DELOS
NETWORK OF EXCELLENCE ON DIGITAL LIBRARIES

DELOS Task 3.8: Description, Matching and Retrieval by Content of 3D Objects

University of Florence (UNIFI-MICC), Italy
Institut National de Recherche en Informatique (INRIA), France
University of Modena and Reggio Emilia (UNIMORE), Italy
University of Amsterdam (UVA), Netherlands

All Tasks Meeting January 2006

Task objectives

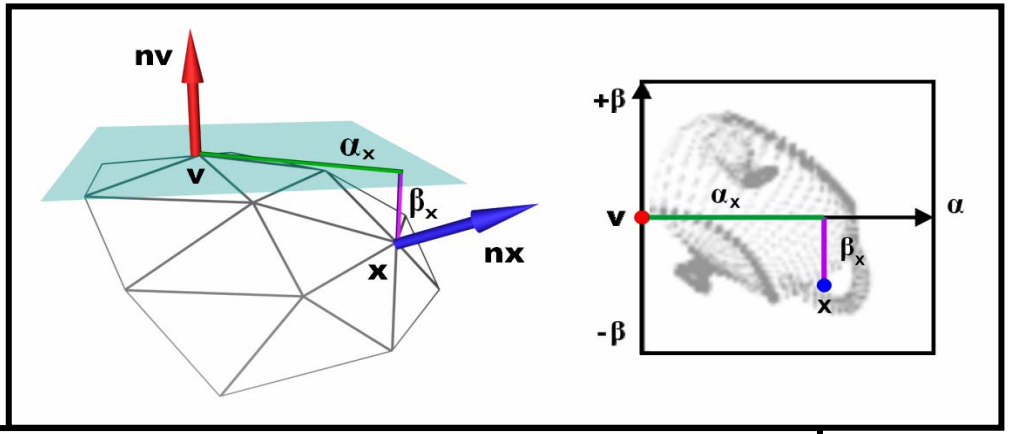
- The goal of this task is to develop solutions to support retrieval by content of 3D objects
  - The project aims at the definition of new view- and structural- based 3D object descriptors, the investigation of new models for indexing and similarity matching for searching in 3D objects digital libraries, the development of new solutions and metaphors for querying libraries of 3D objects.
  - The theoretical investigation of these models will end up with the design and development of prototype systems.

Activities

- Project activities have addressed so far:
  - Definition of a new 3D object representation based on ***spin image signatures*** (*view and structural based*) and development of a prototype retrieval system (UNIFI-MICC).
  - Development of a new 3D object representation ***enhanced depth buffer based*** descriptor (*view based*) and development of a prototype retrieval system (INRIA).
  - Implementation of the ***MPEG-7 3D object*** descriptor (UNIMORE).
  - Solutions for **3D objects reconstruction** from video data (UvA).

Spin image signatures descriptor (UNIFI-MICC)

- Spin images are 2D representations of 3D objects that maintain some properties of the vertices of the mesh. One vertex of the mesh is used as a reference. Positions of the other vertices are calculated as viewed from the reference vertex.



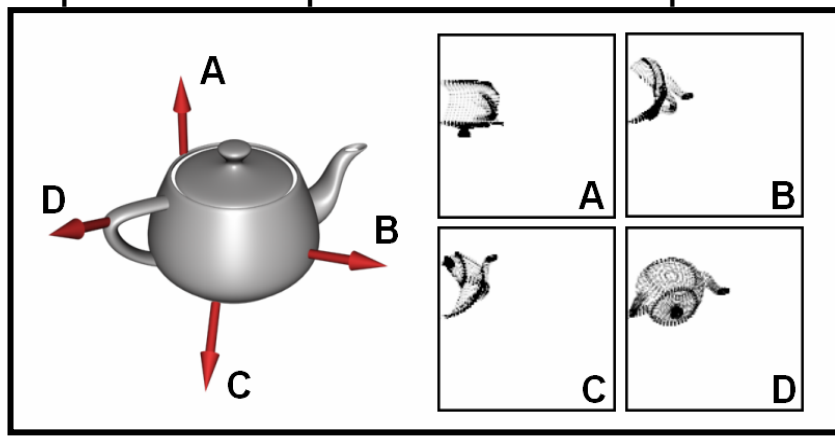
$$S_v : R^3 \rightarrow R^2 \qquad S_v(x) \rightarrow (\alpha_x, \beta_x)$$

$$\alpha_x = \sqrt{\|x - v\|^2 - (nv \cdot (x - v))^2} \qquad \beta_x = nv \cdot (x - v)$$

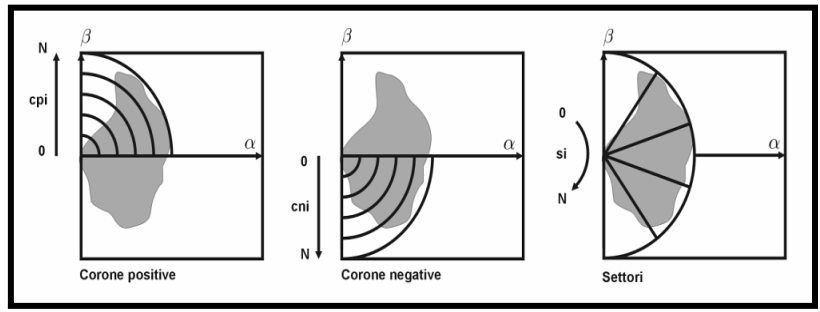
- For each vertex, *V* identifies the plane normal to *NV*, α β are the coordinates of the projection of a mesh vertex on this plane

Spin image signatures descriptor (UNIFI-MICC)

- Several spin images are needed to have a complete description of a 3D shape.
  - In that spin images use an object-centered coordinate system, they are invariant to rigid transformations, and have limited sensibility to variations of position of mesh vertices.

