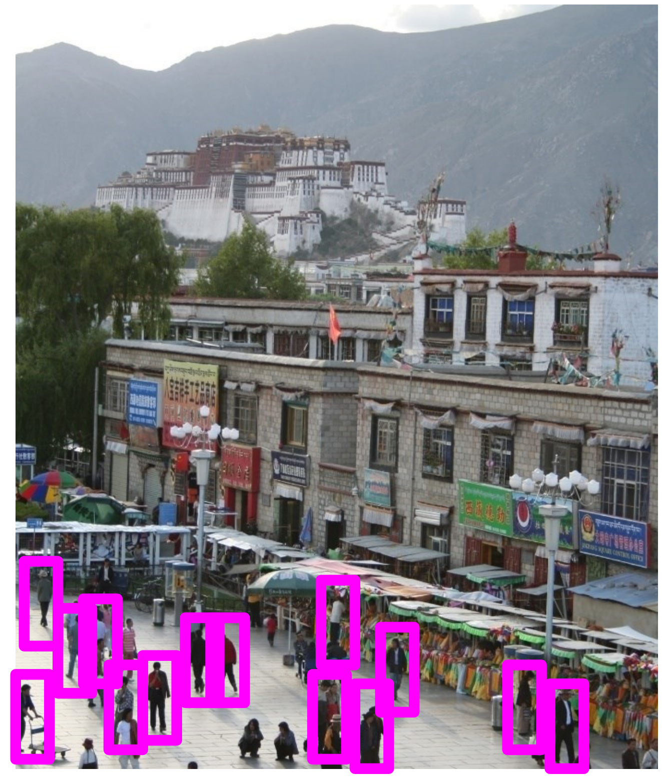
Advanced Computer Vision

Unsupervised Learning
(A Review)

Recognition

Often needs machine learning
for compact descriptions of the visual world.

Scene recognition
- City/forest/factory/…



Find pedestrians

MACHINE LEARNING CRASH COURSE

未来媒体研究中心
CENTER FOR FUTURE MEDIA

电子科技大学
University of Electronic Science and Technology of China

Our approach

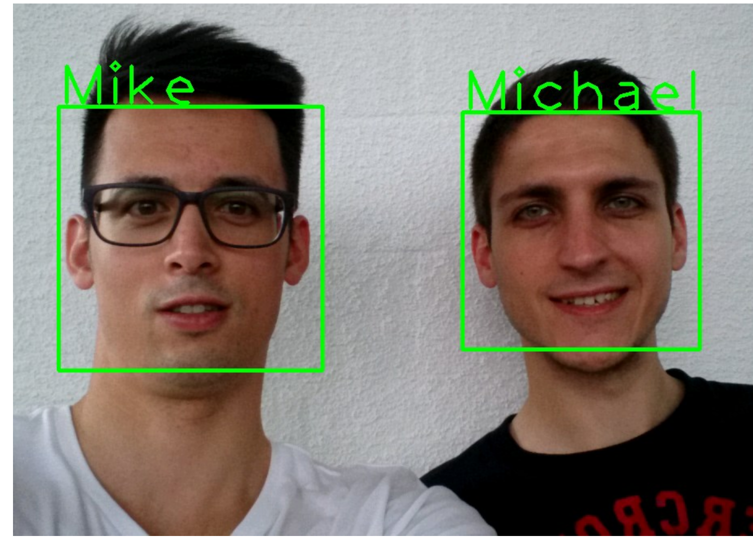
- We will look at ML as a tool. We will not detail the underpinnings of each learning method.
- Please take a machine learning course if you want to know more!

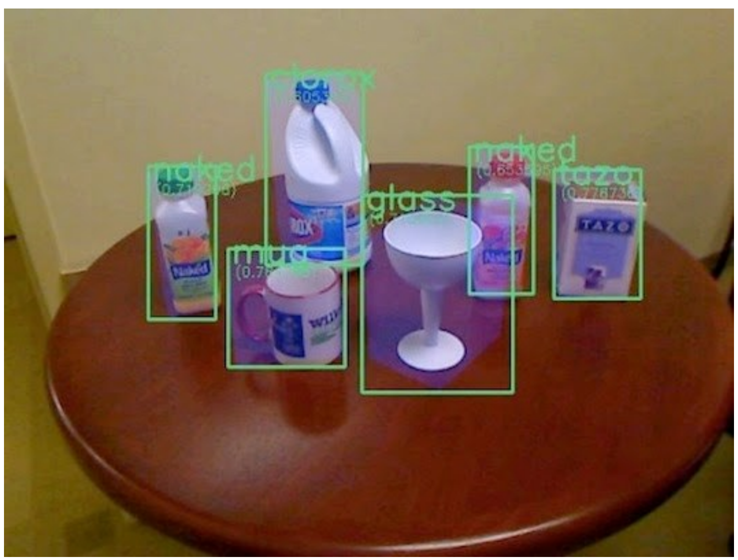
Machine Learning

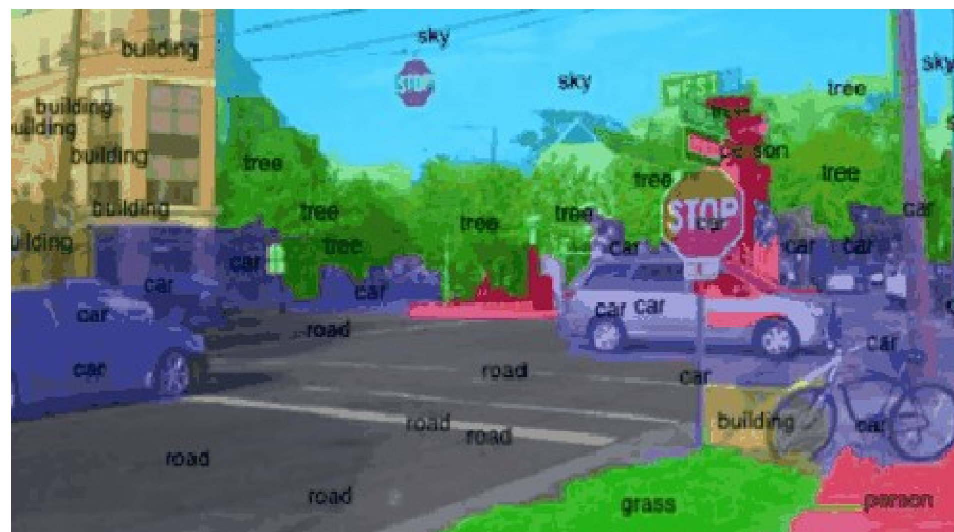
- Learn from and make predictions on data.
- Arguably the greatest export from computing to other scientific fields.
- We will look at ML as a tool. We will not detail the underpinnings of each learning method.
- Please take a machine learning course if you want to know more!

