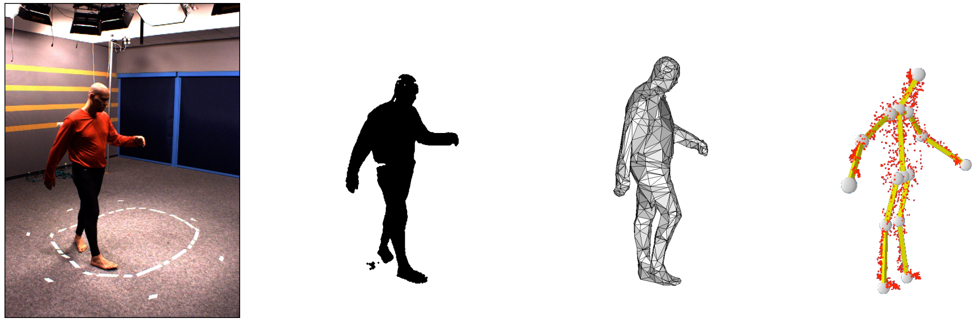
3D Scene Modeling Using Multiple Cameras

Edmond Boyer

Universités de Grenoble - INRIA Rhone-Alpes



Context

- Computer Vision: to give visual analysis capabilities to computers.

- 3D Scene Modeling: analysis of 3D scenes composed of real objects, possibly moving and deforming
  1. Shape Modeling.
  2. Motion Modeling.
  3. Motion Semantic Modeling (e.g. actions, activities).

- Applications :

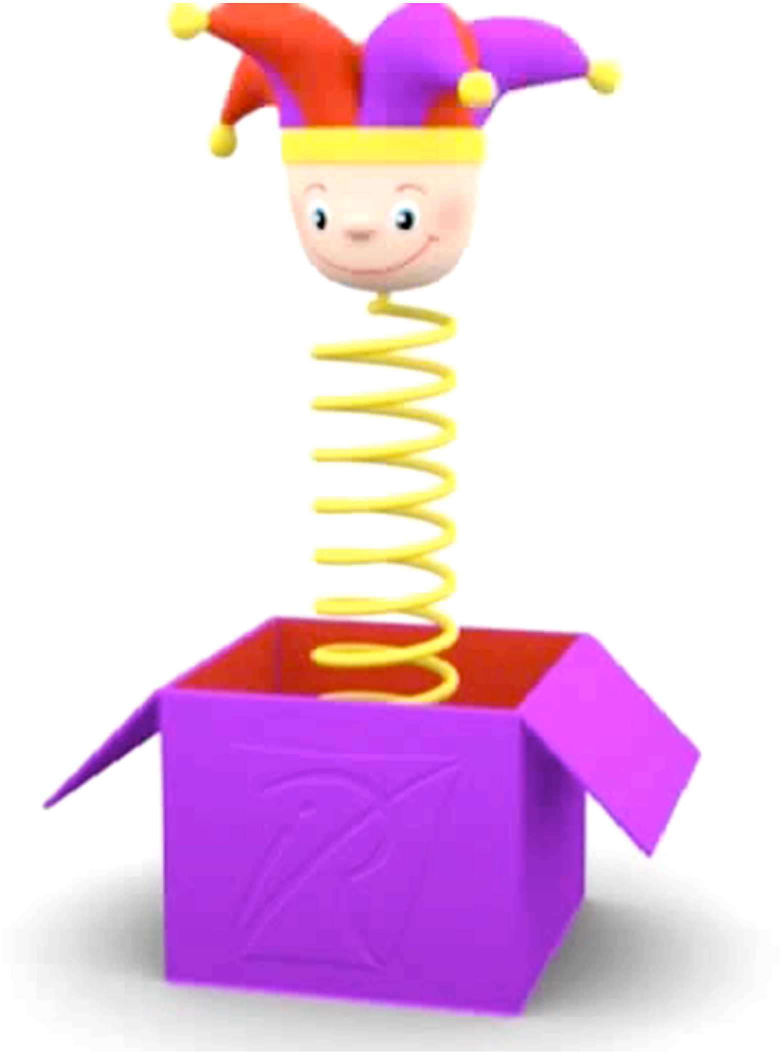
  3D Contents Production: TV3D, Virtual Reality, Virtualization, Interactions.
  Intelligent Environments: smart rooms,
  Surveillance.

Example : shape modeling

Example : motion modeling

Tony Hawk

Example : interactions



Outline

- Shape Modeling:
  - Acquisition Systems.
  - Geometric Approaches.
  - Photometric and geometric Approaches.

- Motion Modeling:
  - Marker Based Approaches.
  - Marker less Approaches.

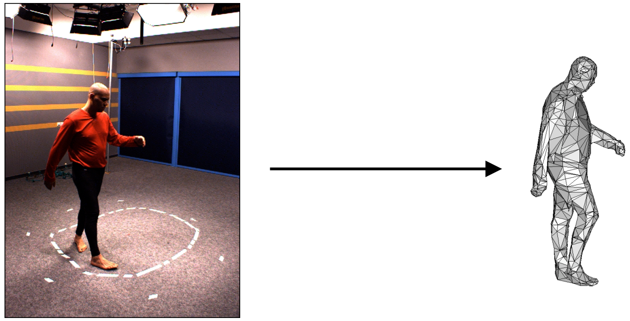
- Action Modeling:
  - Model Based Approaches.
  - Holistic Approaches.

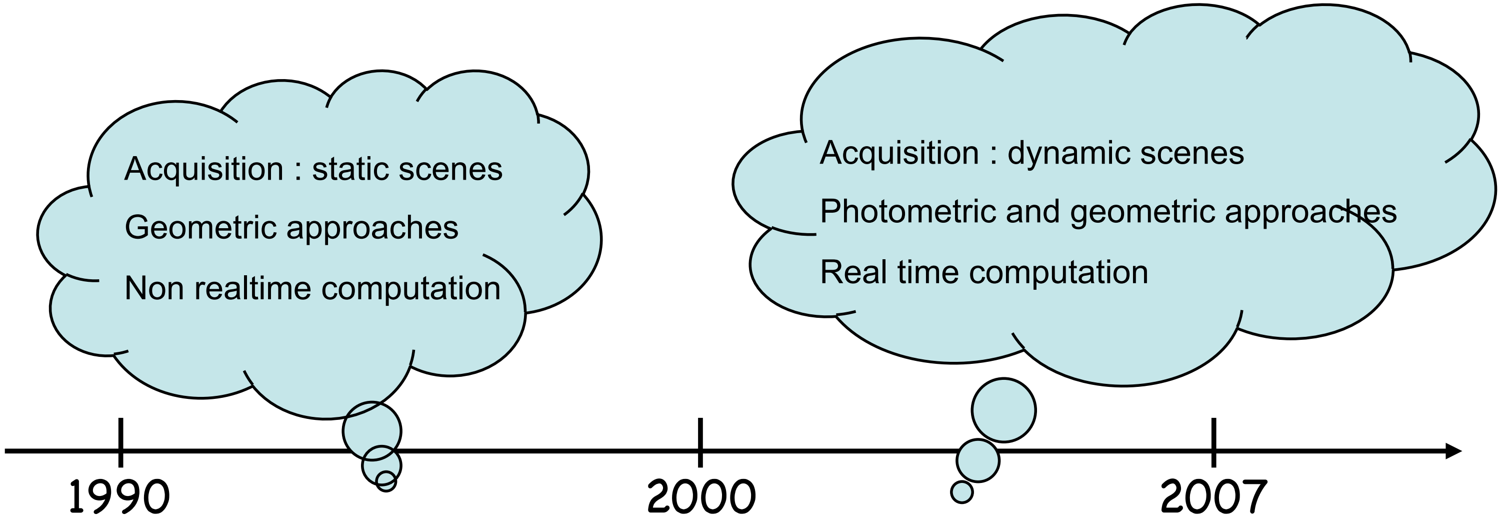
Outline

- Shape Modeling:
  - Acquisition Systems.
  - Geometric Approaches.
  - Photometric and geometric Approaches.
- Motion Modeling:
  - Marker Based Approaches.
  - Marker less Approaches.
- Action Modeling:
  - Model Based Approaches.
  - Holistic Approaches.

Shape Modeling

- Issues:
  - Model Quality.
  - Data acquisition.
  - Complexity.





Acquisition : static scenes

Geometric approaches

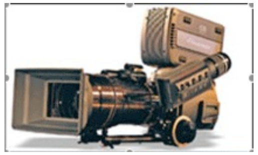
Non realtime computation

Acquisition : dynamic scenes

Photometric and geometric approaches

Real time computation

Acquisition Systems



HD Cameras

+ quality
- cost, complexity

USB Cameras

+ cost, complexity
- quality

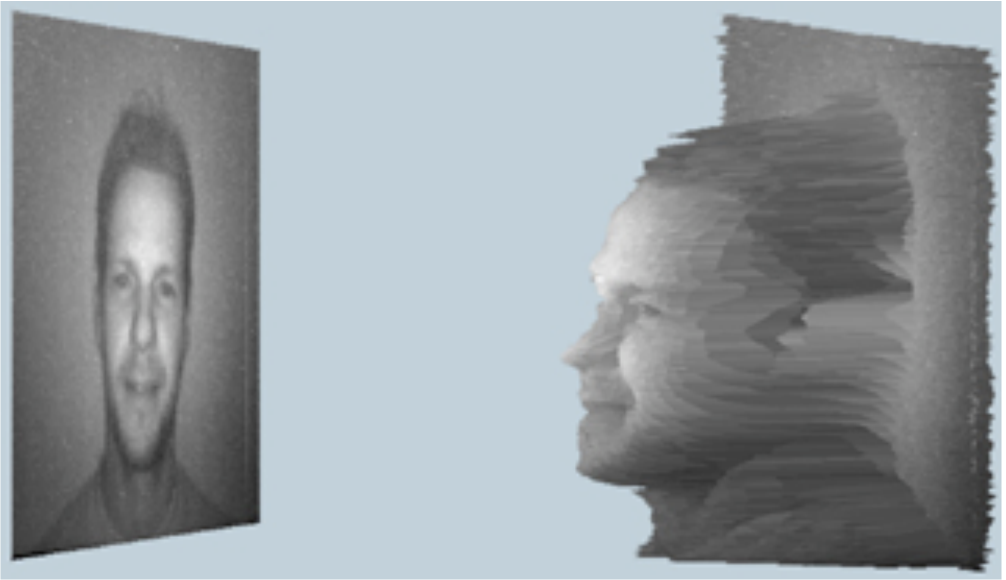
Firewire Cameras

+ cost, complexity
-+ quality

Issues:

- Synchronisation: improve significantly the modeling.
- Calibration: complex with lots of cameras.
- Environnement: a controled environnement is frequently required.
- Computation: important power computation required for complex models or real time applications.

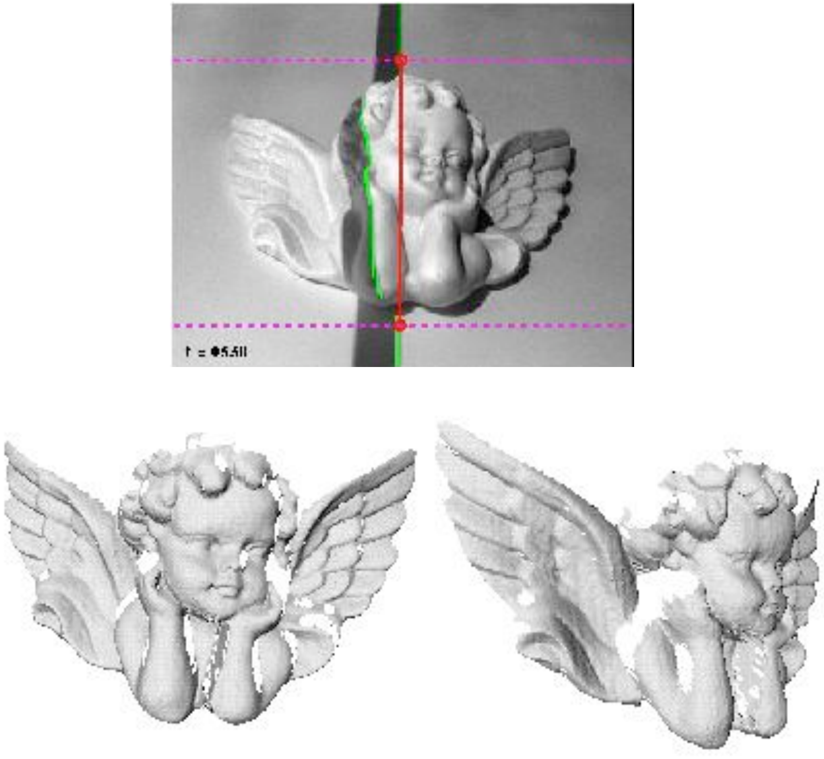
2D + Depth Systems



Time of flight cameras
or Stereo cameras ≠ 3D Modeling

Structured Light Systems

Technical University of
Braunschweig
Institute for Robotics and
Process Control



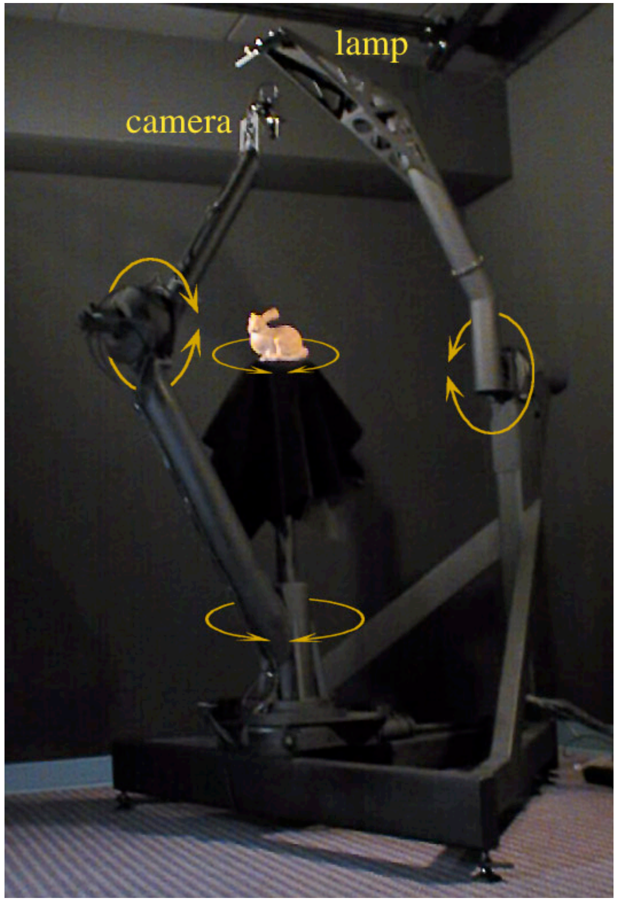
Jean-Yves Bouguet and Pietro Perona
Caltech

Turn Tables

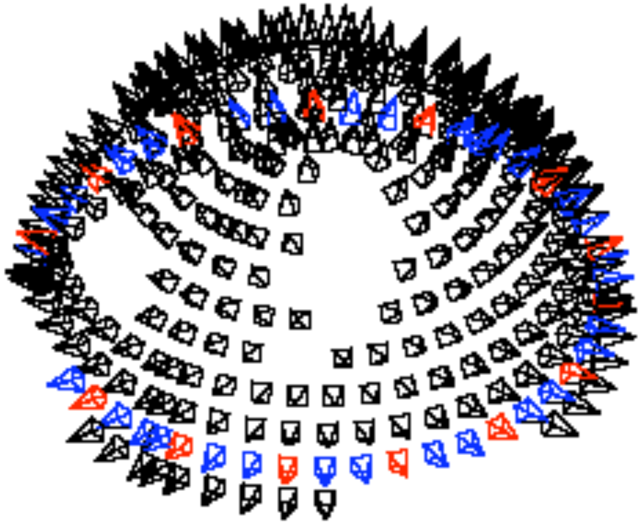
[Ortery technologies]

- Single Camera
- Calibration simplified
- Static System

Acquisition System



Stanford Spherical Gantry
http://www-graphics.stanford.edu/projects/gantry/



-> Middlebury Univ. 3D Datasets

Grimage Platform

- **Acquisition** : 6-10 caméras FireWire (SONY DFW-VL500, DFW-X700, AVT Marlin). Synchronisation hardware (trigger signal).

- **Visualisation :** wall screen (2mx2.7m, 4096x3072 pixels) 16 projectors Sanyo LCD.

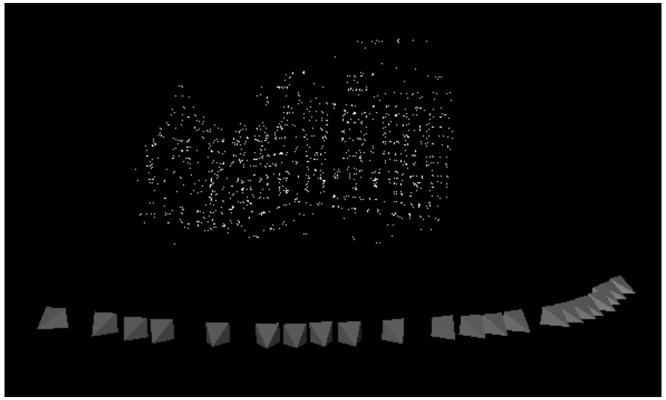
- **Computation:** PC cluster, 11 dual-xeon et 16 dual opteron connected through a double Ethernet gigabit.

- **Mobile Version:** 4-6 Firewire cameras, 6 mini-PCs and 1 laptop.

Geometric Modeling

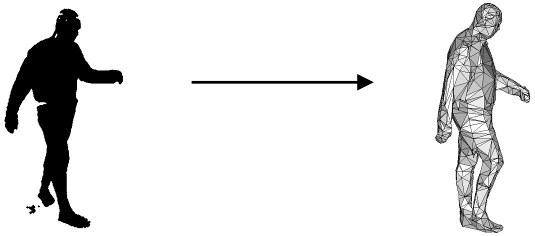
- No *a priori* knowledge
  otherwise matching between model and observations.

- 2D primitive extraction:
  - points, lines -> 3D point clouds

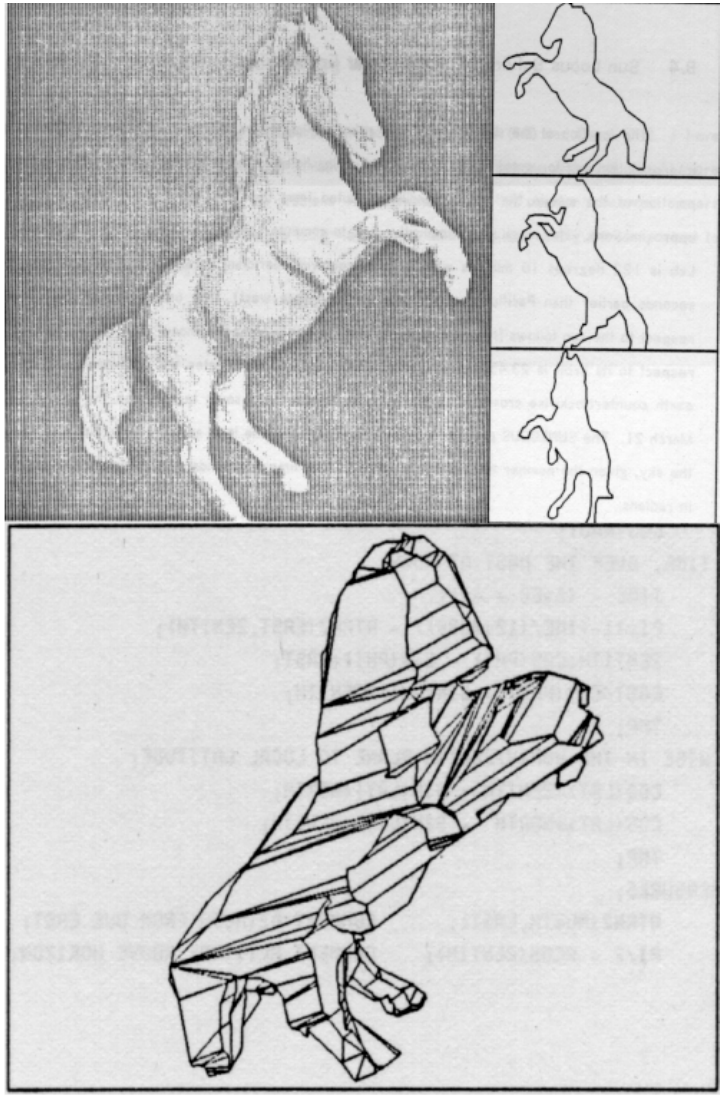


[M. Pollefeys, 1998]

  - Régions (silhouettes) -> surfaces, volumes

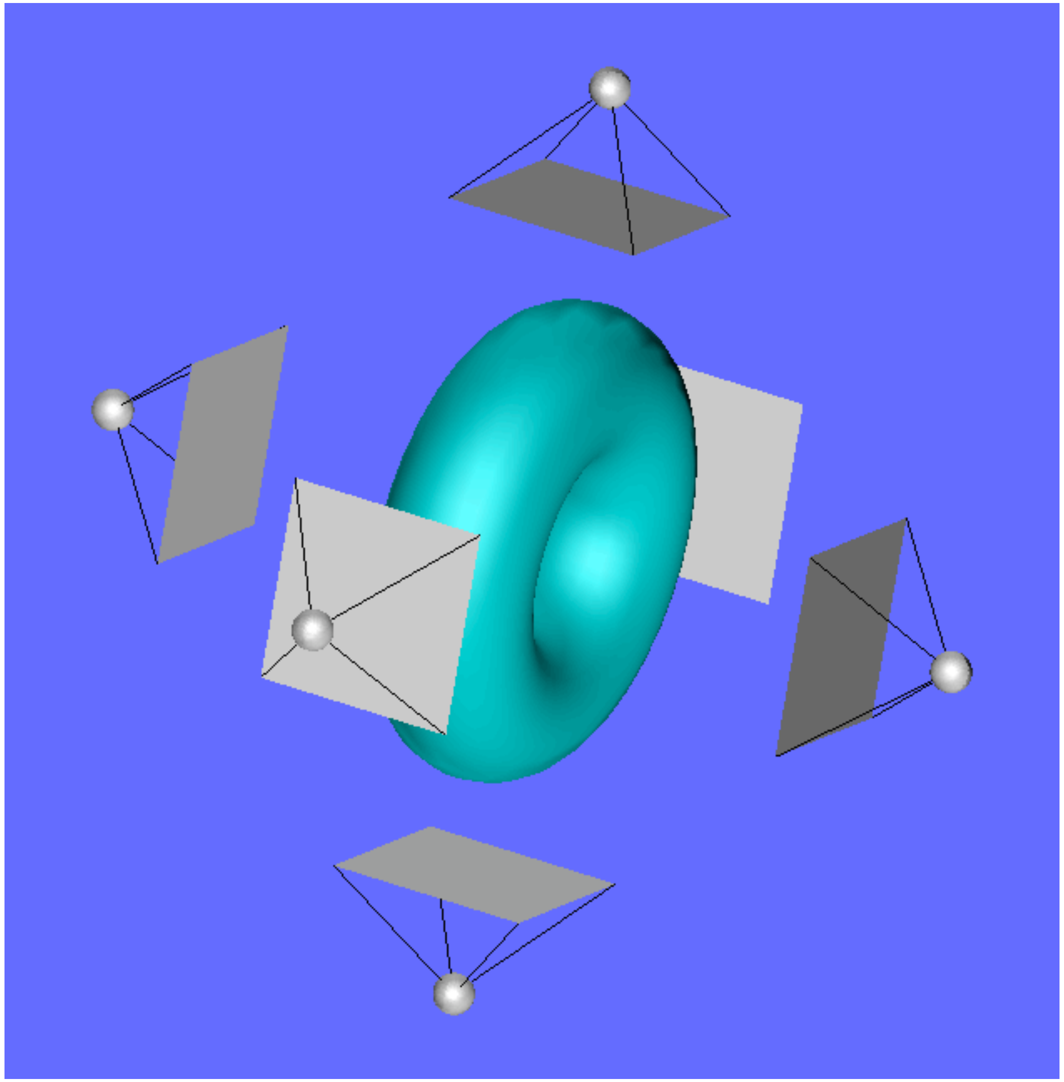


Silhouette Based Approaches

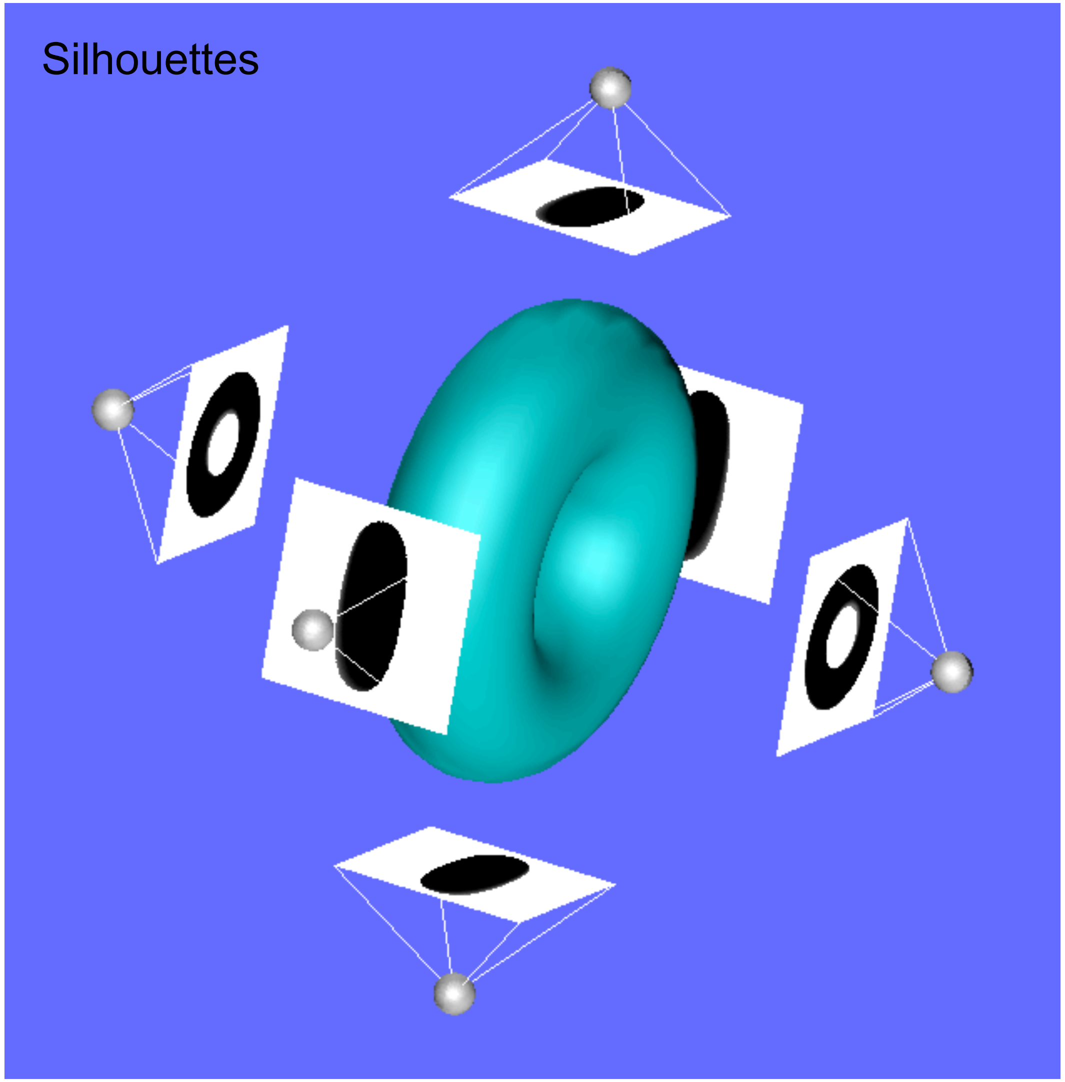


[B.W. Baumgart 1975, Stanford]

