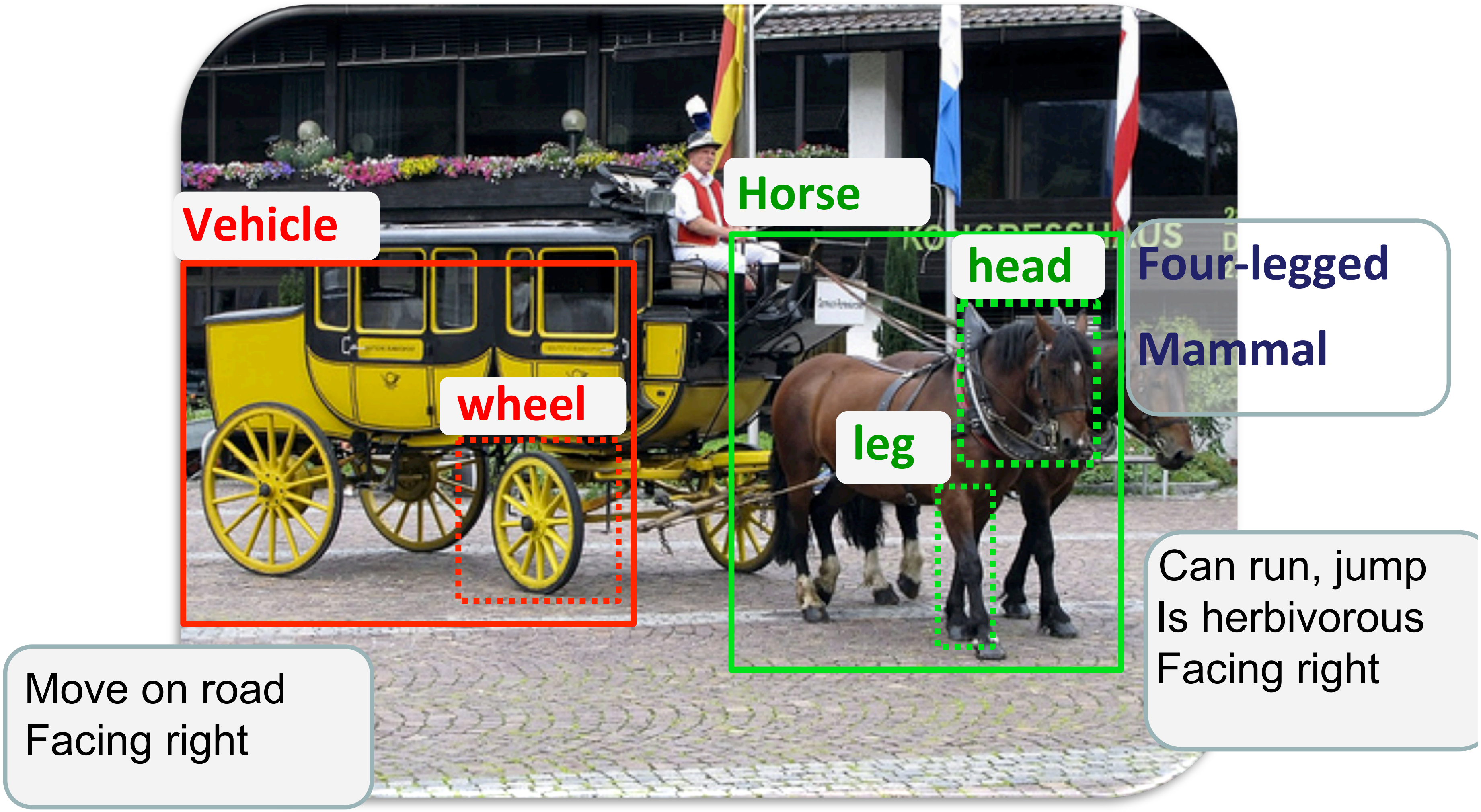
6.S093: Visual Recognition through Machine Learning Competition

Image Representation

Acknowledgment: Many slides from Antonio Torralba, James Hays, Derek Hoeim, and Pete Barnum.

Image Recognition

Image Recognition



Our competition

Vehicle

Horse

Our competition

PLEASE FORM A TEAM UPTO 3 MEMBERS

Vehicle

Horse

Image Representation

- Image is merely a set of numbers

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

Image Representation

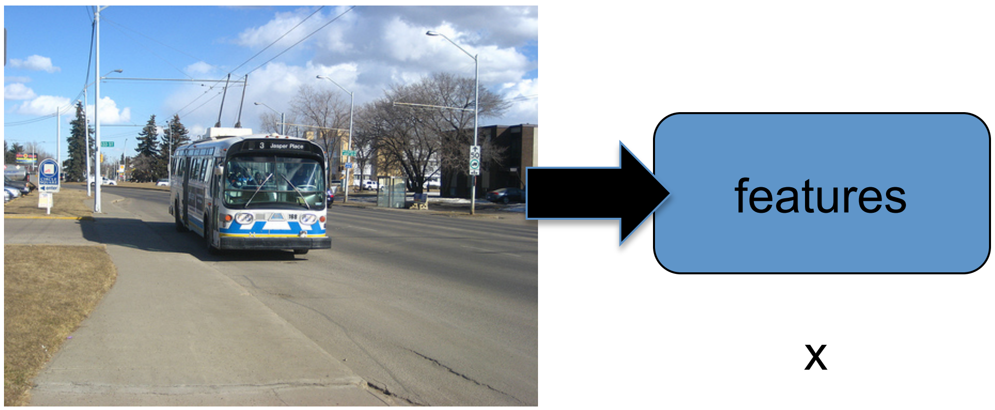
- Image is merely a set of numbers

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

- Images have to be summarized to be more meaningful for machines to interpret.