ML for Computer Vision

- Face Recognition
- Object Classification
- Scene Segmentation





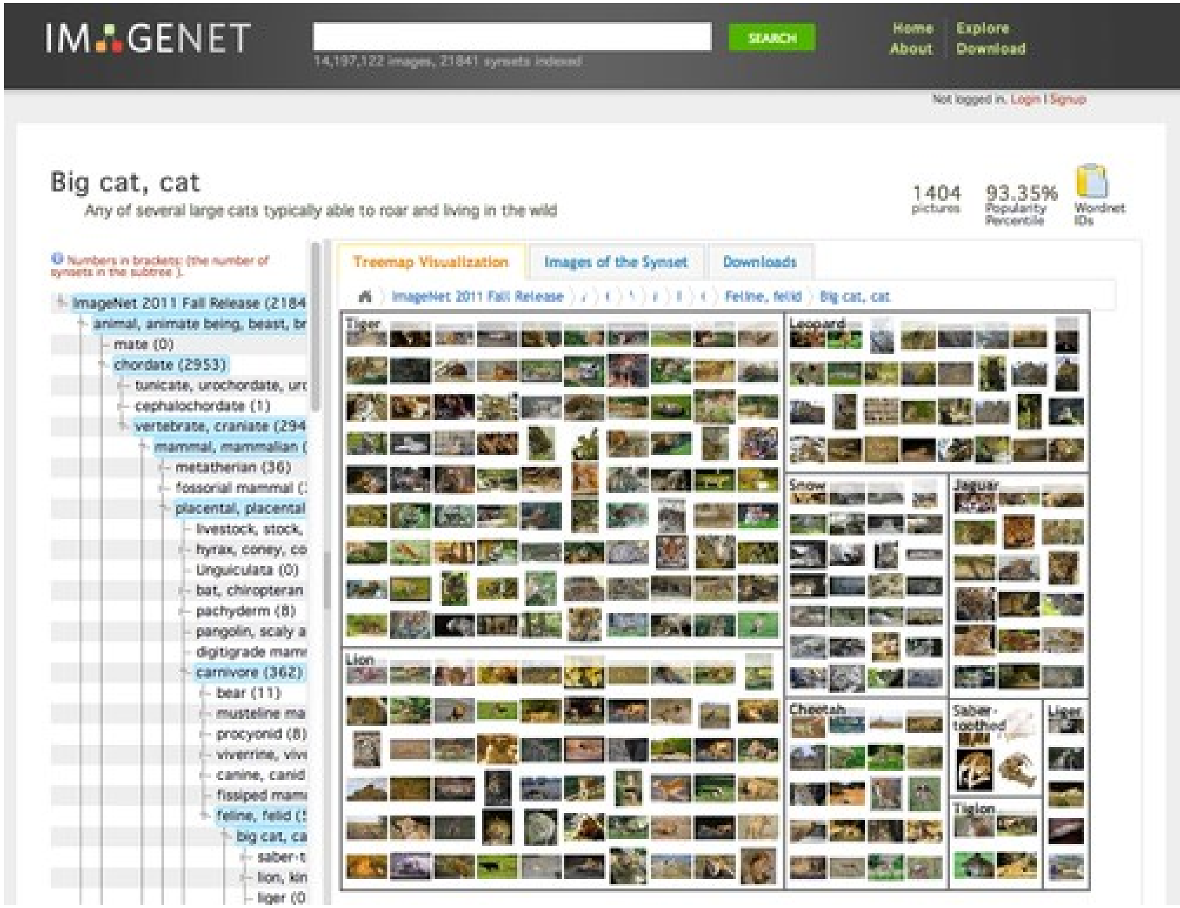


Data, data, data!

- Norvig – “The Unreasonable Effectiveness of Data” (IEEE Intelligent Systems, 2009)
  - “... invariably, simple models and a lot of data trump more elaborate models based on less data”

