Features

CMP719– Computer Vision
Pinar Duygulu
Hacettepe University

Digital Color Images

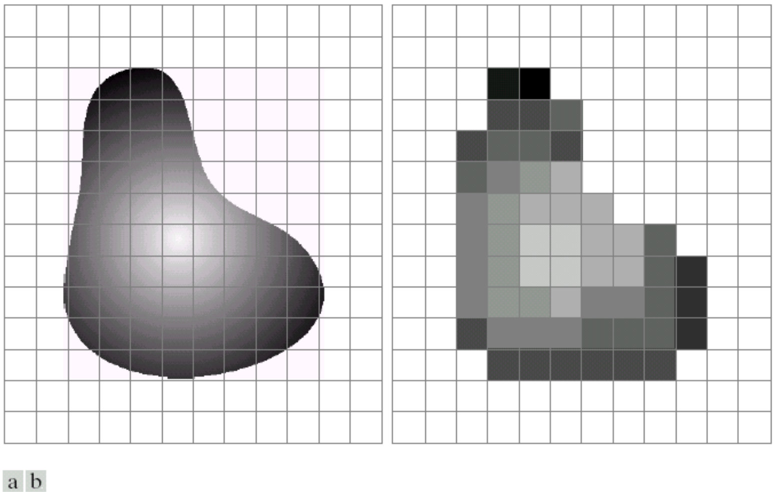
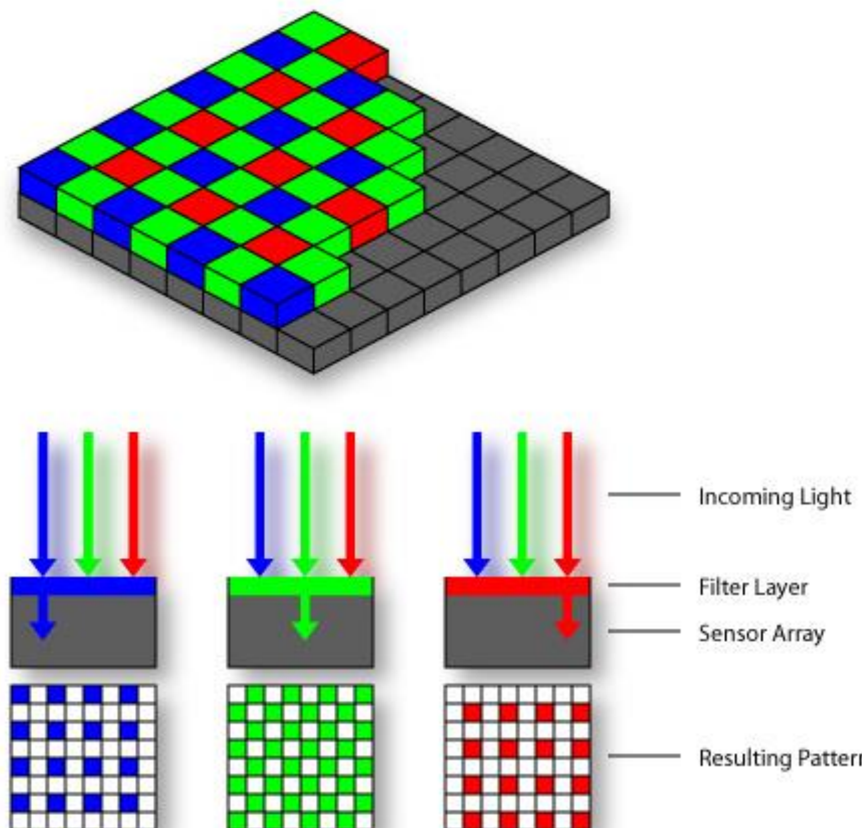
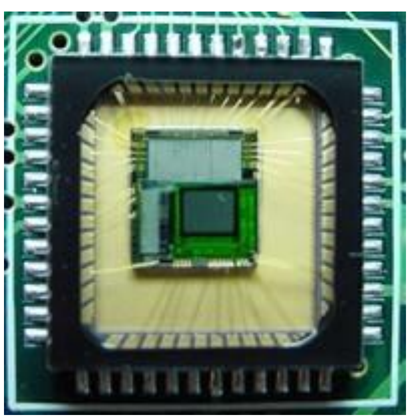


FIGURE 2.17 (a) Continuos image projected onto a sensor array. (b) Result of image sampling and quantization.

