Region Segmentation

Computer Vision (CS0029)

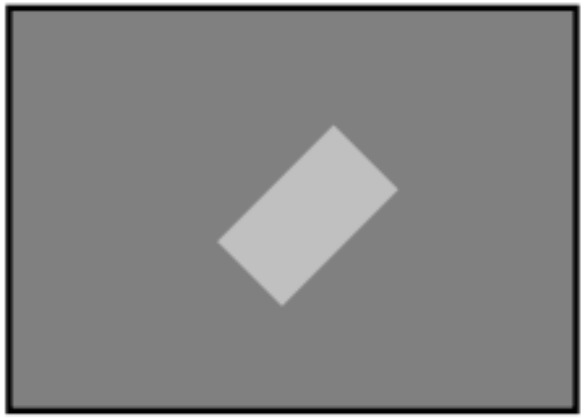
Region Segmentation

- Segment image into set of regions occupied by objects
  - Then perform region shape analysis/recognition
- Group those pixels that are similar in some properties
  - Intensity, color, texture, motion, etc.
- Topic
  - Background subtraction
  - Morphology
  - Region growing

Simple Segmentation

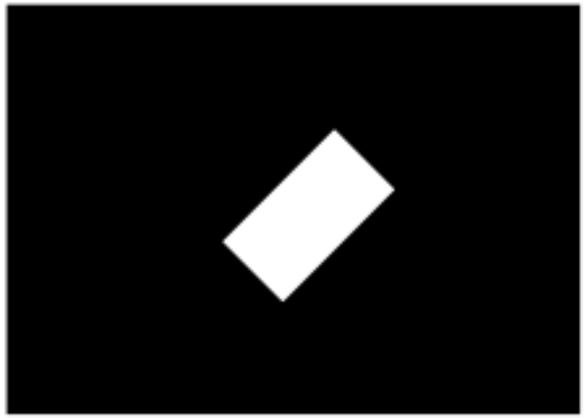
- If object pixels in grayscale image have a higher/lower intensity than background pixel
  - Thresholding image yields binary image
    - E.g. object:1, background 0

Original image I



$$B[x,y] = \begin{cases} 1 & \text{if } I[x,y] > T \\ 0 & \text{otherwise} \end{cases}$$

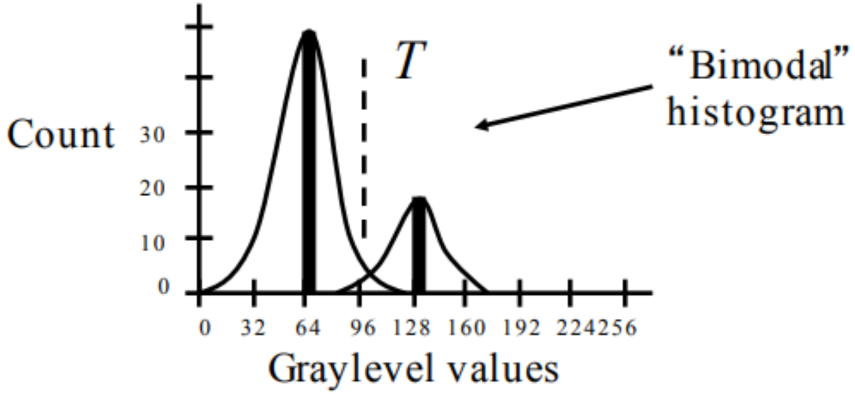
Binary image B



Thresholds and Histograms

- Distribution of graylevels can be used to determine binary threshold

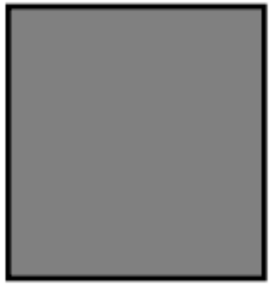
- Histogram graphs number of pixel in the image with a particular graylevel, as a function of the possible graylevels
  - Find peaks and set threshold between peaks
  - Idea: the threshold is chosen such that the division in the histogram yields the largest reduction in standard deviation of the pixel intensities



Background Subtraction I

- If object is not uniformly brighter or darker than the background?
- Get the background image
- Subtract object image from background image
  - Difference highlights the object
  - Use abs or square of difference

Background R



Input I

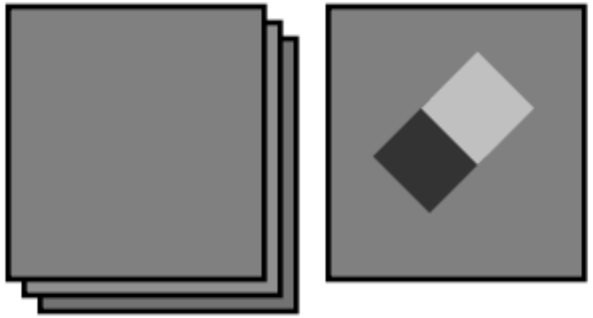
$$B[x,y] = \begin{cases} 1 & \text{if } \left| I[x,y] - R[x,y] \right| > T \\ 0 & \text{otherwise} \end{cases}$$

