

SSP 100 - 10

How Computers See

09 / 24 / 2026

Instructor: Michael Eckmann

Today's Topics

- Questions? / Comments?
- Histograms
- Distance between histograms
- Uses in Content Based Image Retrieval

Looking Ahead

- I will give you some readings by 10 a.m. tomorrow dealing with the recent technical topics:
 - Binary numbers
 - Convolution
 - For edge detection and blurring and other purposes
 - Histograms and their use in content-based image retrieval
- Next topics to be covered including class discussion:
 - more high-level than technical like the recent material

Image Retrieval

- Suppose you have a large number of images stored and you have a particular “query” image.
- Goal:
 - find the image or images that are most similar to the query image
 - Or better yet, output a sorted list of all the images stored that range from most similar to the query image down to least similar

Histogram

- Let's consider first a 1 channel grayscale image, where each pixel is a value from 0 .. 255
- A 256 bin histogram of that image would be a structure of 256 cells, each of which holds the number of pixels for each of the values.
- Example on the board.

Histogram

- Instead of 256 bins we could make fewer bins like say 16 bins.
- Bin 0 holds a count of how many pixels in range 0..15
- Bin 1 holds a count of how many pixels in range 16..31
- Bin 2 holds a count of how many pixels in range 32..47
- ...
- Bin 14 holds a count of how many pixels in range 224..239
- Bin 15 holds a count of how many pixels in range 240..255

histograms

- So, a histogram of a greyscale (aka intensity) image is a function
 - whose domain is a set of grey values and
 - And whose value at a grey value is the number of pixels in the image with that grey value
- For a greyscale image
 - we could make 256 bins, where each bin is one grey value or
 - if we want 32 bins, then each bin will be associated with $256/32 = 8$ grey values. 0-7 in bin0, 8-15 in bin1, ..., 248-255 in bin31.
 - our histogram will be a function whose domain is the bins and whose value is the number of pixels in the image with any of the grey values corresponding to that bin

Histogram

- If we are considering a histogram of an RGB image, we could have over 16 million bins, 1 bin for each different RGB value.
- But just like gray value 26 and 27 look about the same brightness to our eyes, say $\text{RGB}=(120,30,20)$ looks about the same color to us as $\text{RGB}=(121,28,22)$.
- Example of 16 bins on the next slide (R and B divided into 2 groups and G divided into 4 groups)

