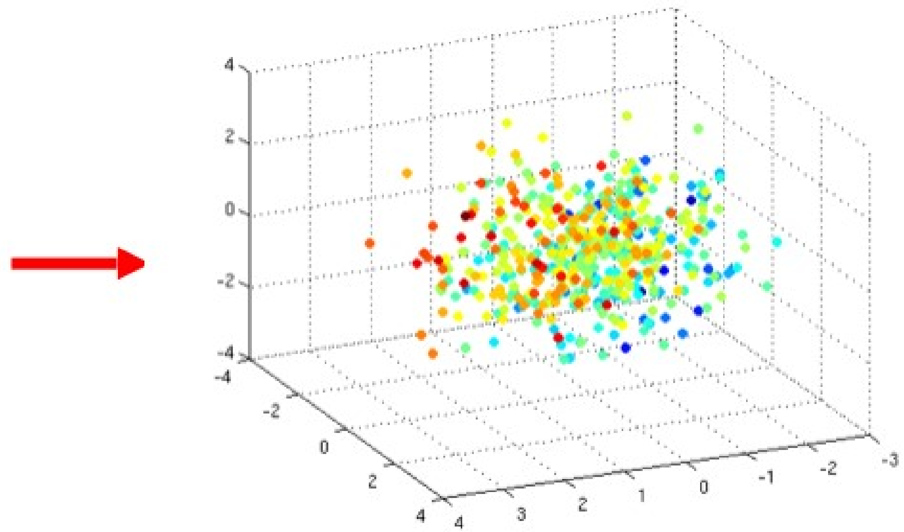
Multiple Segmentations

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

One way to think about segmentation is **Clustering**

- Pixels are points in a (high-dimensional) feature space, e.g.
  - color: 3D
  - color + location: 5D
- Cluster pixels into segments



One way to think about segmentation is **Clustering**

Clustering: group together similar data points and represent them as a single *entity*

Clustering is an unsupervised learning method

Key Challenges:

1) What makes two points/images/patches similar?

2) How do we compute an overall grouping from pairwise similarities?

Slide: Derek Hoiem

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Distance vs Similarity Measures

Let x and x’ be two objects from the dataset.
The distance or similarity between x and x’ is a real number, dist(x, x’) or sim(x, x’)

- The Euclidian distance is defined as

$$dist(x, x') = \sqrt{\sum (x_i - x'_i)^2}$$

- In contrast, cosine similarity measure would be

$$\begin{aligned} sim(x, x') &= \cos(\theta) \\ &= \frac{x^\top x'}{\|x\| \cdot \|x'\|} \\ &= \frac{x^\top x'}{\sqrt{x^\top x}\sqrt{x'^\top x'}}. \end{aligned}$$

Desirable Properties of a Clustering Algorithm

1. Scalability in terms of both time and space
2. Ability to deal with different data types
3. Minimal requirements for domain knowledge to determine algorithm parameters
   - Don't need to know how many objects there are or what those object categories will be.
4. Interpretability and usability are optional
   - Incorporation of user-specified constraints

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

General ideas

- Bottom-up clustering
  - pixels belong together because they are locally coherent
- Top-down clustering
  - pixels belong together because they lie on the same visual entity (object)

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Clustering algorithms

- Agglomerative clustering
- K-means
- Mean-shift clustering

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

- Visual Recognition Tasks
- Introduction to segmentation and clustering
- Agglomerative clustering
- K-means clustering
- Mean-shift clustering
- Efficient Graph-based image segmentation

Reading: Forsyth Chapter 9

D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, TPAMI 2002

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

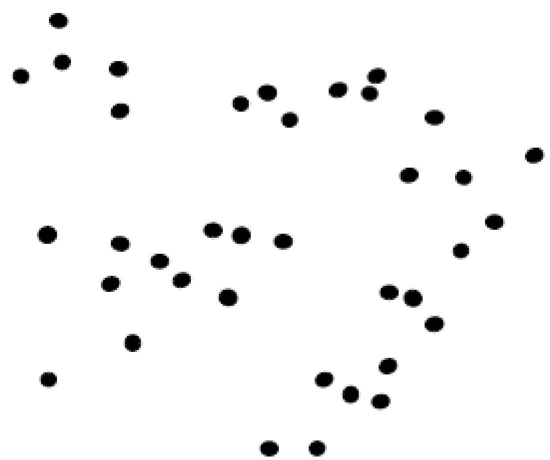
Agglomerative Hierarchical Clustering - Algorithm

1. Initially each item $x_1, \ldots, x_n$ is in its own cluster $C_1, \ldots, C_n$.
2. Repeat until there is only one cluster left:
3. Merge the nearest clusters, say C_i and C_j.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Agglomerative Hierarchical Clustering - Algorithm

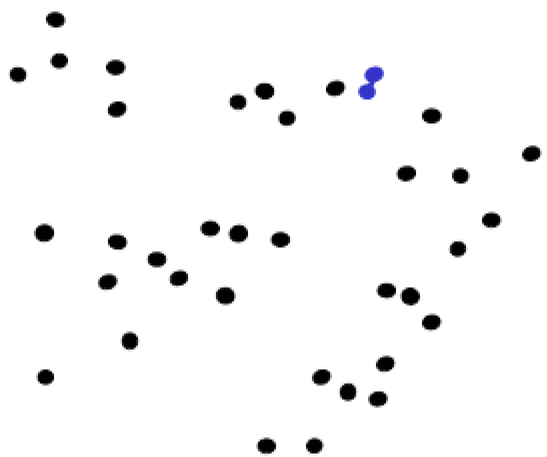


1. Say "Every point is its own cluster"

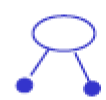
Slide credit: Andrew Moore

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Agglomerative Hierarchical Clustering - Algorithm



1. Say "Every point is its own cluster"
2. Find "most similar" pair of clusters

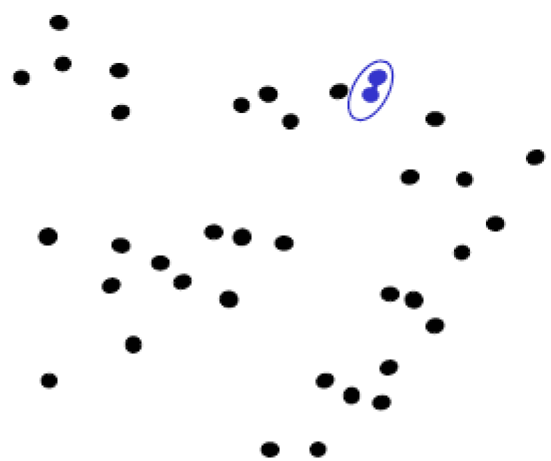


Slide credit: Andrew Moore

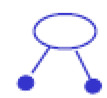
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Agglomerative Hierarchical Clustering - Algorithm



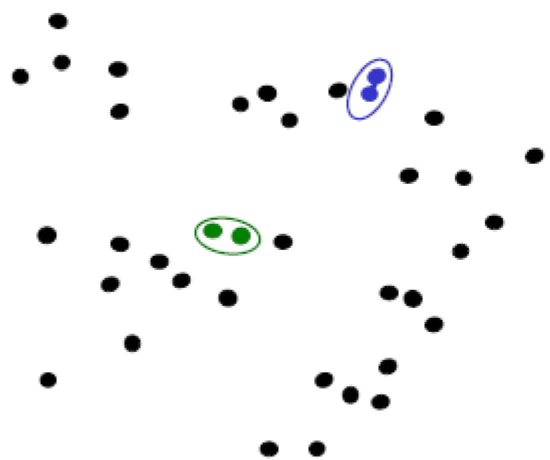
1. Say "Every point is its own cluster"
2. Find "most similar" pair of clusters
3. Merge it into a parent cluster



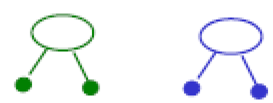
Slide credit: Andrew Moore

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Agglomerative Hierarchical Clustering - Algorithm



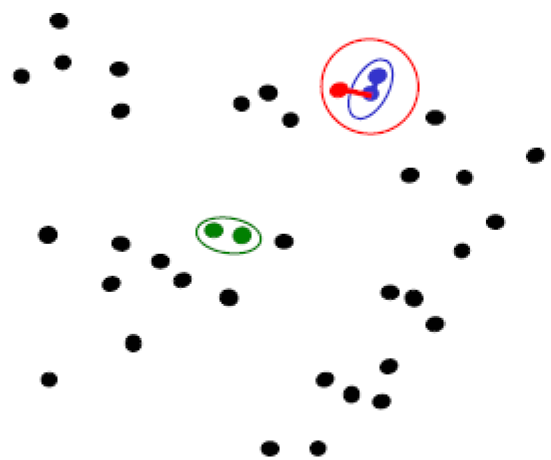
1. Say "Every point is its own cluster"
2. Find "most similar" pair of clusters
3. Merge it into a parent cluster
4. Repeat



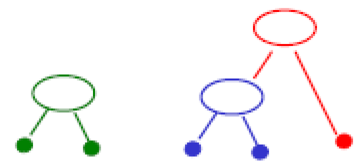
Slide credit: Andrew Moore

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Agglomerative Hierarchical Clustering - Algorithm

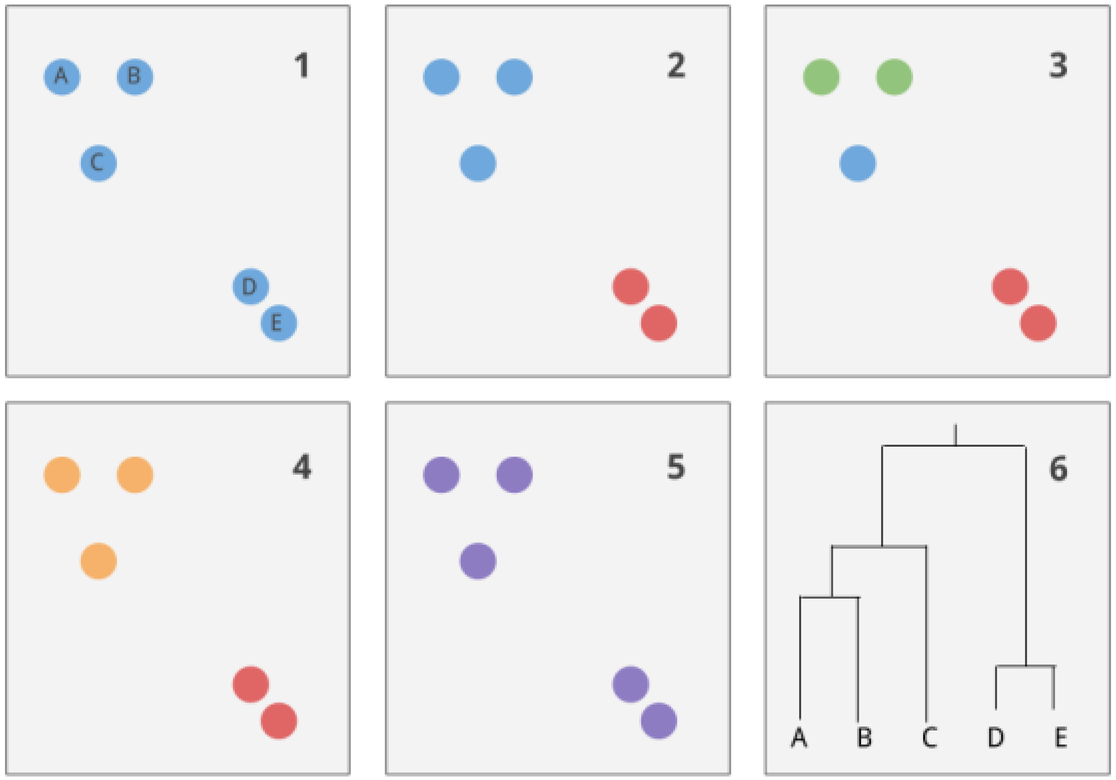


1. Say "Every point is its own cluster"
2. Find "most similar" pair of clusters
3. Merge it into a parent cluster
4. Repeat



Slide credit: Andrew Moore

Agglomerative clustering example



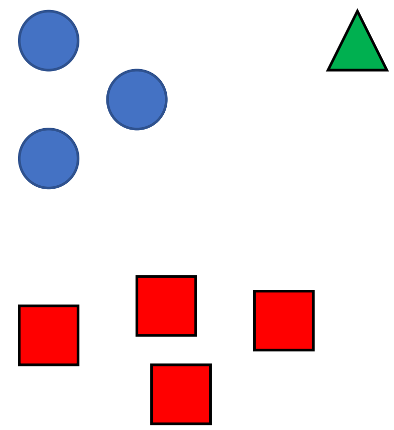
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Agglomerative clustering

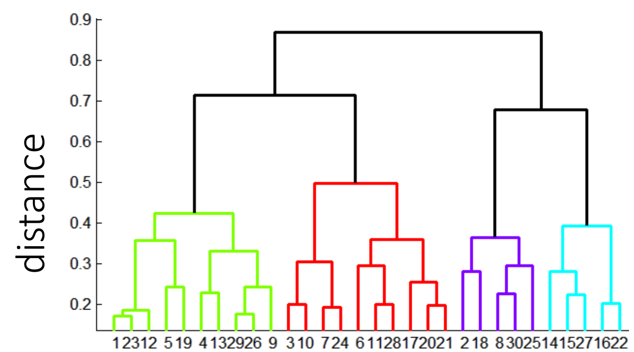
How to define cluster similarity?

- average distance between points
- maximum distance
- minimum distance
- distance between means or medoids



How many clusters?

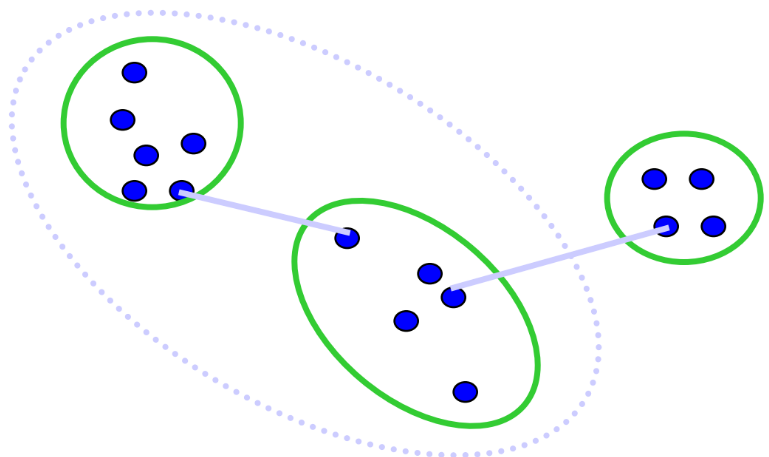
- Clustering creates a dendrogram (a tree)
- Threshold based on max number of clusters or based on distance between merges



Different measures of nearest clusters

Single Link

- Distance between clusters is the minimum distance between their points



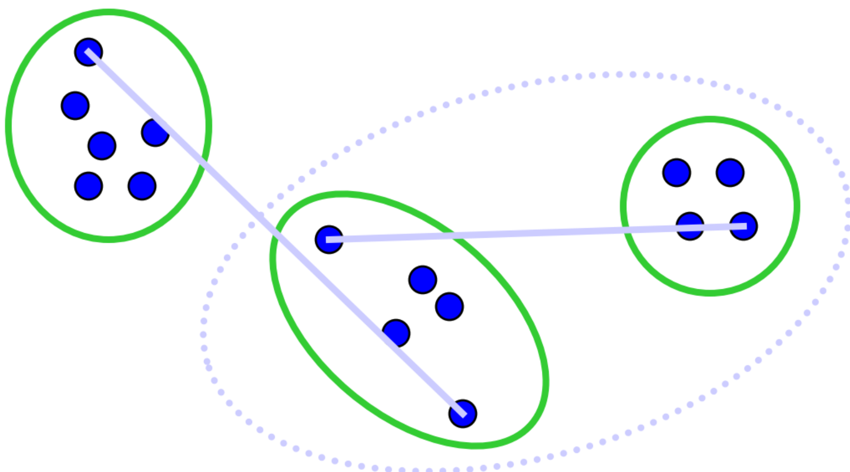
Long, skinny clusters

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Different measures of nearest clusters

Complete Link

- Distance between clusters is the maximum distance between their points



Tight clusters

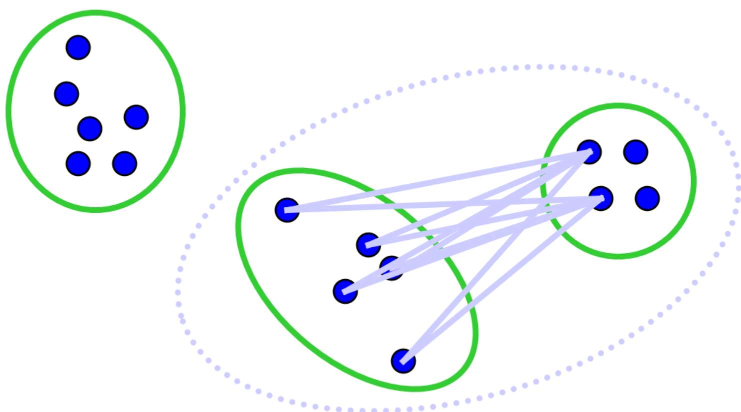
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Different measures of nearest clusters

