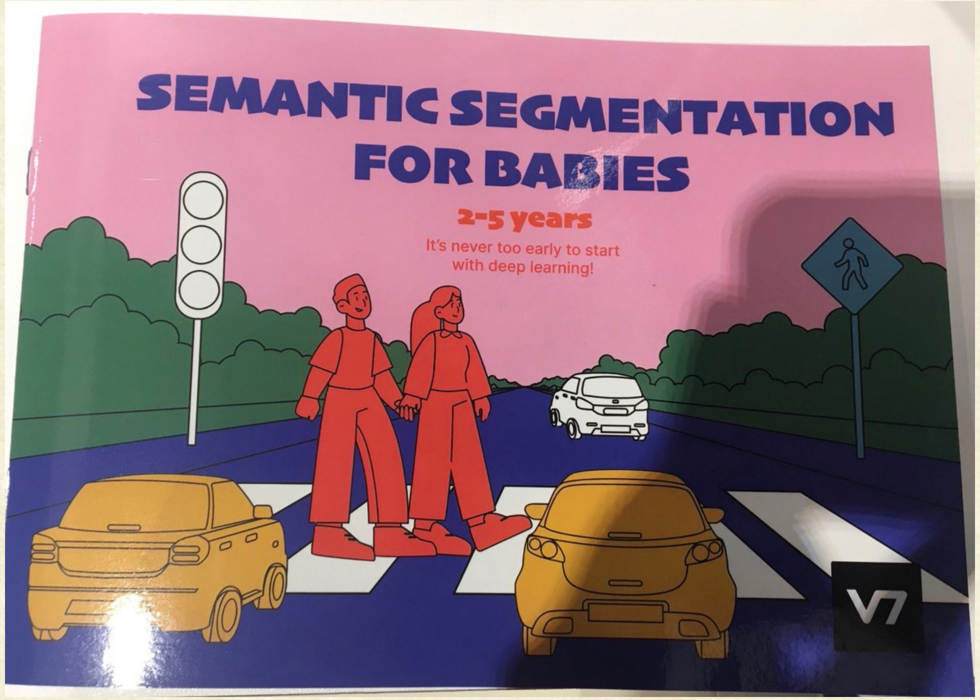
CS7.505: Computer Vision

Spring 2024: Segmentation as Pixel Labelling



Makarand Tapaswi
Katha AI Lab, CVIT,
IIIT Hyderabad

Anoop Namboodiri
Biometrics and Secure ID Lab, CVIT,
IIIT Hyderabad

Three "Urges" on seeing a Picture*

1. **To group** proximate and similar parts of the image into meaningful "regions".

   Called segmentation in computer vision.

2. **To connect to memory** to recollect previously seen "objects".

   Called recognition in computer vision.

3. **To measure** quantitative aspects such as number and sizes of objects, distances to/between them, etc.

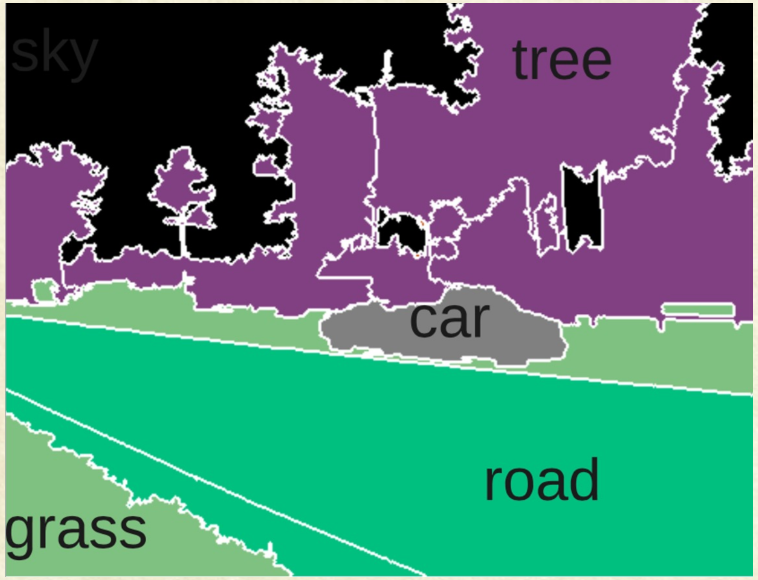
   Called reconstruction in computer vision.

**Jitendra Malik; Mysore Park, Dec. 2011*

Urge to Group

- We don’t see individual pixels (like the computer does!)
- We see groups of pixels together
- What is the basis for “correct” grouping?

Urge to Group



- Group similar pixels together as objects.
- Group semantically meaningful pixels together as objects.
- Is appearance similarity the same as semantic similarity?

Segmentation

- Dividing an image into semantically meaningful regions.

