INF 5300 15.1.14
Introduction and a taste of the course

Lecturers:

- Are Jensen
- Anne Schistad Solberg

Contact information

- Are Jensen
  - On IFI wednesdays, room 4457
    Email: arej@ifi.uio.no
- Anne Schistad Solberg
  - Room 4458
  - anne@ifi.uio.no , 22852435

Course hightlights

- Two main parts:
  - Computer vision
  - Pattern recognition
- One mandatory exercise
  - Individual themes
  - Can be linked to your master topic if requested
  - Deadline can be fitted to your schedule if it fits with the course schedule.
- Lab exercise sessions
- Oral exam if less than 10 approx. students

Lecture plan

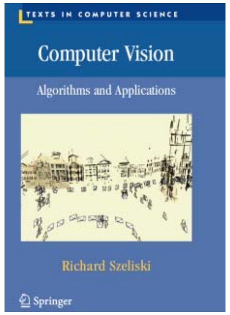
- 22.01: Introduction to regularization (Anne)
- 29.01: Active contour models (Anne)
- 05.02: Markov random fields and contextual models (Anne)
- 12.02: Lab on active contours/Markov fields (Anne)
- 26.02: Extracting good features for tracking/geometry (Anne)
- 05.03: Lab (Anne)
- 12.03: Feature selection (Are)
- 19.03: Linear feature transforms (Are)
- 26.03: Lab (Are)
- 02.04: Image registration (Anne)
- 09.04: Motion estimation (Anne)
- 30.04: Support vector machines
- 07.05: Lab

Curriculum

- Lecture notes are most important!
- Some lectures are based on:
  - ”Pattern Recognition” by S. Theodoridis and K. Koutroumbas:
  - (A paper copy of selected sections will be provided)
- Many of the lectures are based on:
  - «Computer vision: Algorithms and applications» by R. Szeliski

Computer Vision book

- See http://szeliski.org/Book/
- See webpage for ordering
- A PDF of book is also available



Every picture tells a story

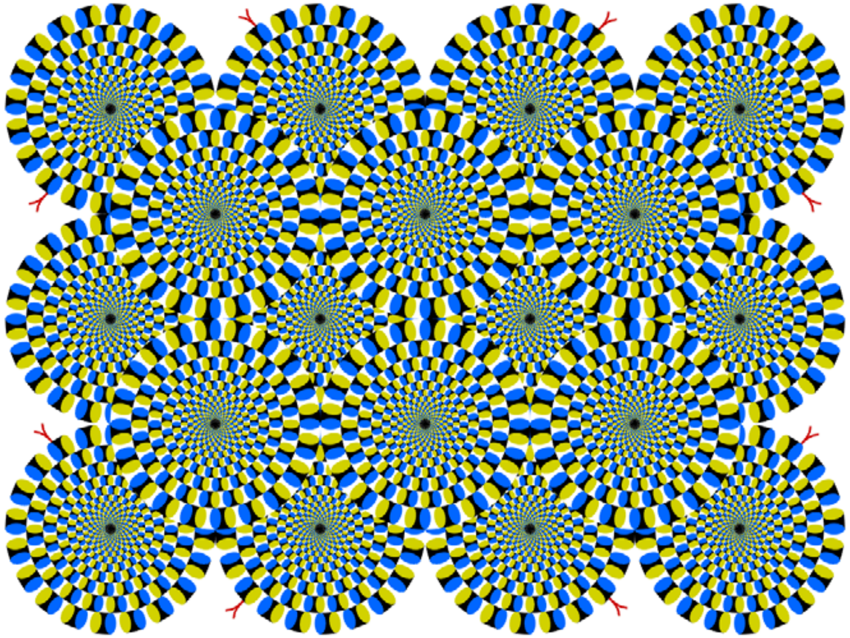
- Goal of computer vision is to write computer programs that can interpret images

Can computers match (or beat) human vision?