ImageNet

- Images for each category of WordNet
- 1000 classes
- 1.2mil images
- 100k test



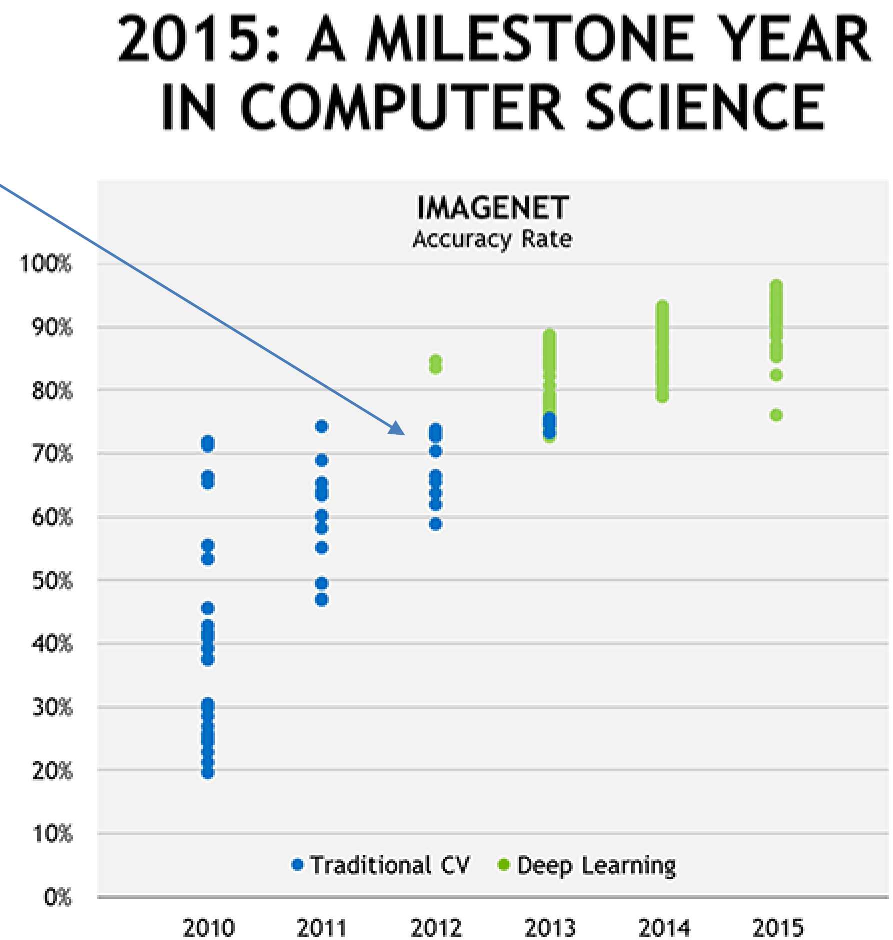
- Top 5 error

IMAGENET


mite
container ship
motor scooter
leopard
mite
black widow
cockroach
tick
starfish
container ship
lifeboat
amphibian
fireboat
drilling platform
motor scooter
go-kart
moped
bumper car
golfcart
leopard
jaguar
cheetah
snow leopard
Egyptian cat
grille
mushroom
cherry
Madagascar cat
convertible
grille
pickup
beach wagon
fire engine
agaric
mushroom
jelly fungus
gill fungus
dead-man's-fingers
dalmatian
grape
elderberry
ffordshire bullterrier
currant
squirrel monkey
spider monkey
titi
indri
howler monkey

ImageNet Competition

- Krizhevsky, 2012
- Google, Microsoft 2015
  - Beat the best human score in the ImageNet challenge.



Machine Learning Problems

	Supervised Learning	*Unsupervised Learning*
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

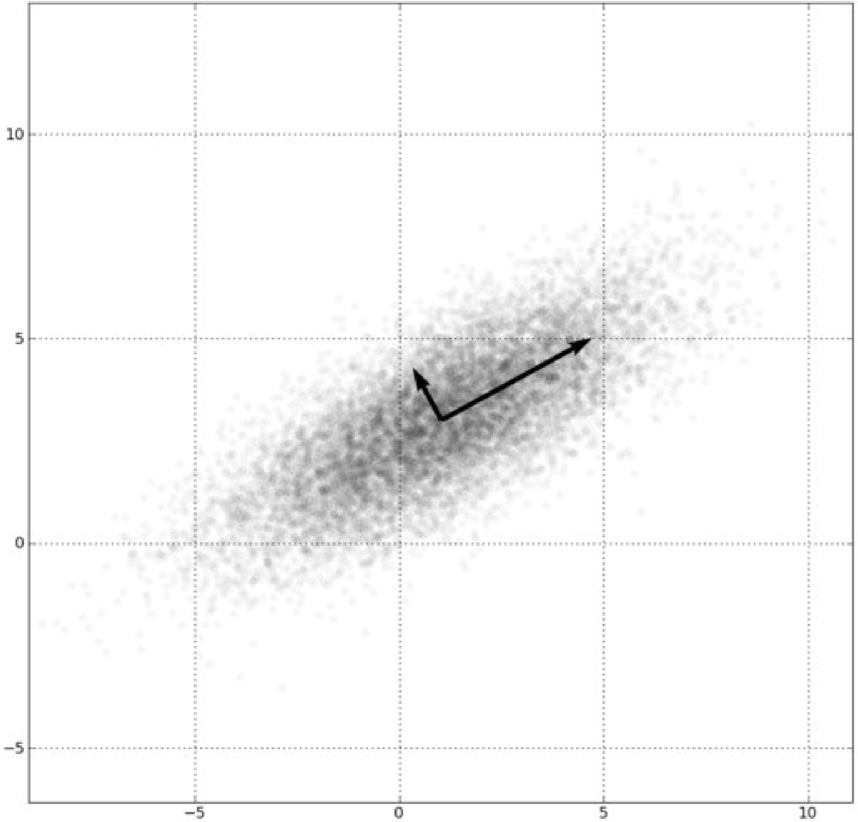
Machine Learning Problems

	Supervised Learning	Unsupervised Learning
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

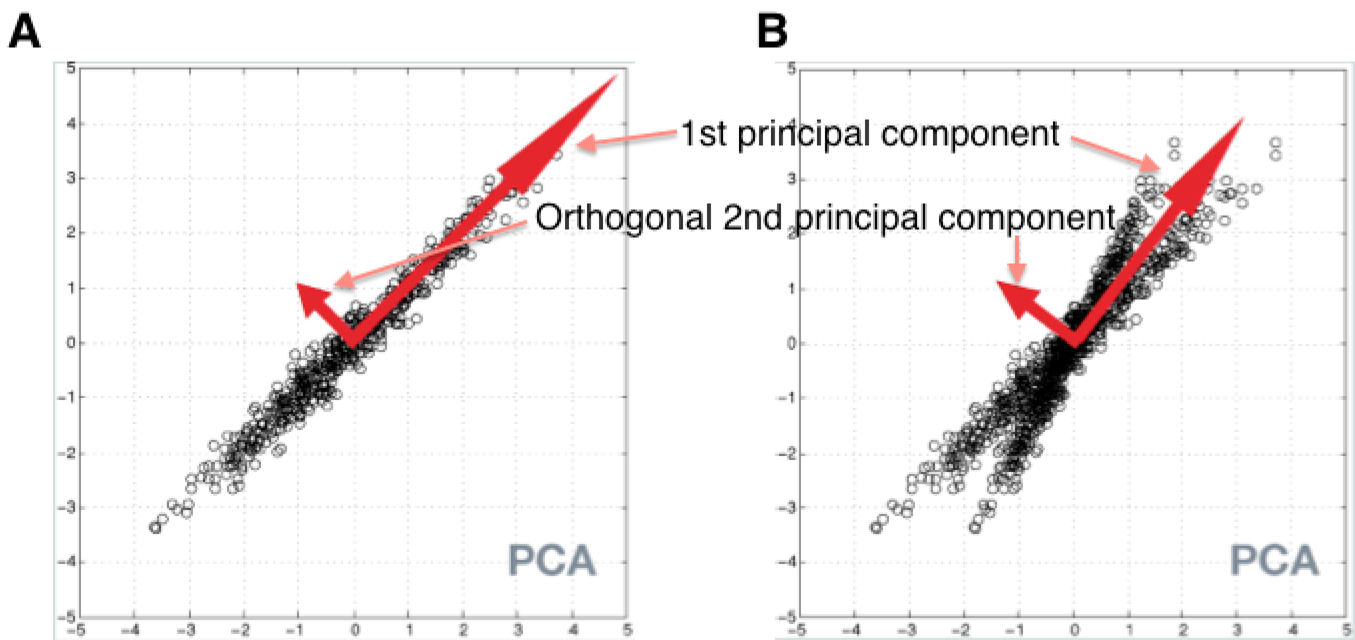
Dimensionality Reduction

- **PCA**, ICA, LLE, Isomap

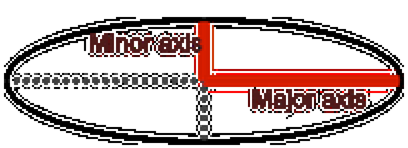
- Principal component analysis
  - Creates a basis where the axes represent the dimensions of variance, from high to low.
  - Finds correlations in data dimensions to produce *best possible* lower-dimensional representation based on linear projections.