Average Link

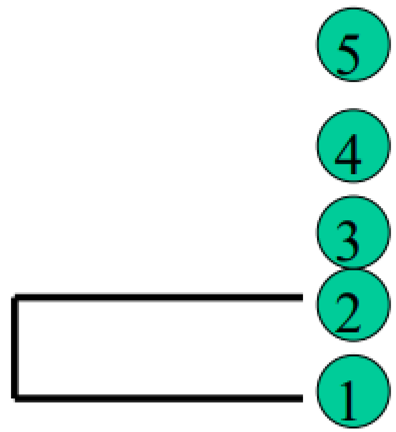
- Distance between clusters is the average distance between their points



Robust against noise

Example – single link

	1	2	3	4	5
1	0				
2	2	0			
3	6	3	0		
4	10	9	7	0	
5	9	8	5	4	0



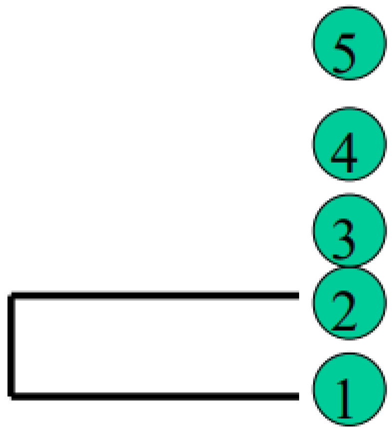
Example – single link

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

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

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

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

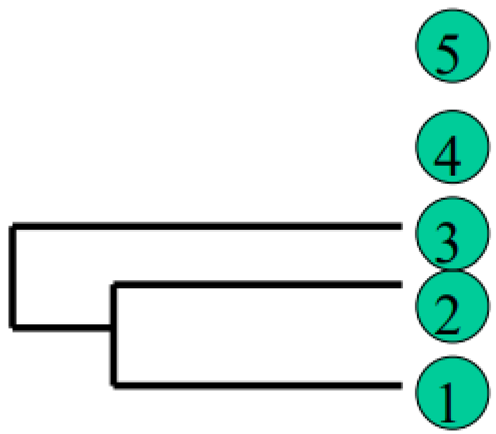


Example – single link

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

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

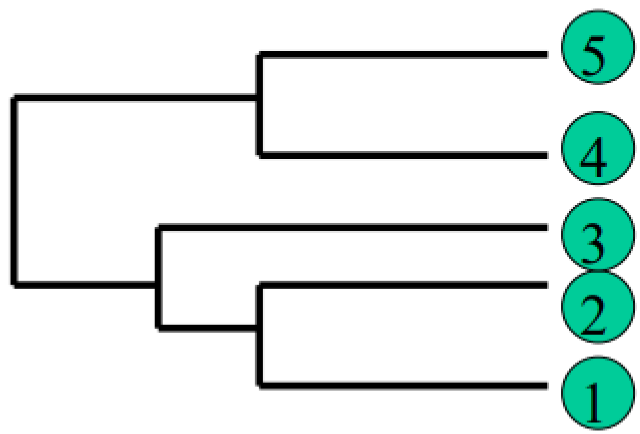
$$d_{(1,2,3),5} = \min\{d_{(1,2),5}, d_{3,5}\} = \min\{8,5\} = 5$$



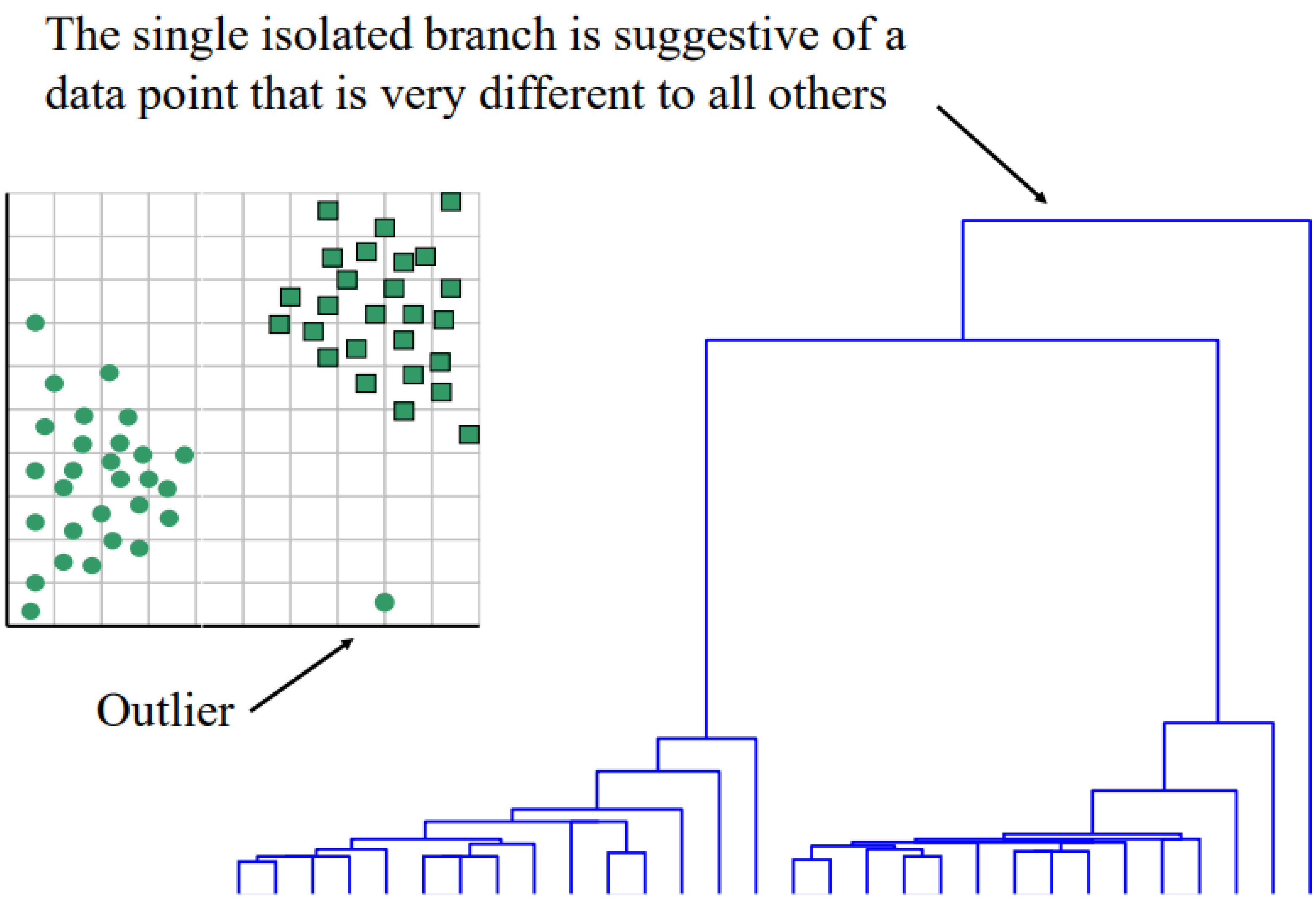
Example – single link

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

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



Outliers



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Conclusions: Agglomerative Clustering

Good

- Simple to implement, widespread application.
- Clusters have adaptive shapes.
- Provides a hierarchy of clusters.
- Can avoid specifying number of clusters in advance.

Bad

- May have imbalanced clusters.
- Still have to choose number of clusters or threshold to use them.
- Does not scale well. Runtime of $O(n^3)$.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

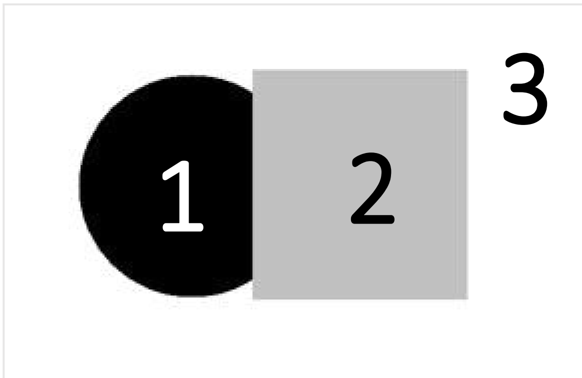
- Visual Recognition Tasks
- Introduction to segmentation and clustering
- Agglomerative clustering
- K-means clustering
- Mean-shift clustering
- Efficient Graph-based image segmentation

Reading: Forsyth Chapter 9

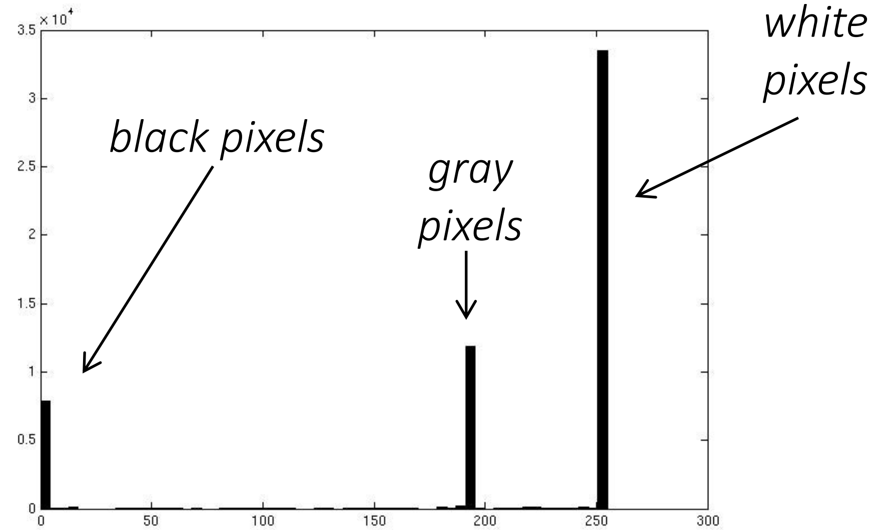
D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, TPAMI 2002

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Image Segmentation: Toy Example



input image

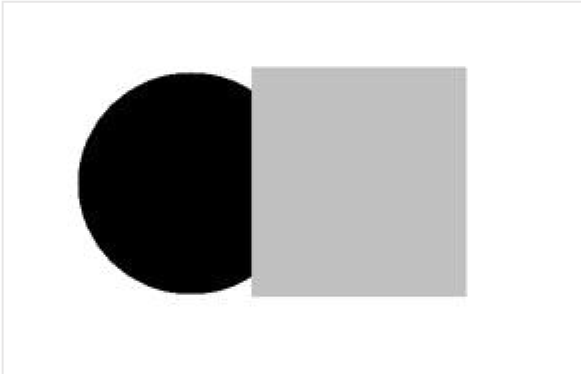


- These intensities define the three groups.
- We could label every pixel in the image according to which of these primary intensities it is.
  - i.e., segment the image based on the image intensity feature.
- What if the image isn't quite so simple?

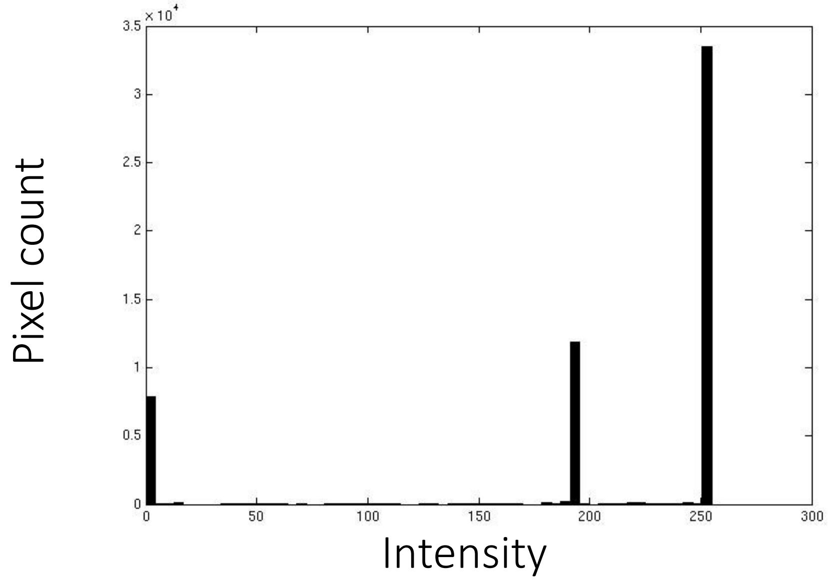
Slide credit: Kristen Grauman

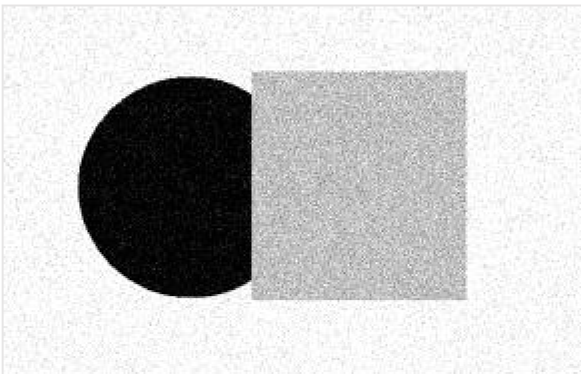
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

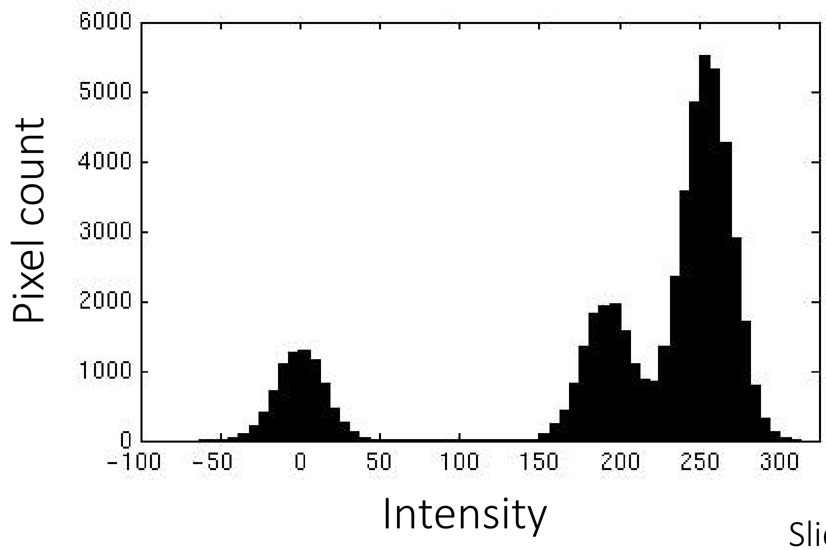


Input image

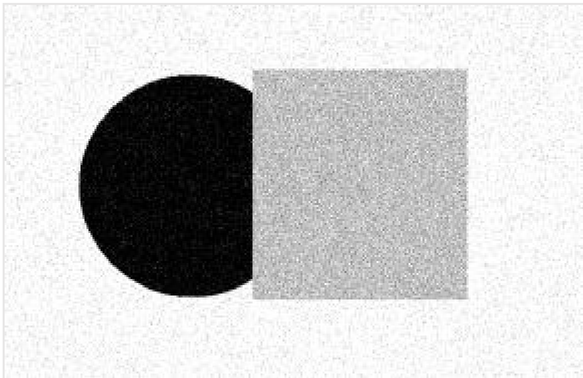




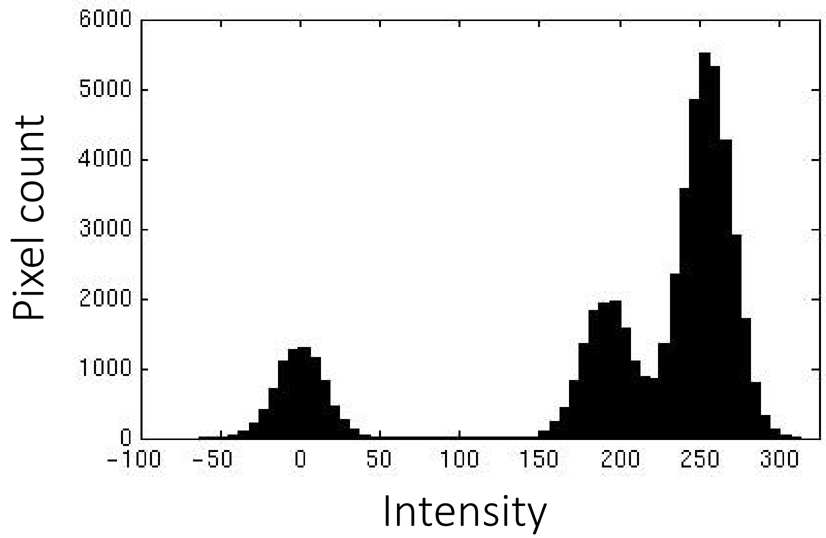
Input image



Slide credit: Kristen Grauman

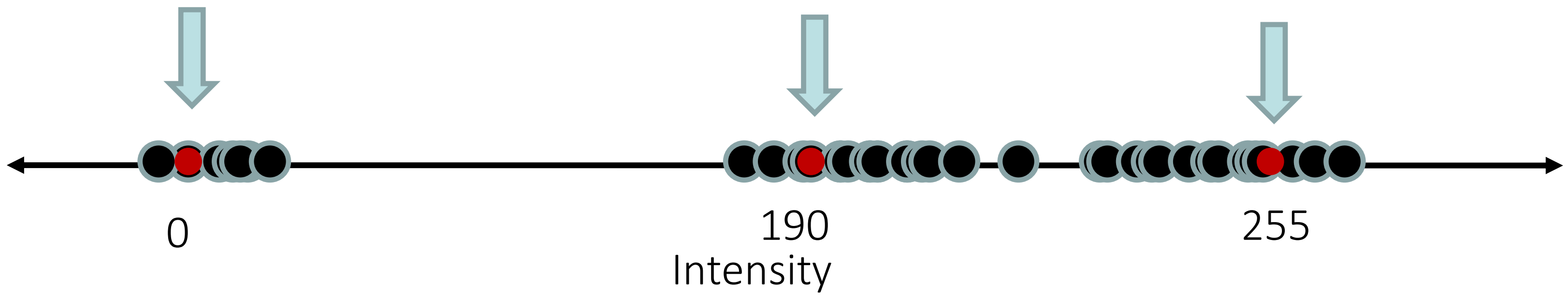


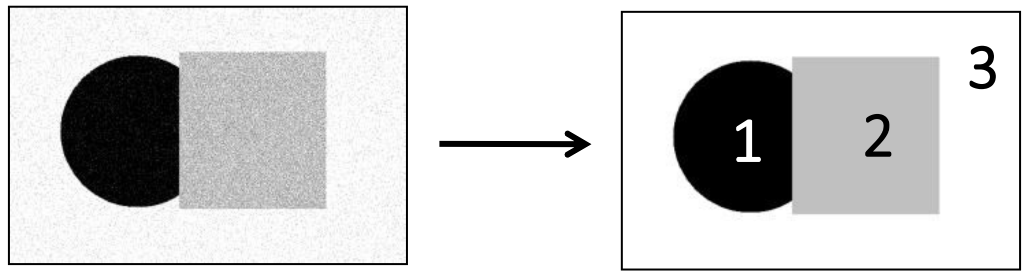
Input image



- Now how to determine the three main intensities that define our groups?
- We need to cluster

