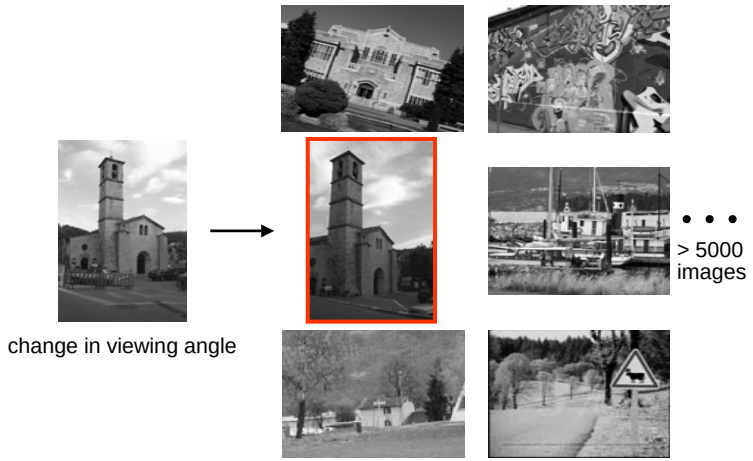
Object recognition

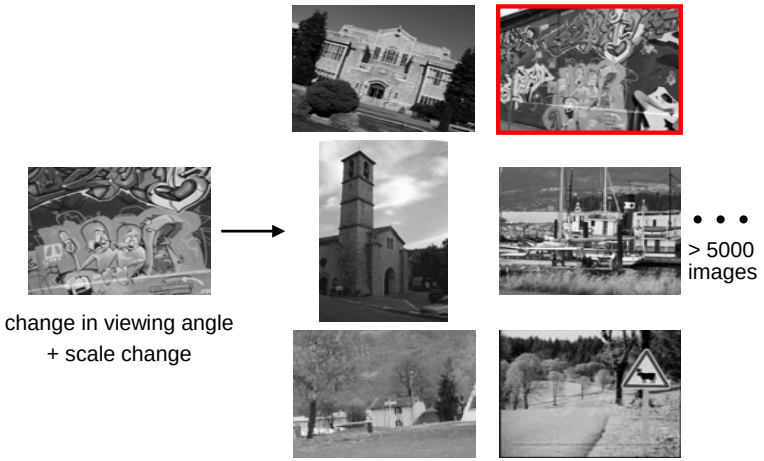
Image retrieval based on voting



Matches

22 correct matches

Image retrieval based on voting



Matches

33 correct matches

Outlier rejection (Lowe'99)

- Need to recognize clusters of just 3 consistent features among many feature match hypotheses
- RANSAC would be hopeless
- Use generalized Hough transform
  - Vote for each potential match according to model ID and pose
  - Insert into multiple bins to allow for error in similarity approximation
  - Using a hash table instead of an array avoids need to form empty bins or predict array size

Model verification

- Examine all clusters in Hough space with at least 3 features
- Perform least-squares affine fit to model
- Discard outliers and add features
- Evaluate correctness of a transformation based on number of features

Solution for affine parameters

- Affine transform of [x,y] to [u,v]:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

- Rewrite to solve for transform parameters:

$$\begin{bmatrix} x & y & 0 & 0 & 1 & 0 \\ 0 & 0 & x & y & 0 & 1 \\ & & \cdots & & & \\ & & \cdots & & & \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} u \\ v \\ \vdots \end{bmatrix}$$

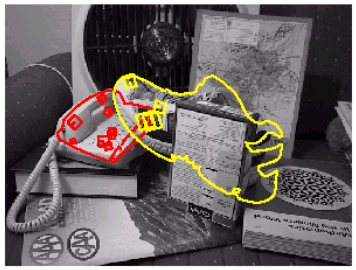
Affine transformations per object

- Affine fit approximates perspective projection
- Only 3 points are needed for recognition