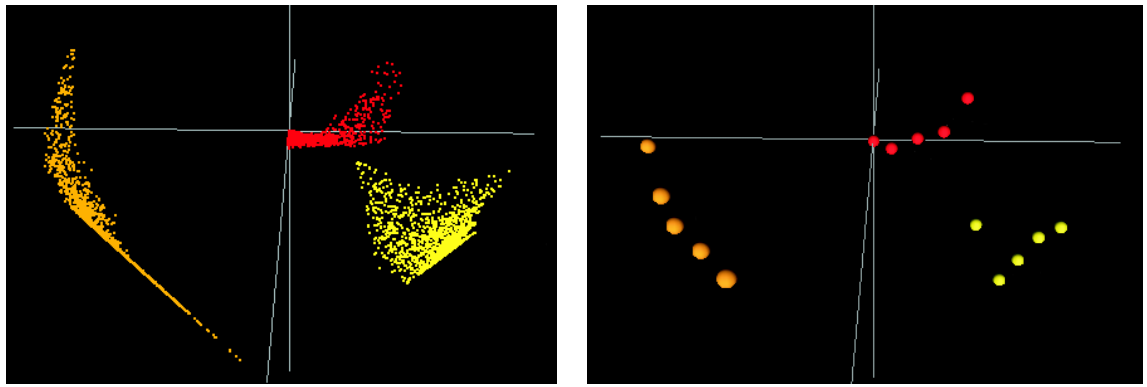


- For the purpose of content-based retrieval, we need to identify a small number of spin images that are sufficient to provide a complete description of the properties of the 3D shape. A spin image can be mapped into a vector descriptor that captures its salient properties.
  - Spin images descriptors are 18-dimensional feature vector, each element containing the fraction of the spin image area that is included within a circular section or a sector..

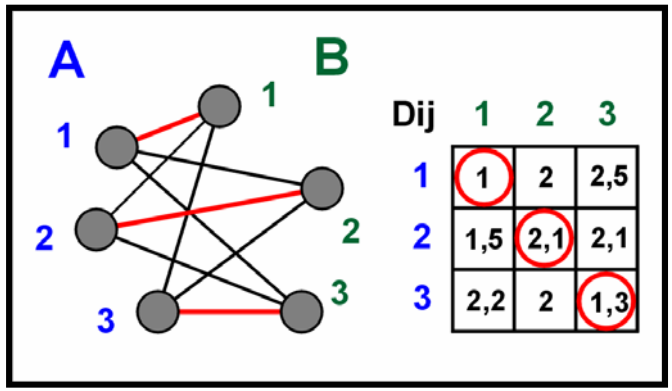


Spin image signatures CBR (UNIFI-MICC)

- Spin image descriptors are clustered using *fuzzy-c-means* so that a small number of spin image descriptors are obtained for each 3D object.



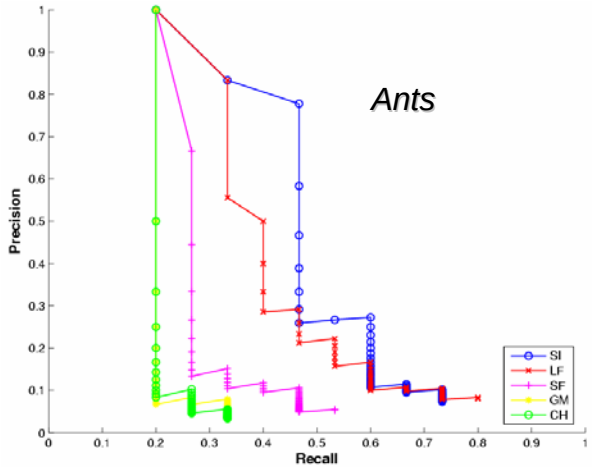
- Given two 3D object models, similarity can be checked considering the distances between the centers of the corresponding spin image clusters.

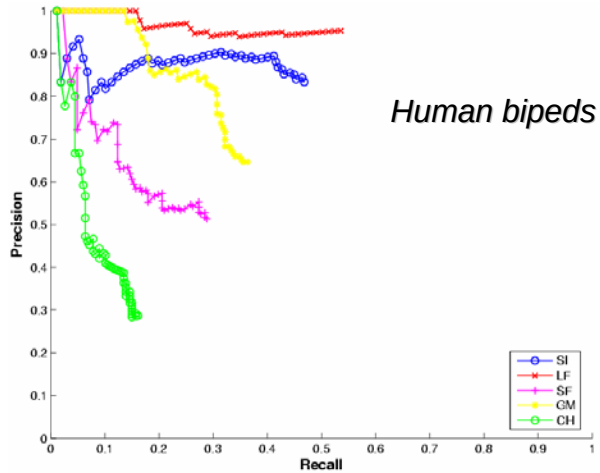


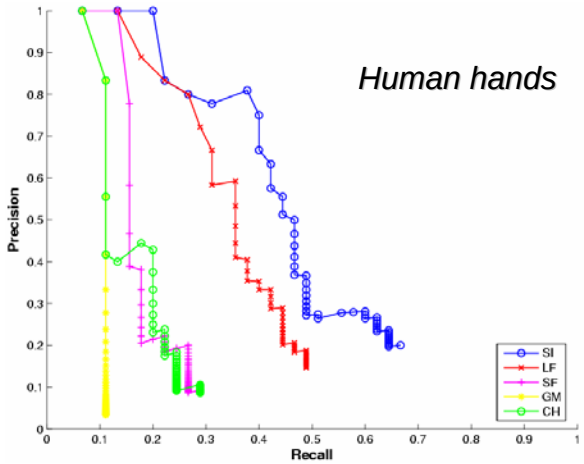
$$Dij = \sum_{k} \left| Ci_k - Cj_k \right|$$