Slide credit: Kristen Grauman





- Goal: choose three “centers” as the representative intensities and label every pixel according to which of these centers it is nearest to.
- Best cluster centers are those that minimize Sum of Square Distance (SSD) between all points and their nearest cluster center c_i:

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

Slide credit: Kristen Grauman

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Clustering for Summarization

Goal: cluster to minimize variance in data, given clusters

$$c^*,\ \delta^* = \arg\min_{c,\delta} \frac{1}{N} \sum_{j}^{N} \sum_{i}^{K} \delta_{ij} \left(c_i - x_j\right)^2$$

Cluster center → c_i

Data → x_j

Whether x_j is assigned to c_i → δ_{ij}

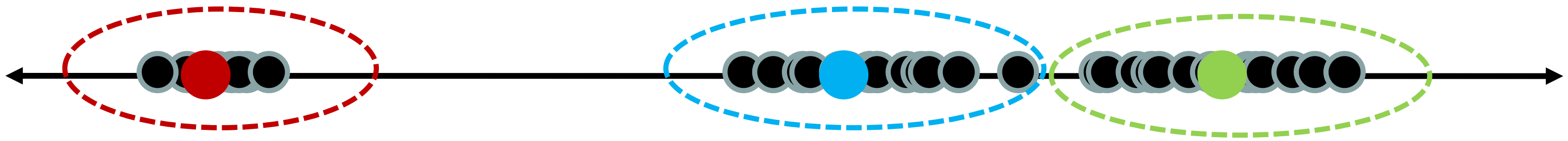
Slide credit: Derek Hoiem

Co-financed by the European Union
Connecting Europe Facility

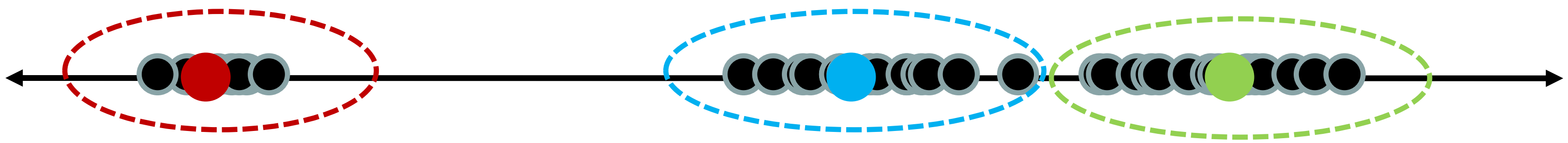
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Clustering

- With this objective, it is a “chicken and egg” problem:
  - If we knew the *cluster centers*, we could allocate points to groups by assigning each to its closest center.



  - If we knew the *group memberships*, we could get the centers by computing the mean per group.



Slide credit: Kristen Grauman

K-means clustering

1. Initialize ($t = 0$): cluster centers $c_1, \ldots, c_K$

2. Compute δ^t : assign each point to the closest center

- δ^t denotes the set of assignments for each x_j to cluster c_i at iteration t

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

3. Compute c^t: update cluster centers as the mean of the points

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

4. Update $t = t + 1$, Repeat Step 2-3 till stopped

Slide credit: Derek Hoiem

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

K-means clustering

1. Initialize ($t = 0$): cluster centers $c_1, \ldots, c_K$
   - Commonly used: random initialization
   - Or greedily choose K to minimize residual
2. Compute δ^t: assign each point to the closest center
   - δ^t denotes the set of assignments for each x_j to cluster c_i at iteration t
   - Typical distance measure:
     - Euclidean
     - Cosine

$$\delta^t = \underset{\delta}{\operatorname{argmin}} \frac{1}{N} \sum_j^N \sum_i^K \delta_{ij}^{t-1} \left(c_i^{t-1} - x_j \right)^2$$

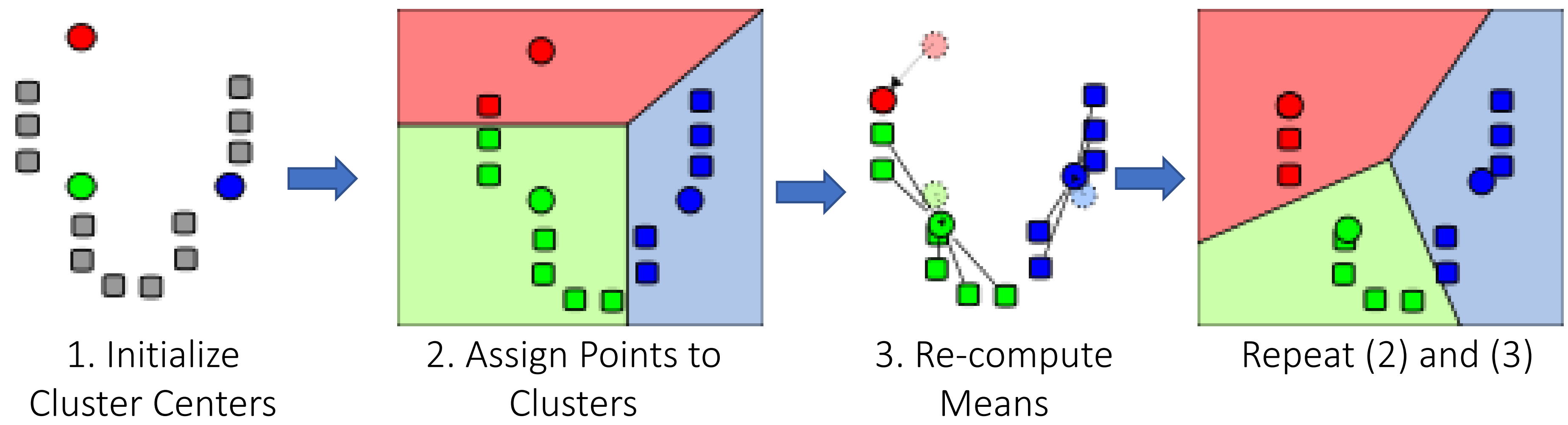
3. Compute c^t: update cluster centers as the mean of the points

$$c^t = \underset{c}{\operatorname{argmin}} \frac{1}{N} \sum_j^N \sum_i^K \delta_{ij}^{t} \left(c_i^{t-1} - x_j \right)^2$$

4. Update $t = t + 1$, Repeat Step 2-3 till stopped
   - c^t doesn't change anymore.

Slide credit: Derek Hoiem

K-means clustering



Demo

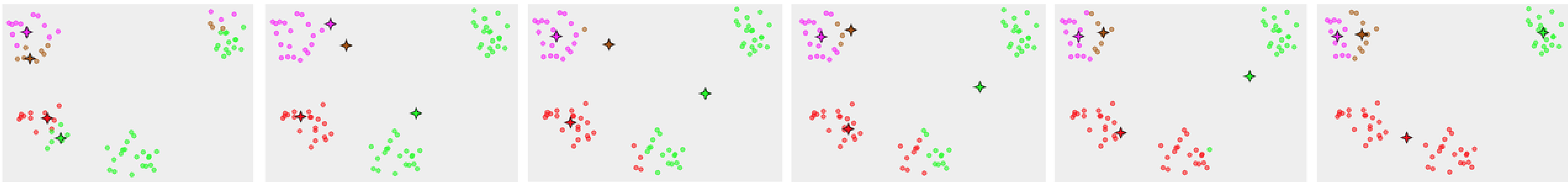
https://stanford.edu/class/engr108/visualizations/kmeans/kmeans.html

Illustration Source: wikipedia

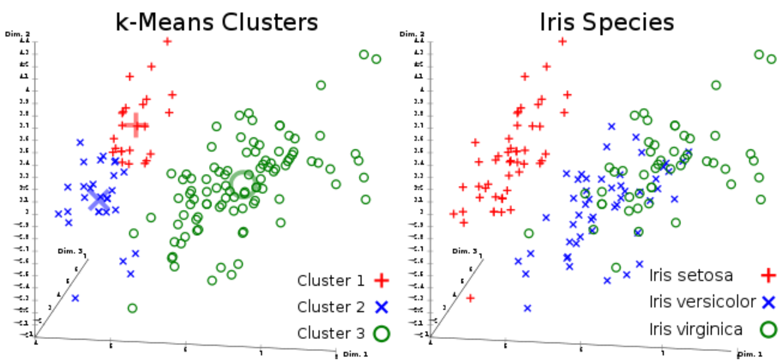
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

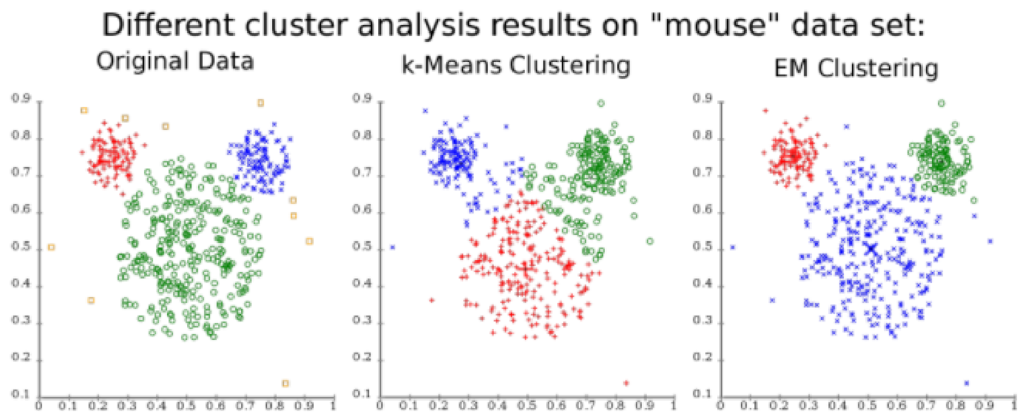
K-means clustering

- Converges to a *local minimum* solution
    - Initialize multiple runs



- Better fit for spherical data





- Need to pick K (# of clusters)

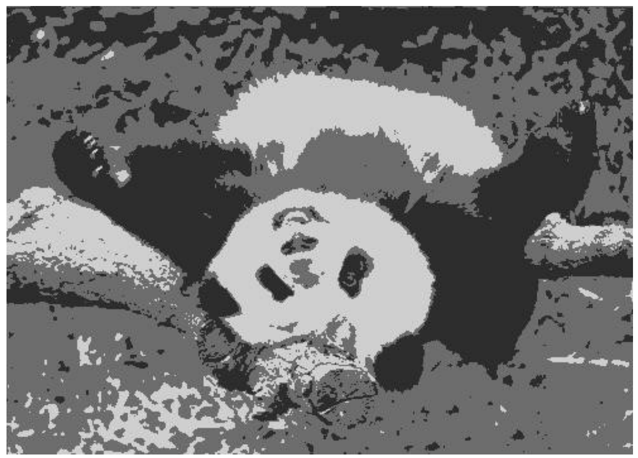
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Segmentation as Clustering

Original image

2 clusters



3 clusters

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Feature Space

- Depending on what we choose as the *feature space*, we can group pixels in different ways.

- Grouping pixels based on intensity similarity

- Feature space: intensity value (1D)

Slide credit: Kristen Grauman

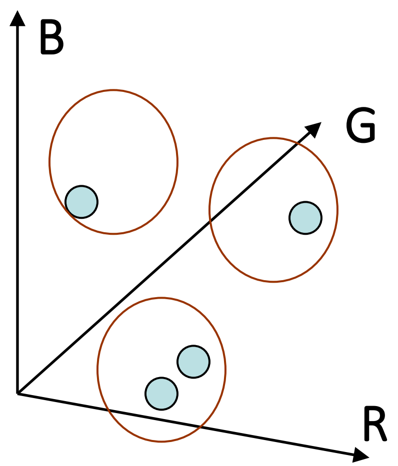
Co-financed by the European Union
Connecting Europe Facility

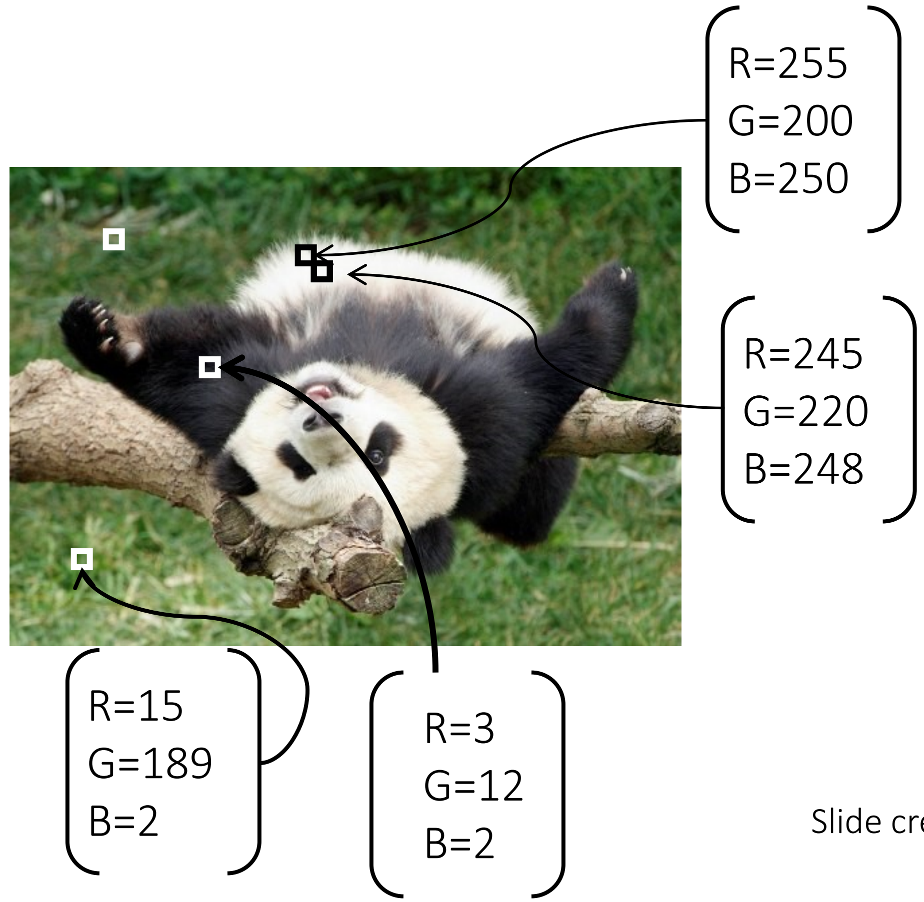
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Feature Space

- Depending on what we choose as the *feature space*, we can group pixels in different ways.

- Grouping pixels based on color similarity





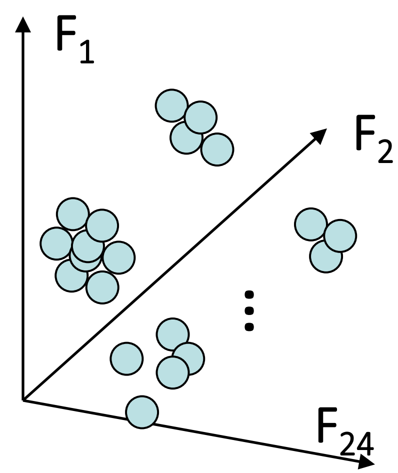
- Feature space: color value (3D)

Slide credit: Kristen Grauman

Feature Space

- Depending on what we choose as the *feature space*, we can group pixels in different ways.

