The Terrapins

Computer Vision Laboratory
University of Maryland
Justin Domke

Cornelia Fermüller

Yi Li

Yiannis Aloimonos

Michi Nishigaki

Kostas Bitsakos

Xiaodong Yu

Alap Karapurkar

and some others
A challenge for Understanding Vision

- Navigation
- Recognition
- Memory
- Active Vision
The approach

- Navigation
- Recognition
Navigation

• Build map using SLAM from laser

Set way points
Recognition Strategy

• Proper Nouns:
  discriminative features
  and geometric transform
  very few internet images (3)

• General Nouns:
  shape descriptor
  20-30 internet images
SIFT

- Dense feature points
- Usually correct matches
- Poor at too much distortion
MSERs

- Sparse keypoints
- Usually correct matches
- Great at affine invariance
Recognition Strategy for Proper Nouns

1. **Image**
2. **Compute Homography**
3. **Feature matching using SIFT**

- **Match with MSER** (affine invariant)
- **Unwarp using Edges**

Examples
The matching Process

Initial matches

20 matches

Final matches

45 matches
Another Example
Matches of planar objects

book "Vision and Art: the biology of seeing" by Livingstone found in images/cap055.jpg with 96 matches
General nouns

• Segmentation
  for the ground from trinocular stereo
  for the upper camera from color

• Shape description
  using adjacent line segments
Segmentation from depth information

Estimate the transformation of the ground plane between the different cameras.
Estimation of the ground plane homography
The descriptor

- Fit edges to small lines
- Adjacent lines: encode the relative coordinates w.r.t pivot point.
  - C / Z shape
  - Y shape
The codebook for the descriptor

• The advantage of the codebook
  – Generic
  – Quantization -> fast

• generate the codebook
  – A large dataset
  – Extract descriptor
  – Cluster the descriptor
Classifier: Support Vector Machine

• Suppose we have $N$ classes
• For each class, we train 1 SVM using images from this class vs other classes.
• Result: $N$ SVM classifiers (linear classifier in high dimensional space)
Example: Apply this descriptor to natural images
Result
Future steps

• Taking images: Segmentation into surfaces. Combine geometry (local occlusion information from motion and/or stereo) with edge information

• Recognition: surface boundaries, symmetry information