PCA



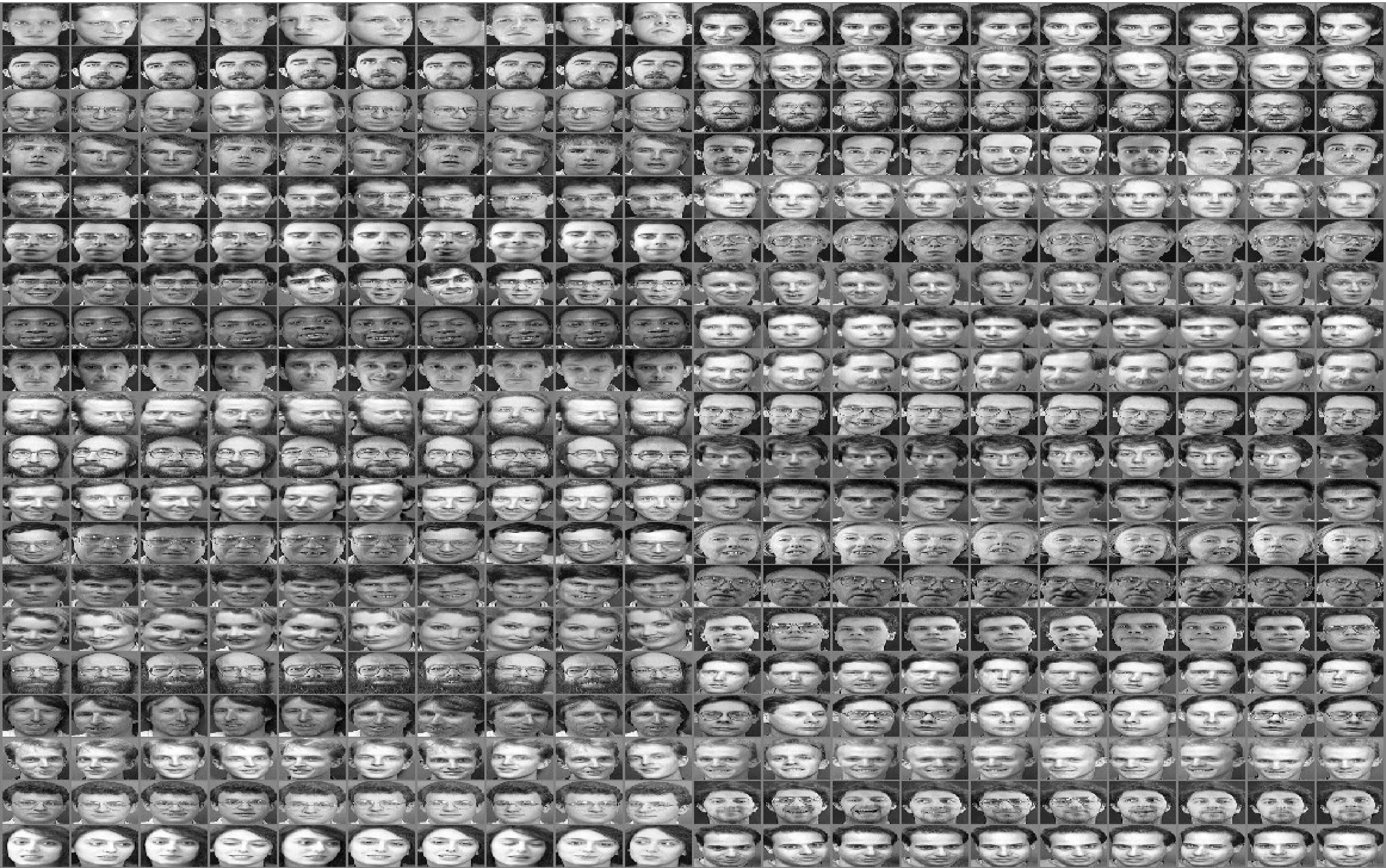
(Figure adapted from C. Beckmann, Oxford FMRIB)



未来媒体研究中心 CENTER FOR FUTURE MEDIA | 电子科技大学 University of Electronic Science and Technology of China