Types of Segmentation

- Classification-based
  - Label pixels based on region properties
  - Label each pixel based on object models
- Region-based
  - Region growing and splitting
- Boundary-based
  - Find edges in the image and use them as region boundary
- Motion-based
  - Group pixels that have consistent motion (e.g., move in the same direction)

Segmentation by Pixel Classification

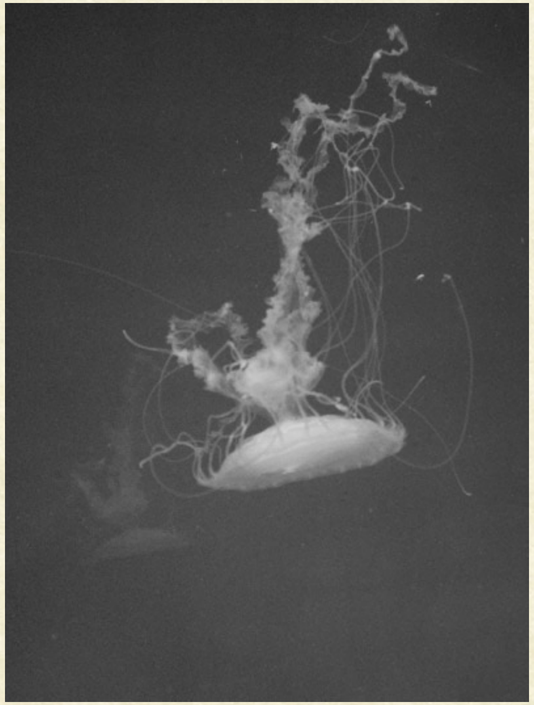
Two Primary Challenges:

1. How to use object / background properties to decide on pixel label?
   - e.g., Ducks are white and yellow, while background is green and brown
2. How to ensure that regions are continuous regions?
   - Avoid fragmentation of object regions

Thresholding

Decide each pixel to be part of an object or background depending on its gray value

$$t(m,n) = \begin{cases} 1 & \text{if } u(m,n) > T \\ 0 & \text{if } u(m,n) \leq T \end{cases}$$



Original

Thresholded (T=95)

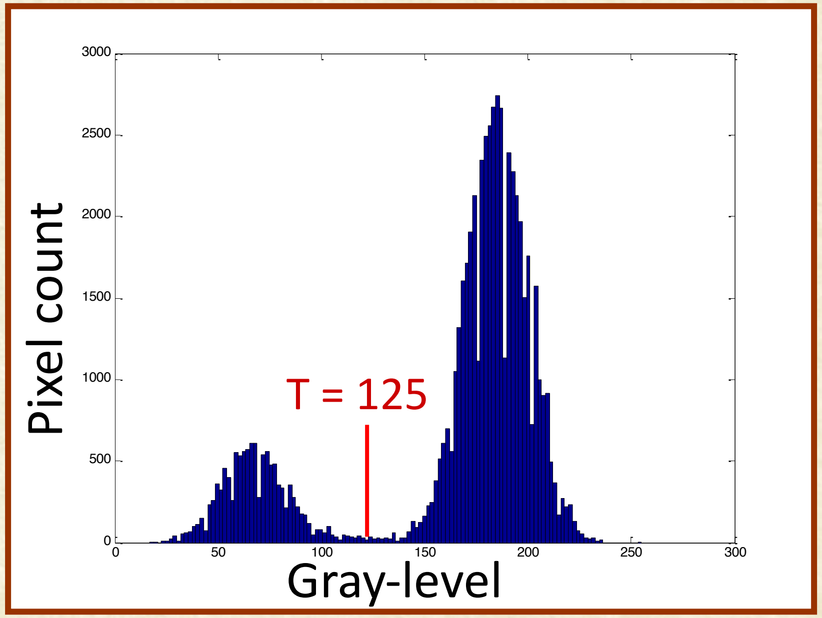
Types of Thresholding

- Global
  - A single threshold is used for the whole image
  - How to determine the threshold?
- Adaptive (Local)
  - Decide the threshold for every pixel depending on its neighborhood
  - How to define the threshold function?

Histogram

- A count of pixels of each graylevel (or range of graylevels) in an image

Grayscale Image



Histogram

Thresholded Image

Original

Thresholded (T=125)

