Computer Vision – TP8
Advanced Segmentation

Miguel Tavares Coimbra

U. PORTO FC

Outline

- Segmentation by Clustering
- Segmentation by Fitting
- Semantic Segmentation

Topic: Segmentation by Clustering

- Segmentation by Clustering
- Segmentation by Fitting
- Semantic Segmentation

What is ‘Segmentation’? (again?)

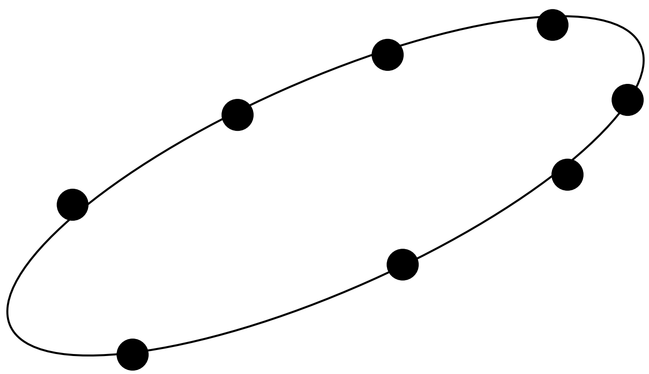
- Traditional definition:

  “Separation of the image in different areas“
  - Decompose an image into “superpixels”
  - Colour and texture coherence.

- Aren’t there other ways to look at the ‘Segmentation’ concept?

Other ‘*Segmentation*’ problems

- Fitting lines to edge points



We can’t see this as ‘separating an image in different areas’!

- Fitting a fundamental matrix to a set of feature points

This one is complicated!
Check Forsyth and Ponce, chap.14

Segmentation as Clustering

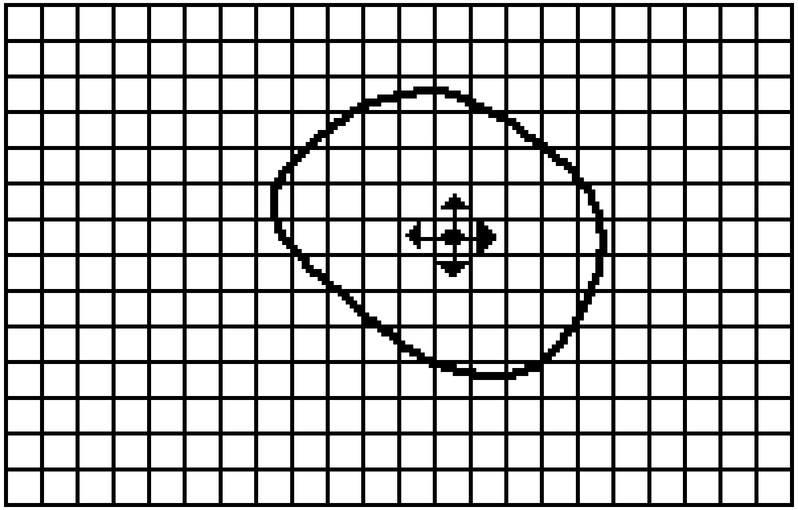
- Tries to answer the question:

  "Which components of the data set naturally belong together?"

- Two approaches:
  - Partitioning
    - Decompose a large data set into pieces that are 'good' according to our model
  - Grouping
    - Collect sets of data items that 'make sense' according to our model

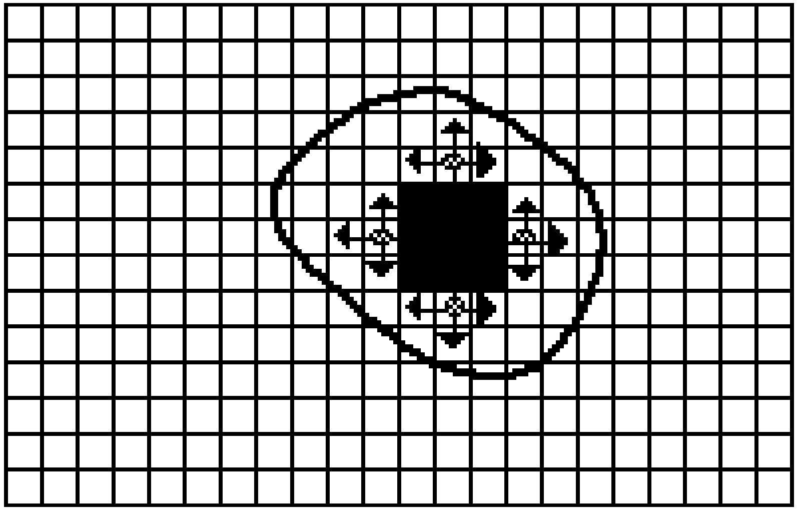
Simple clustering

- Two natural types of clustering:
  - Divisive clustering
    - Entire data set is regarded as a cluster
    - Clusters are recursively split
  - Agglomerative clustering
    - Each data item is a cluster
    - Clusters are recursively merged
- Where have I seen this before?



(a) Start of Growing a Region

Split and Merge (Region-based segmentation) is in fact a *clustering* algorithm



(b) Growing Process After a Few Iterations

Generic simple clustering algorithms

- Divisive Clustering
  - Construct a single cluster containing all points
  - While the clustering is not satisfactory
    - Split the cluster that yields the two components with the largest inter-cluster distance
  - end
- Agglomerative Clustering
  - Make each point a separate cluster
  - Until the clustering is satisfactory
    - Merge the two clusters with smallest inter-cluster distance
  - end

Which inter-cluster distance?

What does this mean?

Simple clustering with images

- Some specific problems arise:
  - Lots of pixels! Graphical representations are harder to read
  - Segmentation: It is desirable that certain objects are connected. How to enforce this?
  - When do we stop splitting/merging process?
- Complex situations require more complex clustering solutions!

K-means Clustering

- What if we know that there are *k* clusters in the image?
- We can define an *objective function*!
  - Expresses how good my representation is
- We can now build an algorithm to obtain the *best* representation

Caution! "*Best*" given my objective function!

K-means Clustering

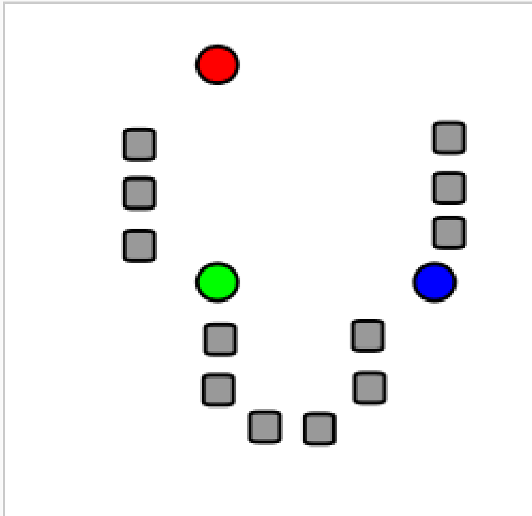
- Assume:
  - We have k clusters
  - Each cluster i has a centre c_i
  - Element j to be clustered is described by a feature vector x_j
- Our objective function is thus:

What does this mean?

