Introduction to Computer Vision

Michael J. Black

Project Ideas
Dates

11/13 Proposals due
- 1 page write-up
- summary and goals (problem/approach)
- what are the key references
- where will you get data and how will you evaluate your method?

12/14 Projects due
Eigen faces
Detect faces
Mahalanobis distance
Masked by skin detections
If you work on face detection

• You must use color in some way (e.g. color eigenspace)
• You must search across a range of scales.
• You must try it on some image data that was not collected in class.
• Use the Moghaddam and Pentland paper.
• Extra: try gender recognition.
Space Carving


• Images from ETH-80 database.
Sources of data

http://www.vision.ethz.ch/projects/categorization/eth80-db.html
Image segmentation

Normalized cuts:
http://www-2.cs.cmu.edu/~jshi/Grouping/
Face gender classification?

http://vis-www.cs.umass.edu/lfw/
13233 labeled faces of 5749 people found with Viola-Jones detector (Adaboost).
Use names to get gender: (http://www.gpeters.com/names/baby-names.php?)
More face images

There are many other face databases on the web. Eg:

http://titan.cog.brown.edu:8080/TarrLab/face-place
Kalman Filter Tracker

• Real-time tracker using a PC camera.
Mean-shift tracking

http://www.caip.rutgers.edu/~comanici/Papers/KernelTracking.pdf
Facial expressions

• Affine head tracking
• analysis of motions of facial features
Recognizing Facial Expressions using Local Parametric Motion Models

M. Black, Xerox PARC
Y. Yacoob, U of Maryland
EigenTracking
(Black and Jepson)

Combines affine motion estimation with PCA representation to allow tracking of deforming objects.

Data:
http://www.cs.brown.edu/~black/images.html

http://www.cs.brown.edu/~black/eigenTrack.html
EigenTracking:
Robust Affine Matching Using a View-Based Representation

Submitted to ECCV’96

Confidential: For Review Only
http://www.robots.ox.ac.uk/~c bibby/research_pwp.shtml
Tadpole tracker

- Contact black or moldovan for data.
(a) Example image with predicted position plotted
(b) Motion model-based distribution
(c) Pixel intensity distribution
(d) Combined distribution
(e) Example image with 200 particles
Rodent tracker

Contact: Prof. Russ Church (Russell_Church@brown.edu)
Active Shape/Appearance Models

http://www.isbe.man.ac.uk/~bim/
Lot’s of support code and data on web.
Stereo

- http://vision.middlebury.edu/stereo/
Dense Optical Flow or Stereo

Data and ground truth flow:
http://vision.middlebury.edu/flow/
Same for stereo
http://vision.middlebury.edu/stereo/

Deqing Sun
Pedestrian Detectors


Figure 6. Our HOG detectors cue mainly on silhouette contours (especially the head, shoulders and feet). The most active blocks are centered on the image background just outside the contour. (a) The average gradient image over the training examples. (b) Each “pixel” shows the maximum positive SVM weight in the block centered on the pixel. (c) Likewise for the negative SVM weights. (d) A test image. (e) Its computed R-HOG descriptor. (f,g) The R-HOG descriptor weighted by respectively the positive and the negative SVM weights.
Non-linear diffusion

Image denoising.

Original Image

Edge Enhancing Diffusion
http://www-mount.ee.umn.edu/~guille/inpainting.htm
-this turns out to be too hard and the best thing to do is a non-linear diffusion method.

Image Inpainting

Since 1699, when French explorers landed at the great bend of the Mississippi River and celebrated the first Mardi Gras in North America, New Orleans has brewed a fascinating melange of cultures. It was French, then Spanish, then French again, then sold to the United States. Through all these years, and even into the 1900s, others arrived from everywhere: Acadians (Cajuns), Africans, Indige--
Colorization

http://www.cs.huji.ac.il/~yweiss/Colorization/
Super-resolution

Super-resolution
More Project “Ideas”

Temporal model of mouth motions (HMM) for recognition.
More advanced machine learning method for mouth or person detection
  - AdaBoost
  - support vector machines
Bayesian image denoising
Stereo
Space carving from silhouettes
Grab-cut
Moghaddam mixture model for face/mouth recognition.