Imperial College London

9 - Segmentation

Prof Peter YK Cheung

Dyson School of Design Engineering

URL: www.ee.ic.ac.uk/pcheung/teaching/DE4_DVS/
E-mail: p.cheung@imperial.ac.uk

What is Segmentation?

- Segmentation is the **division** of an image into **spatial region**.
- This is achieved by **grouping pixels with similar visual characteristics** into regions.
- If R represent the entire spatial region, and R_i, R_j represents two separate regions i and j, image segmentation is def $\bigcup_{i=1}^{n} R_i = R$. he following five conditions:

a) Every pixel must be in the region R:

b) R_i is a region where pixels are connected (i.e. in connected set).

c) Region R_i must be disjoint: $R_i \cap R_j = \emptyset$ for all i and $j, i \neq j$.

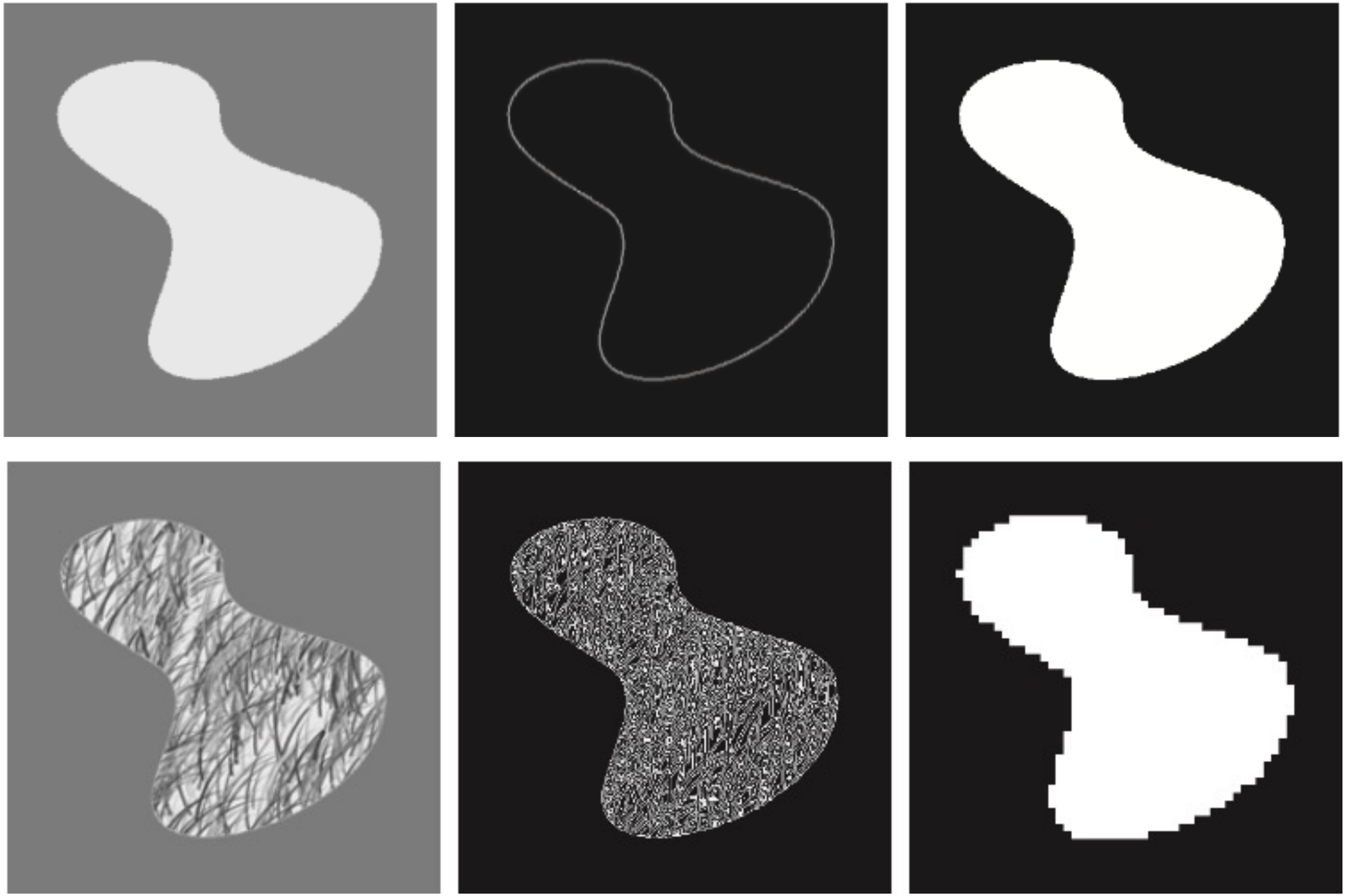
d) All pixels must share some common properties Q: $Q(R_i) = TRUE$ for all pixels in R_i.

e) Two adjacent regions R_i and R_j must be different in some properties Q.

Segmentation based on discontinuity

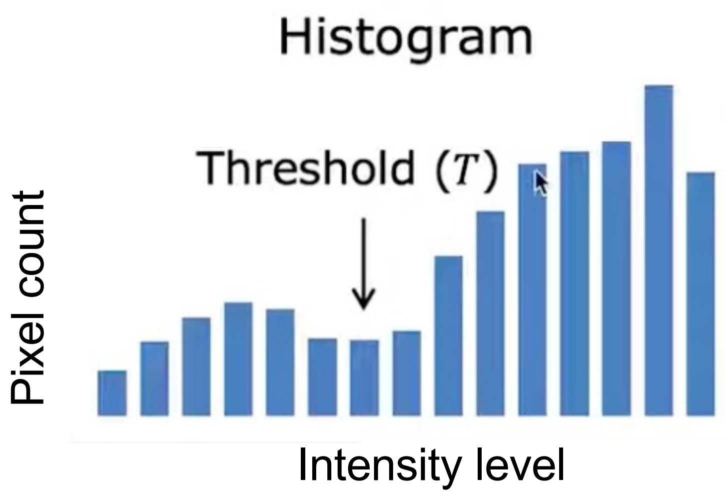
Intensity based segmentation

Texture based Segmentation (8 x 8 patch)