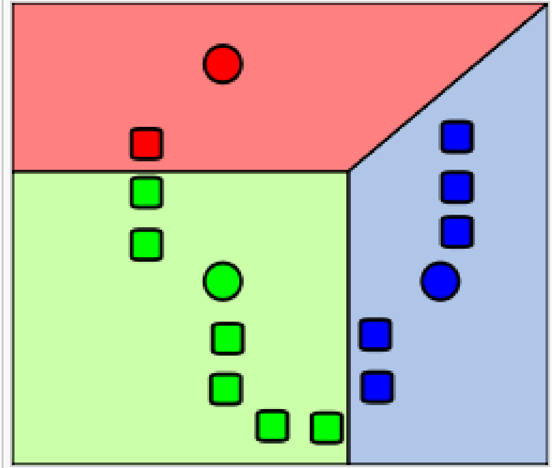
$$\Phi(clusters, data) = \sum_{i \in clusters} \left\{ \sum_{j \in cluster(i)} (x_j - c_i)^T (x_j - c_i) \right\}$$

Iteration step

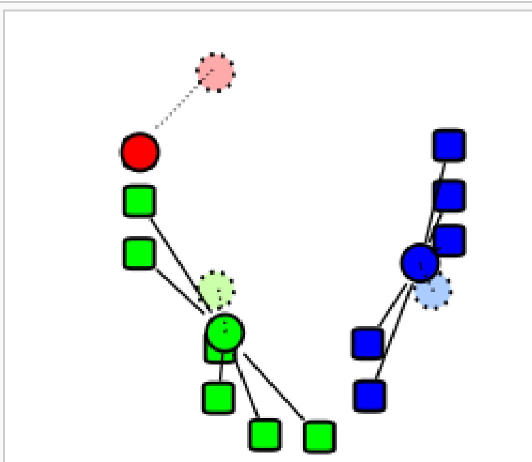
- Too many possible allocations of points to clusters to search this space for a minimum
- Iterate!
  - Assume cluster centres are known and allocate each point to the closest cluster centre
  - Assume the allocation is known and choose a new set of cluster centres. Each centre is the mean of the points allocated to that cluster



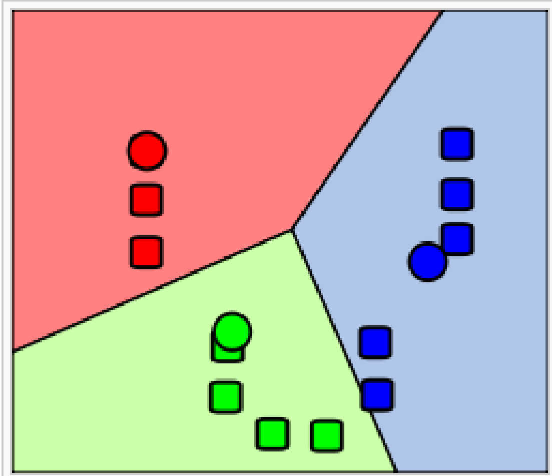
Shows the initial randomized centroids and a number of points.



Points are associated with the nearest centroid.



Now the centroids are moved to the center of their respective clusters.



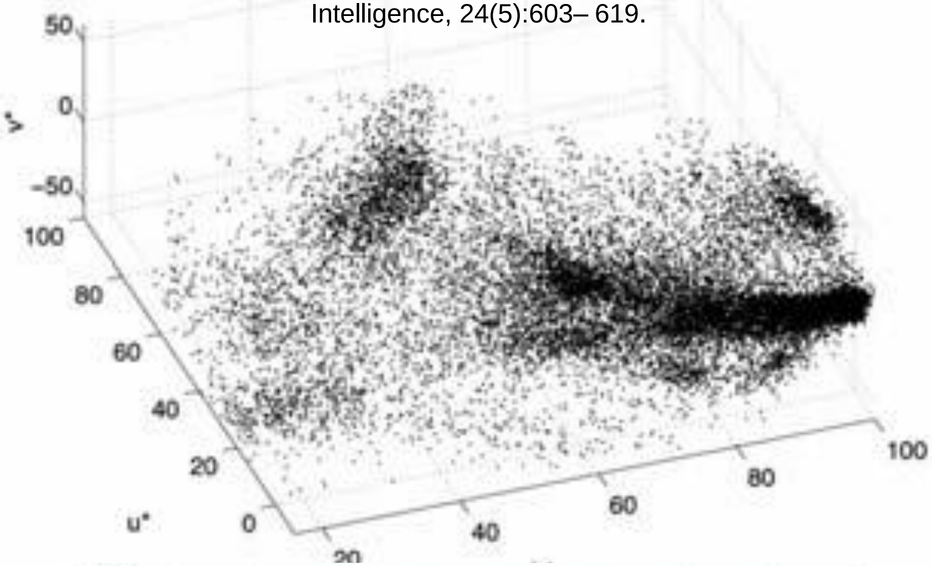
Steps 2 & 3 are repeated until a suitable level of convergence has been reached.

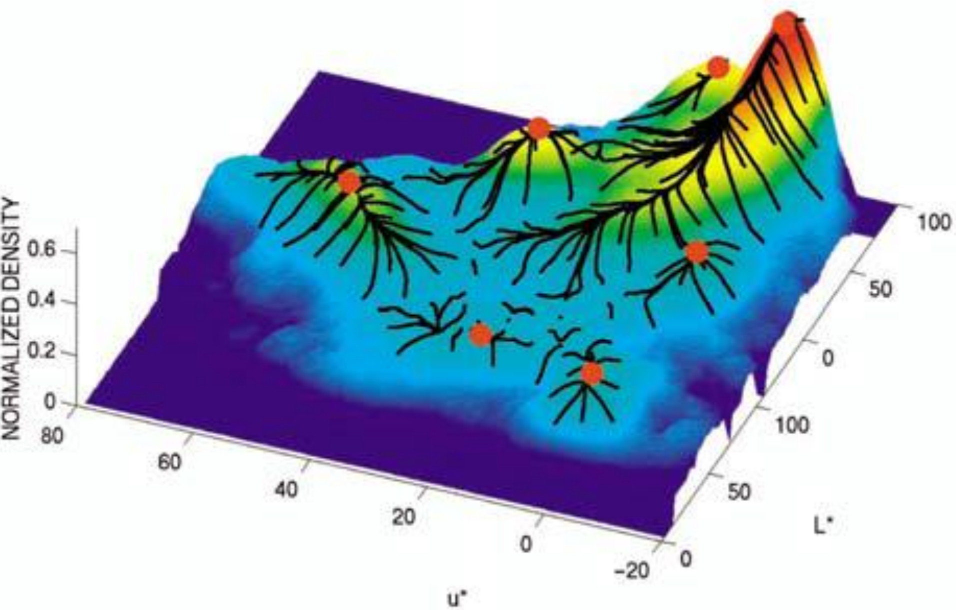
[Wikipedia]

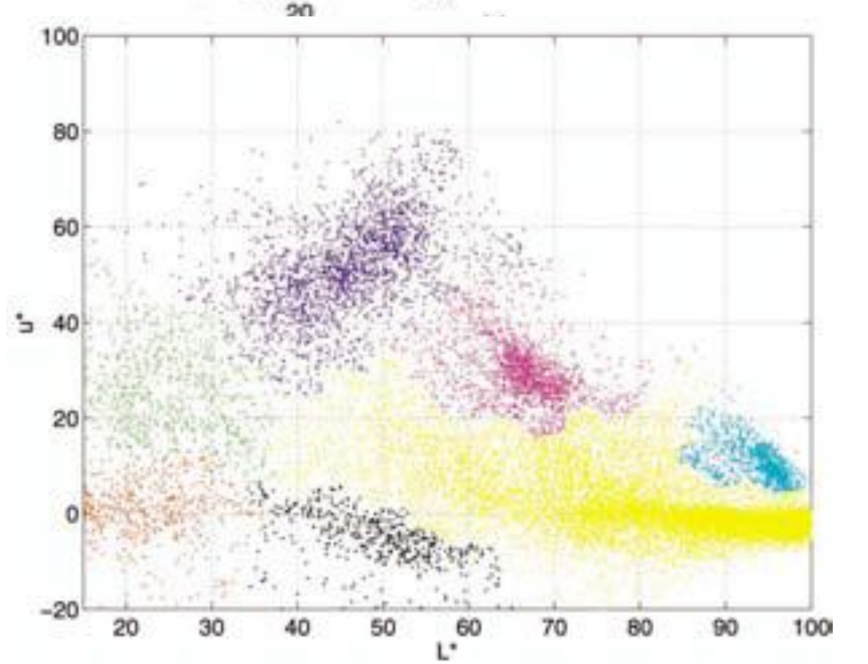
Mean Shift

- K-means:
  - Segments our feature space (and not the image!) into clusters (and not regions)
  - Uses a parametric model for its distributions (e.g. Gaussians), whose locations (centers) and shape (covariance) can be estimated
- Mean shift:
  - Segments an image (and not the feature space) into regions
  - Uses a non-parametric model (simply tries to find distribution peaks)

Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5):603– 619.







How does it work?