- Grouping pixels based on texture similarity



Filter bank of
24 filters

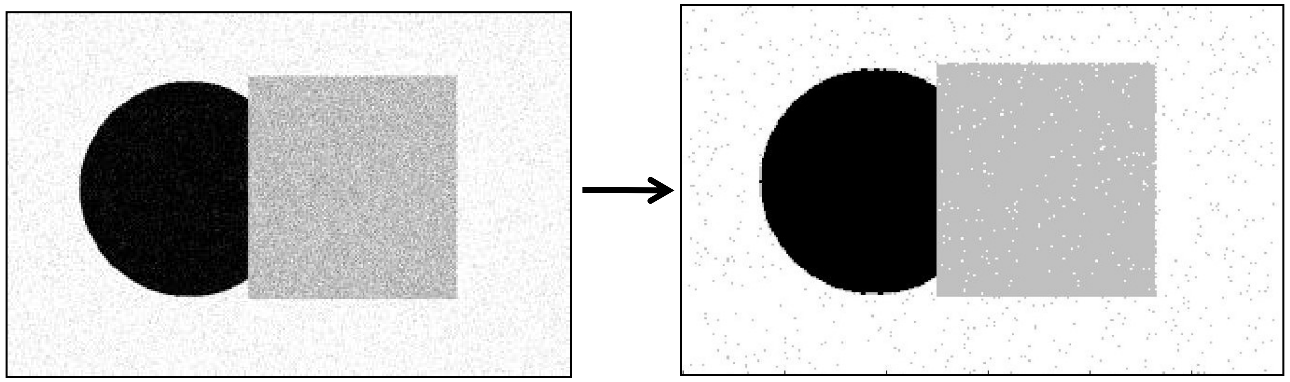
- Feature space: filter bank responses (e.g., 24D)

Slide credit: Kristen Grauman

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

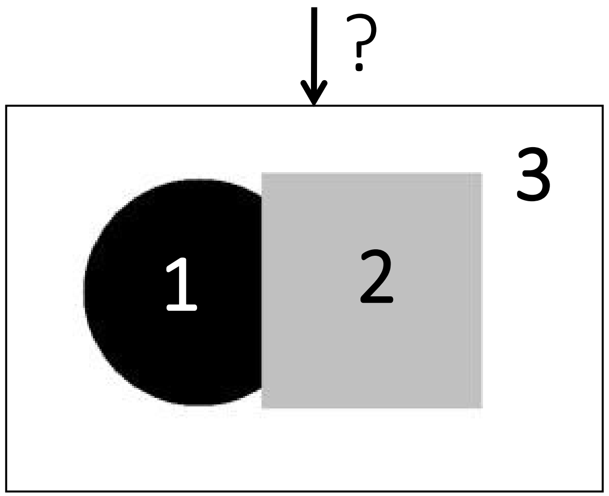
Smoothing Out Cluster Assignments

- Assigning a cluster label per pixel may yield outliers:



Original

Labeled by cluster center's intensity



- How can we ensure they are spatially smooth?

Slide credit: Kristen Grauman

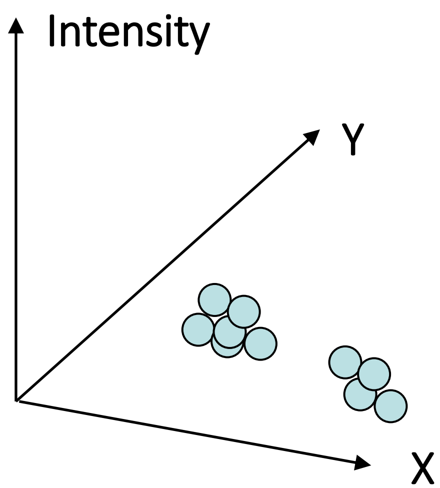
Co-financed by the European Union
Connecting Europe Facility

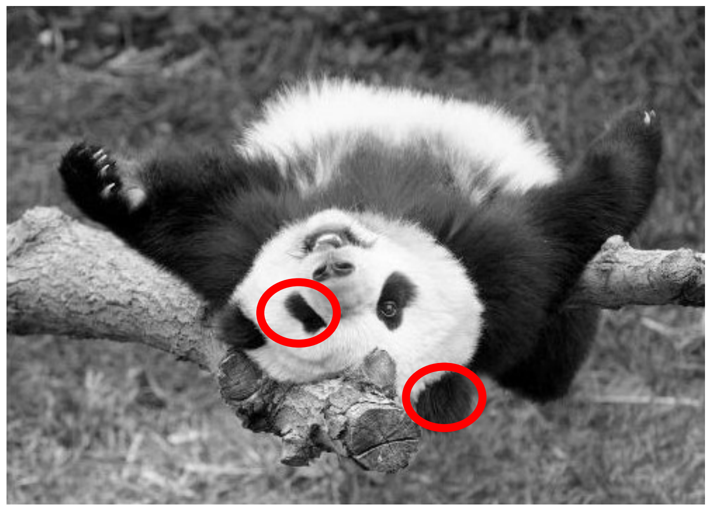
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Segmentation as Clustering

- Depending on what we choose as the *feature space*, we can group pixels in different ways.

- Grouping pixels based on *intensity+position* similarity





⇒ Way to encode both *similarity* and *proximity.*

Slide credit: Kristen Grauman

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

K-Means Clustering Results

- K-means clustering based on intensity or color is essentially vector quantization of the image attributes
  - Clusters don't have to be spatially coherent

Image | Intensity-based clusters | Color-based clusters

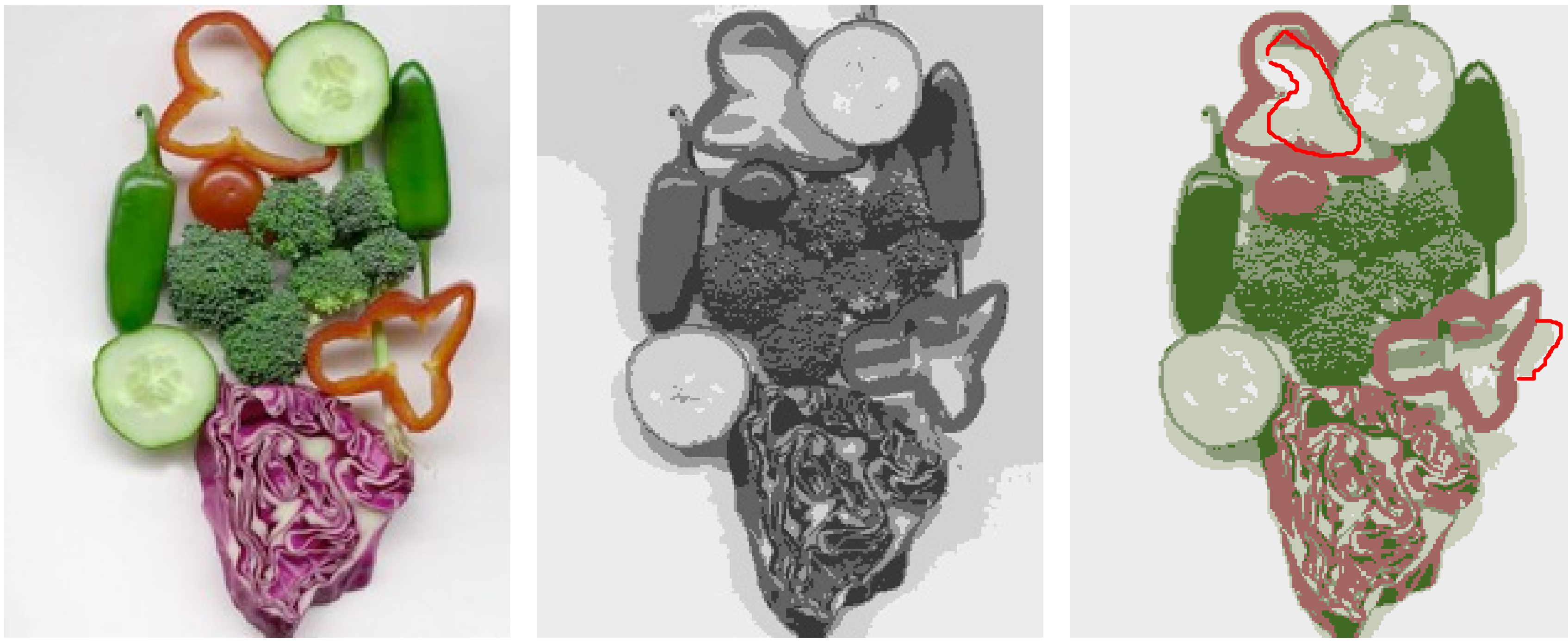


Image source: Forsyth & Ponce

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

K-Means Clustering Results

- K-means clustering based on intensity or color is essentially vector quantization of the image attributes
  - Clusters don't have to be spatially coherent
- Clustering based on (r,g,b,x,y) values enforces more spatial coherence

Co-financed by the European Union
Connecting Europe Facility

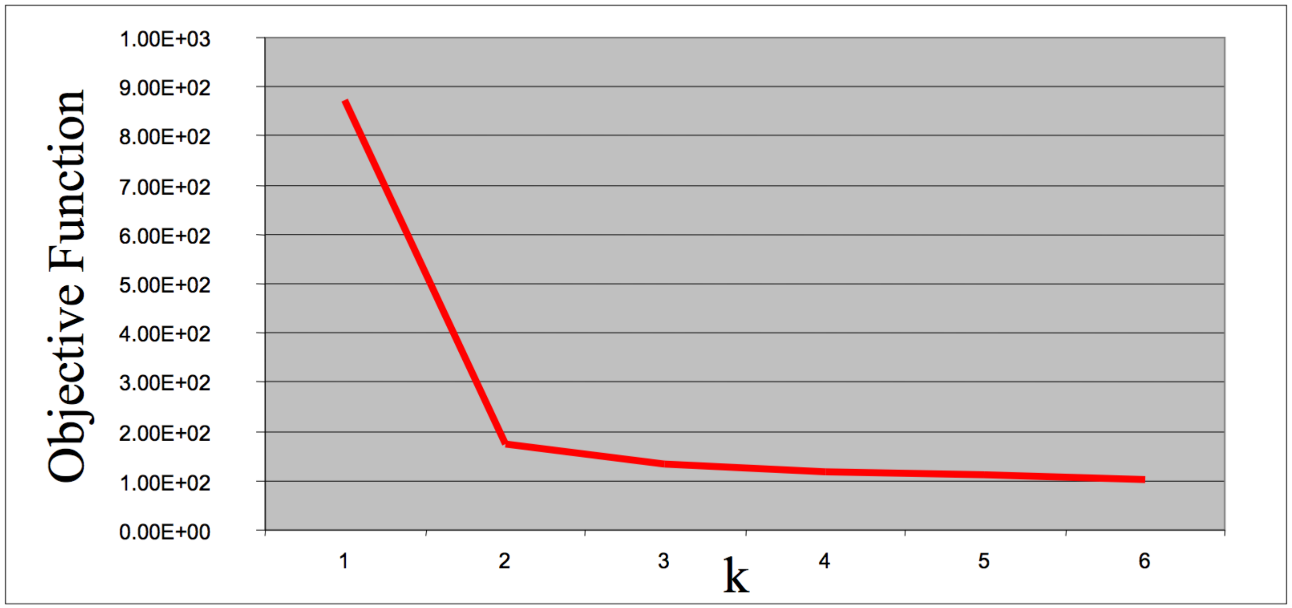
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

How to choose the number of clusters?

Try different numbers of clusters in a validation set and look at performance.

We can plot the objective function values for k equals 1 to 6…

The abrupt change at k = 2, is highly suggestive of two clusters in the data. This technique for determining the number of clusters is known as “knee finding” or “elbow finding”.



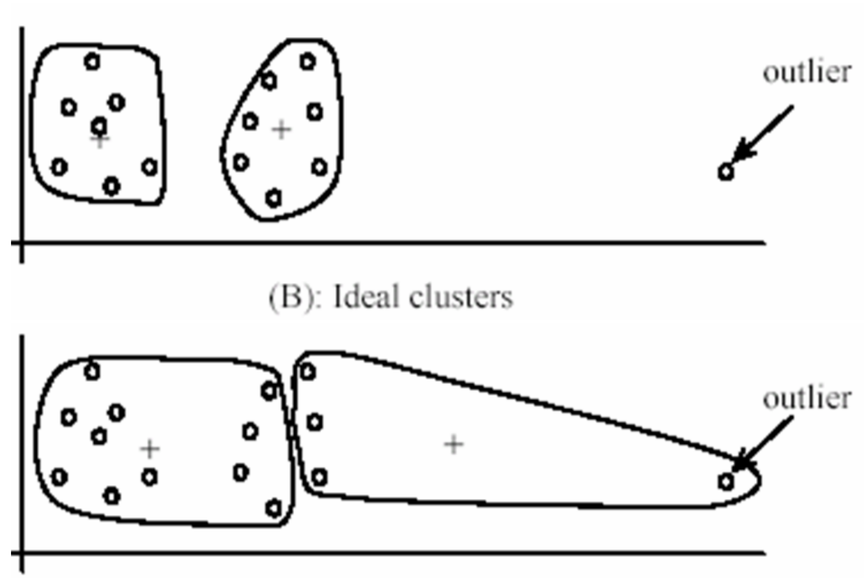
Slide credit: Derek Hoiem

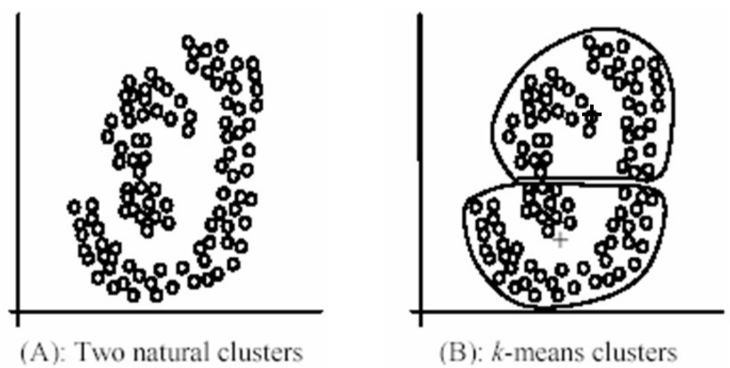
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

K-Means pros and cons

- Pros
  - Finds cluster centers that minimize conditional variance (good representation of data)
  - Simple and fast, Easy to implement
- Cons
  - Need to choose K
  - Sensitive to outliers
  - Prone to local minima
  - All clusters have the same parameters (e.g., distance measure is non-adaptive)
  - Distance computation in N-dimensional space could be slow





Today's Agenda

- Visual Recognition Tasks
- Introduction to segmentation and clustering
- Agglomerative clustering
- K-means clustering
- Mean-shift clustering
- Efficient Graph-based image segmentation

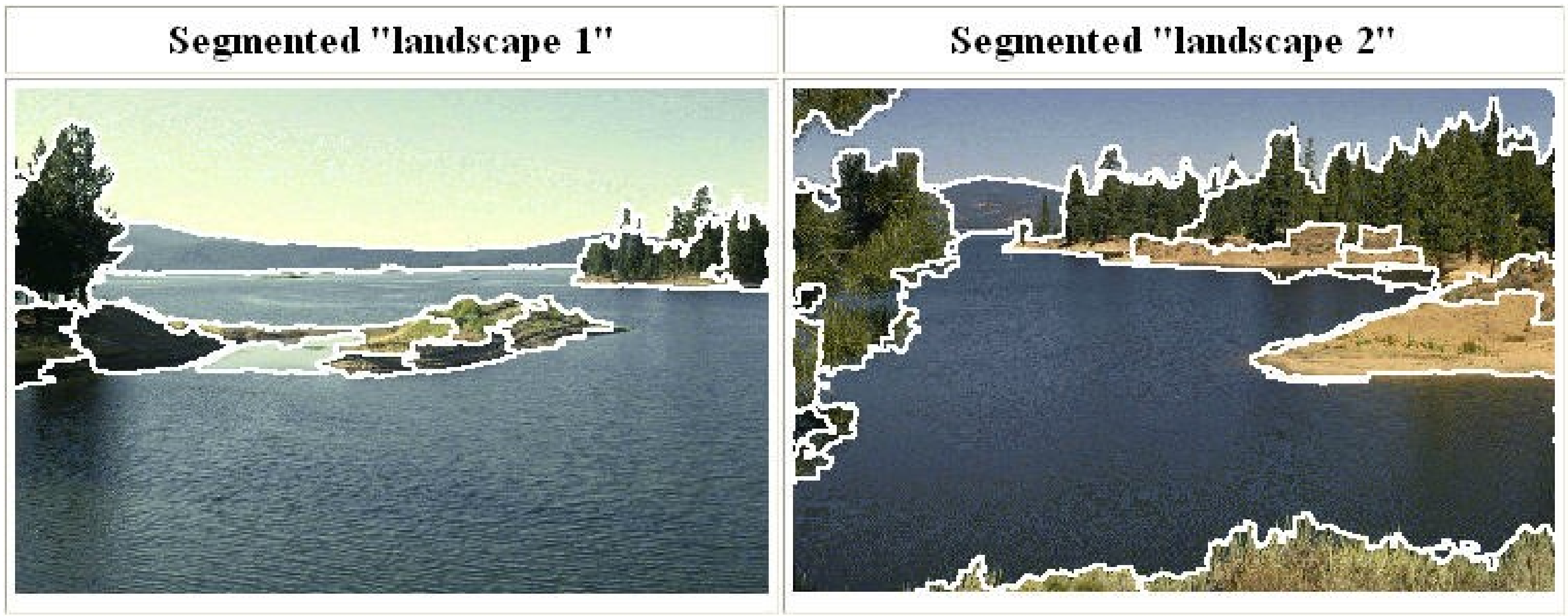
Reading: Forsyth Chapter 9

D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, TPAMI 2002

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Mean-Shift Segmentation

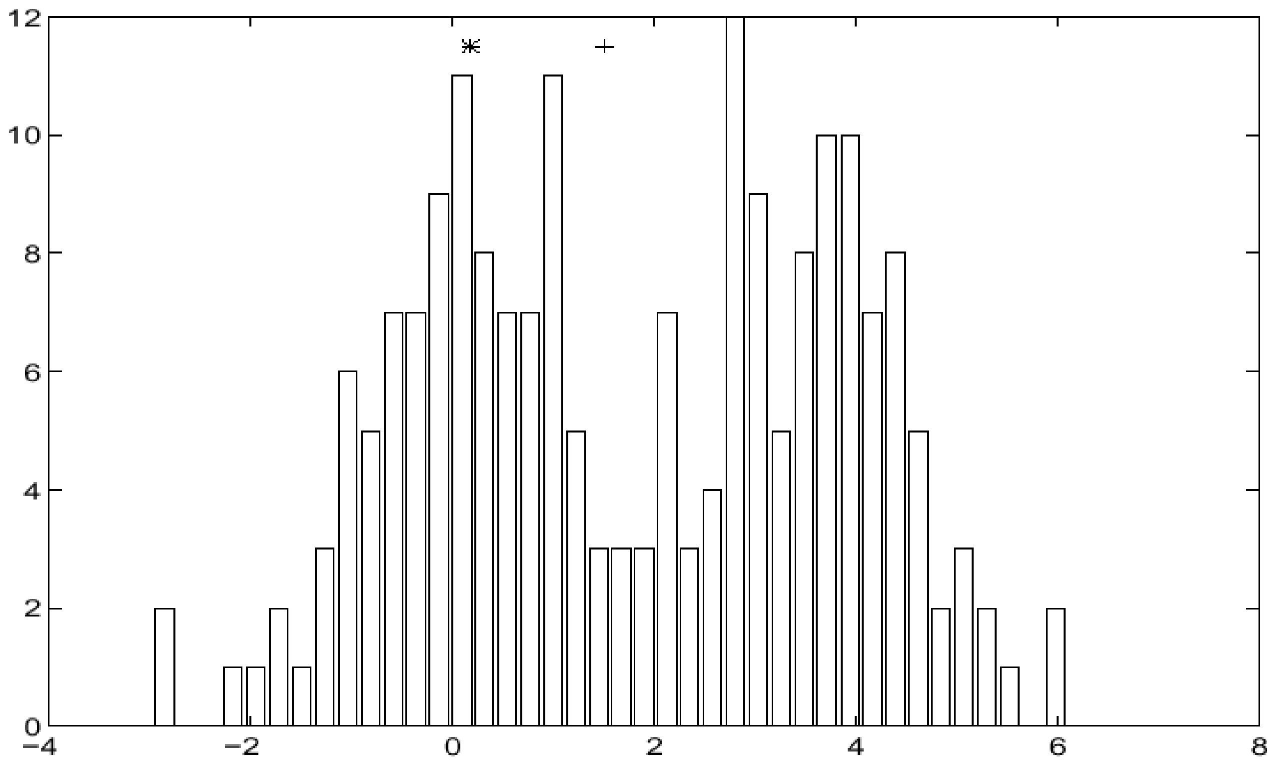
- An advanced and versatile technique for clustering-based segmentation



