Image Segmentation

Joseph Huber
Phil Vaccaro

THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

April 10, 2018

Questions

1. What is the main conceptual goal in performing image segmentation?

1. How does one method of edge detection work?

1. What does the Hough Transform do?

Joseph Huber

- Started Undergrad in Spring 2014 in Material Science
- Switched to Computer Engineering 5yr BS/MS program
- I graduate from Undergrad this semester
- Interned at ORNL in 2016 and part-time in 2017
- Interned at Siemens Medical Imaging from 2017 to 2018

- Academic Interests
  - High Performance Computing and Optimization
  - FPGAs
  - Digital Image Processing
  - Functional Programming

Joseph Huber

- Born in Knoxville Tennessee
  - 5th Generation Tennessean
- I have an identical twin
- Other Interests
  - Singing lessons at UTK
  - Cycling
  - Bowling

Philip Vaccaro

- Graduated with B.S. in Computer Science from UTK in 2014
- Interned at Siemens Medical Imaging in Knoxville during undergrad
- After graduation worked for a year as a Software Engineer at Siemens Corporate Research in Princeton, NJ
- Began M.S. in Computer Science at UTK in 2015
- Worked as a Co-op student at Sandia National Labs during Spring and Summer 2016 semesters
- Defended thesis last semester and became a concurrent M.S./PhD student

Philip Vaccaro

- Born in Nashville, Tennessee
- Played baseball as a pitcher through high school
    - Won Little League State Championship for Tennessee and played in Regionals
- Moved from Tennessee → New Jersey → Tennessee → New Mexico → Tennessee, all since 2014
- I live with my girlfriend Chloe by Lake Loudon with our two cats (Ivy and Toula), two dogs (Macy and Milo), and gecko (Echo)

Outline

- Brief History
- Overview of Image Segmentation
- Overview of Digital Image Processing
  - Color spaces
  - Convolution
- Overview of some Image Segmentation Techniques
  - Thresholding
  - Edge Detection
  - Clustering
- Segmentation by Pixel coloring
- Results & Conclusion
- References

History of Digital Image Processing

- Bartlane cable picture transmission system
  - Developed by Harry G. Bartholomew and Maynard D. McFarlane in 1920
  - Sent digitized newspaper images from New York to London using telegraph cable
  - Images encoded as 5, later 15, levels of gray on a punch card and reassembled on the other side
- Russell A. Kirsch developed and scanned the first digital image in 1957
  - Identified intensities and stored them in a computer
- Improving computer technology allowed for the storage and processing of digital images

Image Processing Pipeline Overview

1. Assemble and/or reconstruct raw imaging data
2. Correct or transform image data to match application requirements
3. Understand and/or classify images via techniques like image segmentation

Image Segmentation: Overview

- Separates images into foreground and background components
- Goal is to assign meaning to components through recognition
- Must have strong techniques for identifying these image segments

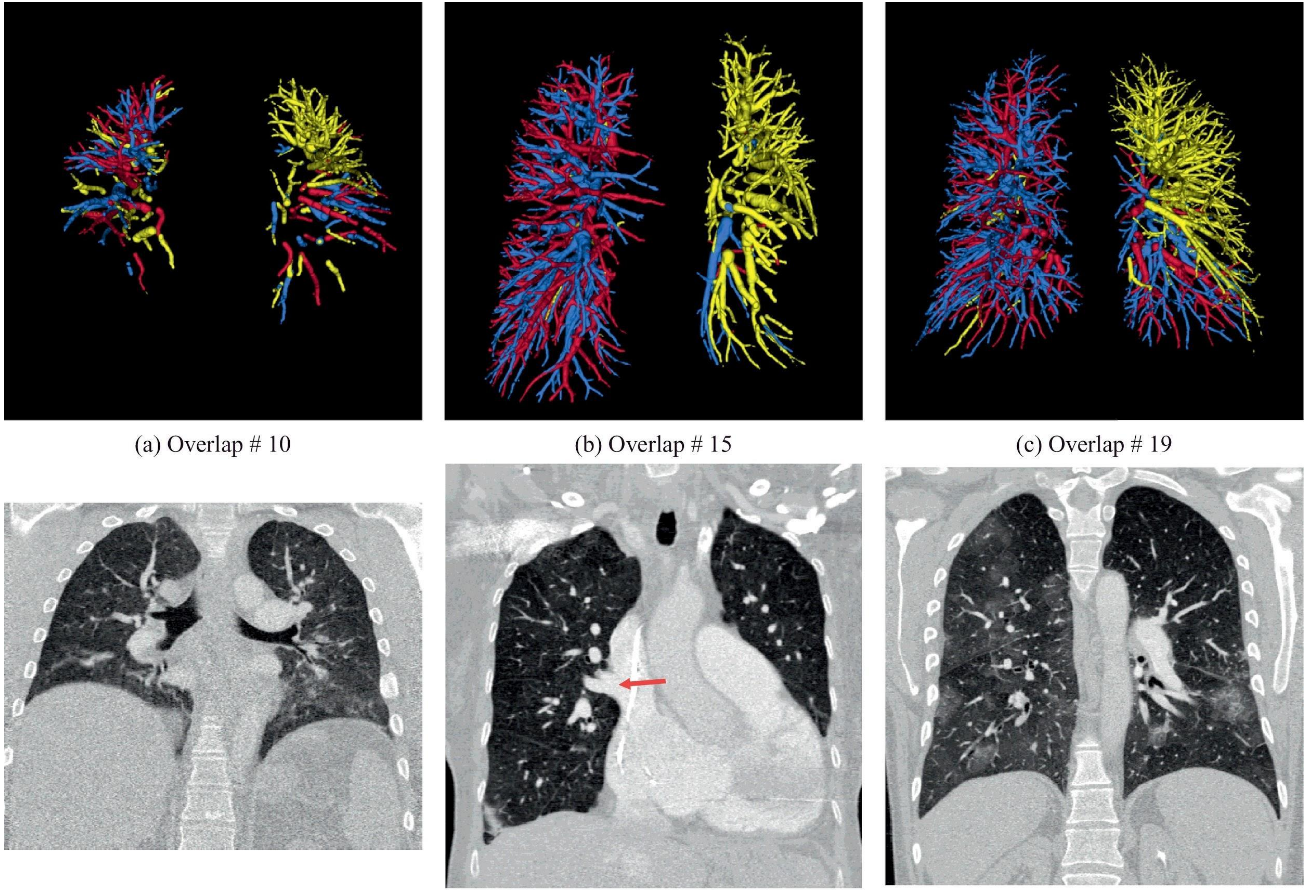
Image Segmentation Techniques

- Thresholding gray level intensities
- Region Growing
- Segmenting based on surfaces
- Edge detection
- Clustering Techniques
- Linear Programming approaches