- Each pixel 'finds' its nearest distribution peak by 'climbing uphill' the image's kernel density function f(x)
- The gradient of f(x) defines the direction for this 'climb', by defining mean-shift vectors
- Pixels that 'climb' to the same peak form a region

One-dimensional example

Szeliski, “Computer Vision: Algorithms and Applications”, Springer, 2011

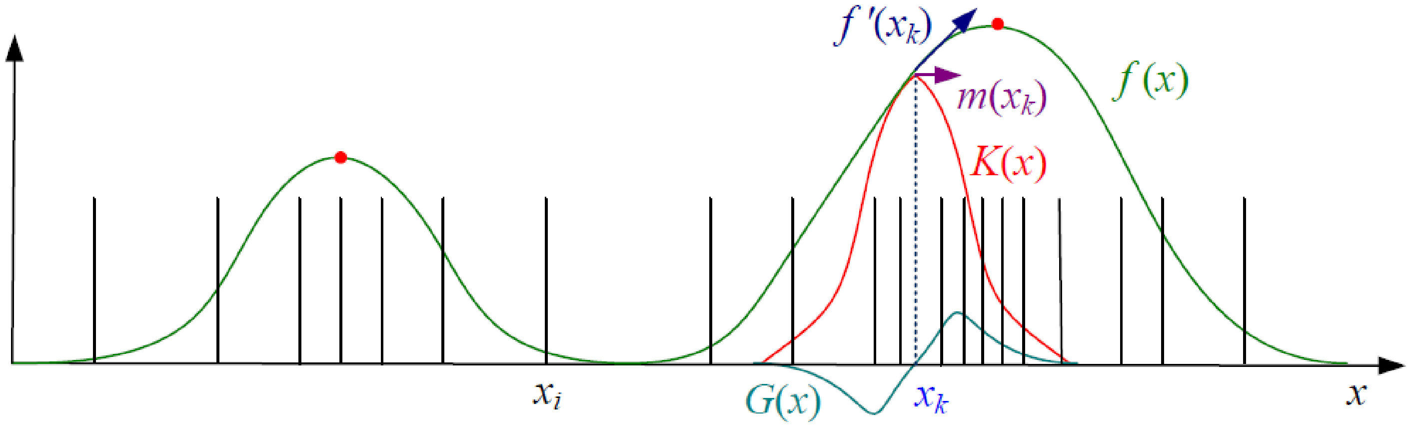
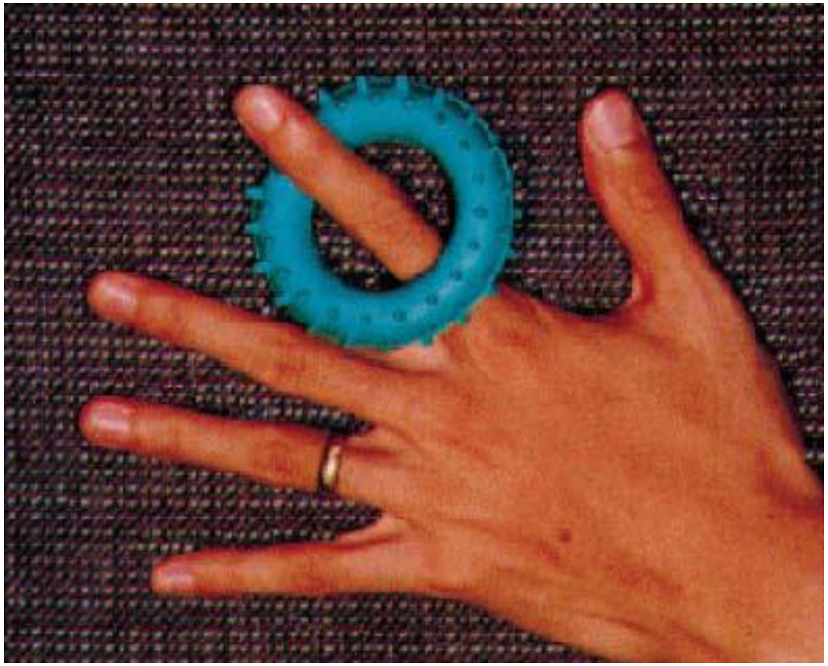
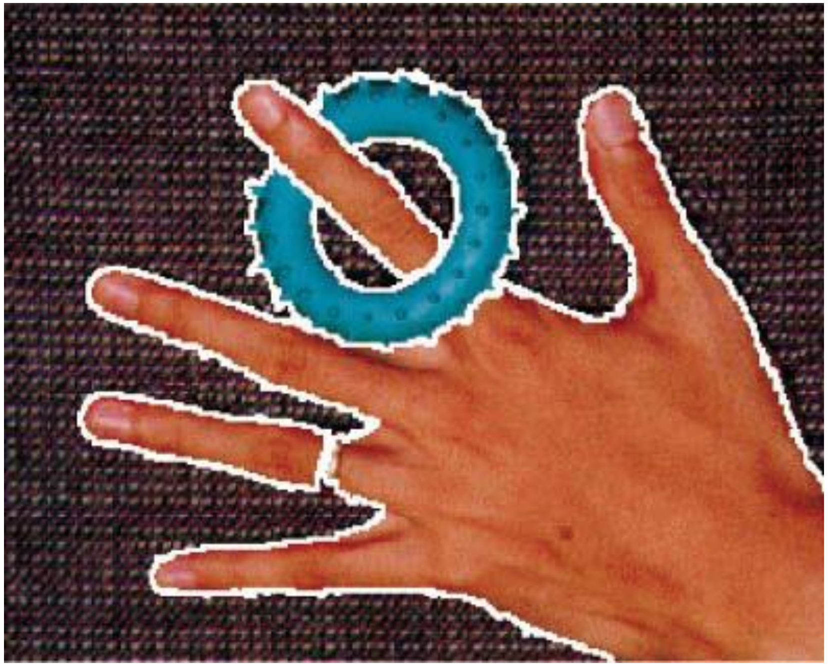


Figure 5.17 One-dimensional visualization of the kernel density estimate, its derivative, and a mean shift. The kernel density estimate $f(x)$ is obtained by convolving the sparse set of input samples x_i with the kernel function $K(x)$. The derivative of this function, $f'(x)$, can be obtained by convolving the inputs with the derivative kernel $G(x)$. Estimating the local displacement vectors around a current estimate x_k results in the mean-shift vector $m(x_k)$, which, in a multi-dimensional setting, point in the same direction as the function gradient $\nabla f(\boldsymbol{x}_k)$. The red dots indicate local maxima in $f(x)$ to which the mean shifts converge.

Mean Shift

Comaniciu, D. and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5):603– 619.





Normalized Cuts

- Clustering can be seen as a problem of "*cutting graphs into good pieces*"
- Data Items
  – Vertex in a weighted graph
  – Weights are large if elements are similar
- Cut edges
  – Cut edges with small weights
  – Keep connected components with large interior weights

Regions!