Bayer Filter

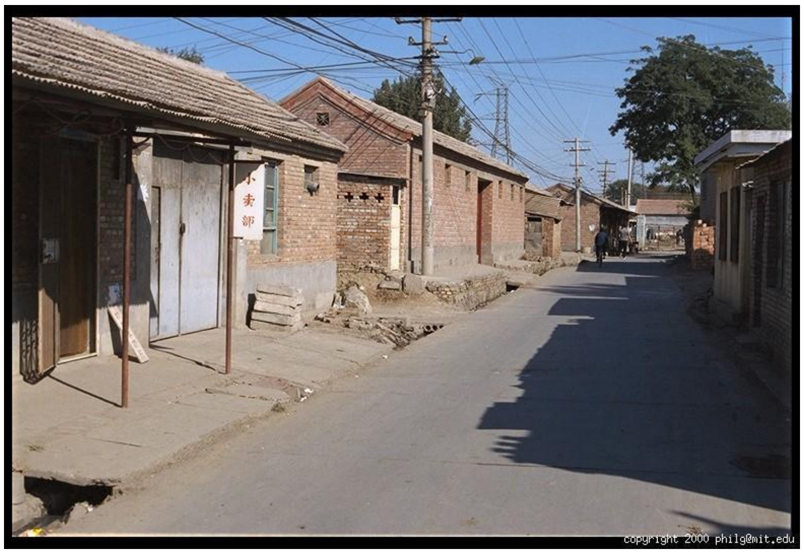




CMOS sensor

Slide credit: Derek Hoiem

Color Image



R

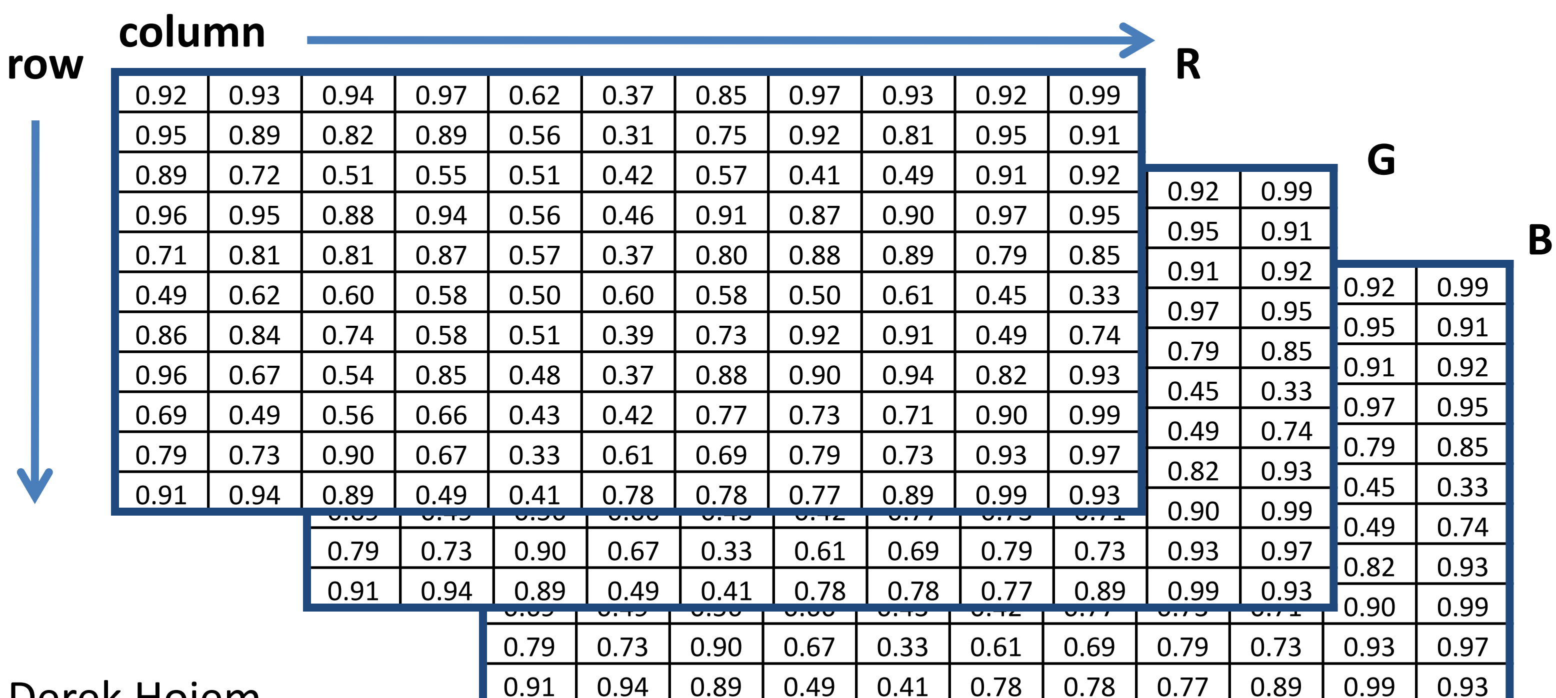
G

B

Slide credit: Derek Hoiem

Images in Matlab

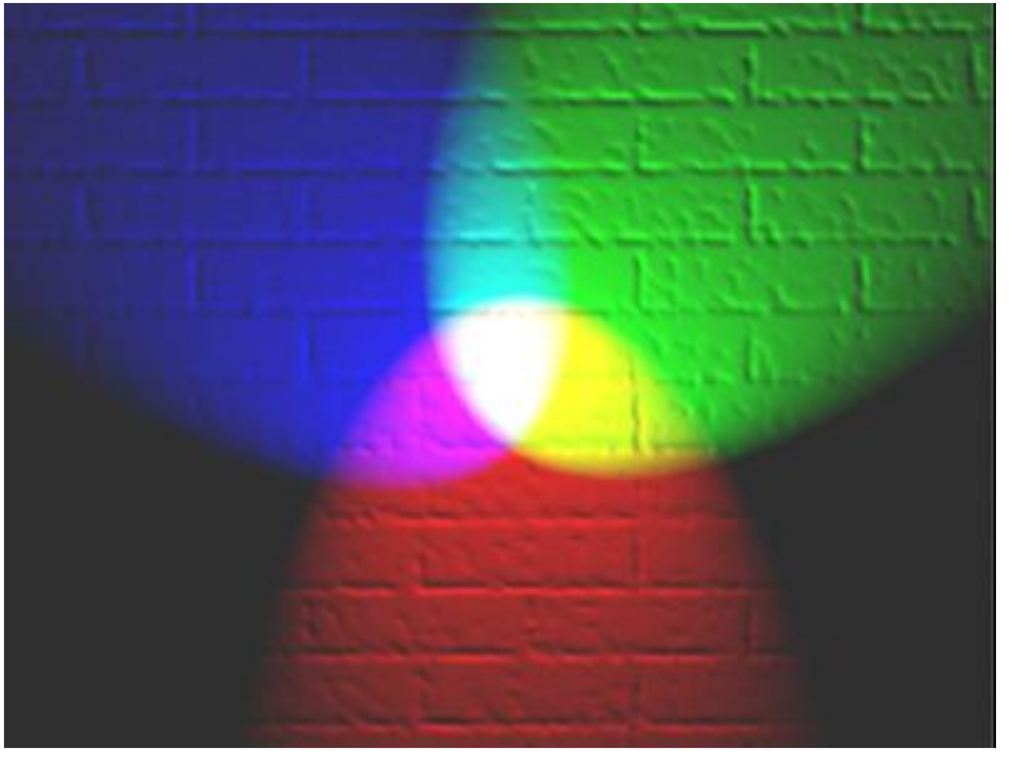
- Images represented as a matrix
- Suppose we have a NxM RGB image called "im"
  - im(1,1,1) = top-left pixel value in R-channel
  - im(y, x, b) = y pixels down, x pixels to right in the bth channel
  - im(N, M, 3) = bottom-right pixel in B-channel
- imread(filename) returns a uint8 image (values 0 to 255)
  - Convert to double format (values 0 to 1) with im2double



Slide credit: Derek Hoiem

Color spaces

- How can we represent color?