Automatic Thresholding

1. Select an initial estimate of T

2. Segment the image using T. Compute the mean gray values of the two regions, μ_1 and μ_2

3. Set the new threshold $T=(\mu_1+\mu_2)/2$

4. Repeat 2 and 3 until T stabilizes

Assumptions: normal distribution, low noise

Extensions

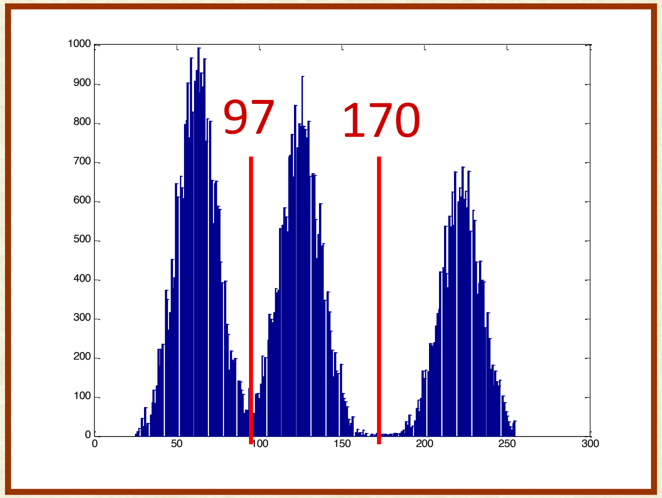
- Multiple Thresholds
  - Find multiple peaks and valleys in the gray level histogram

- Multi-spectral Thresholding
  - In color images, one could use different thresholds for each of the color channels

One might set all the background pixels to black, while leave the foreground at the original value so that the information is not lost.

Multiple Thresholds

Original

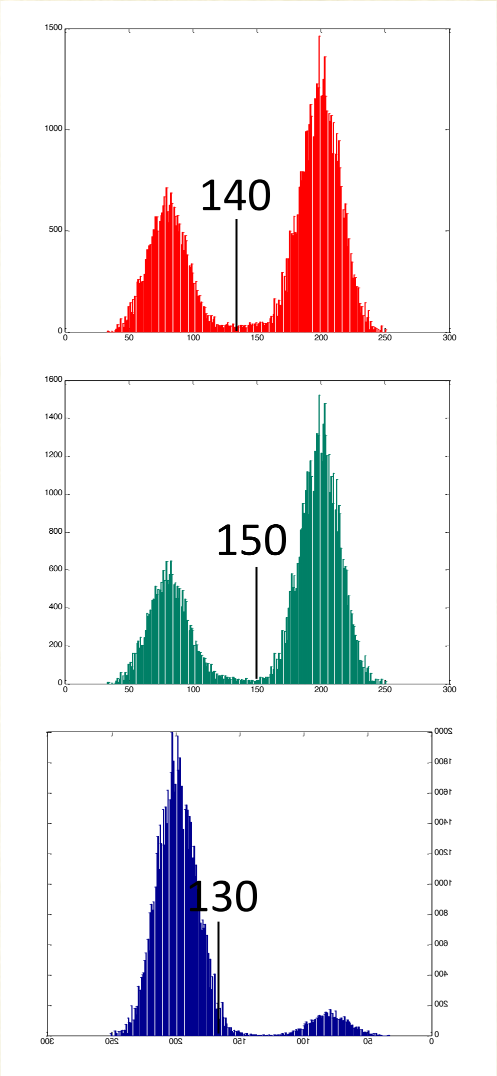


Histogram

Thresholded

Multi-spectral Thresholding

Original

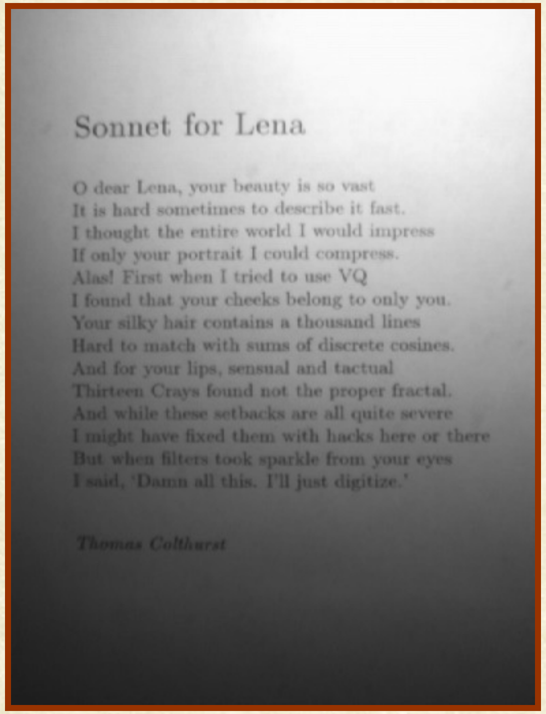


Histograms

Thresholded

Adaptive Thresholding

- Adaptive thresholding changes the threshold dynamically over the image. This can accommodate strong illumination gradients and shadows