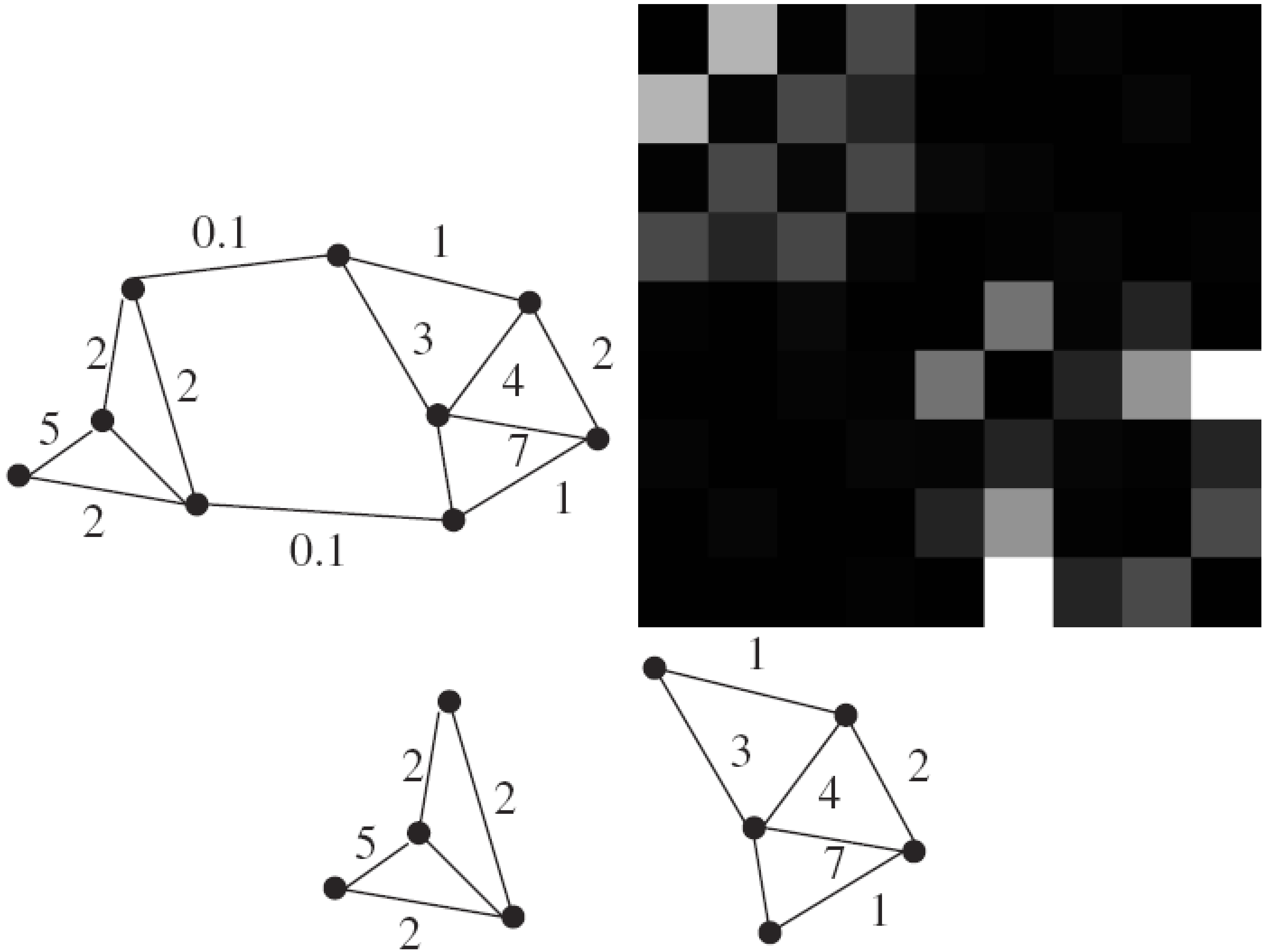


Figure 16.16. On the **top left**, a drawing of an undirected weighted graph; on the **top right**, the weight matrix associated with that graph. Larger values are lighter. By associating the vertices with rows (and columns) in a different order, the matrix can be shuffled. We have chosen the ordering to show the matrix in a form that emphasizes the fact that it is very largely block-diagonal. The figure on the **bottom** shows a cut of that graph that decomposes the graph into two tightly linked components. This cut decomposes the graph's matrix into the two main blocks on the diagonal.

Graphs and Clustering

- Associate each element to be clustered with a vertex on a graph
- Construct an edge from every element to every other
- Associate a weight with each edge based on a similarity measure
- Cut the edges in the graph to form a good set of connected components

Weight Matrices

- Typically look like block diagonal matrices
- Why?
  - Interclusters similarities are strong
  - Intracluster similarities are weak
- Split a matrix into smaller matrices, each of which is a block
- Define *Affinity Measures*

More on this

- Affinity measures
  - – Affinity by Distance
  - – Affinity by Intensity
  - – Affinity by Colour
  - – Affinity by Texture

Want to know more?
Check out: Forsyth and Ponce, Section 14.5

- Popular method: Normalized cuts

Jianbo Shi and Jitendra Malik, "Normalized Cuts and Image Segmentation", IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 22, No. 8, August 2000

Topic: Segmentation by Fitting

- Segmentation by Clustering
- Segmentation by Fitting
- Semantic Segmentation

Fitting and Clustering

- Another definition for segmentation:
  - Pixels belong together because they conform to some model
- Sounds like “Segmentation by Clustering”...
- Key difference:
  - The model is now **explicit**

We have a mathematical model for the object we want to segment.
Ex: A line

Hough Transform

- Elegant method for direct object recognition
- Edges need not be connected
- Complete object need not be visible
- Key Idea: Edges **VOTE** for the possible model

Image and Parameter Spaces

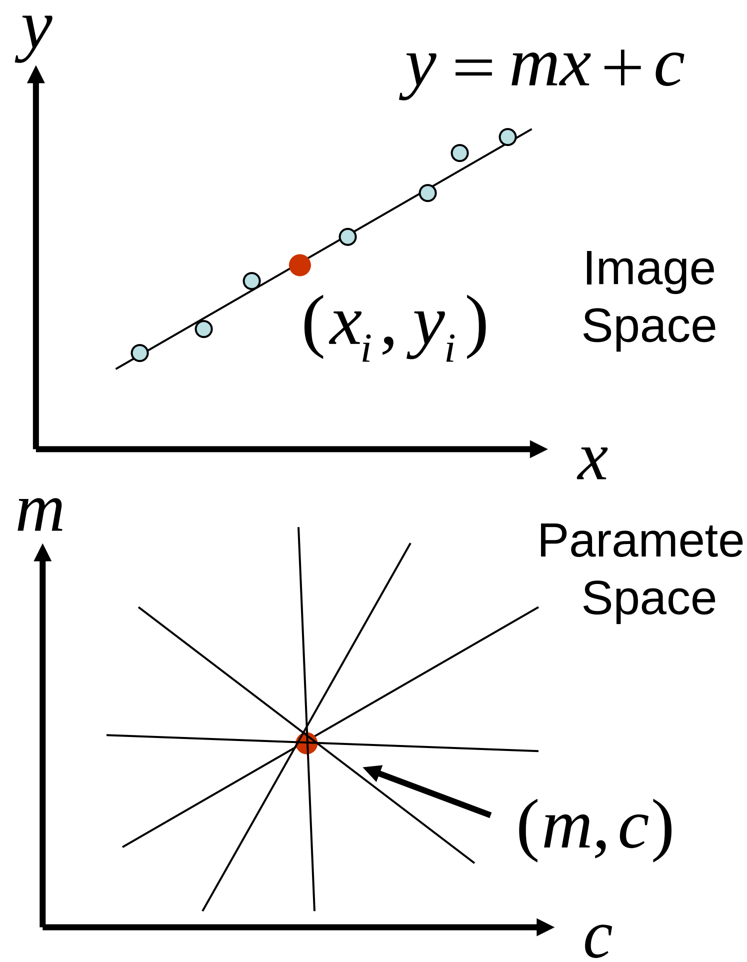
Equation of Line: $y = mx + c$

Find: (m, c)

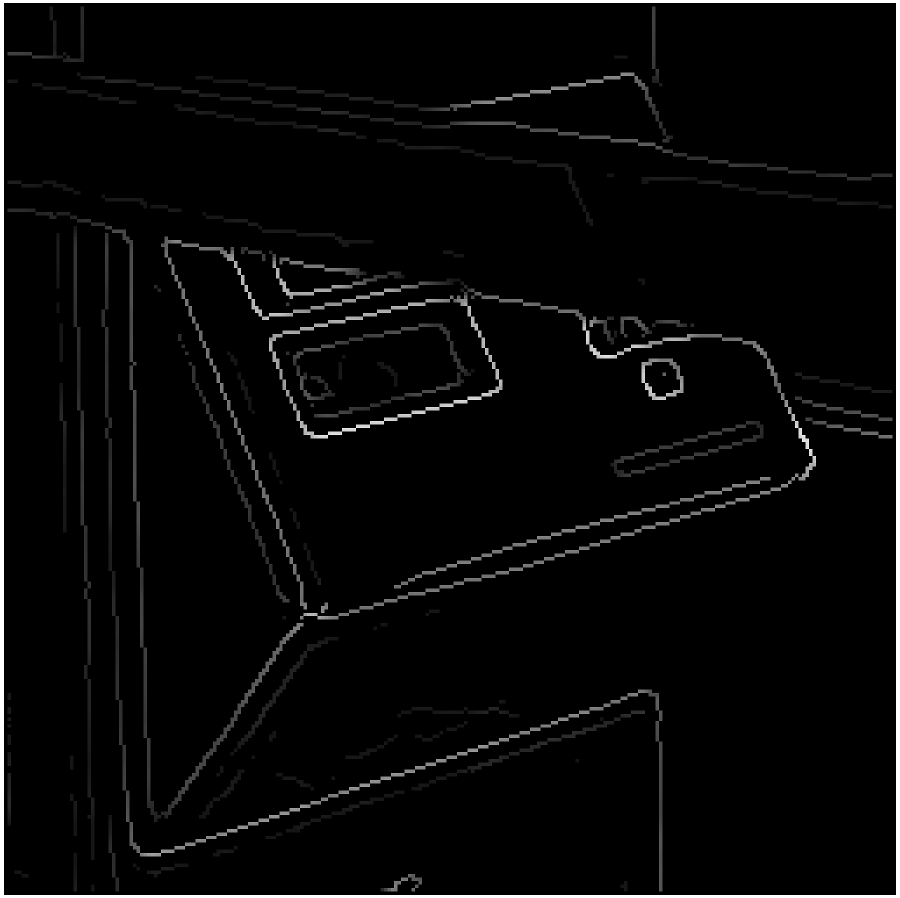
Consider point: (x_i, y_i)