Binary B

Background Subtraction II

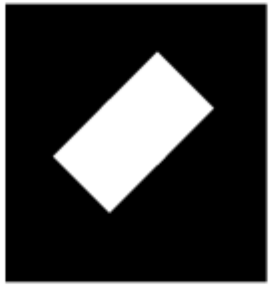
- How to choose the threshold?
  - Want the difference to be above noise level in images
- Obtain sequence of background images
  - Compute intensity mean(μ) and standard deviation(σ) for each pixel (across/through the background images)
- Check statistical distance (#stdev) of object pixels from background image

Backgrounds Input I



$$B[x,y] = \begin{cases} 1 & \text{if } \dfrac{\left(I[x,y]-\mu[x,y]\right)^2}{\sigma^2[x,y]} > T^2 \\ 0 & \text{otherwise} \end{cases}$$

Binary B



Background Subtraction III

- Color image?
  - $\sqrt{(I_R[x,y]-R_R[x,y])^2+(I_G[x,y]-R_G[x,y])^2+(I_B[x,y]-R_B[x,y])^2} > T$
  - Euclidean distance in 3D (RGB space)

- $\frac{(I_C[x,y]-\mu_C[x,y])^2}{\sigma_C^2[x,y]} > T_C^2$
- AND/OR result for each color channel

Better Model

- Use the “covariance” of color information

RGB vector: $C_{x,y} = \left[I_R[x,y], I_G[x,y], I_B[x,y] \right]^T$

Background model

$$\left\{ \begin{array}{ll} \text{Mean vector:} & M_{x,y} = \frac{1}{N} \sum_i C_{x,y}^i \\ \text{Covariance matrix:} & K_{x,y} = \frac{1}{N} \sum_i (C_{x,y}^i - M_{x,y})(C_{x,y}^i - M_{x,y})^T \end{array} \right.$$

$$B[x,y] = \begin{cases} 1 & \text{if } [C_{x,y} - M_{x,y}]^T K_{x,y}^{-1} [C_{x,y} - M_{x,y}] > T^2 \\ 0 & \text{otherwise} \end{cases}$$

Multi-dimensional Mahalanobis Distance

Shadows

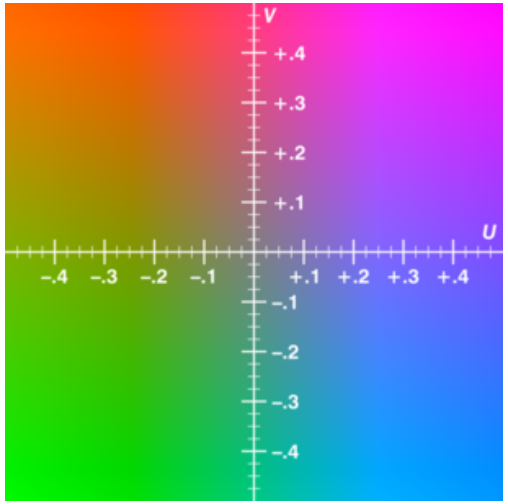
- May not want to select shadows in image as part of the object
  - Shadows due to change (lower) in intensity
  - But, still same chromaticity

- Normalize brightness of RGB space
  - $R^n = \frac{R}{R+G+B}$ $G^n = \frac{G}{R+G+B}$ $B^n = \frac{B}{R+G+B}$

- Or convert RGB to YUV, HSV, or YIQ
  - Intensity and chrominance are decoupled
  - Perform previous method using chrominance channels

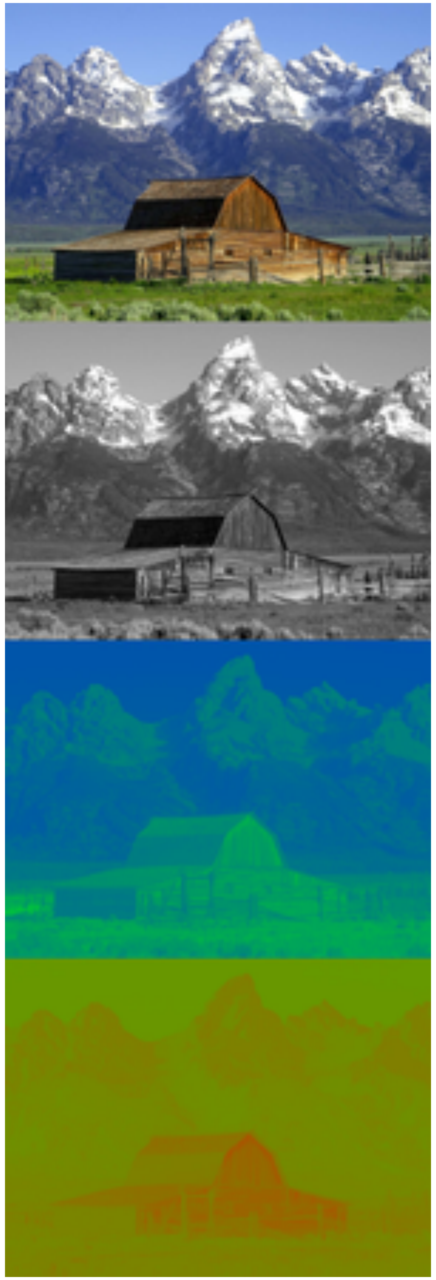
YUV Space

- Y: brightness
- U: Chrominance
- V: Chroma
- YUV color space
  - Decouples intensity and color
    - Y component provides all monochrome/intensity information
    - U and V components carry the color information