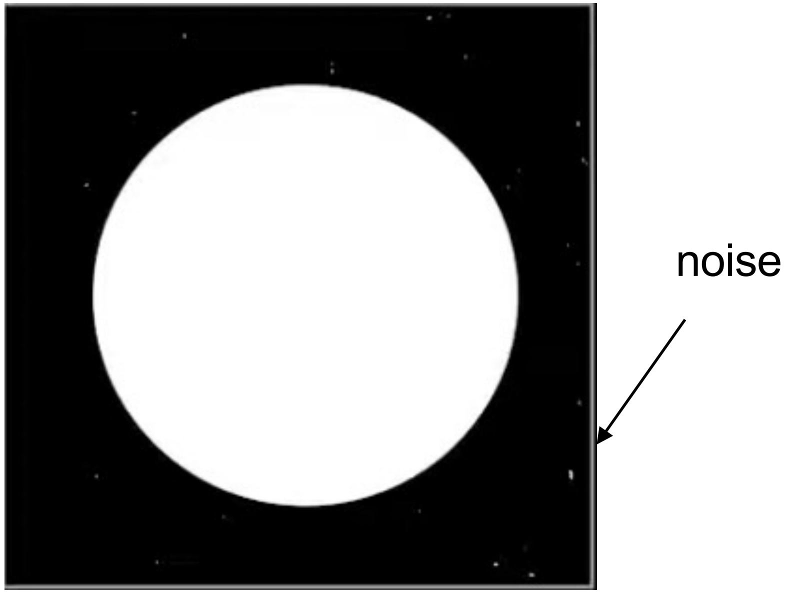


Segmentation using Thresholding

$f(x,y)$

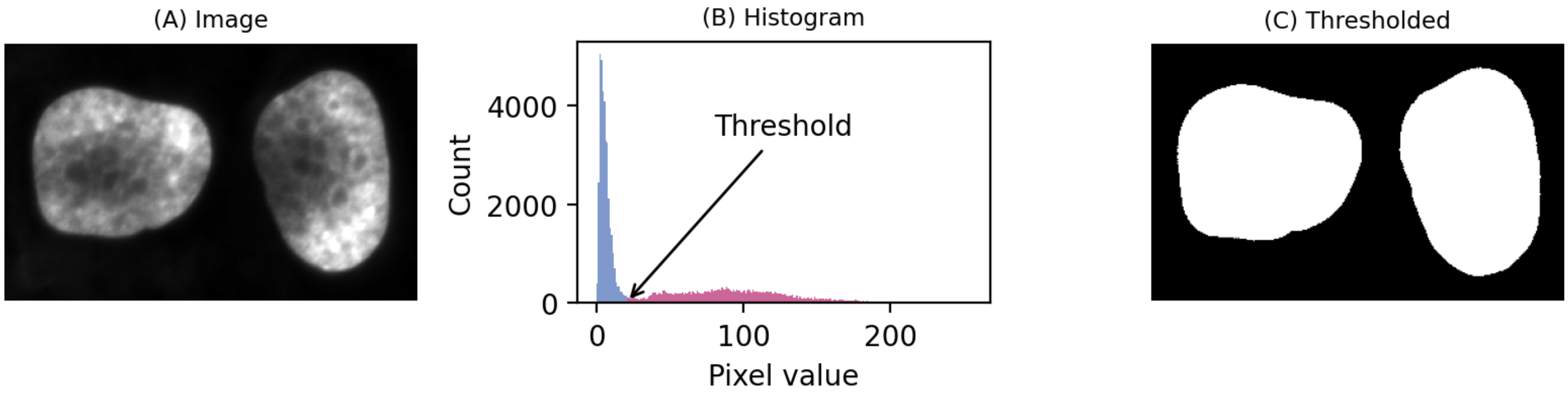


Binary image



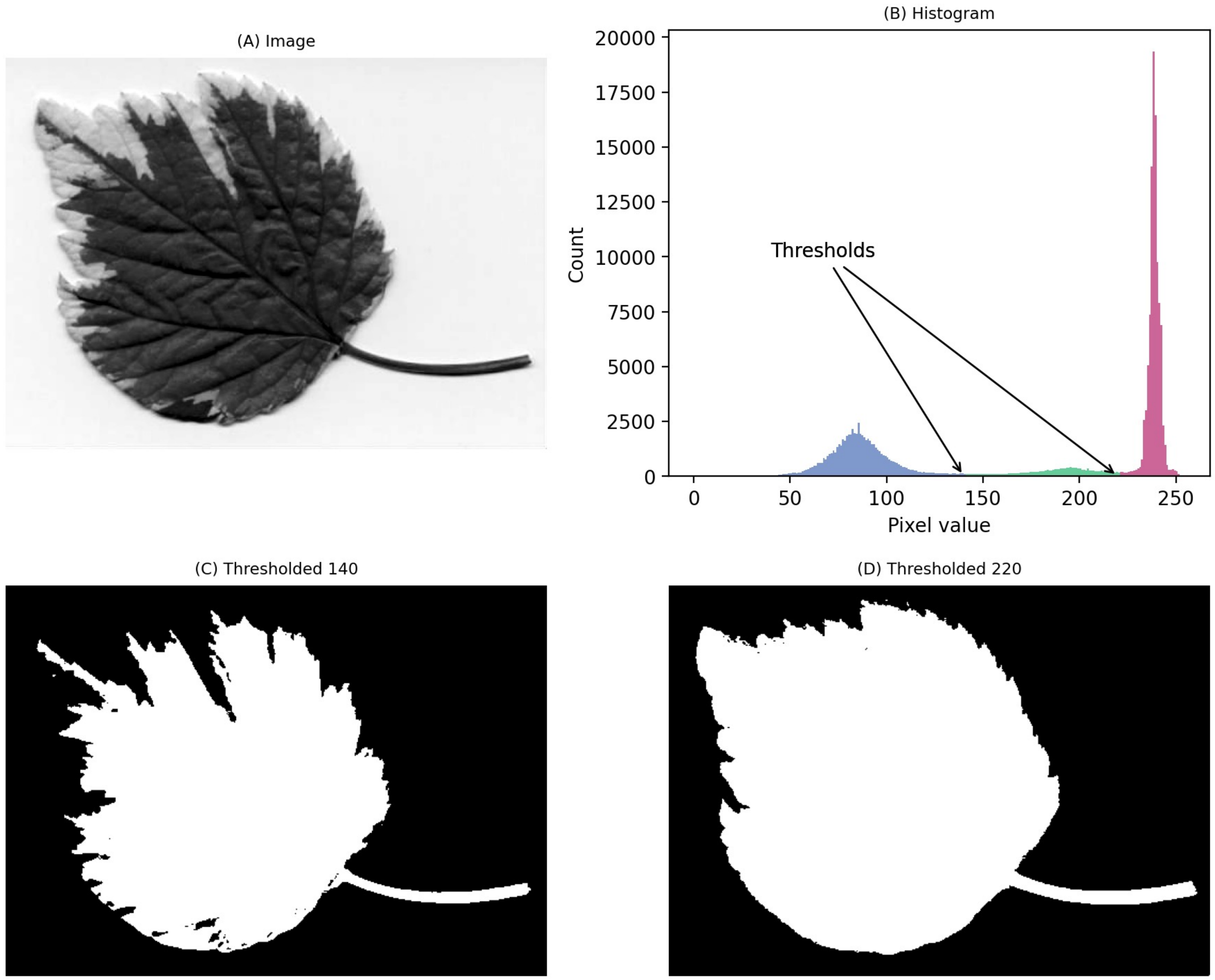
$g(x,y)$

Histogram can be a guide



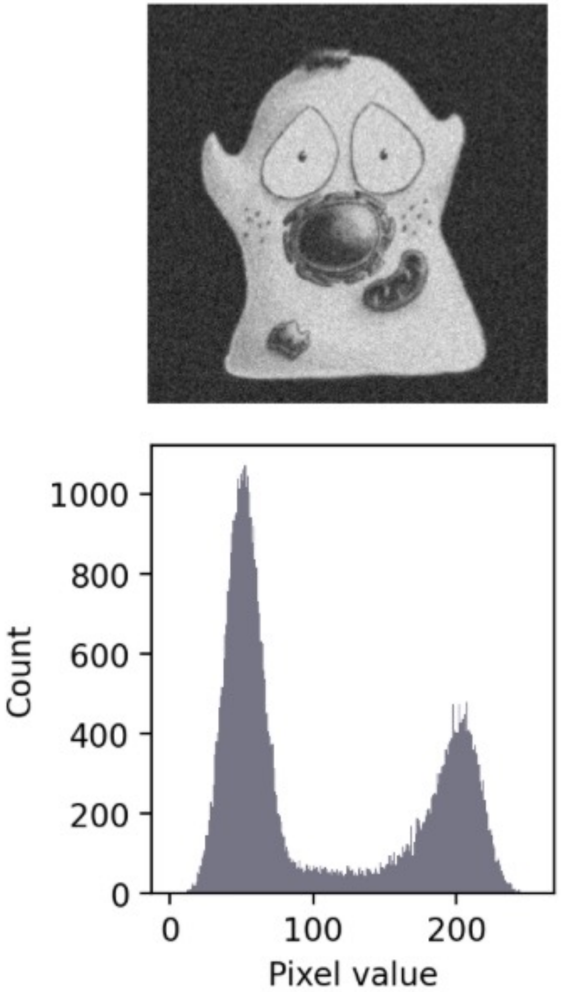
- The left image is to be segmented.
- Histogram is bimodal - the background and the object occupying two distinct regions.
- Putting the threshold between the two “hills” will easily segment the image.

What if it is not bimodal?