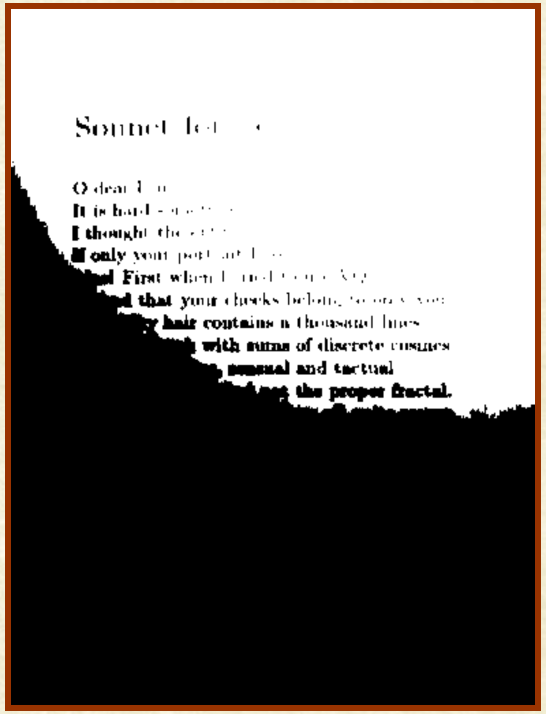


Sonnet for Lena

O dear Lena, your beauty is so vast
It is hard sometimes to describe it fast.
I thought the entire world I would impress
If only your portrait I could compress.
Alas! First when I tried to use VQ
I found that your cheeks belong to only you.
Your silky hair contains a thousand lines
Hard to match with sums of discrete cosines.
And for your lips, sensual and tactual
Thirteen Crays found not the proper fractal.
And while these setbacks are all quite severe
I might have fixed them with hacks here or there
But when filters took sparkle from your eyes
I said, 'Damn all this. I'll just digitize.'

Thomas Colthurst

Original

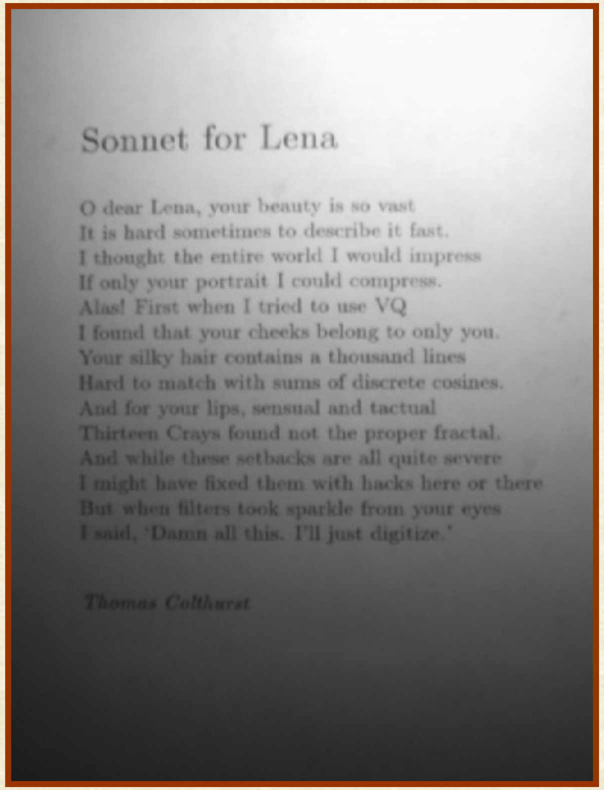


Single Threshold

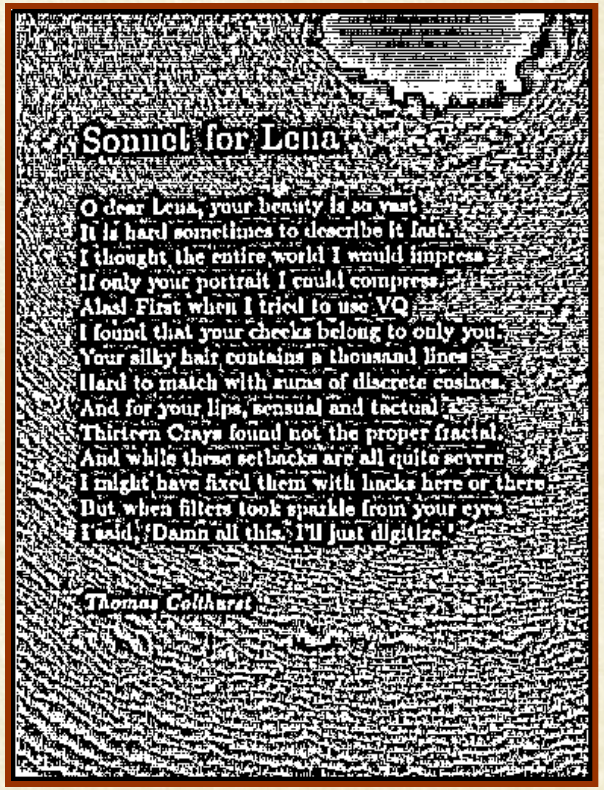
http://www.dai.ed.ac.uk/HIPR2

Adaptive Thresholding

- Set the threshold as mean of pixels (gray values) in a neighborhood (say 7x7)