D. Comaniciu and P. Meer, [Mean Shift: A Robust Approach toward Feature Space Analysis](), TPAMI 2002

Slide credit: Svetlana Lazebnik

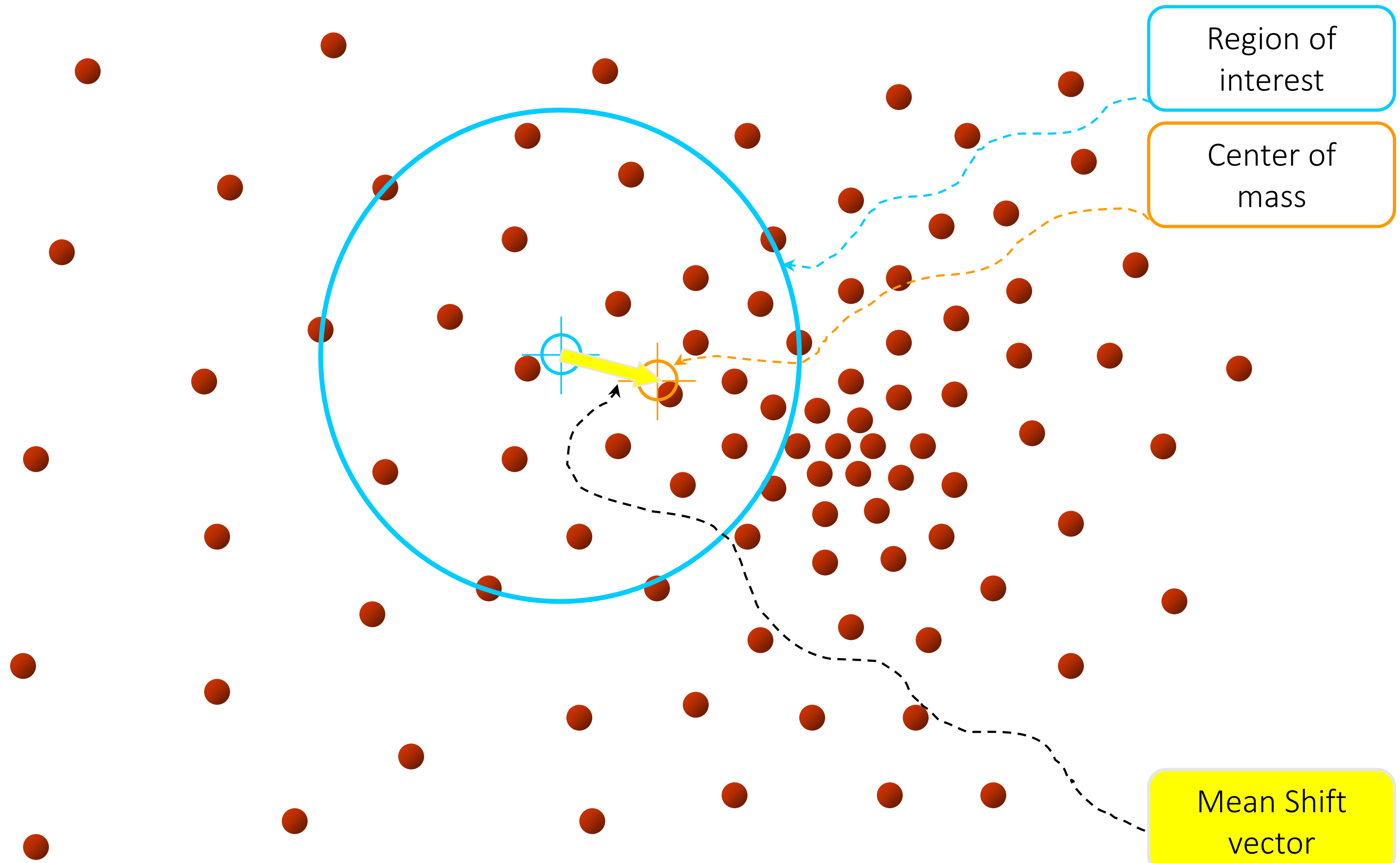
Mean-Shift Algorithm



- Iterative Mode Search
  1. Initialize random seed, and window W
  2. Calculate center of gravity (the "mean") of W: $\sum_{x \in W} xH(x)$
  3. Shift the search window to the mean
  4. Repeat Step 2 until convergence

Slide credit: Steve Seitz

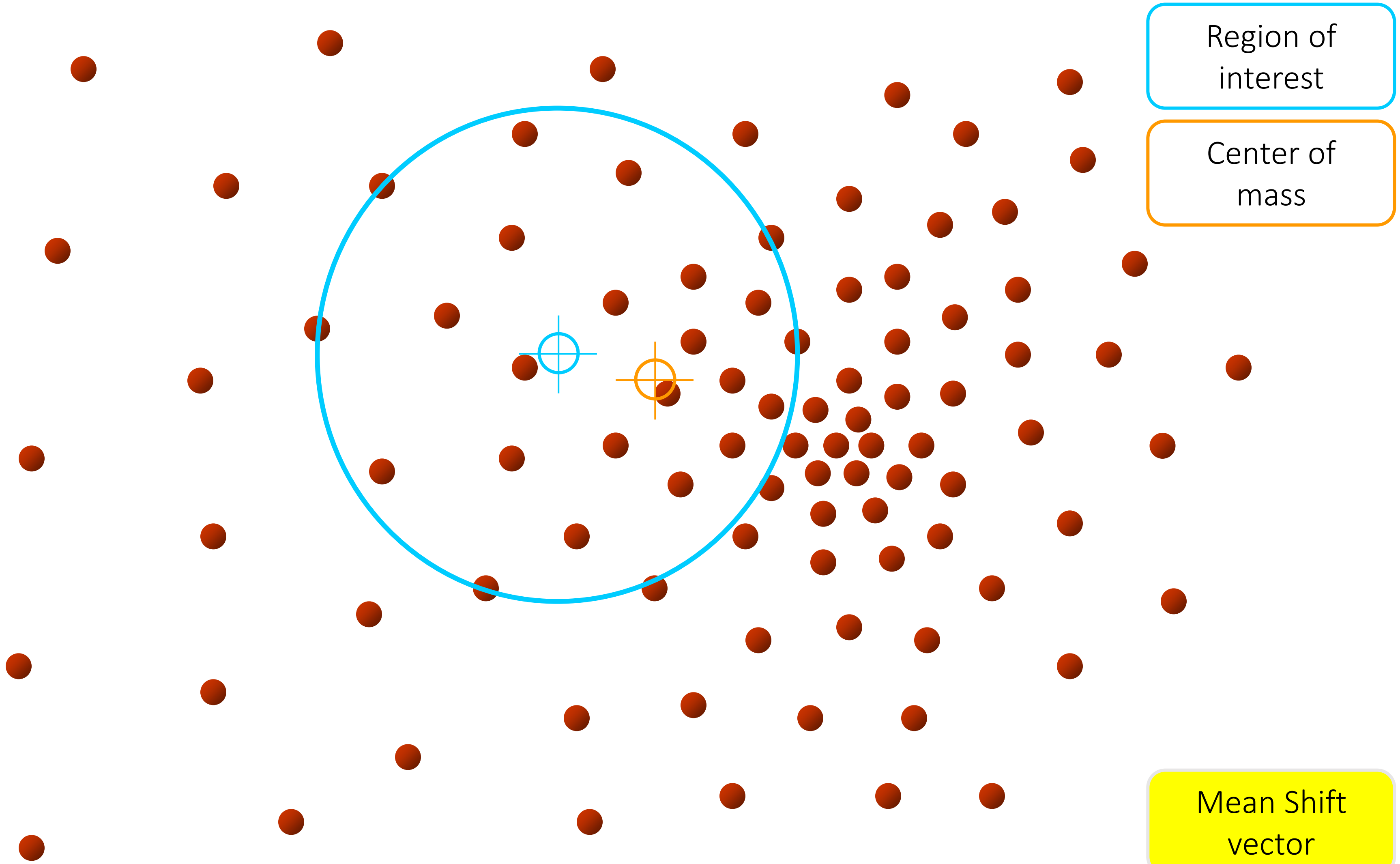
Mean-Shift



Slide credit: Y. Ukrainitz & B. Sarel

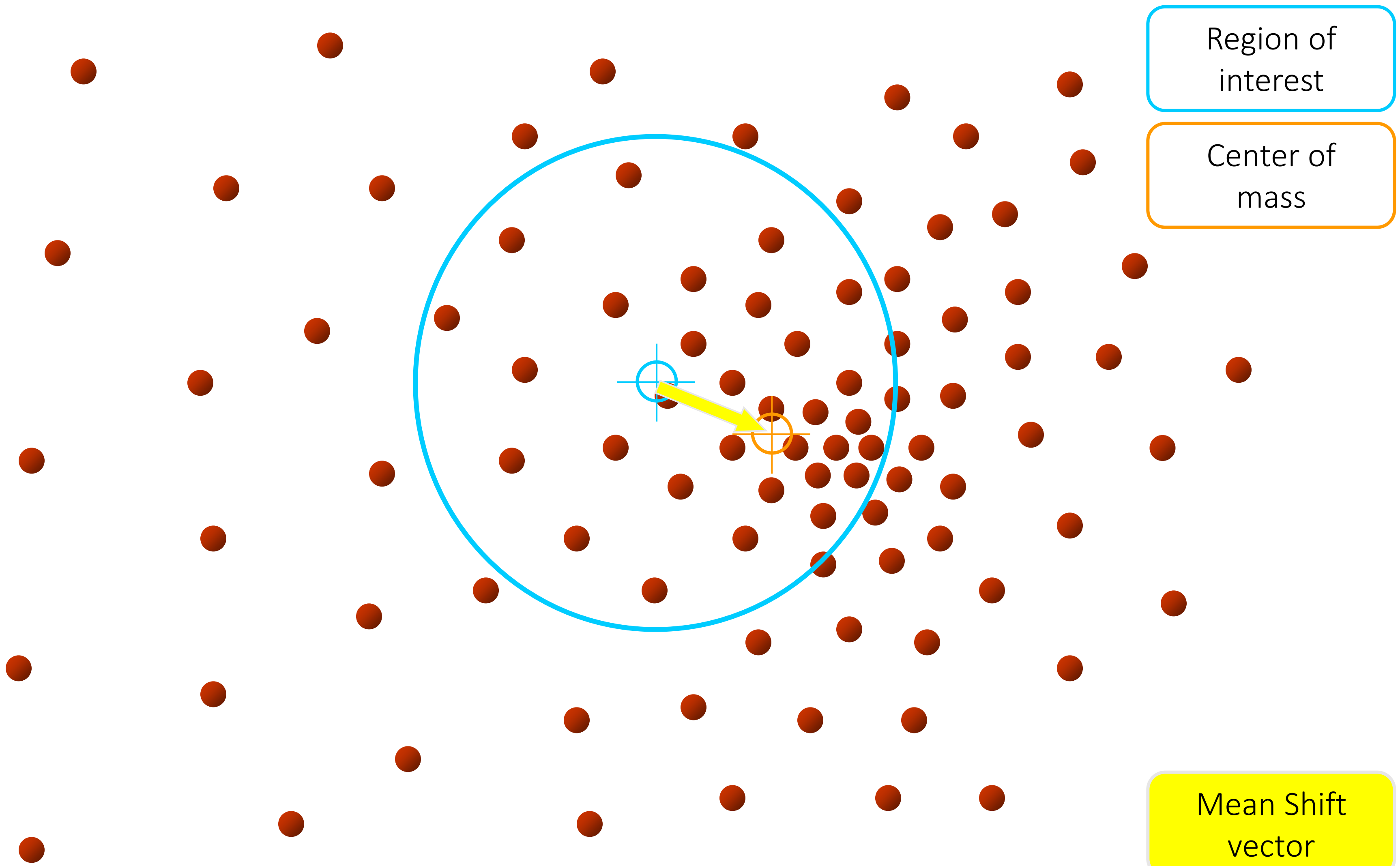
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Mean-Shift