From: A REVIEW ON IMAGE SEGMENTATION TECHNIQUES, NIKHIL R. PAL and SANKAR K. PAL

Image Segmentation: Applications

- Medical images can be segmented to reveal obscure anatomical structure
  - CT, MRI, Compounded Ultrasound
- Segmented images can be used in pursuit of less invasive procedures
  - Multi-modal registration techniques
- Facial recognition and object detection

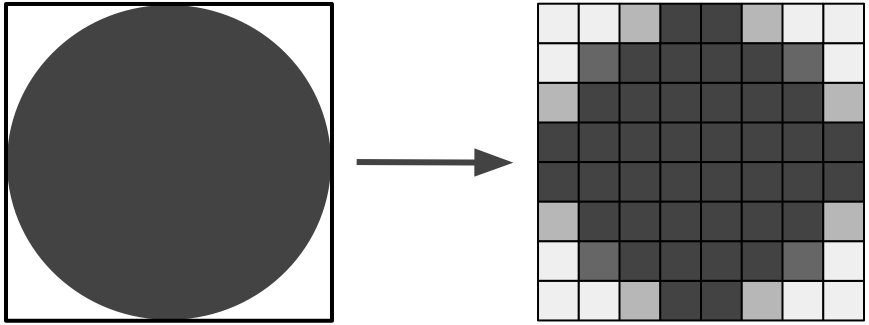


(a) Overlap # 10 (b) Overlap # 15 (c) Overlap # 19

(d) Slice # 10 (e) Slice # 15 (f) Slice # 19

Digital Image Processing

- An Image is defined as a two-dimensional function of intensity f(s, t), with spatial coordinates s and t
- A digital image is an image such that x, y, and f[x, y] are *finite* and *discrete*

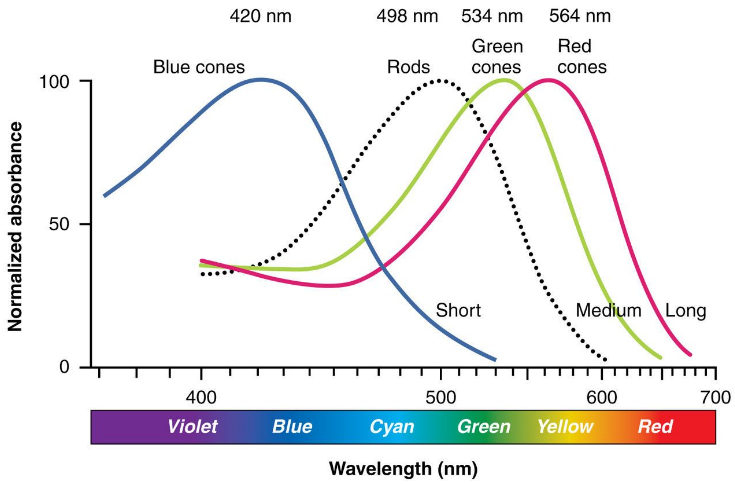


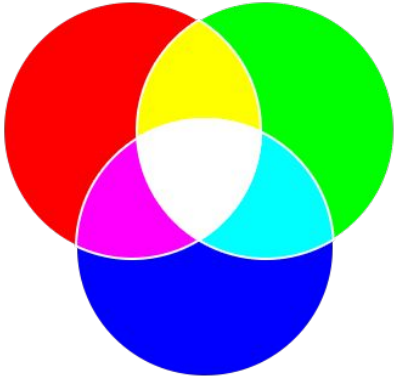
Digital Image Processing

- The goal of digital image processing is to enhance or extract useful information from a digital image.

Color Spaces

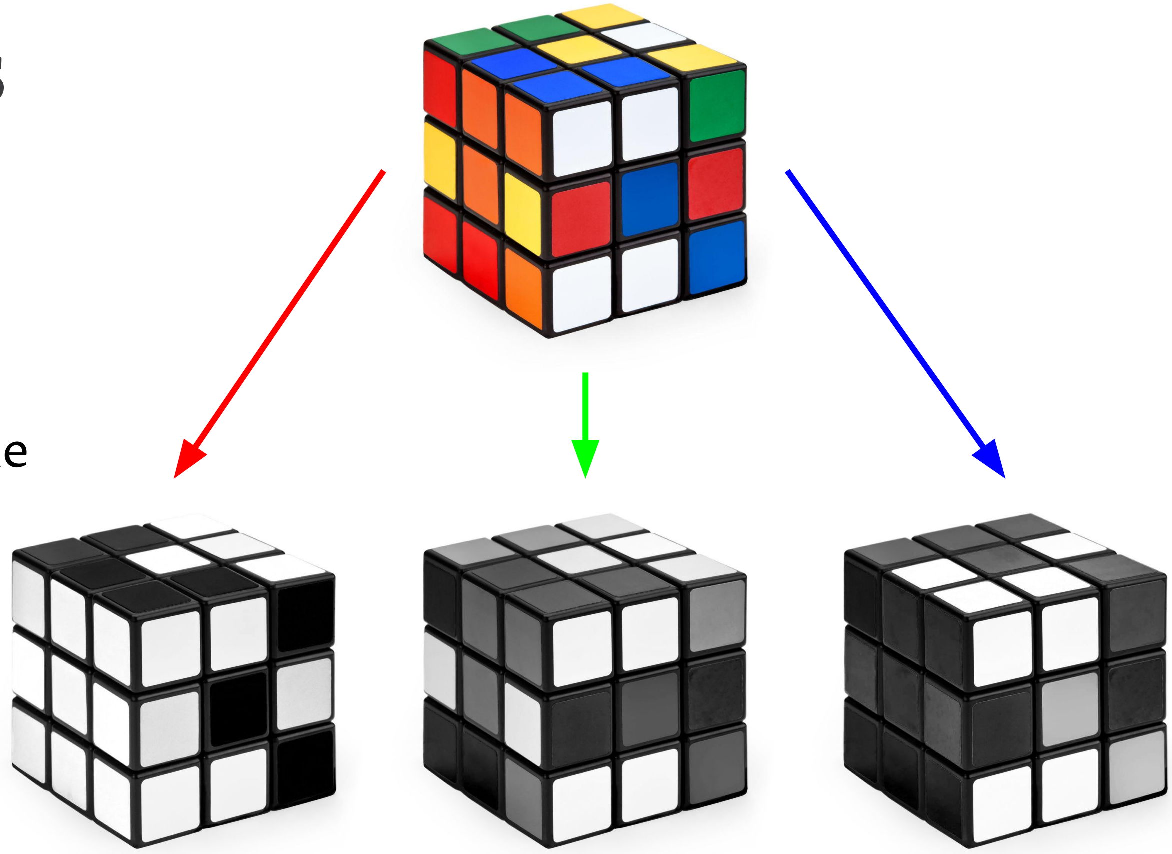
- Color spaces describe a mathematical model to reproduce colors
- Developed around the physics of human vision
- The span of the color space is called the gamut
- Many different color work well or poorly for certain applications
- Colors are typically normalized in [0,1] but quantized before display