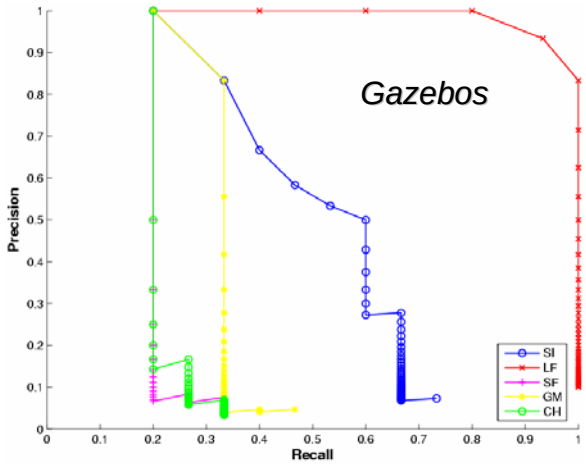
$$M_d(A,B) = \sum D_{coppie}$$

Precision–Recall on the Princeton Shape Benchmark (UNIFI)









SI: spin images

LF: light field

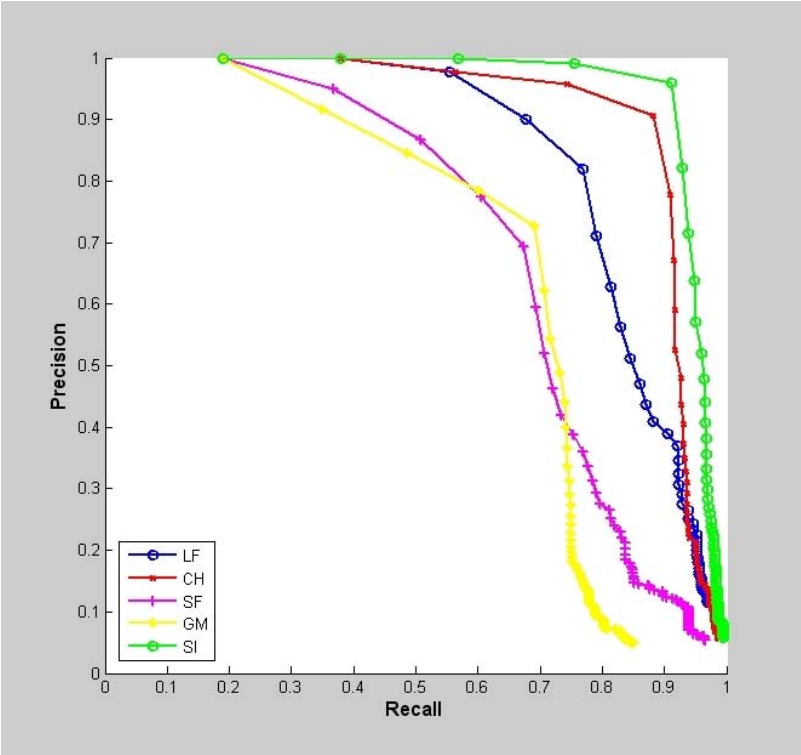
SF: shape function

GM: geometric moments

CH: curvature histo

Precision–Recall on the Princeton Shape Benchmark (UNIFI)

Average values



SI: spin images

LF: light field

SF: shape function

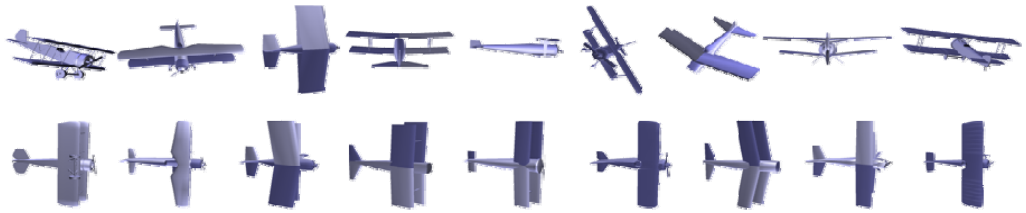
GM: geometric moments

CH: curvature histo

Retrieval results examples (UNIFI-MICC)

Enhanced Depth Buffer based descriptor (INRIA)

- Analysis of the existing 3D shape descriptors based on 2D views of 3D models. Depth buffer based approaches, where a 3D model is described by depth-buffer images.
- Implementation of “Continuous” PCA [Vranic 2004].



More robust than standard PCA.
Application of PCA on all points of the object’s surface

- Implementation of DBD (Depth Buffer-Based Descriptors) [Vranic 2004]
- Introduction of a new descriptor Enhanced DBD.

Enhanced Depth Buffer based descriptor (INRIA)

- Including in the descriptor the diversity of information contained in the depth buffer images results in a more precise representation.

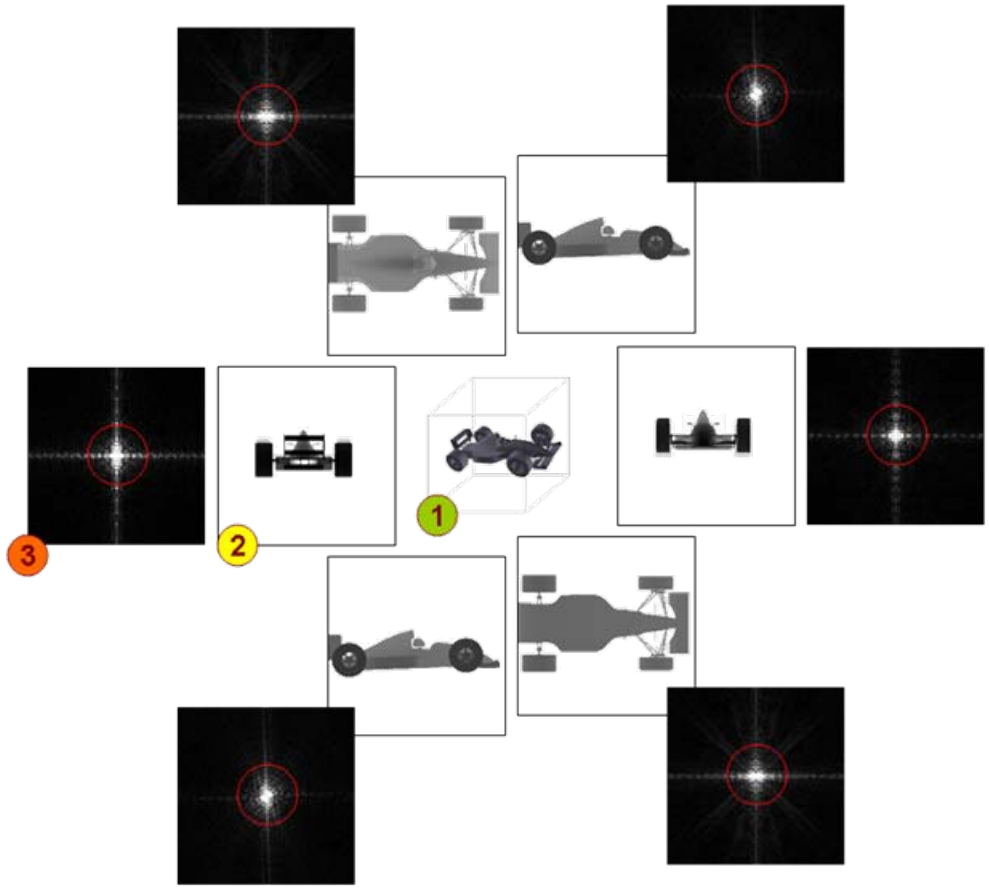
Extended Bounding Box

↓

6 Depth-Buffer Images

↓

FFT Transform



Enhanced Depth Buffer based descriptor (INRIA)

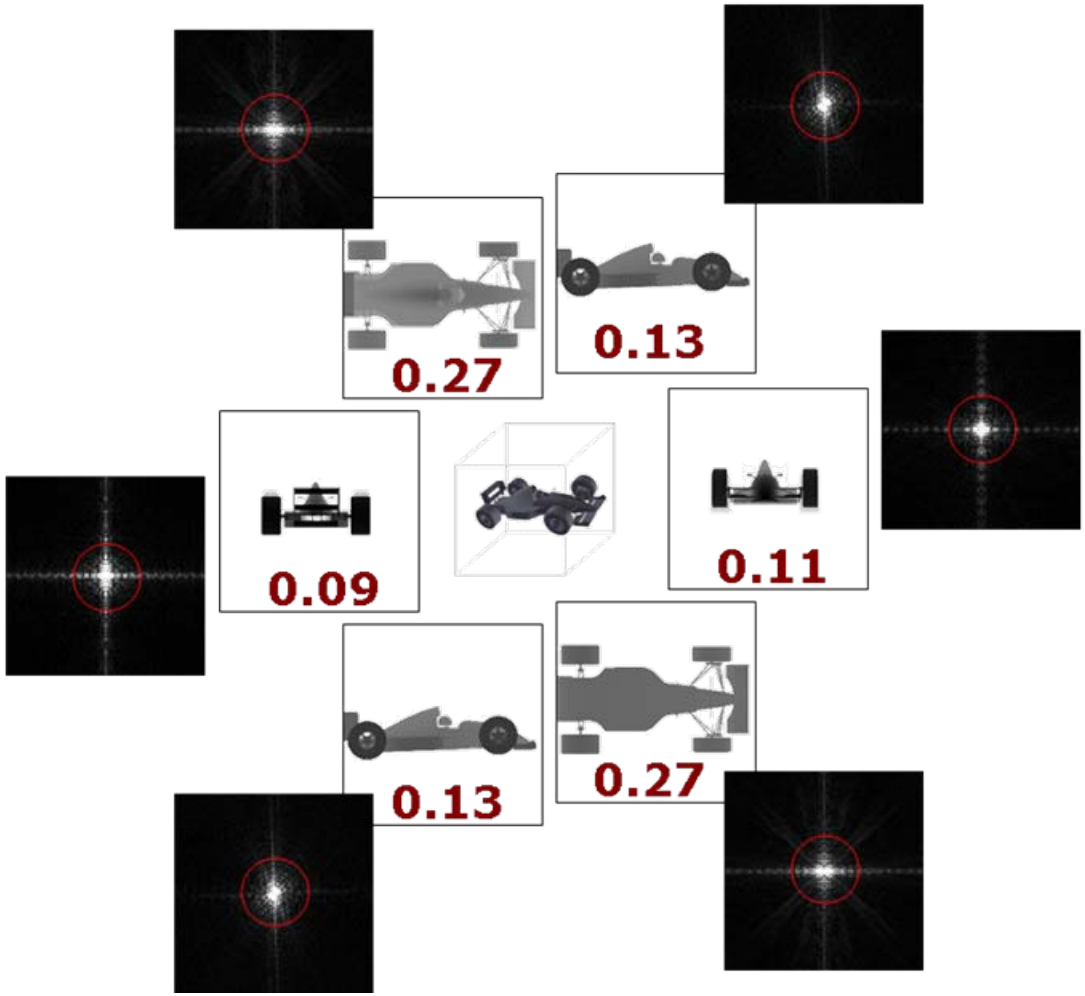
Relevance indices

↓

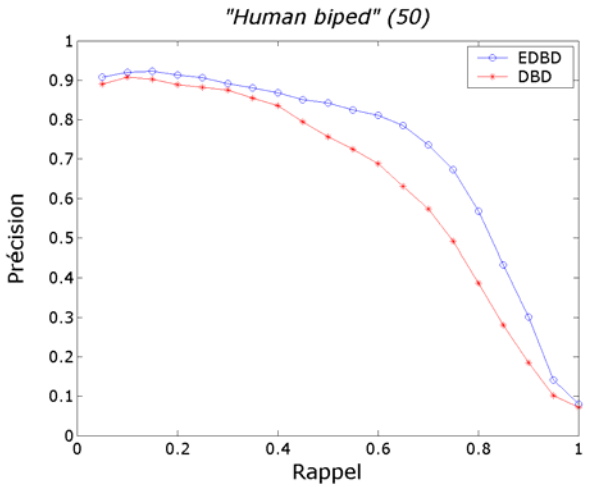
associated to depth buffer images

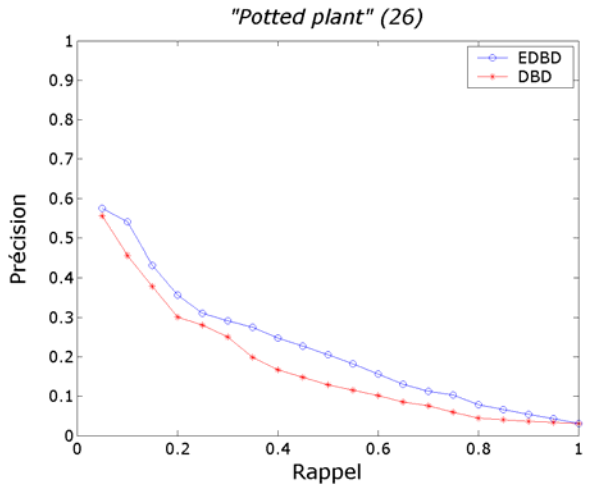
↓

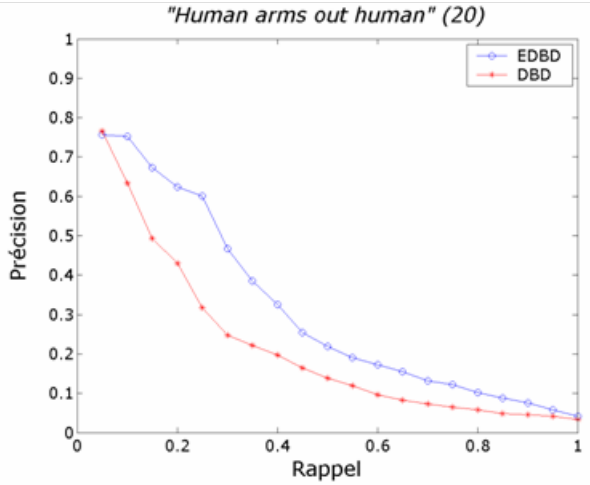
used to estimate the degree of similarity between 3D shapes

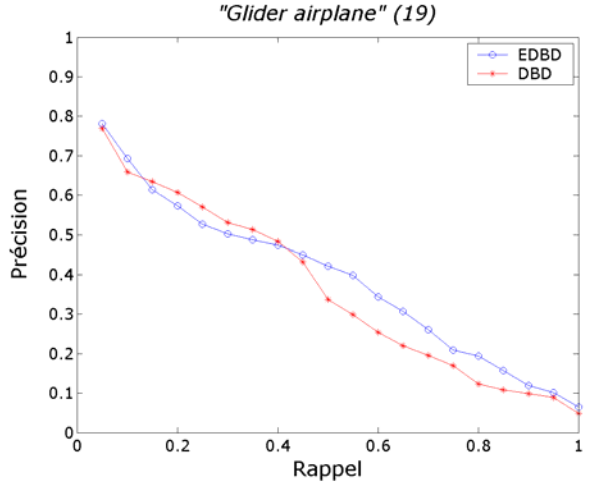


Precision–Recall on the Princeton Shape Benchmark (INRIA)









Retrieval result (INRIA)

DBD

✓ 0.0518	✓ 0.0546	✓ 0.0590
✓ 0.0603	✓ 0.0623	X 0.0625
X 0.0666	X 0.0672	X 0.0673

EDBD

✓ 0.0099	✓ 0.0099	✓ 0.0105
✓ 0.0107	✓ 0.0113	✓ 0.0114
X 0.0116	X 0.0117	✓ 0.0118

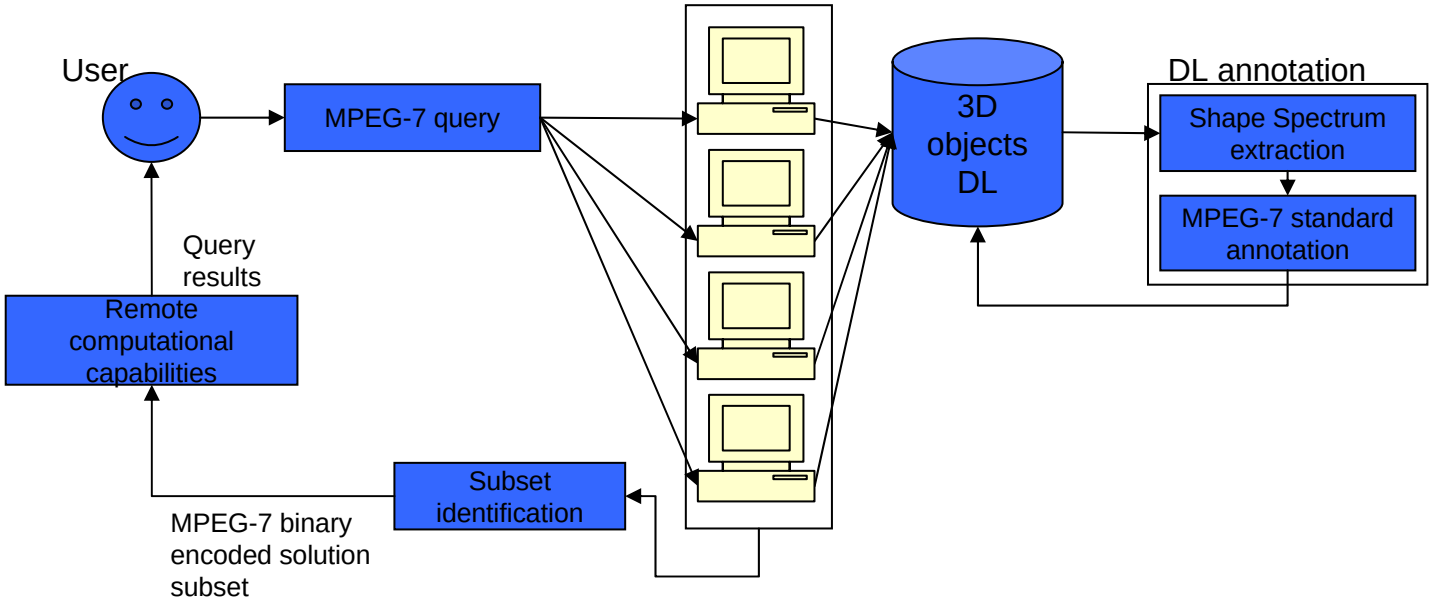
INRIA platform

DBD

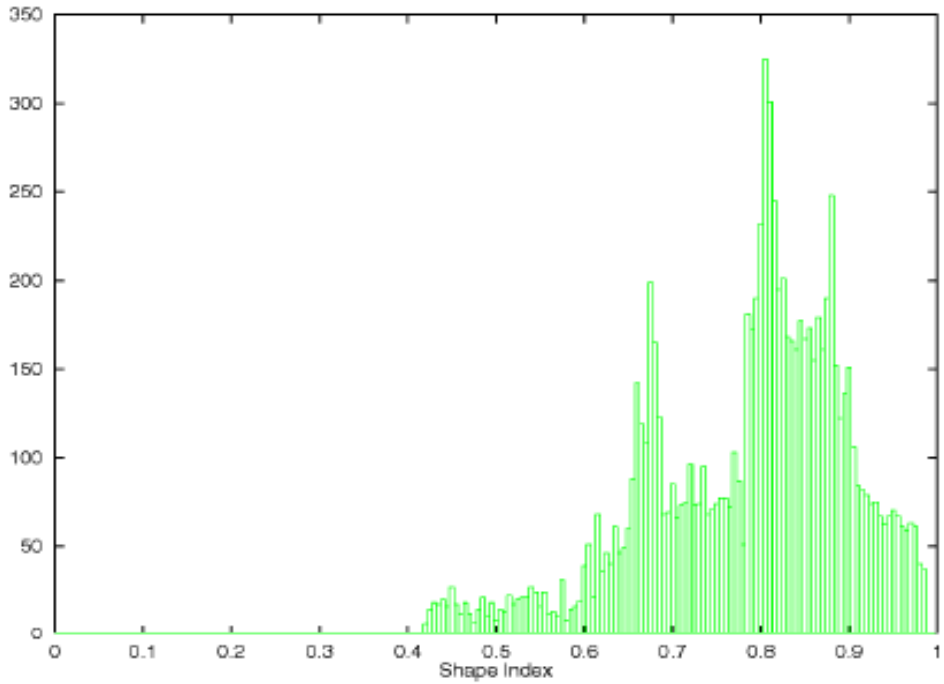
EDBD

MPEG-7 3D descriptor (UNIMORE)

- Fully compliant MPEG-7 description of 3D models.
- Test of the retrieval performance against.
- Implementation a web based system, which employs MPEG-7 descriptors and supports remote querying and retrieval with computationally distributed comparisons.



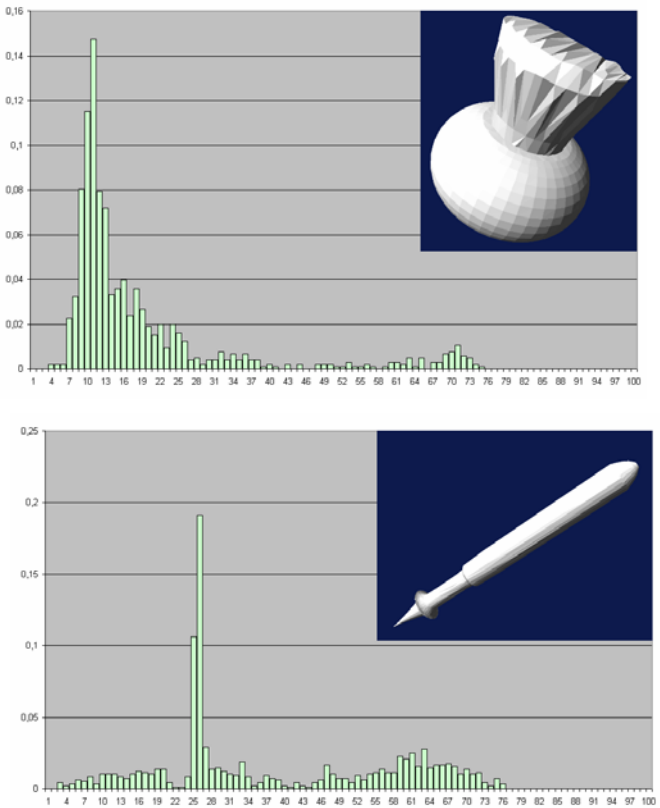
MPEG-7 shape descriptors (UNIMORE)



$$I_p = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_p^1 + k_p^2}{k_p^1 - k_p^2}$$

- Intrinsic shape description for 3D mesh models.
- Exploitation of local attributes of the 3D surface, extracted from the principal curvatures.
- The shape index value is in the interval [0,1] and is not defined for planar surfaces.
- The shape spectrum of the 3D mesh is the histogram of the shape indices calculated over the entire mesh.

Example: shape spectrums and MPEG-7 description (UNIMORE)



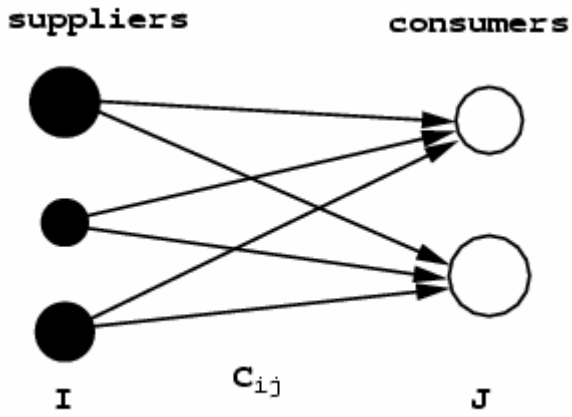
- Example MPEG-7 output description

```
<?xml version="1.0" encoding="UTF-8"?>
<Mpeg7 xmlns="urn:mpeg:mpeg7:schema:2001"
      :xsi="http://www.w3.org/2001/XMLSchema-instance"
      :mpeg7="urn:mpeg:mpeg7:schema:2001"
      :schemaLocation="urn:mpeg:mpeg7:schema:2001 Mpeg7-2001.xsd">
      <Description xsi:type="ContentEntityType">
            <MultimediaContent xsi:type="MultimediaType">
                  <Multimedia>
                        <MediaLocator>
                              <MediaUri>Statua3D.wrl</MediaUri>
                        </MediaLocator>
                        <MediaSourceDecomposition>
                              <Segment xsi:type="StillRegion3DType">
                                    <VisualDescriptor xsi:type="Shape3DType">
                                          <Spectrum>12 130 148 22 300</Spectrum>
                                          <PlanarSurfaces>0</PlanarSurfaces>
                                          <SingularSurfaces>0</SingularSurfaces>
                                    </VisualDescriptor>
                              </Segment>
                        </MediaSourceDecomposition>
                  </Multimedia>
            </MultimediaContent>
      </Description>
</Mpeg7>
```

Earth Movers distance for 3D object comparison (UNIMORE)

- Comparison of histograms with different number of bins.

- Perceptually significant distance needed: from the Transportation Problem the problem can be stated as: “find the least expansive flow of goods from supplier to consumer (each consumer-supplier chain has an associated cost to deliver one unit of goods)”.