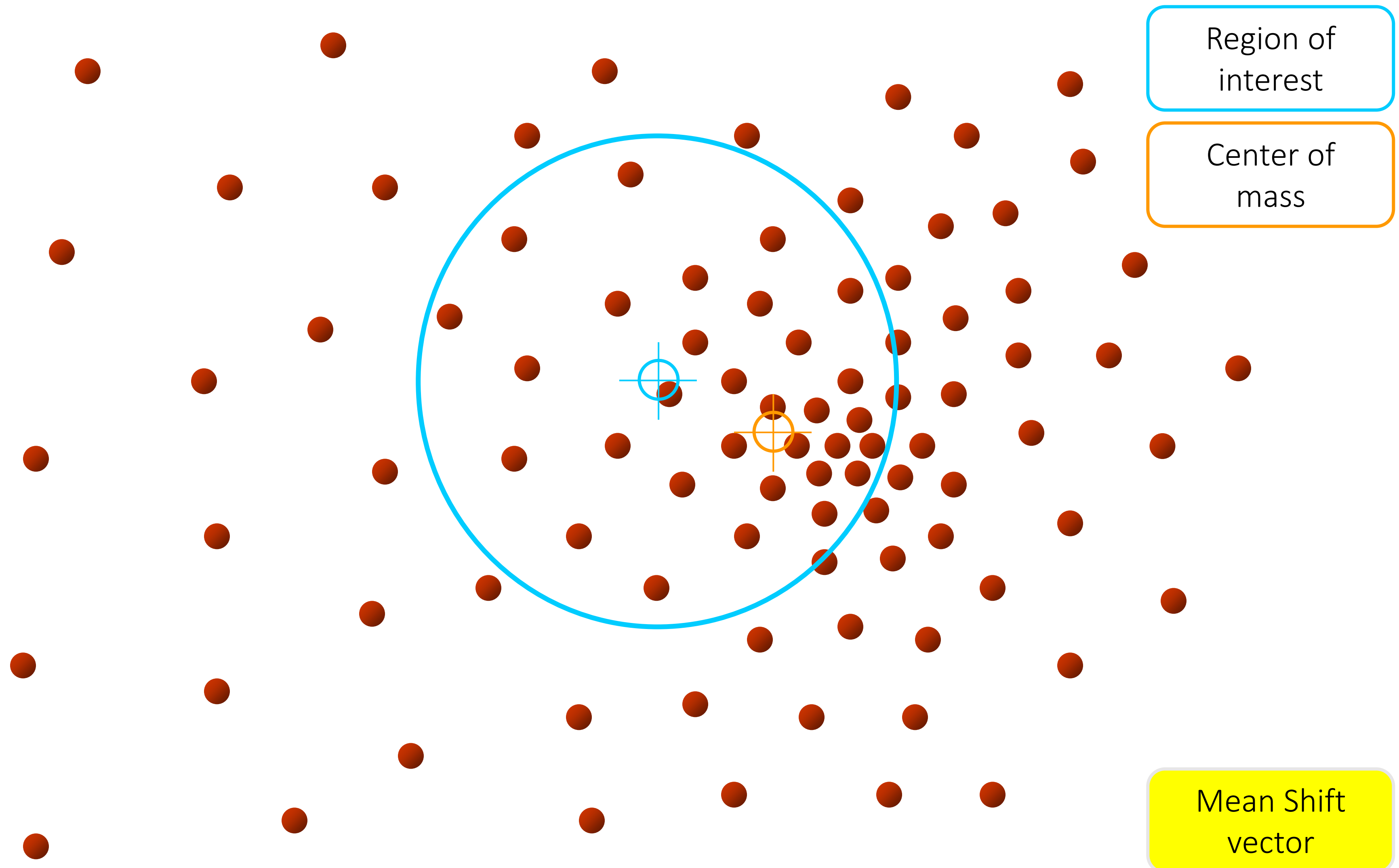
Slide credit: Y. Ukrainitz & B. Sarel

Mean-Shift



Slide credit: Y. Ukrainitz & B. Sarel

Mean-Shift

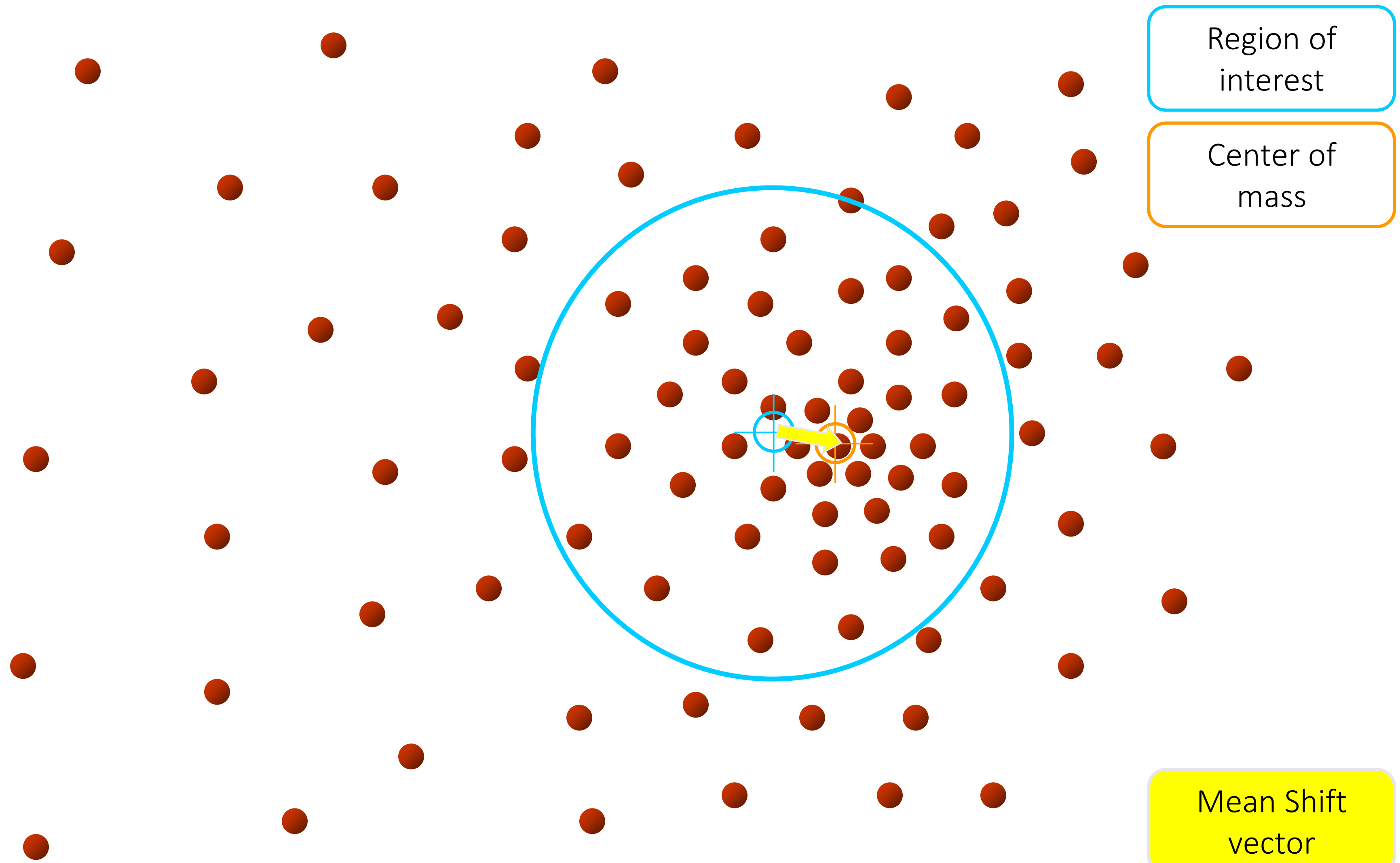


Slide credit: Y. Ukrainitz & B. Sarel

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

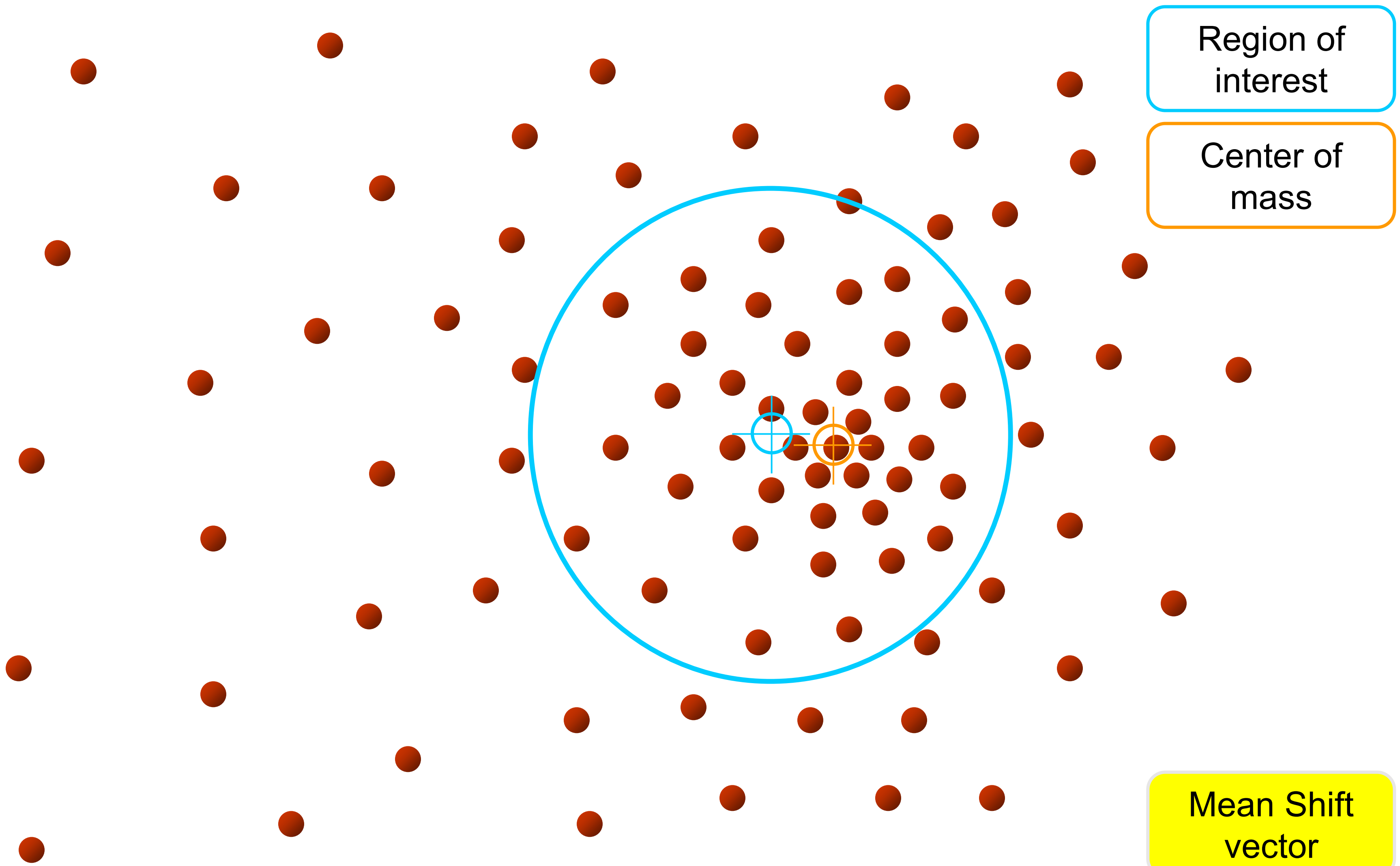
Mean-Shift



Slide credit: Y. Ukrainitz & B. Sarel

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

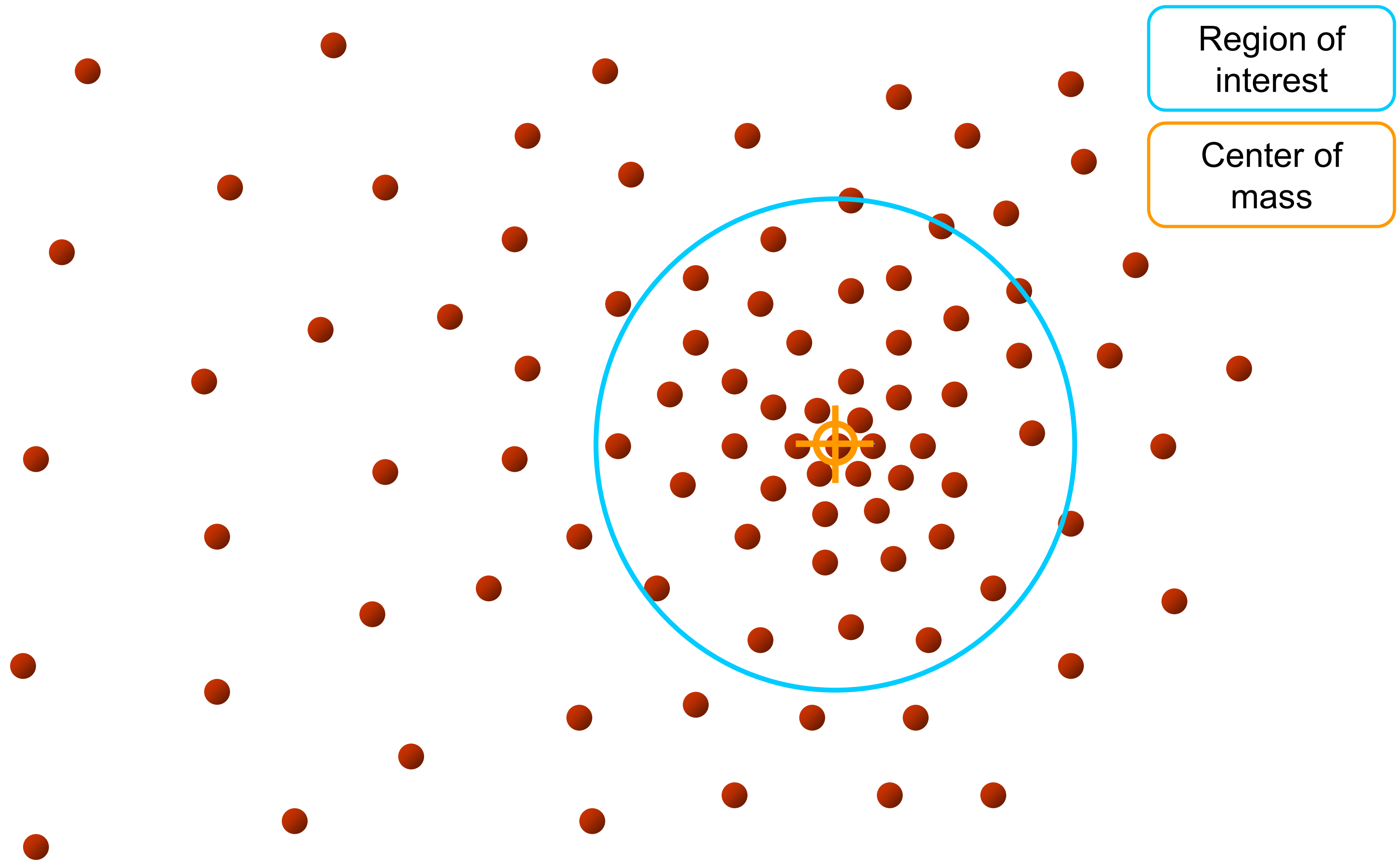
Mean-Shift



Slide credit: Y. Ukrainitz & B. Sarel

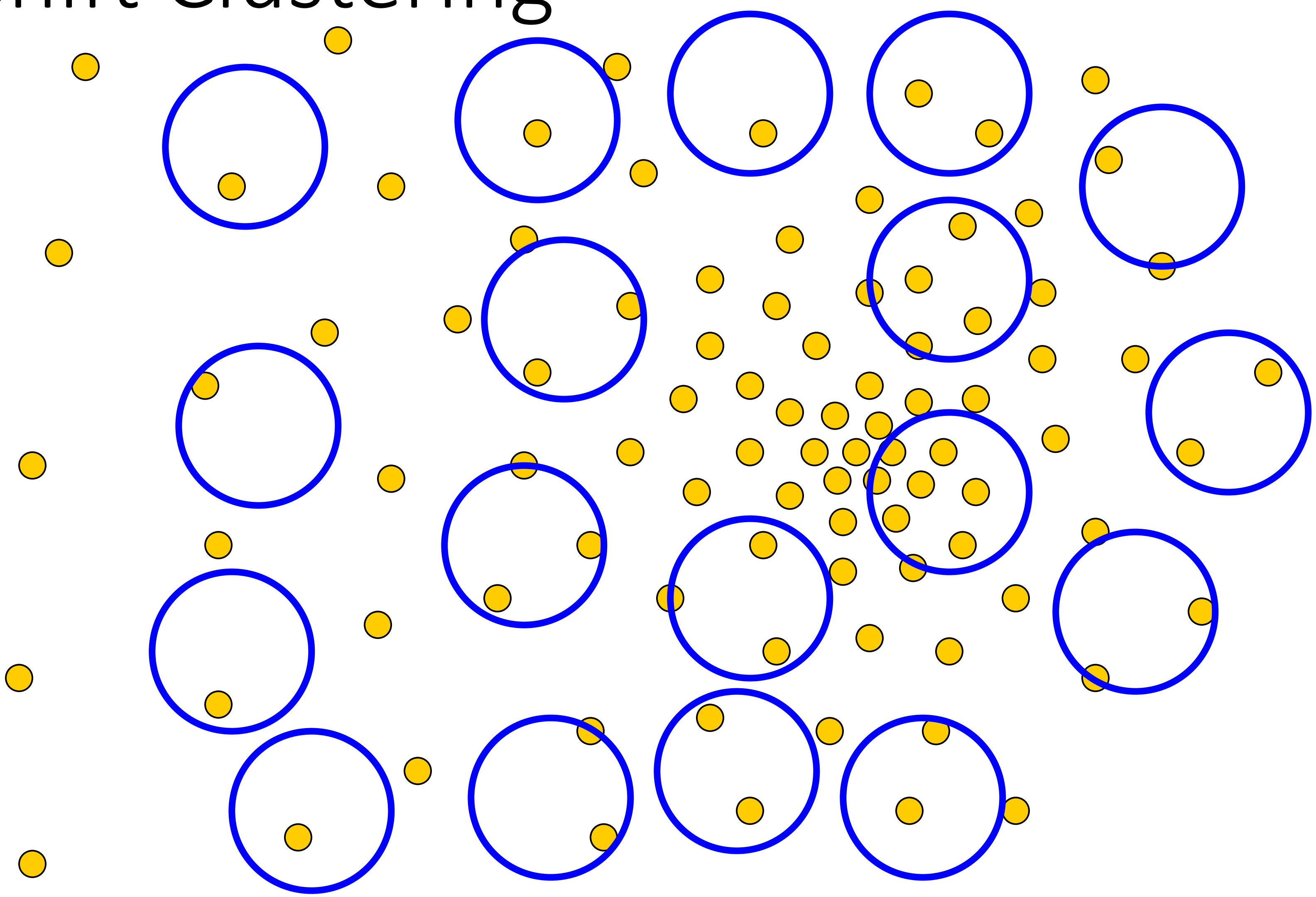
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Mean-Shift