Original



Adaptive Threshold

http://www.dai.ed.ac.uk/HIPR2

Adaptive Thresholding

- Thresholding using Mean-C

  Set cxc image regions of uniform graylevel to background

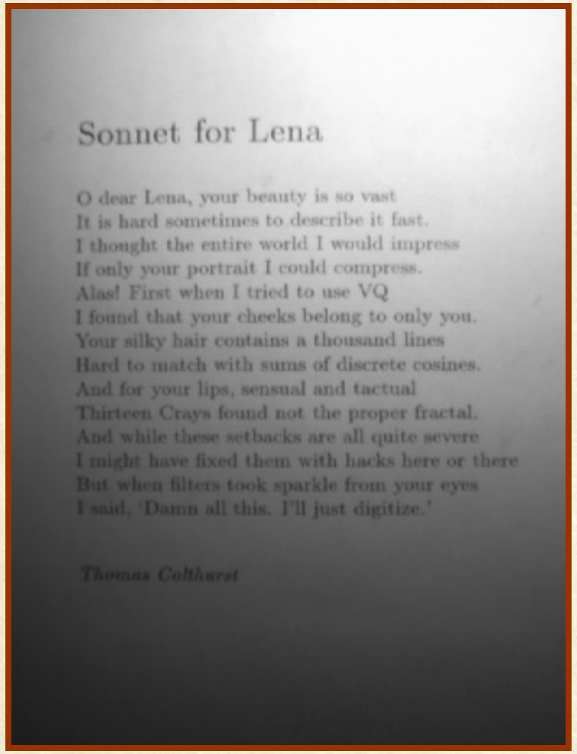
- Chow and Kaneko

1. Apply the mean operator (low pass filter)
2. Subtract original image from the “mean”mage
3. Threshold image in step 2
4. Invert the result

C.K. Chow and T. Kaneko Automatic Boundary Detection of the Left Ventricle from Cineangiograms, Comp. Biomed. Res.(5), 1972, pp. 388-410.