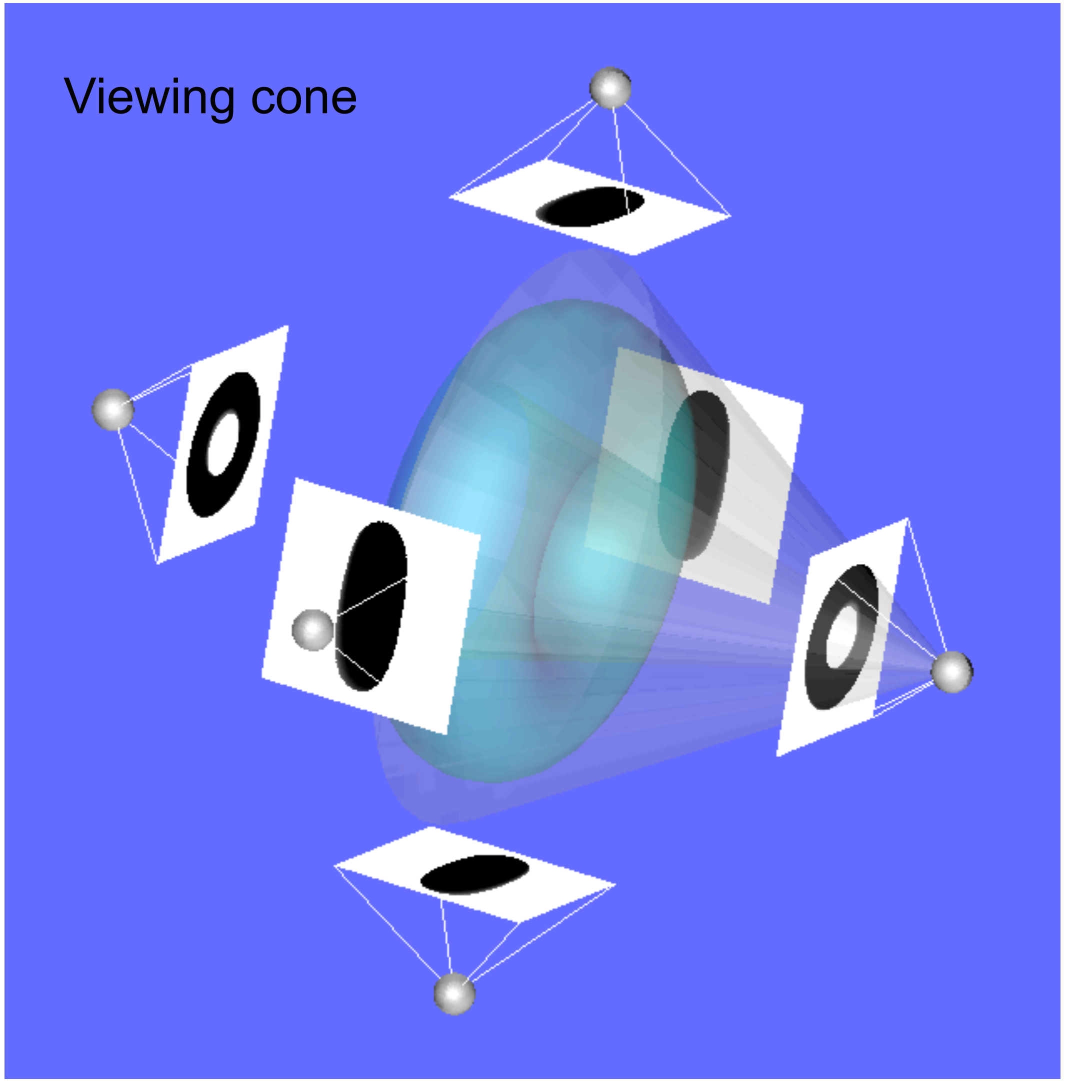
Silhouette Based Approaches principle



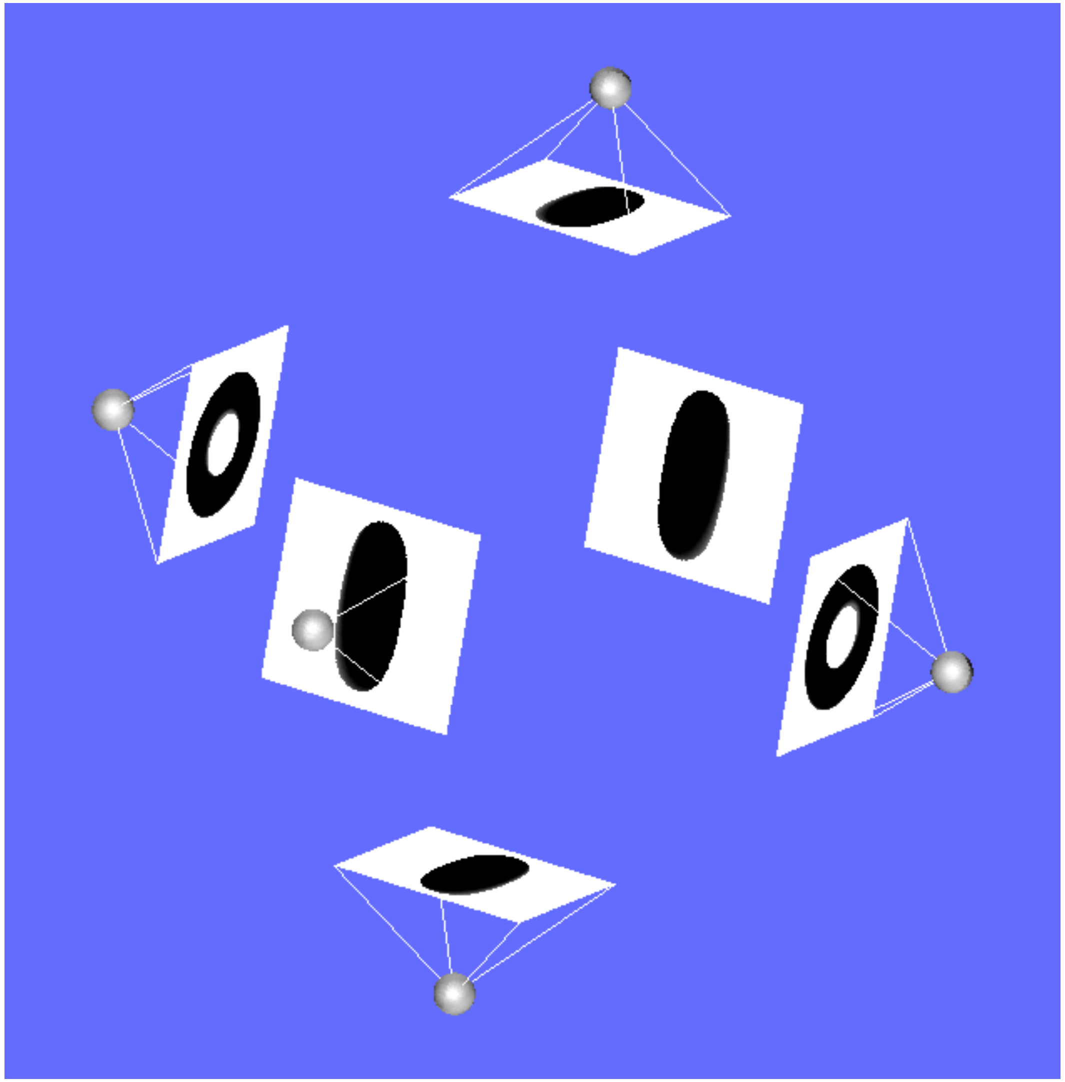
Silhouette Based Approaches principle



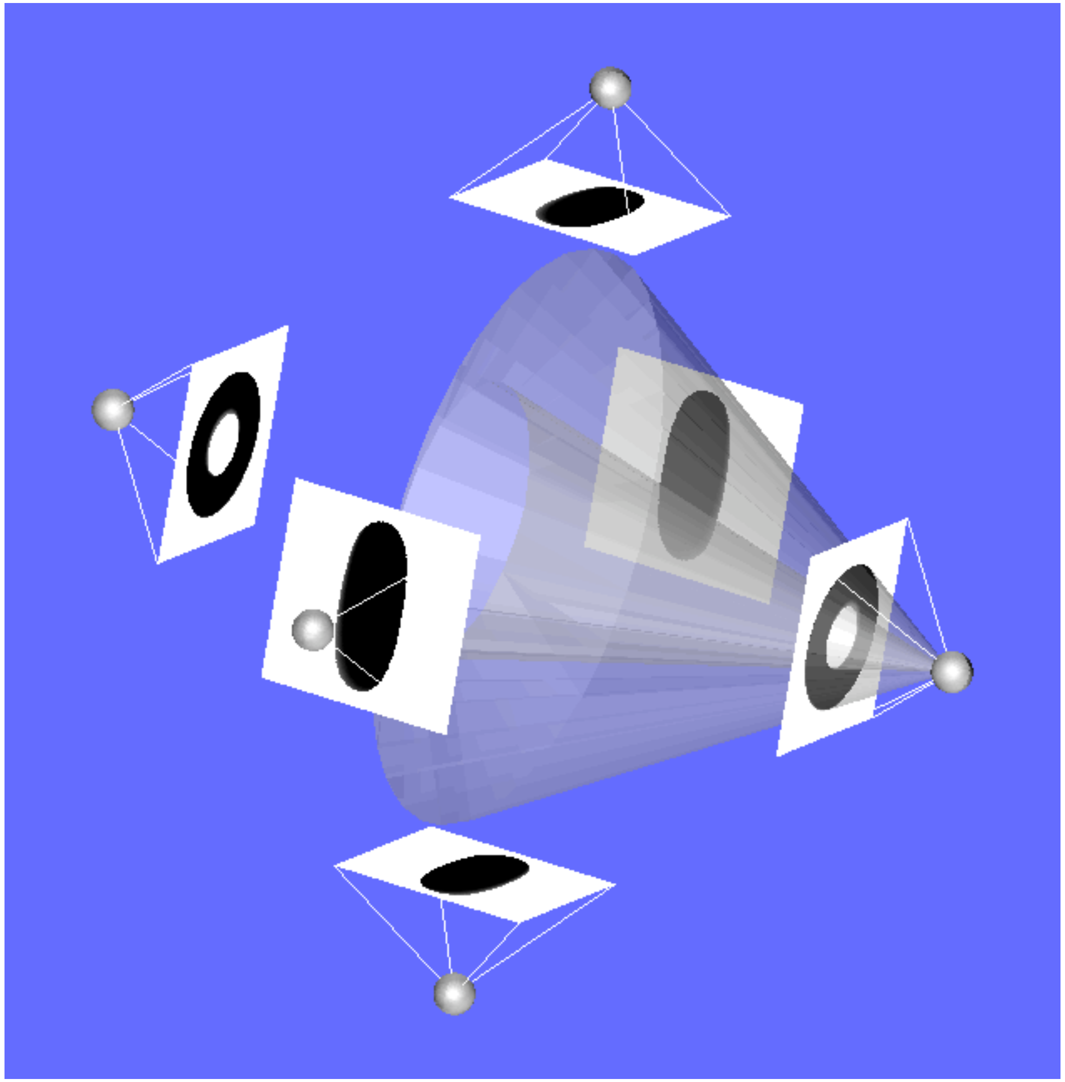
Silhouette Based Approaches principle



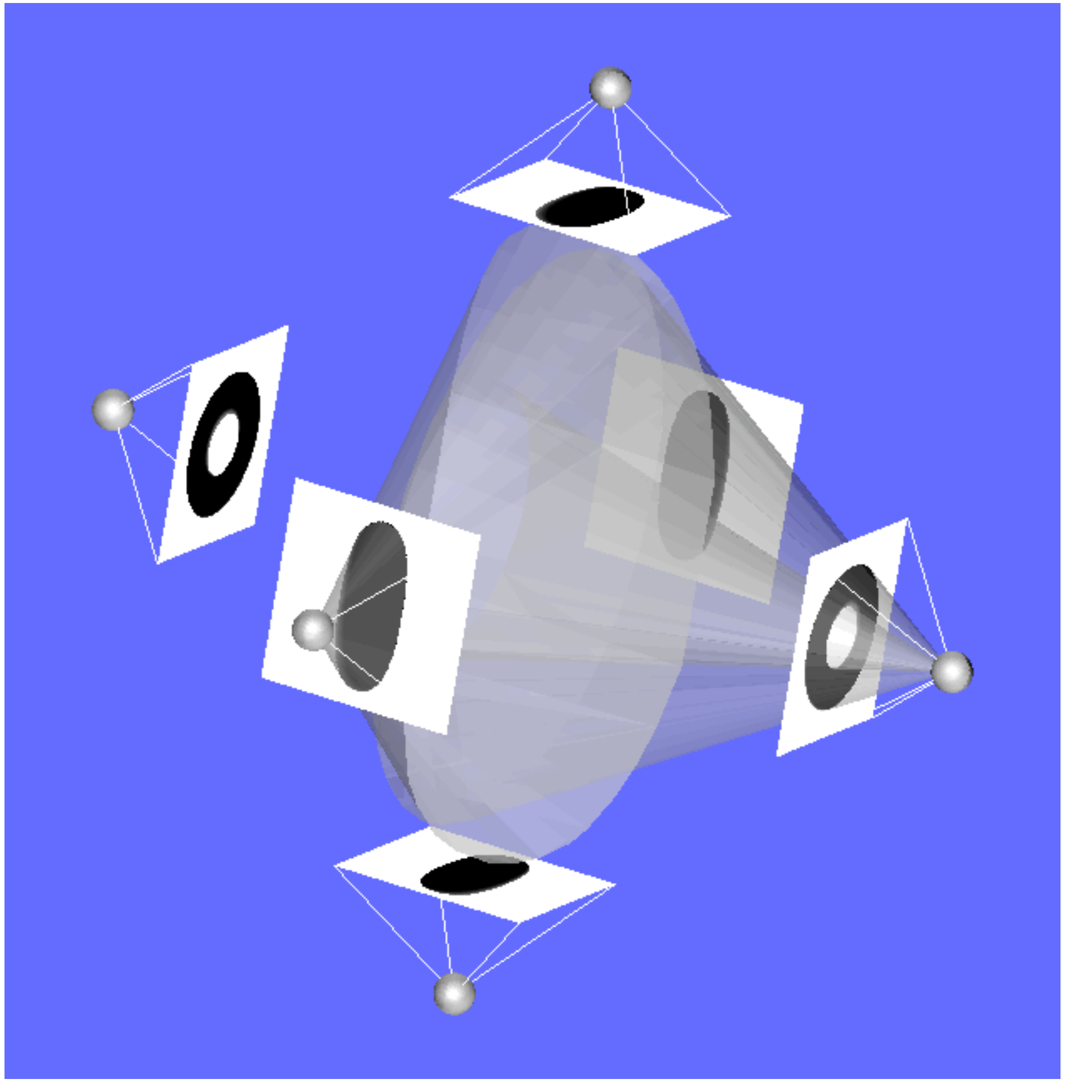
Silhouette Based Approaches principle



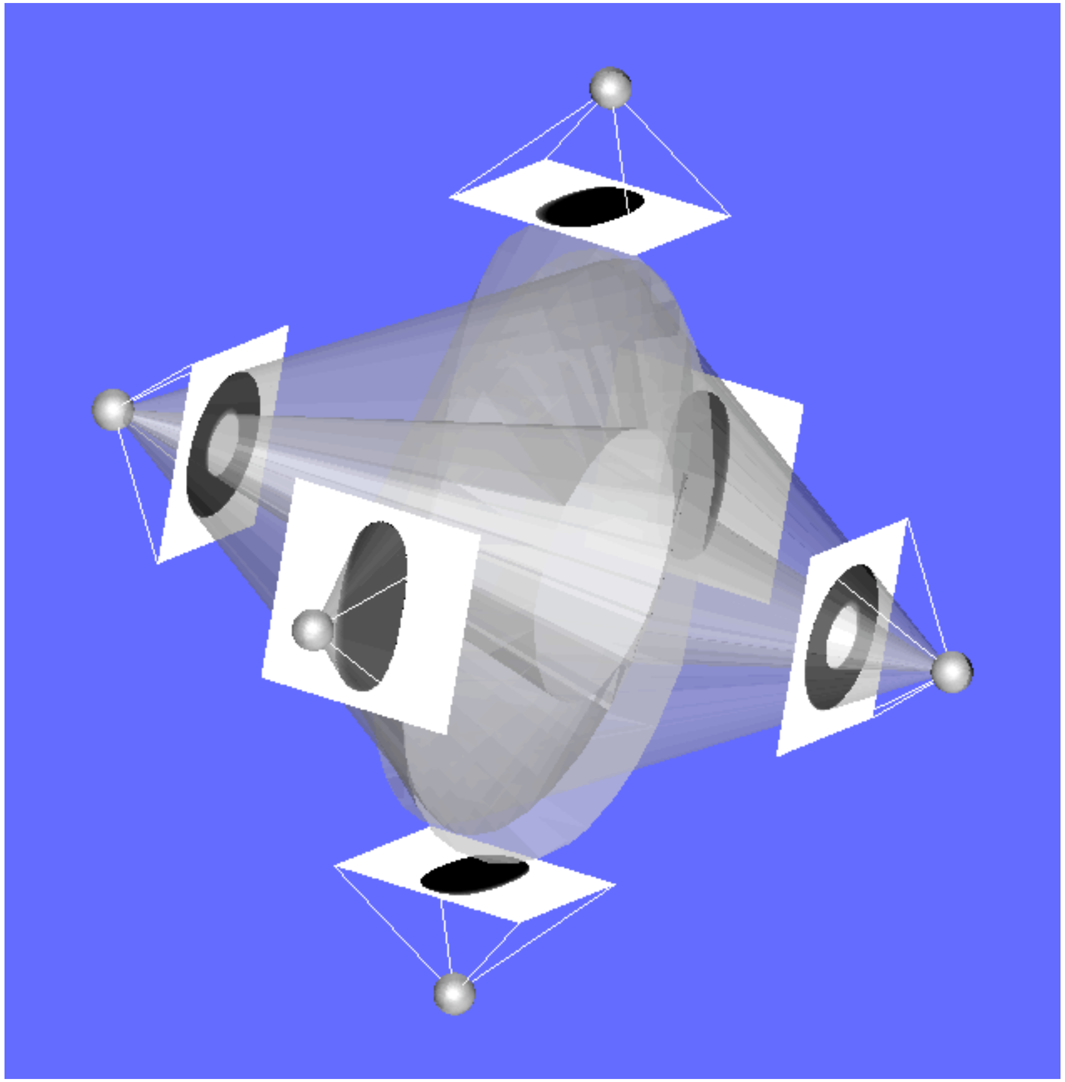
Silhouette Based Approaches principle



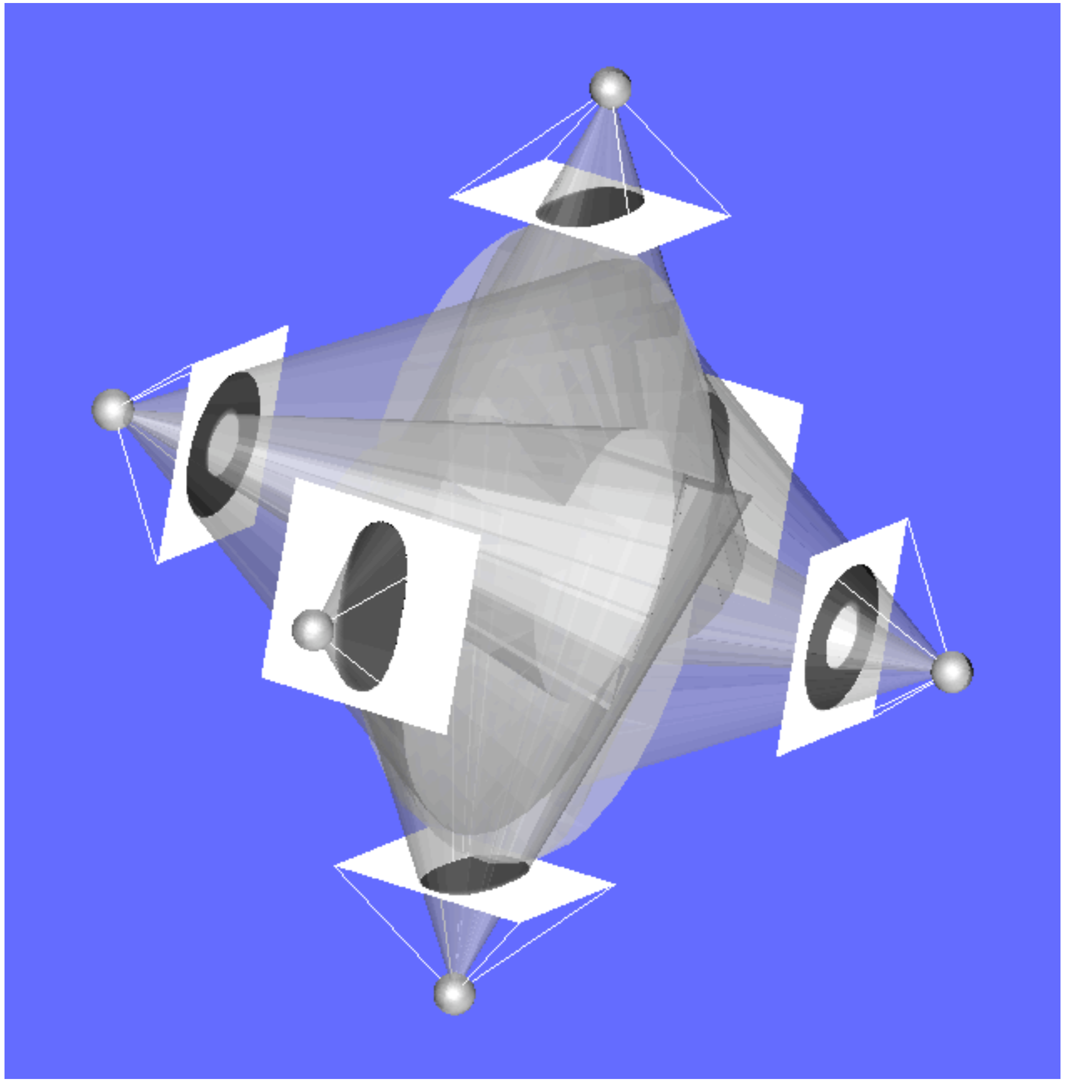
Silhouette Based Approaches principle



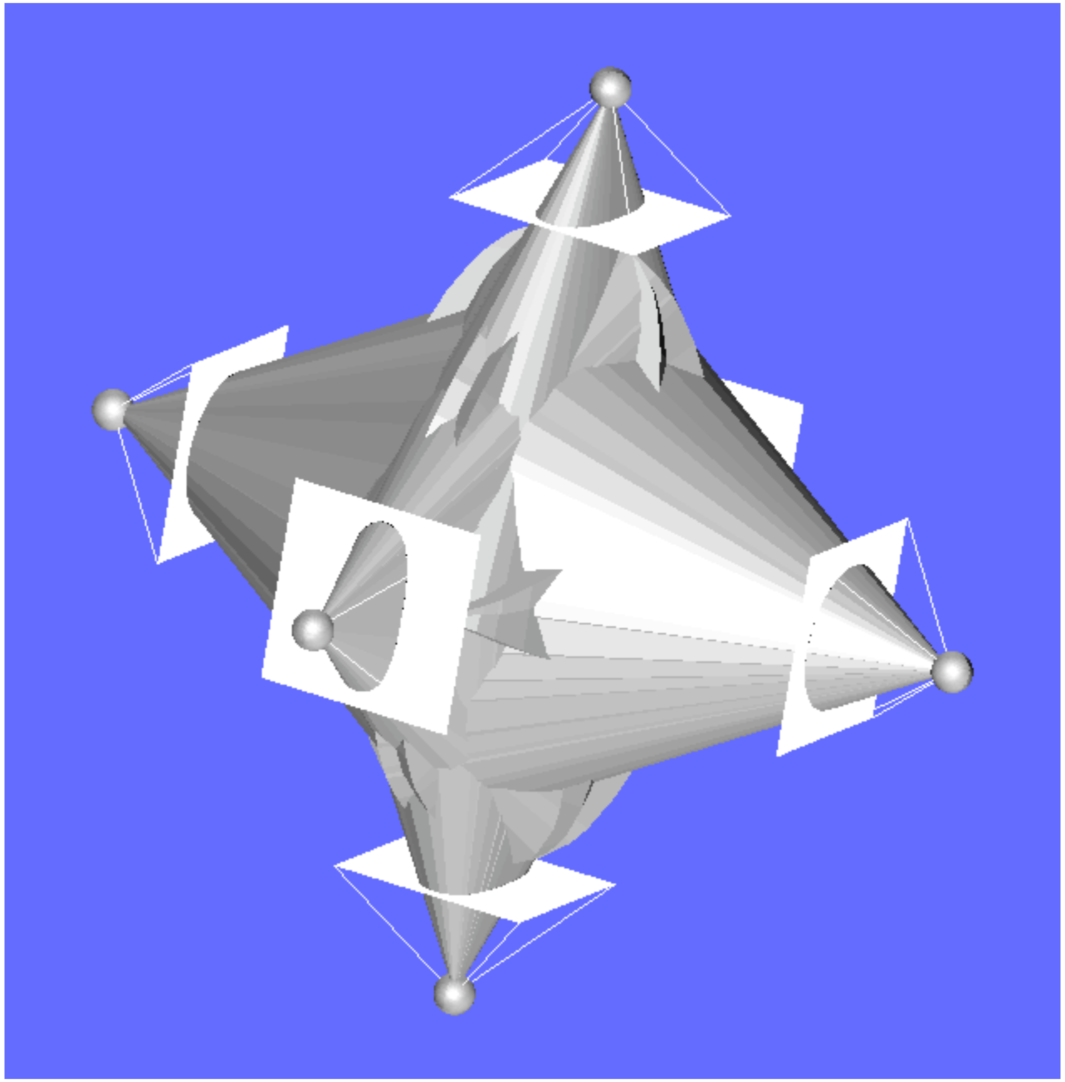
Silhouette Based Approaches principle



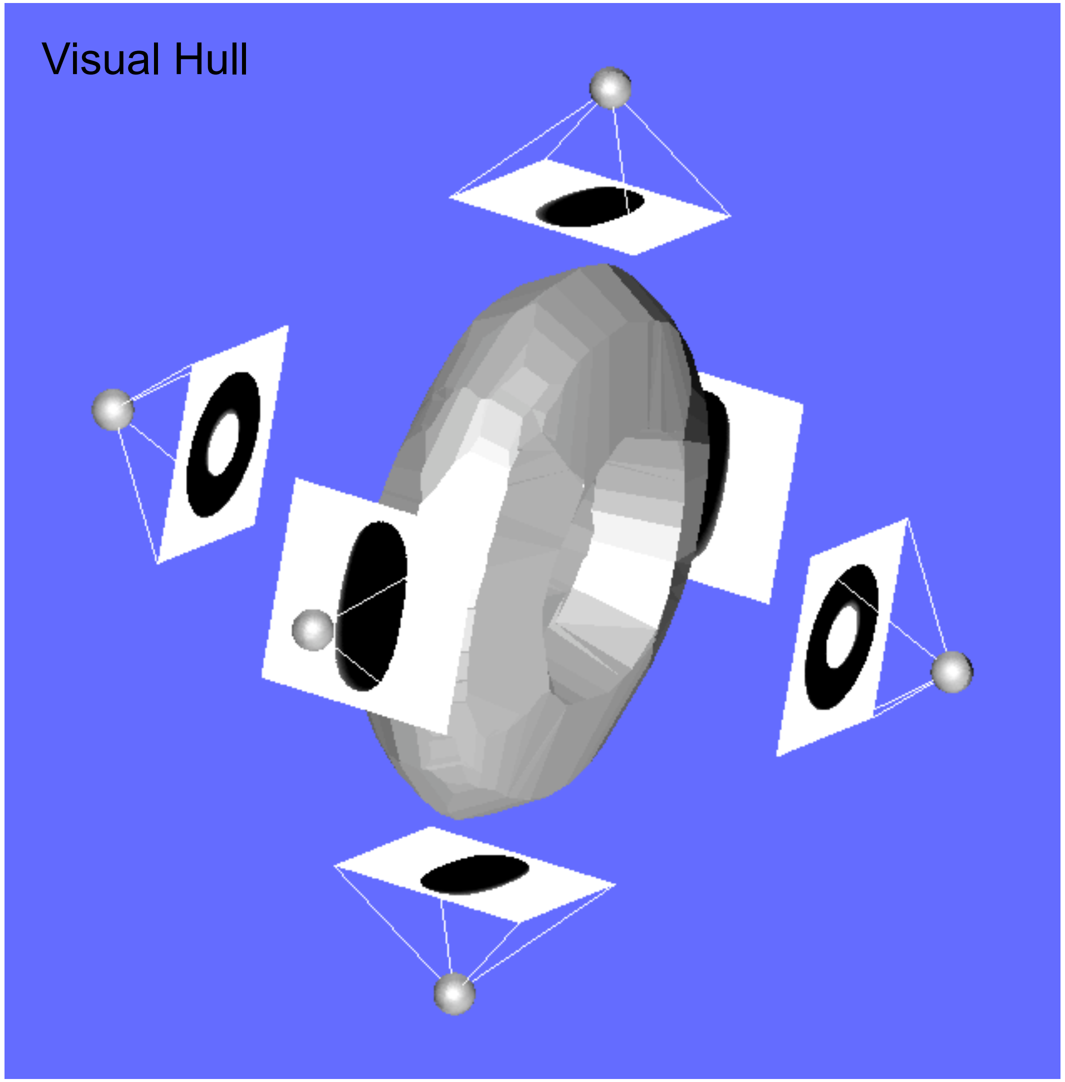
Silhouette Based Approaches principle



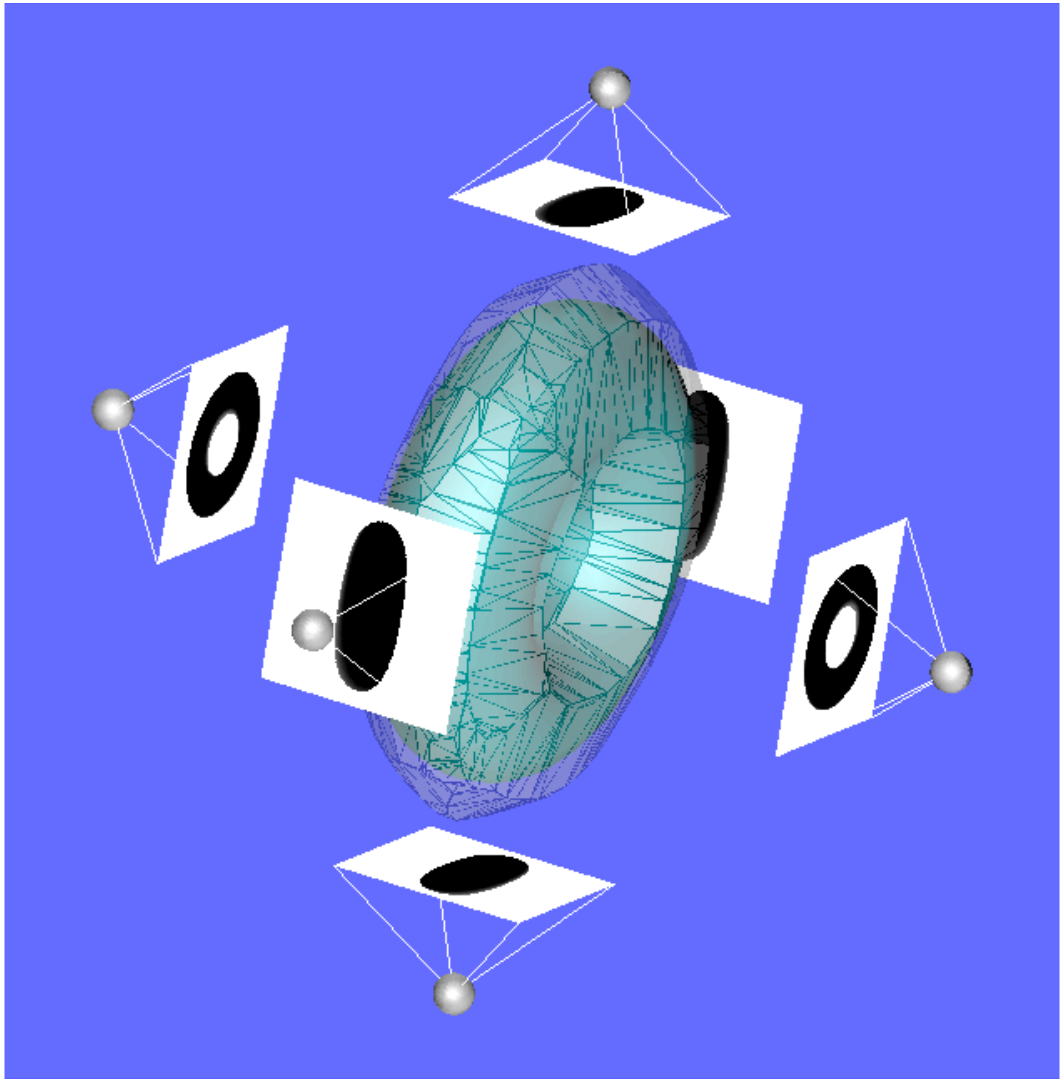
Silhouette Based Approaches principle



Silhouette Based Approaches principle



Silhouette Based Approaches
principle

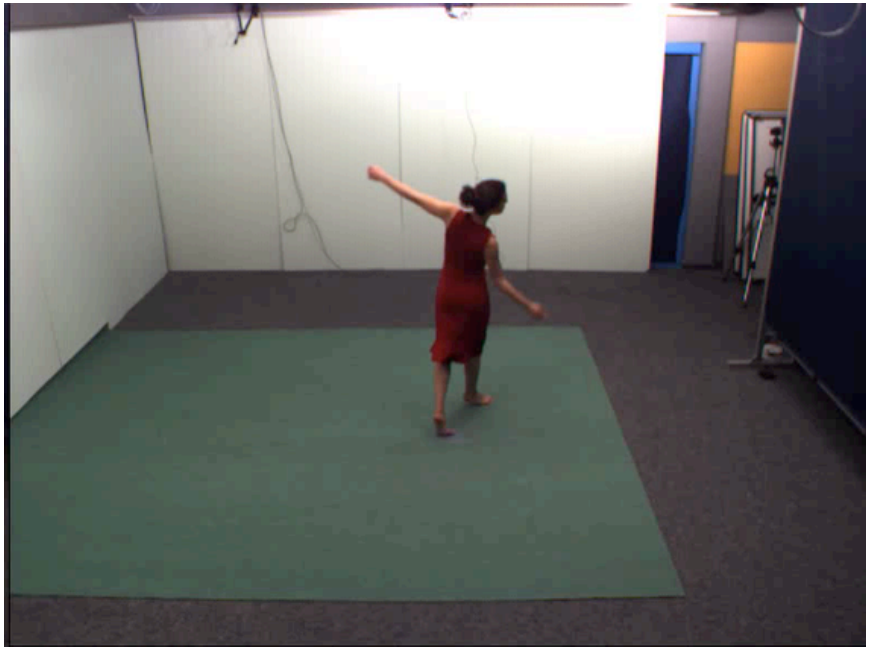


Geometric Modeling
Visual Hull

Formal Definition [Laurentini'94] : The visual Hull is the maximal volume coherent with (that entirely projects onto) a set of silhouettes.

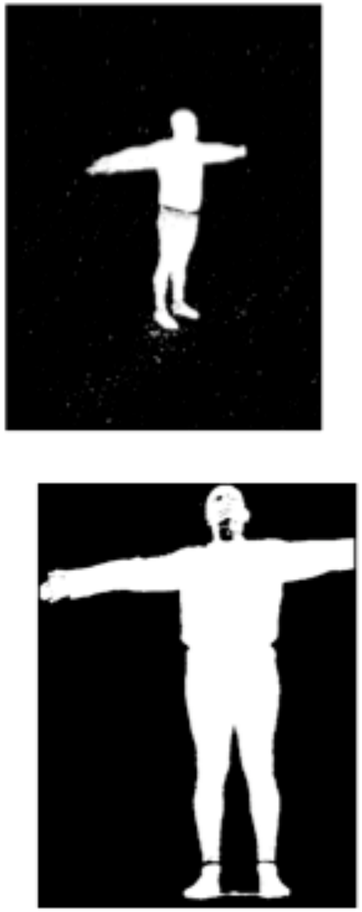
- The visual hull is the intersection of viewing cones.
- In the case of infinite viewpoints outside the convex hull of the observed objects, the visual hull corresponds to the observed surfaces without their concavities.
- In the discrete case (polygonal silhouettes, finite nb of viewpoints), the visual hull is a polyhedron.

Geometric Modeling
Visual Hull



[Grimage 2005]

Geometric Modeling
Silhouette Extraction