$$
\begin{aligned}
Y &= 0.299 * R + 0.587 * G + 0.114 * B \\
U &= -0.169 * R - 0.331 * G + 0.5 * B + 128 \\
V &= 0.5 * R - 0.419 * G - 0.081 * B + 128
\end{aligned}
$$

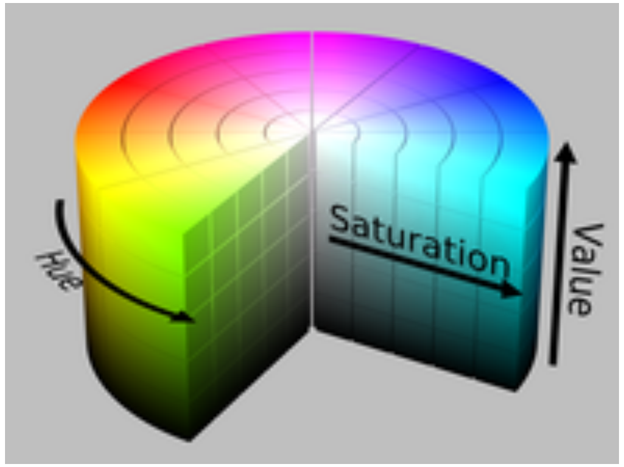
$$
\begin{aligned}
R &= Y + 1.13983 * (V - 128) \\
G &= Y - 0.39465 * (U - 128) - 0.58060 * (V - 128) \\
B &= Y + 2.03211 * (U - 128)
\end{aligned}
$$

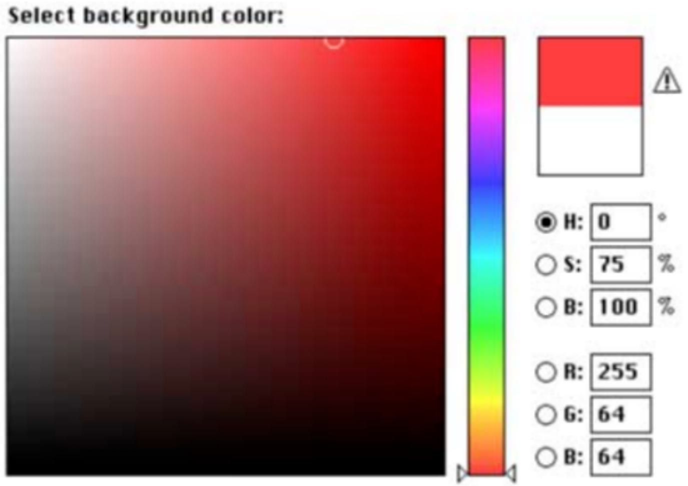


*https://zh.wikipedia.org/wiki/YUV

HSV Space

- Hue, saturation and lightness





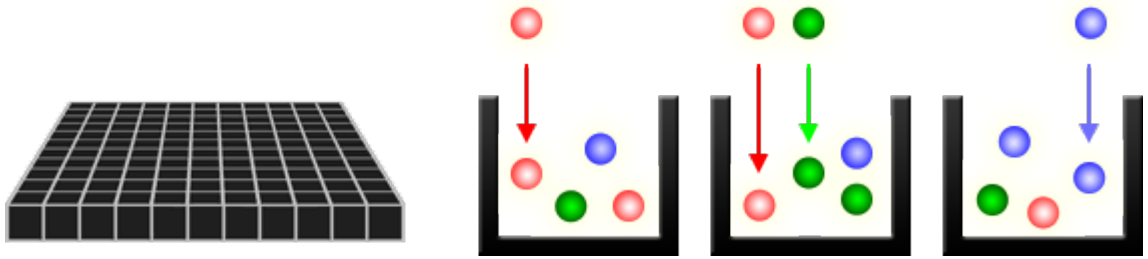
Color conversion equation: https://en.wikipedia.org/wiki/HSL_and_HSV#To_RGB