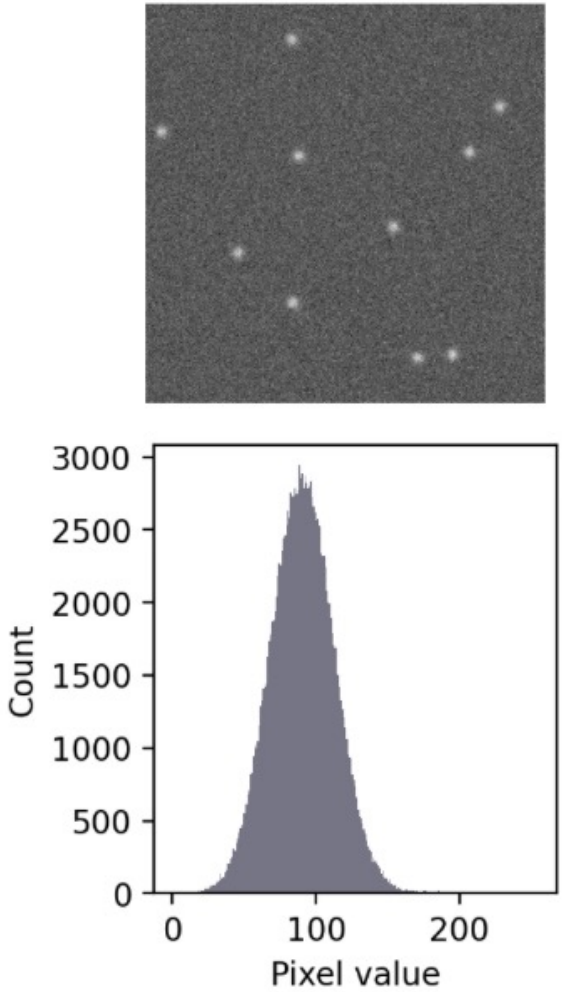


Three Different Type of Images & Histograms

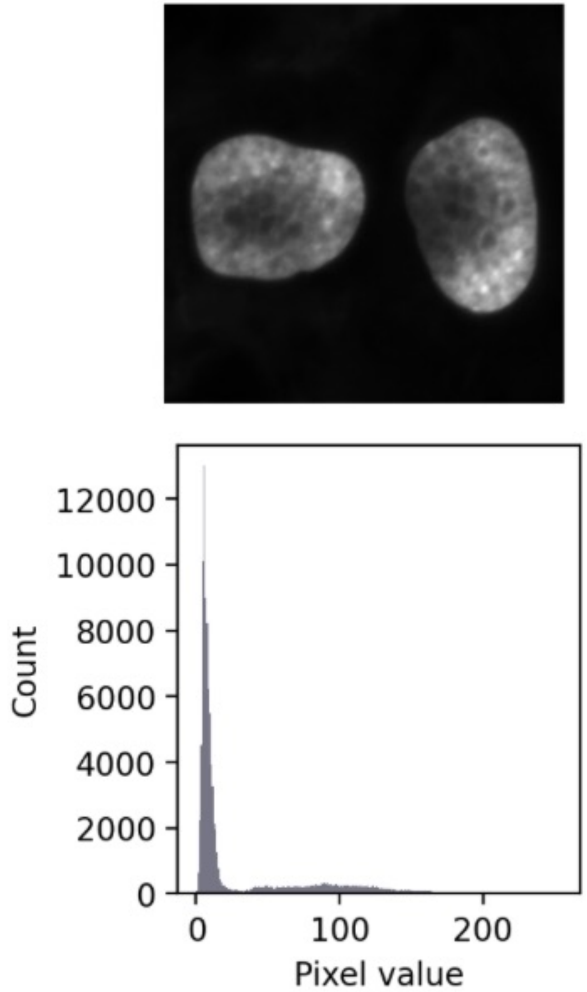
Bimodal



Unimodal

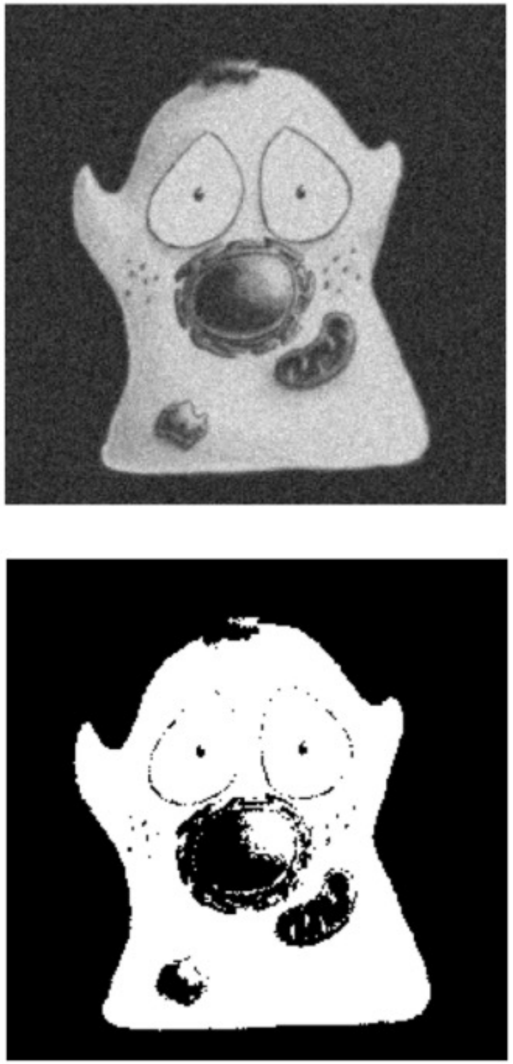


Clear background
& object

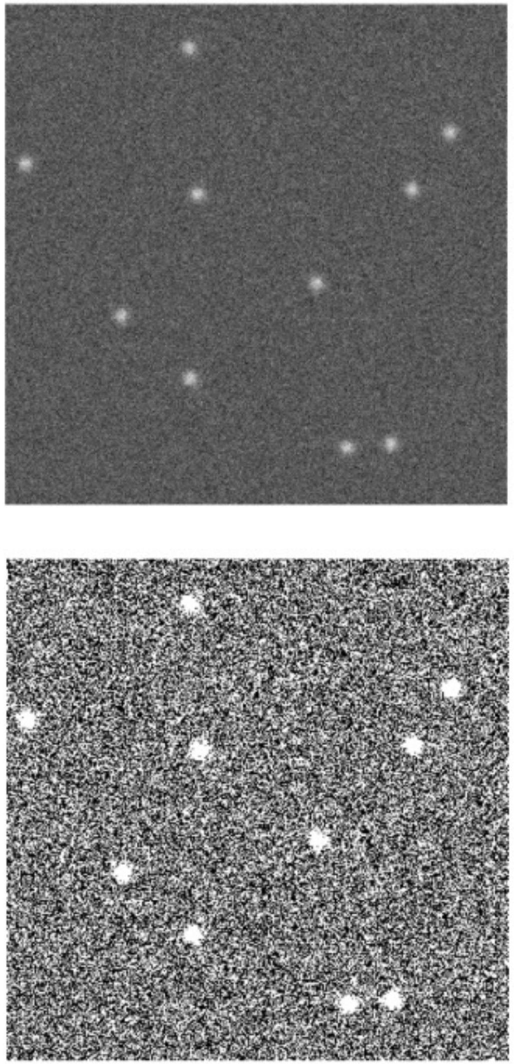


Otsu’s method to find threshold

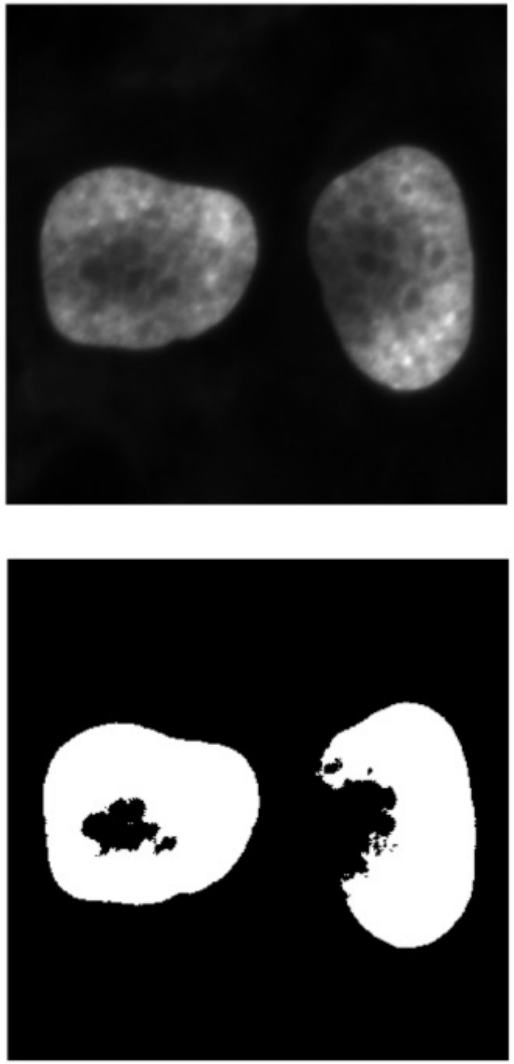
Threshold = 123



Threshold = 90



Threshold = 59



Variable Thresholding based on local statistics

- Threshold $T_{xy} = a\sigma_{xy} + bm_{xy}$, where σ_{xy} is local standard deviation, m_{xy} is local mean intensity, and a, b are positive constants.