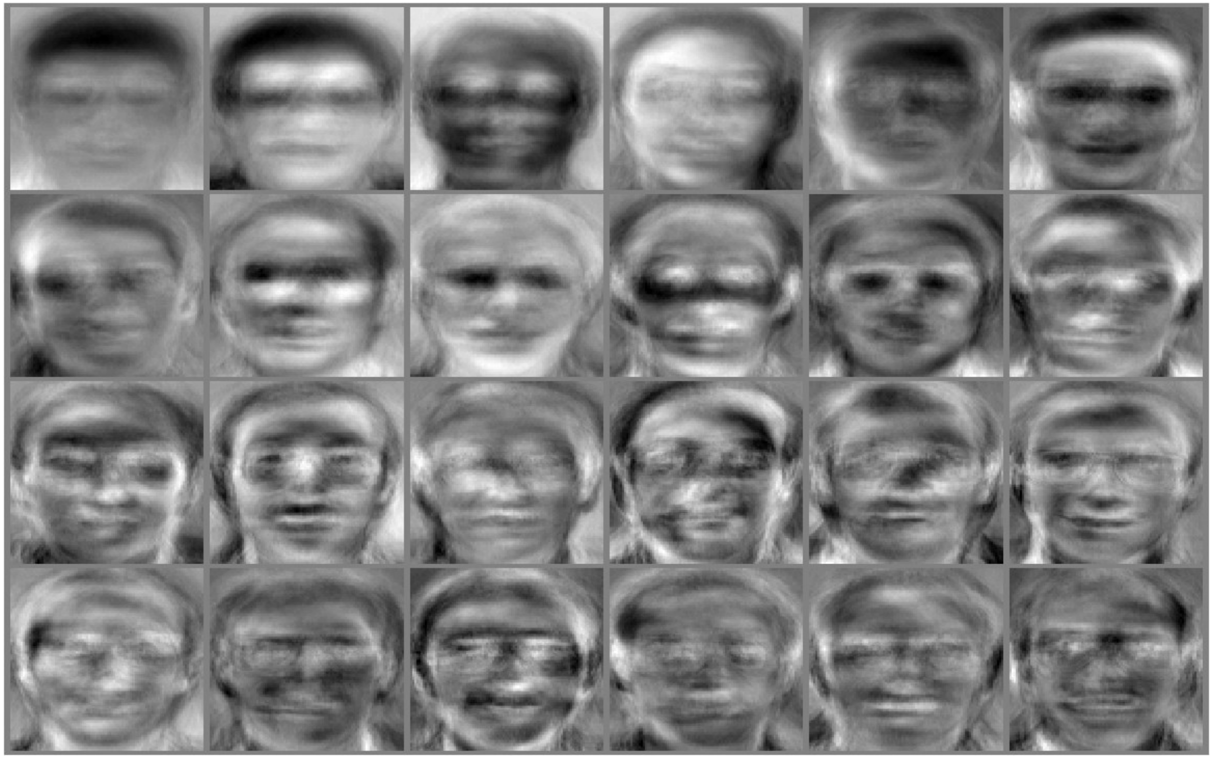
Eigenfaces

The ATT face database (formerly the ORL database), 10 pictures of 40 subjects each



Eigenfaces

Mean face



Basis of variance (eigenvectors)

M. Turk; A. Pentland (1991). "Face recognition using eigenfaces" (PDF).
Proc. IEEE Conference on Computer Vision and Pattern Recognition. pp. 586–591.

Machine Learning Problems

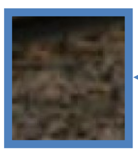
	Supervised Learning	*Unsupervised Learning*
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

Clustering example: image segmentation

Goal: Break up the image into meaningful or perceptually similar regions

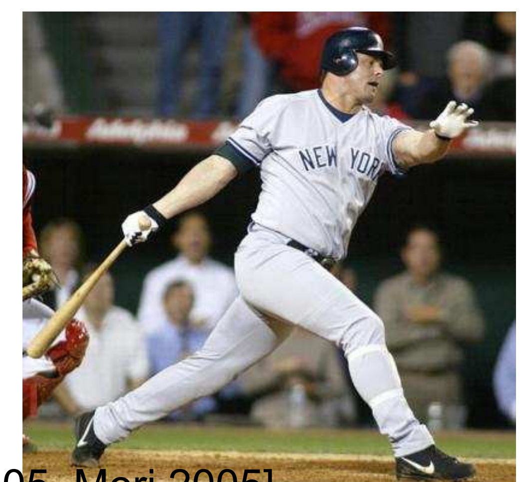
Segmentation for feature support or efficiency

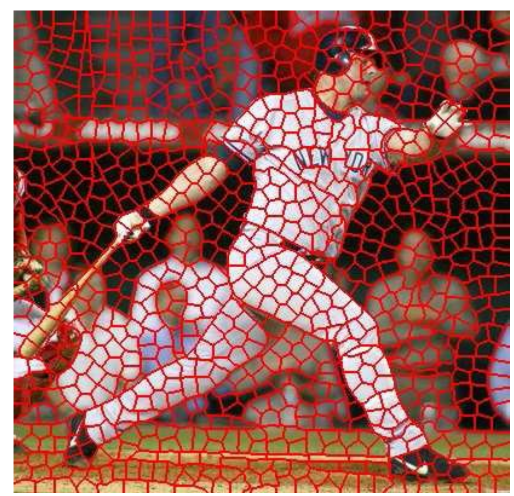
50x50
Patch



50x50
Patch

[Felzenszwalb and Huttenlocher 2004]





Superpixels!

[Hoiem et al. 2005, Mori 2005]

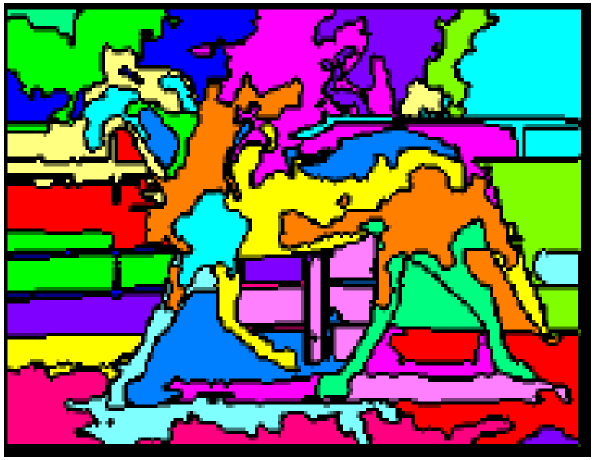
[Shi and Malik 2001]

Segmentation as a result

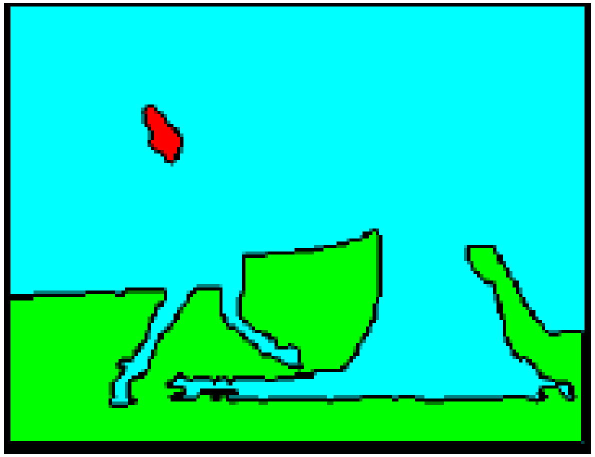
未来媒体研究中心 CENTER FOR FUTURE MEDIA | 电子科技大学 University of Electronic Science and Technology of China