Blue/Green Screening

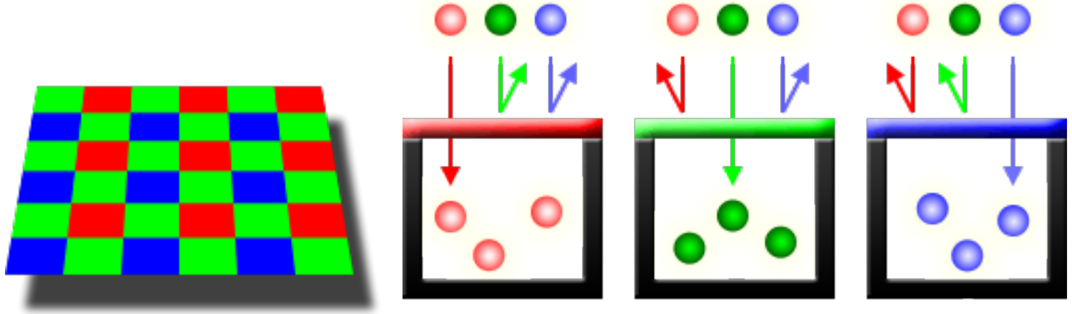
- A subject is photographed in front of a evenly bright and pure blue/green background
- The compositing process, whether photographic or electronic, replaces all the blue/green in the picture with another image, known as the background plate

Why Green?

- Choose the color is rarely in the foreground
- Why not red, orange or yellow?
  - Human skin color is in between yellow and orange
- Video cameras most sensitive to green channel
  - Green channel is cleanest channel in most digital cameras today
  - Bayer Pattern filtering
    - Pixel array records twice as many green pixel as red or blue (recording resolution is double)