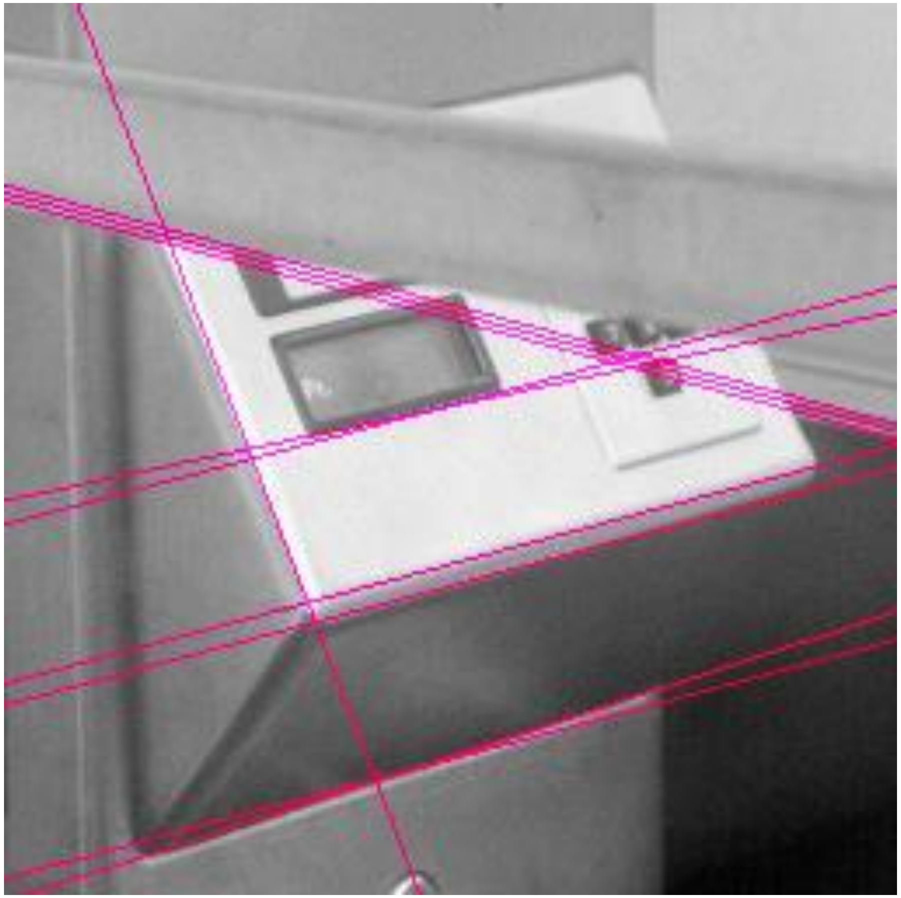
$$y_i = mx_i + c \quad or \quad c = -x_i m + y_i$$