Slide credit: Y. Ukrainitz & B. Sarel

Mean-Shift Clustering



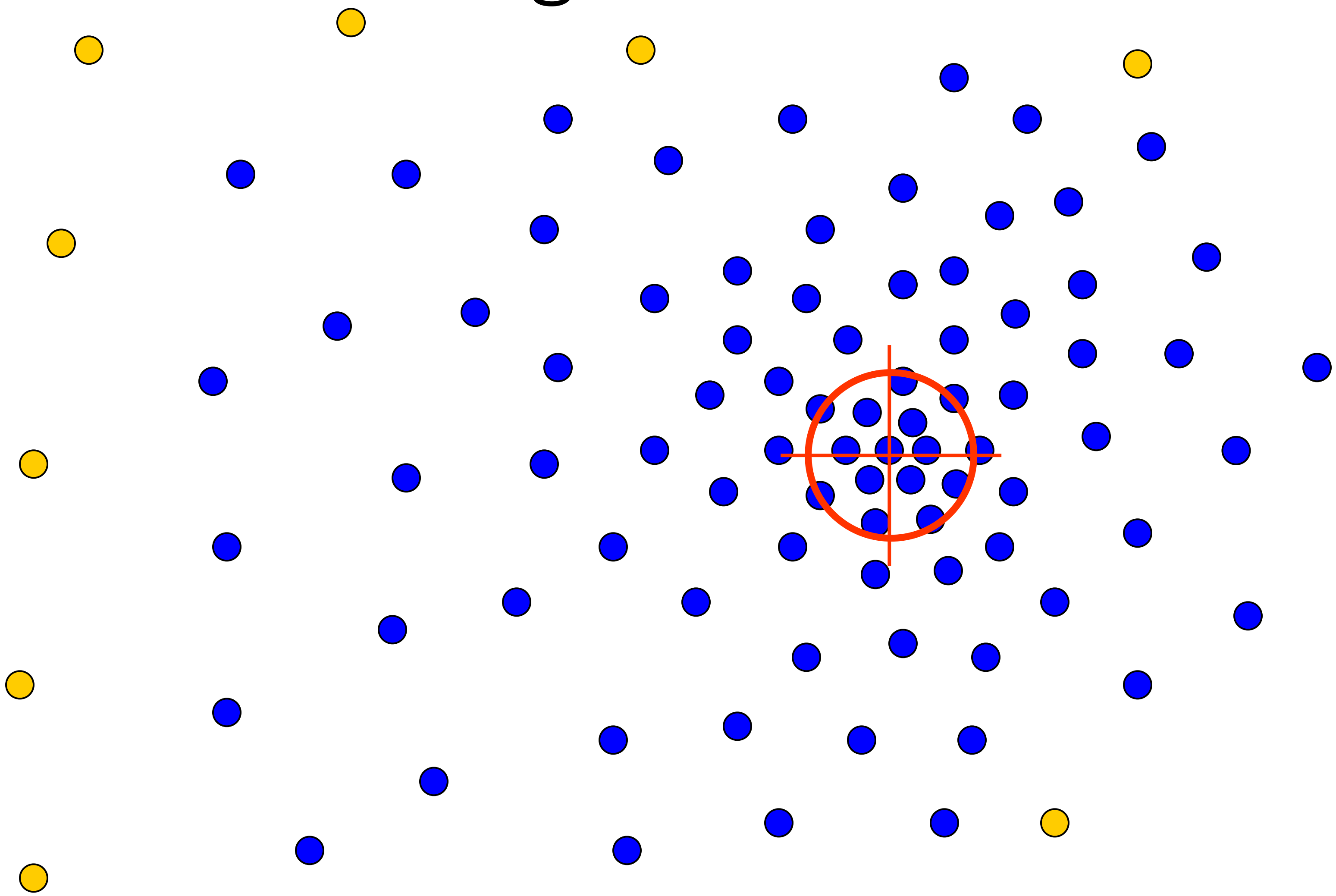
Initialize multiple windows in the space

Run the procedure in parallel

Slide credit: Y. Ukrainitz & B. Sarel

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Mean-Shift Clustering



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

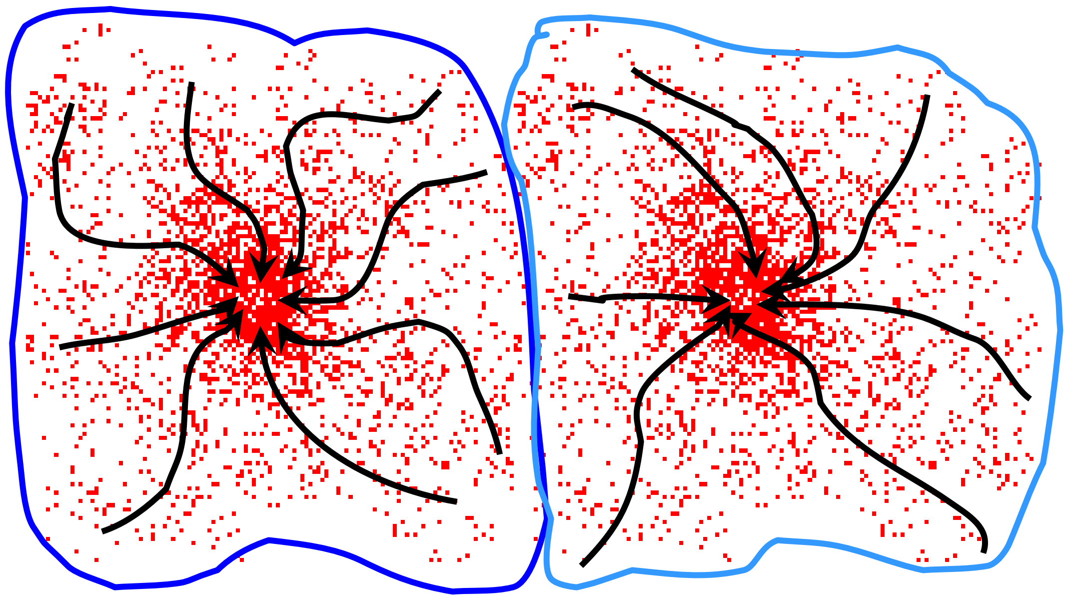
Slide credit: Y. Ukrainitz & B. Sarel

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Mean-Shift Clustering

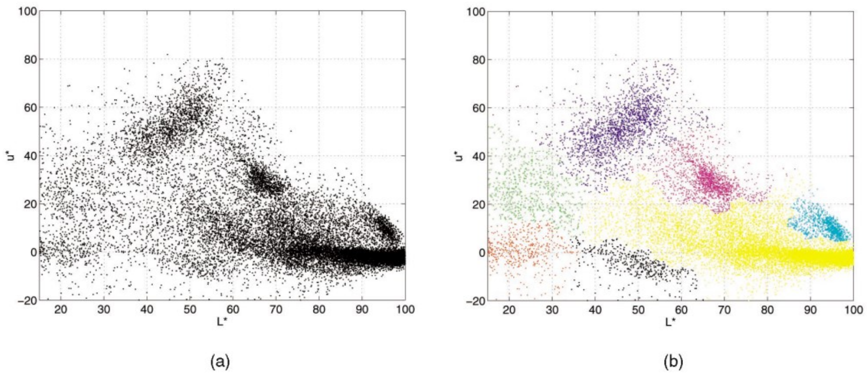
- Cluster: all data points in the attraction basin of a mode
- Attraction basin: the region for which all trajectories lead to the same mode

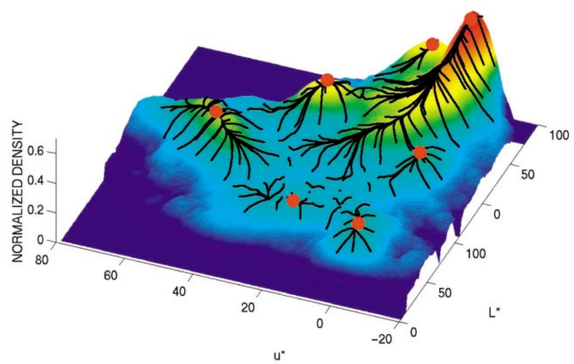


Slide credit: Y. Ukrainitz & B. Sarel

Mean-Shift Clustering/Segmentation

- 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

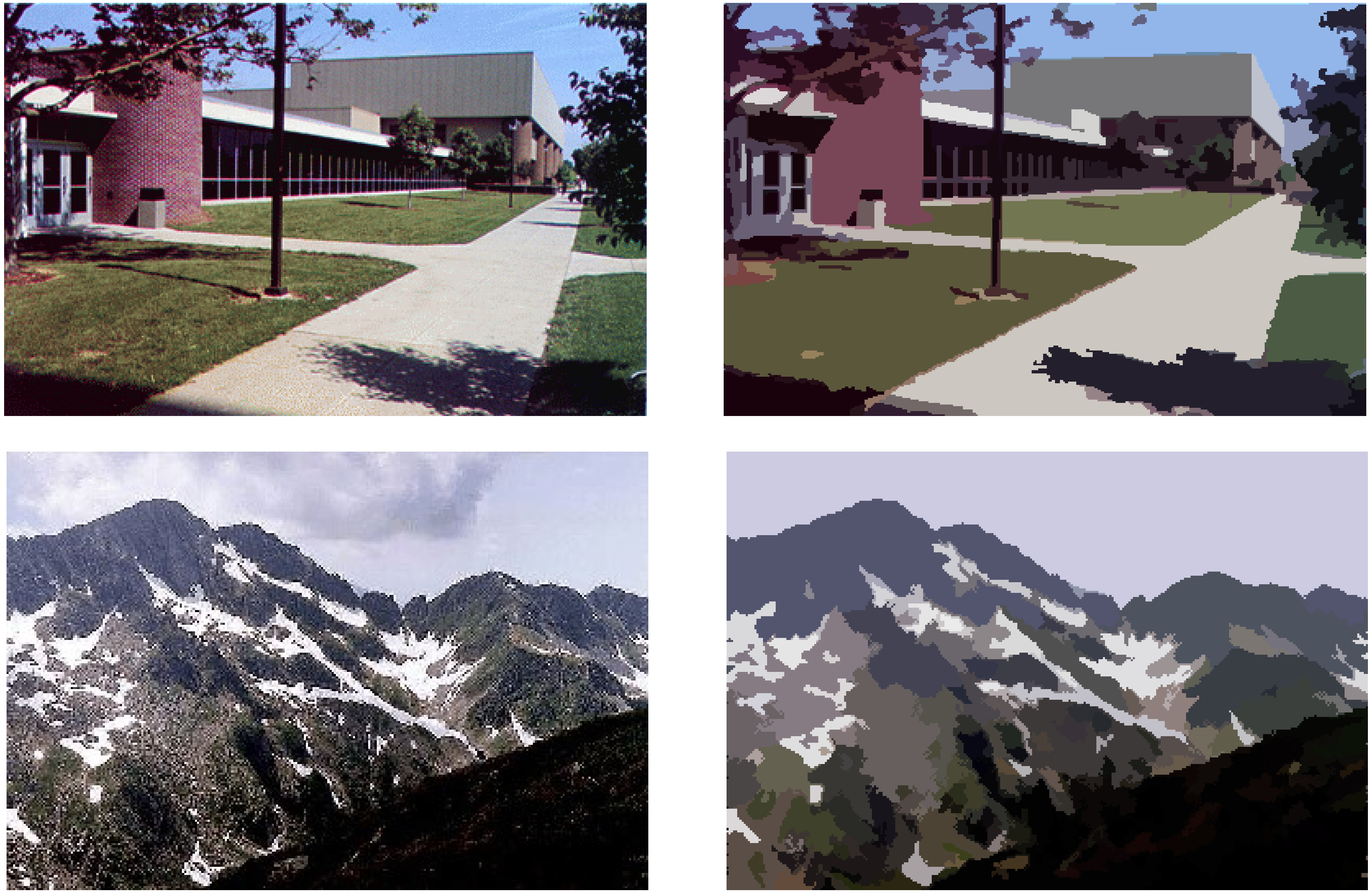




Slide credit: Svetlana Lazebnik

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

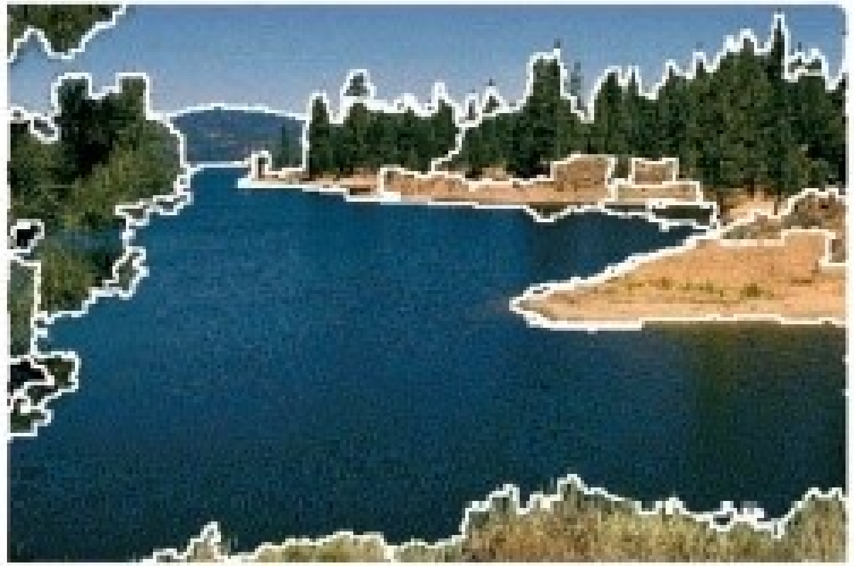
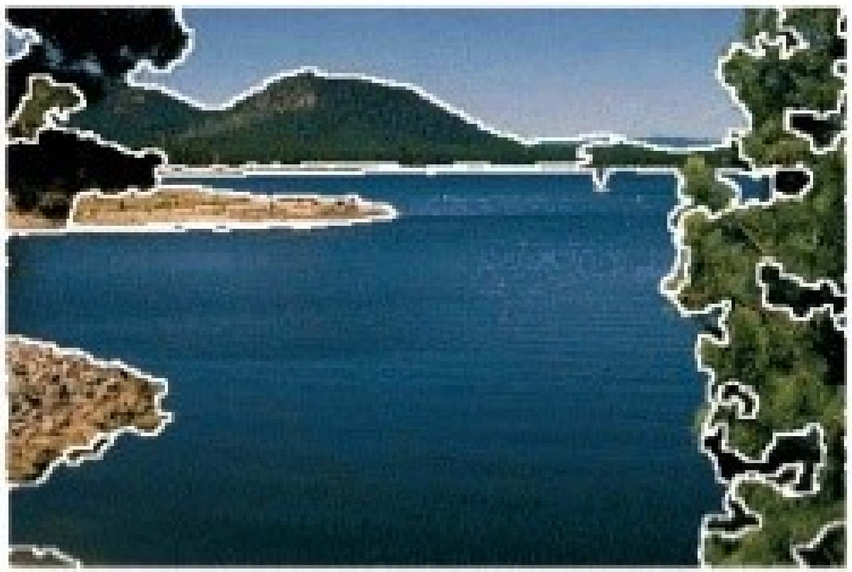
Mean-Shift Segmentation Results



Slide credit: Svetlana Lazebnik

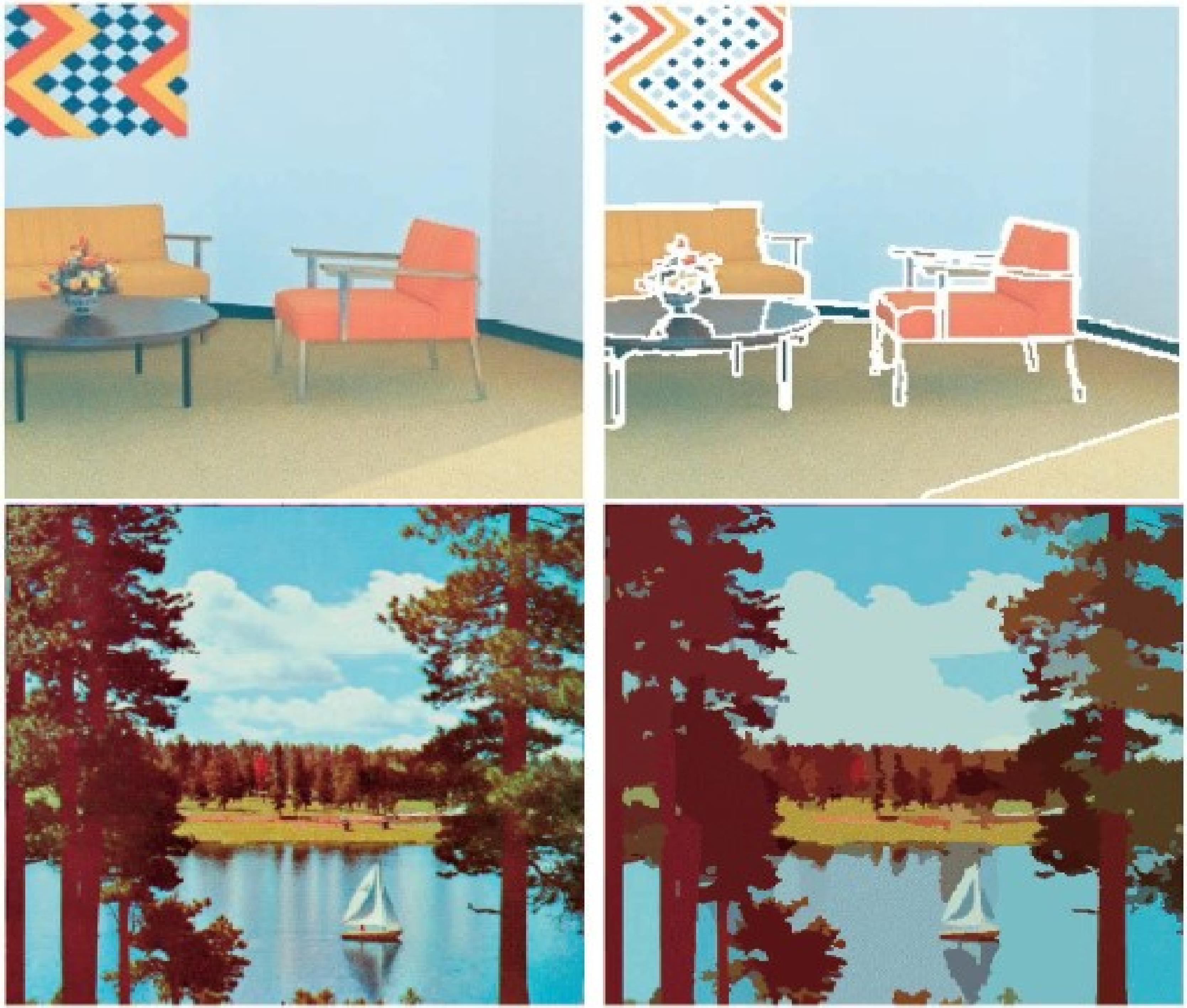
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