Color Spaces

- Color images are represented by components inside a color space
- Certain applications like photography require color close to human vision while others like medical don't
- Examples
- sRGB, Lab, HSL, Gray



Convolution

- many operations can be expressed as a convolution between an image f[x,y] and a kernel k[x,y] to produce an image g[x,y]
- Convolution shifts one function through another function, returning the *area* under the product at each point

$$g[x,y] = f[x,y] \circledast k[x,y] = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} k[m,n] f[x-m, y-n]$$

- The infinite sum can be simplified if f[x,y] or k[m,n] are known to be zero at certain locations
- Kernels we use are flipped before performing the convolution

Kernels

- Kernels describe operating on adjacent pixels and adding the results together
- Let k[x,y] be a kernel with values between -1 and 1 and zero elsewhere

$$g[x,y] = \sum_{n=-1}^{1} \sum_{m=-1}^{1} k[m,n] f[x-m, y-n]$$

$$\begin{array}{cccc} & -1 & 0 & 1 \\ -1 & a & b & c \\ 0 & d & e & f \\ 1 & g & h & i \end{array} \circledast \begin{array}{cccc} & 0 & 1 & 2 \\ 0 & 1 & 2 & 3 \\ 1 & 4 & 5 & 6 \\ 2 & 7 & 8 & 9 \end{array}$$

- Expansion of convolution around the center
- It looks backwards, so the kernels we use are flipped prior to convolution

$$g[1,1] = 9a + 8b + 7c + 6d + 5e + 4f + 3g + 2f + \mathrm{i}$$

Image Segmentation

- Subdivide a picture into its constituent regions or objects
- Divide the region R into n subregions with the following properties
    - All the subregions add up to the original full image
    - Each subregion is completely distinct from others
    - Each subregion is connected in some way
    - There is some predicate that is true for each subregion
    - That same predicate is false for adjacent subregions

Global Threshold

- Segment image into different levels of intensity using a threshold intensity T

$$g(x, y) = \begin{cases} 1, \text{if } f(x, y) > T \\ 0, \text{if } f(x, y) \leq T \end{cases}$$

- The value of T can either be specified or calculated to optimally segment the image
- Can be split into multiple threshold levels called *multiple threshold*

Otsu's Method

- Developed by Nobuyuki Otsu
- Find an ideal threshold value T in a bimodal image
- Compute a histogram of the image
- The optimal value will be the threshold that maximizes the variance between the pixels to the left and right of the threshold value

$$\text{Maximize } \sigma_b^2(T) = \omega_1(T)\omega_0(T)[\mu_0(T) - \mu_1(T)]^2$$

- ω
  - Number of pixels above and below threshold T in the histogram
- μ
  - The average intensity of the pixels above and below threshold

Otsu's Method

the reader whose ear has long been accustomed to the haunting melodies of the original. But that Mallarmé's translations are read and enjoyed in France is in itself a testimony to the innate beauty, the residual charm, of Poe's poetic structures, when bereft of those formal elements in which their beauty and charm have hitherto been thought so largely to consist. It is hard to think of *The Raven* or of *Ulalume* without those interreticulations of sound and form, those reciprocities of repetition and parallelism, which in Poe's hands fused them into artistic unity; but in Mallarmé's versions, however exquisite the prose, it is prose still. That Poe has stood the test is a noteworthy tribute to the intrinsic worth and fundamental texture of his poetic material. One unexpected result of Mallarmé's work has been to put Poe, in the eyes of Frenchmen at least, side by side with Whitman in the ranks of the *vers librists.* Strange bedfellows, these! "Yet it is true," says Caroline Ticknor,[12] "that Mallarmé's translations of Poe set the pace for the new school from which the exponents of *vers libre* assuredly derive their inspiration."

Of the many French biographies of Poe the most elaborate is that by Lauvrière.[13] But it is a study in pathology. Scholarly, painstaking, accurate, and even sympathetic in its statement of facts, its inferences do not carry conviction. *Morbidité, aliénisme, dégénérescence, décadence*—these do not belong to Poe. They can be read into his life and genius only by a studied selection of incidents and an equally studied

[12] *Poe's Helen* (1916), New York, p. 276.
[13] *Edgar Poe, sa vie et son oeuvre,* Paris, 1904. The book contains 732 pages.

18 EDGAR ALLAN POE

the reader whose ear has long been accustomed to the haunting melodies of the original. But that Mallarmé's translations are read and enjoyed in France is in itself a testimony to the innate beauty, the residual charm, of Poe's poetic structures, when bereft of those formal elements in which their beauty and charm have hitherto been thought so largely to consist. It is hard to think of *The Raven* or of *Ulalume* without those interreticulations of sound and form, those reciprocities of repetition and parallelism, which in Poe's hands fused them into artistic unity; but in Mallarmé's versions, however exquisite the prose, it is prose still. That Poe has stood the test is a noteworthy tribute to the intrinsic worth and fundamental texture of his poetic material. One unexpected result of Mallarmé's work has been to put Poe, in the eyes of Frenchmen at least, side by side with Whitman in the ranks of the *vers librists.* Strange bedfellows, these! "Yet it is true," says Caroline Ticknor,[12] "that Mallarmé's translations of Poe set the pace for the new school from which the exponents of *vers libre* assuredly derive their inspiration."