Parameter space also called Hough Space

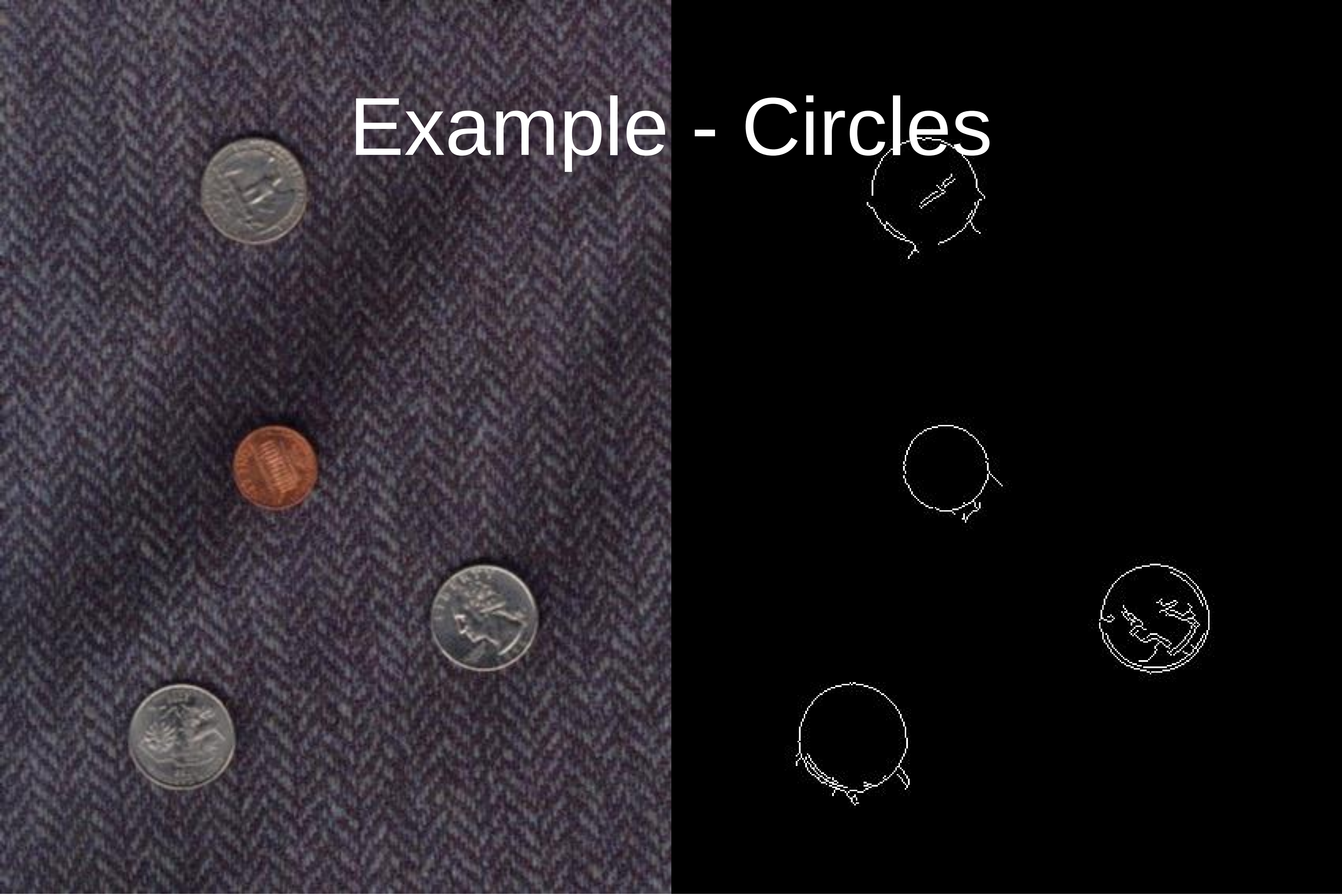


Example - Lines





Example - Circles

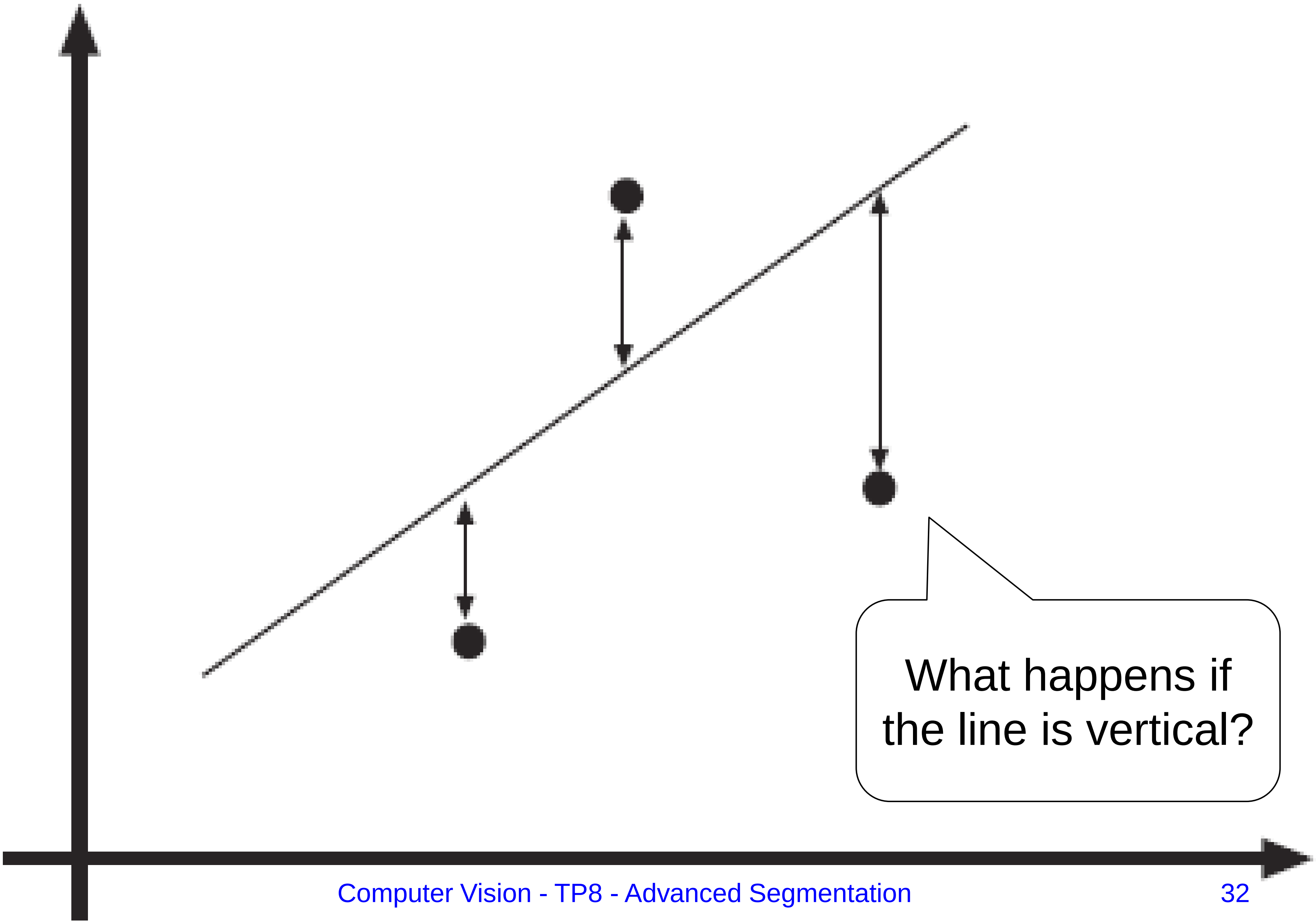


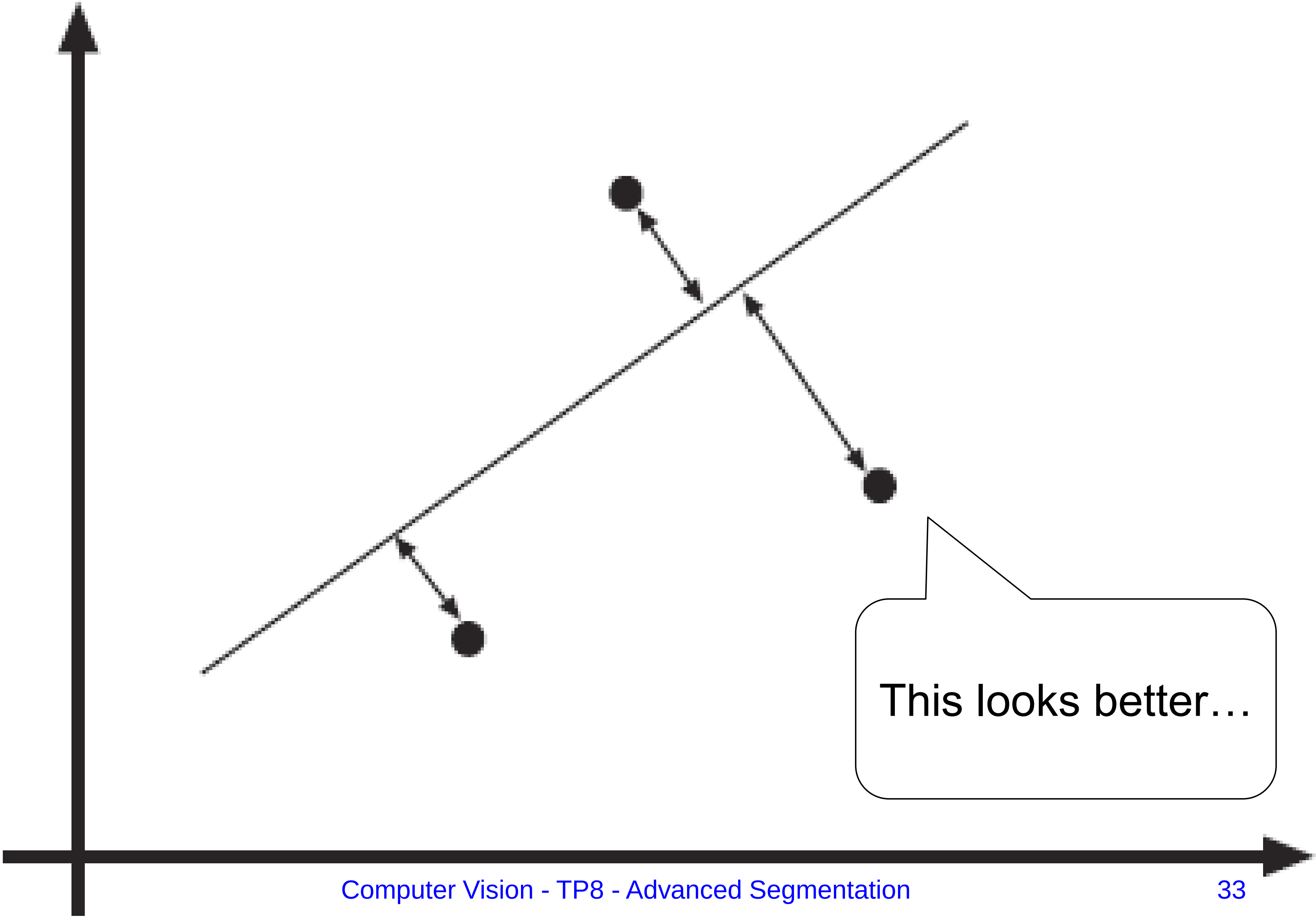
Least Squares Line Fitting

- Popular fitting procedure
- Simple but biased (why?)
- Consider a line:

$$y = ax + b$$

- What is the line that best predicts all observations (x_i, y_i)?
  - Minimize: $\sum_i (y_i - ax_i - b)^2$


What happens if the line is vertical?


This looks better…

Total Least Squares

- Works with the actual distance between the point and the line (rather than the vertical distance)
- Lines are represented as a collection of points where:

$$ax + by + c = 0$$

- And:

$$a^2 + b^2 = 1$$

Again... Minimize the error, obtain the line with the 'best fit'.