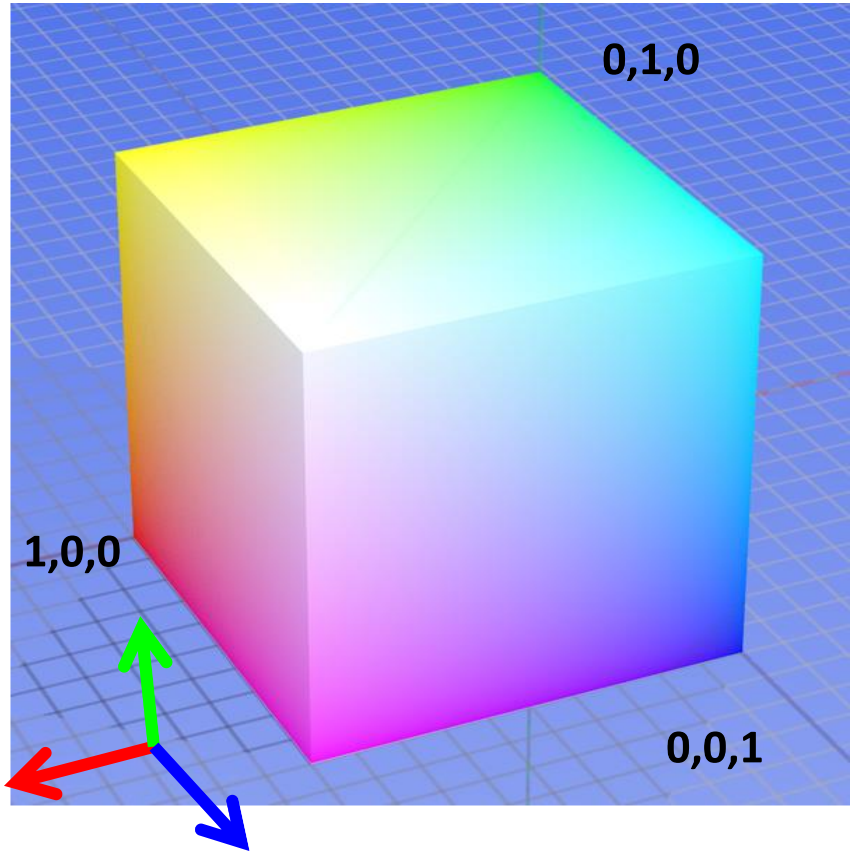


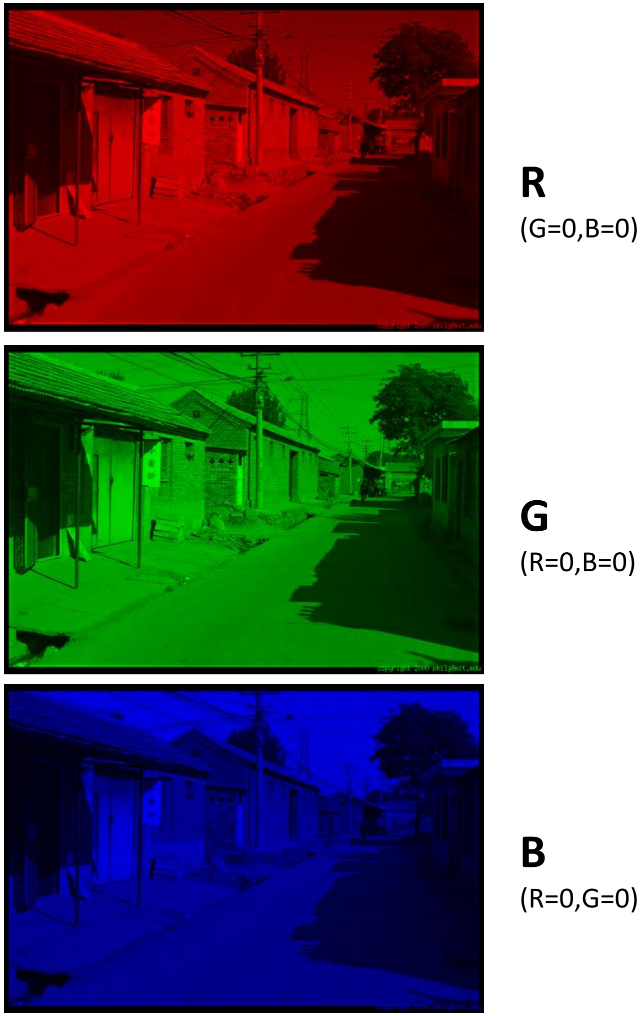
Slide credit: Derek Hoiem

http://en.wikipedia.org/wiki/File:RGB_illumination.jpg

Color spaces: RGB

Default color space





Some drawbacks

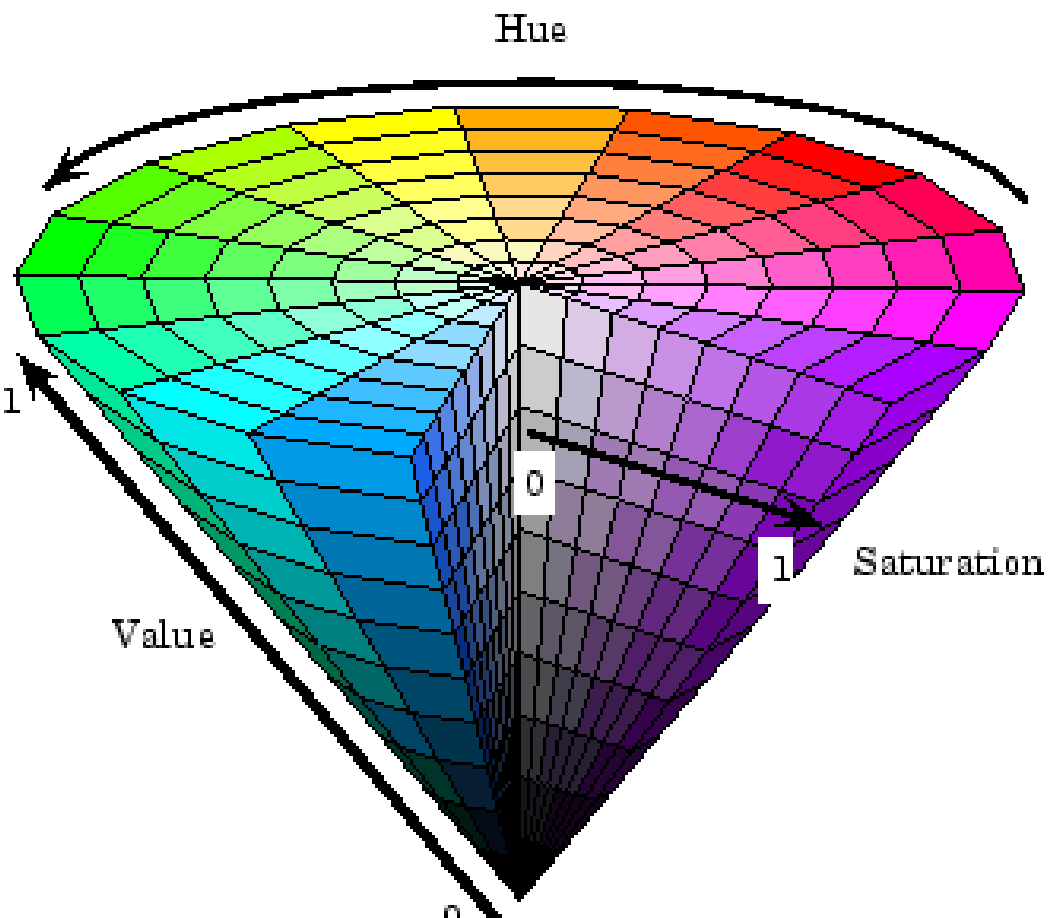
- Strongly correlated channels
- Non-perceptual

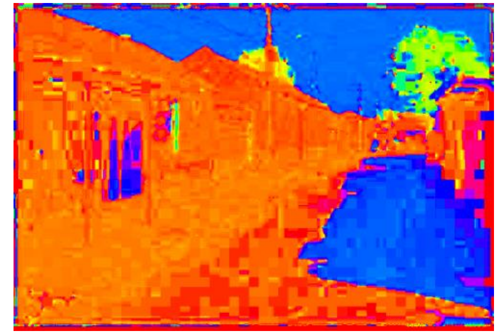
Slide credit: Derek Hoiem

Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

Color spaces: HSV

Intuitive color space





H
(S=1,V=1)

S
(H=1,V=1)

V
(H=1,S=0)

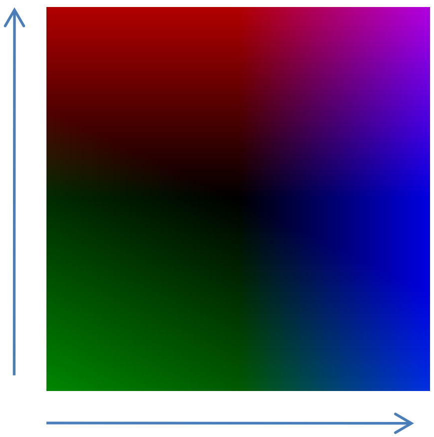
Slide credit: Derek Hoiem

Color spaces: YCbCr

Fast to compute, good for compression, used by TV

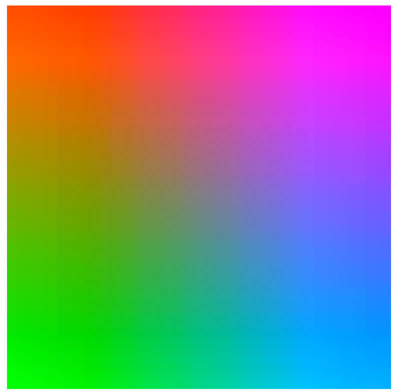
Y=0

Cr



Cb

Y=0.5



Y=1

Y
(Cb=0.5,Cr=0.5)

Cb
(Y=0.5,Cr=0.5)

Cr
(Y=0.5,Cb=05)

$$Y' = 16 + \frac{65.738 \cdot R'_D}{256} + \frac{129.057 \cdot G'_D}{256} + \frac{25.064 \cdot B'_D}{256}$$