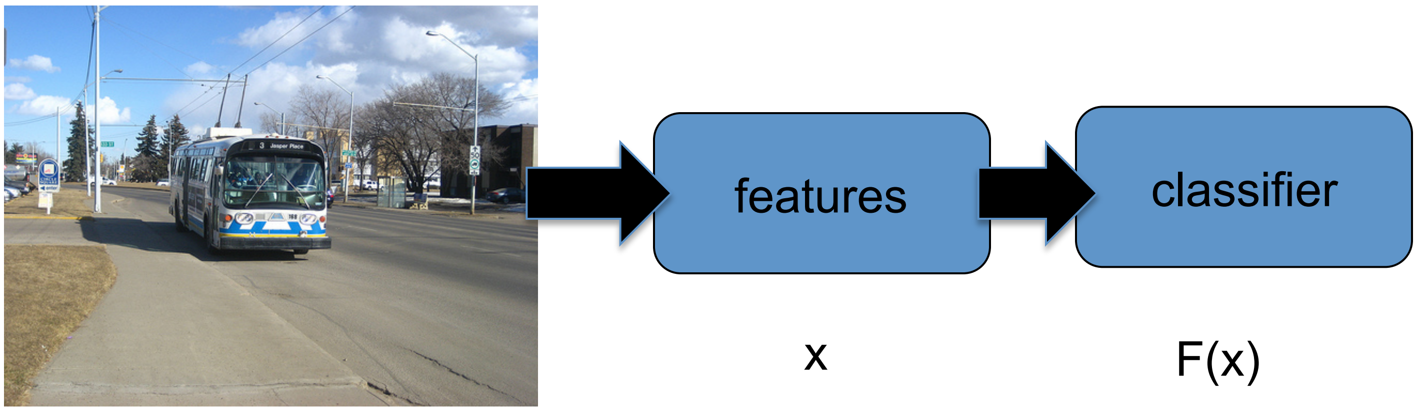
Typical Recognition System

Typical Recognition System



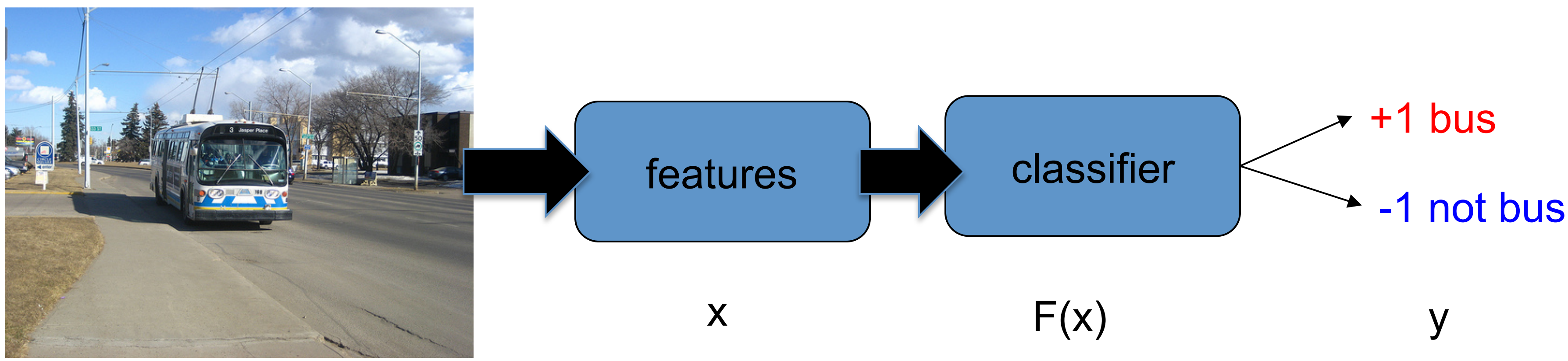
- Extract features from an image

Typical Recognition System



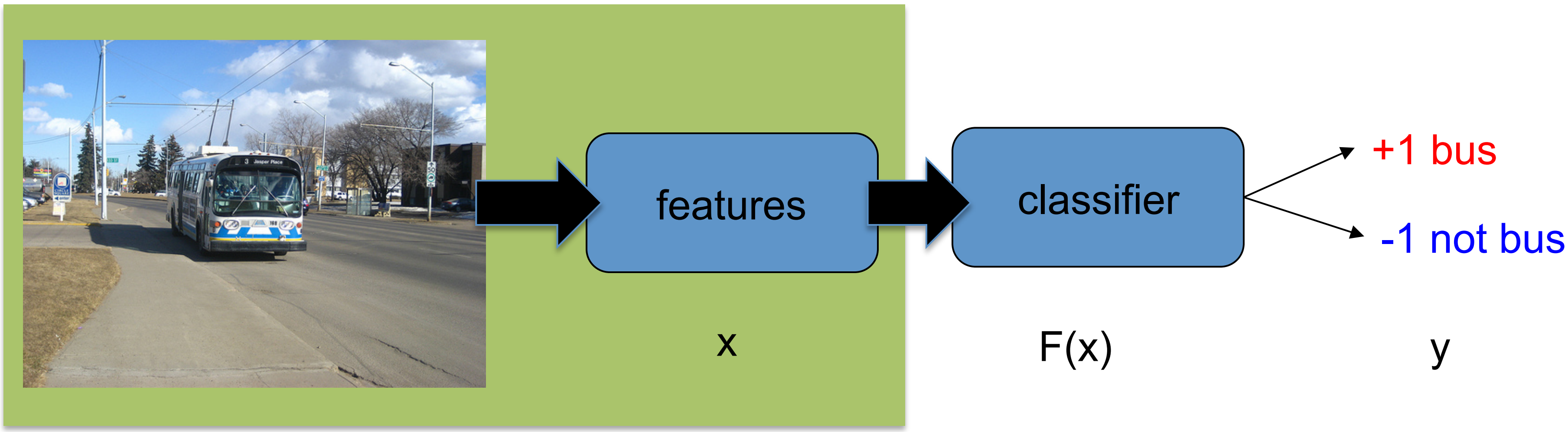
- Extract features from an image
- A classifier will make a decision based on extracted features

Typical Recognition System



- Extract features from an image
- A classifier will make a decision based on extracted features

Typical Recognition System



- Extract features from an image
- A classifier will make a decision based on extracted features

Features

Color histogram — Texture histogram

0.2	0.7	0.1		•••		0.1	0.4	0.2	0.9

0.4	0.1	0		•••		0.9	0.7	0.2	0.4

0.8	0.9	0.7		•••		0	0	0.1	0.7

Different types of representation

- Low-level
  - Color, HOG, SIFT
- Mid-level
  - Part-based, Region-based
- High-level
  - Object presence, object relationship, scene class