Point correspondence

- We can estimate a line but, **which points are on which line**?
- Usually:
  - We are fitting lines to edge points, so…
  - Edge directions can give us hints!
- What if I only have isolated points?
- Let's look at two options:
  - Incremental fitting
  - Allocating points to lines with K-means

Incremental Fitting

- Start with connected *curves* of edge points
- Fit *lines* to those points in that curve
- Incremental fitting:
  - Start at one end of the *curve*
  - Keep fitting all points in that curve to a line
  - Begin another line when the fitting deteriorates too much
- Great for closed curves!

```
Put all points on curve list, in order along the curve
empty the line point list
empty the line list

Until there are two few points on the curve
  Transfer first few points on the curve to the line point list
  fit line to line point list

  while fitted line is good enough
    transfer the next point on the curve
    to the line point list and refit the line
  end

  transfer last point back to curve
  attach line to line list
end
```

K-means allocation

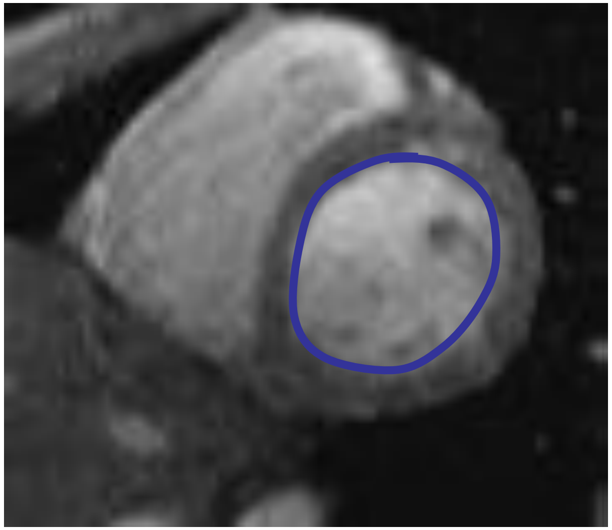
- What if points carry no hints about which line they lie on?
- Assume there are *k* lines for the *x* points.
- Minimize: $$\sum_{\text{lines}} \sum_{\text{points}} dist(\text{line}, \text{point})^2$$
- Iteration:
  - Allocate each point to the closest line
  - Fir the best line to the points allocated to each line

```
Hypothesize k lines (perhaps uniformly at random)
or
hypothesize an assignment of lines to points
and then fit lines using this assignment

Until convergence
  allocate each point to the closest line
  refit lines
```

Active Contours

- Goal: evolve the contour to fit exact object boundary



- How?
  - Reward solutions next to high image gradients
  - Punish solutions that deform shape too much
  - Iteratively find the ‘best’ solution to these requirements

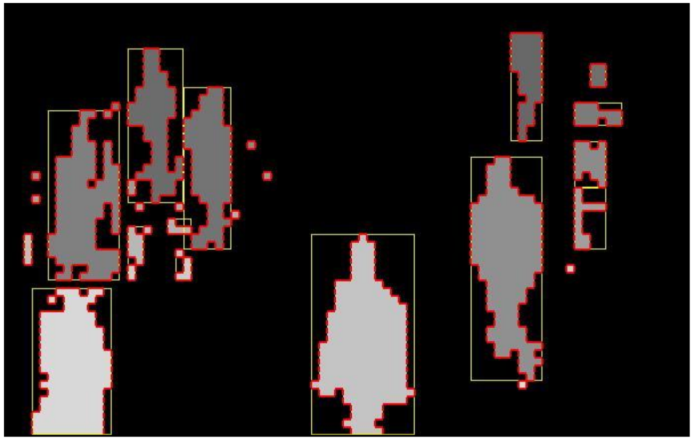
Topic: Semantic Segmentation

- Segmentation by Clustering
- Segmentation by Fitting
- Semantic Segmentation

Remember ‘Segmentation’?

- Separation of the image in different areas
  - Objects
  - **Areas with similar visual** or semantic **characteristics**

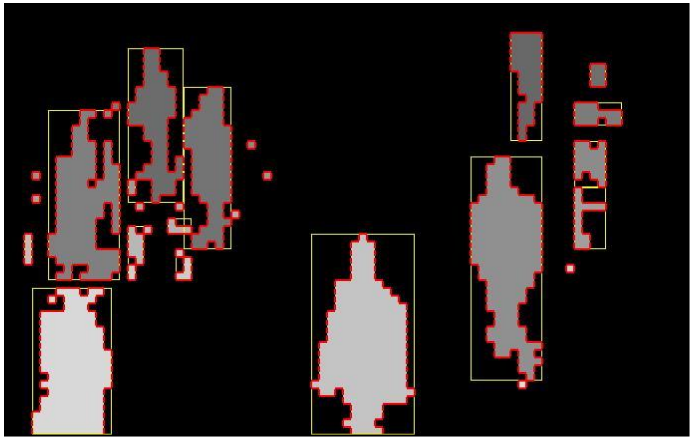
First form regions based on visual characteristics, then find the semantics of each region



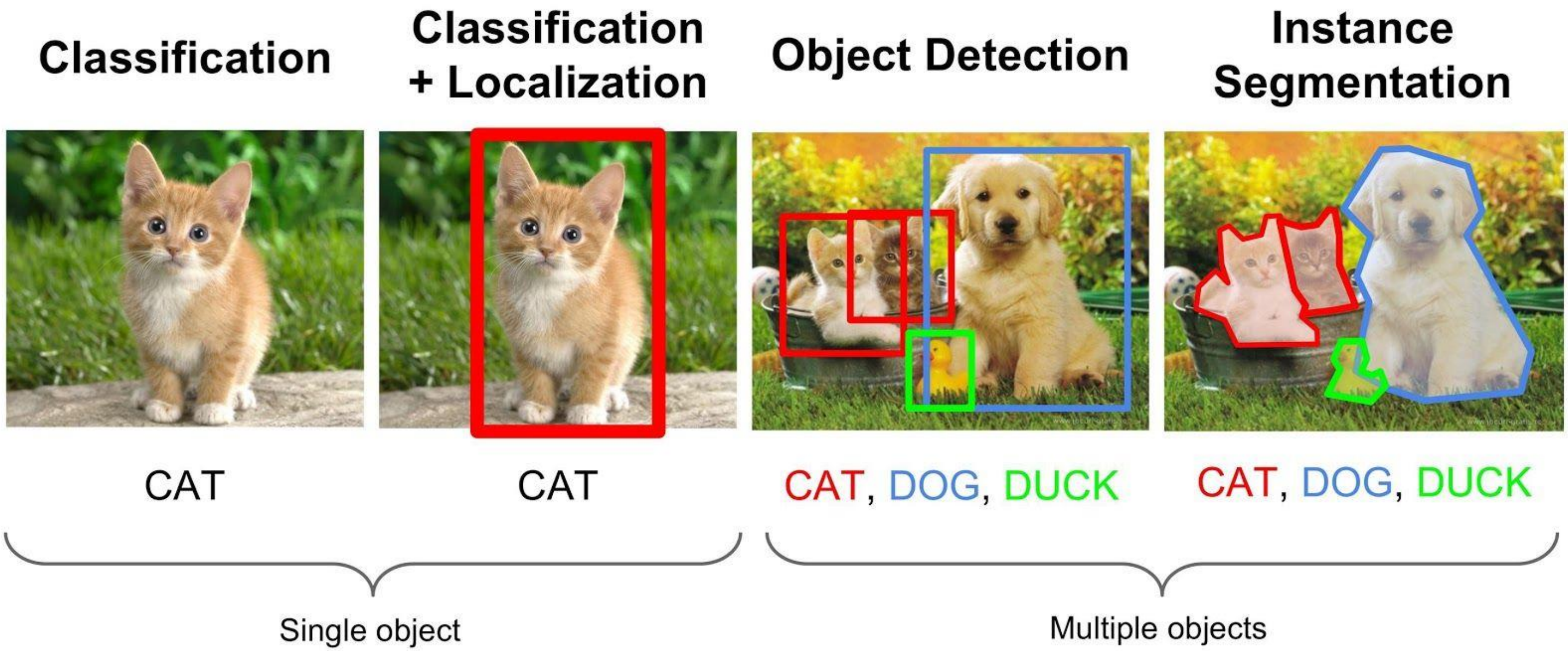
Semantic Segmentation

- Separation of the image in different areas
  - Objects
  - **Areas with similar** visual or **semantic characteristics**

First classify each pixel, and only then form regions (much harder!!)