GrabCut, Rother et al. 2004

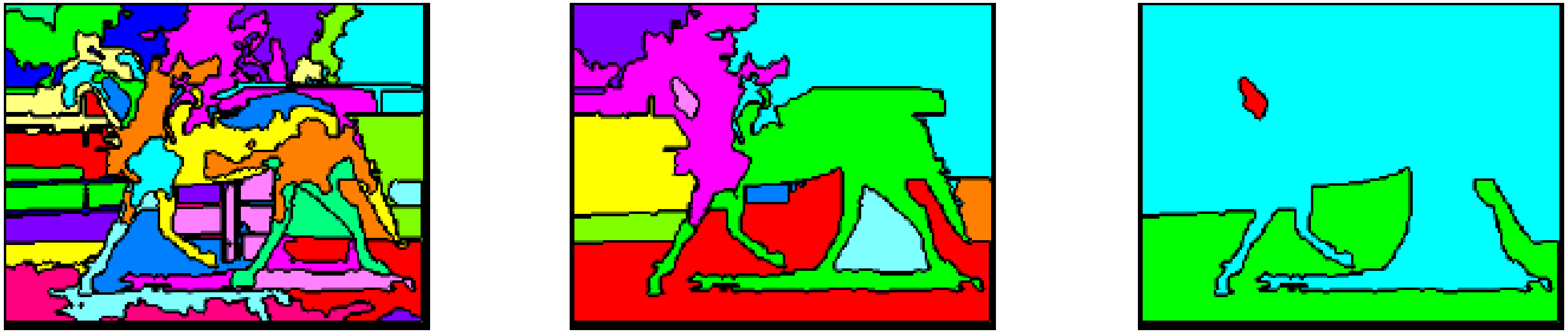
Types of segmentations



Oversegmentation



Undersegmentation



Hierarchical Segmentations

Clustering

Group together similar 'points' and represent them with a single token.

Key Challenges:

1) What makes two points/images/patches similar?

2) How do we compute an overall grouping from pairwise similarities?

Why do we cluster?

- **Summarizing data**
  - Look at large amounts of data
  - Patch-based compression or denoising
  - Represent a large continuous vector with the cluster number
- **Counting**
  - Histograms of texture, color, SIFT vectors
- **Segmentation**
  - Separate the image into different regions
- **Prediction**
  - Images in the same cluster may have the same labels

K-means algorithm

1. Randomly select K centers

2. Assign each point to nearest center

3. Compute new center (mean) for each cluster

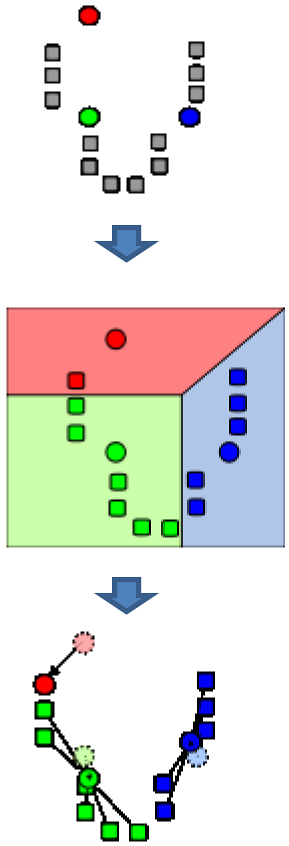


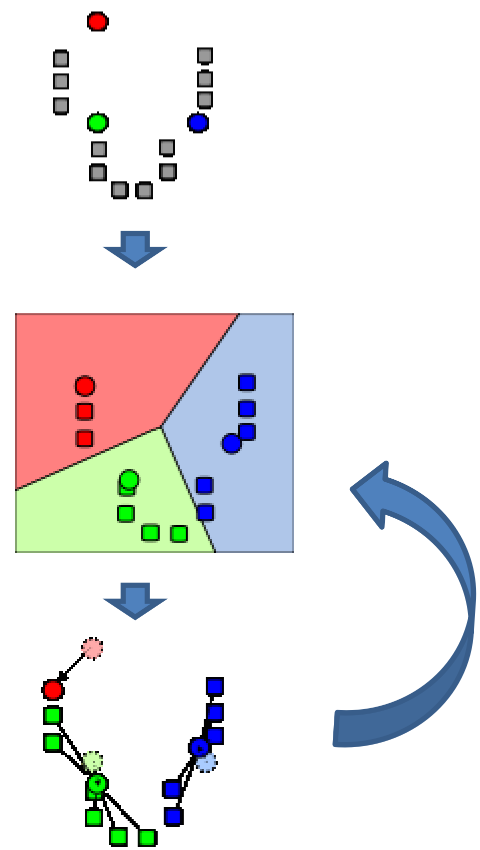
Illustration: http://en.wikipedia.org/wiki/K-means_clustering

K-means algorithm

1. Randomly select K centers

2. Assign each point to nearest center

3. Compute new center (mean) for each cluster



Back to 2

Illustration: http://en.wikipedia.org/wiki/K-means_clustering

K-means

1. Initialize cluster centers: $\mathbf{c}^0$; t=0

2. Assign each point to the closest center

$$\boldsymbol{\delta}^t = \underset{\boldsymbol{\delta}}{\operatorname{argmin}} \tfrac{1}{N} \sum_j^N \sum_i^K \delta_{ij} \left(\mathbf{c}_i^{t-1} - \mathbf{x}_j\right)^2$$