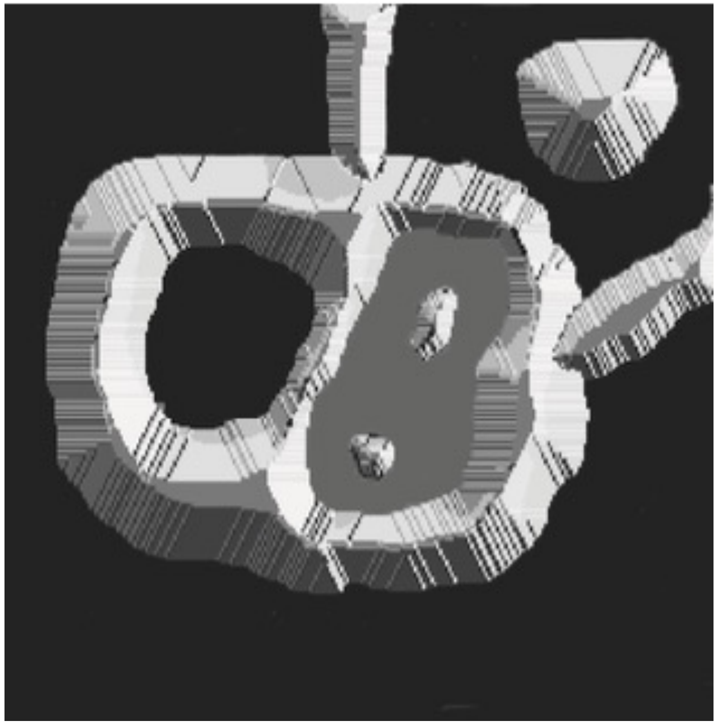
Image of yeast cells

Using dual global thresholds

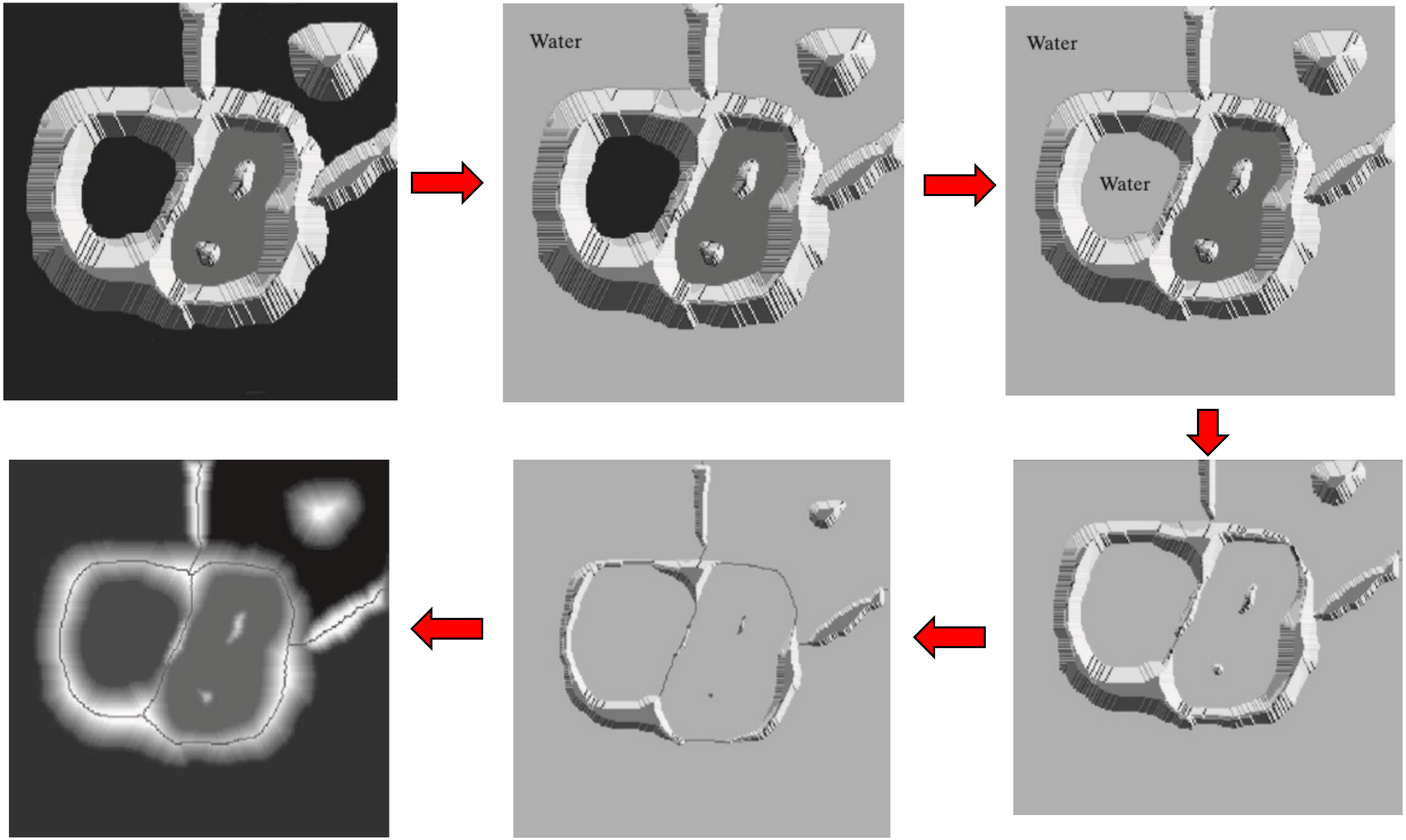
Using local thresholds

Segmentation using Watershed Transform (1)

- A watershed is the ridge that divides areas drained by different river systems.
- A catchment basin is the geographical area draining into a river or reservoir.
- The segmented regions are the basins that catch the rainwater as water rises.
- The watershed ridge line partition the image into regions thus achieving segmentation.



Segmentation using Watershed Transform (2)

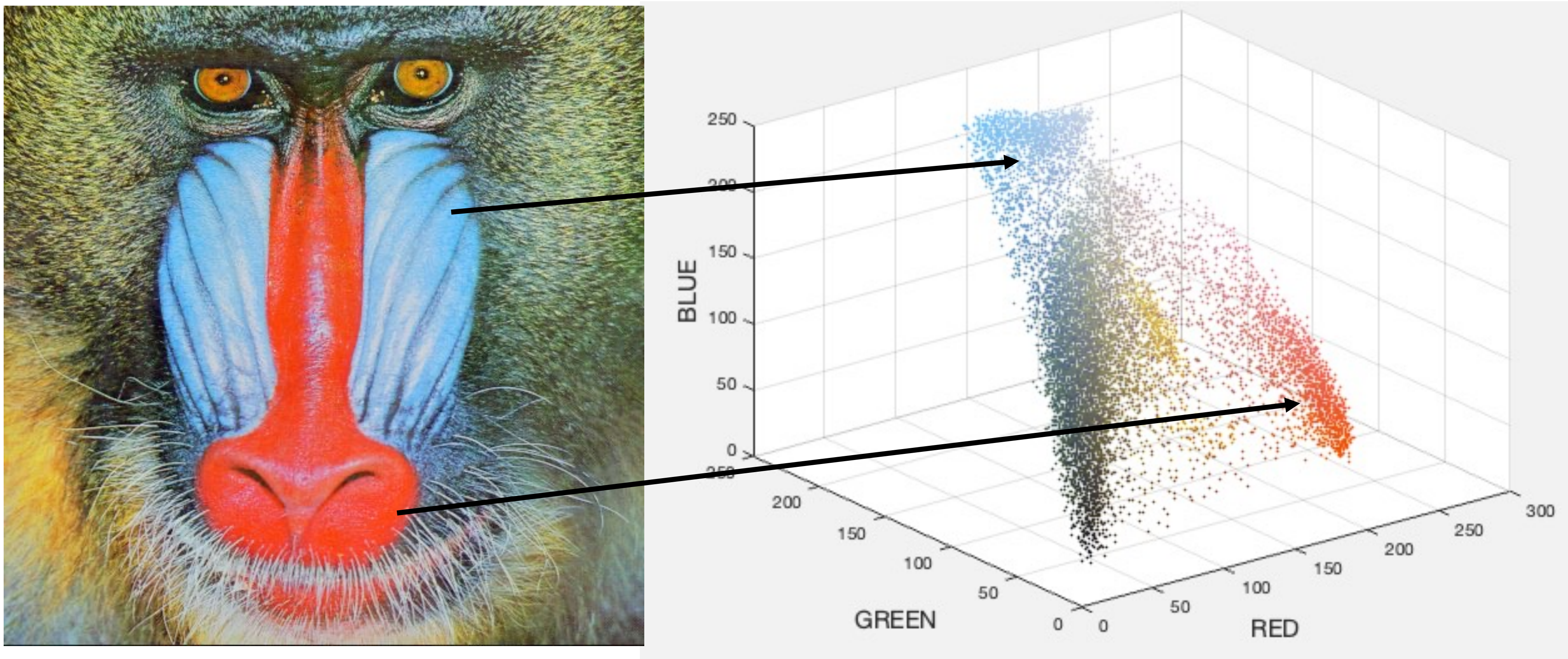


Segmentation based on clustering

- Visual characteristics: Intensity, colour, position, texture, motion, depth
- Each pixel as feature vector: [R, G, B, x, y …..]

Colour image

Pixel colour in feature vector [R G B]

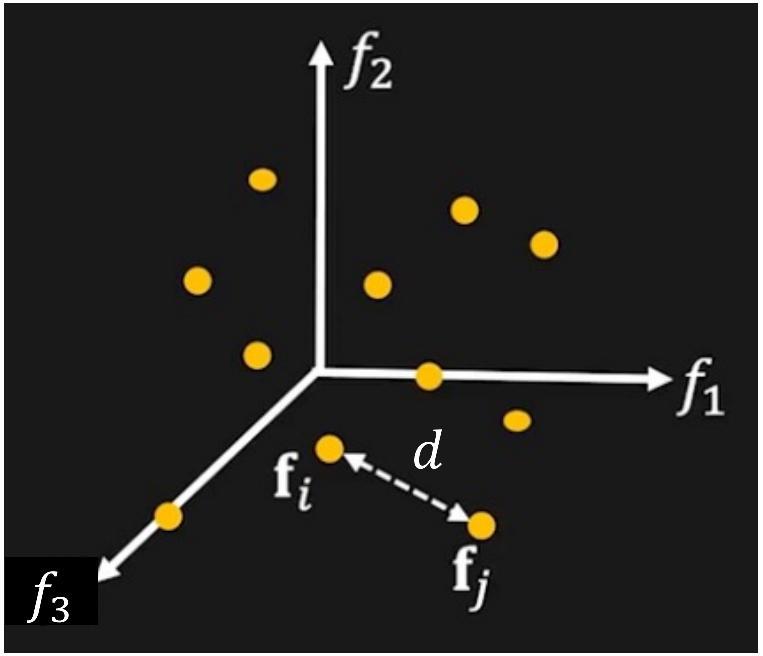


Pixel Similarity Measure

- Let f_i and f_j be the feature vector for pixels i and j respectively.
- The Euclidian distance between f_i and f_j is given by:

$$d(f_i, f_j) = \sqrt{\sum_k (f_{ik} - f_{jk})^2}$$

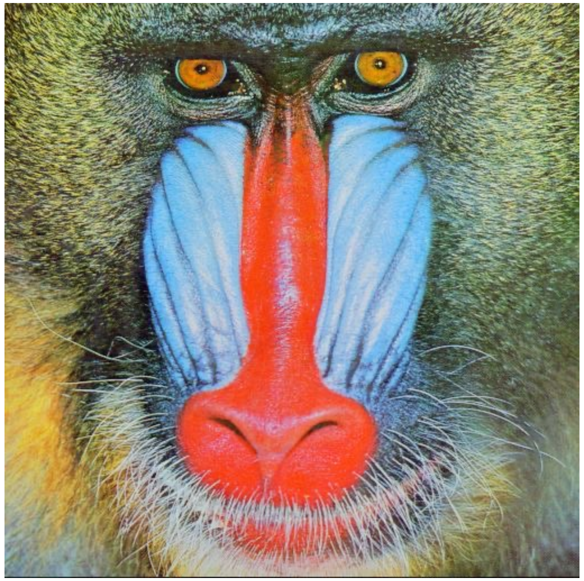
- Smaller the value of d, the greater the similarity.



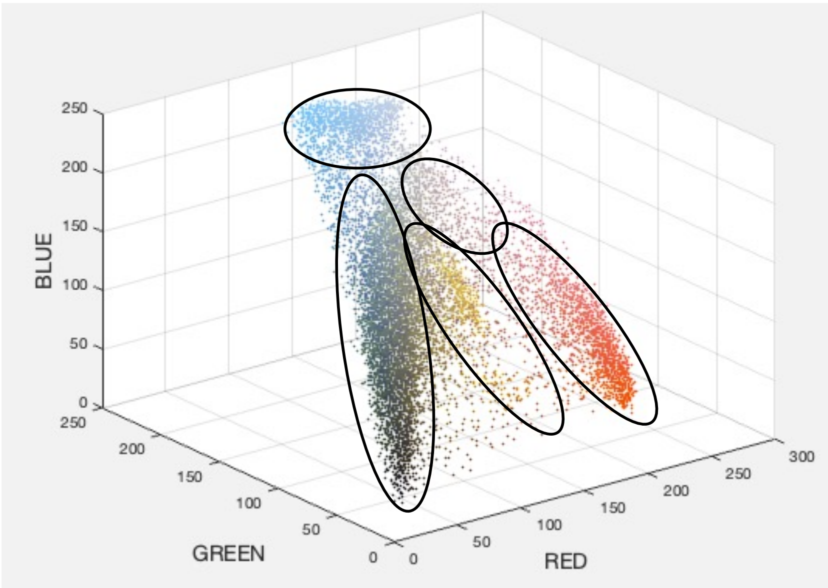
Segmentation by Clustering [R G B] vectors

- Group together those pixels that have high similarity in colour (i.e. short $d(r, g, b)$ distances) to form clusters.
- Assign a "mean" colour to each cluster.
- Now we have a segmented image by colour.

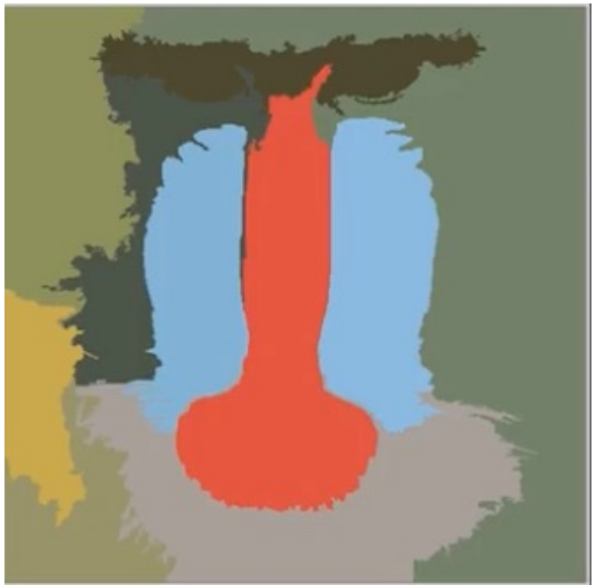
Original image



Features divided into 5 clusters

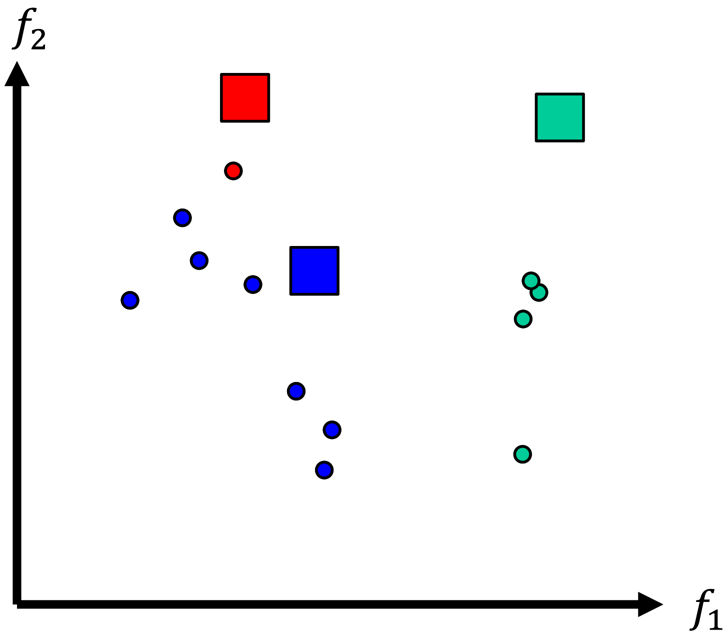


Segmented by colour



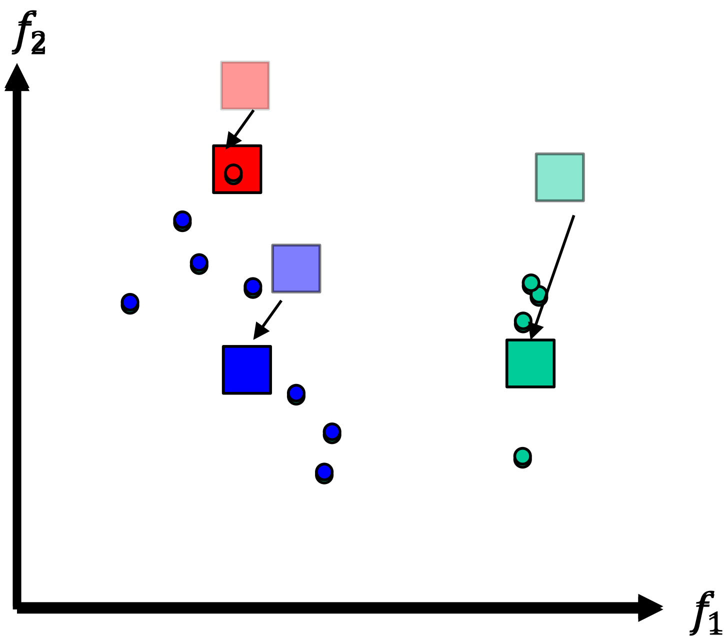
Segmentation by k-Means Clustering (1)

- Segment image into k clusters using pixel characteristics. Assume k = 3
- Step 1: Generate 3 random initial means (centroids) in feature space.
- Step 2: Cluster each pixel to the mean with shortest distance to form 3 clusters as shown.