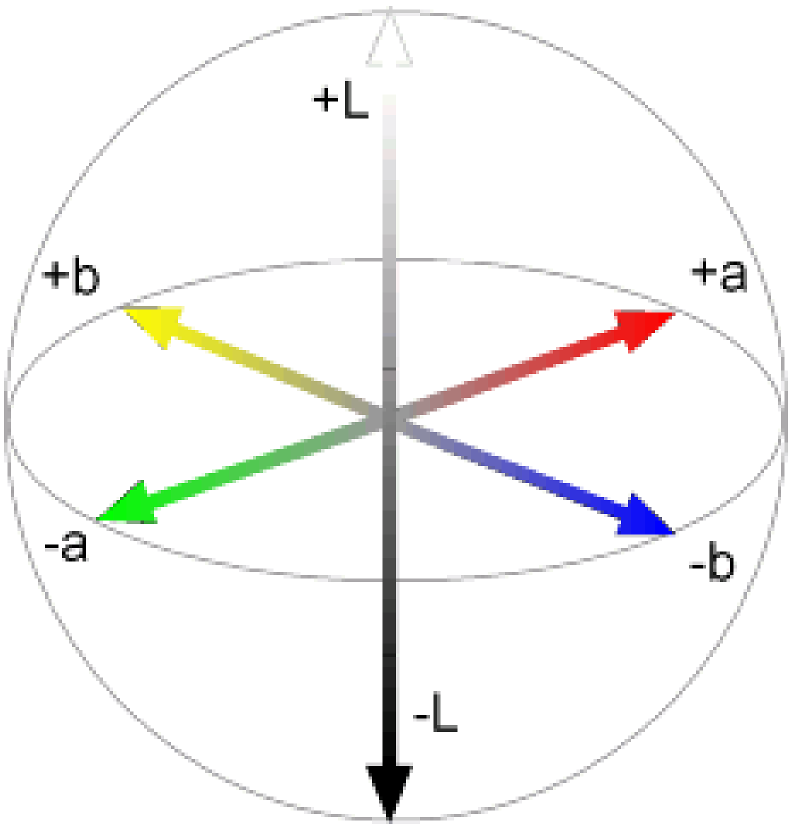
$$C_B = 128 + \frac{-37.945 \cdot R'_D}{256} - \frac{74.494 \cdot G'_D}{256} + \frac{112.439 \cdot B'_D}{256}$$

$$C_R = 128 + \frac{112.439 \cdot R'_D}{256} - \frac{94.154 \cdot G'_D}{256} - \frac{18.285 \cdot B'_D}{256}$$

Slide credit: Derek Hoiem

Color spaces: CIE L*a*b*

"Perceptually uniform" color space



L
(a=0,b=0)

a
(L=65,b=0)

b
(L=65,a=0)

Luminance = brightness
Chrominance = color

Slide credit: Derek Hoiem

Which contains more information?

(a) **intensity** (1 channel)

(b) **chrominance** (2 channels)