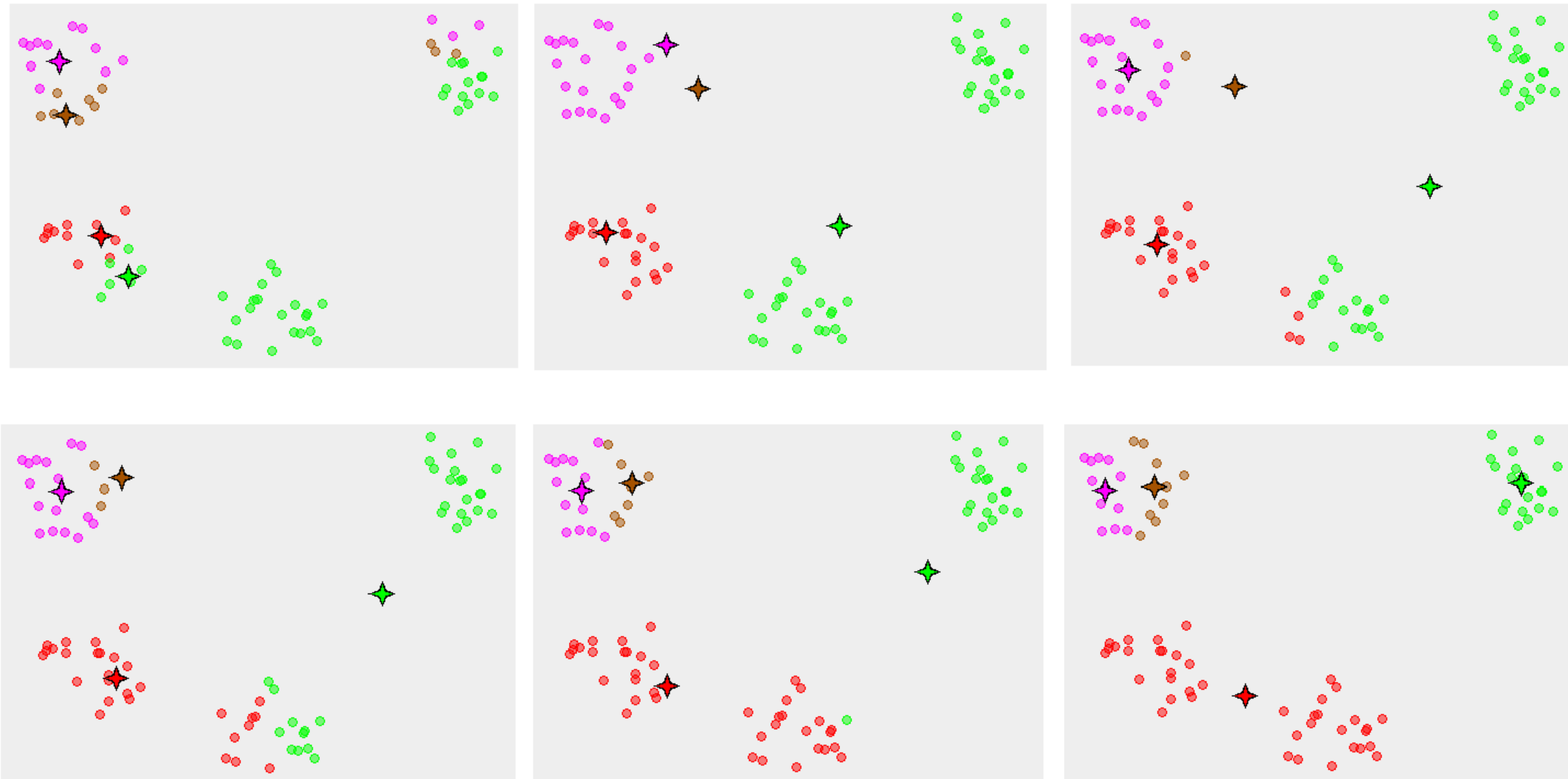
3. Update cluster centers as the mean of the points

$$\mathbf{c}^t = \underset{\mathbf{c}}{\operatorname{argmin}} \tfrac{1}{N} \sum_j^N \sum_i^K \delta_{ij}^t \left(\mathbf{c}_i - \mathbf{x}_j\right)^2$$

4. Repeat 2-3 until no points are re-assigned (t=t+1)

Slide: Derek Hoiem

K-means convergence



Think-Pair-Share

- What is good about k-means?
- What is bad about k-means?
- Where could you apply k-means?

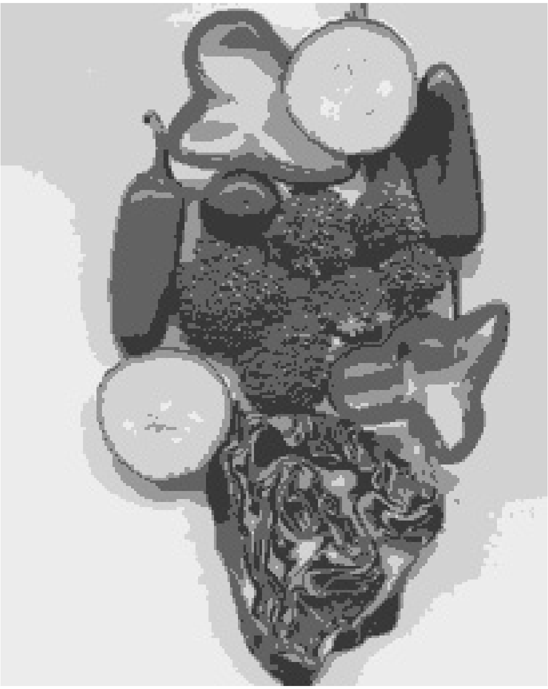
K-means: design choices

- Initialization
  - Randomly select K points as initial cluster center
  - Or greedily choose K points to minimize residual
- Distance measures
  - Traditionally Euclidean, could be others
- Optimization
  - Will converge to a *local minimum*
  - May want to perform multiple restarts

K-means clustering using intensity or color

Image

Clusters on intensity



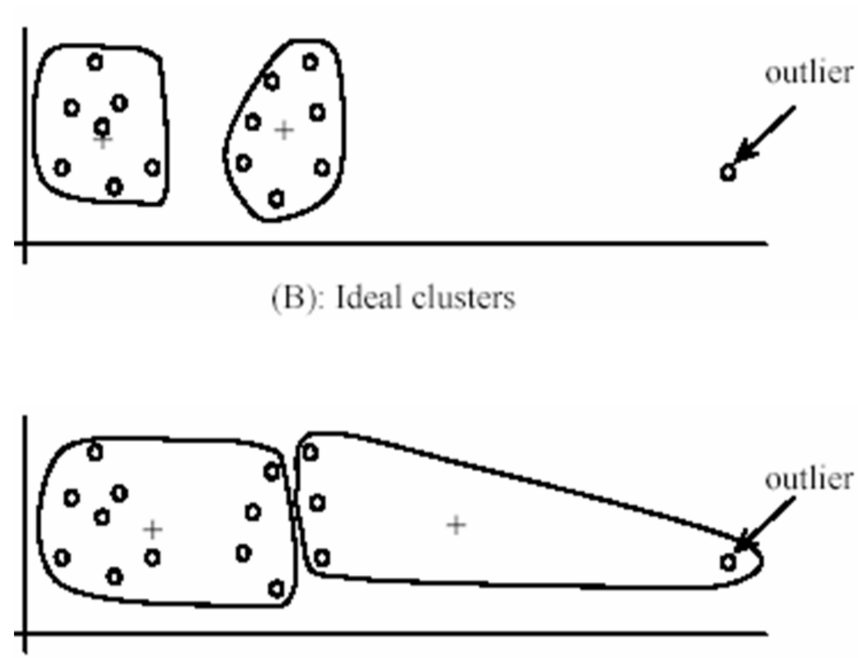
Clusters on color

How to choose the number of clusters?