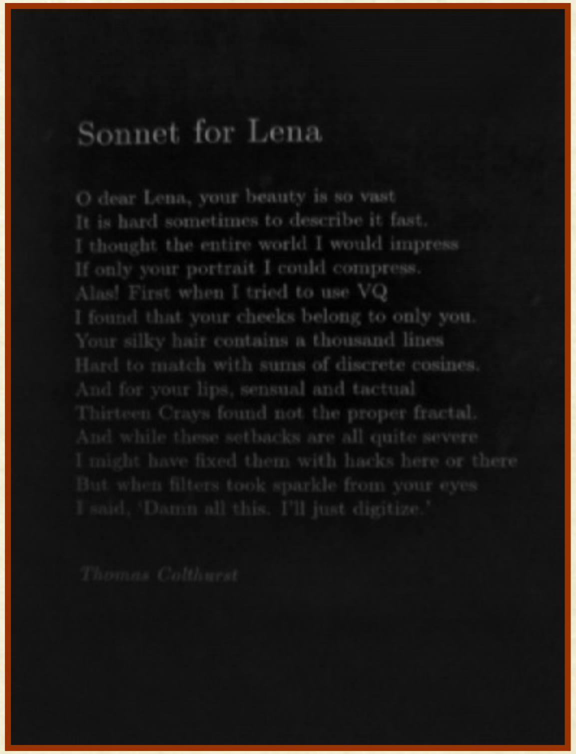
Adaptive Thresholding

<u>Chow & Kaneko Thresholding:</u>



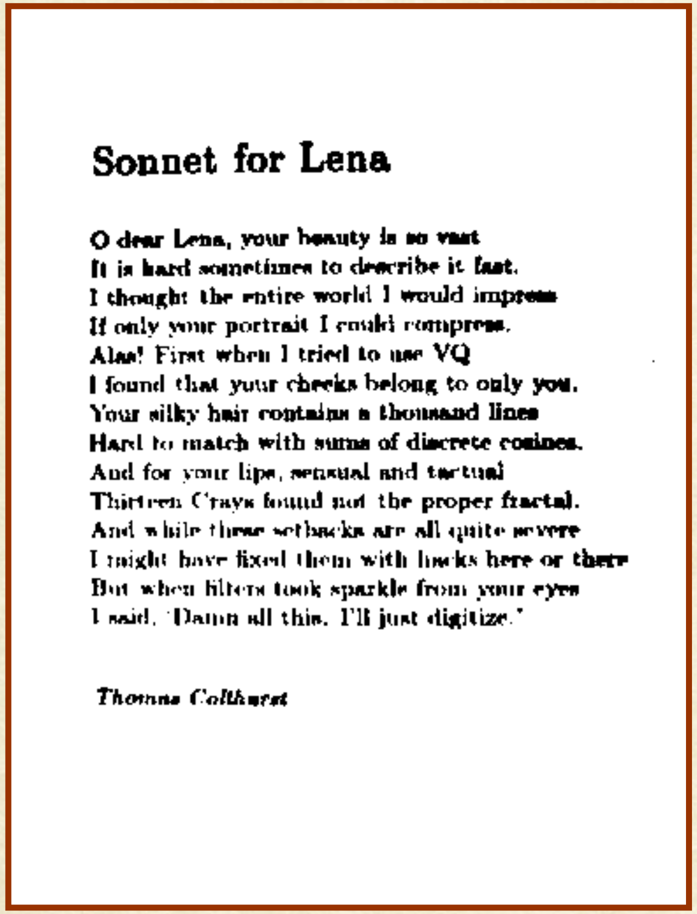
Original

Low-pass filtered



Difference

Adaptive Thresholding Results

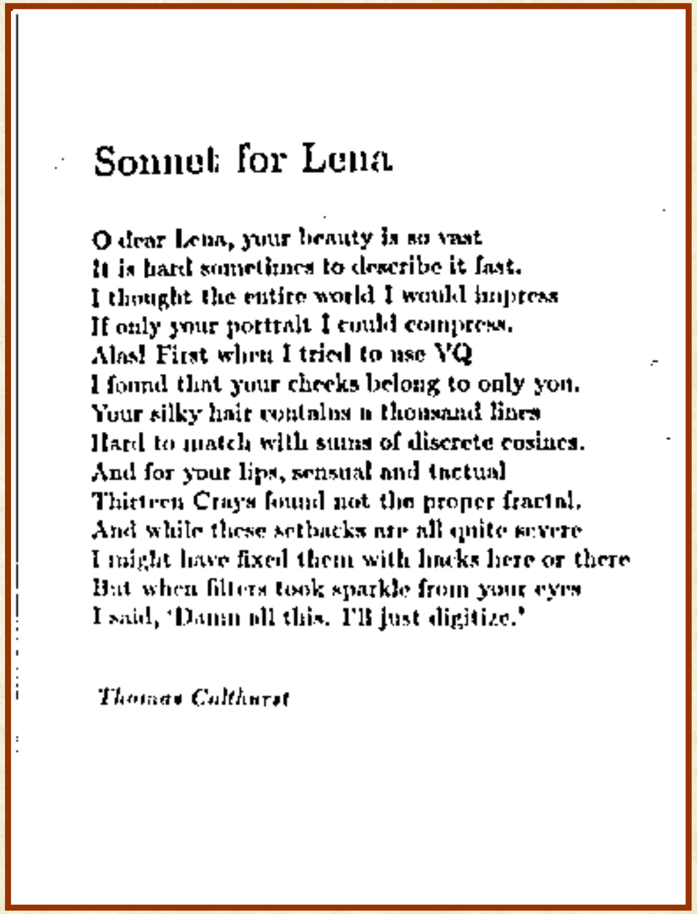


Sonnet for Lena

O dear Lena, your beauty is so vast
It is hard sometimes to describe it fast.
I thought the entire world I would impress
If only your portrait I could compress.
Alas! First when I tried to use VQ
I found that your cheeks belong to only you.
Your silky hair contains a thousand lines
Hard to match with sums of discrete cosines.
And for your lips, sensual and tactual
Thirteen Crays found not the proper fractal.
And while these setbacks are all quite severe
I might have fixed them with hacks here or there
But when filters took sparkle from your eyes
I said, 'Damn all this. I'll just digitize.'

Thomas Colthurst

Chow & Kaneko Thresholding



Sonnet for Lena

O dear Lena, your beauty is so vast
It is hard sometimes to describe it fast.
I thought the entire world I would impress
If only your portrait I could compress.
Alas! First when I tried to use VQ
I found that your cheeks belong to only you.
Your silky hair contains a thousand lines
Hard to match with sums of discrete cosines.
And for your lips, sensual and tactual
Thirteen Crays found not the proper fractal.
And while these setbacks are all quite severe
I might have fixed them with hacks here or there
But when filters took sparkle from your eyes
I said, 'Damn all this. I'll just digitize.'