- Yes and no (but mostly no!)
  - humans are much better at “hard” things
  - computers can be better at “easy” things

Human perception has its shortcomings…

Sinha and Poggio, *Nature*, 1996



Copyright A.Kitaoka 2003

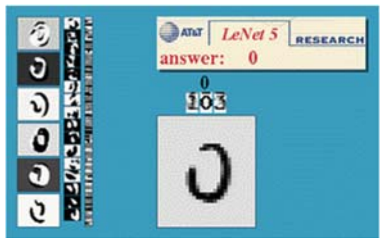
Current state of the art

- The next slides show some examples of what current vision systems can do

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



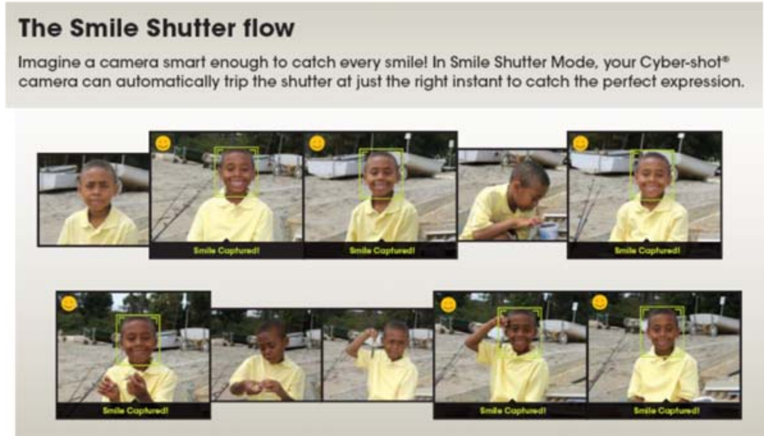
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

Smile detection?



Sony Cyber-shot® T70 Digital Still Camera

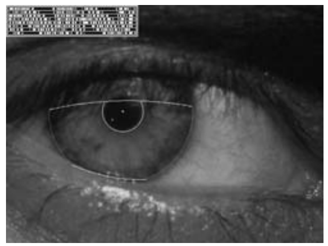
Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics

"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… "

Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story



Login without a password…

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Object recognition (in mobile phones)

- This is becoming real:
  - Lincoln Microsoft Research
  - Point & Find, Nokia

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

Slide content courtesy of Amnon Shashua

Smart cars

- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.
  - Video demo

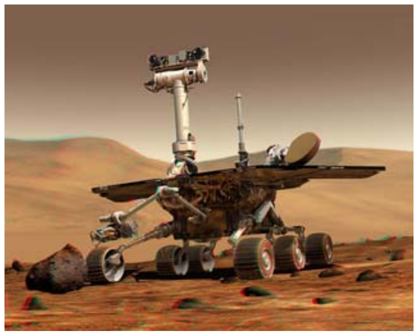
Vision in space

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

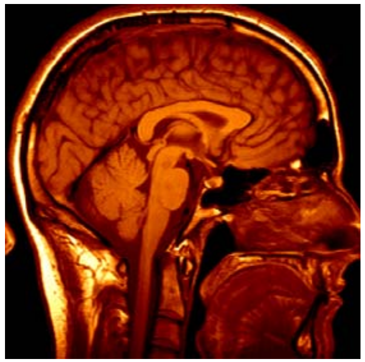
Robotics



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

http://www.robocup.org/

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Current state of the art

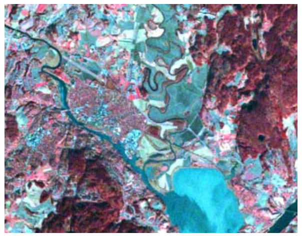
- You just saw examples of current systems.
  – Many of these are less than 5 years old
- This is a very active research area, and rapidly changing
  – Many new apps in the next 5 years
- To learn more about vision applications and companies
  – David Lowe maintains an excellent overview of vision companies
    - http://www.cs.ubc.ca/spider/lowe/vision.html

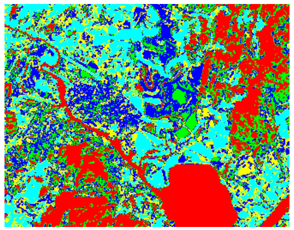
A taste of regularization

- In segmentation, we often want objects with smooth boundaries.
- Regularization is the process of contraining a an algorithmg to produce smooth objects/boundaries.
- We can regularize simple classification using contextual models
  – The class labels for a pixel will be modelled as a function of its neighbors.
- We can regularize the shape of the boundaries
  – Example: snakes
- We can regularize the estimated motion of objects in e.g. a video.

Background – contextual classification

- An image normally contains areas of similar class
  – neighboring pixels tend to be correlated.
- Classified images based on a non-contextual model often contain isolated misclassified pixels (or small regions).
- How can we get rid of this?
  – Majority filtering in a local neighborhood
  – Remove small regions by region area
  – Relaxation (Kittler and Foglein – see INF 3300 Lecture 23.09.03)
  – Bayesian models for the joint distribution of pixel labels in a neighborhood.
- How do we know if the small regions are correct or not?
  – Look at the data, integrate spatial models in the classifier.