Image representation is used for

- Object recognition
- Image alignment and building panoramas
- 3D reconstruction
- Motion tracking
- Indexing and database retrieval
- Robot navigation

Right features depend on the goal

- Object
  - Local shape info, shading, shadows, texture
- Scene
  - Linear perspective, gradients
- Material properties
  - Color, texture
- Motion
  - Optical flow, tracked points

General Principles of Representation

- Coverage
  - Ensure that all relevant info is captured
- Concision
  - Minimize number of features without sacrificing coverage
- Directness
  - Ideal features are independently useful for prediction

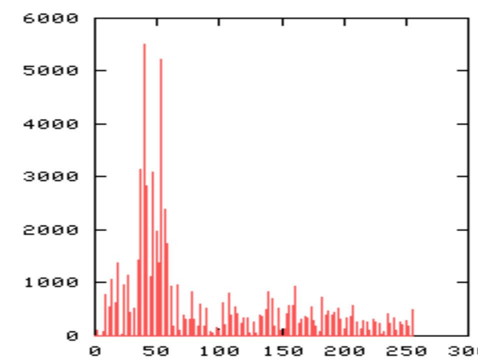
Image representations

- Templates
  - Intensity, gradients, etc.

Image Intensity

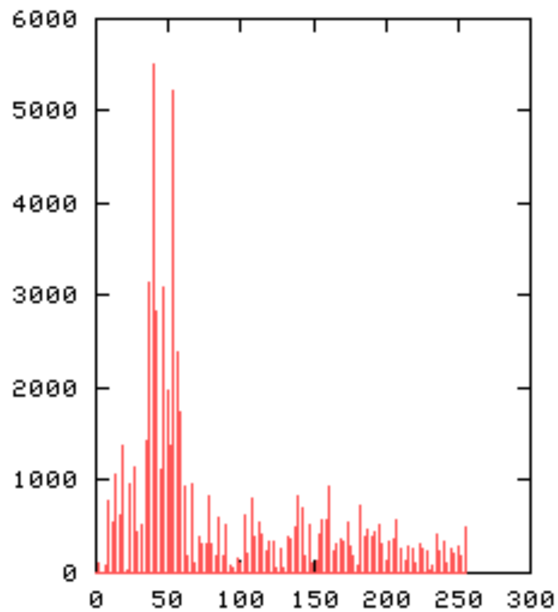
Gradient template

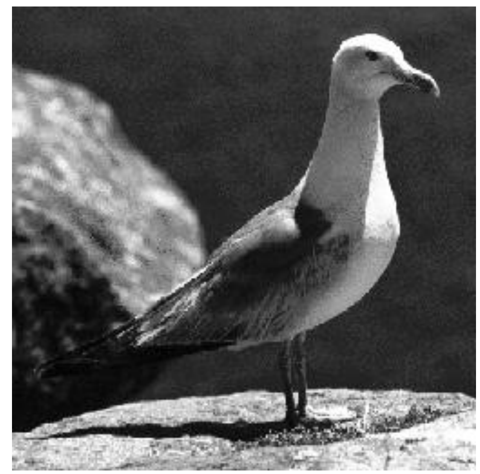
- Histograms
  - Color, texture, SIFT descriptors, etc.



