Foundation and core

Rick Szeliski &
Andrew Zisserman

The Need ….

George Santayana:

"Those who cannot remember the past are condemned to repeat it"

- We see in other communities their rediscovery of results and methods already well known in computer vision, e.g. in multi-media retrieval

- It is even worse if this happens in our own community

Outline

1. Some successes and areas where there has been significant progress

2. 20 things every CV researcher should know

3. Dissemination

Multiple View Geometry

One of the success stories of Computer Vision

- 1990

  estimate epipolar geometry of two images using a calibration rig

- 2000

  automatic estimation of cameras and structure from 1000s of frames of uncontrolled video footage

Two advances made this possible:

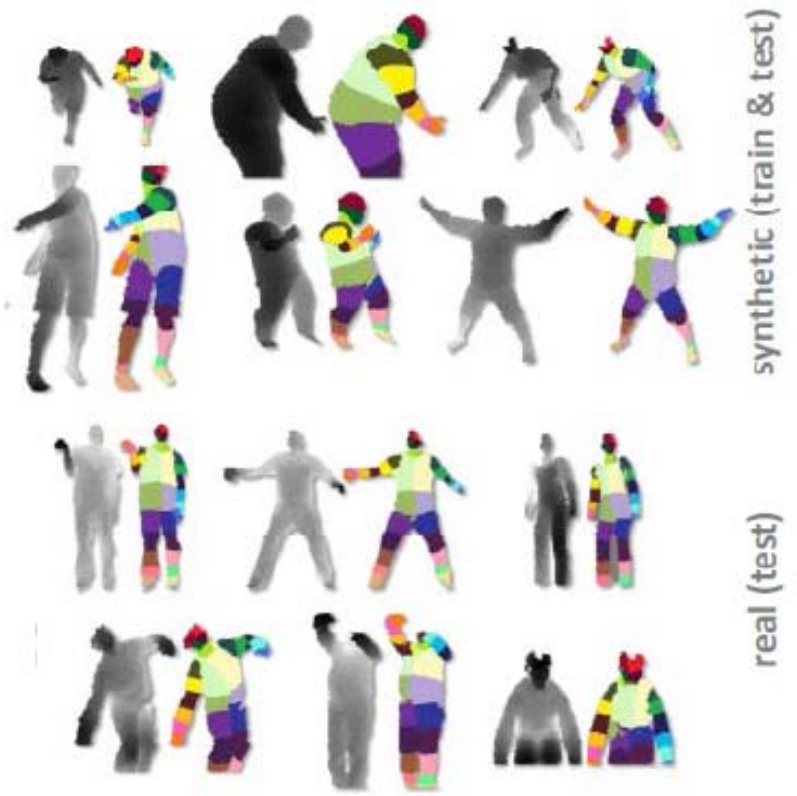
1. Understanding the (projective) geometry of multiple views
   - fundamental matrix, trifocal tensor …
2. Automatic estimation of this geometry from images
   - development of robust estimation algorithms

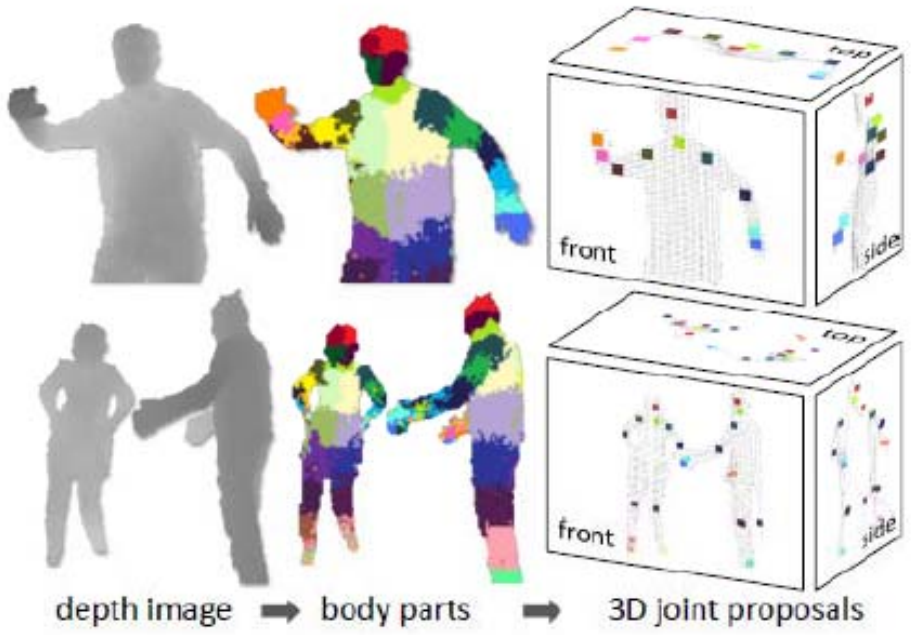
Computer vision in the movies

Match mover/camera tracking

PhotoSynth/PhotoTourism

Kinect





J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman and A. Blake. Real-Time Human Pose Recognition in Parts from Single Depth Images. **Best paper award at CVPR 2011**

What works reasonably well?