Cavity array Light cavities

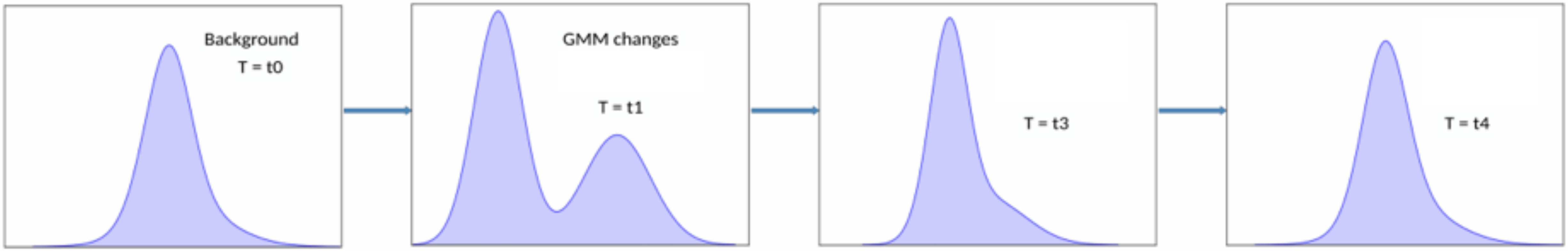


Color filter array Photosities with color filter

*https://www.cambridgeincolour.com/tutorials/camera-sensors.htm

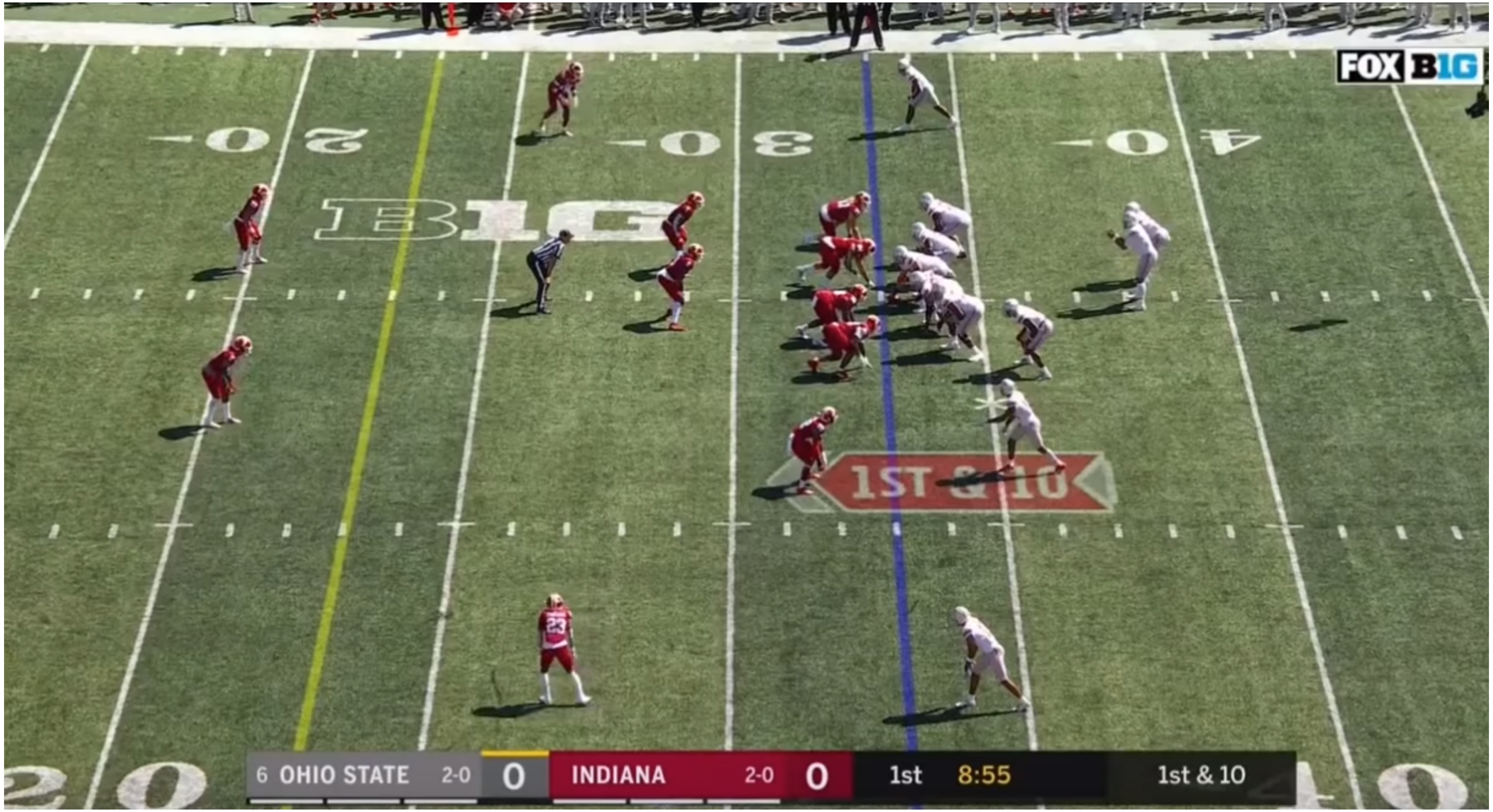
Multi-Model Background Modeling

- Static single distribution background models often too simplistic for long duration videos
- Have to consider for dynamic factors
  - Illumination changes
  - Moving scene elements
  - Objects that are stationary for significant durations
- Need to adapt overtime
- Use temporally adaptive Gaussian Mixture Model (GMM) for each pixel



SporTVision: First-Down Line in American Football

https://www.youtube.com/watch?v=1Oqm6eO6deU

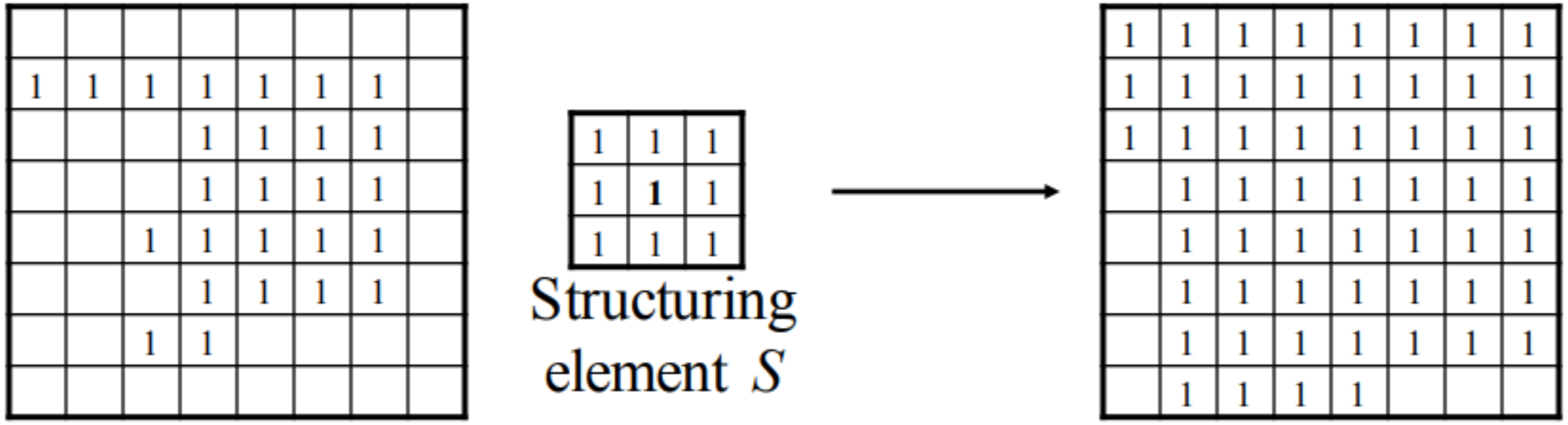


Binary Image Morphology

- Useful once extract binary regions
  - E.g., from background subtraction
- Basic operations of binary morphology
  - Dilation
  - Erosion
  - Closing
  - Opening
- Binary image B
- Structuring element S
  - Smaller binary image or mask