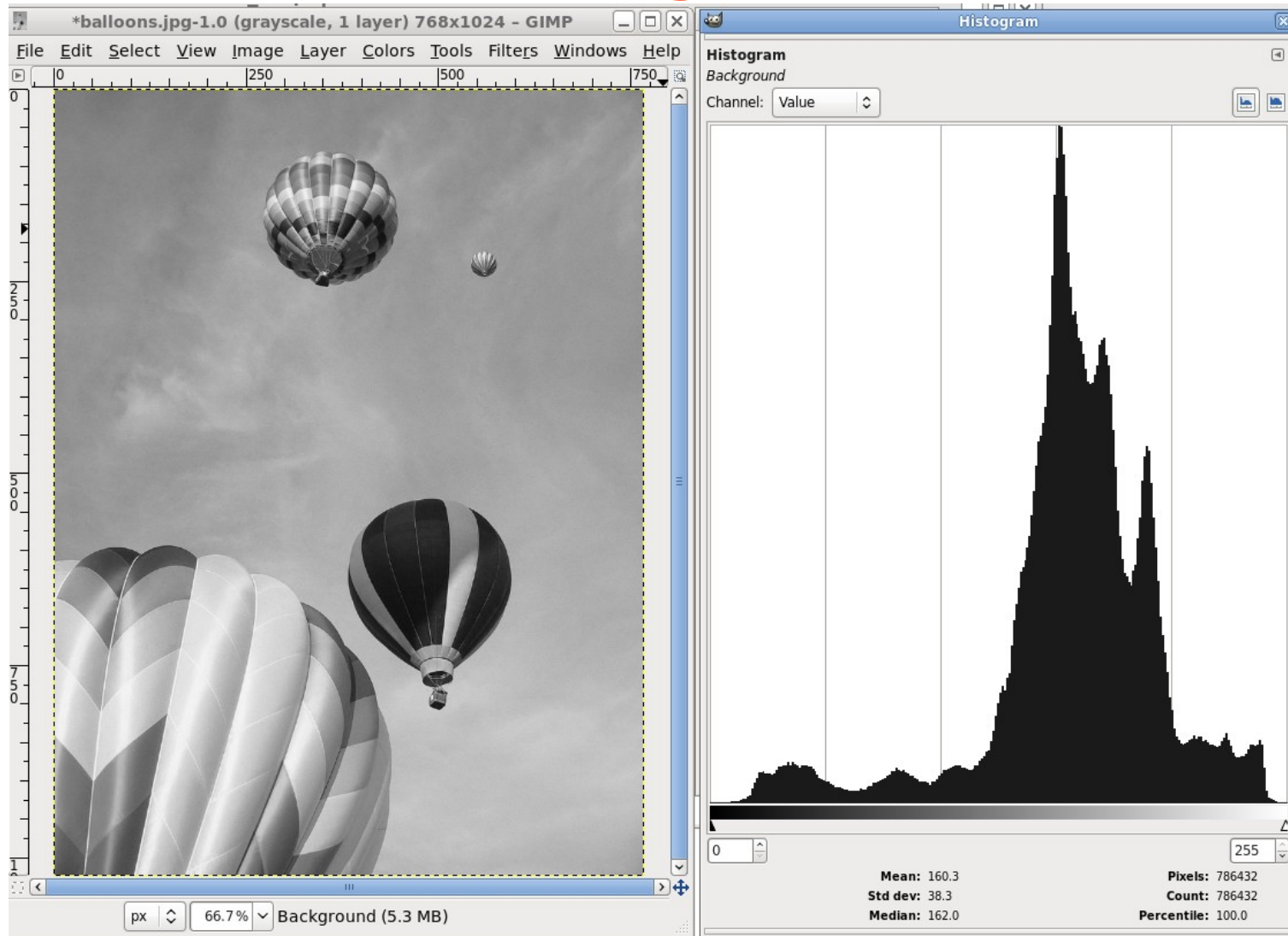
histograms

- For color histograms, since we have multiple channels, we need a way to combine ranges of all the channels into a bin. Consider these 16 bins for a color image histogram:
- Bin0 those with R in 0-127, G in 0-63, B in 0-127
- Bin1 those with R in 0-127, G in 0-63, B in 128-255
- Bin2 those with R in 0-127, G in 64-127, B in 0-127
- Bin3 those with R in 0-127, G in 64-127, B in 128-255
- Bin4 those with R in 0-127, G in 128-191, B in 0-127
- Bin5 those with R in 0-127, G in 128-191, B in 128-255
- Bin6 those with R in 0-127, G in 192-255, B in 0-127
- Bin7 those with R in 0-127, G in 192-255, B in 128-255
- Bin8 those with R in 128-255, G in 0-63, B in 0-127
- Bin9 those with R in 128-255, G in 0-63, B in 128-255
- Bin10 those with R in 128-255, G in 64-127, B in 0-127
- Bin11 those with R in 128-255, G in 64-127, B in 128-255
- Bin12 those with R in 128-255, G in 128-191, B in 0-127
- Bin13 those with R in 128-255, G in 128-191, B in 128-255
- Bin14 those with R in 128-255, G in 192-255, B in 0-127
- Bin15 those with R in 128-255, G in 192-255, B in 128-255

histograms

- A histogram then is a mapping from the domain (the bins) to the range (the number of pixels corresponding to a particular bin)

histograms



Michael Eckmann - Skidmore
College - SSP 100 - Fall 2026

histograms

- Histograms describe the *global* intensity (or color) content of an image.
- Note well --- a histogram ignores where pixels are in the image
- very different images can have the exact same histogram

Determining similarity

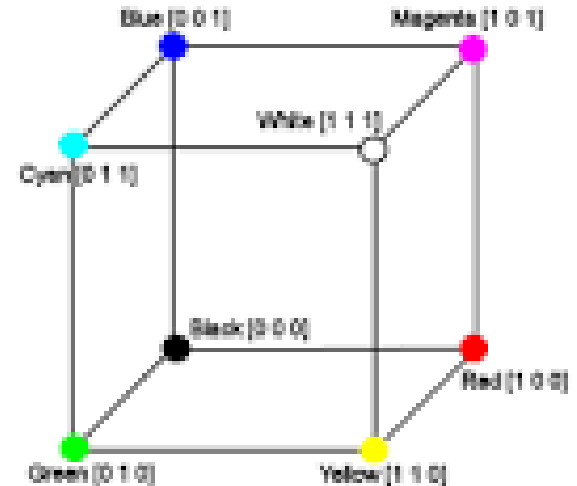
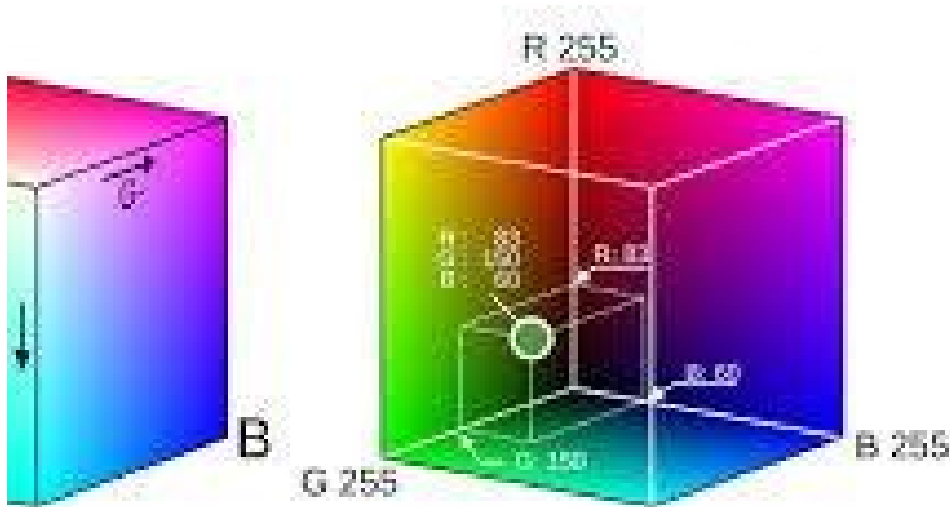
- Many problems in vision require us to compare how similar/dissimilar images are to each other.
- Trying to compare the raw image pixels to each other to determine how similar or dissimilar images are is not going to yield very good results.
- Instead we represent the images in some way. Then we compare how similar/dissimilar their *representations* are.