- Average of features

Histograms



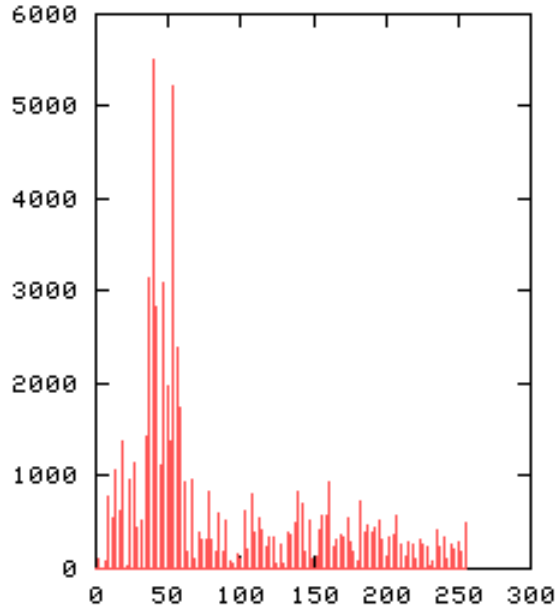


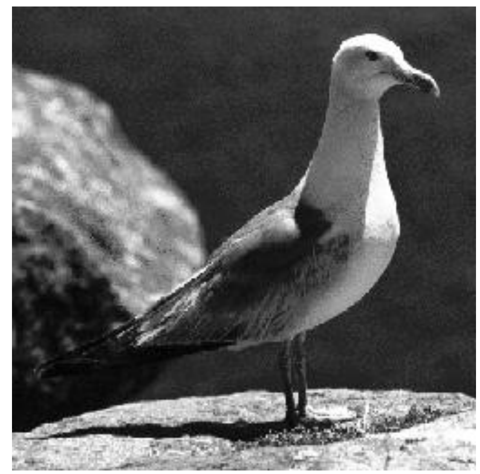
Global histogram

- Represent distribution of features
  - Color, texture, depth, …

Images from Dave Kauchak

Histograms





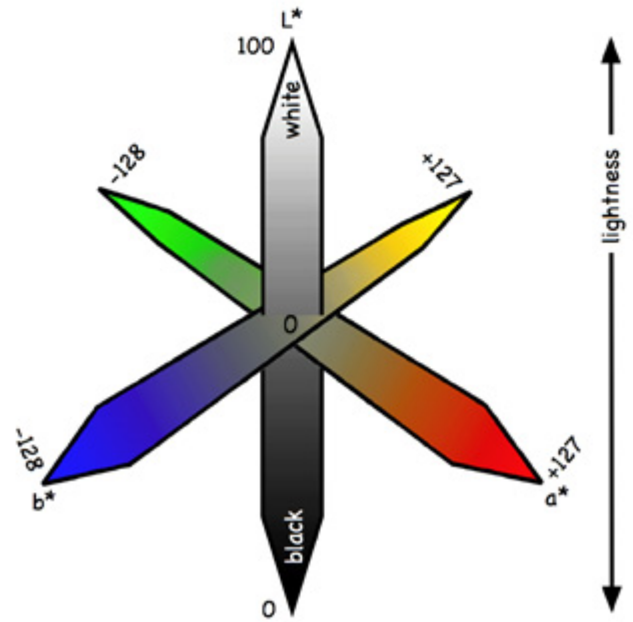
Global histogram

- Represent distribution of features
  - Color, texture, depth, …
- For example, how many red color pixels are there?

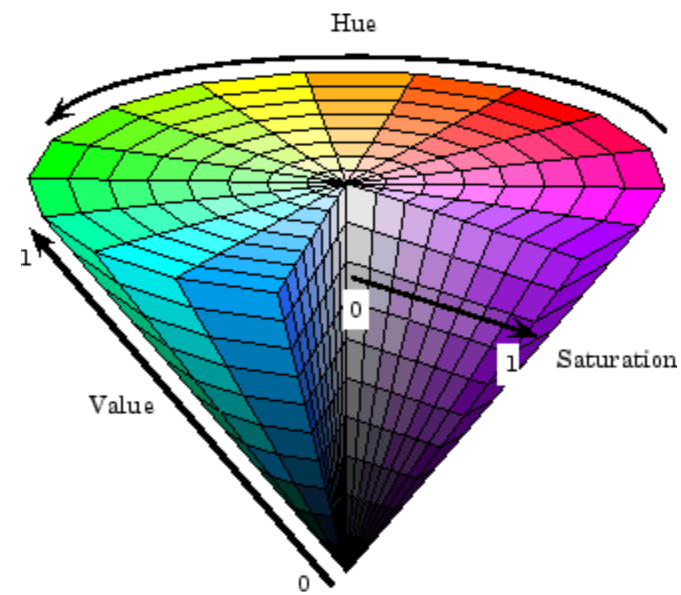
Images from Dave Kauchak

What kind of things do we compute histograms of?