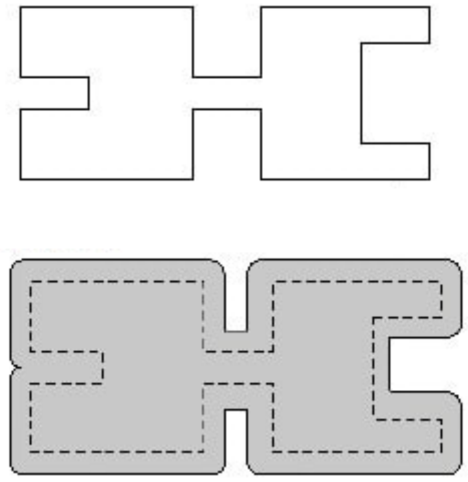
Dilation

- Dilation operation enlarges a binary region
- Convolve S throughout B
  - If center of S touches binary 1-pixel in B, then S is OR-ed with region in B and place into output image



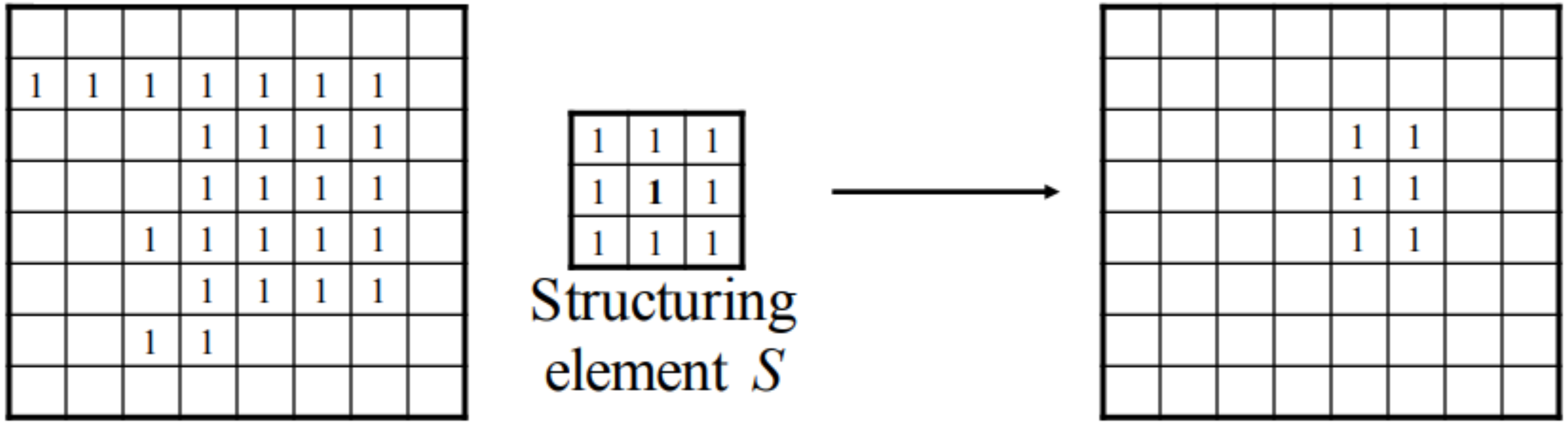
Binary image B

Dilation of B

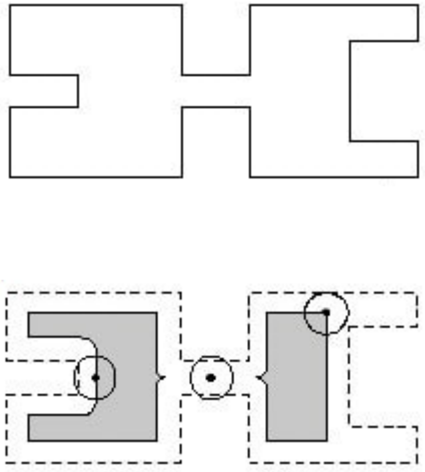


Erosion

- Erosion operation decreases size of binary region
- Convolve S throughout B
  - If every 1-pixel of S touches a binary 1-pixel in B, then center pixel of S is OR-ed with corresponding pixel in B and place into output image