Markov Random Fields

- Used a lot in computer vision and graphics:
- stereo matching
- image segmentation
- image blending
- texture synthesis
- image restoration



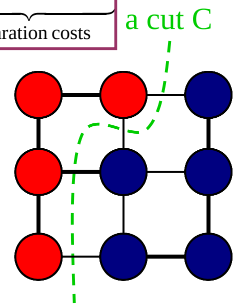


Markov Random Fields

- We want to minimize the energy function $E(f)$

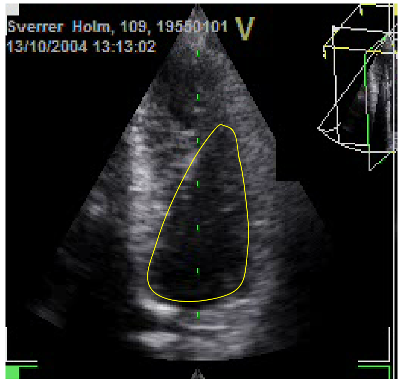
$$\arg\min_f \underbrace{\sum_p D(p, f_p)}_{\text{assignment costs}} + \underbrace{\sum_{p,q \in N} V(f_p, f_q)}_{\text{separation costs}}$$

- with spatially varying smoothness/interaction potentials $V_{pq}(f_p,f_q)$



Snakes – active contours

Example – segmenting ultrasound images of the hearth

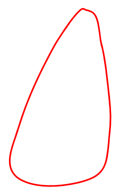


Find the border of the left ventricle

- 2D views have partly discontinuous border
- Noisy image

The initial idea: Snakes

- An active contour (snake) is a set of points which aims to enclose a target feature.
- Snakes are model-based methods for localization and tracking of image structures.
- The snake is defined as an energy minimizing contour (often defined using splines).
- The energy of the snake depends on its shape and location within the image.
- Snakes are attracted to image boundaries through forces.

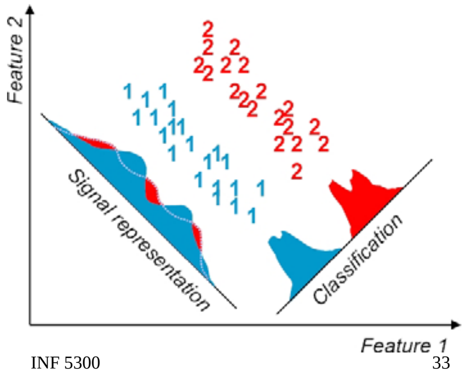


Back to pattern recognition: Feature selection

- Given a large set of N features, how do we select the best subset of m features?
  - How do we select m?
  - Finding the best combination of m features out a N possible is a large optimization problem.
  - Full search is normally not possible.
  - Suboptimal approaches are often used.
  - How many features are needed?
- Alternative: compute lower-dimensional projections of the N-dimensional space
  - PCA
  - Fisher's linear discriminant
  - Projection pursuit and other non-linear approaches

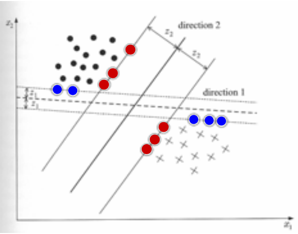
Linear feature transforms

- Principal components analysis (PCA)
  - - signal representation, unsupervised
  - Minimize the mean square representation error
- Linear discriminant analysis (LDA)
  - -classification, supervised
  - Maximize the distance between the classes



Classification by Support Vector Machines

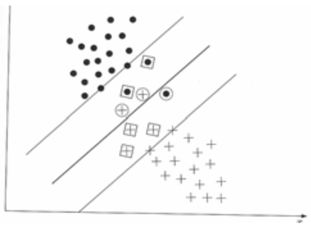
- SVM defines the decision boundary from the samples closest to the other classes.
- If a problem is linearly separable, many hyperplanes can divide the classes.
- In SVM the hyperplane is selected as the one the the largest distance to the other class.



The support vectors for hyperplane 1 are the blue circles.
The support vectors for hyperplane 2 are the red circles.

SVM: The nonseparable case

- If the two classes are nonseparable, a hyperplane satisfying the conditions $w^Tx-w_0=\pm1$ cannot be found.
- A penalty will now be given to the samples closest to the boundary.
- A magic trick solves the classification problem in a higher dimensional space.



☐ Correctly classified

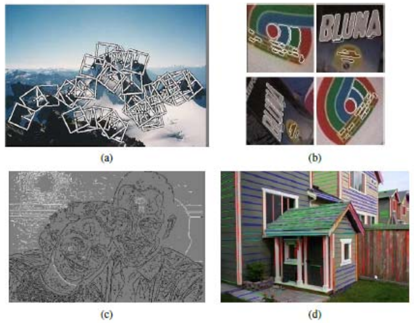
◯ Erroneously classified

Which features are good for tracking?

- In many applications, we need to match or track different parts of the image.
- Applications are e.g.
  - Tracking in video
  - Motion estimation
  - Multi-camera views

Features for tracking

- Should we track:
  - Pixel patches?
  - Regions?
  - Edges?
  - Lines?
  - Is multiscale needed?



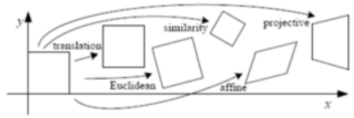
Motion estimation

- Motion estimation in e.g. medical ultrasound images gives information about the object (e.g. how the hearth behaves).
- Motion estimation in a video sequence can be due to camera instability and might be removed.
- Visual motion indicates the dynamics in the scene.
- Geometrical motion models are used in parametric motion estimation.

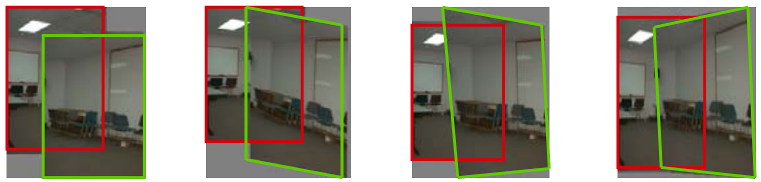
Patch matching

- How do we determine correspondences?
  - *block matching* or *SSD* (sum squared differences)
- How do we ensure a smooth motion estimate?

Parametric motion models

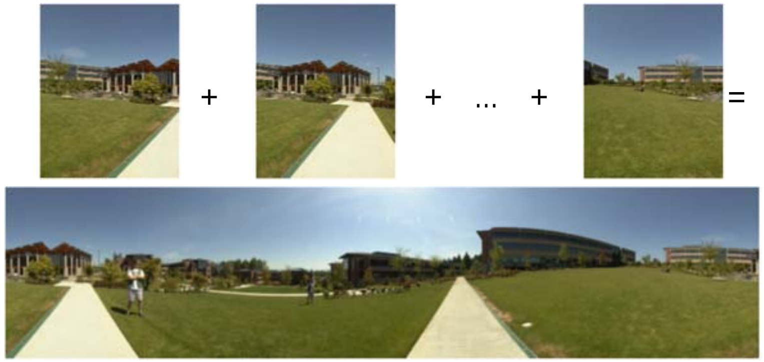


Translation **Affine** **Perspective** **3D rotation**



2 unknowns **6 unknowns** **8 unknowns** **3 unknowns**

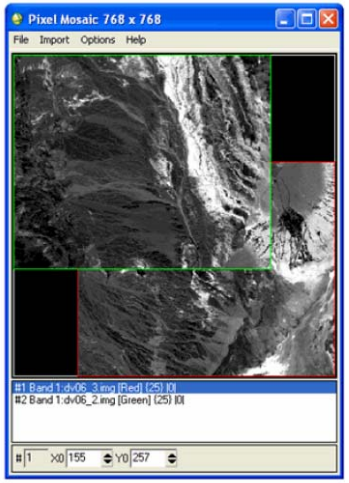
Image Mosaics

- 

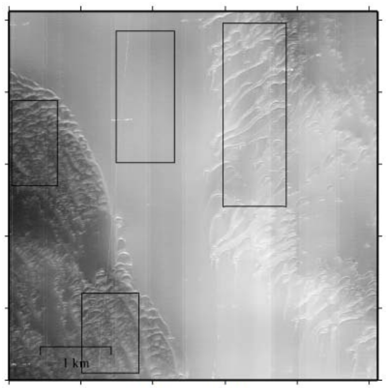
Goal: Stitch together several images into a seamless composite

Applications of stitching: Panoramic images

Applications of stitching: remote sensing



Applications of stitching: seafloor mapping



(a) Original DTM

Curriculum for next lecture:

- Have a brief look at section 3.7 Global optimization in Szeliski
- Active contours will be based onon Chapter 6 from Feature Extraction & Image Processing by Nixon and Aguado. You can find it at
- /ifi/asgard/k00/inf5300/pensum-artikler/activecontour_kap6.pdf