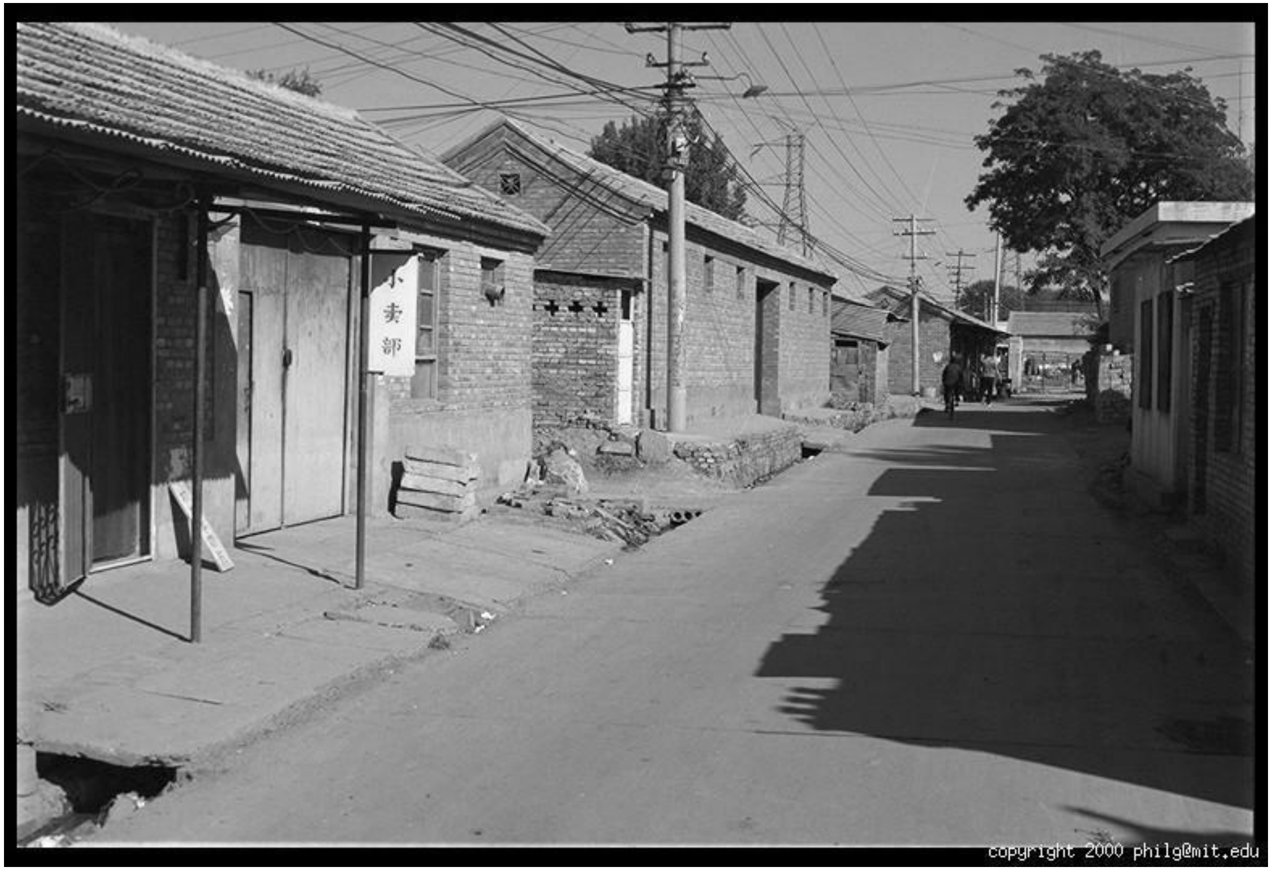
Slide credit: Derek Hoiem

Most information in intensity

Only color shown – constant intensity

Slide credit: Derek Hoiem

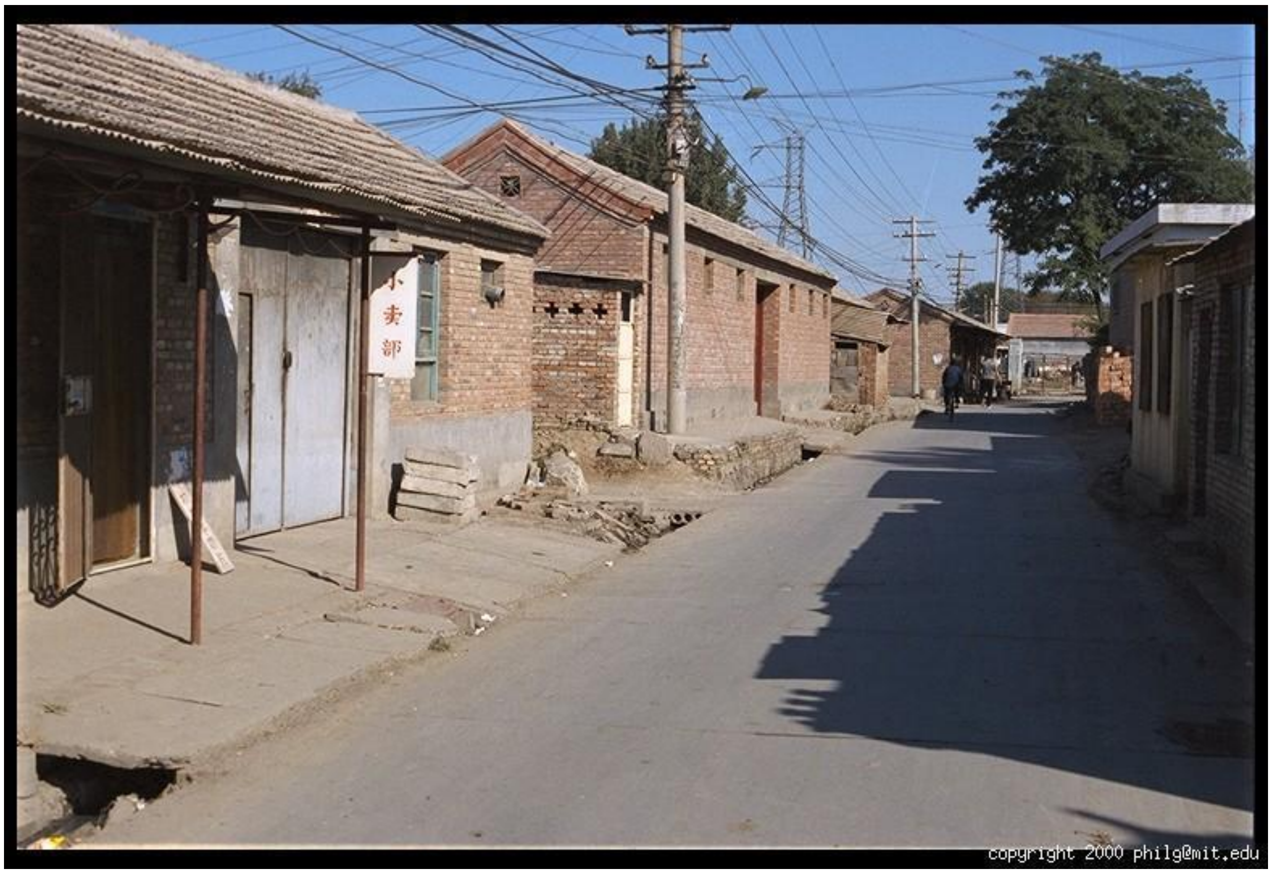
Most information in intensity



Only intensity shown – constant color

Slide credit: Derek Hoiem

Most information in intensity



Original image

Slide credit: Derek Hoiem

Development of Low-Level Image Features

1999

Classical features

- Raw pixel
- Histogram feature
  - Color histogram
  - Edge histogram
- Frequency analysis
- Image filters
- Texture features
  - LBP
- Scene features
  - GIST
- Shape descriptors
- Edge detection
- Corner detection

Local Descriptors

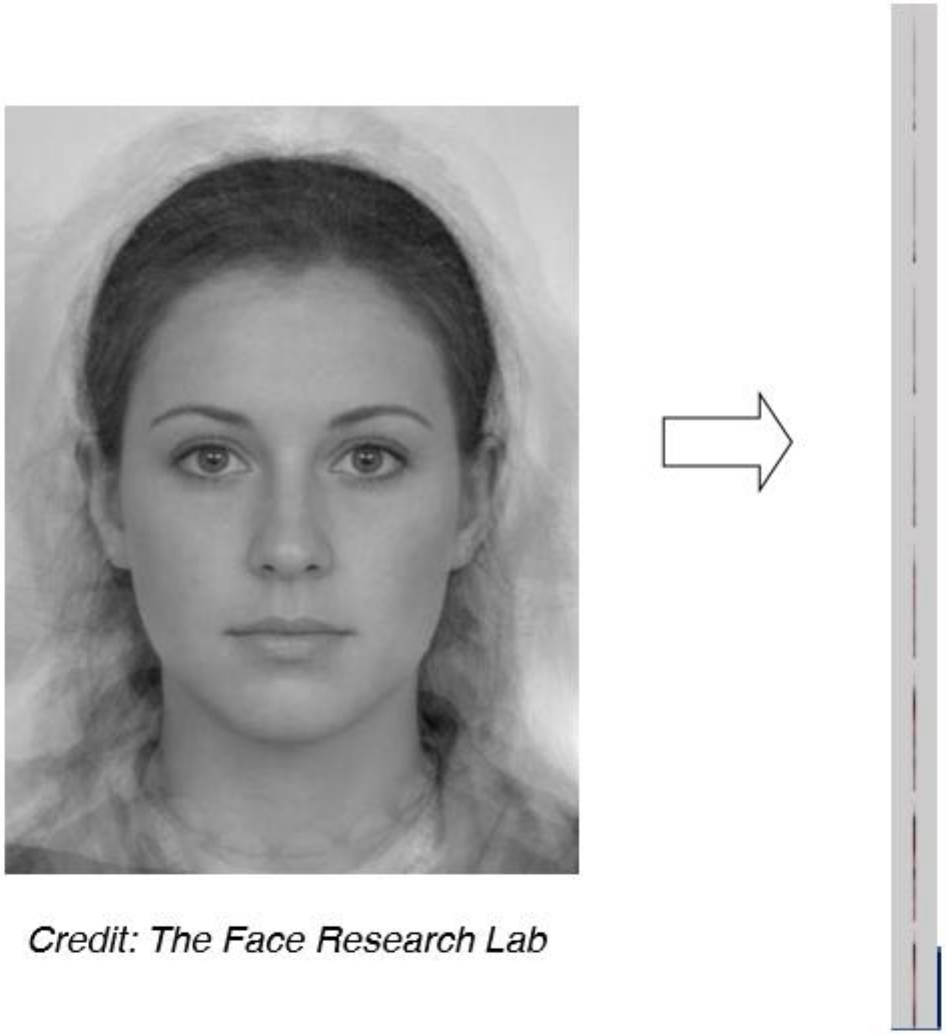
- SIFT
- HOG
- SURF
- DAISY
- BRIEF
- ...

Local Detectors

- DoG
- Hessian detector
- Laplacian of Harris
- FAST
- ORB
- ...

Slide credit: Liangliang Cao

Concatenating Raw Pixels As 1D Vector



Credit: The Face Research Lab

Slide credit: Liangliang Cao

Concatenated Raw Pixels

Famous applications (widely used in ML field)

- Face recognition

- Hand-written digits

Pictures courtesy to Sam Roweis

Slide credit: Liangliang Cao

Tiny Images

Antonio Torralba et al. proposed to resize images to 32x32 color thumbnails, which are called "tiny images"

office

waiting area

dining room

dining room

- Related applications
  - Scene recognition
  - Object recognition

Fast speed with limited accuracy

Torralba et al, MIT CSAIL report, 2007

Slide credit: Liangliang Cao

Problem of raw-pixel based representation

- Rely heavily on good alignment
- Assume the images are of similar scale
- Suffer from occlusion
- Recognition from different view point will be difficult