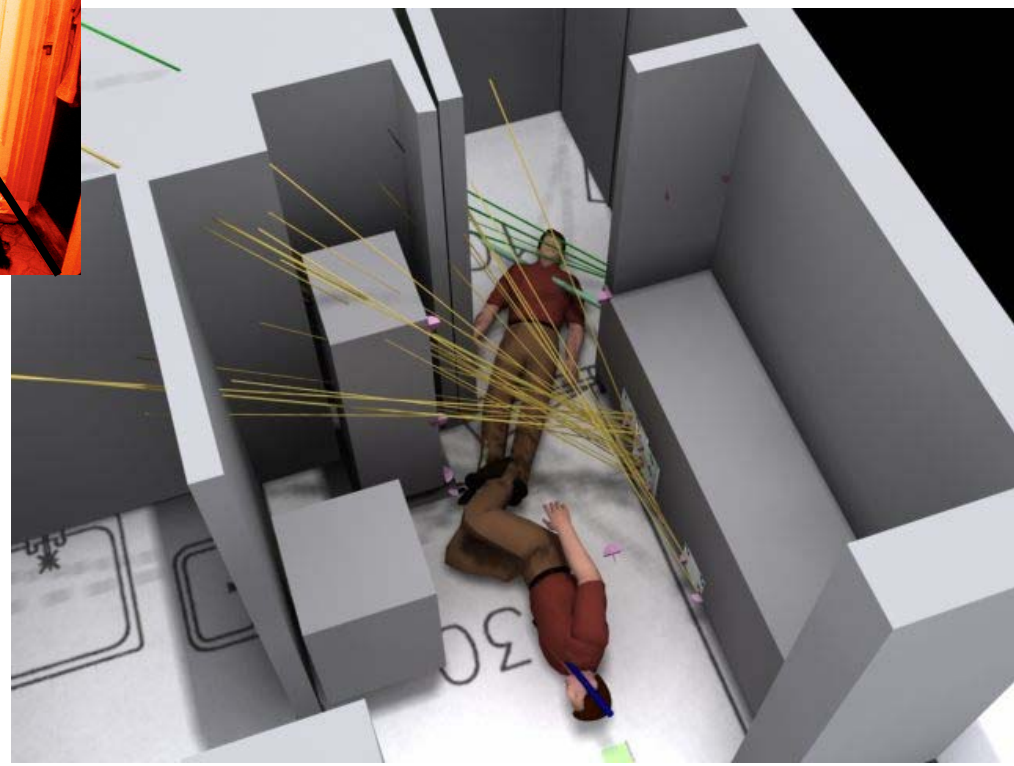
Current and future work

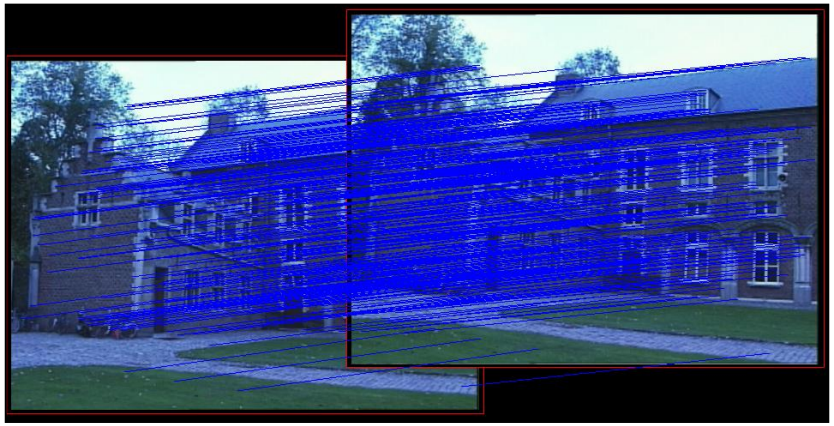
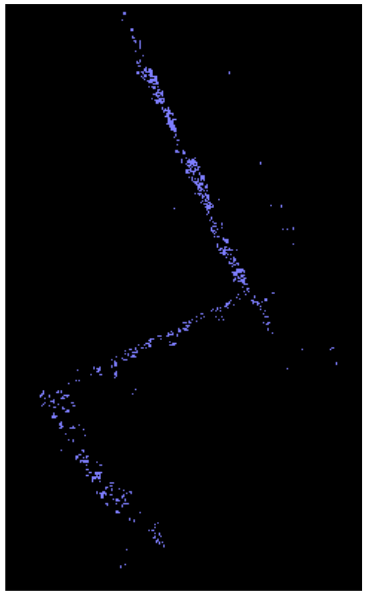
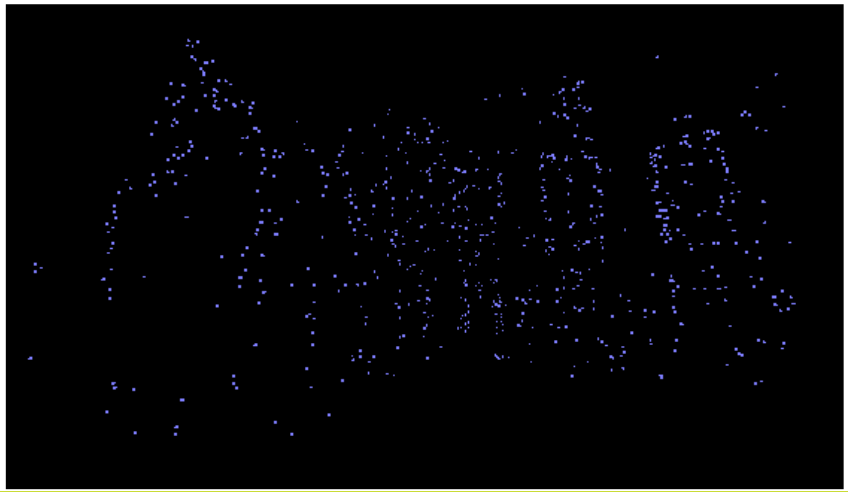
- **Database building**
  completed available locally and at the test-bed DELOS site
- **3D models descriptors**
  partially completed; enhancements and new solutions to be developed
- **Retrieval**
  experiments and comparisons to be developed over all the participants
- **Demonstrators**
  partially available; final versions and integration using a common interface to be developed

List of scientific publications available

3D models reconstruction (UvA)

- Goal: Reconstruction of 3D models from video data.



Simulated crime scene (UvA)

Preliminary results (UvA)