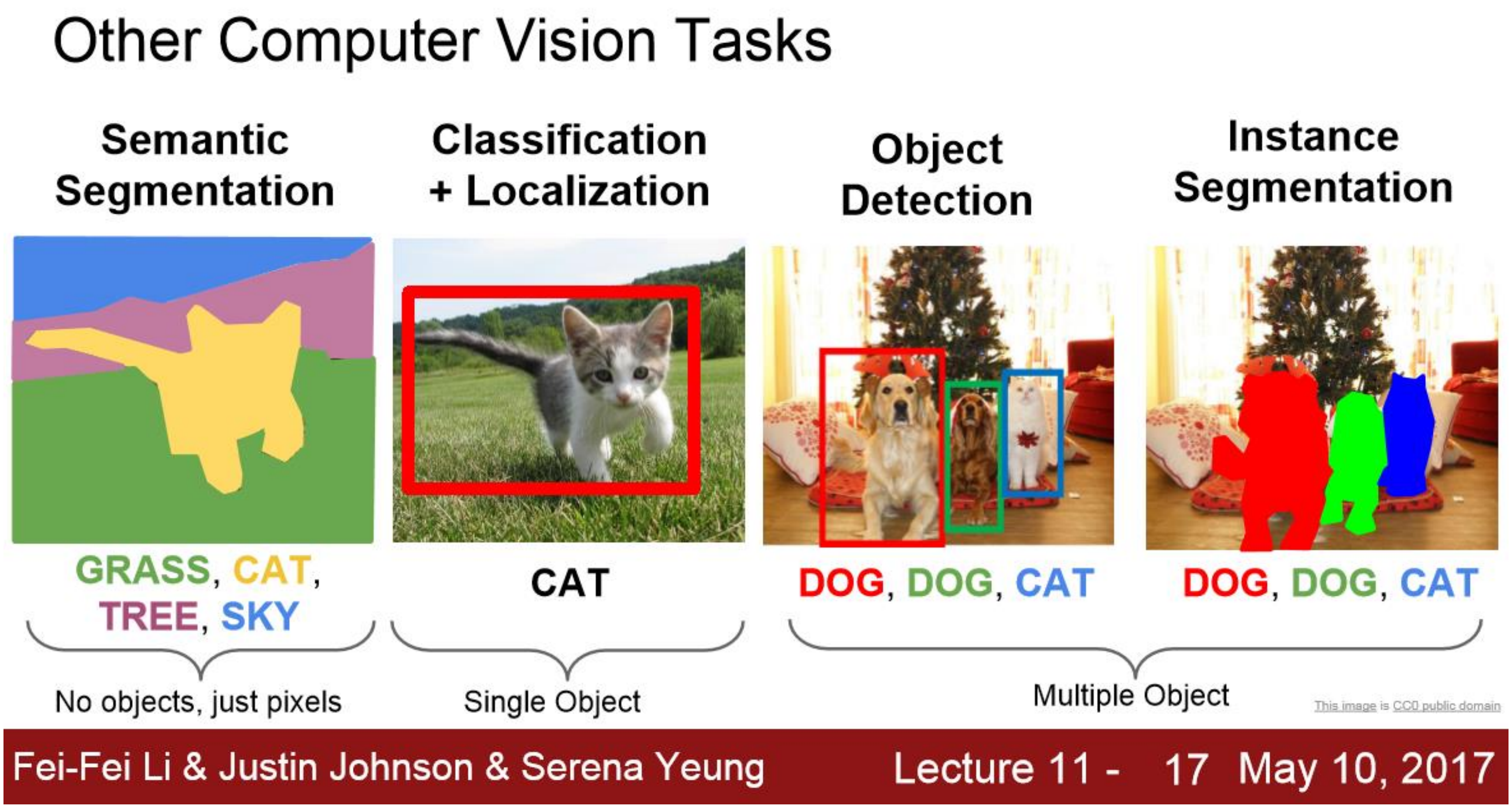


Classification and Segmentation



Source http://cs224d.stanford.edu/index.html

Semantic Segmentation



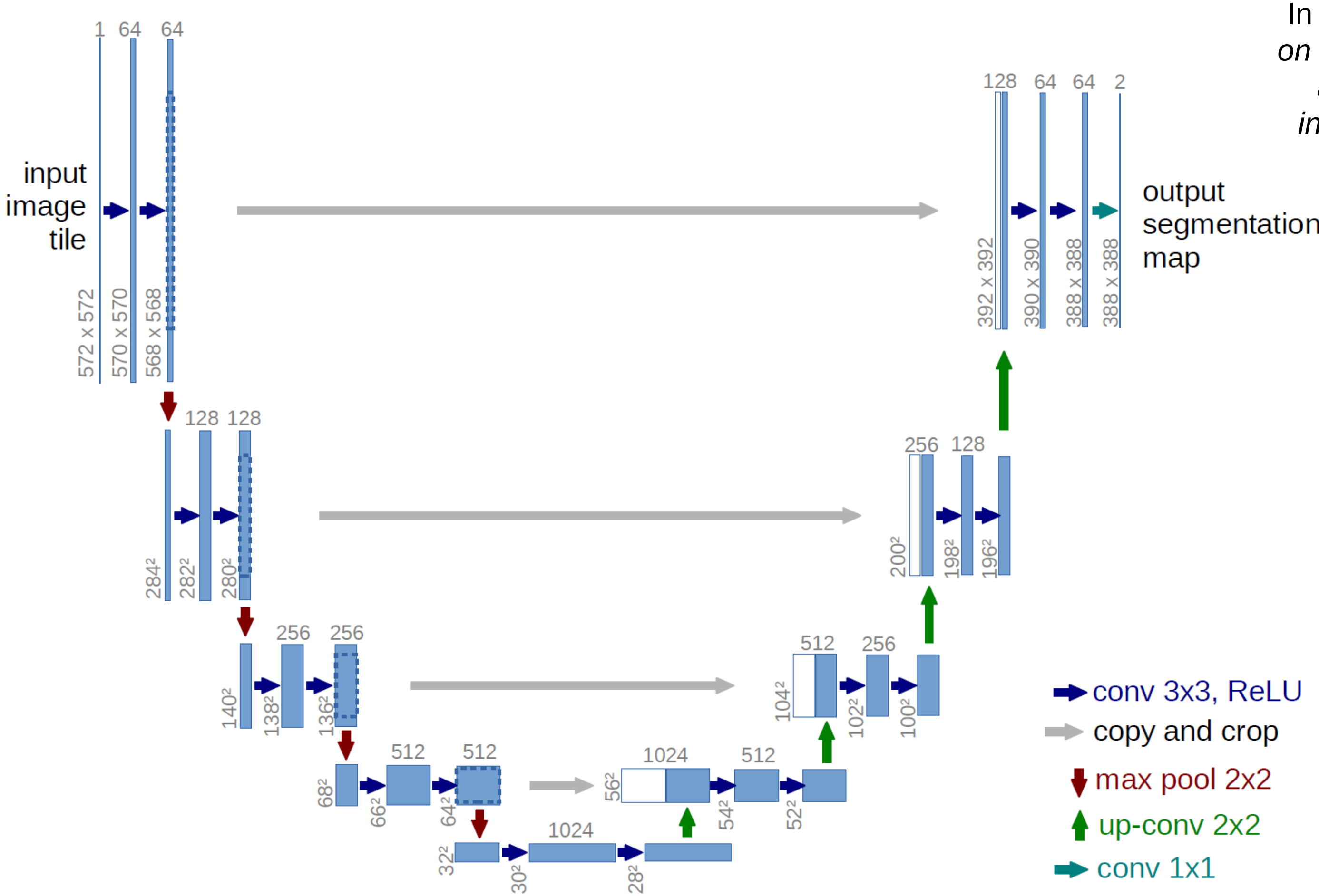
Source http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf

Semantic Segmentation

- Requires sophisticated pixel-level classification algorithms to be effective
- Powerful data-based approach to segmentation
- Fueled by recent advances in deep neural networks, such as U-NET

U-Net

O. Ronneberger, P. Fischer, and T. Brox. "U-net: Convolutional networks for biomedical image segmentation." In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234-241. Springer, Cham, 2015.



- Encoder-decoder structure

Resources

- Szeliski, "Computer Vision: Algorithms and Applications", Springer, 2011
  - Chapter 5 – "Segmentation"