- Color



L*a*b* color space

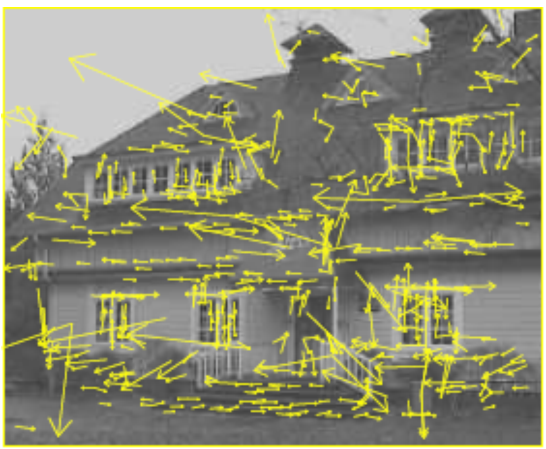


HSV color space

- Texture (filter banks or HOG over regions)

What kind of things do we compute histograms of?

- Histograms of descriptors



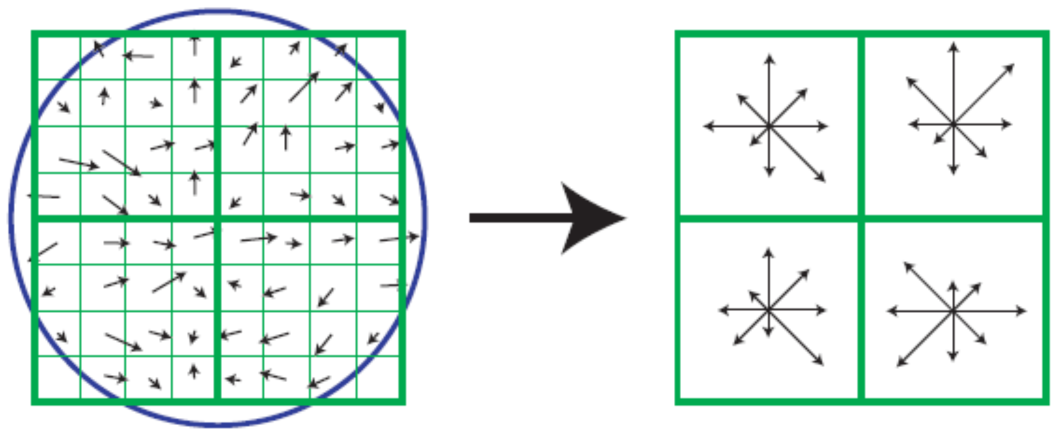


Image gradients

Keypoint descriptor

SIFT – Lowe IJCV 2004

- “Bag of words”

Histograms

Histograms

Histograms

Histogram of color will do a great job!

Histograms

Histograms

Histograms

Histogram of color will be poor.