Similarity

- To motivate how you might compare histograms for similarity, consider how you tell if a point in 2 dimensional space is closer to a second point or a third point.
- Let me draw on the board.
- Closer implies similarity, further implies dissimilarity.

Similarity / Distance

- Consider 3 dimensional space like the RGB cube.
- How might you tell how close two RGB colors (that is 2 points in 3d space) are?



Distance

- This works for any number of dimensional space (not only 2 or 3).
- For example, we can compute the distance between two 16 bin histograms in the same way.
- This distance is called the Euclidean distance, also known as L2 distance.
- Another distance measure is L1 (sometimes called taxicab or Manhattan distance).

Distance

- Consider the 3 histograms on the board.
- What is the Euclidean distance between the first and each of the others?
- Does that seem fair?

Distance

- Fewer bins could help and still use Euclidean and get reasonable answers most of the time.
- Or a different distance measure that takes into account that nearby bins are more similar than bins further away.
- One such distance measure is called the Earth Mover's Distance. Measures how much work it takes to convert one histogram into another.

Distance measures

- the L1 distance (aka Manhattan or Taxicab) is the sum of the absolute values of the differences of each dimension
- the L2 distance (aka Euclidean) is the square root of the sum of squares of the differences of each dimension
- These are easy to see in 2 dimensions.
- Say we have $p = (p_1, p_2)$ and $q = (q_1, q_2)$ then the L1 distance between them is: $\text{abs}(p_1 - q_1) + \text{abs}(p_2 - q_2)$
- The L2 distance between them is: $\text{sqrt}((p_1 - q_1)^2 + (p_2 - q_2)^2)$

Distance measures

- In general for n dimensions,
- if $p=(p_1, p_2, \dots, p_n)$ and $q = (q_1, q_2, \dots, q_n)$
- the L1 distance between them is:
$$\text{abs}(p_1-q_1) + \text{abs}(p_2-q_2) + \dots + \text{abs}(p_n-q_n)$$
- The L2 distance between them is:
$$\text{sqrt}((p_1-q_1)^2 + (p_2-q_2)^2 + \dots + (p_n-q_n)^2)$$

Distance measures

- Let's do an example where the vectors have 5 values (5 dimensions)
- Example: $p = (5,3,2,6,9)$ and $q=(4,4,4,7,6)$ and $r = (10,6,4,5,0)$
- Let's see whether p is closer to q or r using L1
- Example calculation:
- L1 distance between p and $q = |5-4| + |3-4| + |2-4| + |6-7| + |9-6| = 1+1+2+1+3 = 8$
- L1 distance between p and $r = |5-10| + |3-6| + |2-4| + |6-5| + |9-0| = 5+3+2+1+9 = 20$
- So p is closer to q than r when considering L1 distance
- In other words we determined that p is more similar to q than it is to r
- Let's do the same with L2 distance

histograms

- The overall idea is that similar images will have similar histograms and the distance between their histograms will be small, whereas dissimilar images will have very dissimilar histograms and the distance between their histograms will be large.
- But remember a histogram just counts up the number of pixels of certain ranges of pixel values and disregards the where the pixels are, so the histogram representation ignores where the pixels are in the image.

histograms

- Because we typically need to compare images that are different sizes (made up of different numbers of pixels) doing the distances on histograms of different sized images doesn't make sense.
- Example on the board

histograms

- Solution: rather than compute distance between their histograms as described earlier (as a value in a bin as the number of pixels) we instead use normalized histograms before computing distance.
- Normalized histograms are those that have each bin value from the original histogram divided by the total number of pixels in the image. Therefore the sum of all the values of a normalized histogram will equal 1. Each bin stores the fraction of the total pixels that have that bin grey value.

Content Based Image Retrieval

- Let me show you a demo of a program that computes the histogram of each image in a folder and does similarity distances of the normalized histograms and shows the top five most similar images and the five least similar images.
- Texture and Shape
- Addition of a spatial component

Recap

Histograms

- What they are.
- Why use them?
- What are they capturing and what are they ignoring
- Distance measures for similarity
- Why use normalized histogram vs. regular histogram

Looking Ahead

- I will give you some readings by 10 a.m. tomorrow dealing with the recent technical topics:
 - Binary numbers
 - Convolution
 - For edge detection and blurring and other purposes
 - Histograms and their use in content-based image retrieval
- Next topics to be covered including class discussion:
 - more high-level than technical like the recent material