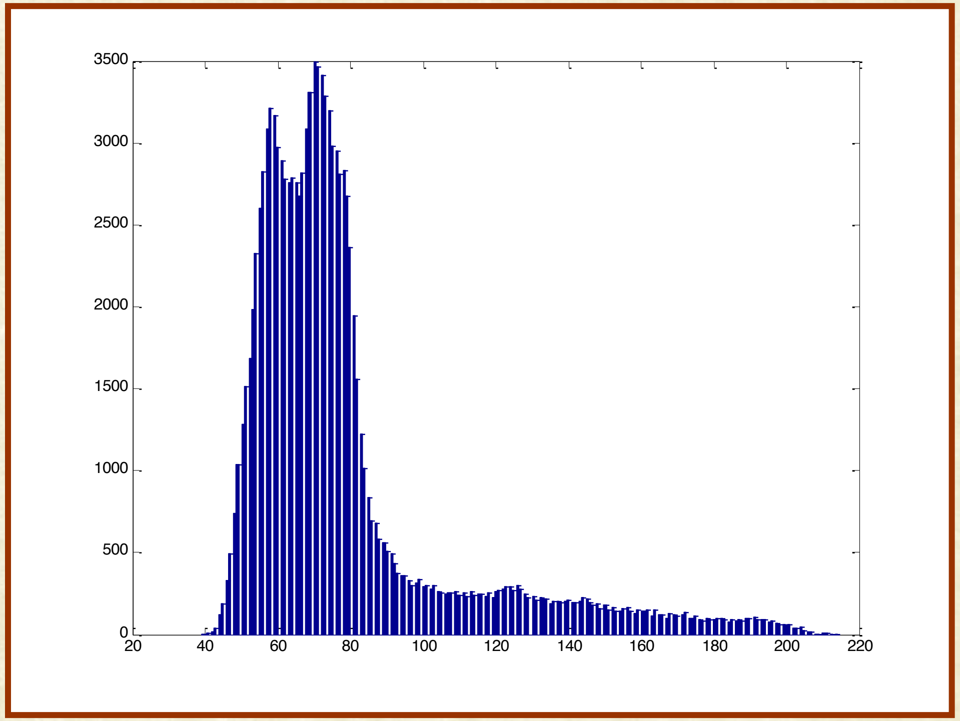
Thomas Colthurst

Mean-C (10) Thresholding

Optimal Thresholding

- The graylevel histogram is approximated using a mixture of two gaussian distributions and set the threshold to minimize the segmentation error

Grayscale Image



Histogram

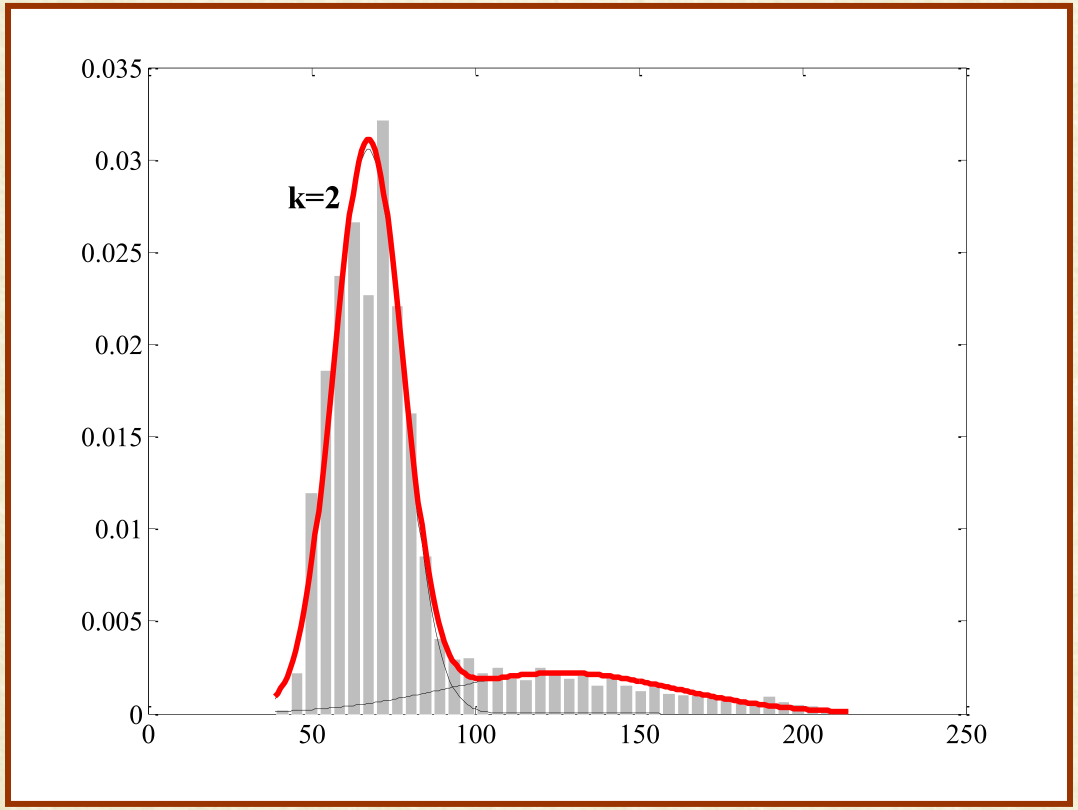
Gaussian Mixture Estimation by EM

- Obj: $N(\mu_1, \sigma_1) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{(x-\mu_1)^2}{2{\sigma_1}^2}}$