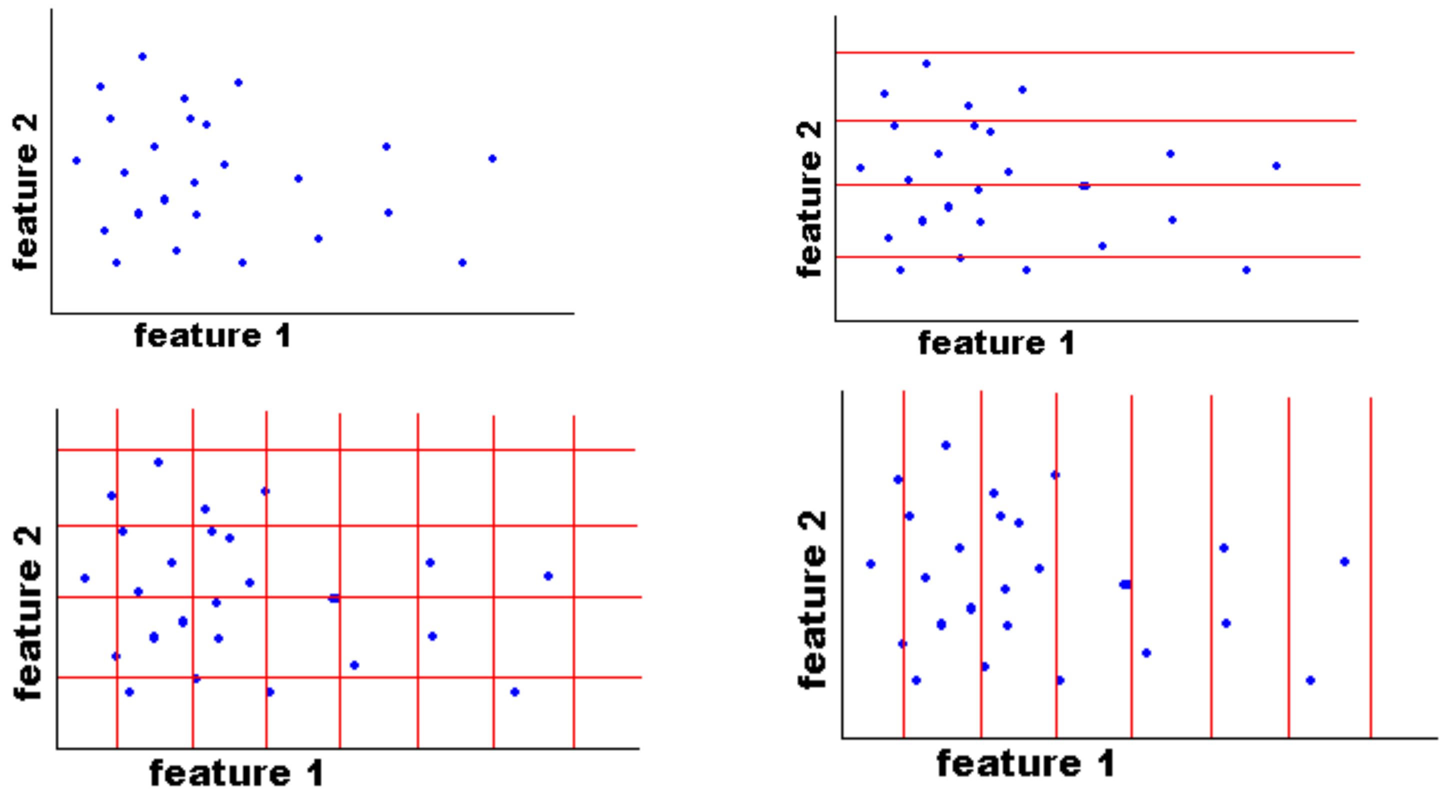
Histograms

Histogram of color will be poor.

How about histogram of orientation??

Image Representations: Histograms

Histogram: Probability or count of data in each bin



Images from Dave Kauchak

Histograms: Implementation issues

- Quantization
  - Grids: fast but applicable only with few dimensions
  - Clustering: slower but can quantize data in higher dimensions

Few Bins
Need less data
Coarser representation

Many Bins
Need more data
Finer representation

- Matching
  - Histogram intersection or Euclidean may be faster
  - Chi-squared often works better
  - Earth mover's distance is good for when nearby bins represent similar values

Histogram of Oriented Gradient

- One of the most successful features for object recognition and detection

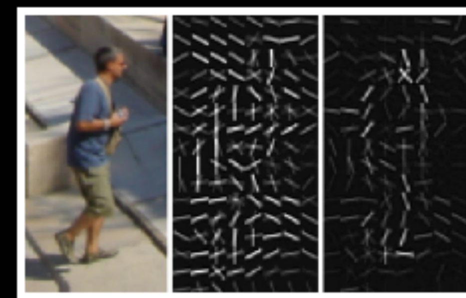
- Dalal, N. and Triggs, B. “Histograms of oriented gradients for human detection”