- Validation set
  - Try different numbers of clusters and look at performance
    - When building dictionaries (discussed later), more clusters typically work better.

K-Means pros and cons

- Pros
  - Finds cluster centers that minimize conditional variance (good representation of data)
  - Simple and fast*
  - Easy to implement
- Cons
  - Need to choose K
  - Sensitive to outliers
  - Prone to local minima
  - All clusters have the same parameters (e.g., distance measure is non-adaptive)
  - *Can be slow: each iteration is O(KNd) for N d-dimensional points

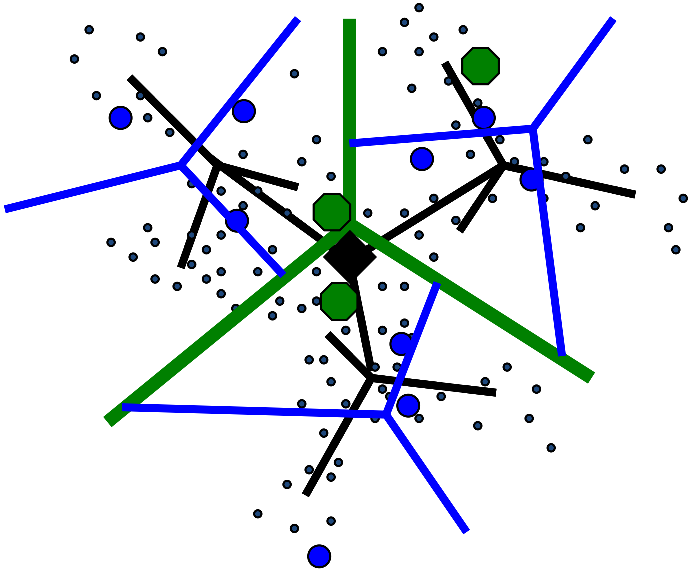


- Usage
  - Cluster features to build visual dictionaries

Building Visual Dictionaries

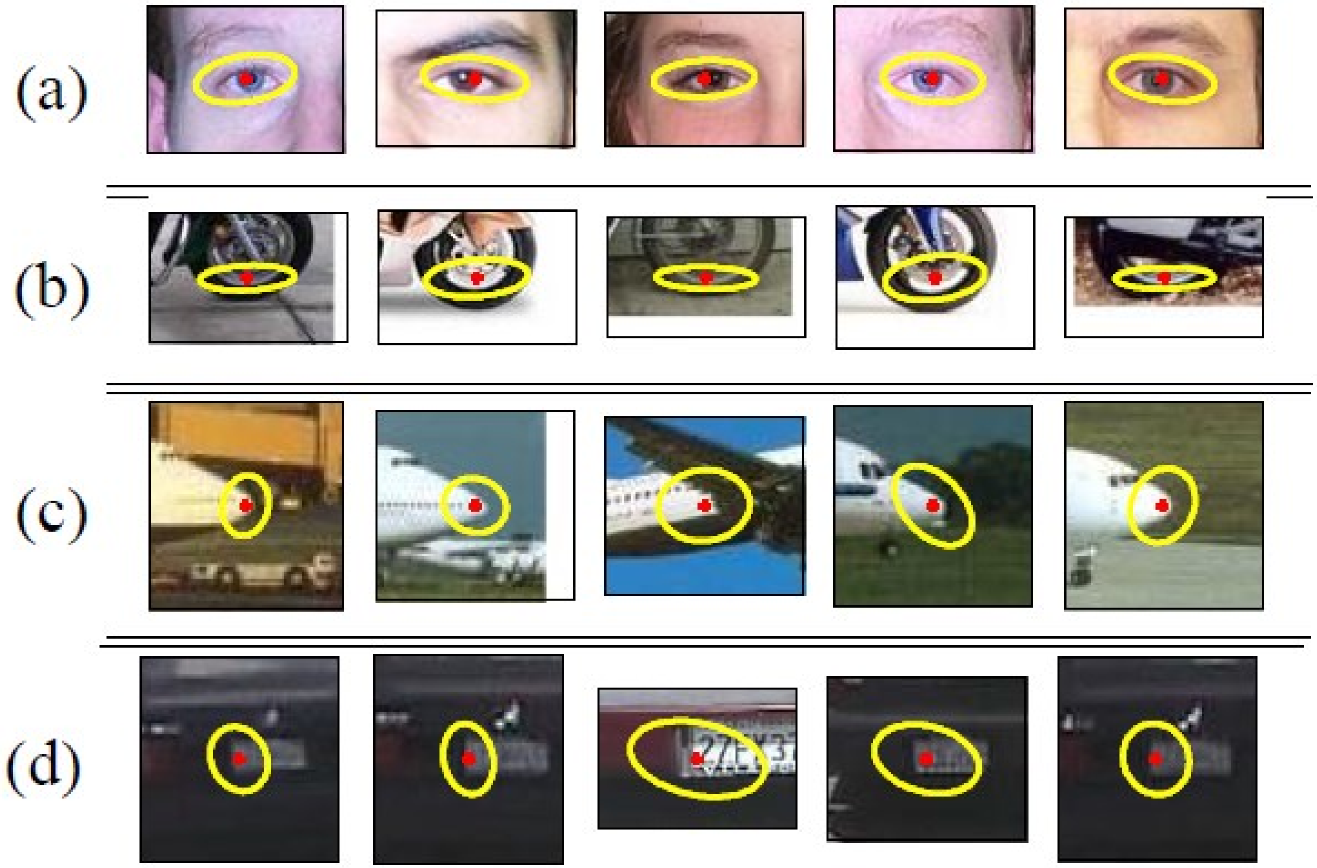
1. Sample features from a database
   - E.g., 128 dimensional SIFT vectors

2. Cluster to build dictionary
   - Cluster centers are the dictionary words



3. To match new features, assign to the nearest cluster to save rebuilding dictionary

Examples of learned codewords



Most likely codewords for 4 learned "topics"
EM with multinomial (problem 3) to get topics

http://www.robots.ox.ac.uk/~vgg/publications/papers/sivic05b.pdf

Sivic et al. ICCV 2005

Machine Learning Problems

	Supervised Learning	*Unsupervised Learning*
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction