CS543 - Computer Vision

D.A. Forsyth
Administrative matters

- Instructor: D.A. Forsyth, 3310 Siebel Hall, daf@uiuc.edu
- Office hours: 14h30-15h30 Tue, Thur, or walk in
- TA: Alex Sorokin, syrnick@gmail.com
- Evaluation: Homeworks, projects
Changes

• More homework
• More structured projects
  • list available shortly; choose teams, project
  • option of change later
• New book chapters, slides
Topics

- Overview, Cameras, Shading
- Color, Texture
- Edges, Corners, Segmentation
- Structure and Motion
- Features, Matching and Recognition
- People (finding, tracking, understanding)
Example problems

- **Obstacle avoidance**
  - A cricketer avoids being hit in the head (->) (<-)
  - the gannet pulls its wings in in time, by measuring time to contact

- **Reconstructing representations of the 3D world**
  - from multiple views
  - from shading
  - from structural models, etc

- **Recognition**
  - draw distinctions between what is seen
    - is it soggy?
    - will it eat me?
    - can I eat it?
    - is it a cat?
    - is it my cat?
3D Reconstruction

• Cues:
  • variation in appearance in multiple views
    • stereo
    • motion
  • (possibly) shading
  • (possibly) contour
  • texture
  • rich body of geometric theory

• Issues:
  • correspondence
  • noise
  • the nature of the representation
Structure from motion
Work by Paul Debevec and Jitendra Malik
Mosaics

Figure from “Quicktime VR - An Image-Based Approach to Virtual Environment Navigation”
S.E. Chen, SIGGRAPH 95
Correspondence

- Local representation of image properties make things easier
  - identify points which are easily localised
    - corners
    - which lie on edges
  - compare with points in next image
    - points which “look similar” may well match
  - search radius is constrained by geometry
    - in ways we will not discuss
Local Representations

- What do edge responses look like nearby?
  - SIFT features

- What is the “general pattern” of grey levels?
  - Statistics of filters
Lowe’s SIFT features

Fig 7 from:
Distinctive image features from scale-invariant keypoints
Texture

• Describe by statistics of filter outputs
• Issues
  • describing textures
  • for region segmentation
  • inferring shape from texture
  • synthesizing textures
Shape from texture

- The foreshortening of textures gives us a cue to shape
  - But how?
    - Details remain rather murky
Some applications of recognition

- Digital libraries
  - Find me the pic of JFK and Marilyn Monroe embracing
  - NCMEC
- Surveillance
  - Warn me if there is a mugging in the grove
- HCI
  - Do what I show you
- Military
  - Shoot this, not that
What are the problems in recognition?

- Which bits of image should be recognised together?
  - Segmentation.
- How can objects be recognised without focusing on detail?
  - Abstraction.
- How can objects with many free parameters be recognised?
  - No popular name, but it’s a crucial problem anyhow.
- How do we structure very large modelbases?
  - again, no popular name; abstraction and learning come into this
Segmentation

- Which image components “belong together”?
- Belong together = lie on the same object
- Cues
  - similar colour
  - similar texture
  - not separated by contour
  - form a suggestive shape when assembled
Interactive segmentation
Segmentation

Efficient Graph-Based Image Segmentation
Pedro F. Felzenszwalb and Daniel P. Huttenlocher
International Journal of Computer Vision, Volume 59, Number 2, September 2004
Matching templates

- Some objects are 2D patterns
  - e.g. faces
- Build an explicit pattern matcher
  - discount changes in illumination by using a parametric model
  - changes in background are hard
  - changes in pose are hard
http://www.ri.cmu.edu/projects/project_320.html
Matching by searching relations

After Felzenszwalb and Huttenlocher, 03
Constellations of parts

Fischler & Elschlager 1973

Yuille ‘91
Brunelli & Poggio ‘93
Lades, v.d. Malsburg et al. ‘93
Cootes, Lanitis, Taylor et al. ‘95
Amit & Geman ‘95, ‘99
Agarwal & Roth ‘02
Typical models

Motorbikes

Spotted cats

Figure after Fergus et al, 03; see also Fergus et al, 04
Words and pictures

Annotation results in complementary words and pictures

Query on

“Rose”

and

Example from Berkeley Blobworld system
Train a system of svm classifiers, one per word but penalize that matrix for rank, after Rennie+Srebro 05