We want more powerful features for real-world problems like the following

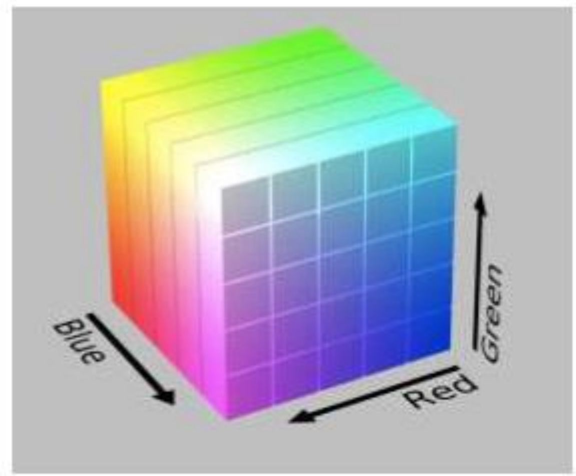
Slide credit: Liangliang Cao

Color Histogram

Each pixel is described by a vector of pixel values

Distribution of color vectors is described by a histogram

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix}$$



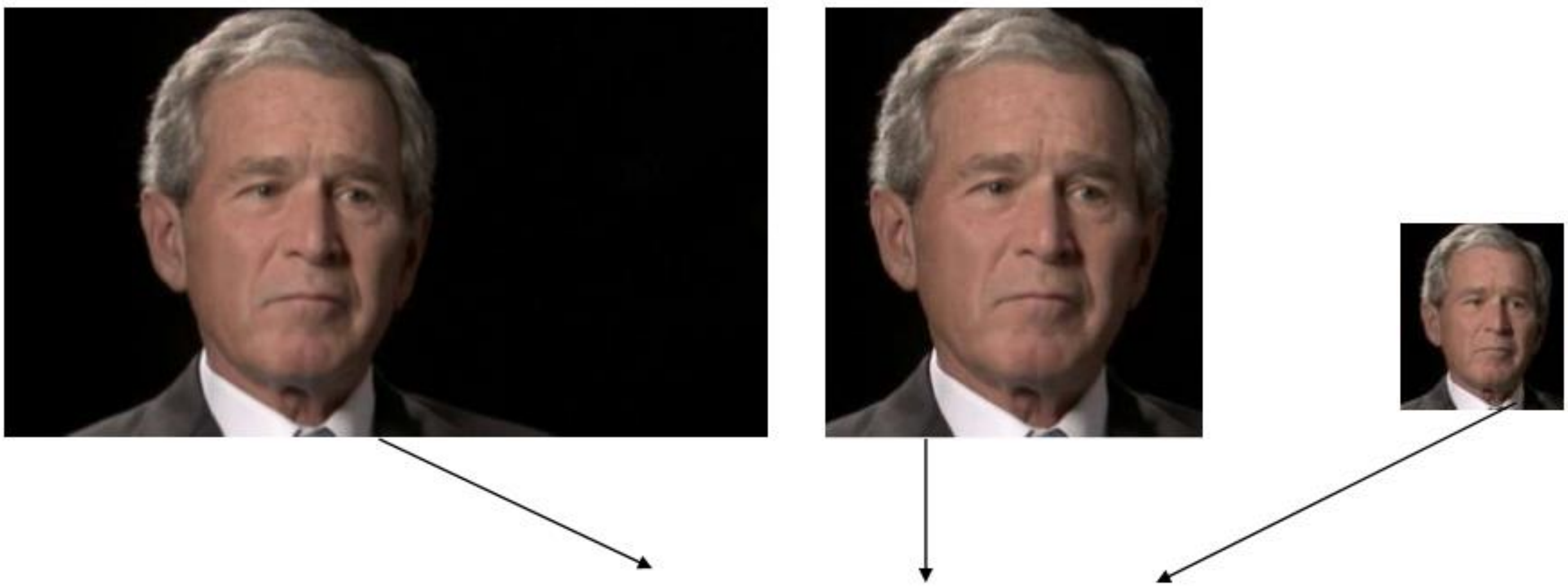
Note: There are different choices for color space: RGB, HSV, Lab, etc.
For gray images, we usually use 256 or fewer bins for histogram.

[Swain and Ballard 91]

Slide credit: Liangliang Cao

Benefits of Histogram Representations

No longer sensitive to alignment, scale transform, or even global rotation

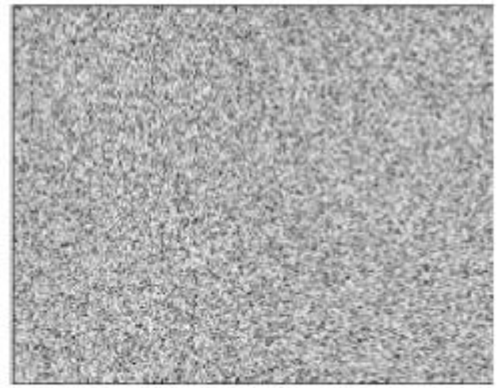


Similar color histograms (*after normalization*)