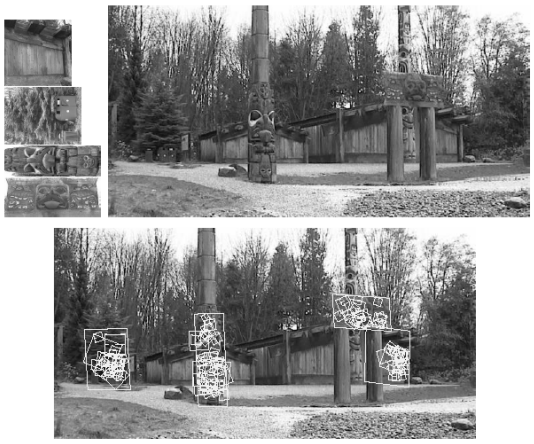
3D Object Recognition

- Only 3 keys are needed for recognition, so extra keys provide robustness
- Affine model is no longer as accurate

Recognition under occlusion

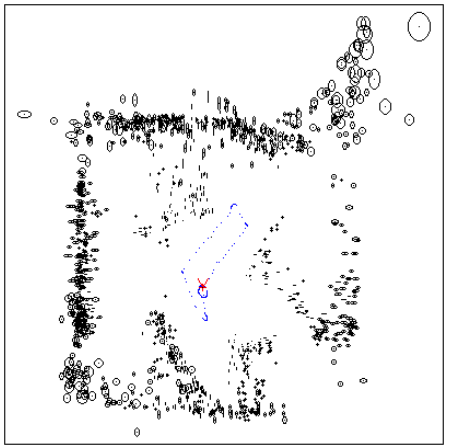


Location recognition (Lowe'01)



Robot Localization (Lowe'01)

Map continuously built over time

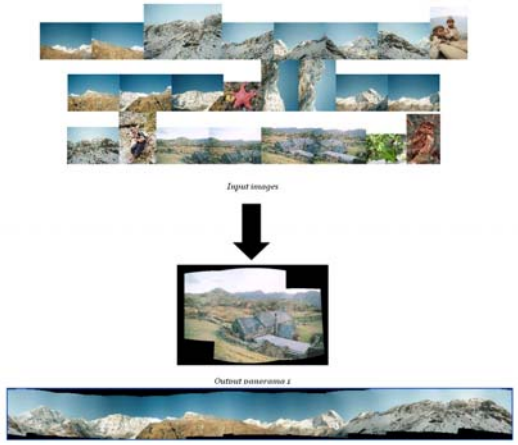


Recognizing Panoramas (Brown & Lowe'03)

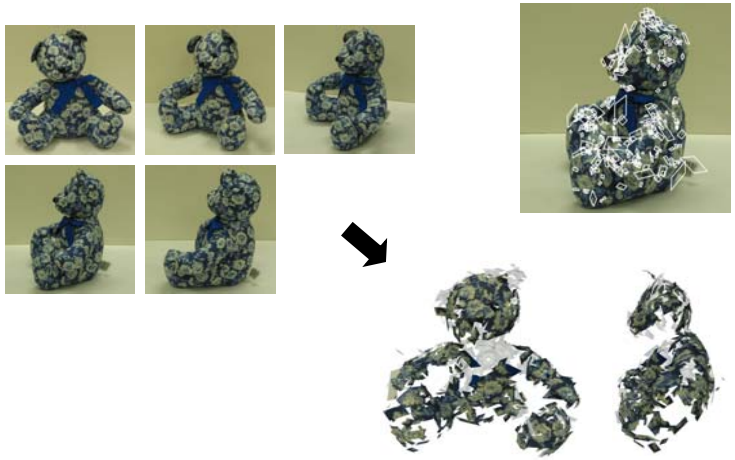
- Recognize overlap from an unordered set of images and automatically stitch together
- SIFT features provide initial feature matching
- Image blending at multiple scales hides the seams

Panorama automatically assembled from 143 images

Multiple panoramas from an unordered image set



3D Recognition

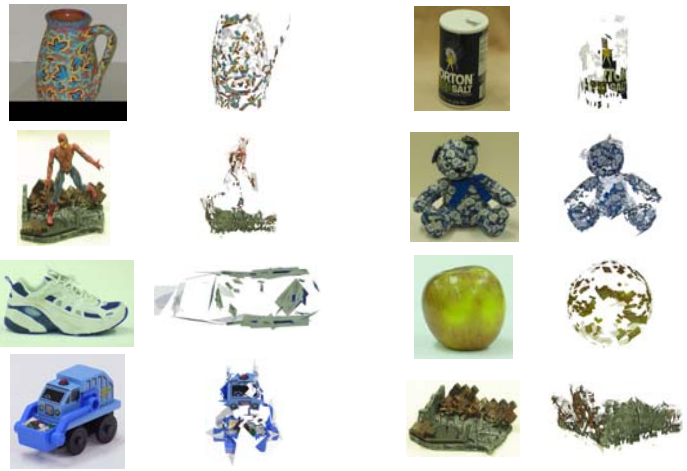


3D object recognition using affine regions

- 3D objcts are never planar in the large, but they are always planar in the small
  => representation : local invariants and their spatial layout
- Recognition uses appearance and geometric constraints
- Improved recognition performance due to
  – additional constraints for verification
  – more compact representation of the model

[3D object modeling and recognition using affine-invariant patches and multi-view spatial constraints, F. Rothganger, S. Lazebnik, C. Schmid and J. Ponce, CVPR'03 & IJCV'05]

Object gallery



Recognition & Pose Recovery

Recognition & Pose Recovery

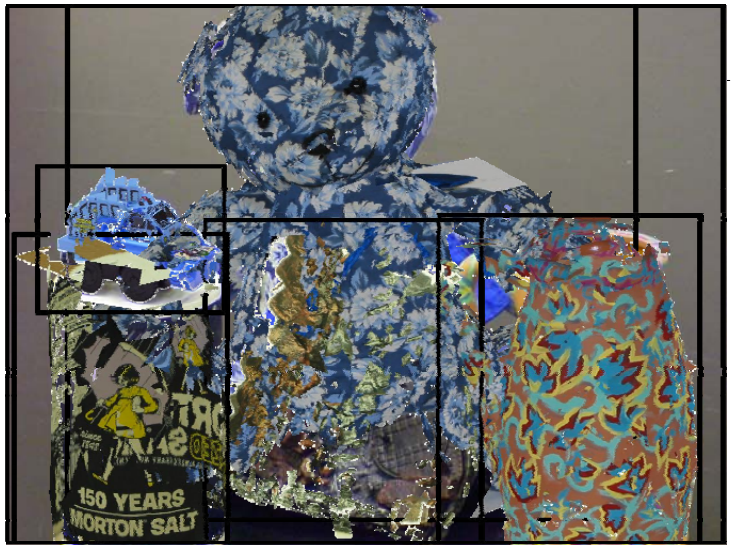
initial matches based on appearance

Recognition & Pose Recovery

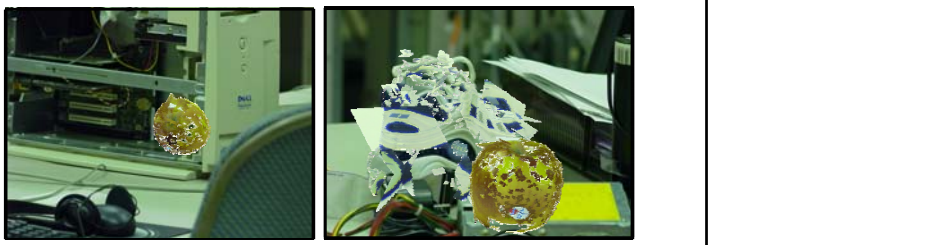
matches verified with geometrical constraints

Recognition & Pose Recovery

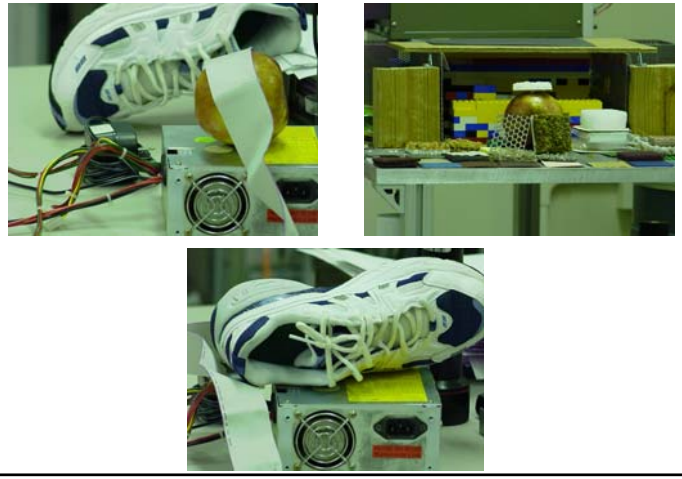
recovered pose



More recognition results



Failure cases

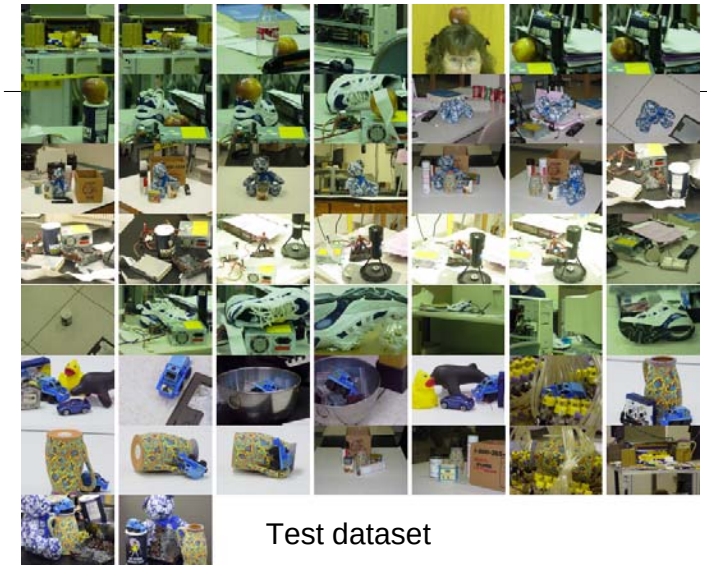


Comparison of 6 recognition algorithms

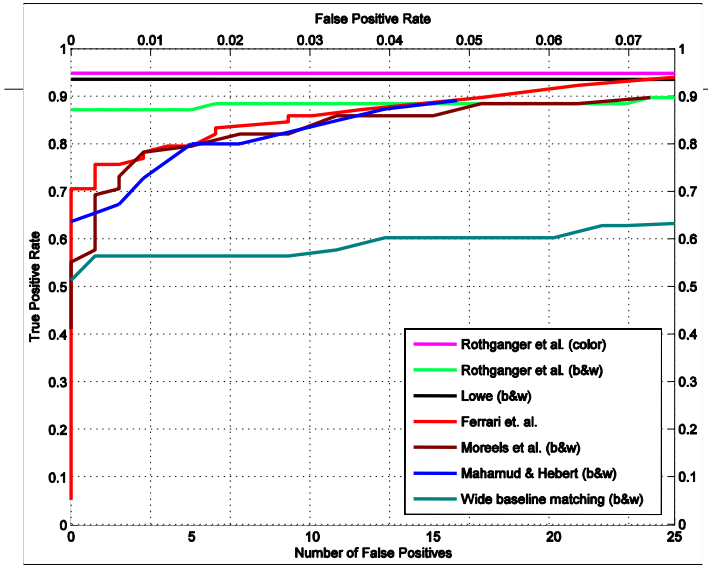
1. Rothganger, Lazebnik, Schmid & Ponce (CVPR'03) → 3D model

2. Lowe (ICCV'99, IJCV'04)
3. Mahamud & Hebert (CVPR'03)
4. Moreels, Maire & Perona (ECCV'04)
5. Ferrari, Tuytelaars & Van Gool (ECCV'04)
6. Matching (baseline method)

→ Views used independently



Test dataset



Video recognition

Image retrieval demonstration

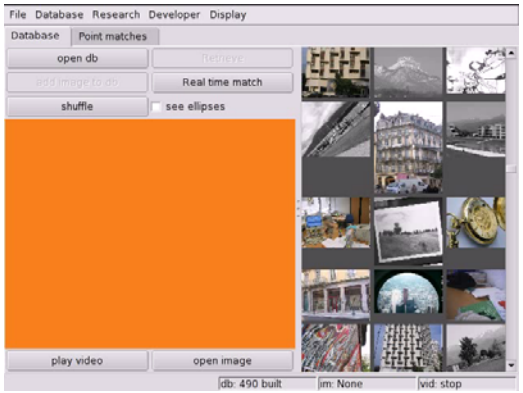


Image retrieval demonstration