More Results


Slide credit: Svetlana Lazebnik

More Results



This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Mean-Shift pros and cons

- Pros
  - General, application-independent tool
  - Model-free, does not assume any prior shape (spherical, elliptical, etc.) on data clusters
  - Just a single parameter (window size h)
    - h has a physical meaning (unlike k-means)
  - Finds variable number of modes
  - Robust to outliers
- Cons
  - Output depends on window size
  - Window size selection is not trivial
  - Computationally (relatively) expensive (~2s/image)
  - Does not scale well with dimensionality of feature space

Slide credit: Svetlana Lazebnik

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

- Visual Recognition Tasks
- Introduction to segmentation and clustering
- Agglomerative clustering
- K-means clustering
- Mean-shift clustering
- Efficient Graph-based image segmentation

Reading: Forsyth Chapter 9

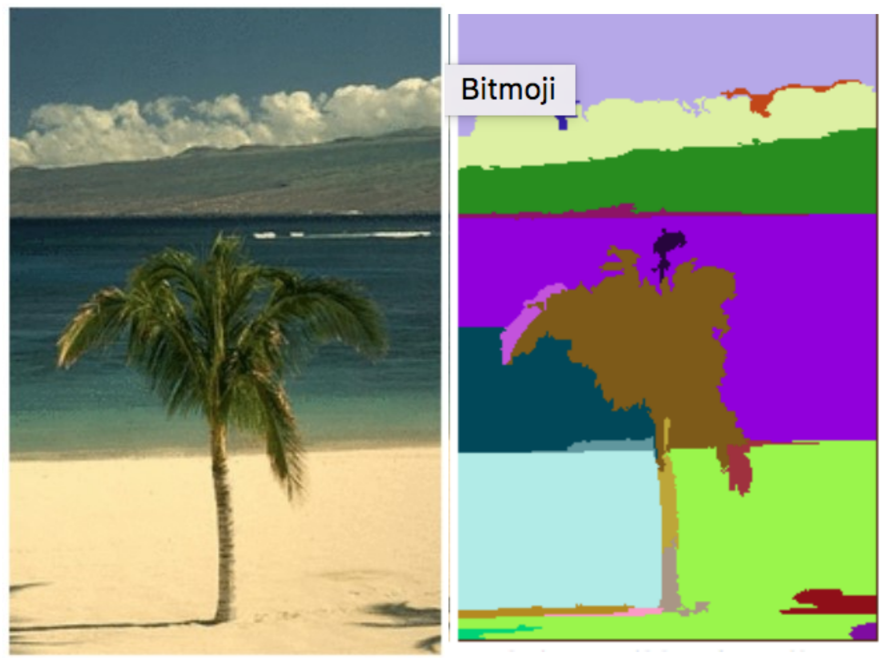
D. Comaniciu and P. Meer, Mean Shift: A Robust Approach toward Feature Space Analysis, TPAMI 2002

Co-financed by the European Union
Connecting Europe Facility

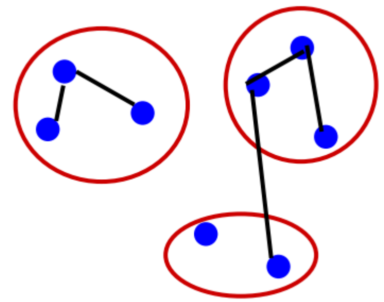
This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Efficient Graph-based Image Segmentation

Oversegmentation algorithm introduced by *Felzenszwalb and Huttenlocher* in the paper titled *Efficient Graph-Based Image Segmentation*



Problem Formulation

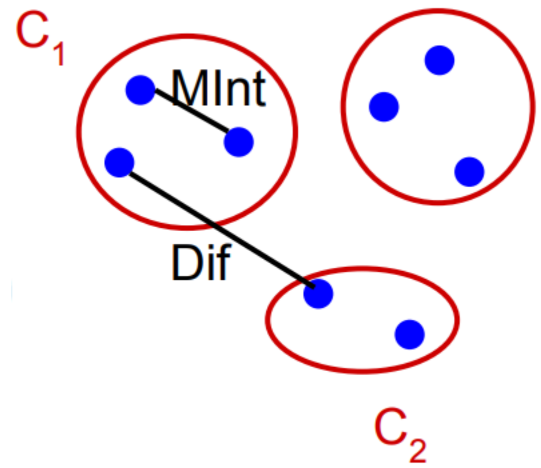


- Graph G = (V, E)
- V is set of nodes (i.e. pixels)
- E is a set of undirected edges between pairs of pixels
- w(vi, vj) is the weight of the edge between nodes vi and vj .
- S is a segmentation of a graph G such that G' = (V, E') where E' ⊂ E.
- S divides G into G' such that it contains distinct clusters C.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Predicate for segmentation



- Predicate D determines whether there is a boundary for segmentation.

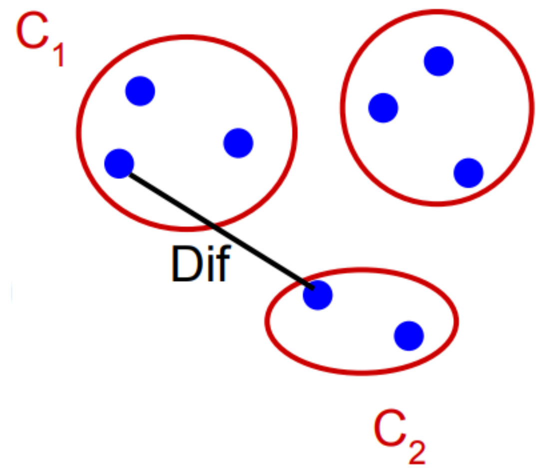
$$Merge(C_1, C_2) = \begin{cases} True & \text{if } dif(C_1, C_2) < in(C_1, C_2) \\ False & otherwise \end{cases}$$

Where

- dif(C1 , C2) is the difference between two clusters.
- in(C1 , C2) is the internal difference in the clusters C1 and C2

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Predicate for Segmentation



- Predicate D determines whether there is a boundary for segmentation.

$$Merge(C_1, C_2) = \begin{cases} True & \text{if } dif(C_1, C_2) < in(C_1, C_2) \\ False & otherwise \end{cases}$$

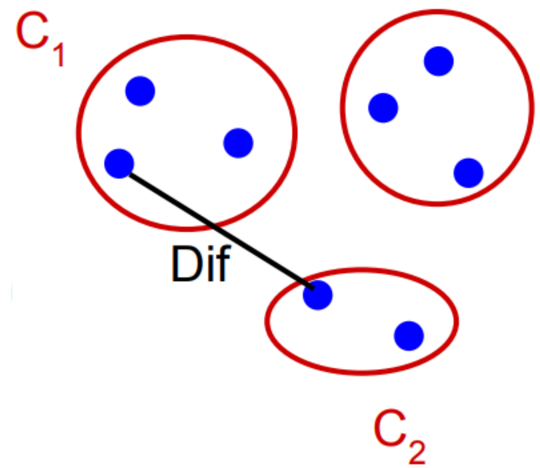
$$dif(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2, (C_1, C_2) \in E} w(v_i, v_j)$$

The difference between two components is the minimum weight edge that connects a node vi in cluster C1 to node vj in C2

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Predicate for Segmentation



- Predicate D determines whether there is a boundary for segmentation.

$$Merge(C_1, C_2) = \begin{cases} True & if\ dif(C_1, C_2) < in(C_1, C_2) \\ False & otherwise \end{cases}$$

$$dif(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2, (C_1, C_2) \in E} w(v_i, v_j)$$

$$in(C_1, C_2) = \min_{C \in \{C_1, C_2\}} \left[\max_{v_i, v_j \in C} \left[w(v_i, v_j) + \frac{k}{|C|} \right] \right]$$

in(C1, C2) is the maximum weight edge that connects two nodes in the same component.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Predicate for Segmentation

- k/|C| sets the threshold by which the components need to be different from the internal nodes in a component.
- Properties of constant k:
  - If k is large, it causes a preference for larger components.
  - k does not set a minimum size for components.

Algorithm for Segmentation

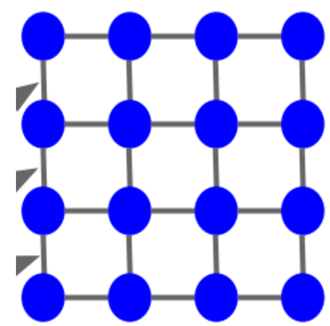
The input is a graph $G = (V, E)$, with n vertices and m edges. The output is a segmentation of V into components $S = (C_1, \ldots, C_r)$.

0. Sort E into $\pi = (o_1, \ldots, o_m)$, by non-decreasing edge weight.
1. Start with a segmentation S^0, where each vertex v_i is in its own component.
2. Repeat step 3 for $q = 1, \ldots, m$.
3. Construct S^q given S^{q-1} as follows. Let v_i and v_j denote the vertices connected by the q-th edge in the ordering, i.e., $o_q = (v_i, v_j)$. If v_i and v_j are in disjoint components of S^{q-1} and $w(o_q)$ is small compared to the internal difference of both those components, then merge the two components otherwise do nothing. More formally, let C_i^{q-1} be the component of S^{q-1} containing v_i and C_j^{q-1} the component containing v_j. If $C_i^{q-1} \neq C_j^{q-1}$ and $w(o_q) \leq MInt(C_i^{q-1}, C_j^{q-1})$ then S^q is obtained from S^{q-1} by merging C_i^{q-1} and C_j^{q-1}. Otherwise $S^q = S^{q-1}$.
4. Return $S = S^m$.

Features and weights

How to build the graph ? Two options:

1. Grid-graph: Every pixel is connected to its 8 neighboring pixels and the weights are determined by the difference in intensities.



2. NN-graph: Project every pixel into **feature space** defined by
   - (x, y, r, g, b).
   - Weights between pixels are determined using L2 (Euclidian) distance in feature space.
   - Edges are chosen for only top ten nearest neighbors in feature space to ensure run time of O(n log n) where n is number of pixels.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Results

With 8-neighbor grid graph
Edge weight: intensity difference

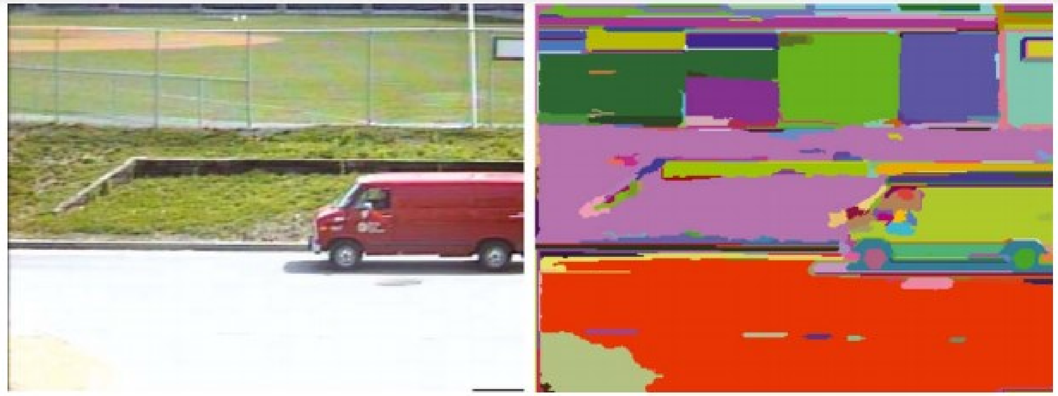


Figure 2. A street scene (320 × 240 color image), and the segmentation results produced by our algorithm ($\sigma = 0.8, k = 300$).

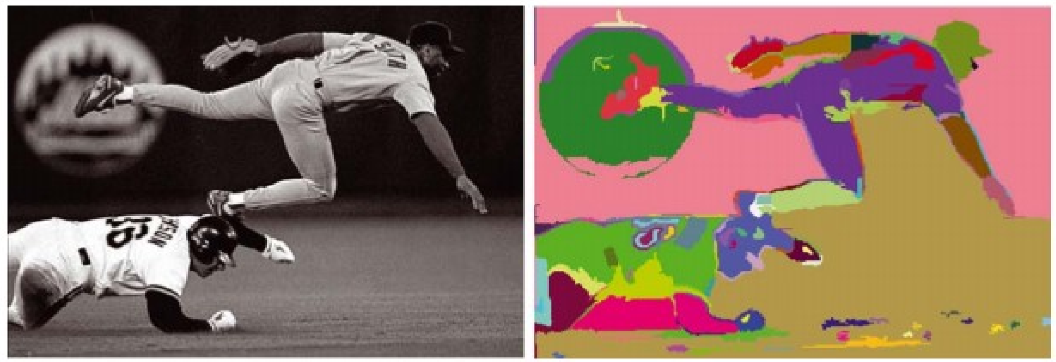


Figure 3. A baseball scene (432 × 294 grey image), and the segmentation results produced by our algorithm ($\sigma = 0.8, k = 300$).

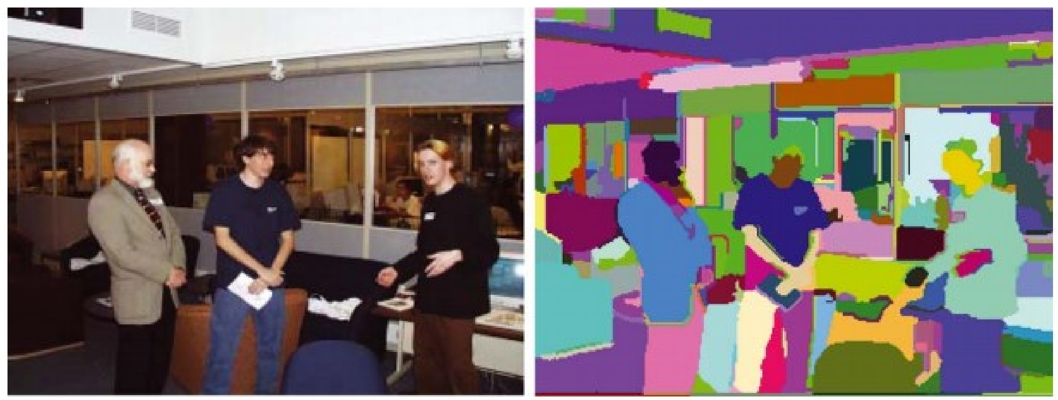


Figure 4. An indoor scene (image 320 × 240, color), and the segmentation results produced by our algorithm ($\sigma = 0.8, k = 300$).

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Results

With nearest neighbor graph
Edge weight: L2 distance in feature space

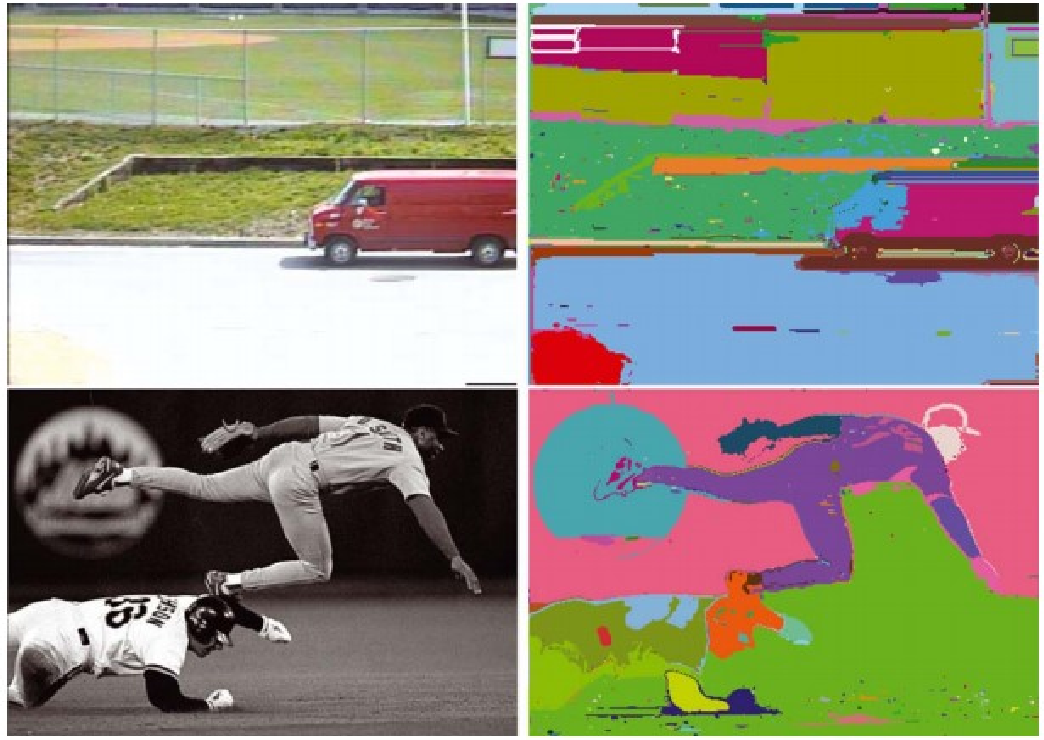


Figure 7. Segmentation of the street and baseball player scenes from the previous section, using the nearest neighbor graph rather than the grid graph ($\sigma = 0.8$, $k = 300$).

Figure 8. Segmentation using the nearest neighbor graph can capture spatially non-local regions ($\sigma = 0.8$, $k = 300$).

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Results – close up

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

Visual Computing Group

CYENS CENTRE OF EXCELLENCE

Thank you.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423