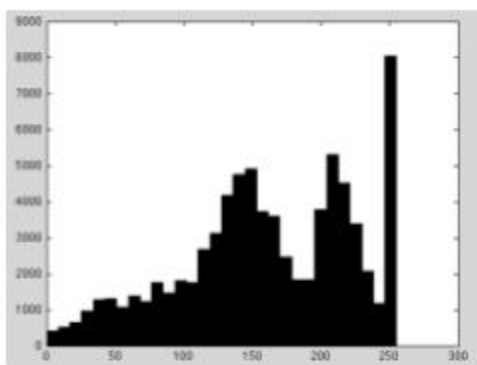
Slide credit: Liangliang Cao

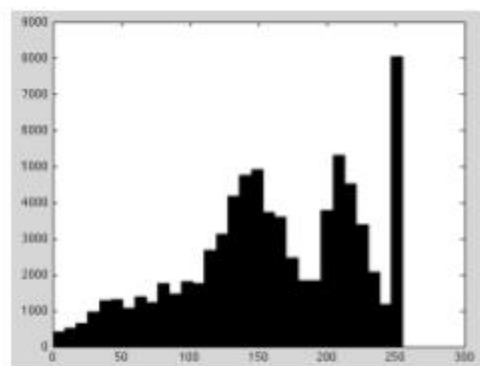
Limitation of Global Histogram

Global histogram has no location information at all

They're equal in terms of global histogram



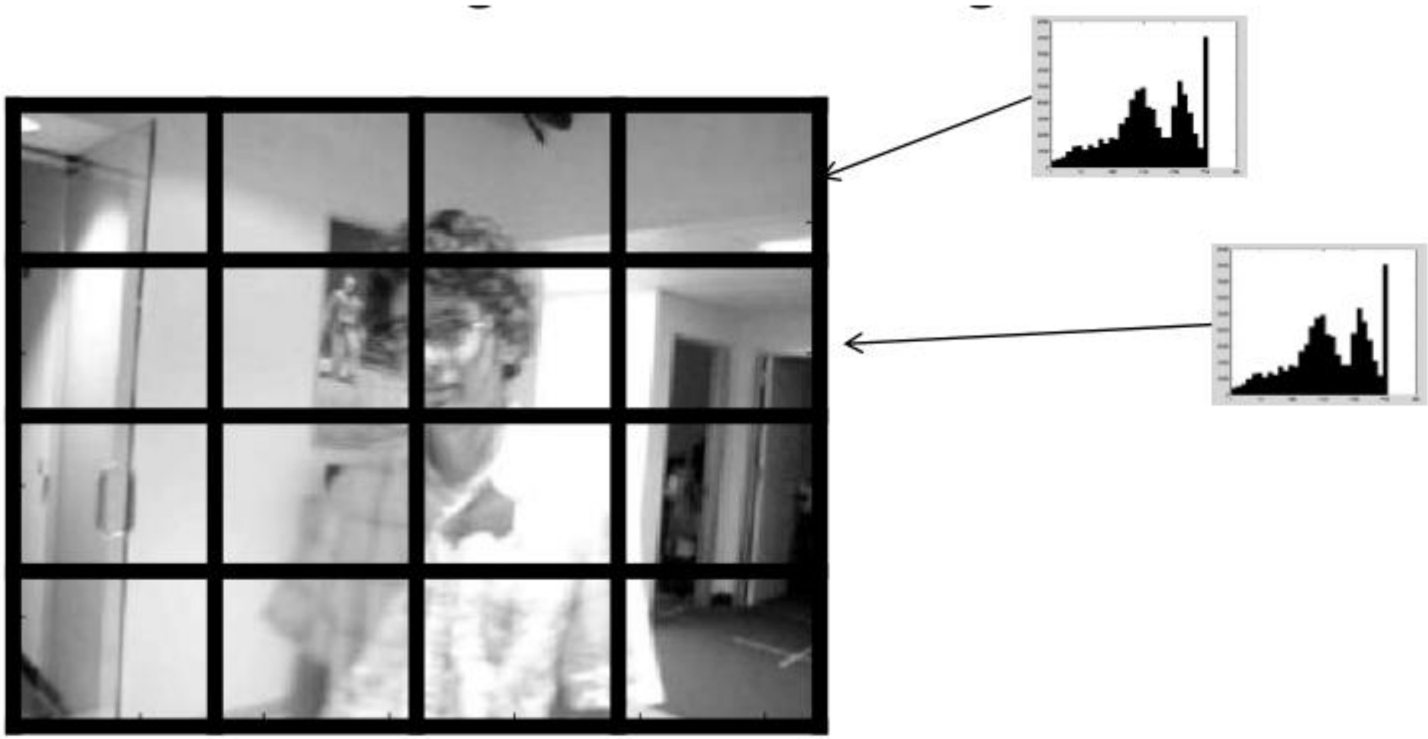


Example courtesy to Erik Learned-Miller

Slide credit: Liangliang Cao

Histogram with Spatial Layout

- Concatenated histogram for each region

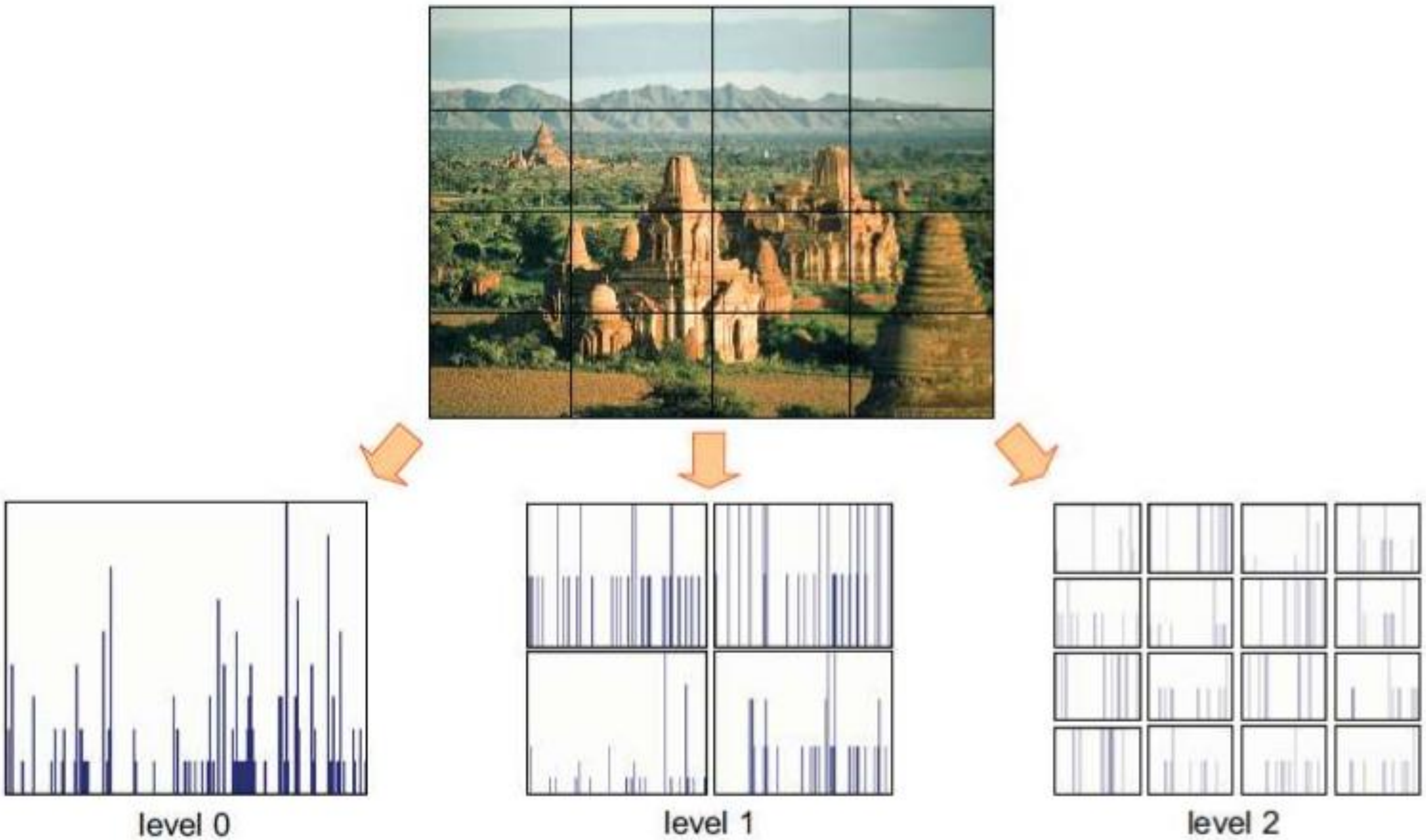


Example courtesy to Erik Learned-Miller

Slide credit: Liangliang Cao

Spatial Pyramid Matching

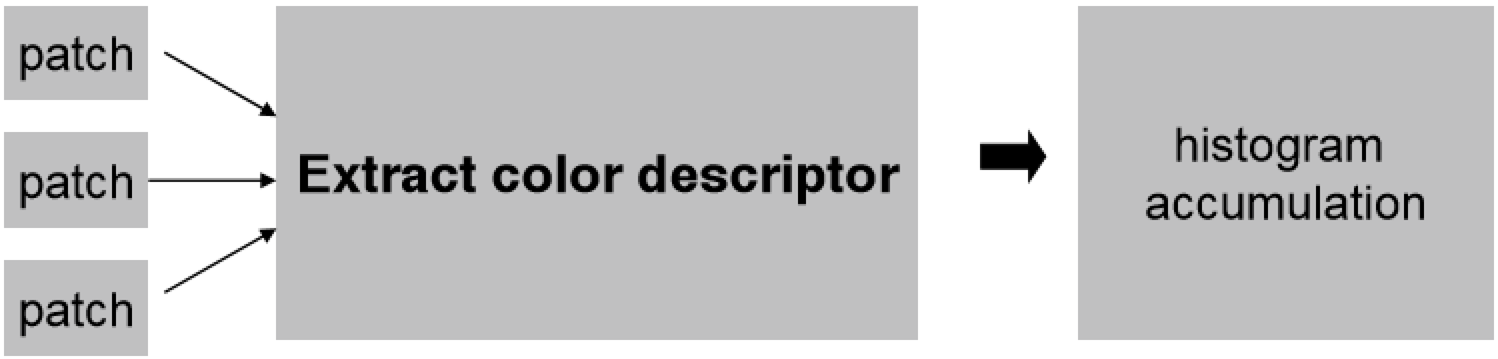
- Lazebnik, Schmid and Ponce, CVPR'06



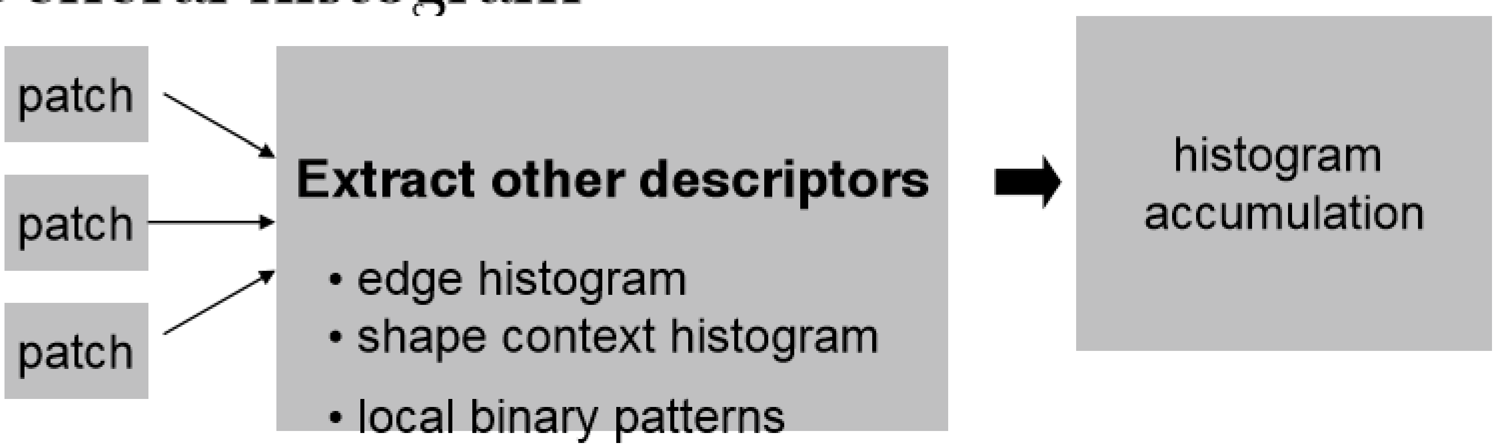
Slide credit: Liangliang Cao

General Histogram Representation

- Color histogram



- General histogram



Slide credit: Liangliang Cao

Local Binary Pattern (LBP)