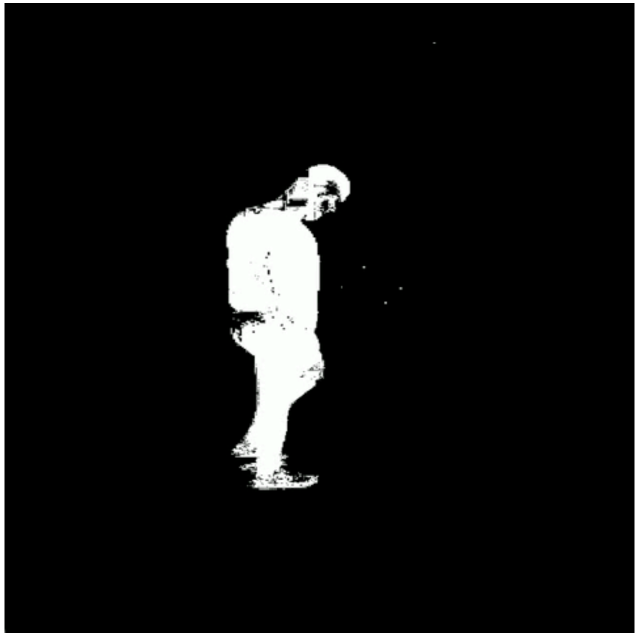
- Silhouettes are regions in the images where object of interest project.
- Silhouettes are estimated using low-level processes.
- Silhouettes give information on the observed surfaces.

Geometric Modeling
Silhouette Extraction

- Segmentation :
  - Manual.
  - Automatic :
    - Contour tracking
    - Chroma keying (blue or green background != skin color)
    - Background subtraction (static background)

Geometric Modeling
Silhouette Extraction

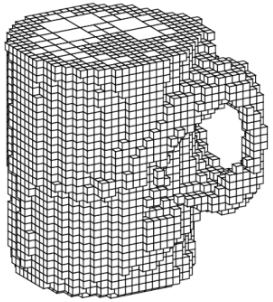
- Background subtraction:
  - Statistical background model
    - Gaussian
    - Gausian mixtures
    - Non parametric: histograms.
- Issues:
  - Image digitalization (noise);
  - Color ambiguities between background and foregroud objects;
  - Luminosities changes, etc.

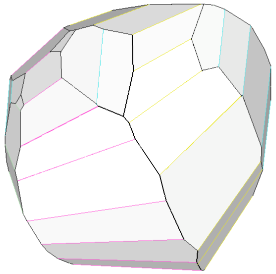


Geometric Modeling
Estimating the visual hull

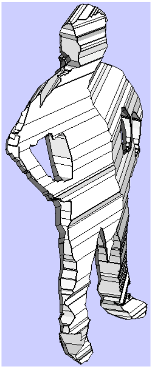
Volumetric Approaches

– voxels





Surface based Approaches



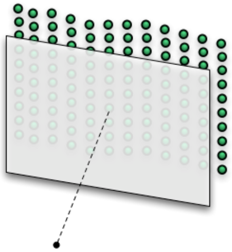
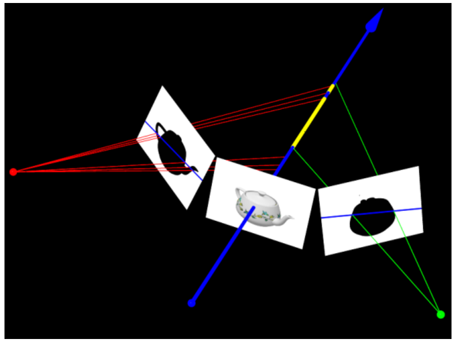
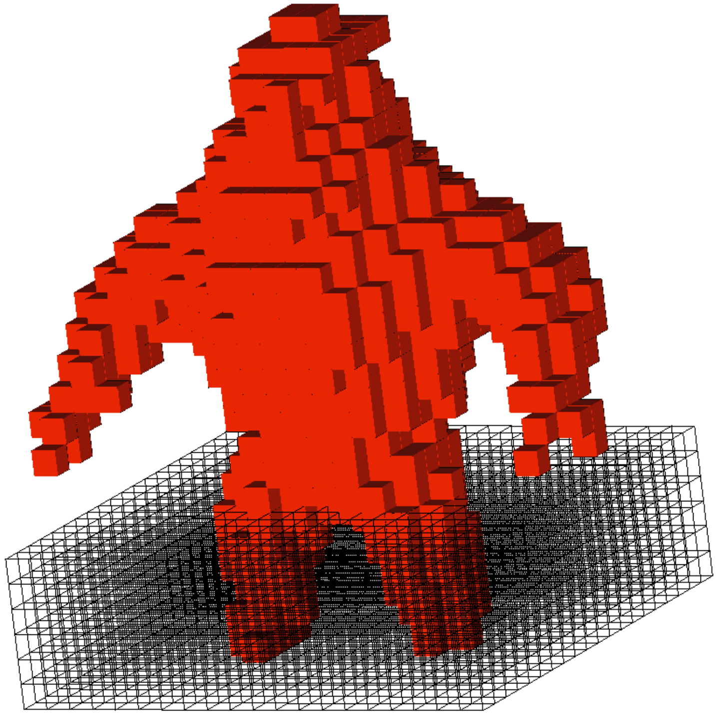


Image based Approaches



Geometric Modeling-VH
Voxel based approaches

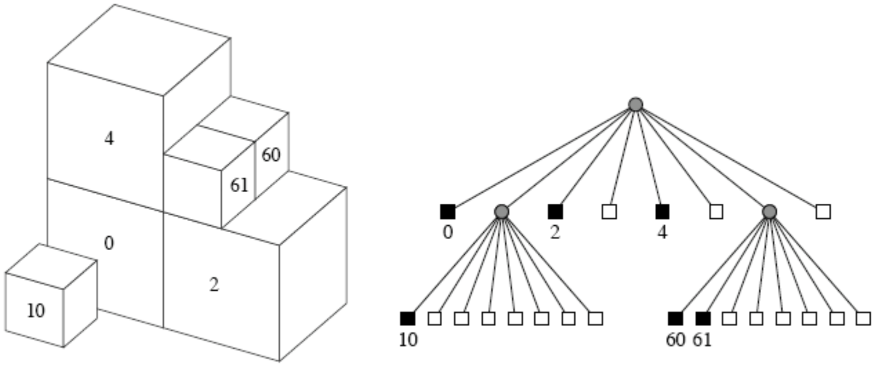


- Elimination criterion for voxels

  -> project outside a silhouette

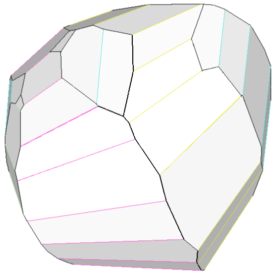
- Tradeoff precision/complexity

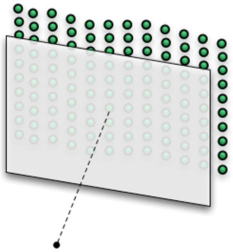
Geometric Modeling-VH
Voxel based approaches



- Adaptive data structures : octrees [Szeliski 93]

Geometric Modeling
Estimating the visual hull



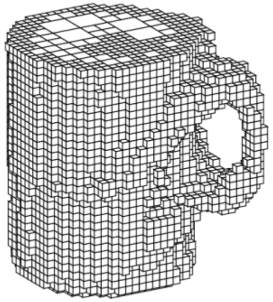


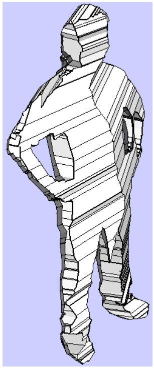
Volumetric Approaches

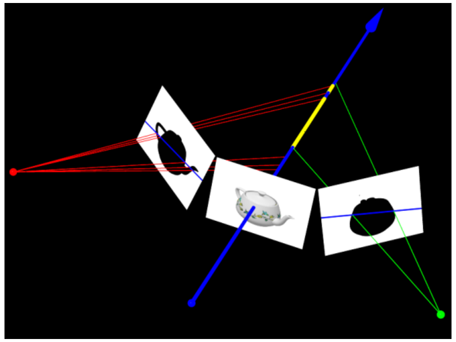
– voxels

Surface based Approaches

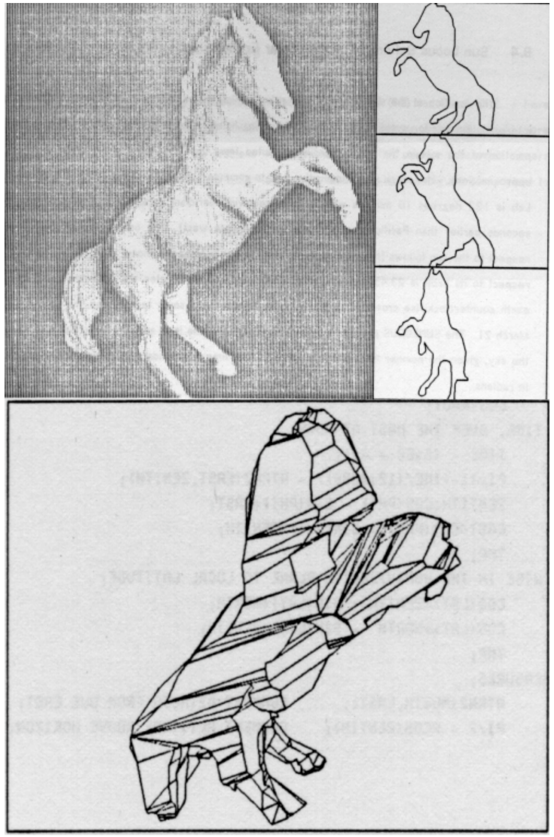
Image based Approaches





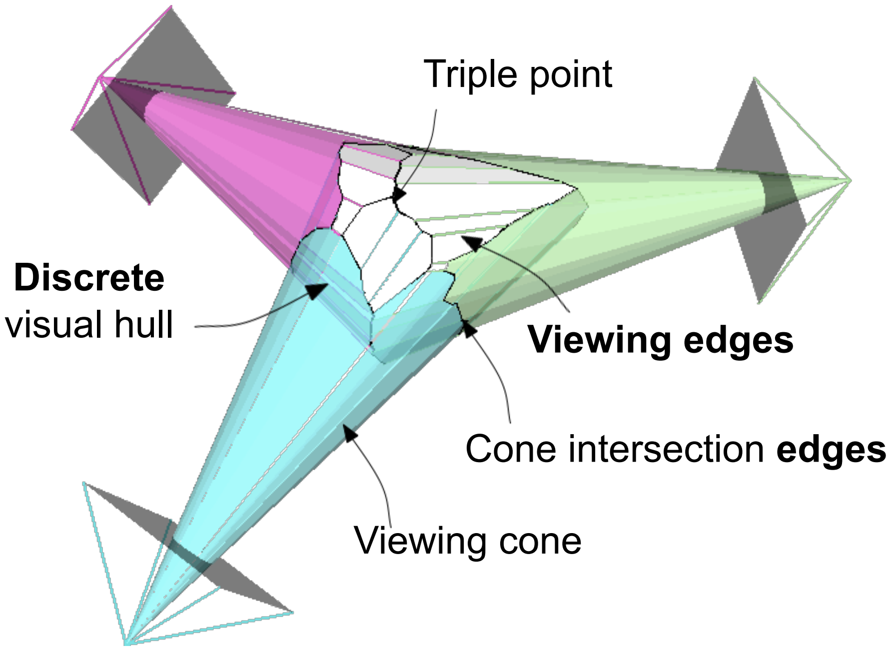


Geometric Modeling-VH
Surface based approaches



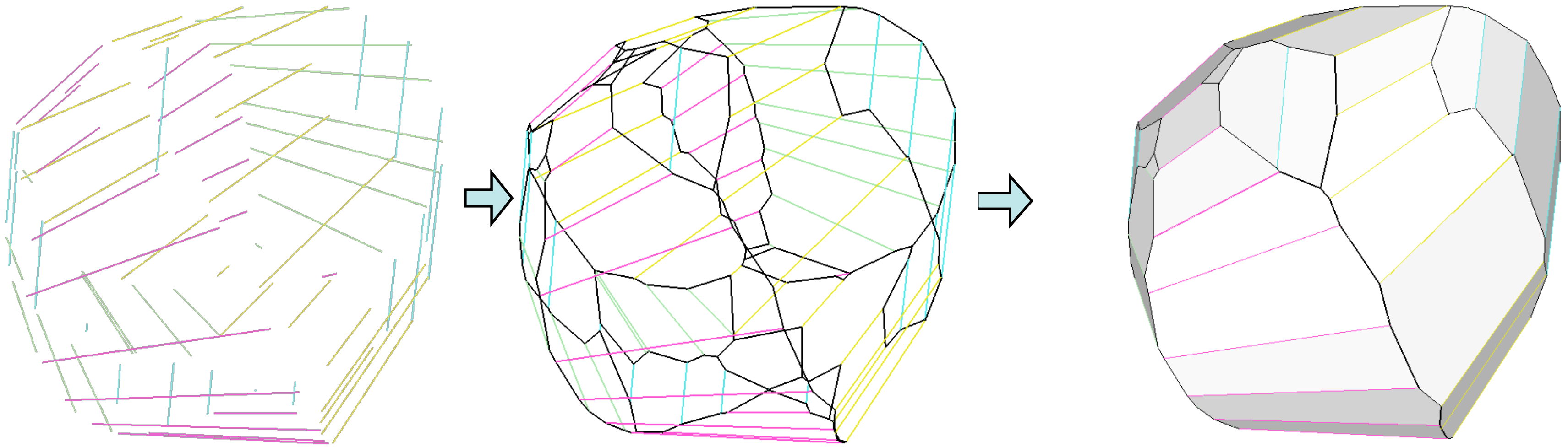
- B.W. Baumgart 74:
  - 1st approach, CSGtype (constructive solid geometry);
  - polygonal contours;
  - visual hull computed as the intersection of the viewing cones.

Geometric Modeling-VH
Surface based approaches



- Franco-Boyer 03:
  - The identification of the exact structure of the discrete visual hull allows exact computations.

Geometric Modeling-VH
Surface based approaches



The case of a sphere observed from 4 viewpoints (4 colors)

3 computation steps:

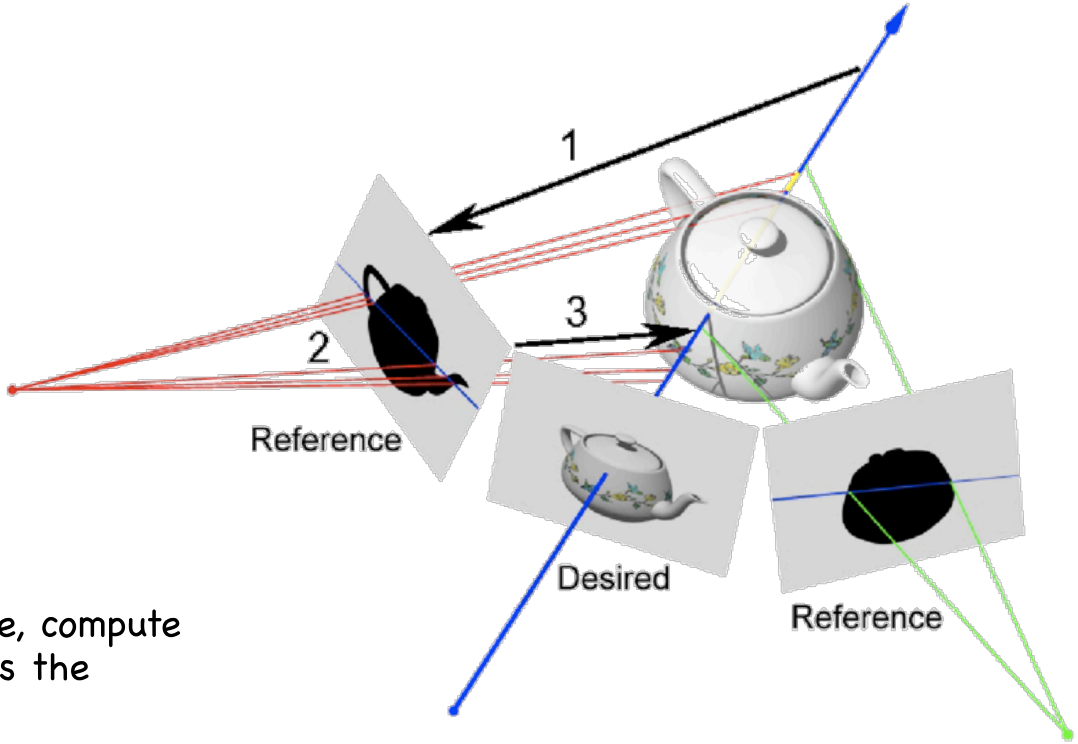
1. Viewing segments(contributions of a silhouette vertex to the visual hull)
2. The mesh connecting viewing segments.
3. Facets by going along the oriented mesh.

Geometric Modeling-VH
Computation complexity

- Polyedron intersection computation is expensive.

- Efficient algorithms:
  - Discretisation;
  - Compute in 2D.

Geometric Modeling-VH
Image based approaches



[Matusik&al. Siggraph 00]

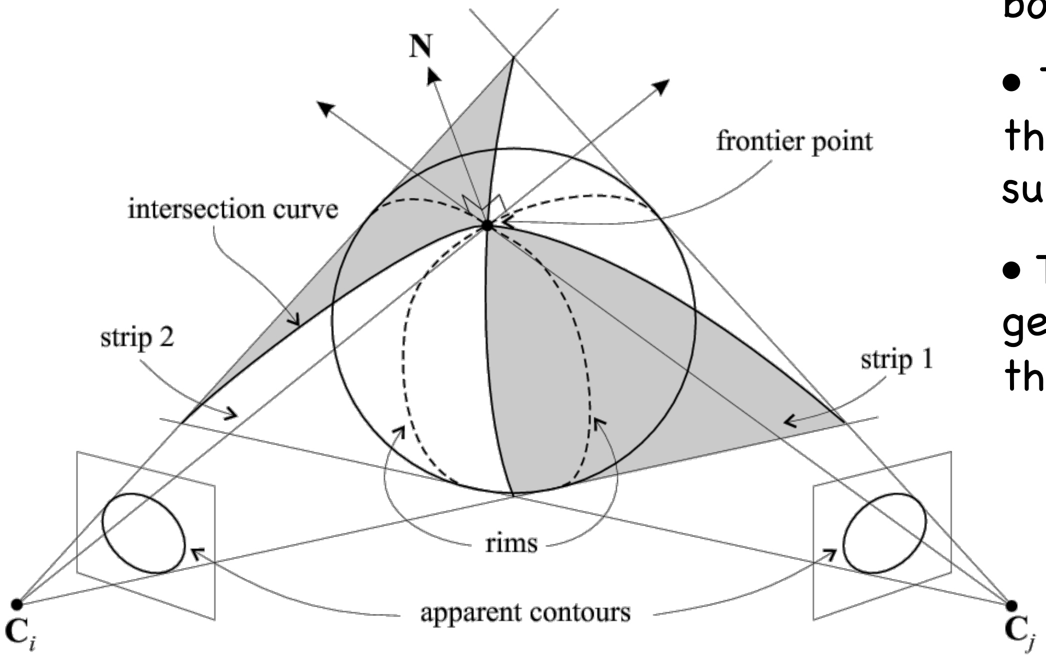
- For each pixel of the desired image, compute whether the viewing line intersects the visual hull.
- The result is an image and not a 3D model.

Geometric Modeling-VH
Image based approaches

Image-Based
Visual Hulls

[Matusik&al. Siggraph 00]

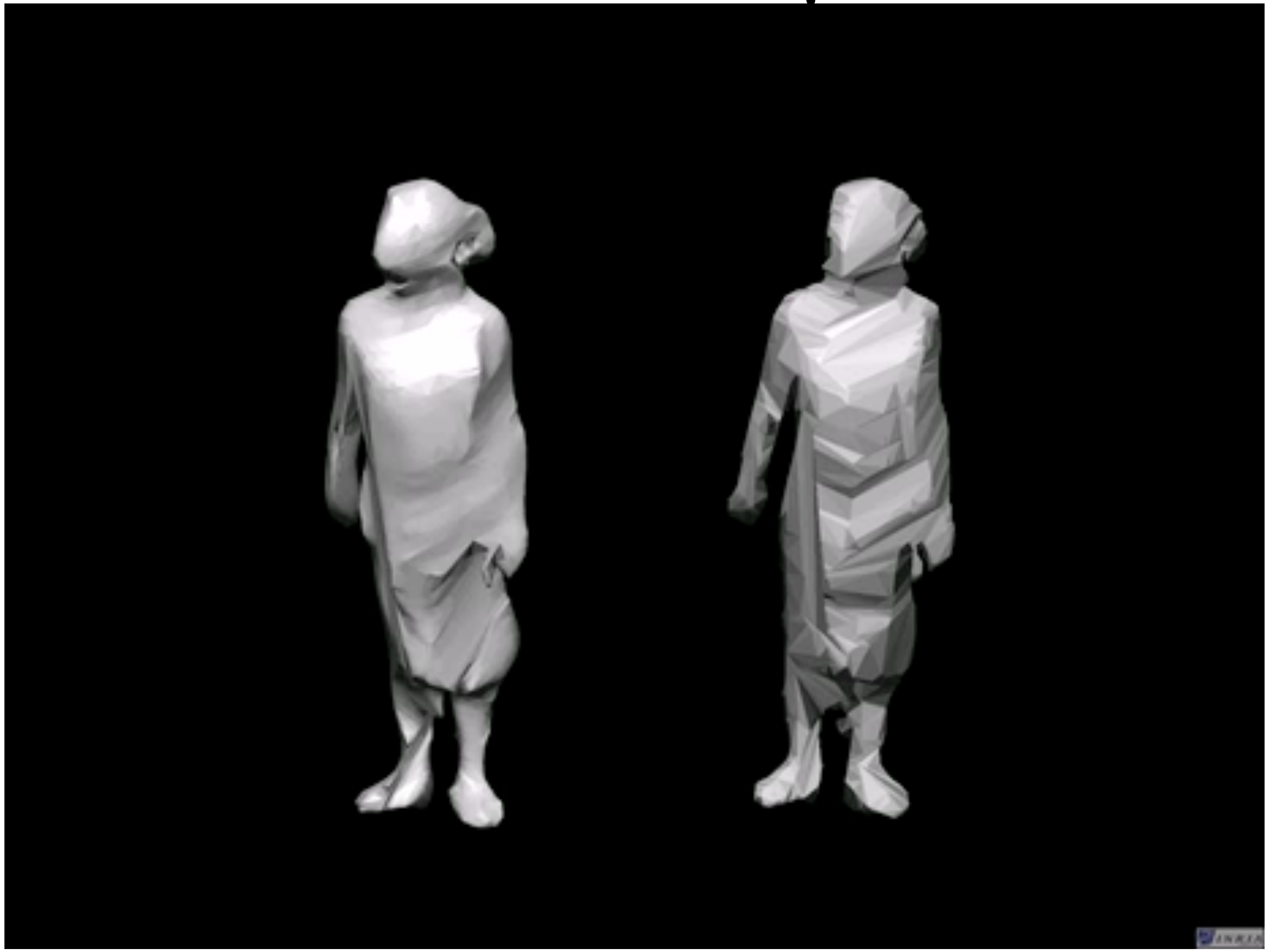
Geometric Modeling
Visual Shapes



[Boyer et al. 2006]

- The Visual hull is an extended bounding box.
- The idea is to search for a surface inside that bounding box and tangent to it. Tehs surfaces define the visual shapes.
- The visual shapes are, in general, a better geometric approximation of the observed than the visual hull.

Geometric Modeling
Visual Shapes



[Grimage 2005]

Geometric Modeling
Probabilistic Approaches

- The approaches mentioned before assume exact calibration and exact silhouettes (deterministic approaches). Such data are not exact in practice (i.e. noisy).

2 images of background probabilities (white prob = 0, black prob = 1)

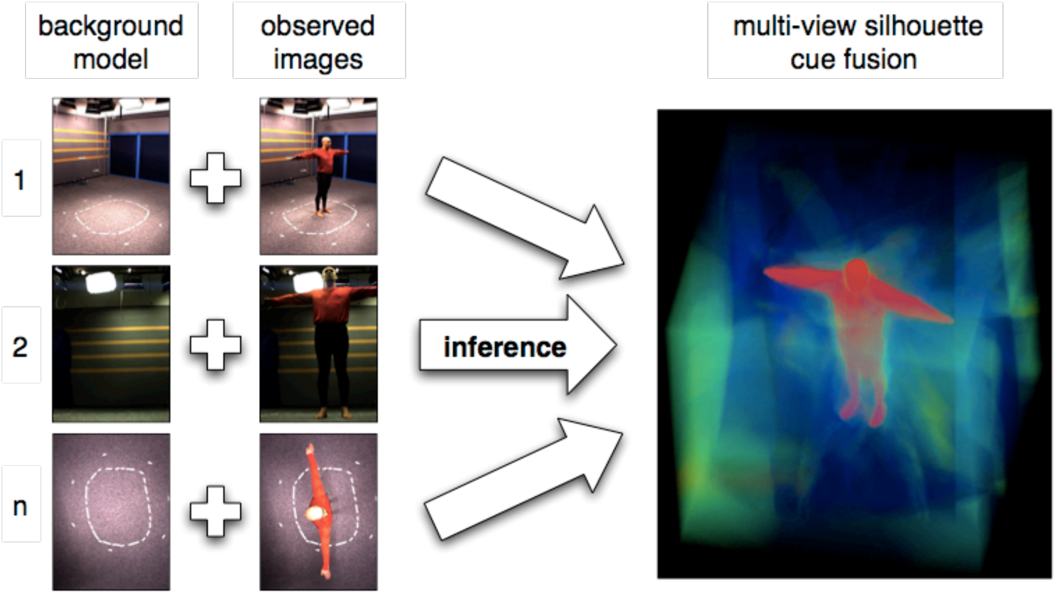
Geometric Modeling
Probabilistic Approaches

[Franco and Boyer 2005]

- Idea: we wish to find the content of the scene from images, as a probability grid
- Modeling the forward problem - explaining image observations given the grid state - is easy. It can be accounted for in a *sensor model*.
- Bayesian inference enables the formulation of our initial inverse problem from the sensor model
- Simplification for tractability: independent analysis and processing of voxels

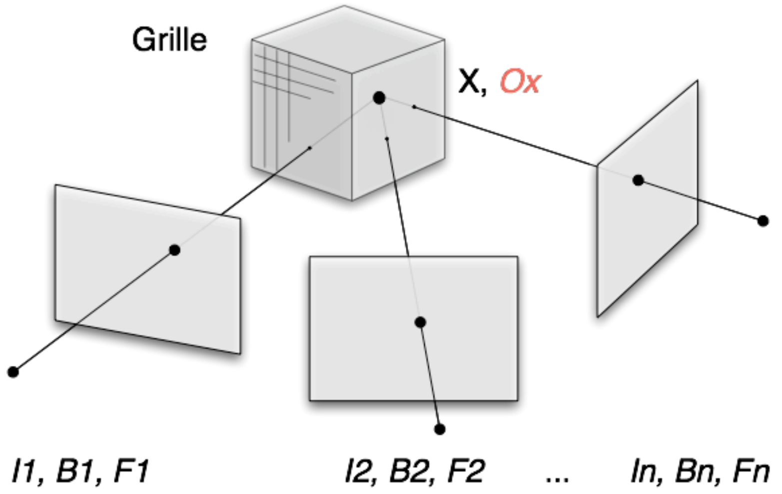
Geometric Modeling
Probabilistic Approaches

- Unreliable silhouettes: do not make decision about their location
- Do sensor fusion: use all image information simultaneously



Geometric Modeling
Probabilistic Approaches

Grid



Sensor model:

$$P(I \mid O_X) = \sum_F P(I \mid F, B) P(F \mid O_X)$$

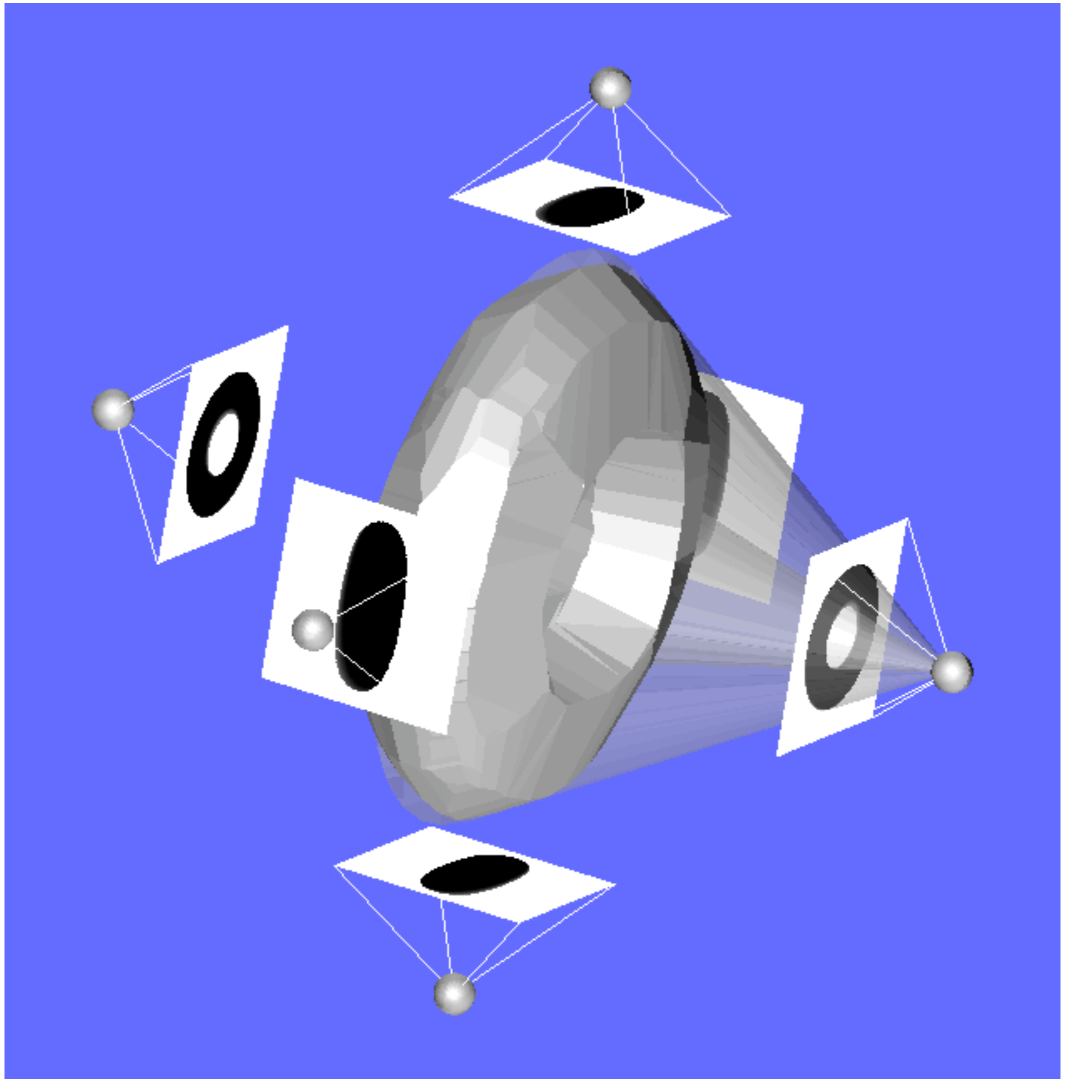
Inference:

$$P(O_X \mid I) = \frac{\prod_{img,pixel} P(I_{img,pixel} \mid O_X)}{\sum_{O_X} \prod_{img,pixel} P(I_{img,pixel} \mid O_X)}$$

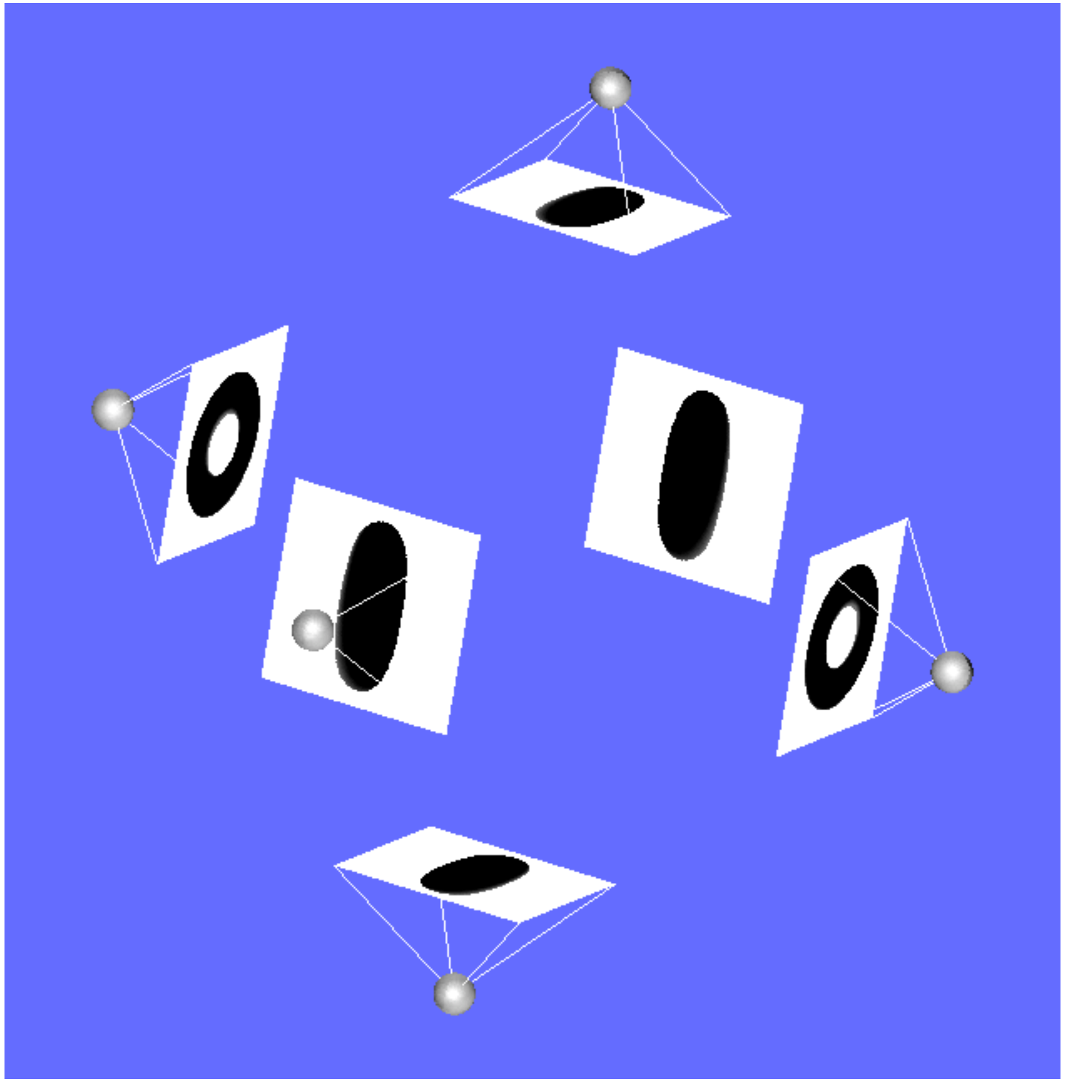
- I: color information in images
- B: background color model
- F: silhouette detection variable (0 or 1): hidden
- O_X: occupancy at voxel X (0 ou 1)

Geometric Modeling
Probabilistic Approaches

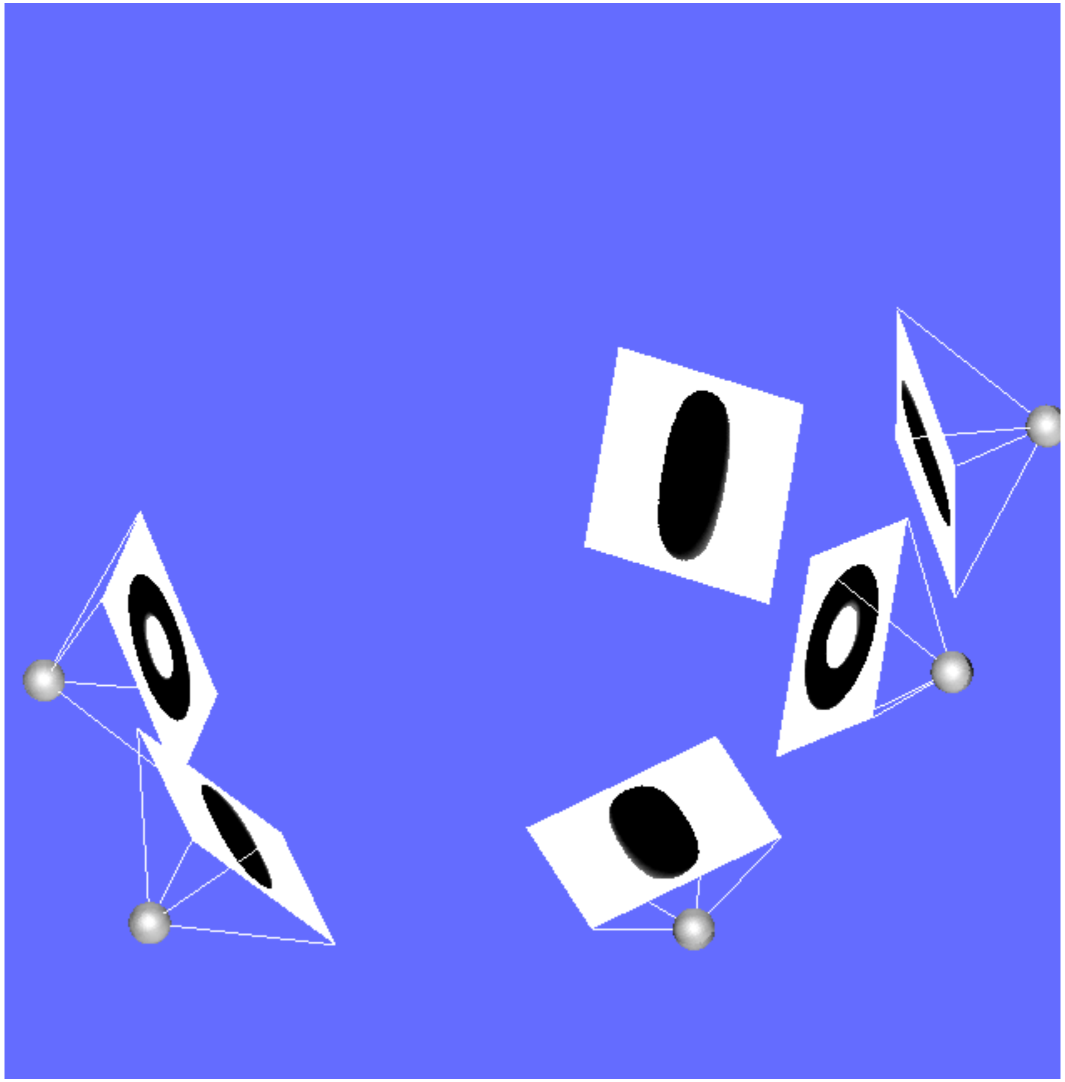
Geometric Modeling
Silhouette consistency



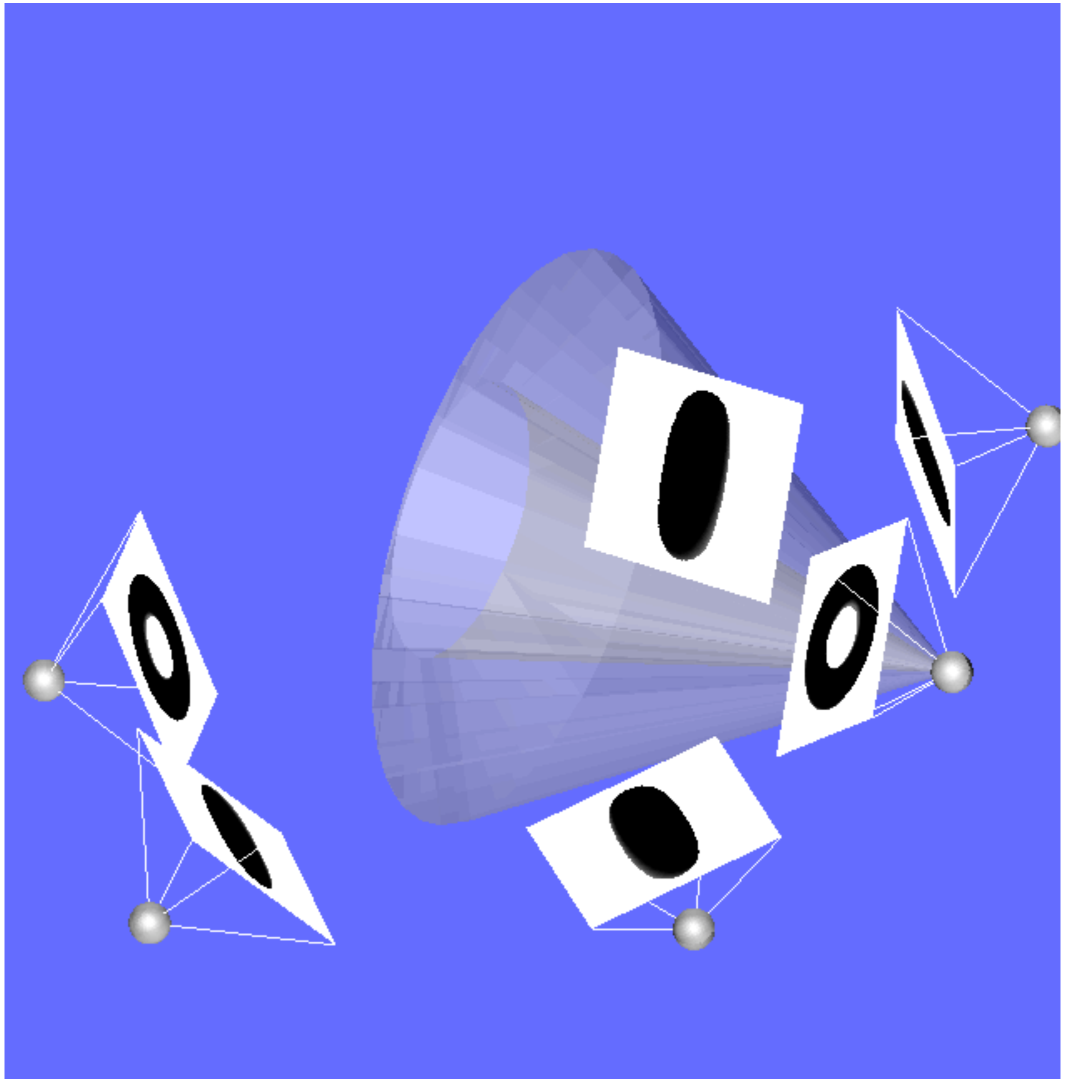
Geometric Modeling
Silhouette consistency



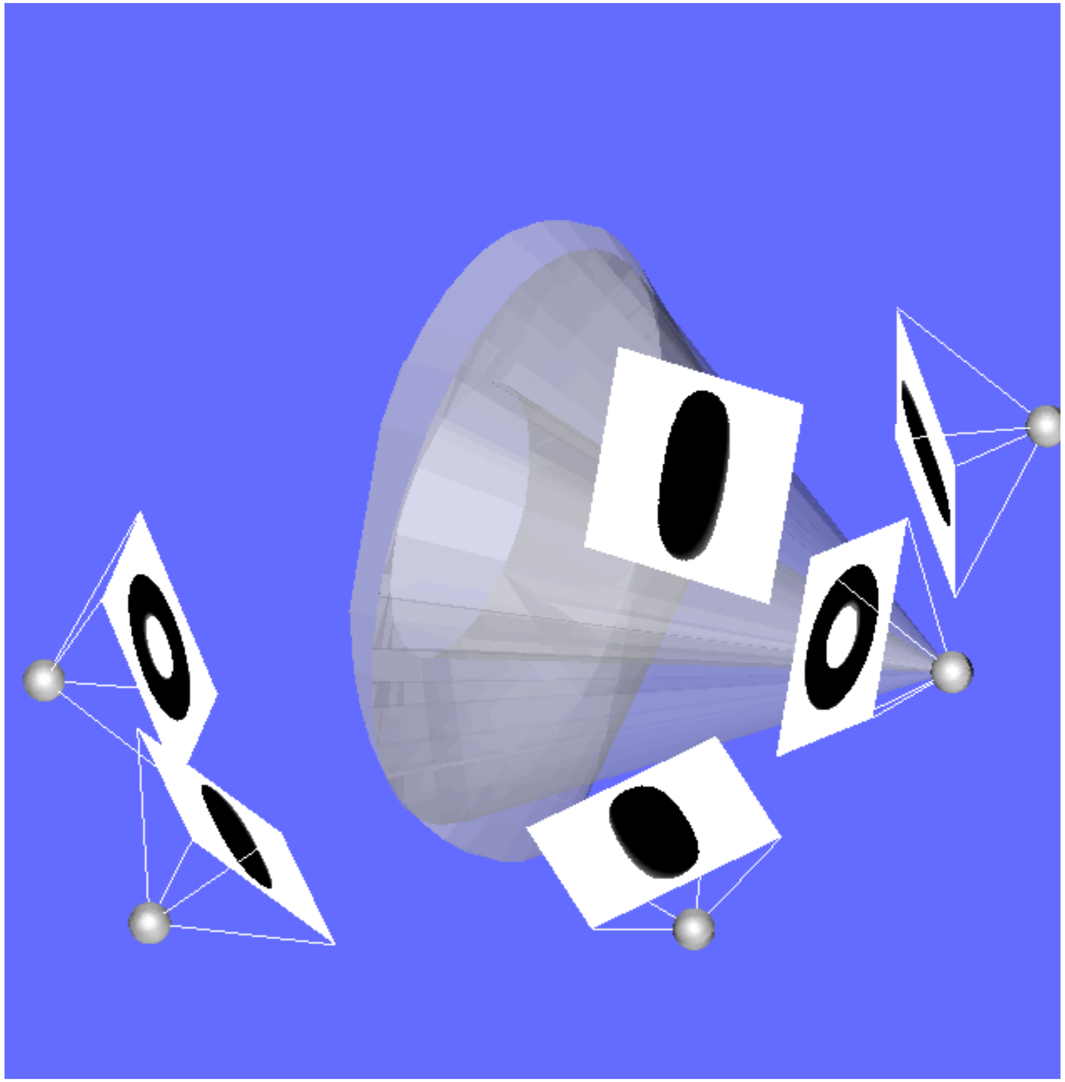
Geometric Modeling Silhouette consistency



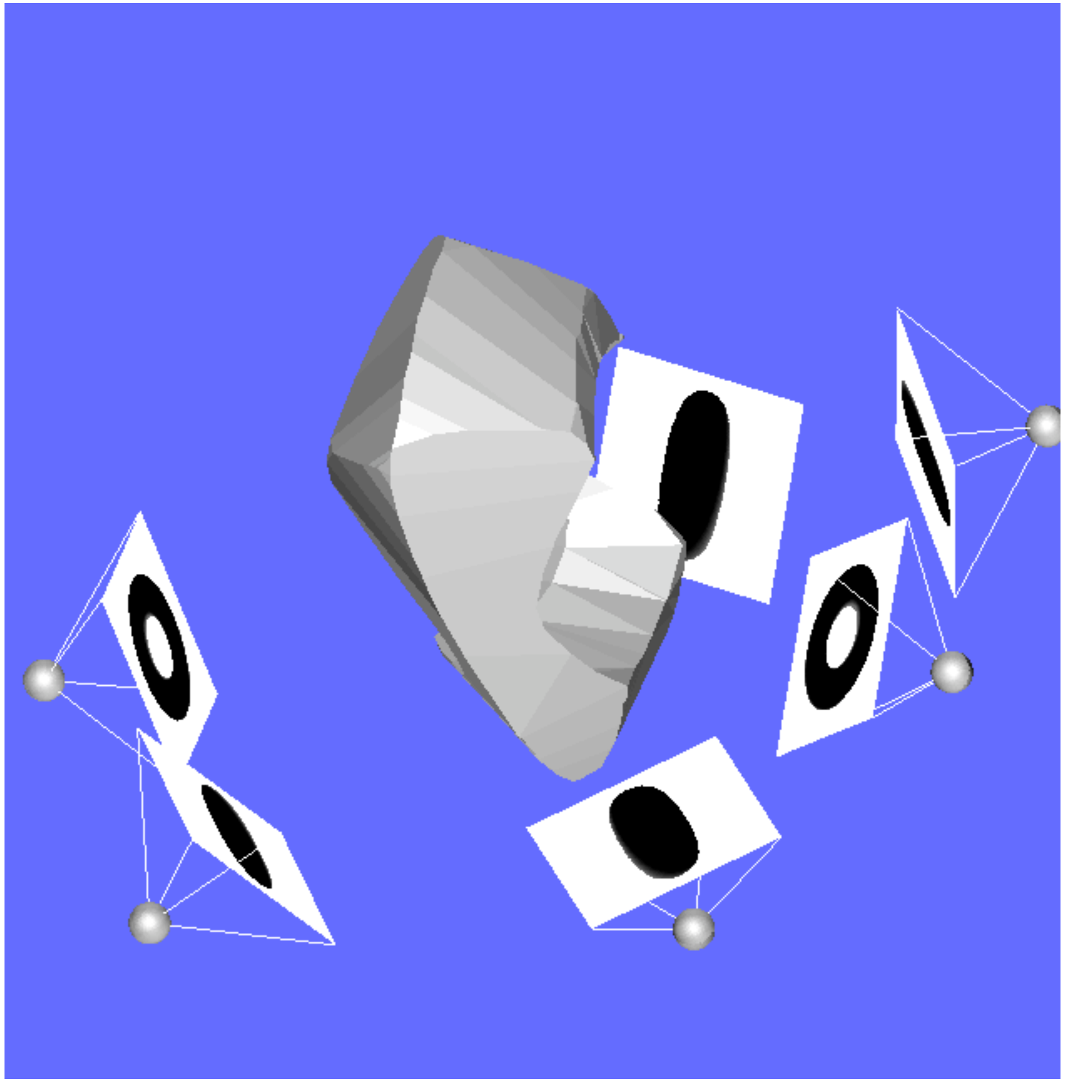
Geometric Modeling
Silhouette consistency



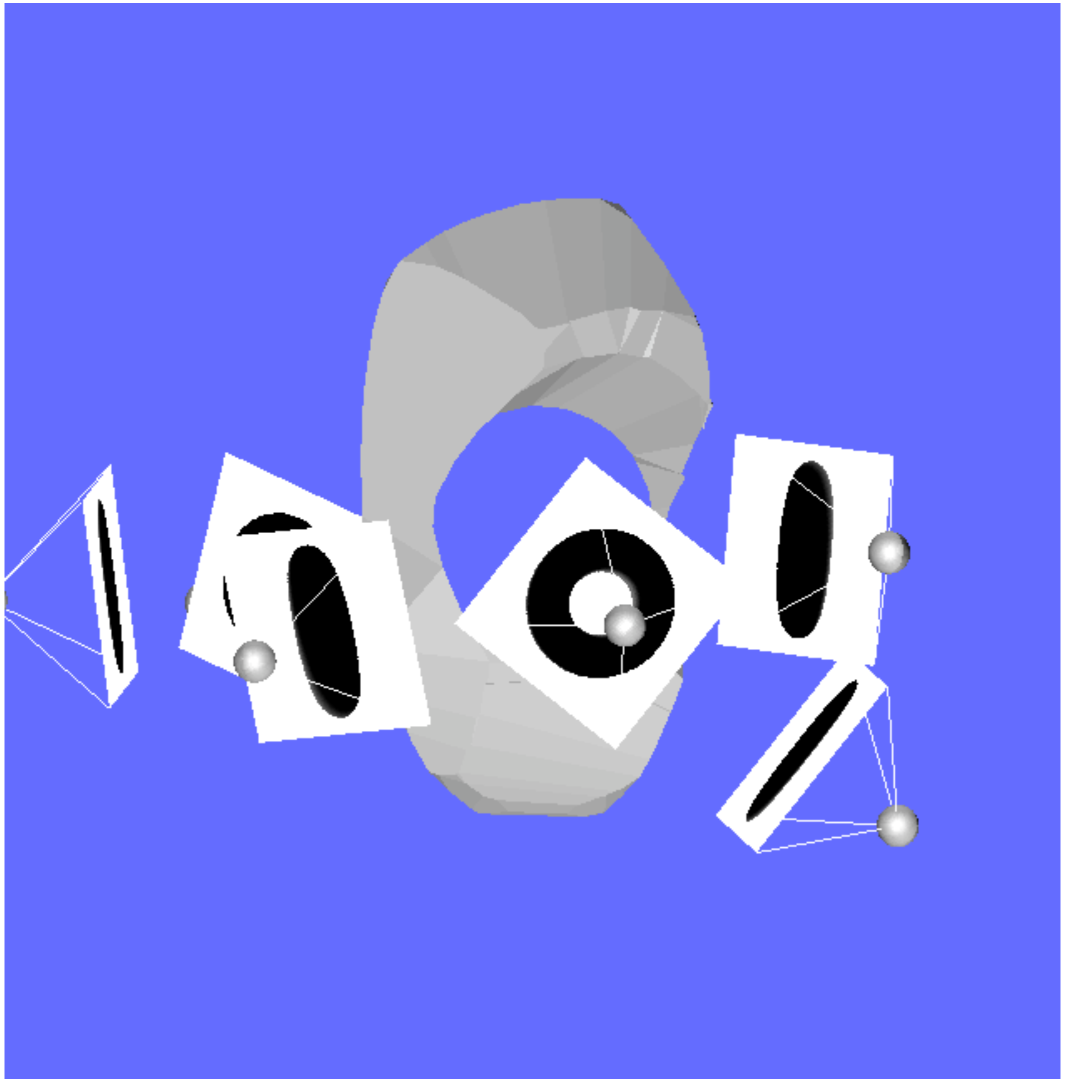
Geometric Modeling
Silhouette consistency



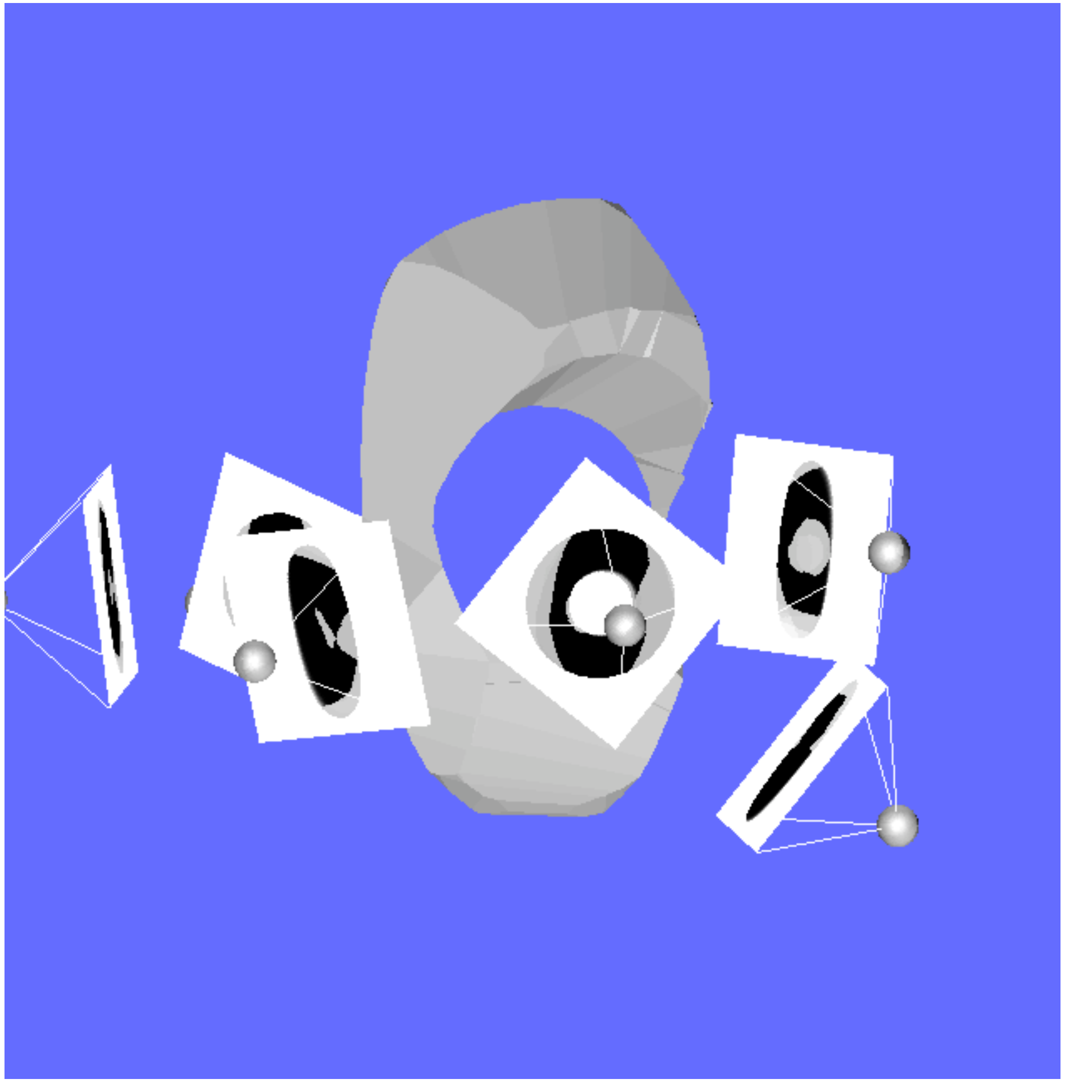
Geometric Modeling
Silhouette consistency