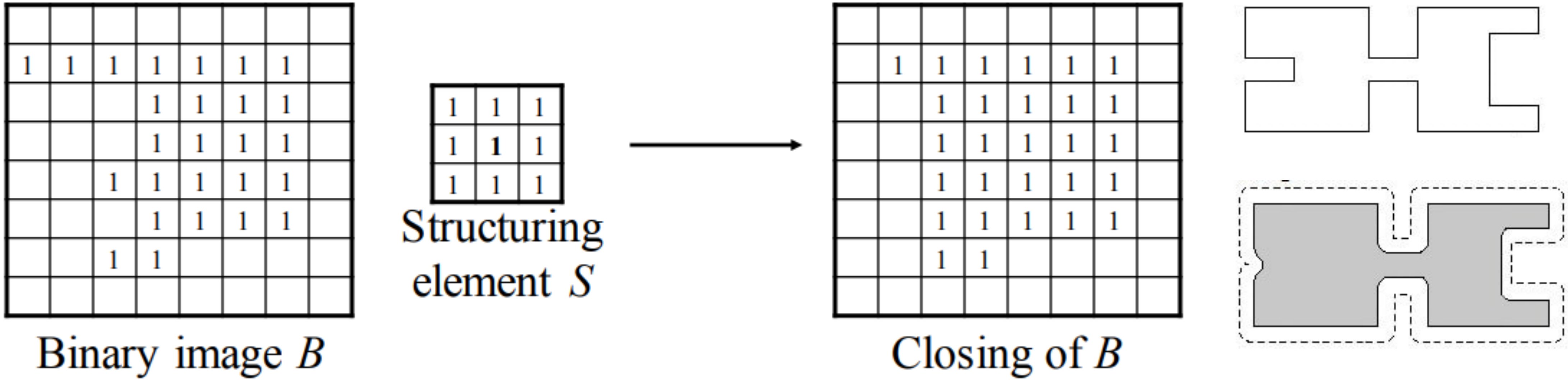


Binary image B Erosion of B



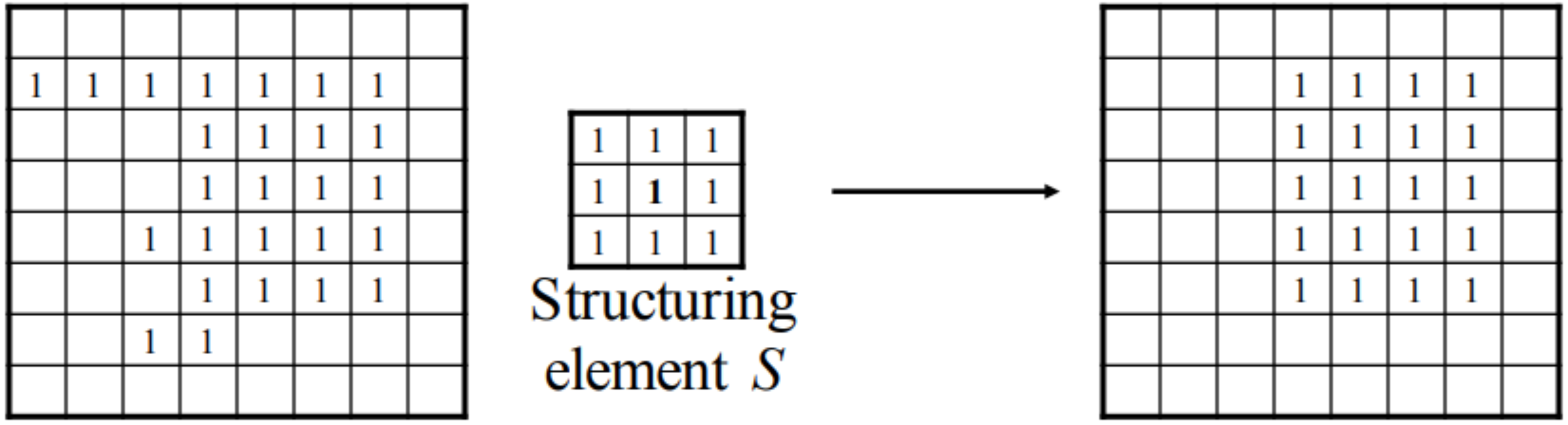
Closing

- Closing operation closes up internal holes and eliminates bayes along boundary
- Perform dilation, then erosion



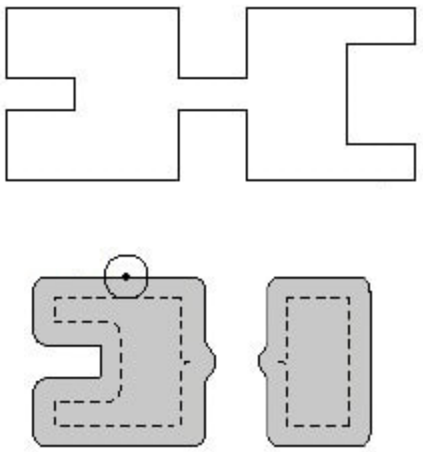
Opening

- Opening operation gets rid of small portions of the binary region that jut out from boundary into background region
- Perform erosion, then dilation