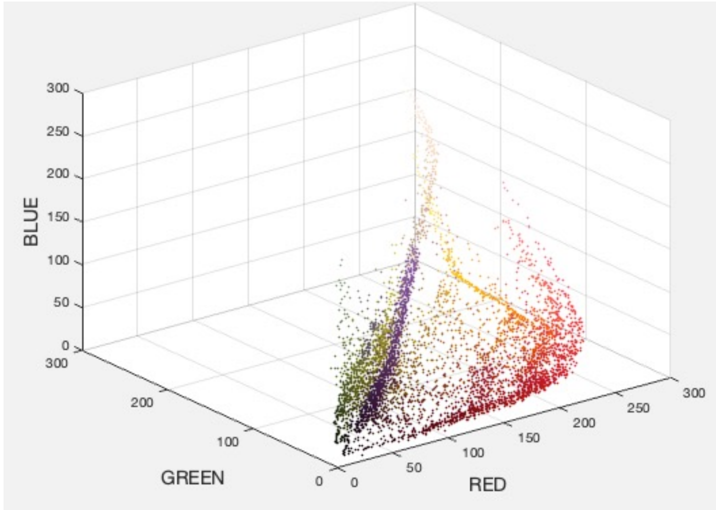


Segmentation by 3-Means Clustering (2)

- Step 3: Recompute the mean of each cluster
- Step 4: Repeat steps 2 and 3 until mean (or centroid) values change below a small margin. Then segmentation is completed and converges to a final solution.
- Needs to determine k. Also need to select initial values.
- Best initial values: Perform k-means clustering on a subset of pixels, and use that solution as initial values.



Example of k-Means Clustering (k = 16)