Of the many French biographies of Poe the most elaborate is that by Lauvrière.[13] But it is a study in pathology. Scholarly, painstaking, accurate, and even sympathetic in its statement of facts, its inferences do not carry conviction. *Morbidité, aliénisme, dégénérescence, décadence*—these do not belong to Poe. They can be read into his life and genius only by a studied selection of incidents and an equally studied

[12] *Poe's Helen* (1916), New York, p. 276.
[13] *Edgar Poe, sa vie et son oeuvre,* Paris, 1904. The book contains 732 pages.

Digitized by Microsoft®

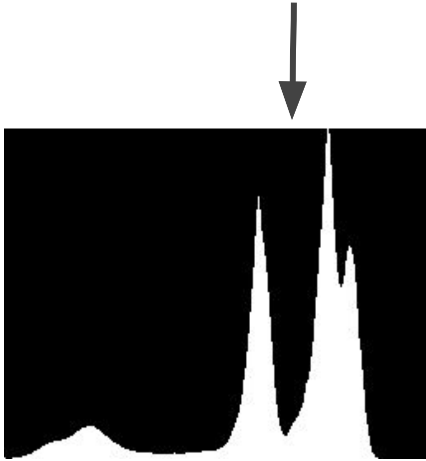
18 EDGAR ALLAN POE

the reader whose ear has long been accustomed to the haunting melodies of the original. But that Mallarmé's translations are read and enjoyed in France is in itself a testimony to the innate beauty, the residual charm, of Poe's poetic structures, when bereft of those formal elements in which their beauty and charm have hitherto been thought so largely to consist. It is hard to think of *The Raven* or of *Ulalume* without those interreticulations of sound and form, those reciprocities of repetition and parallelism, which in Poe's hands fused them into artistic unity; but in Mallarmé's versions, however exquisite the prose, it is prose still. That Poe has stood the test is a noteworthy tribute to the intrinsic worth and fundamental texture of his poetic material. One unexpected result of Mallarmé's work has been to put Poe, in the eyes of Frenchmen at least, side by side with Whitman in the ranks of the *vers librists.* Strange bedfellows, these! "Yet it is true," says Caroline Ticknor,[12] "that Mallarmé's translations of Poe set the pace for the new school from which the exponents of *vers libre* assuredly derive their inspiration."

Of the many French biographies of Poe the most elaborate is that by Lauvrière.[13] But it is a study in pathology. Scholarly, painstaking, accurate, and even sympathetic in its statement of facts, its inferences do not carry conviction. *Morbidité, aliénisme, dégénérescence, décadence*—these do not belong to Poe. They can be read into his life and genius only by a studied selection of incidents and an equally studied

[12] *Poe's Helen* (1916), New York, p. 276.
[13] *Edgar Poe, sa vie et son oeuvre,* Paris, 1904. The book contains 732 pages.

Otsu's Method Effectively segments images with two strong peaks in their histogram

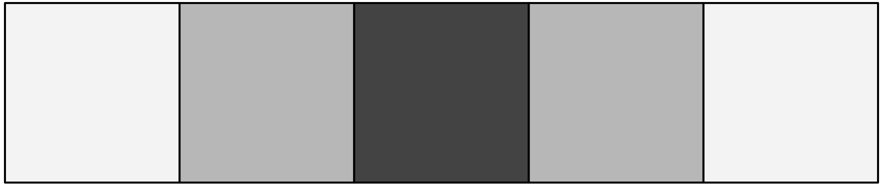


Variable Threshold in Brief

- Global Threshold works poorly with images containing a strong color gradient
- Variable Thresholding allows T to vary over the image
  - Local Thresholding changes the value of T according to its neighbors
  - Adaptive Thresholding changes the value of T as a function of x and y in f[x,y]

Edge Detection

- Segment the image by identifying areas with a large change in intensity
- Link the similar edges together to create closed shapes



- One edge detection method finds changes in intensity by taking derivatives of the image

Edge Detection

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

- The image is discrete so the closest we can get is 1

$$\frac{\partial f}{\partial x} = f[x+1] - f[x]$$

- The second derivative of an image is the Laplacian

$$\nabla^2 f[x,y] = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Edge Detection

- Turn difference equations into kernels

$$\frac{\partial f}{\partial x} = \begin{array}{|c|c|}\hline -1 & 1 \\ \hline\end{array} \qquad \frac{\partial f}{\partial y} = \begin{array}{|c|}\hline -1 \\ \hline 1 \\ \hline\end{array} \qquad \nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- Laplacian kernel

0	1	0
1	-4	1
0	1	0

or

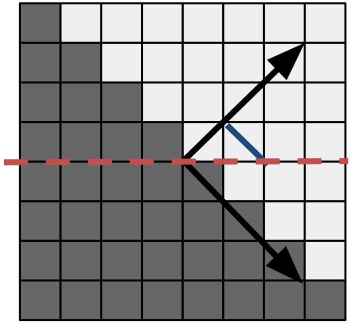
1	1	1
1	-8	1
1	1	1

$$\nabla^2 f[x,y] = f[x+1,y] + f[x,y+1] + f[x-1,y] + f[x,y-1] - 4f[x,y]$$

Edge Detection

- Gradient is the direction and strength of the edges

$$|\nabla f| = \sqrt{g_x^2 + g_y^2} \qquad \angle \nabla f = \tan^{-1} \frac{g_y}{g_x}$$



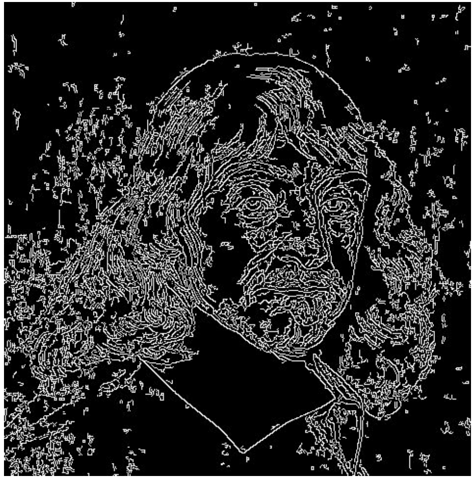
- Gradient vector is normal to the direction of the edge
- Obtain by convolving image with gradient kernel
- Space out to avoid shifting the image if needed

-1	0	1

-1
0
1

Canny Edge Detection in Brief

- Edge detection is very sensitive to noise in the image
- Smooth image by convolution with a Gaussian function
- Thin edges to specified angles and remove weak edges