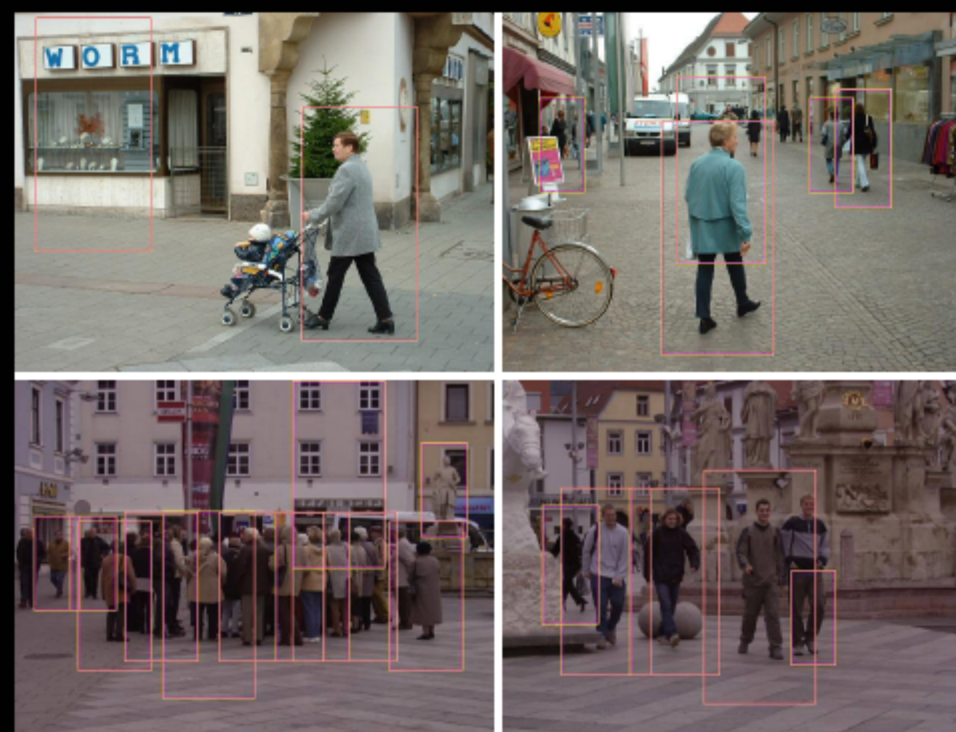
2d Global Detector

Dalal and Triggs, CVPR 2005

- 3-D Histogram of Oriented Gradients (HOG) as descriptors
- Linear SVM for runtime efficiency
- Tolerates different poses, clothing, lighting and background
- Currently works for fully visible upright persons

Importance weight responses





Slides from Sminchisescu

Dynamic Pedestrian Detection

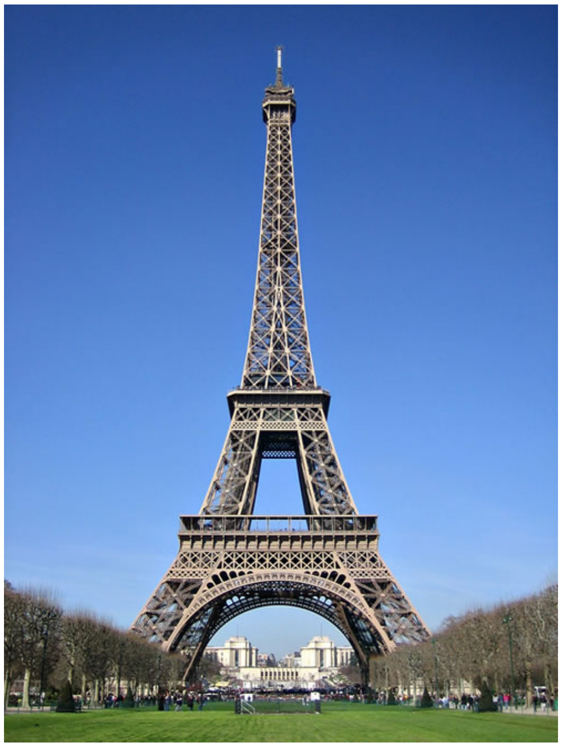
Viola, Jones and Snow, ICCV 2003

- Train using AdaBoost, about 45,000 possible features
- Efficient and reliable for distant detections (20x15), 4fps

Slides from Sminchisescu

Input Image → Compute gradients → Weighted vote into orientation cells → Normalization over spatial blocks → HOG Feature

Input Image → **Compute gradients** → Weighted vote into orientation cells → Normalization over spatial blocks → HOG Feature