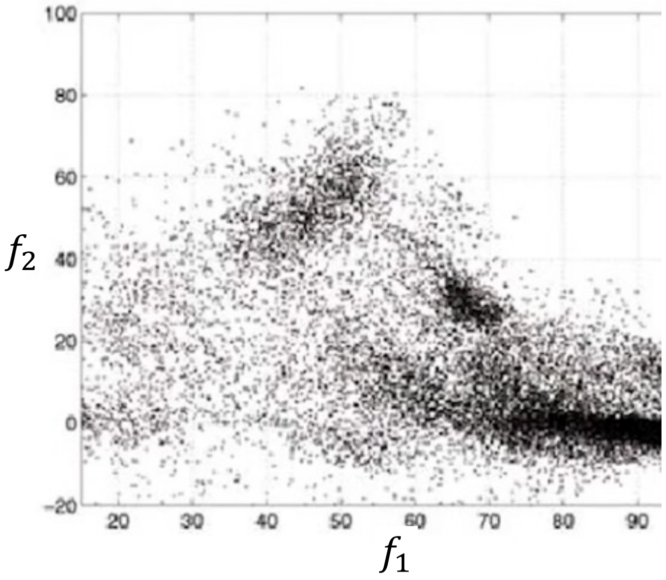


k = 16, {R, G, B} space only

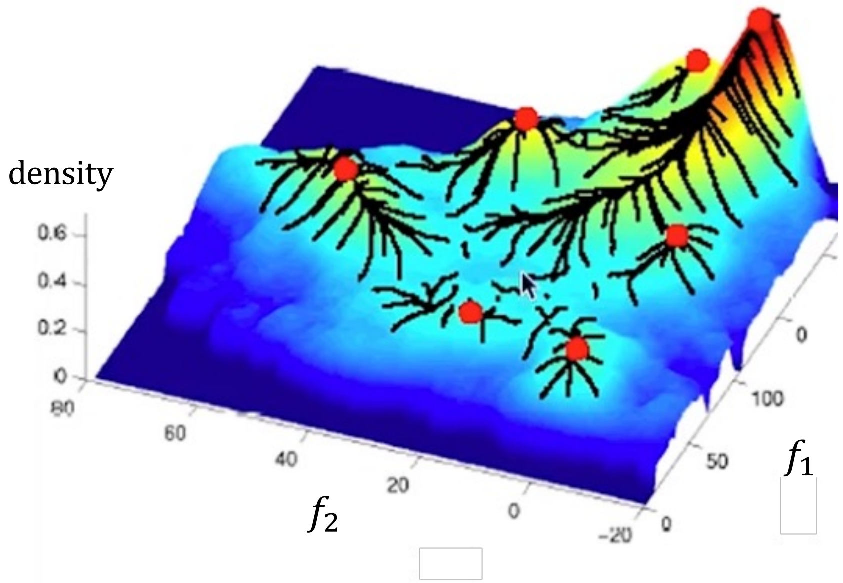
k = 16, {R, G, B, x, y} space

Idea behind Mean Shift Clustering

Pixel feature distribution

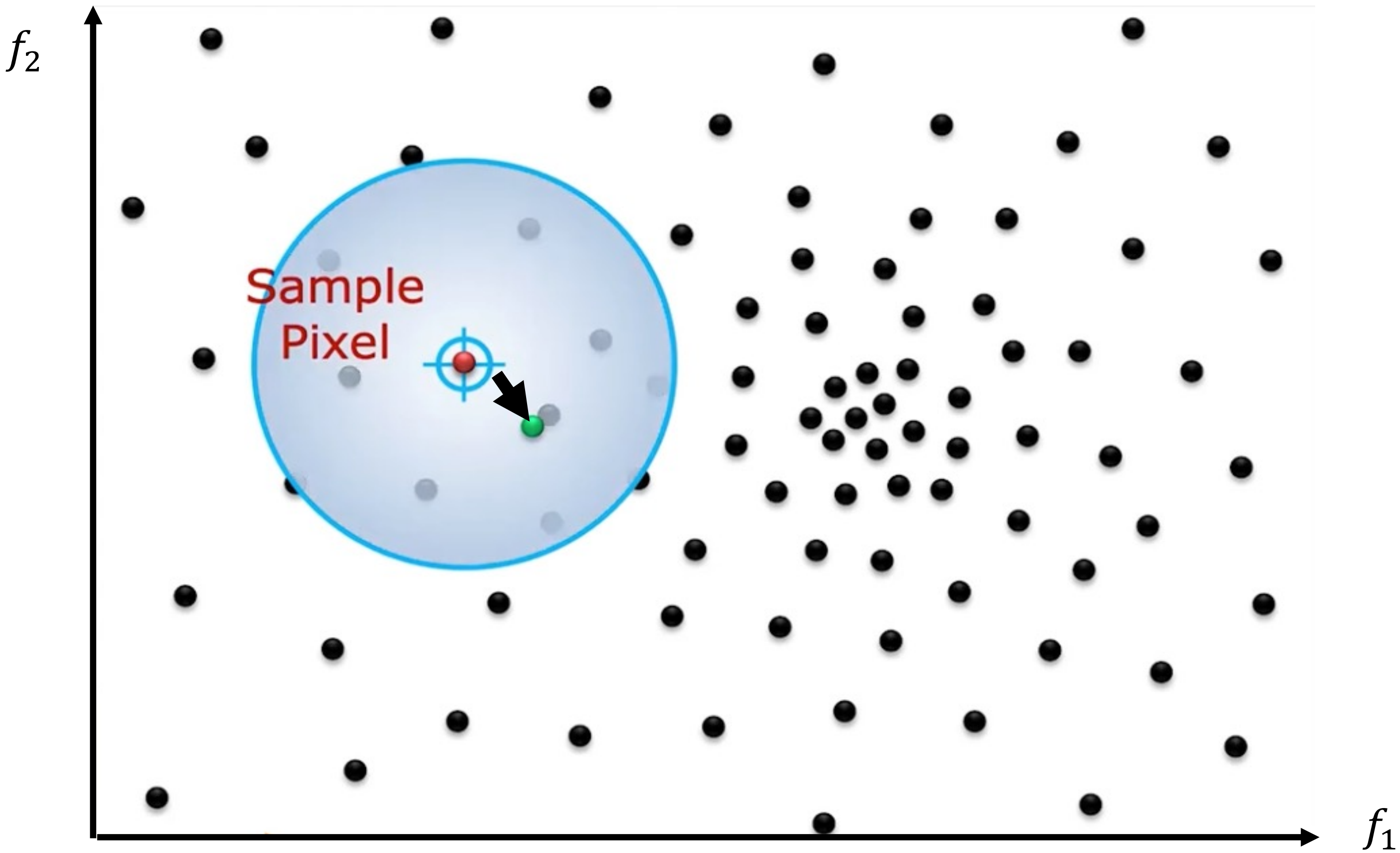


Density (normalized to [0 1])



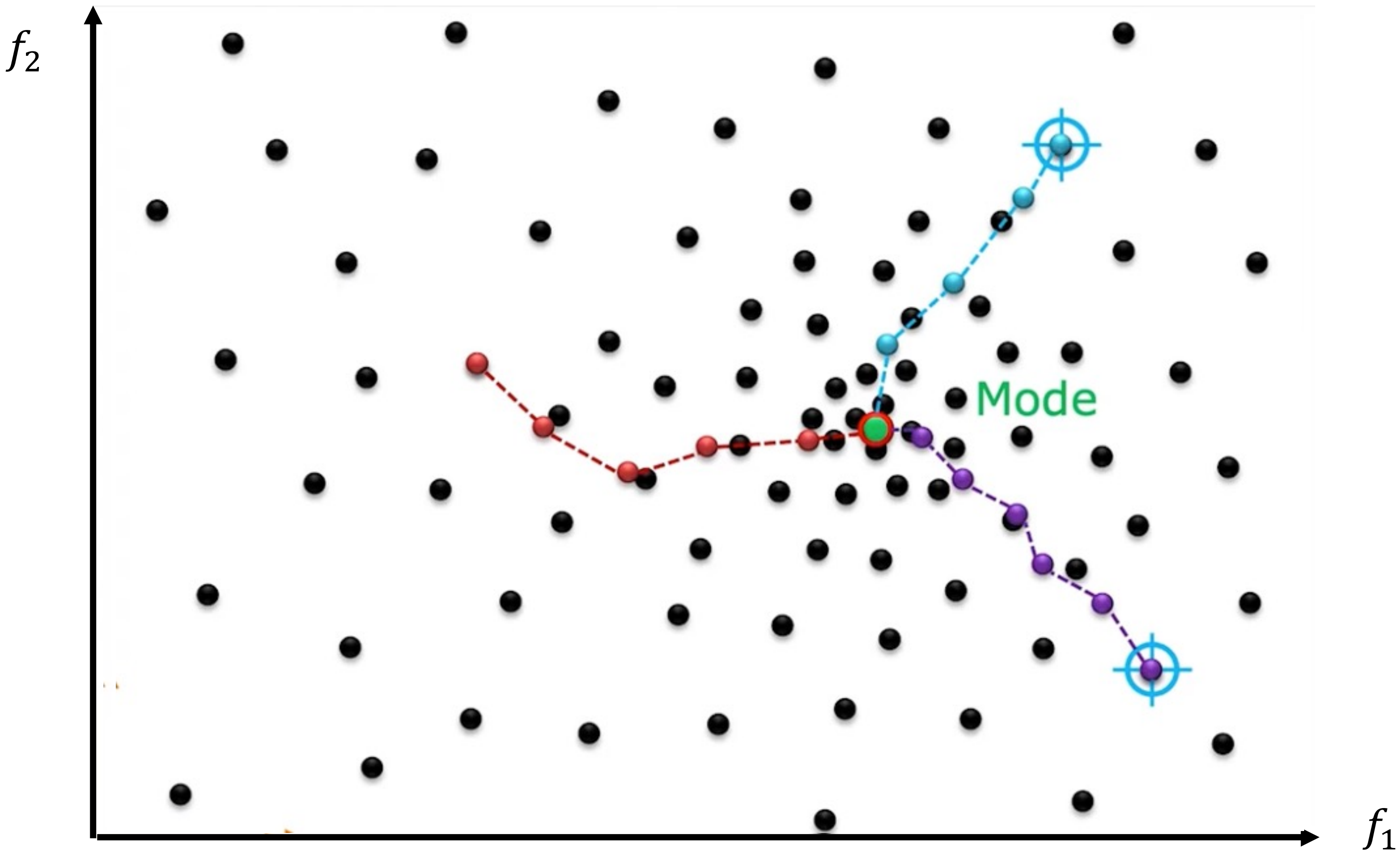
- Mean Shift determines how many clusters automatically depending on image.
- Each hill of the distribution represents a cluster.
- Each peak of the hill represents “mean” or ”centroid” of a cluster.
- Each pixel climbs the steepest hill within its neighbourhood.
- Pixel takes on the the peak of the hill in its cluster.

Mean Shift Hill Climbing



- Consider one pixel (red) in feature space. Form a window centred at pixel with with w.
- Calculate the weighted means of feature parameters to find a new location in green.
- Shift the original pixel in feature space to the new mean (hence name “mean shift”).
- Keep doing this until there is no shift – we are done for this pixel!

Mean Shift Hill Climbing



- Therefore, the original pixel (in red), successfully climb the hill to the MODE (or peak).
- Other pixels nearby also eventually reach the same peak.
- These are ALL belonging the same cluster. They are then assigned the same feature vector as that of the peak.

Example of Mean Shift Clustering

k-Means (k = 16)

Mean Shift ($w = 21$)