Edge Linking

- Edge detection does not often provide a smooth, closed set of edges
- Similar edges can be connected using the direction
- Straightforward solution
  - Calculate the magnitude and angle of the gradient of the image
  - Connect points along the line described by the gradient within a certain distance

Hough Transform

- Patented by Paul Hough in 1962 and later improved by Richard Duda and Peter Hart in 1972
- Very good edge and shape detection by finding shapes that create the edges in the image
- Take multiple points (x, y) and find a shape that best intersects the points
- The shapes that hit the most points simultaneously are dominant features in the final image

Hough Transform

- Try to find a line that hits many points on the edge

$$y = mx + b$$

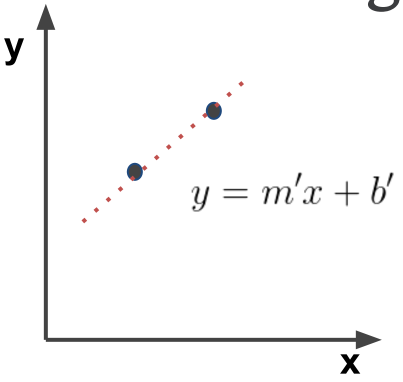
- Infinitely many lines pass through this point
- Express the point in parameter space

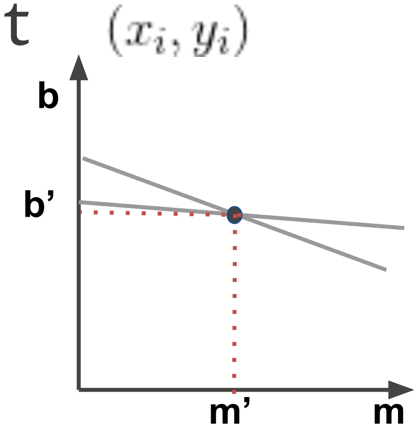
$$b = (-x)m + y$$

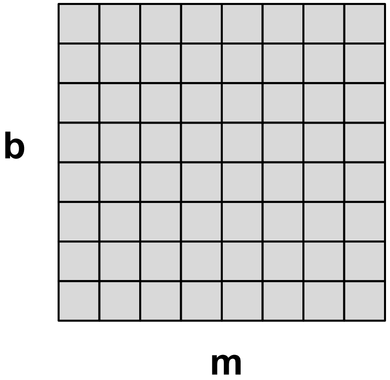
- A point in the image space corresponds to a line in parameter space

Hough Transform

- Vote on discrete points in parameter space for all points, increment that point in an accumulator table
- The points in parameter space with a large vote count are a good fit







Hough Transform

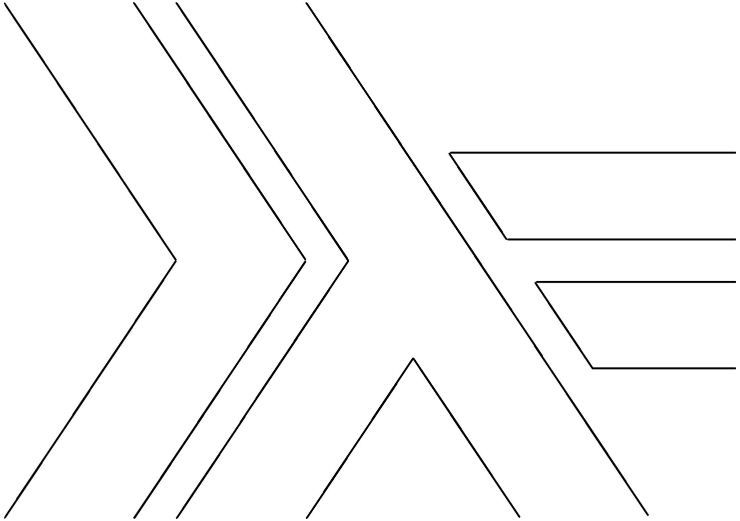
- Line has an infinite slope when the points are connected vertically
- Express the line in polar format

$$\rho = x cos\theta + y sin\theta \qquad y = -\frac{cos\theta}{sin\theta}x + \frac{r}{sin\theta}$$

- Now represents sinusoids in parameter space corresponding to an point obtained from an angle and distance in image space
- Can be applied to circles, parabolas, etc

Hough Transform

- Segments can be constructed through the line equations from the Hough Transform



Clustering

- Segment image by grouping together items with similar characteristics
- Ideally separate the image into several distinct groups
- Many clustering methods
  - Centroids
  - Graphs
- Hard clustering requires exclusive membership
- Soft clustering allows shared membership

K-means Clustering

- Common centroid-based clustering algorithm
- Create K centroids and initialize them to some location
  - Assign each point to its closest centroid using some distance function
    - Typically Euclidean distance
  - Adjust each centroid to the mean of the points that belong to it
  - Repeat until convergence
- NP-Hard problem to converge
  - A few iterations are a good approximation

K-means Clustering

- Clustering on images can be done by intensity or color
- Each data point is a pixel represented by its color space components
- Measure distance using Euclidean distance between the color space components
- Distance in sRGB is not completely accurate

$$D = \sqrt{R^2 + G^2 + B^2} \qquad R, G, B \in [0, 1]$$

- RGB distance is not close to how humans see color distance
- Convert the image to some other color space depending on the application

Image Quantization

- Represent an image using a few colors
- Cluster similar colors together
- Visually similar to segmentation

HSL Quantization (3)

Source Image

sRGB Quantization (8)

Lab Quantization (8)

Watershed Algorithms in Brief

- View the gradient image as a topological surface of slopes in directions
- Watershed visualization
  - Imagine water slowly fills through the bottom of the local minima
  - Whenever the water in distinct basins is about to overflow build a dam between them
  - Continue until only the dams are visible above the water level

Segmentation via Pixel Coloring

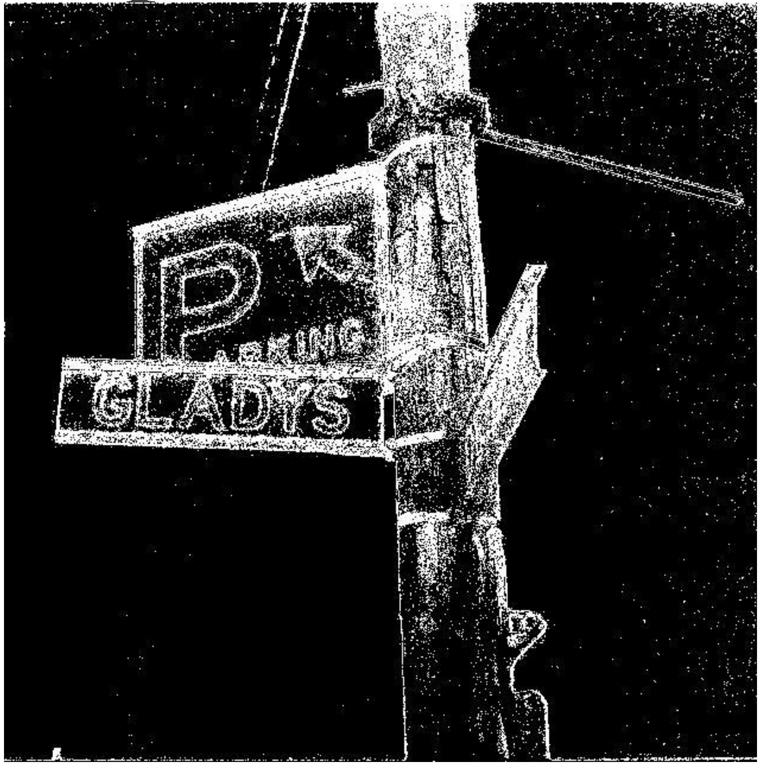
- Simple method for estimating foreground and background of images
- Starts with initial representative pixel
- Compares against all other pixels for uniqueness to decide whether pixel is in the foreground or background

Segmentation Process

- A starting pixel is chosen and is assigned a hash
- All other pixels are iterated through in order to compute their hash values
- Each pixel hash value is compared against all discovered unique pixels
- If a pixel is unique, mark it white, else mark it black

Segmentation Example

Segmentation Example



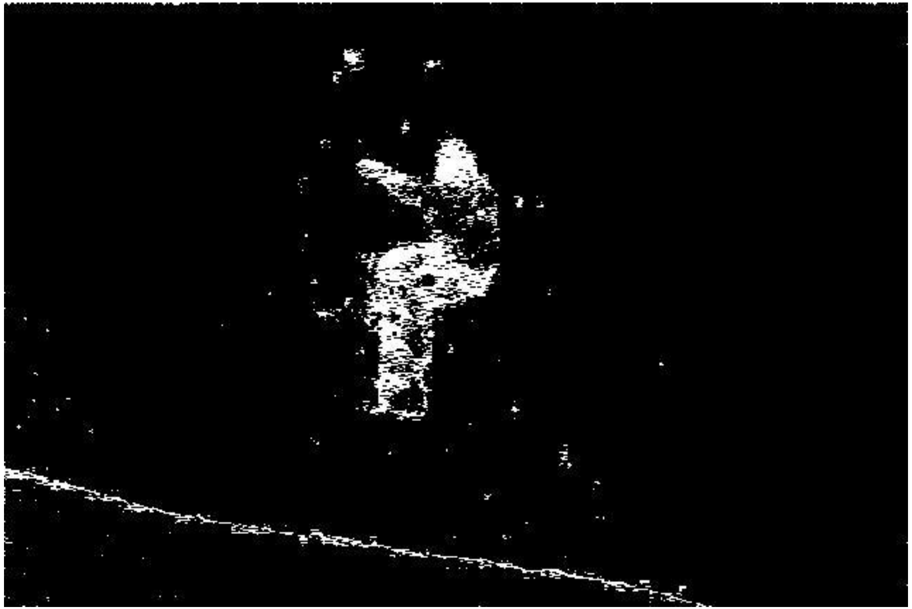
Initial Results

- Classifying pixels by uniqueness initial boundaries
- However, in both examples there is noise present in the resulting image
- Adding a secondary step can be used for a “smoothing” effect