The latent space reveals scenes because it is good at word prediction and takes appearance into account

Loeff Farhadi Forsyth??
It was there and we didn’t

It was there and we predicted it

It wasn’t and we did

Loeff Farhadi Forsyth??
Lexicon building

- In its simplest form, missing variable problem
- Pile in with EM
  - given correspondences, conditional probability table is easy (count)
  - given cpt, expected correspondences could be easy
- Caveats
  - might take a lot of data; symmetries, biases in data create issues

"the beautiful sun"

"le soleil beau"

"sun   sea   sky"

Brown, Della Pietra, Della Pietra & Mercer 93; Melamed 01
Attributes

• Properties shared by many object categories
• Material (like)
  • glass, wood, furry, red, etc.
• Part (like)
  • has wheel, has head, has tail, etc.
• Shape (like)
  • is 2D Boxy, is cylindrical, etc.
How is an object different from typical?

- Pragmatics suggests this is how adjectives are chosen
  - If we are sure it’s a cat, and we know that
    - an attribute is different from normal
    - the detector is usually reliable
  - we should report the missing/extra attribute
Missing attributes

Aeroplane
No “wing”

Car
No “window”

Boat
No “sail”

Aeroplane
No “jet engine”

Motorbike
No “side mirror”

Car
No “door”

Bicycle
No “wheel”

Sheep
No “wool”

Train
No “window”

Sofa
No “wood”

Bird
No “tail”

Bird
No “leg”

Bus
No “door”
Extra attributes

- Bird "Leaf"
- Bus "face"
- Motorbike "cloth"
- DiningTable "skin"
- People "Furn. back"
- Aeroplane "beak"
- People "label"
- Sofa "wheel"
- Bike "Horn"
- Monitor window"
News dataset

- Approx 5e5 news images, with captions
  - Easily collected by script from Yahoo over the last 18 months or so
- Mainly people
  - politicians, actors, sportsplayers
  - long, long tails distribution
- Face pictures captured “in the wild”
- Correspondence problem
  - some images have many (resp. few) faces, few (resp. many) names (cf. Srihari 95)

President George W. Bush makes a statement in the Rose Garden while Secretary of Defense Donald Rumsfeld looks on, July 23, 2003. Rumsfeld said the United States would release graphic photographs of the dead sons of Saddam Hussein to prove they were killed by American troops. Photo by Larry Downing/Reuters
Data examples

Doctor Nikola shows a fork that was removed from an Israeli woman who swallowed it while trying to catch a bug that flew into her mouth, in Poriah Hospital northern Israel July 10, 2003. Doctors performed emergency surgery and removed the fork. (Reuters)

President George W. Bush waves as he leaves the White House for a day trip to North Carolina, July 25, 2002. A White House spokesman said that Bush would be compelled to veto Senate legislation creating a new department of homeland security unless changes are made. (Kevin Lamarque/Reuters)
US President George W. Bush (L) makes remarks while Secretary of State Colin Powell (R) listens before signing the US Leadership Against HIV/AIDS, Tuberculosis and Malaria Act of 2003 at the Department of State in Washington, DC. The five-year plan is designed to help prevent and treat AIDS, especially in more than a dozen African and Caribbean nations (AFP/Luke Frazza).

German supermodel Claudia Schiffer gave birth to a baby boy by Caesarian section January 30, 2003, her spokeswoman said. The baby is the first child for both Schiffer, 32, and her husband, British film producer Matthew Vaughn, who was at her side for the birth. Schiffer is seen on the German television show 'Bet It...?!' ('Wetten Dass...?!') in Braunschweig, on January 26, 2002. (Alexandra Winkler/Reuters)

British director Sam Mendes and his partner actress Kate Winslet arrive at the London premiere of 'The Road to Perdition', September 18, 2002. The film stars Tom Hanks as a Chicago hit man who has a separate family life and co-stars Paul Newman and Jude Law. REUTERS/Dan Chung.
Tracking

- Use a model to predict next position and refine using next image
- Model:
  - simple dynamic models (second order dynamics)
  - kinematic models
  - etc.
- Face tracking and eye tracking now work rather well
- People tracking is hard, but important and do-able
Run