NTNU

Kunnskap for en bedre verden

Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes

TDT 03 - Advanced Topics in Computer Graphics

Presentation by Ruben H. Fagerli

Paper to be summarized

- *Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes [1]*
- By Andrew E. Johnson and Martial Herbert
- Publish year 1999

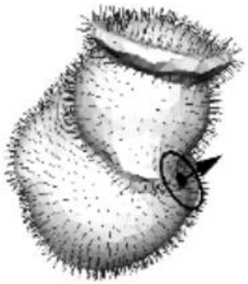
NTNU

Abstract

- 3D shape-based object recognition
- Multiple simultaneous objects
- Allows clutter and occlusion
- Spin image
- Compression

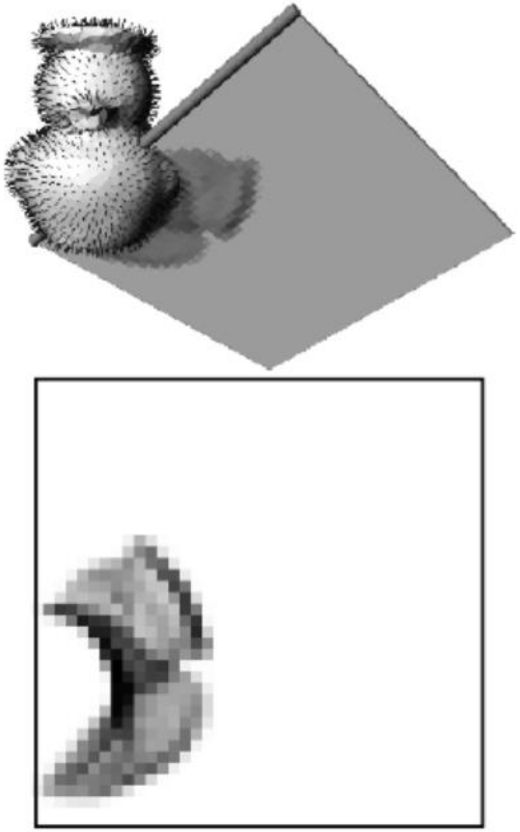
Surface Shape Representation

- 3D points
- Surface normals
- Descriptive image
  - Spin image



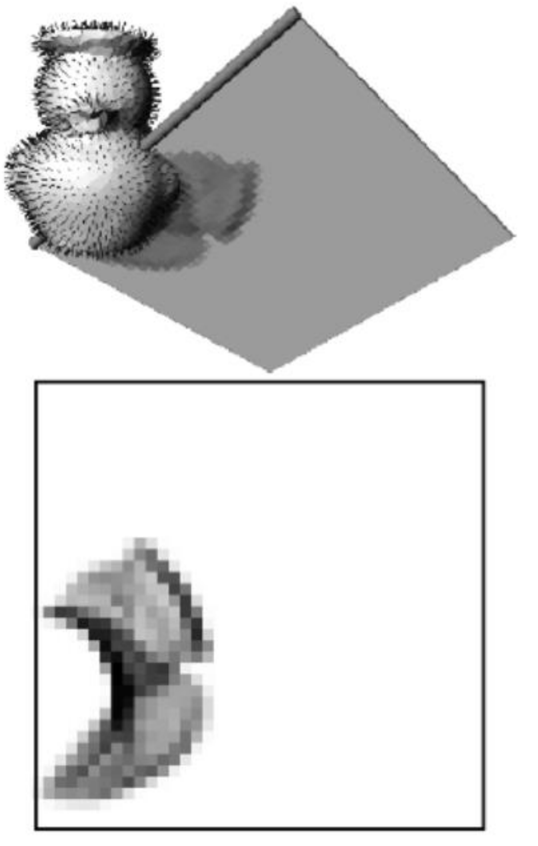
Spin Images

- Imagine a rod sticking out along the *surface normal*
- Centre of rod at *surface plane*



Spin Images

- Imagine a D x D plane of bins spinning around the rod
- Each bin collects all vertrices with:
  - Same radial distance α from the rod
  - Same elevation β from the surface plane



Spin Images - Surface normal [2]

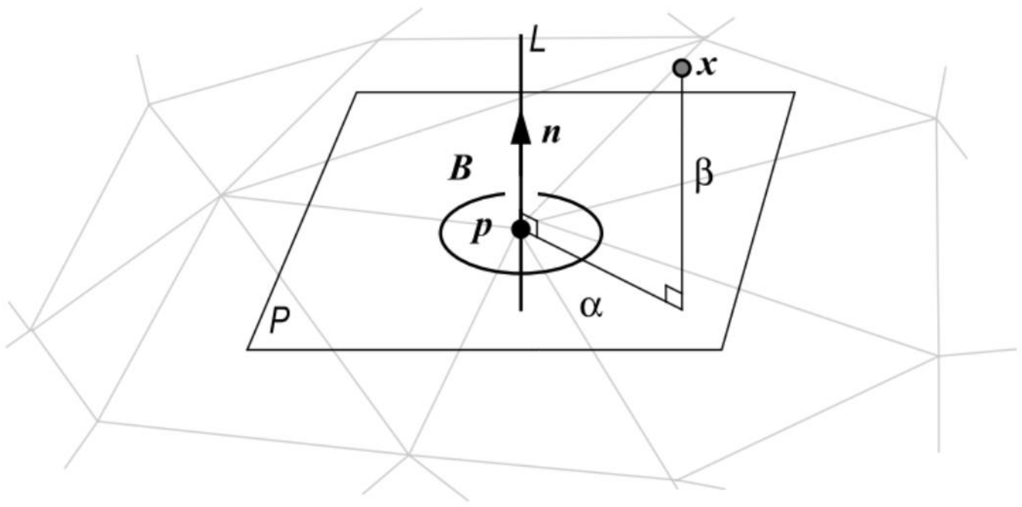
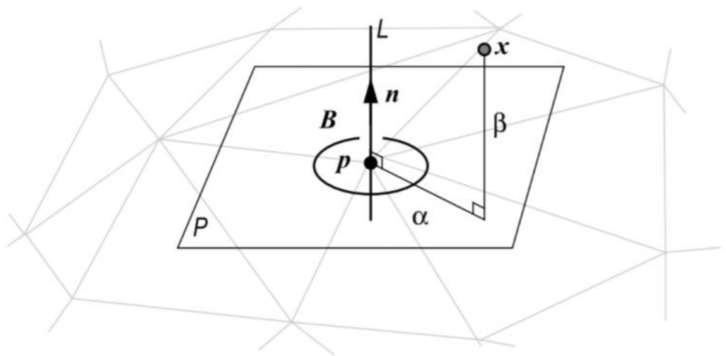


Figure 2-1: An oriented point basis created at a vertex in a surface mesh. The position of the oriented point is the 3-D position of the vertex, and the direction of the oriented point is the surface normal at the vertex. Two coordinates can be calculated given an oriented point: α the radial distance to the surface normal line L and β the axial distance above the tangent plane P.

Spin Images - Surface normal [2]

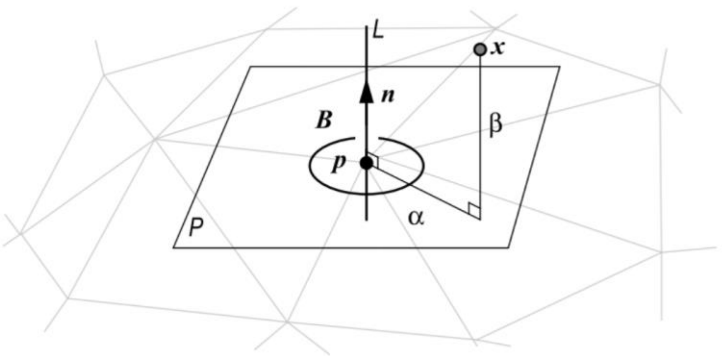
- Looked at citation 15 of the paper [1]
  - Spin-Images: A representation for 3-D Surface Matching [2]
  - More detail on spin images
  - More detail on surface normal/plane
  - Figure here is from that paper



Spin Images - Surface normal [2]

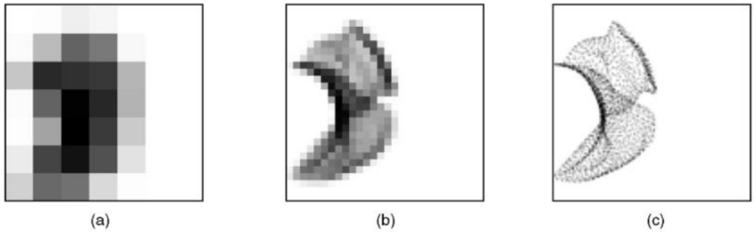
Smallest eigenvalue of inertia matrix of vertex and directly connected vertrices

- Outside of the surface must be determined



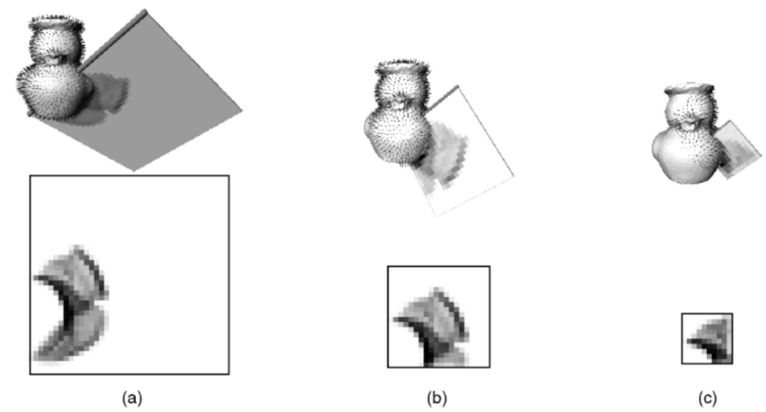
Spin Images - Parameters

- Bin size
  - Matched to mesh resolution
- Image width
  - Equal to height
- Support angle
  - 60 degrees +



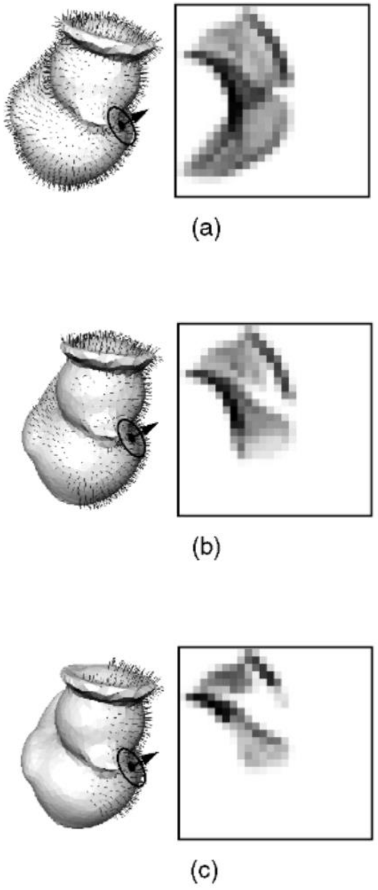
Spin Images - Support Distance

- *Bin size* and *image width* creates *support distance*
  - Amount of space swept out by a spin image
  - On order of model size



Spin Images - Support Angle

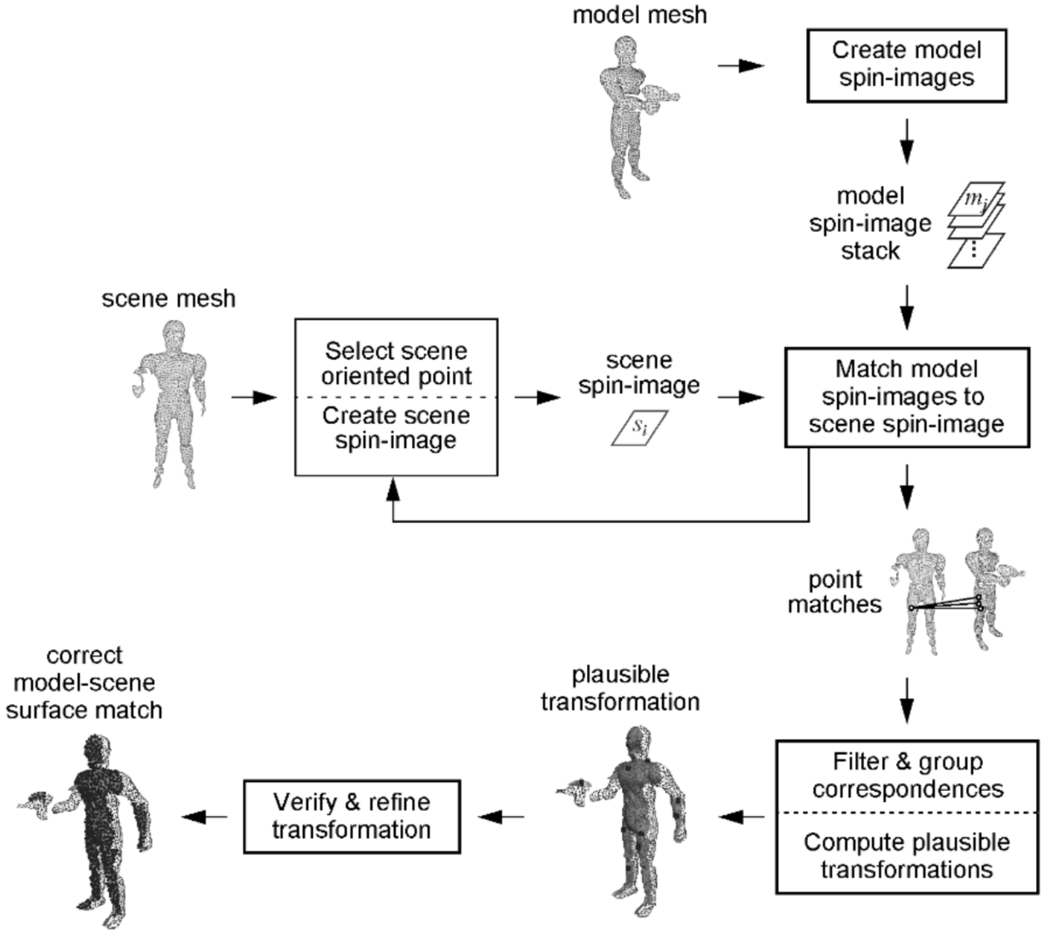
- *Support angle* determines how different *surface normals* can be
  - 180 degrees
  - 90 degrees
  - 60 degrees
- 60 degrees used
  - Less disturbance
  - Still great accuracy



Surface Matching Engine

- Points on different surfaces compared by spin image correlation
- Find ~100 correspondences
- Calculate rigid transformations
- Finalize with iterative closest point (ICP)

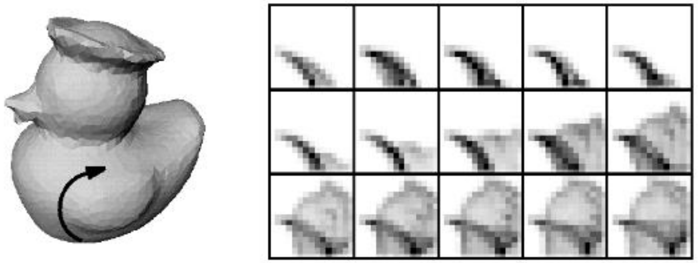
Surface Matching Engine



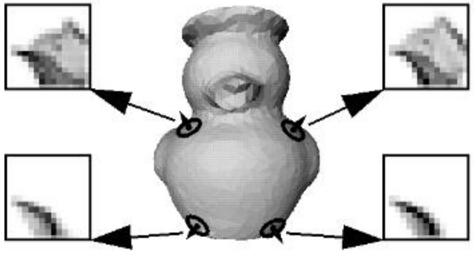
Spin Image - Redundancy

Various effects will cause redundancy:

- *Proximity* will create very simillar results
- *Symmetry* will cause similarity
- *Local similarity* in different objects



(a)



(b)

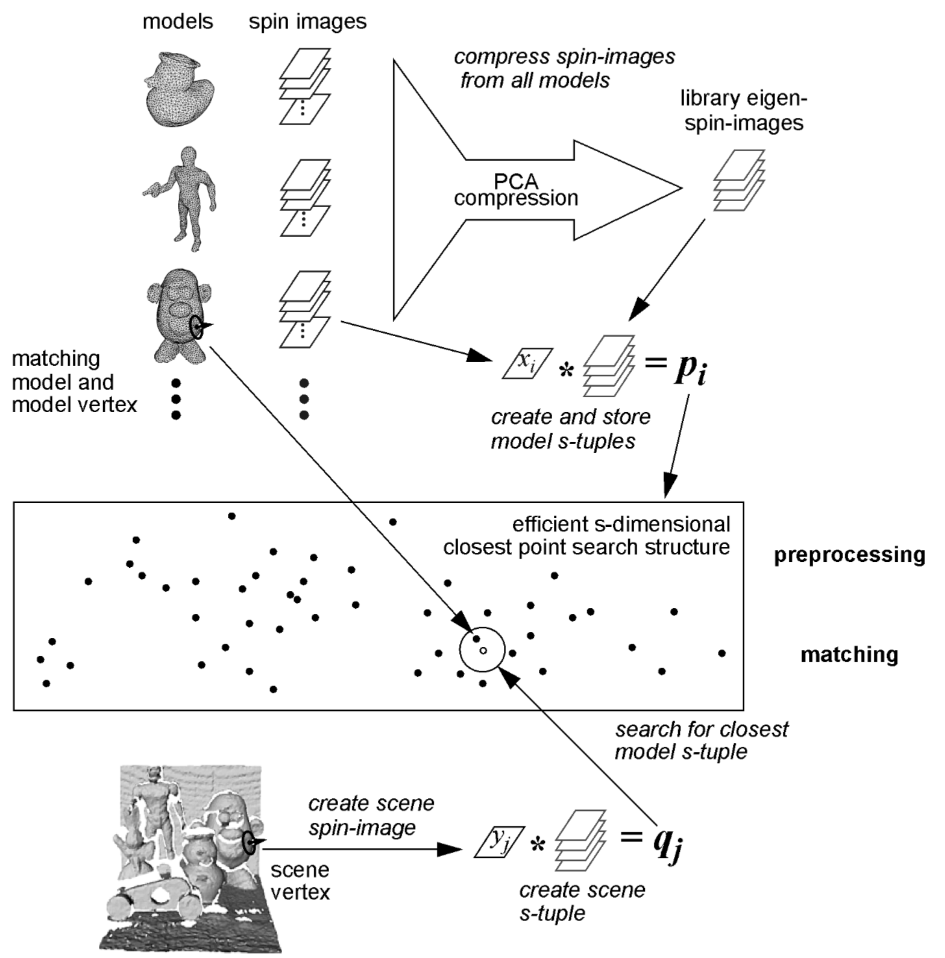
Spin Image Compression

- Native method inefficient
  - ~200 bins **D** per spin image
  - **v** vertrices in each model surface
  - **n** models to compare against
  - Scales linearly with increasing number of models in library
- Compression nessecary
  - Even more so when paper was released in 1999

Spin Image Compression

- **D**-dimensional vectors are compressed to **s**-tuples of eigenvectors by *principal component analysis* (PCA)
- Compression-ratio **s**/**D**
- Low dimension of **s** makes it possible to match in *sublinear time* using efficient closest point search

Compressed spin-image pipeline

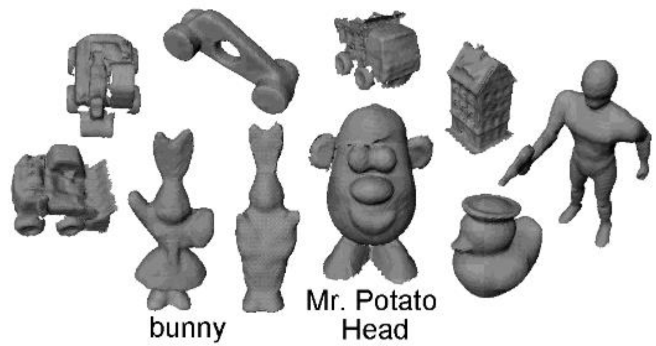


Compression Algorithm

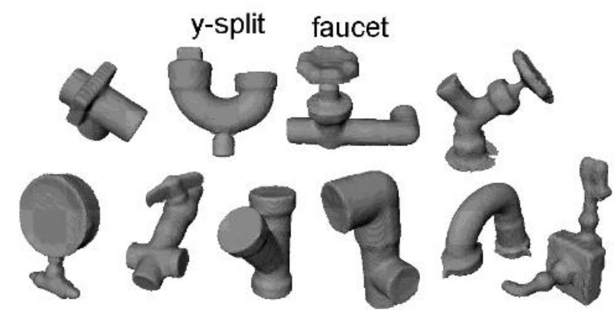
1. Calculate mean $\bar{\boldsymbol{x}}$ of all spin images $\boldsymbol{x_i}$
2. Substract the mean $\bar{\boldsymbol{x}}$ for all images $\boldsymbol{x_i}$ call it $\widehat{\boldsymbol{x}}_{\boldsymbol{i}}$
3. Set $\boldsymbol{S^m}$ to the list of all $\widehat{\boldsymbol{x}}_{\boldsymbol{i}}$ where each column is $\widehat{\boldsymbol{x}}_{\boldsymbol{i}}$
4. Covariance matrix $\boldsymbol{C^m} = \boldsymbol{S^m}(\boldsymbol{S^m})^T$
5. Calculate the eigenvectors $\boldsymbol{\lambda_i^m}$ of $\boldsymbol{C^m}$
6. Project the spin images $\widehat{\boldsymbol{x}}_{\boldsymbol{i}}$ into the subspace of the **s** eigenvectors of largest eigenvalue

Tests and results

- 20 model dataset
  - Toys
  - Plumbing
- Recognition in multi-object scenes
- Compressed vs uncompressed
- Effect of occlusion and clutter

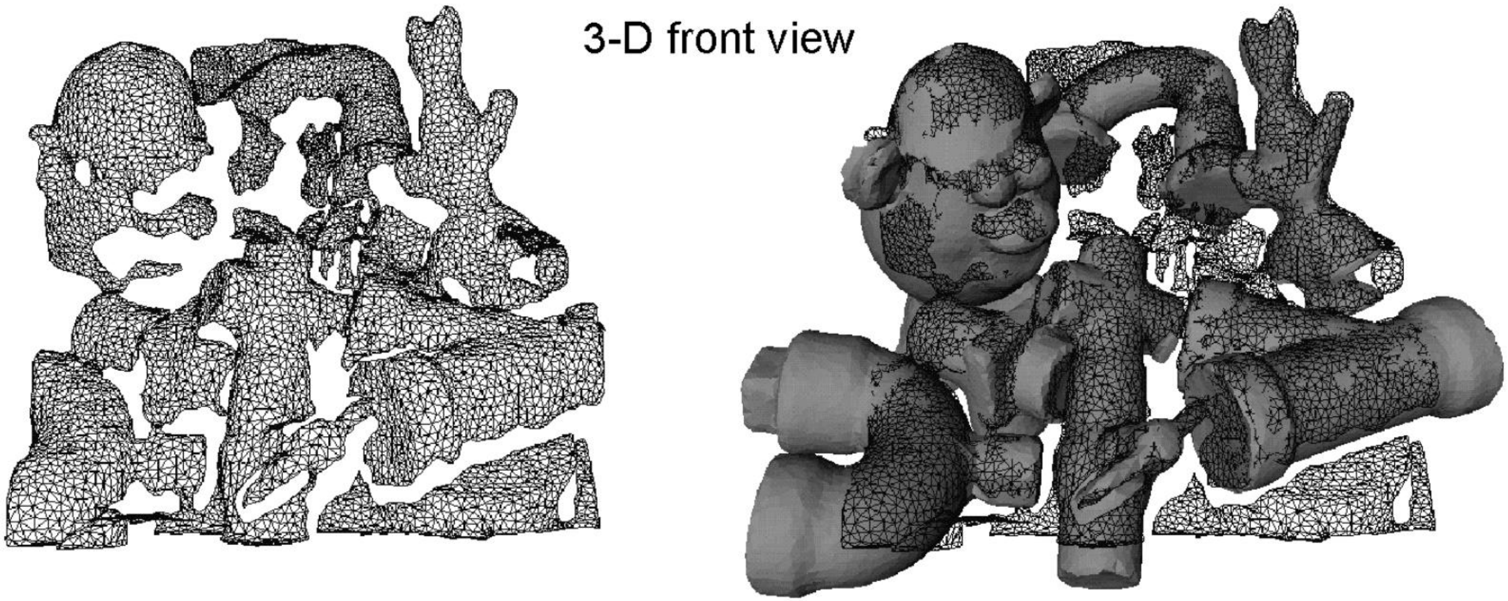


(a)

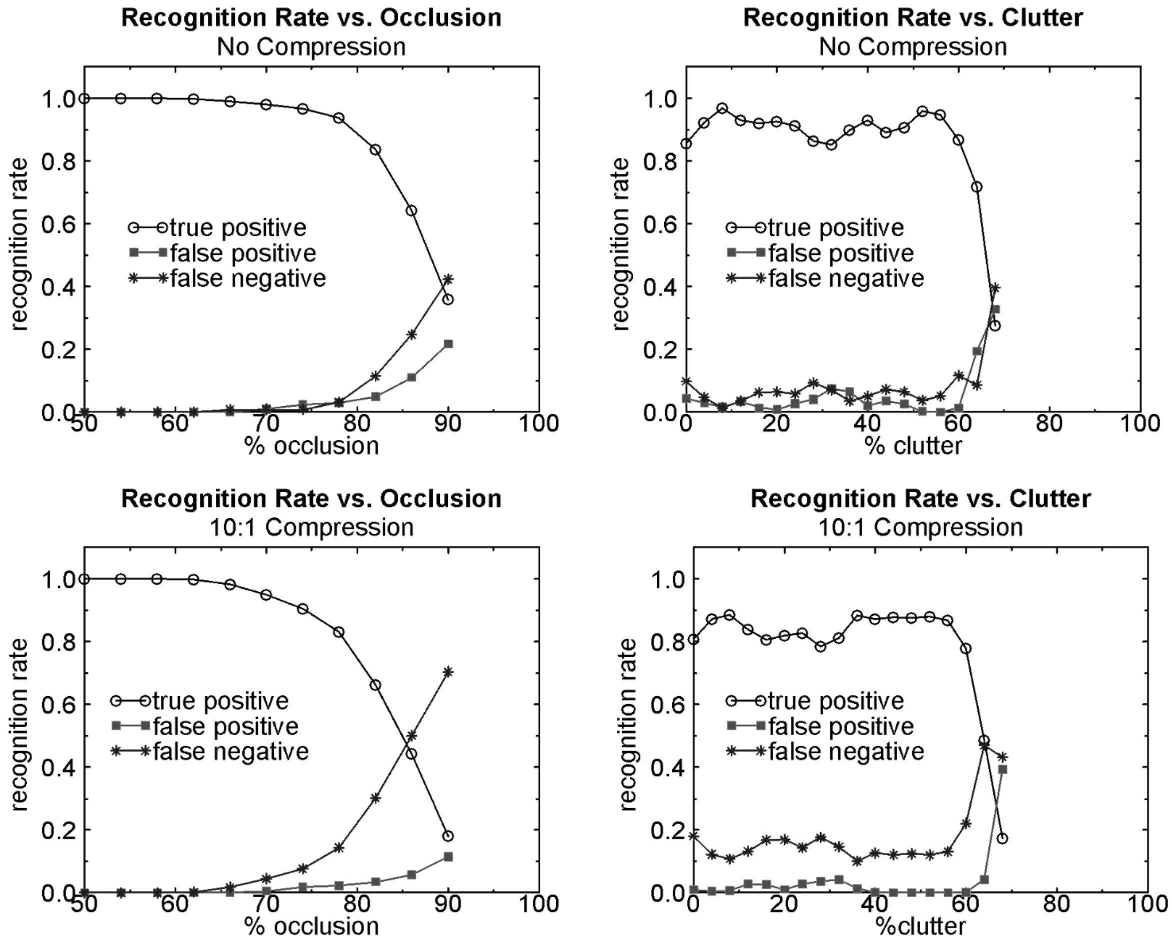


(b)

Results - Visual

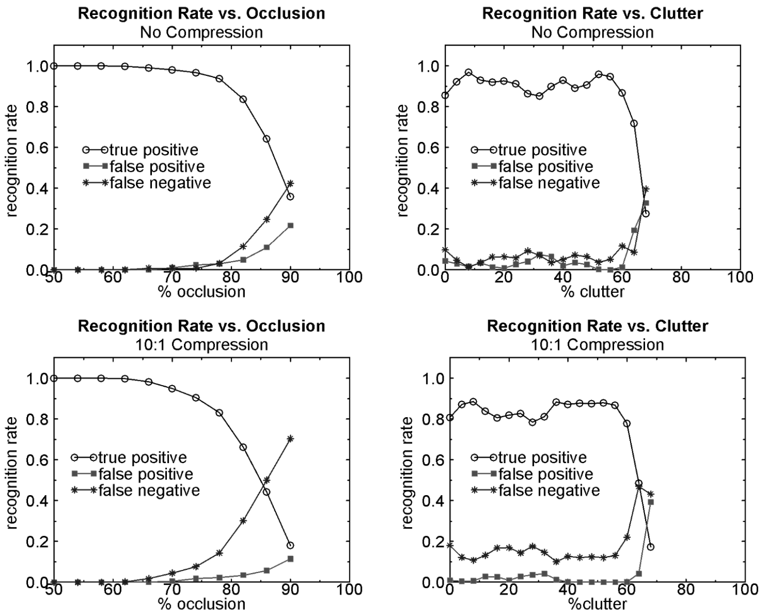


Results - Recognition Rates



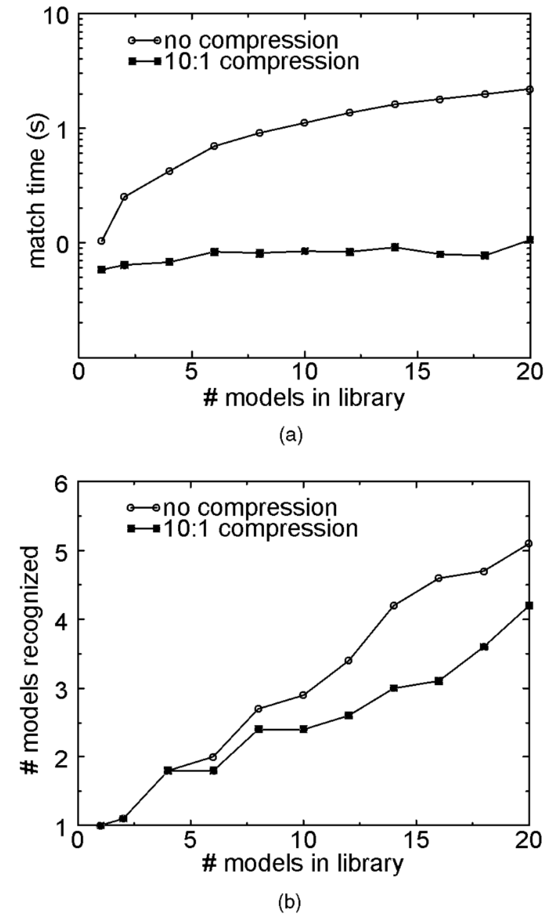
Results - Recognition Rates

- Great results up to 70% occlusion
- Compression gives slightly worse at higher occlusion
- No effect of clutter under 60%



Results - Speed

- Hardware (1999):
  - Silicon Graphics O2
  - 174-MHz R10000 processor
- Compressed solution scales much better with more models
  - 20 times faster with 20 model library

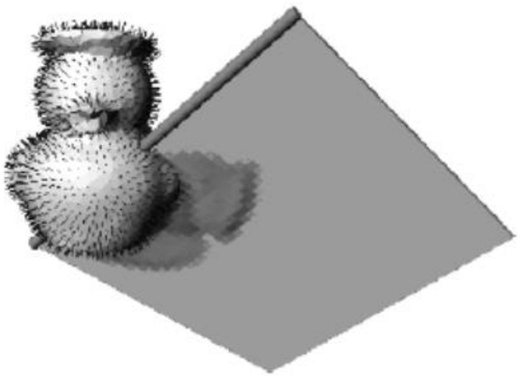


Summing up

- 3D model points
- Surface normal
- Spin Image

- Matching spin images of different surfaces

- Compression of spin image for better scaling



References

[1] A.Johnson and M.Hebert, «Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes», IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 21, No.5, pp.433-449, May 1999

[2] A.Johnson, *Spin-Images: A representation for 3-D Surface Matching, doctoral dissertation*, The Robotics Institute, Carnegie Mellon Univ., 1997