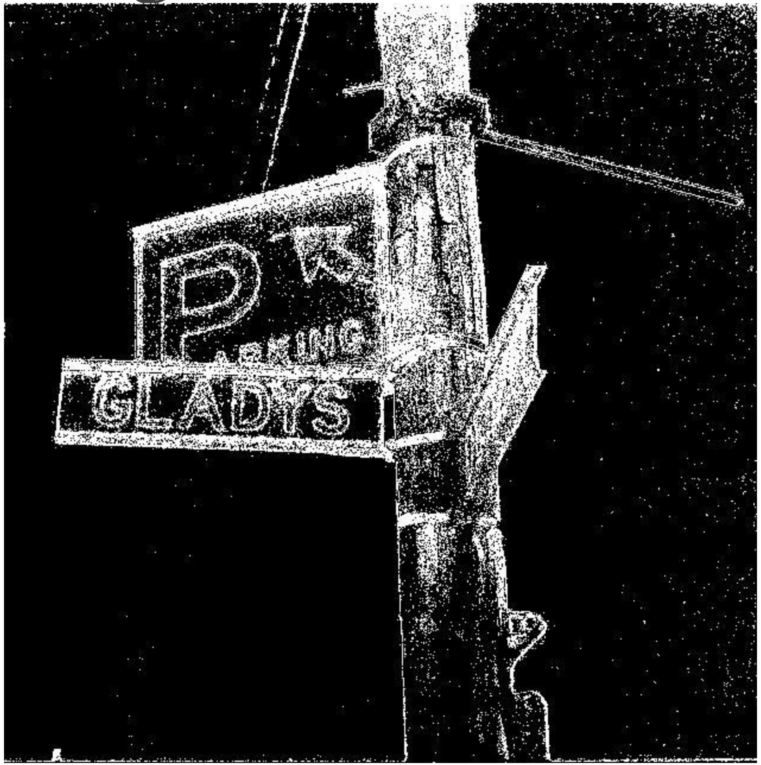
Augmenting With Neighbor Averaging

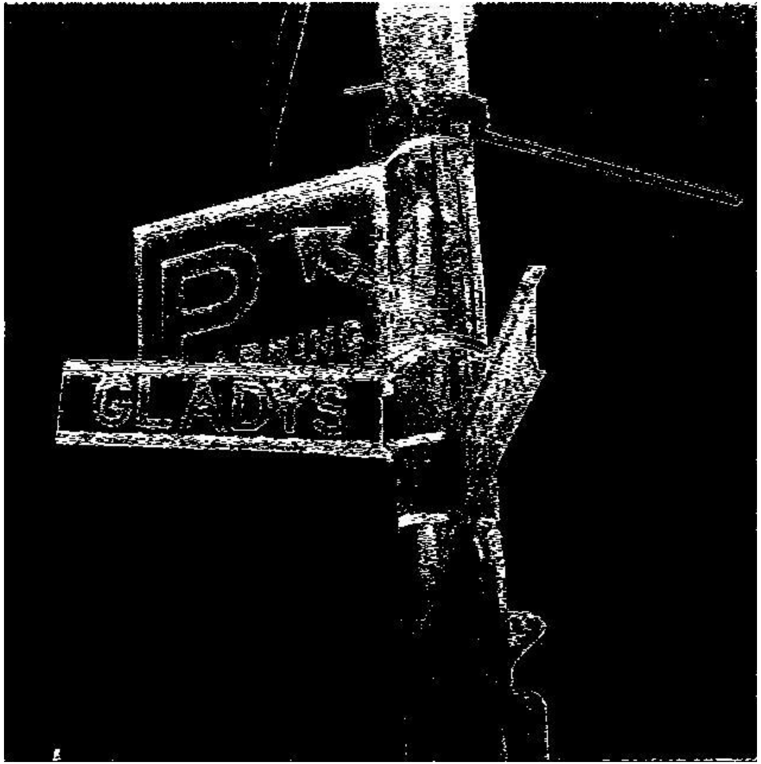
- Another step was applied to get a "smoother" segmented image
- Neighbor pixels of segmentation are averaged to assign each center pixel to be black or white
- Helps to clear up noise

Augmented Example: One Iteration



Augmented Example: One Iteration

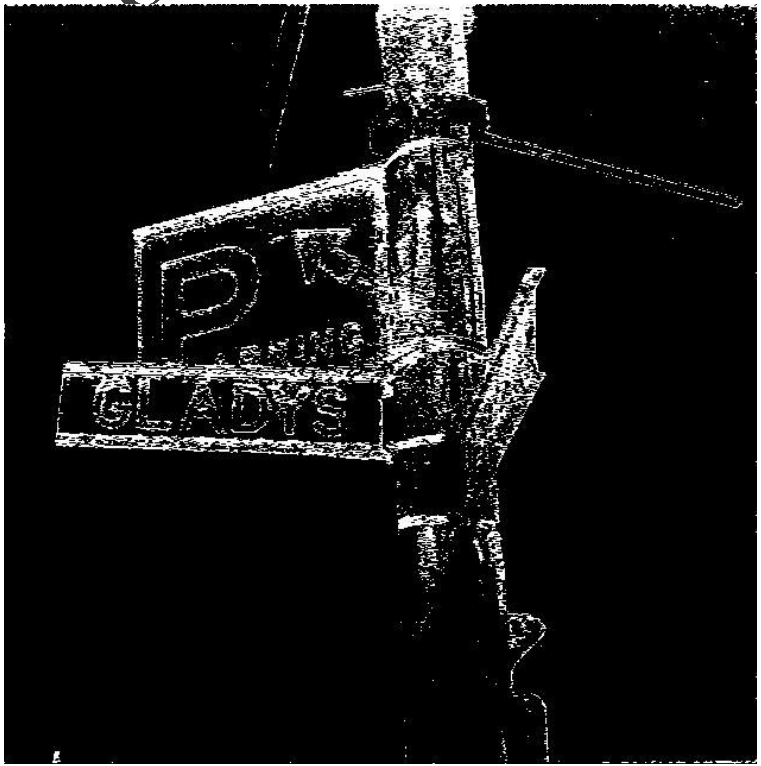


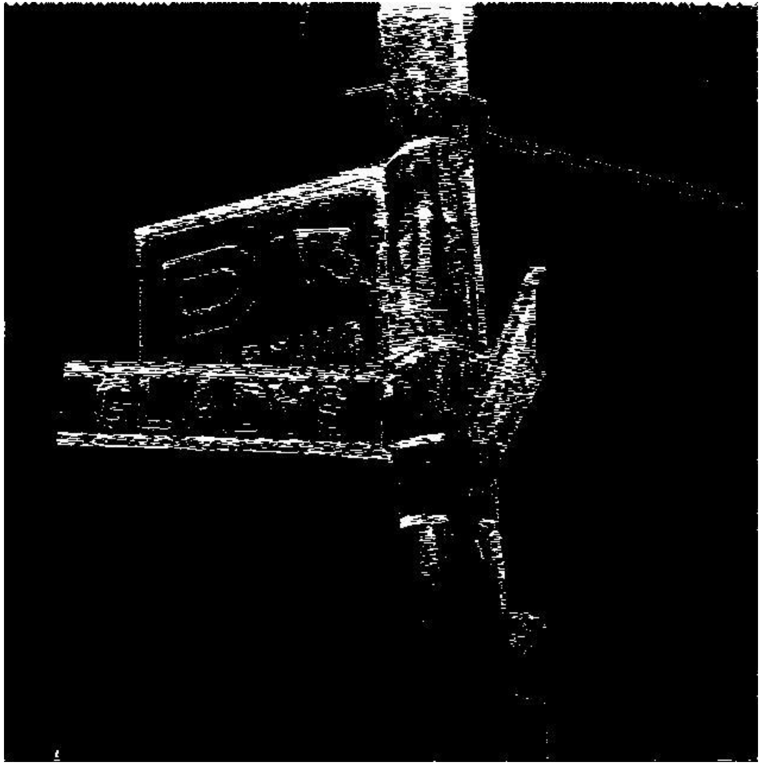


Augmented Results

- By incorporating neighboring pixels, some of the noise has been removed
- Both examples show improvement in terms of edge definition after one iteration
- Provides finer grained boundaries