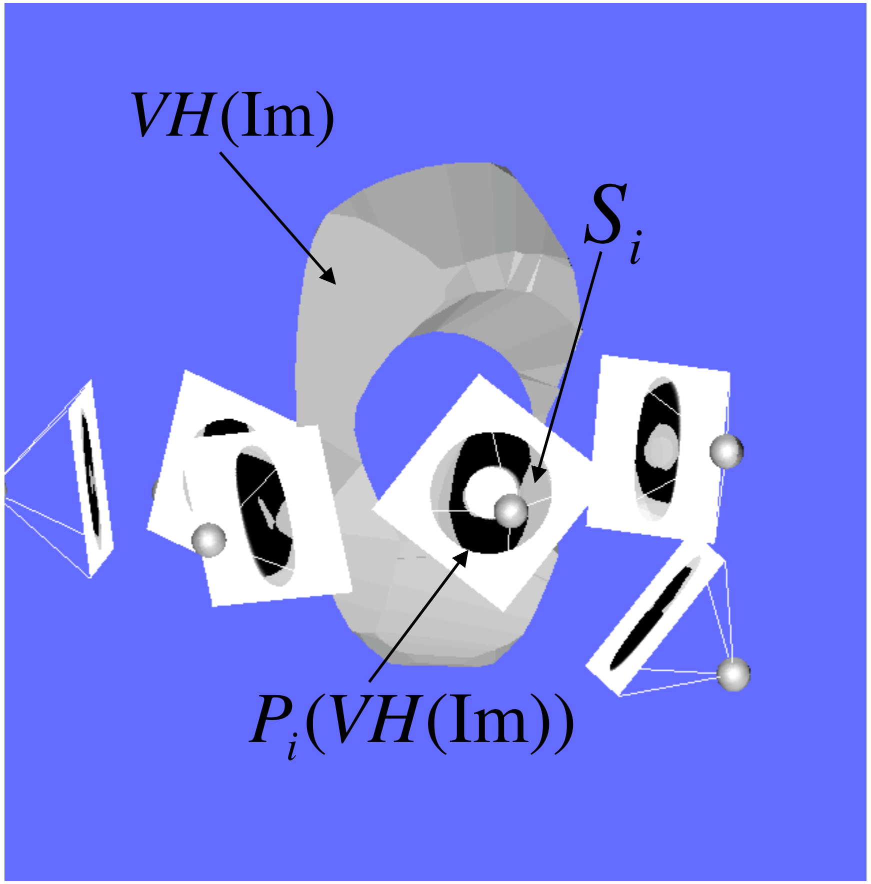
Geometric Modeling
Silhouette consistency



Geometric Modeling
Silhouette consistency



Geometric Modeling
Silhouette consistency

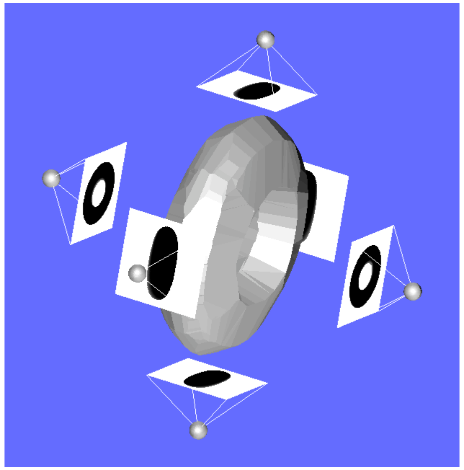


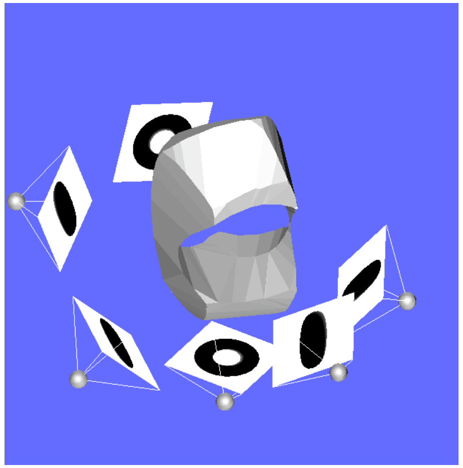
Silhouette consistency constraint:

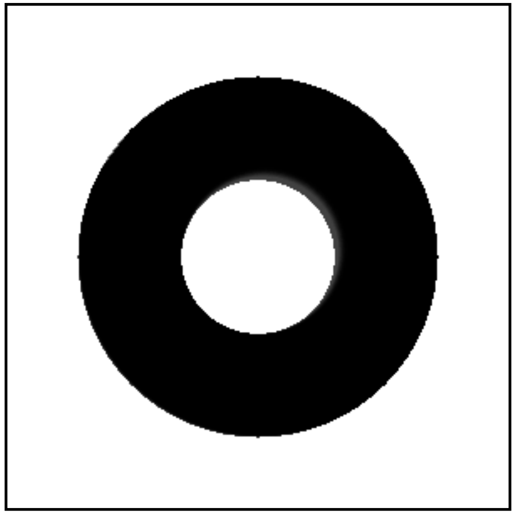
$$\boxed{\bigcup_{\mathrm{Im}}\left[S_i - P_i\left(VH(\mathrm{Im})\right)\right] = \emptyset}$$

- Si is the silhouette in image i;
- Pi is the projection in image i;
- VH(Im) is the visual hull of the set of images Im.

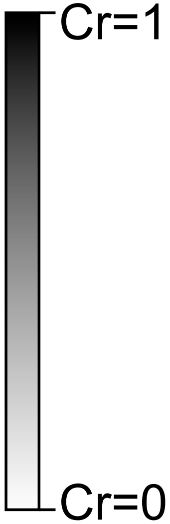
Geometric Modeling
Silhouette consistency

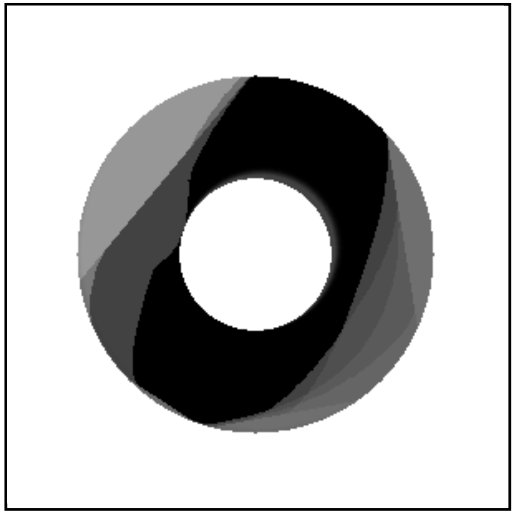






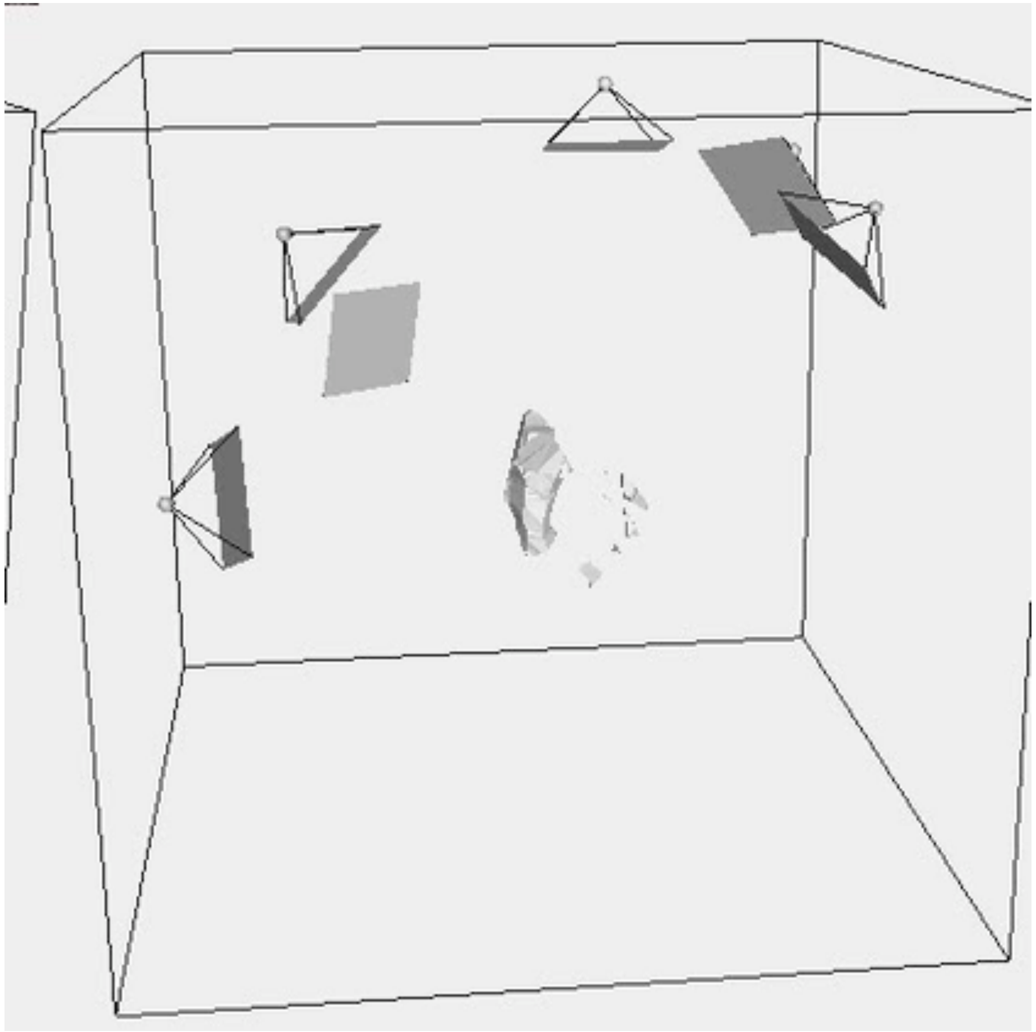
Exact calibration
Exact silhouettes





Modified calibration
Exact silhouettes

<



Photometric and geometric modeling

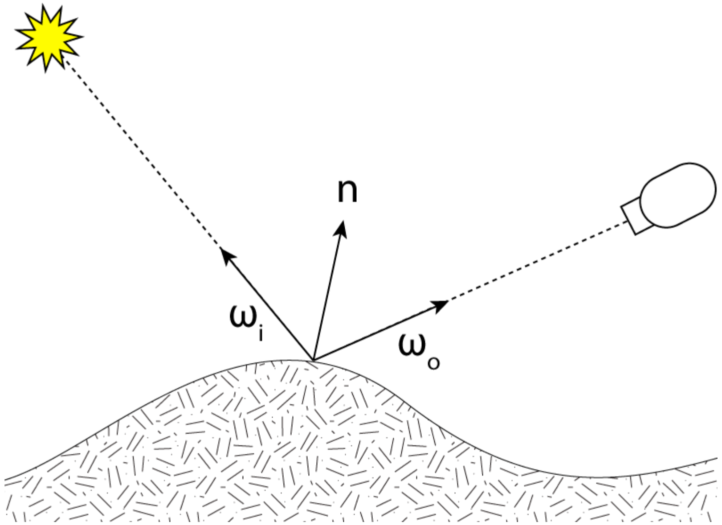
○ Using geometric information only (e.g. locations, orientations) does not allow to model precisely from images. To this purpose photometric information, or appearance, must be considered.

○ The objective is therefore to search for the geometry and the photometry that are consistent with the observations (the images).

Photometric and geometric modeling

Hypothesis :

- The observed surfaces reflect light with respect to a simplified model.
- Diffuse model (lambertian surfaces) and specular.
- constant BRDF:



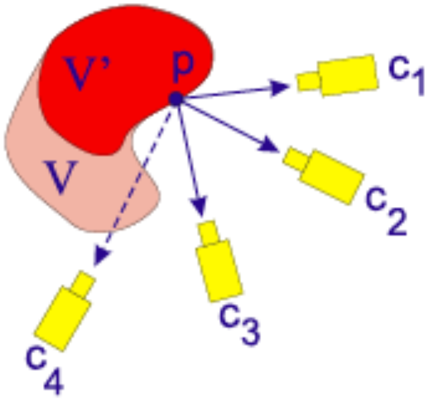
[image from Wikipedia] Bidirectionnal reflectance distribution function (BRDF)

Photometric and geometric modeling

A Lambertian surface appears equally bright in any directions

⟶ photometric criterion:

Projections of a 3D surface point must exhibit similar photometric properties in any view where the point is visible.



[Kutulakos & Seitz IJCV00]

Photometric and geometric modeling

Approaches :

- **Voxel coloring/Space carving** [Seitz & Dyer'97] : voxels not photoconsistent are carved.

Photometric and geometric modeling

Approaches:

- **Multi-View Stereo**

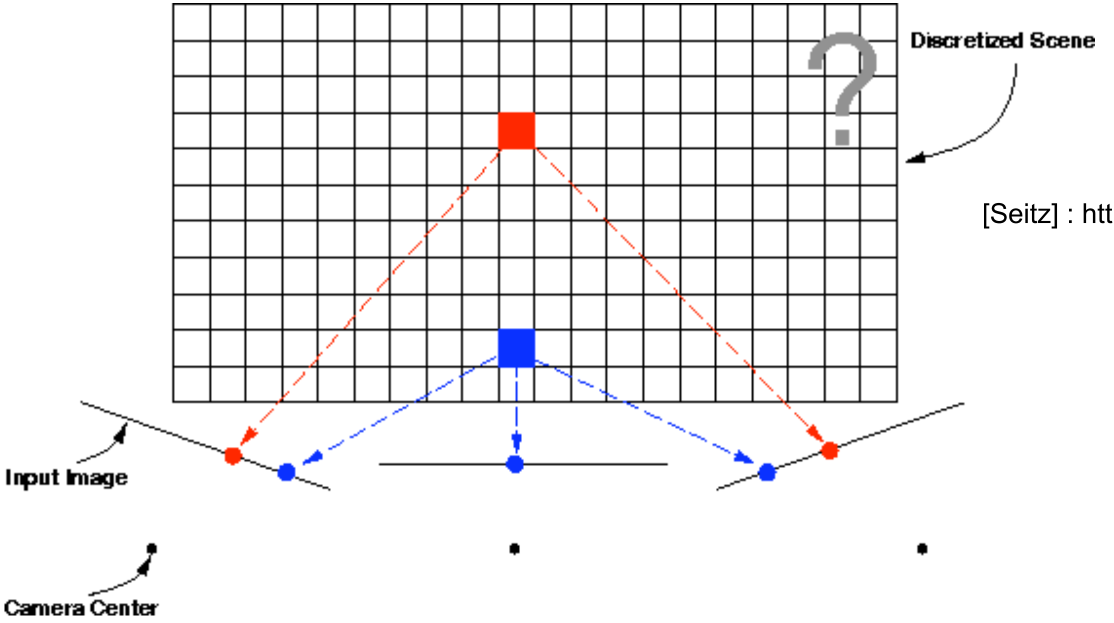
[Hernandez'04,Furukawa'06,Pons'07] : an initial surface estimation is optimized-deformed with respect to a photometric criterion.

    - Lagragian methods: mesh deformations
    - Eulerian methods: discrete grids -> level sets

[Zaharescu'07]

Photometric and geometric modeling

Voxel Coloring

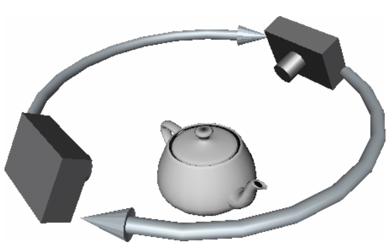


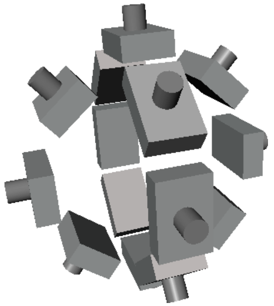
[Seitz] : http://www.cs.cmu.edu/~seitz/vcolor.html

→ voxel visibility, occluders !

Photometric and geometric modeling

Voxel Coloring





Straightforward situations for visibility: the observed scene is outside the camera center convex hull.

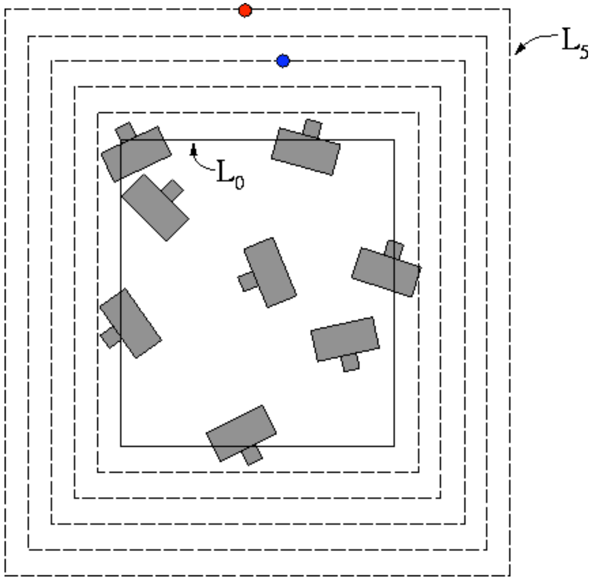
⟶ voxels are eliminated in order of increasing disctances to the camera convex hull

[Seitz] : http://www.cs.cmu.edu/~seitz/vcolor.html

Photometric and geometric modeling

Voxel Coloring

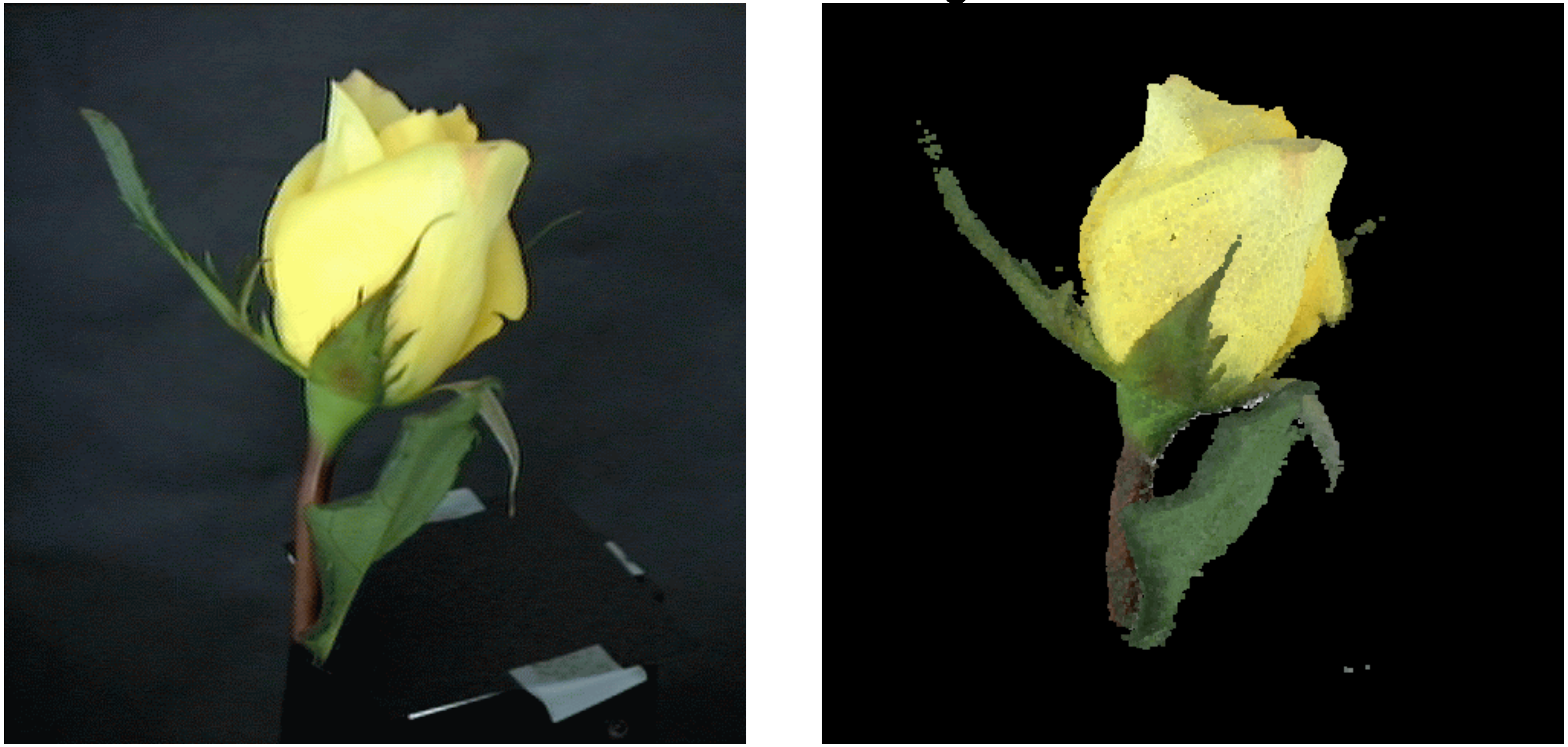
```
S={}  /* initial set of colored voxels is empty  */
for i = 1 to r do   /* traverse each of r layers */
  foreach V in the ith layer of voxels do
    project V into all images where V is visible
    if sufficient correlation of the pixel colors
    then add V to S
```



[Seitz] : http://www.cs.cmu.edu/~seitz/vcolor.html

Photometric and geometric modeling

Voxel Coloring



[Seitz] : http://www.cs.cmu.edu/~seitz/vcolor.html

Photometric and geometric modeling

Voxel Coloring

General case (Space carving) :

```
Voxels= full grid
Do until convergence
foreach V in Voxels do
if V on the surface
  project V into all images where V is visible
   if not sufficient correlation of the pixel colors
    then carve V
```

Complexity !

Photometric and geometric modeling

Voxel Coloring

Sweeping planes

Photometric and geometric modeling

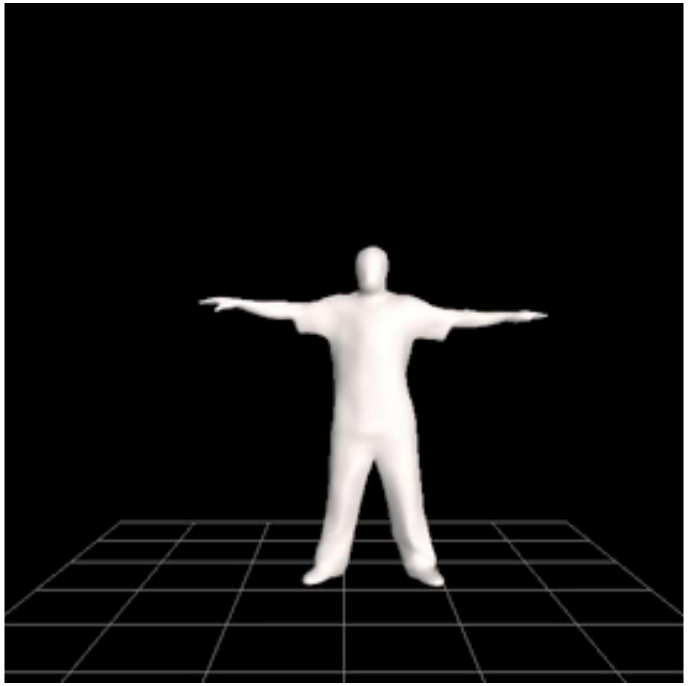
MultiView Stereo

An initial surface estimation is optimized-deformed with respect to a photometric criterion:

- Eulerian methods: values of an implicit function are evolved over a discrete-cartesian grid and with respect to a cost function (e.g. level set approaches)
  - no surface parametrization required
  - discretization issues (tradeoff precision/complexity)
  - better than space carving since voxels are not handled independently
- Lagragian methods: surface deformations in order to decrease a cost function
  - requires a surface parametrization (e.g. a mesh)

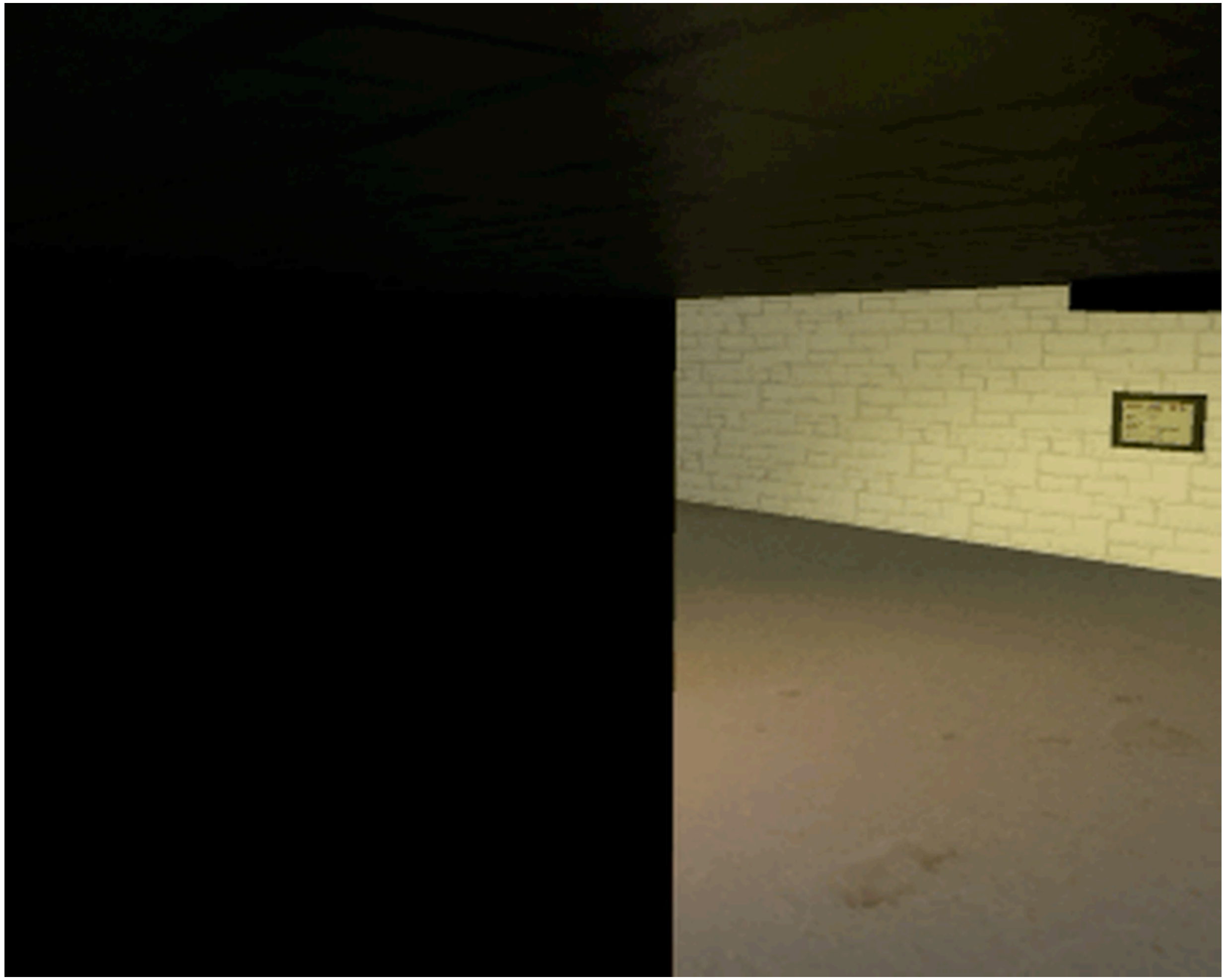
Photometric and geometric modeling

MultiView Stereo



[A. Hilton & J. Starck'06, Surry, UK]