+1	0	-1
+2	0	-2
+1	0	-1

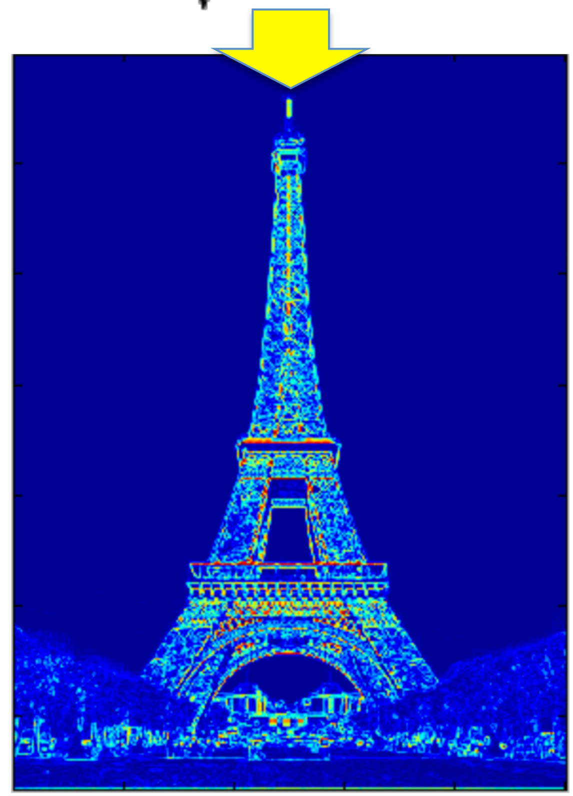
Sobel x

+1	+2	+1
0	0	0
-1	-2	-1

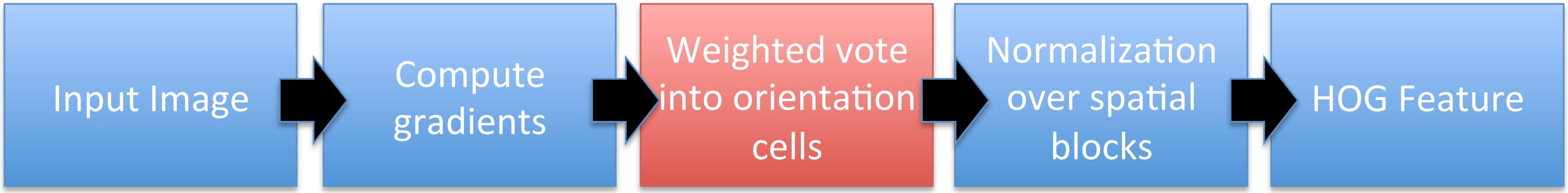
Sobel y

Apply filters to obtain the gradient

$$\mathbf{G} = \sqrt{{\mathbf{G}_x}^2 + {\mathbf{G}_y}^2}$$

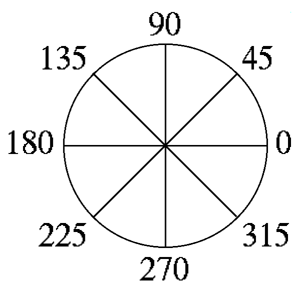


$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

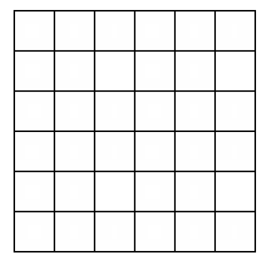


- Histogram of gradient orientations
  - Orientation

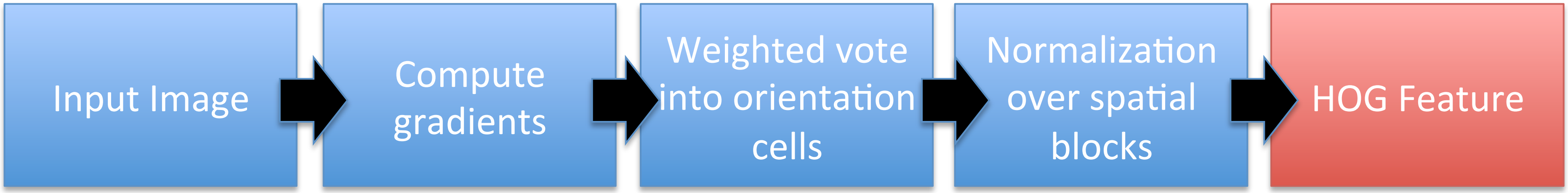
    $$\mathbf{\Theta} = \mathrm{atan2}\left(\mathbf{G}_y, \mathbf{G}_x\right)$$

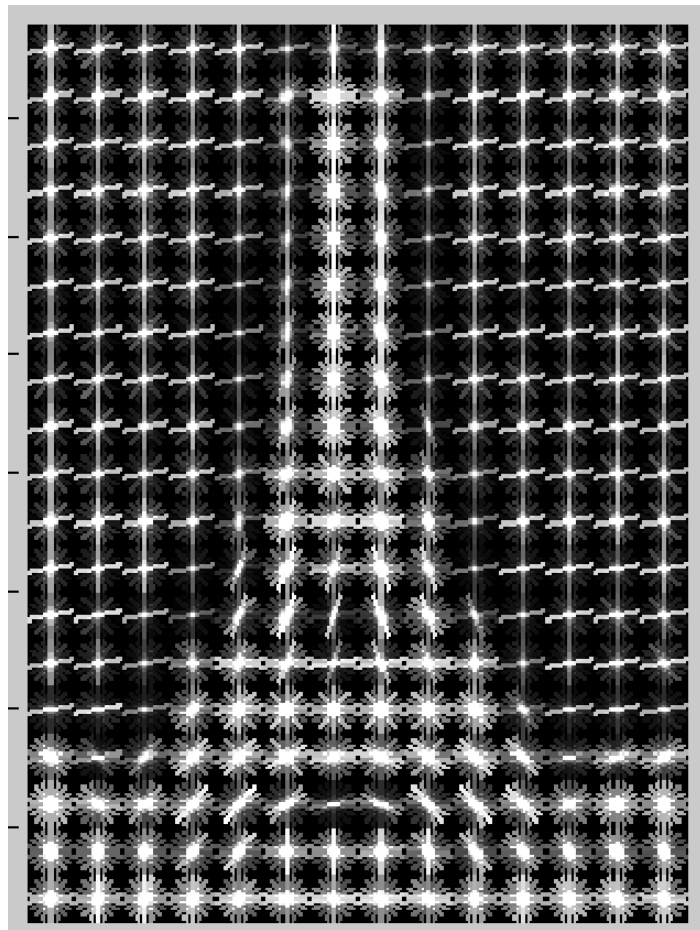
    

  - Position

  – Weighted by magnitude $\mathbf{G} = \sqrt{{\mathbf{G}_x}^2 + {\mathbf{G}_y}^2}$


Input Image
Compute gradients
Weighted vote into orientation cells
Normalization over spatial blocks
HOG Feature



Choosing a descriptor

- Again, need not stick to one
- For object recognition or stitching, HOG or variant (e.g. SIFT) showed a promising result.

Exercise: Simplified HOG

- Exercise 1: Compute gradient orientation and magnitude
- Exercise 2: Compute global histogram of gradient
- Exercise 3: Compute a distance between histograms