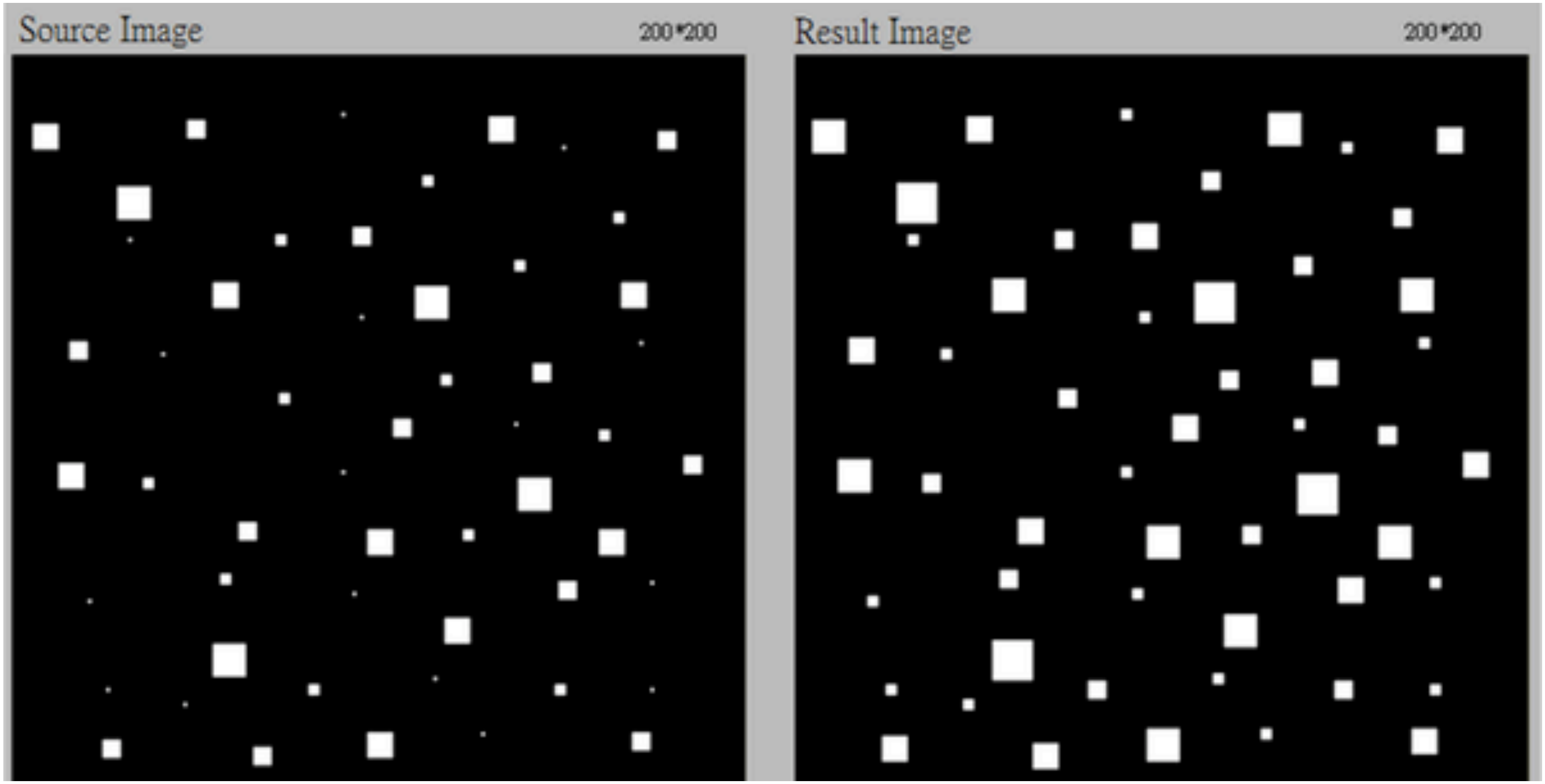


Binary image B

Opening of B



Dilation? Erosion?



Dilation? Erosion?

Raw Binary Image

Dilation? Erosion?

Dilation? Erosion?

Counting Objects

- So far, no notion of a distinct/separate object region
  - No selection of a region of pixels
- Segment binary image (after morphology)
  - Region growing
  - Connected components

Sequential Connected Components

- Steps (2 passes)
  - Scan binary image left-to-right, top-to-bottom
  - If unlabeled 1-pixel encountered, assign new label number according to rules

$$\begin{array}{cc} & 0 \\ 0 & 1 \end{array} \rightarrow \begin{array}{cc} & 0 \\ 0 & L \end{array} \qquad \begin{array}{cc} & 0 \\ L & 1 \end{array} \rightarrow \begin{array}{cc} & 0 \\ L & L \end{array}$$

$$\begin{array}{cc} & L \\ 0 & 1 \end{array} \rightarrow \begin{array}{cc} & L \\ 0 & L \end{array} \qquad \begin{array}{cc} & L \\ M & 1 \end{array} \rightarrow \begin{array}{cc} & L \\ M & L \end{array} \quad (\text{Set } L = M)$$

  - Determine equivalent classes of labels
  - In second pass, assign same label to all elements in equivalent class

Example

Binary image

1	1		1	1	1		1
1	1		1		1		1
1	1	1	1				1
							1
1	1	1	1		1		1
			1		1		1
1	1		1				1
1	1		1		1	1	1

After Pass 1

1	1		2	2	2		3
1	1		2		2		3
1	1	1	1				3
							3
4	4	4	4		5		3
			4		5		3
6	6		4				3
6	6		4		7	7	3

After Pass 2

1	1		**1**	**1**	**1**		3
1	1		**1**		**1**		3
1	1	1	1				3
							3
4	4	4	4		5		3
			4		5		3
6	6		4				3
6	6		4		**3**	**3**	3

Equivalence classes:

(1,2)

(3,7)