Photometric and geometric modeling



[4D View Solutions'07]

Modeling: temporal aspects

The idea is not to model indenpedently over time sequences but to evolve instead a single model

[B. Bickel M. Botsch
R. Angst W. Matusik
M. Otaduy H. Pfister
M. Gross Siggraph'07]

Interests :

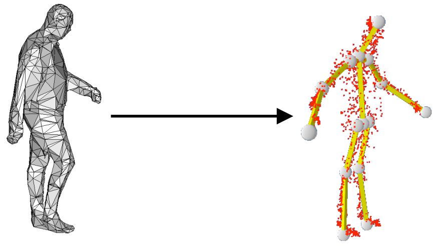
1. Constraint model with nice geometric properties (e.g. correct topology)
2. Temporal correspondences as a byproduct (-> tells about motion)

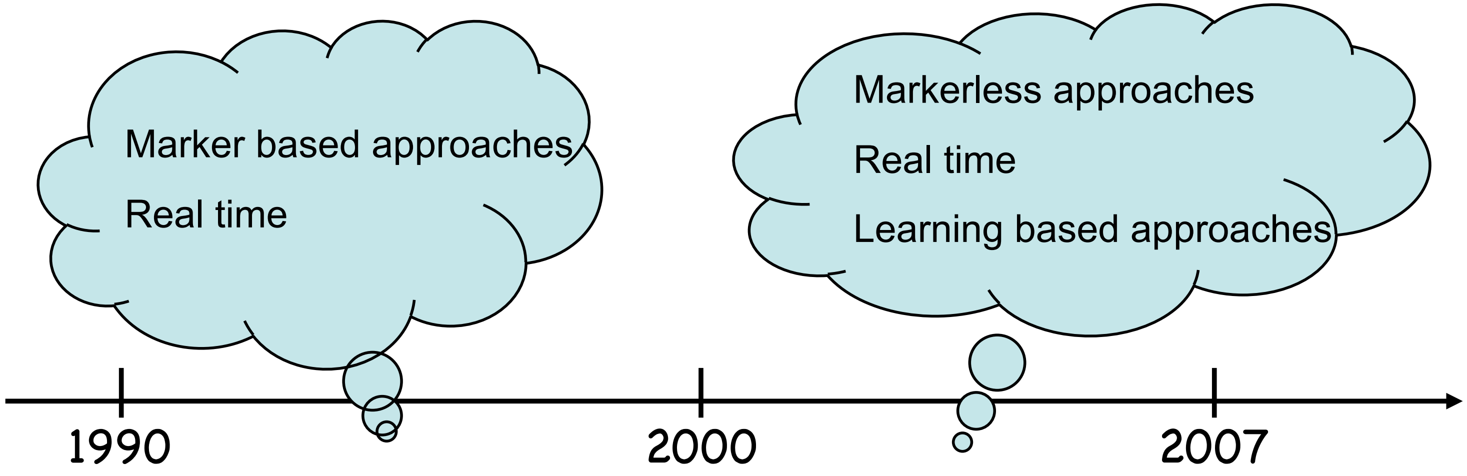
Outline

- Shape Modeling:
  - Acquisition Systems.
  - Geometric Approaches.
  - Photometric and geometric Approaches.
- Motion Modeling:
  - Marker Based Approaches.
  - Marker less Approaches.
- Action Modeling:
  - Model Based Approaches.
  - Holistic Approaches.

Motion Modeling

- Issues:
  - Precision, robustness.
  - *a priori* knowledge (motion model) required.
  - Markers required or not.





Marker based approaches

Real time

Markerless approaches

Real time

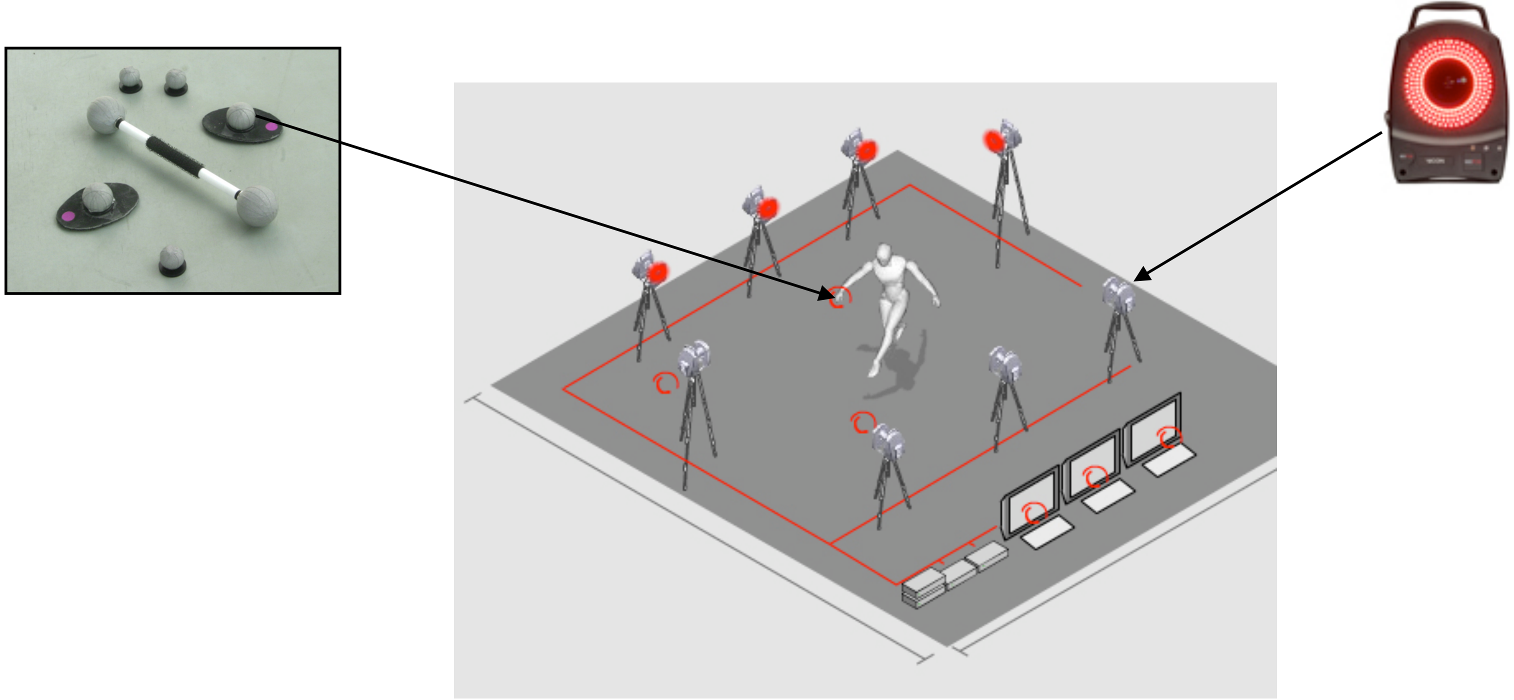
Learning based approaches

Motion Modeling

- Optical methods with markers:
  - Observations: markers are tracked in calibrated images
    - Passive markers: retroreflective markers
    - Active markers: markers emit their own lights for identification (e.g. different pulses per markers).
  - Motion recovery: a motion model (e.g. a skeleton) is matched with the marker locations.
  - Applications: "Motion capture" of bodies (with articuled models), faces (with deformable meshes), etc.
- Non optical methods with markers also exist: inertial or magnetic systems for instance.

Motion Modeling

- Marker based approaches, the Vicon system:
  - High precision cameras (up to 16 Mpixels) and high frequencies (> 400Hz).



Motion Modeling

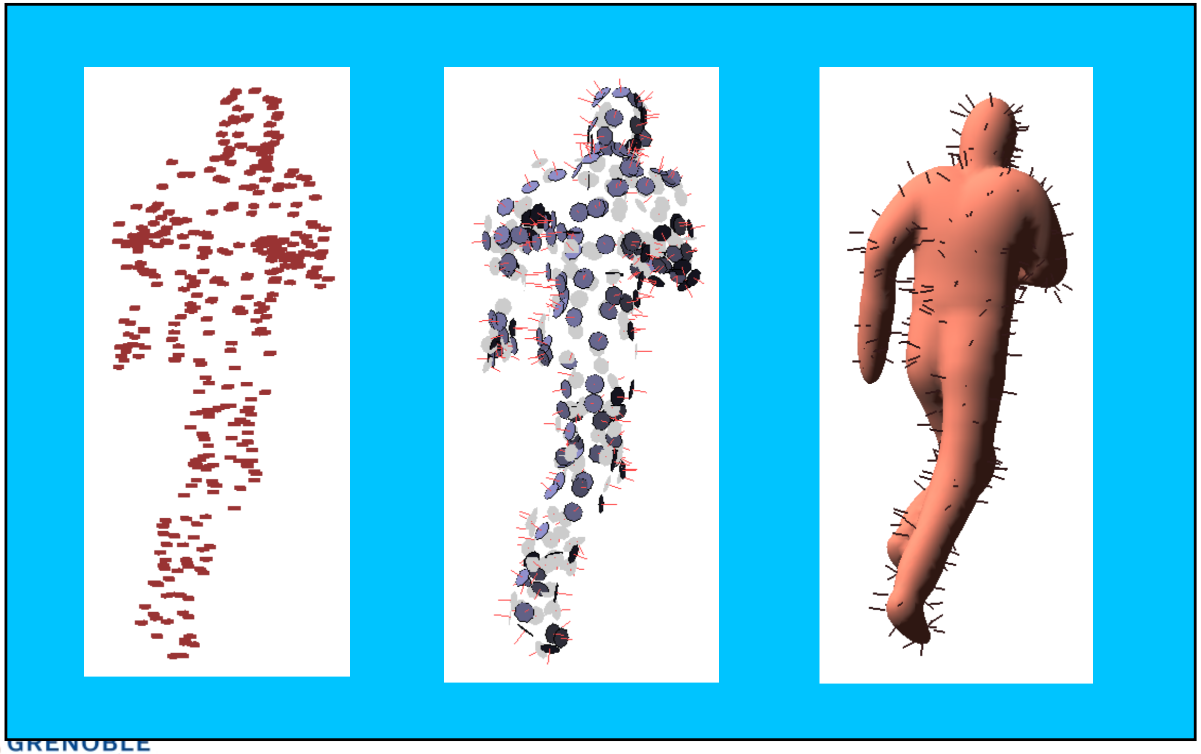
- Markerless approaches:
  - Observations: 2D (contours, silhouettes), 3D (points, shapes) in a single or multiple calibrated images.

  - Motion recovery: the motion model is matched to the observations:
    - parametric motion models: find the best model parameters (e.g. joint angles) such that the model explains the observations.
    - non parametric models: find the best model (in the database) that explains the observations.

  - Applications: motion capture of bodies, faces, etc.

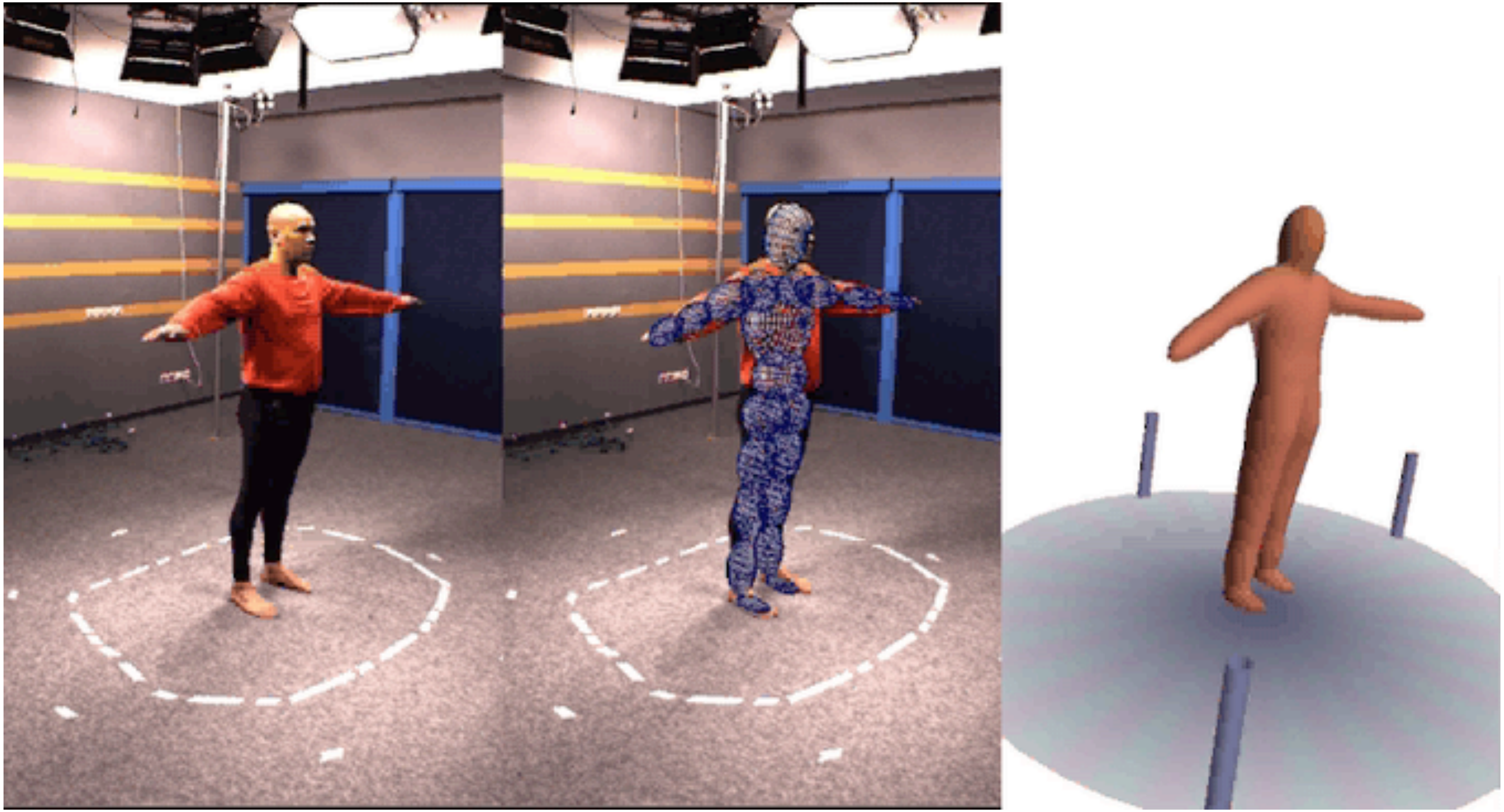
Motion Modeling

- Parametric motion model and 3D observations (locations and surface normals)
- Articulated model composed of ellipsoids.



Motion Modeling

Motion Modeling



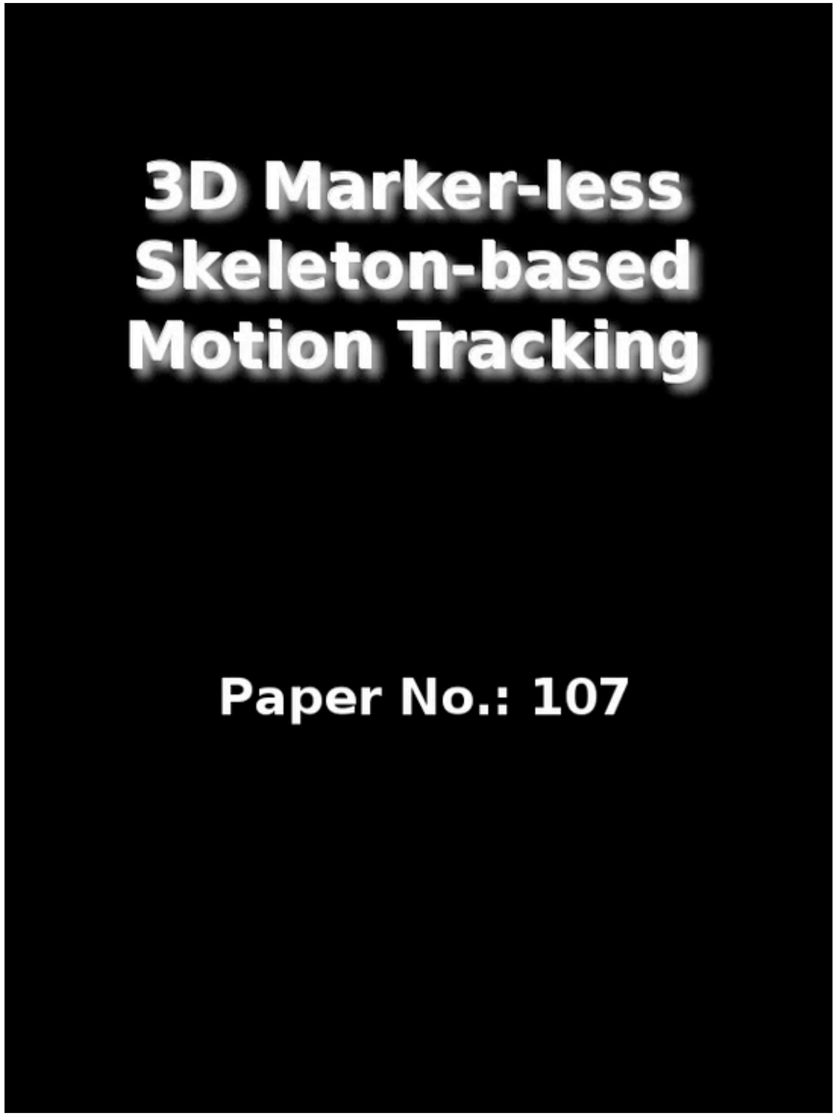
M. Niskanen, E. Boyer and R. Horaud BMVC'05

Motion Modeling

Parametric model approach:

- Skeleton model: focus on the motion not the shape.
- 3 step approach:
  1. Estimate shape.
  2. Compute the medial axis of the estimated shape (discrete version of the medial axis).
  3. Find skeleton parameters such that distances between the medial axis and the skeleton are minimum.

Motion Modeling



3D Marker-less
Skeleton-based
Motion Tracking

Paper No.: 107

C. Ménier, E. Boyer and B. Raffin 3DPVT'06

Motion Modeling

- Learning based approaches: observations (silhouettes for instance) are matched to pre-learned models with known motion.

  ⟶ Allow for non parametric motion models, e.g. motions are represented by a set of key pose transitions.

[Aggarwal & Triggs'04]

Motion Modeling

- Generative approaches (e.g. kinematic models): poses are determined by inference using a parametric model (i.e. continuous).
  - Pros: allow to represent large spaces of motions.
  - Cons: difficult to find a pose without a good initialization
- Non generative approaches: poses are recognized in a learned database.
  - Pros: initialization is not an issue.
  - Cons: limited spaces only can be represented.

Outline

- Shape Modeling:
  - Acquisition Systems.
  - Geometric Approaches.
  - Photometric and geometric Approaches.

- Motion Modeling:
  - Marker Based Approaches.
  - Marker less Approaches.

- Action Modeling:
  - Model Based Approaches.
  - Holistic Approaches.

Action Modeling

Lift Arm

Scratch Head

Sit Down

Walk

⋮

Punch

Action Modeling: Applications

Entertainment

HCI

Surveillance

Ambient Intelligence

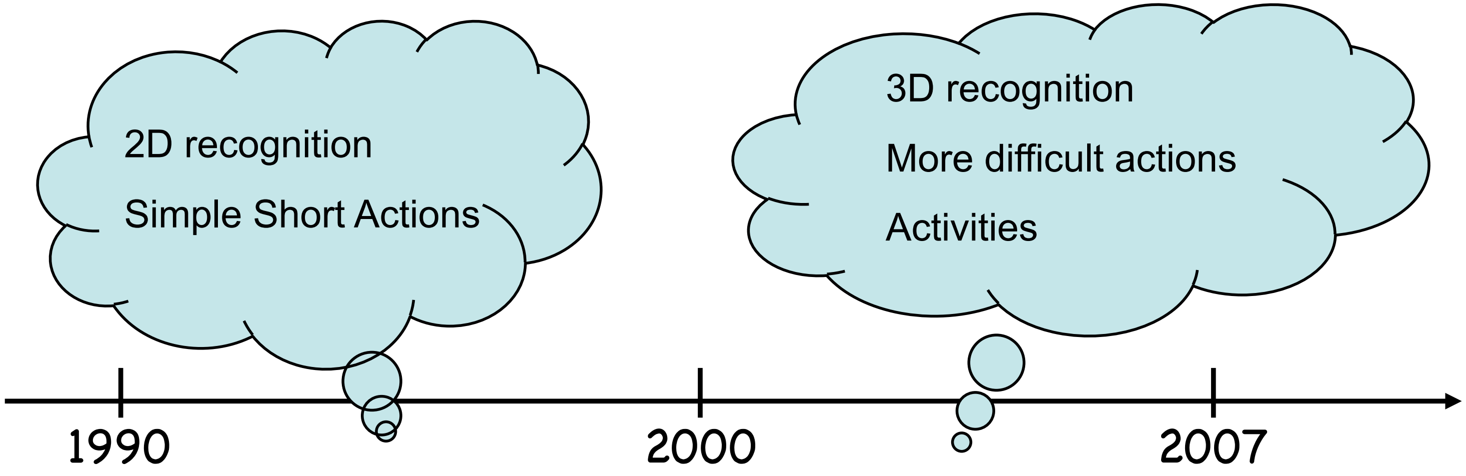
Elderly Monitoring

Sports

Group Actions

Action Modeling

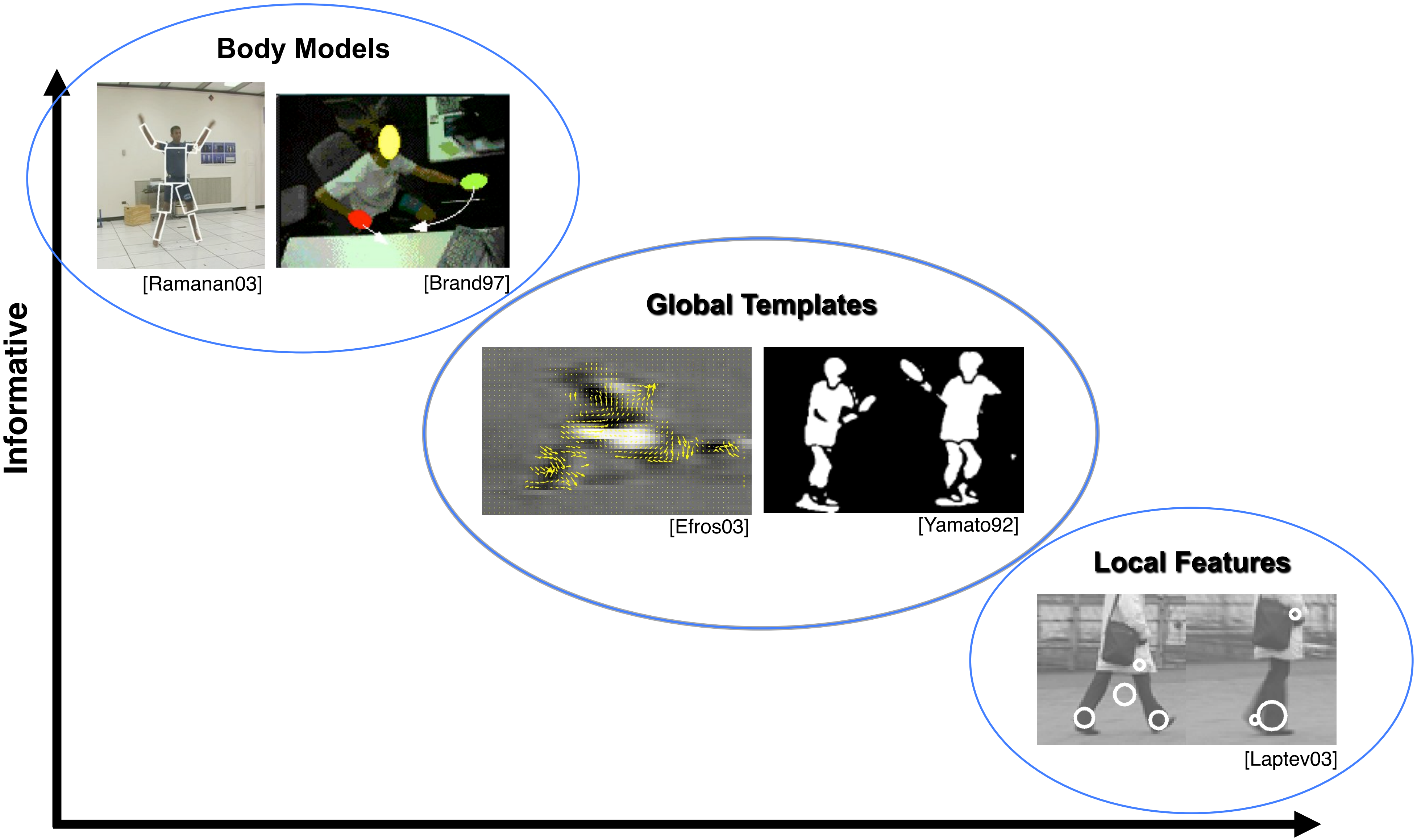
- Issues:
  - What is an action ? Posture and Dynamics
  - Action space dimensions.
  - Viewpoint dependence .
  - Occlusions.
  - Computation time.



