Image Segmentation

Why are we studying this?
What's the connection to recognition?

Scientific motivation:
Most studies focus on isolated object recognition.
But, objects are typically immersed in an environment.

Conventional belief:

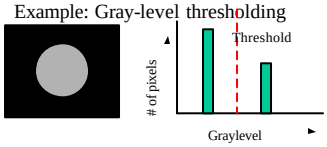
Image → Segmentation → Image Partitioned Into regions → Region to Model matching → Object recognized

Practical motivation:
Digital compositing for special effects in movies.

Basic idea behind current segmentation strategies:

Partition image into clusters of 'similar' pixels.

Example: Gray-level thresholding



All pixels with graylevel A belong to region 1 and all those with graylevel B belong to region 2.

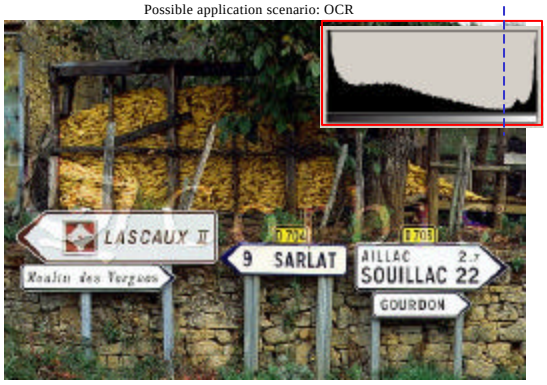
Possible application scenario: OCR

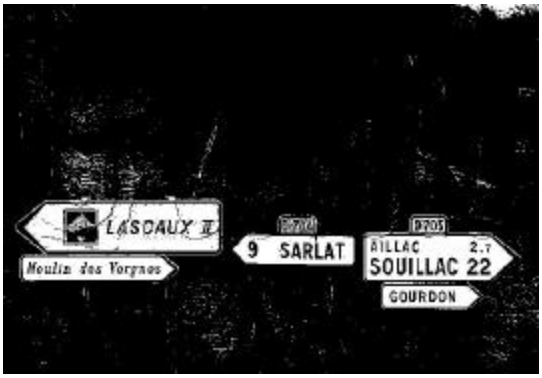
Graylevel histogram should be clearly bimodal, with the modes corresponding to different partitions. The threshold can be set to a value in the valley between the modes.

Iterative estimation of threshold value:
Ti+1 = (mean backgroundi + mean objecti)/2
Iterate until Ti and Ti+1 are sufficiently close.
Initial condition: Background is the four corner points.

Example: Gray-level thresholding

Possible application scenario: OCR

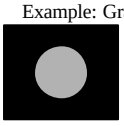


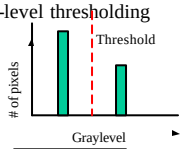


Basic idea behind current segmentation strategies:

Partition image into clusters of 'similar' pixels.

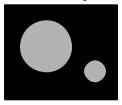
Example: Gray-level thresholding



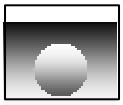


All pixels with graylevel A belong to region 1 and all those with graylevel B belong to region 2.

Problems with this approach:

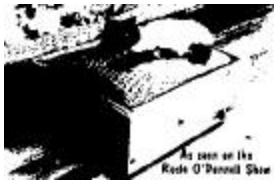


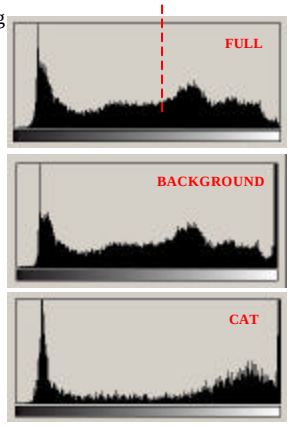
Multiple regions with similar graylevels



Regions with multiple graylevels

Example: Gray-level thresholding



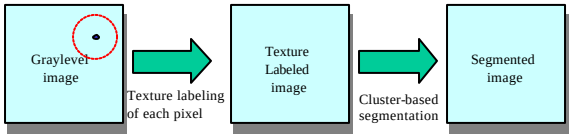


Possible solution:

Spatially varying threshold.

But, what about non-intensity based segmentation problems?

Texture-based segmentation:

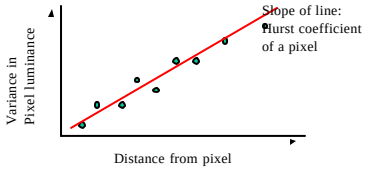


How can we texture-label pixels?

How can we texture-label pixels?

No principled answers yet. Some ad-hoc suggestions…

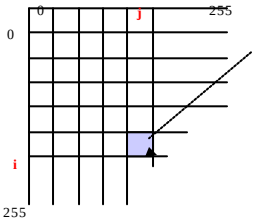
HURST COEFFICIENTS:



How can we texture-label pixels?

HARALICK'S CO-OCCURRENCY MATRICES:

The matrices reflect gray-level variations induced by texture over a region. Haralick demonstrated the approach on a satellite image classification task (8 terrain classes: old residential, new residential, lake, swamp, marsh, Railroad yard, scrub, woodland). Results: 82% classification accuracy.



$P_{Theta, dist}(i, j)$

# of pixel pairs with graylevels i and j at angle Theta and inter-pixel distance dist

Example of a co-occurrency matrix:

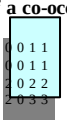
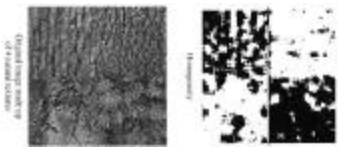


Image
(pixels can take on only 4 gray values)

$$P_{0,1}(i, j) = \begin{vmatrix} 4 & 2 & 1 & 1 \\ 0 & 4 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{vmatrix}$$

Co-occurrency matrices for various thetas and dists capture the textural variations in a region. We can define different attributes as functions of the co-occurrency matrices

'Homogeneity' = sum of diagonal terms / sum of entire matrix



Thresholding on one attribute is a very simple case of clustering into two groups.
In general, there may be many attributes and multiple clusters.

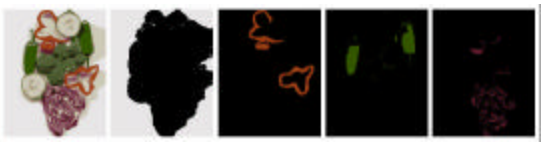
Generalizing the clustering problem of segmentation:

The k-means algorithm:

Arbitrarily choose k data points to act as cluster centers

Until the cluster centers are unchanged:
Allocate each data-point to cluster whose center is nearest

Replace the cluster centers with the mean of the elements in their clusters
end



Open issues:

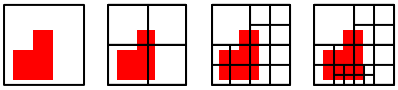
Segmentation via pixel-based attributes or region-based attributes?

Segmentation via clustering in attribute space or image space?

Segmentation by clustering in image space:

…because clustering in attribute space may not reflect clustering in image space. Clustering in attribute space can lead to unconnected regions in image space (but this may be useful for handling occlusions).

Implemented as region splitting (quad trees) or region growing (flood fill)



Splitting regions recursively if they fail a homogeneity test

Future challenge:

Incorporating object-specific knowledge in image segmentation

Application: Image retrieval

Image classification can be thought of as based on clustering using the kinds of color/intensity/texture attributes we have discussed (for instance through the use of global color histograms).

At a finer level, images may be described as sets of regions with different attributes and queries may be based on a subset of these regions.