- Bkg: $N(\mu_2, \sigma_2) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{(x-\mu_2)^2}{2{\sigma_2}^2}}$

- Initialize μ_1, σ_1, μ_2, and σ_2.
- E-Step: Computed the expected pixel label assignments.
  This could be either hard or soft assignment.
- M-Step: Computed Maximum-(Log)Likelihood estimates of the parameters: $\mu_1, \sigma_1, \mu_2, \sigma_2$
- Repeat the E and M steps until convergence

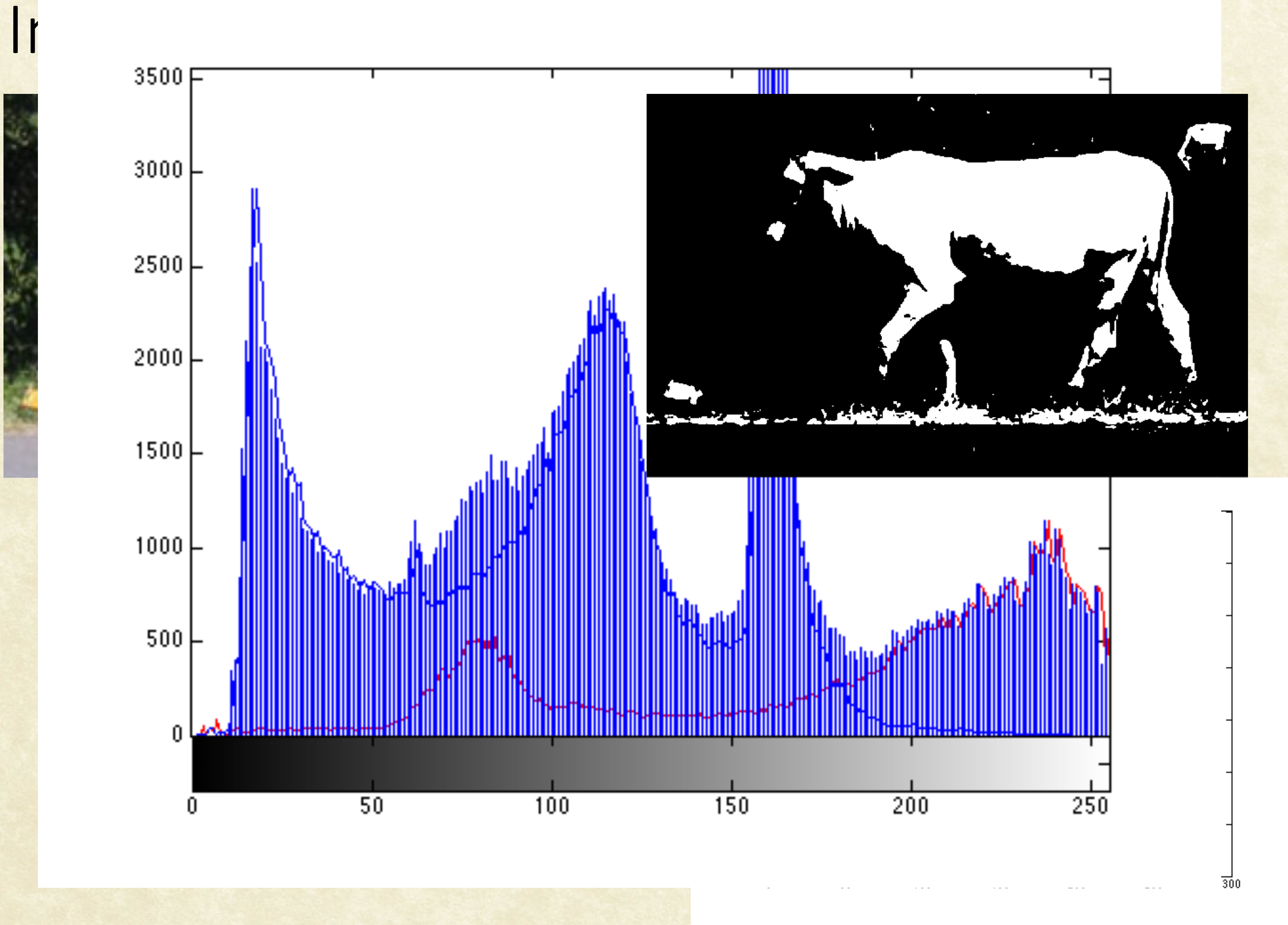
Optimal Thresholding



Histogram with bimodal fit

Thresholded (T=94)

Is In



Segmentation by Pixel Classification

Two Primary Challenges:

1. How to use object / background properties to decide on pixel label?
   - e.g., Ducks are white and yellow, while background is green and brown
2. How to ensure that regions are continuous regions?
   - Avoid fragmentation of object regions