Significant progress and applications:

- Face detection/recognition
- OCR, (isolated) hand writing recognition
- Industrial inspection
- Affine covariant detectors
- Gaze tracking
- Multi-view stereo reconstruction
- Tracking isolated people in video
- Interactive segmentation, e.g. medical
- Image denoising and correction
- Image stitching, HDR
- Fingerprint recognition
- Duplicate image/video search
- Instance recognition for weakly textured object

What works reasonably well continued?

- Active range finding
- Optical flow
- Autonomous driving with LIDAR

20 things everyone should know

- Data mining community. 2005/2006. Identified
  - 10 challenging problems, and
  - 10 most influential algorithms.

- The result was:
  - C4.5, k-Means, SVM, Apriori, EM, PageRank, AdaBoost, kNN, Naive Bayes, and CART
  - **Top 10 algorithms in data mining,** Wu et al., Knowl Inf Syst, 2008

- In Computer Vision??
  - There is so much: so many areas, so many disciplines, so many application areas

What 20 techniques should all computer vision researchers know (short list)?

1. Image formation and optics
2. Image processing, filtering, Fourier analysis
3. Pyramids and wavelets
4. Feature extraction
5. Image matching
6. Bag of words
7. Optical flow
8. Structure from motion
9. Multi view stereo
10. Segmentation
11. Clustering
12. Viola-Jones
13. Bayesian techniques
14. Machine learning
15. RANSAC and robust techniques
16. Numerical methods
17. Optimization
18. Range finding, active illumination
19. Algorithms
20. Graph cuts
21. Dynamic programming
22. Complexity analysis
23. MATLAB **and** C++. **and** assembly (optional: GPU programming)
24. Communication and presentation skills

Following slides collated from poll

Image and features

- NCC
- Interest point operators
- Scale invariant and affine invariant detectors & descriptors
- Scale space
- Image processing, filtering, Fourier analysis
- Pyramids and wavelets
- Edge detection
- Restoration e.g. deblurring, super-resolution
  - Linear, e.g. Wiener filter
  - MRF
  - Non-local means/BM3D/bilateral filter

Segmentation, grouping and tracking

- Segmentation
  - Normalized cuts
- Grouping
  - Hough transforms
- Clustering
  - K-means
  - Mean-shift
  - Pedro-clustering
- Tracking
  - Kalman filter
  - Particle filter

Multi-view: stereo, SFM, flow

- RANSAC and other robust techniques
- Geometry:
  - epipolar geometry (projective and affine)
  - planar homographies
  - Affine camera
- Geometry estimators
  - 8 point algorithm for F
  - 4 point algorithm for H
- Factorization
- Bundle-adjustment
- Flow
  - Horn & Schunck L2
  - Lucas-Kanade
  - L1 regularized

Recognition

- Bag of visual words
- HOG, SIFT, GIST
- Spatial pyramid
- Spatial configurations/Pictorial structures
- Sliding window/jumping window
- Cascades

Others

Machine Learning

- Adaboost
- kNN
- SVM
- Random forest
- PCA, ICA, CCA
- EM
- MIL/Latent-SVM
- Regularization
- HMM
- Graphical & Bayesian models

Optimization

- Classical linear and non-linear
- Graph operations
- Dynamic programming/message passing for MAP, max-marginals
- Graph cuts for binary variable MAP

- Texture synthesis

Who can help?

Dissemination

- Accessible Wikipedia articles for generalists, with high-quality links
- Separate Wiki: CVOnline ???
- Summer schools (especially US), **recorded**
- Tutorial videos
- Online programming assignments (Ted procedures),
- Tutorial courses
- Industry panels @ conferences (will people attend?)
- Introductory video [Serge]

Courses that your computer vision students do/should take

- Typically, they take computer vision directly (see core techniques), sometimes after graphics or image processing.

- Ideally, would like them to take (pre-requisites or co-requisites):

1. Linear algebra, numerical methods, optimization
2. Statistics, Bayesian and robust methods, machine learning
3. Computer graphics **and** image processing
4. Optics, image formation, sensor design
5. Visual perception
6. Programming: parallel languages (MATLAB), efficient languages (C++ and assembly)
7. Technical writing and presentation (all graduate students, possibly UGs as well)

Are current computer vision textbooks sufficient? What is missing?

- No good up-to-date introductory book
- Need algorithm descriptions + pseudo code
- Video lectures, Khan academy (voice-over-pen/tablet drawing), http://www.ai-class.com/, http://robots.stanford.edu/cs221/
- (On-line) books supplemented by online tools, exercises, examples, resources, videos
- Social media + textbooks
- Ads to support free articles (??)