- For each pixel p, create an 8-bit number b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8, where $b_i = 0$ if neighbor i has value less than or equal to p's value and 1 otherwise.

- Represent the texture in the image (or a region) by the histogram of these numbers.

	1	2	3	
	100	101	103	
8	40	50	80	4
	50	60	90	5
	7	6		

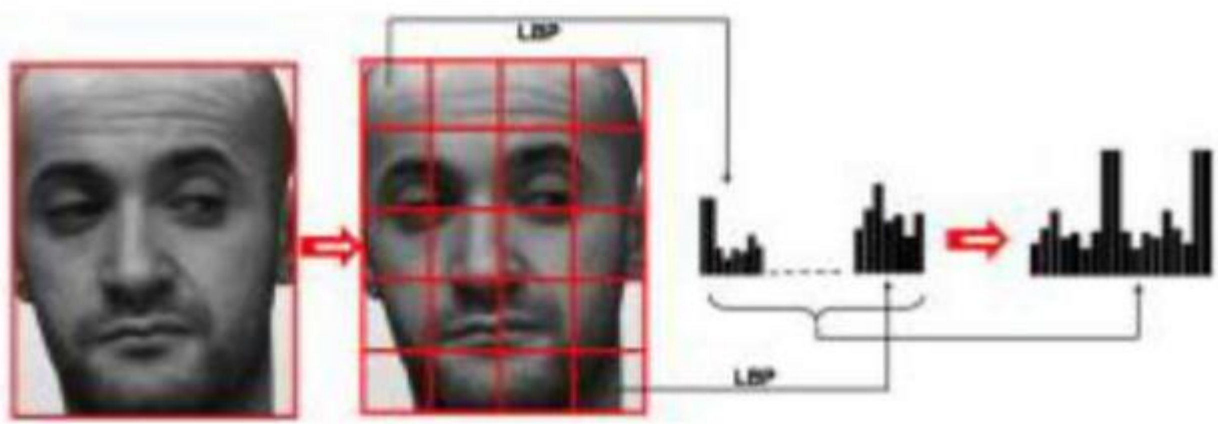
→ 1 1 1 1 1 1 0 0

Ojala et al, PR'96, PAMI'02

Slide credit: Liangliang Cao

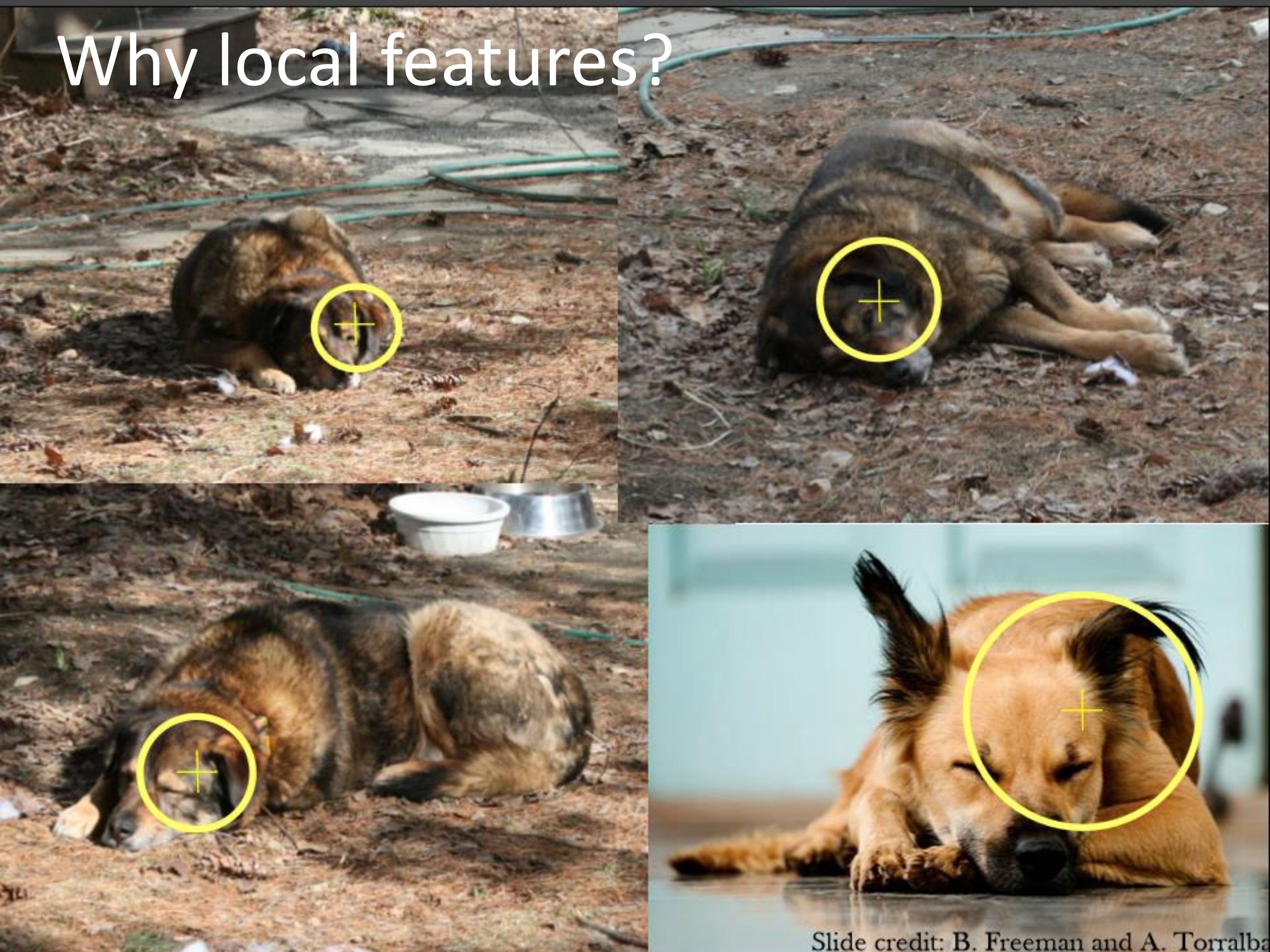
LBP Histogram

- Divide the examined window to cells (e.g. 16x16 pixels for each cell).
- Compute the histogram, over the cell, of the frequency of each "number" occurring.
- Optionally normalize the histogram.
- Concatenate normalized histograms of all cells.



Slide credit: Liangliang Cao

Why local features?



Slide credit: B. Freeman and A. Torralba