Augmented Example: x2 Iterations

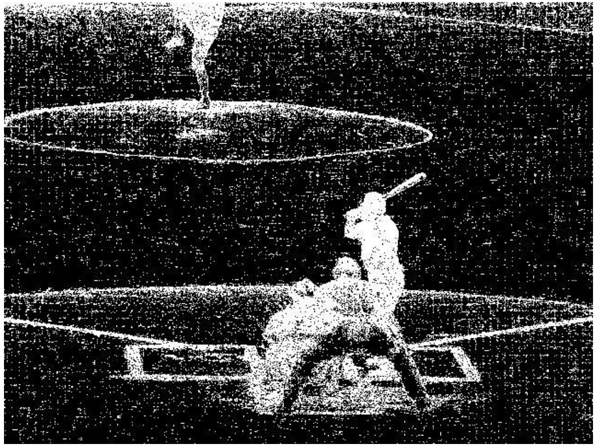


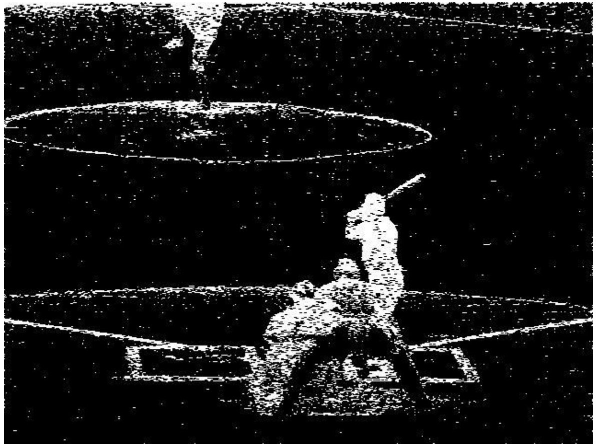


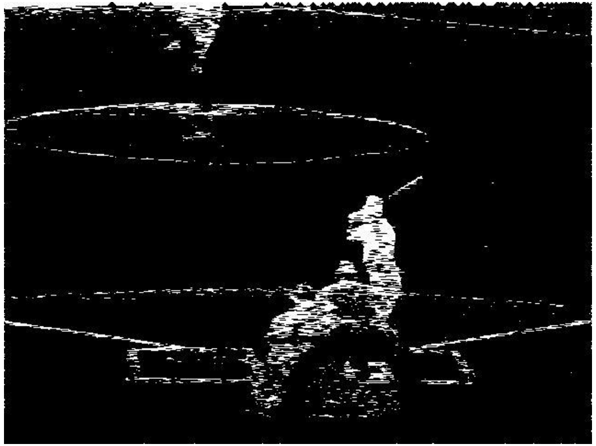
Augmented Results Cont.

- The previous example shows that multiple iterations may not be ideal
- Image quality plays a factor in the ability to adequately assign uniqueness to pixels
- The initial assignment of pixels to white or black is affected by this
- As a result, applying the augmented step involving neighbor pixels seems to remove too much data

x2 Iterations



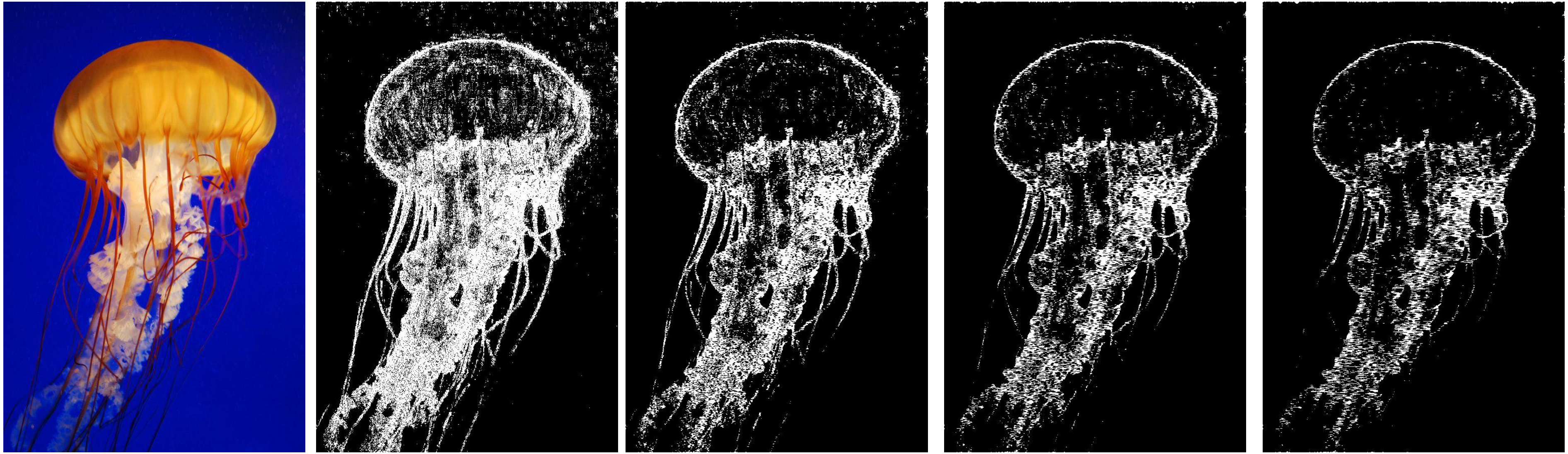




Full Pipeline Results

- Unlike the example with x2 iterations, this x2 iterations example shows a different effect
- The first iteration involving neighbor pixels shows improvement upon the uniqueness-based segmentation
- Furthermore, the second iteration seems to further remove noise from the black areas of the image

x3 Iterations



Full Pipeline Results Cont.

- As with the previous example, this sequence of segmentations shows noise reduction as more iterations are applied
- These results may indicate that image quality and image content play a major role in the effectiveness of these techniques
- Finding the ideal number of iterations is key to removing an acceptable amount of noise

Concluding Remarks

- There are a wide range of image segmentation techniques that have been available for many years
- The appropriate technique is highly dependent upon the application
- Understanding data is key in choosing how to best manipulate the data programmatically

Questions?

Questions

1. What is the main conceptual goal in performing image segmentation?

1. How does one method of edge detection work?

1. What does the Hough Transform do?

References

1. Khan, M. W. (2014). A survey: Image segmentation techniques. International Journal of Future Computer and Communication, 3(2), 89.
2. Kaur, D., & Kaur, Y. (2014). Various image segmentation techniques: a review. International Journal of Computer Science and Mobile Computing, 3(5), 809-814.
3. Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. Pattern recognition, 26(9), 1277-1294.
4. Luptak, A. J. A Fast Method For Simple Image Segmentation.
5. Payer, C., Pienn, M., Bálint, Z., Shekhovtsov, A., Talakic, E., Nagy, E., ... & Urschler, M. (2016). Automated integer programming based separation of arteries and veins from thoracic CT images. Medical image analysis, 34, 109-122.
6. Gonzalez, R. C., & Woods, R. E. (2012). Digital image processing.