Action Modeling

- Strategies :
  - Model Based approaches: Actions are modeled in the parameter space associated to an a priori model.
  - Holistic Approaches: (*Holism [def.]: the tendency in nature to form wholes that are greater than the sum of the parts through creative evolution) :* actions are described with respect to observations without the help of an intermediate model.

Action Modeling: Posture



Body Models

[Ramanan03]

[Brand97]

Global Templates

[Efros03]

[Yamato92]

Local Features

[Laptev03]

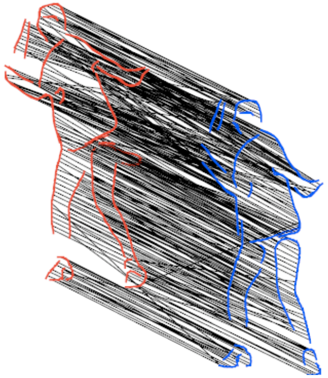
Informative

Practicable

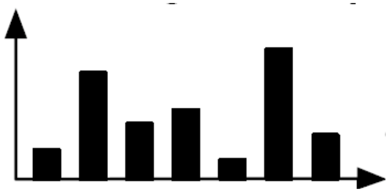
Action Modeling: Dynamics

Efficient ?

Dynamic free

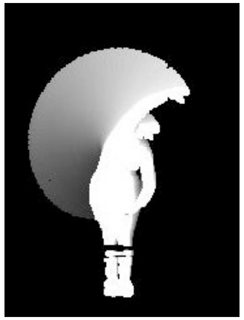


Key-Frames [Carlsson01]



Histogram of Key-Frames [Thurau07]

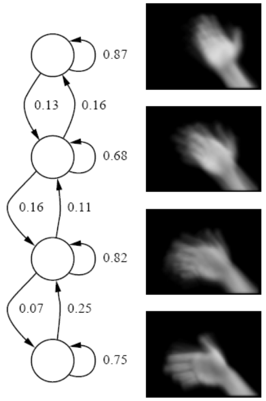
Implicit:
Space-Time Rep.



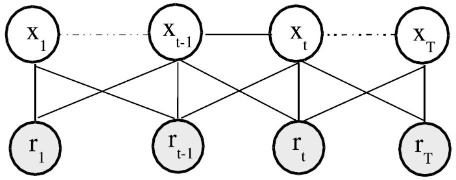
Motion History Images [Bobick96]

Space-Time Volumes [Blank05]

Explicit:
State-transition Model



HMM [Wilson95]

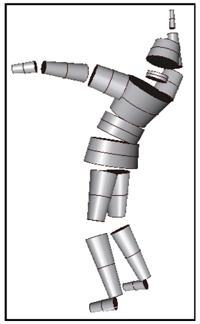


CRF [Sminchisescu05]

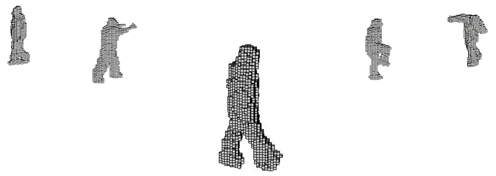
Dynamics

Action Modeling: View-Independence

- Use 3D information

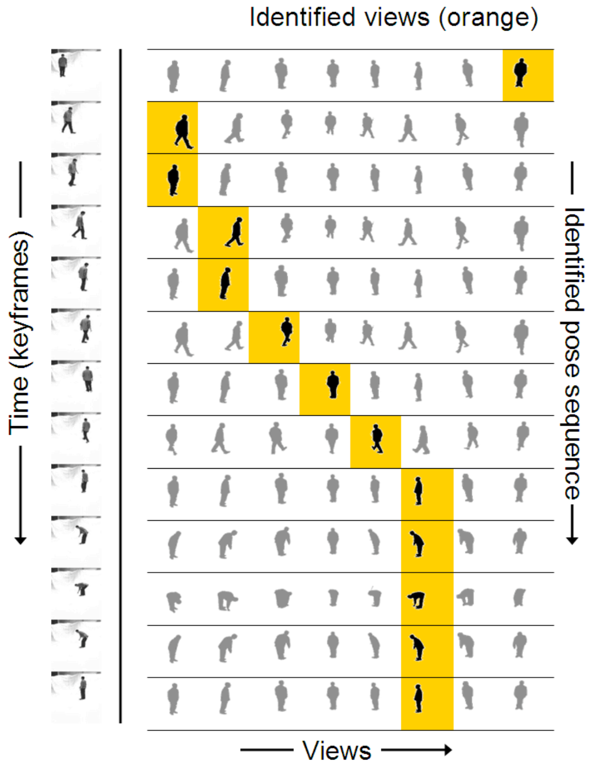


Kinematic model



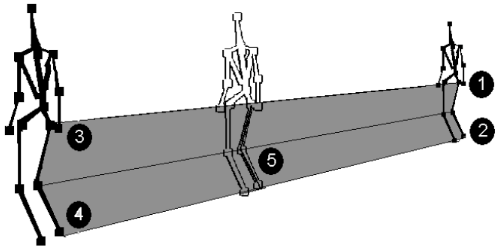
Exemplars [Weinland 07]

- Use multiple views

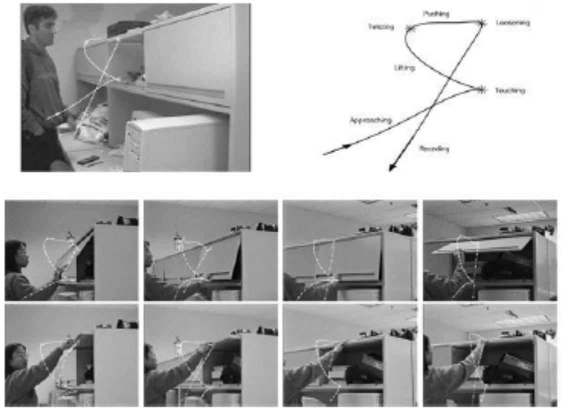


[Ogale04]

- Use View-invariance



Geometrical Invariants computed from 5 points in plane [Parameswaran05]

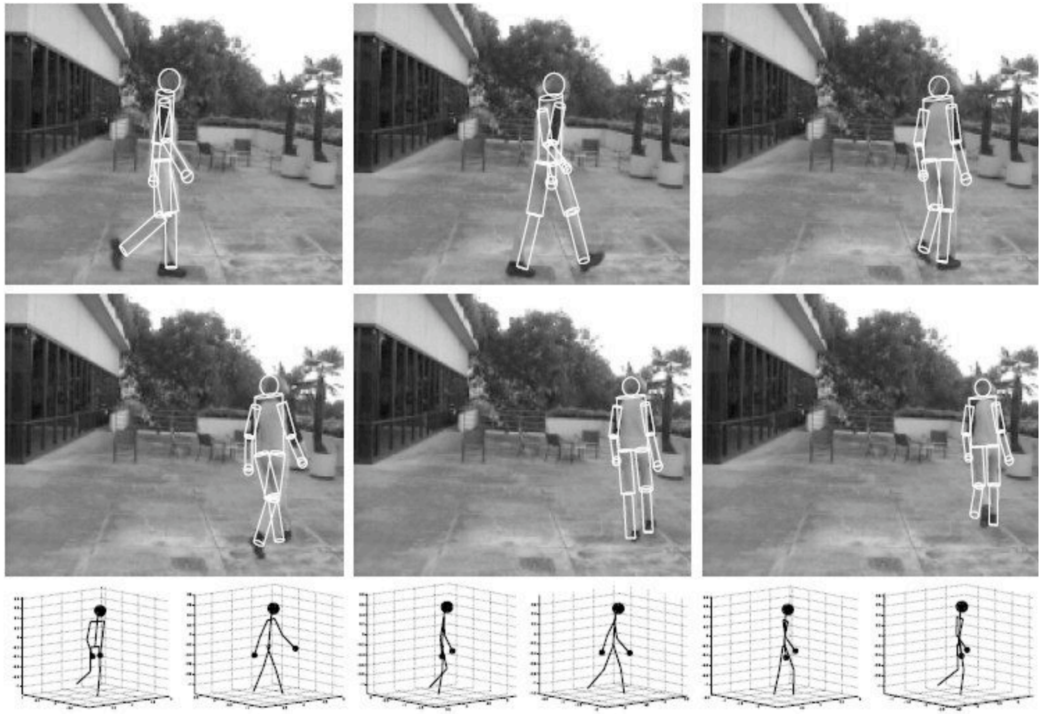


Rank constraints on point correspondences [Rao03]

Action Modeling

Model based approaches:

Intermediate representation, e.g. kinematic model



[Sidenbladh1996]

Recognition in the parameter space

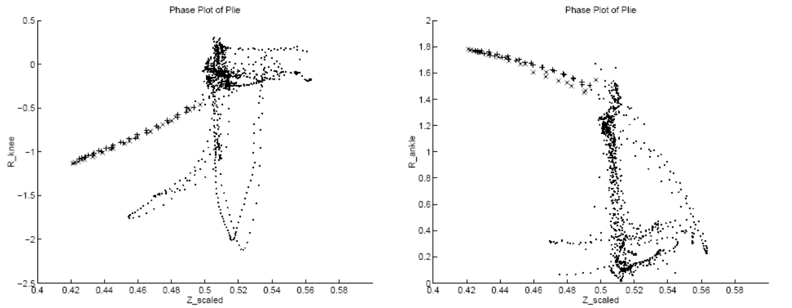
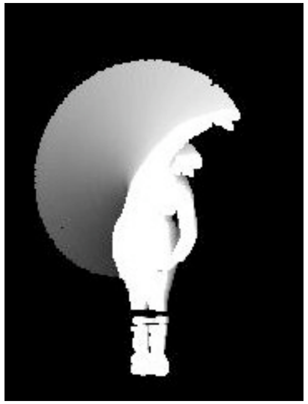


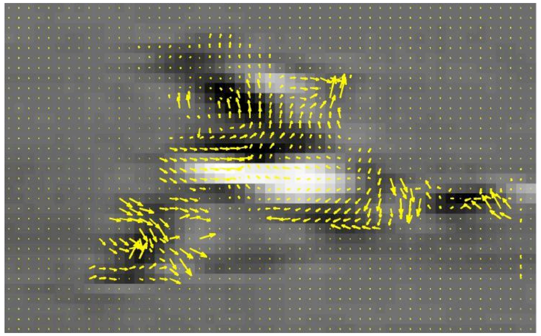
Figure 2: × and + marks time steps during pliés for two dancers; "." marks time steps during other movements. Angles are in radians. These two views of the phase space show how the plié lies in a region separated from the other movements (data subsampled for clarity).

[Campbell95]

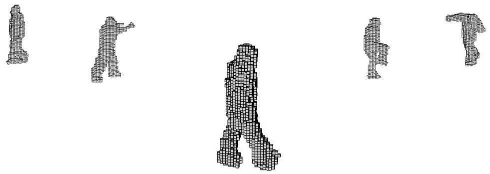
Action Modeling: Holistic approaches



Motion History Images [Bobick96]

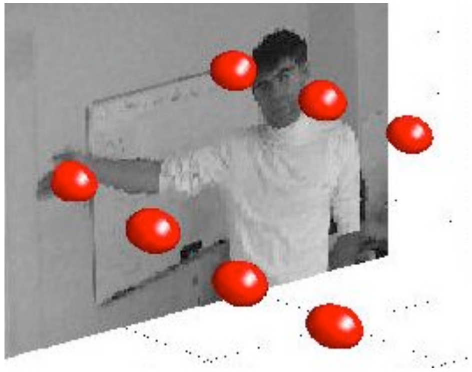


Optical Flow [Efros03]

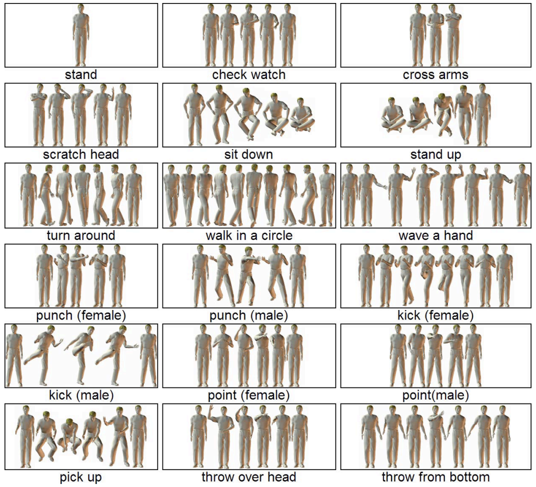


3D Exemplars [Weinland07]

Space-Time Volumes [Blank05]



Space-Time Interest Points [Laptev05]



2D synthetic exemplars [Lv07]

➔ View dependence

Action Modeling

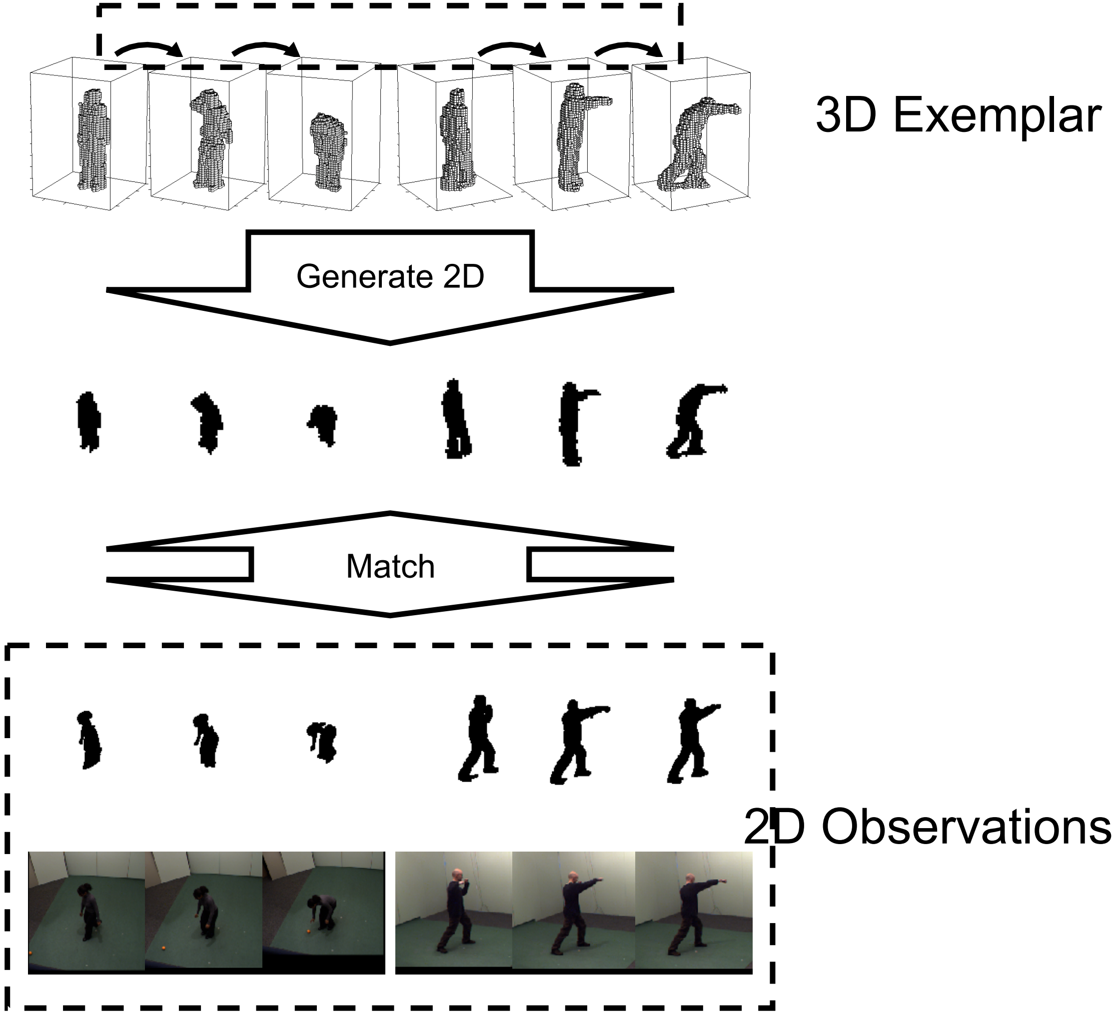
[A.A. Efros, A.C. Berg, G. Mori and J. Malik'03]

Action Modeling

3D Exemplar based approach

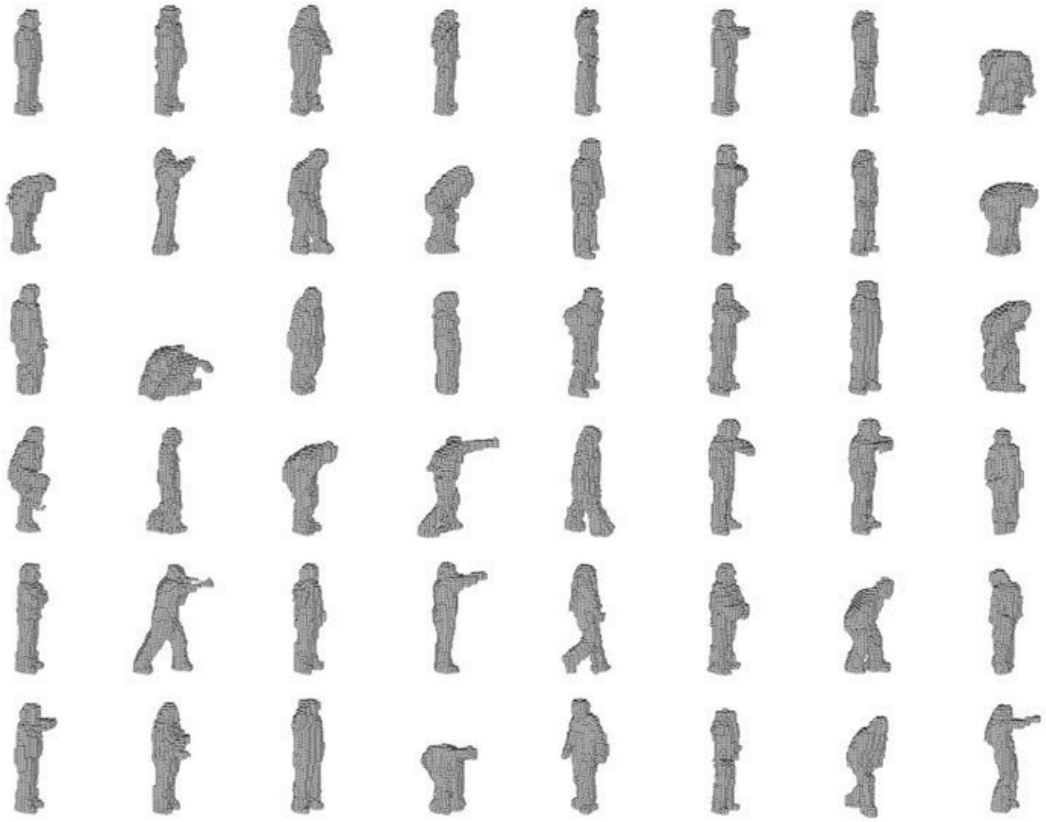
➔ The kinematic model is replaced by learned exemplars.

➔ Generative approach.



[Weinland & al.'07]

Action Modeling



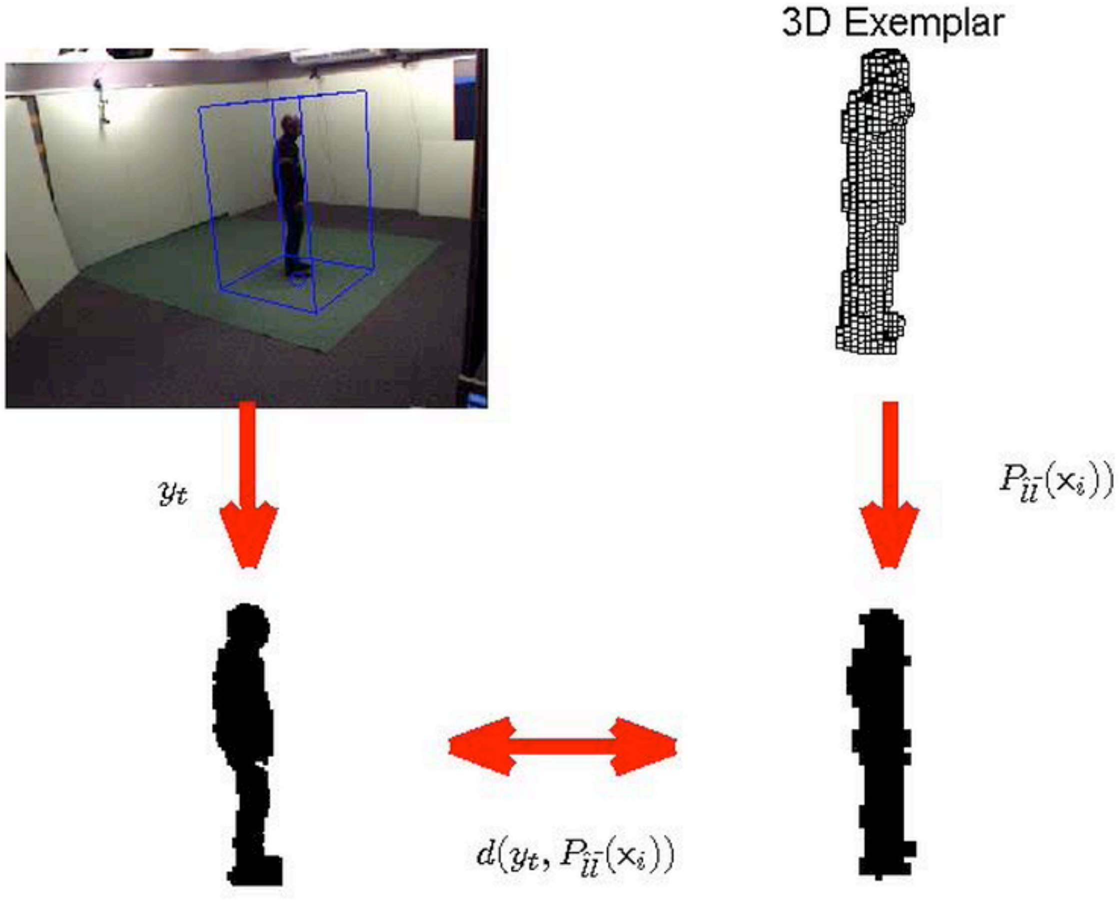
Exemplars associated with a database composed of 11 actions-10acteurs

Distance Function

- Silhouette-to-Silhouette Matching
  - Assumes background subtraction
  - d = Euclidean distance between silhouettes

- Silhouette-to-Edge Matching
  - No background subtraction
  - d = Chamfer Distance between silhouette and edge image

Action Modeling



[Weinland & al.'07]

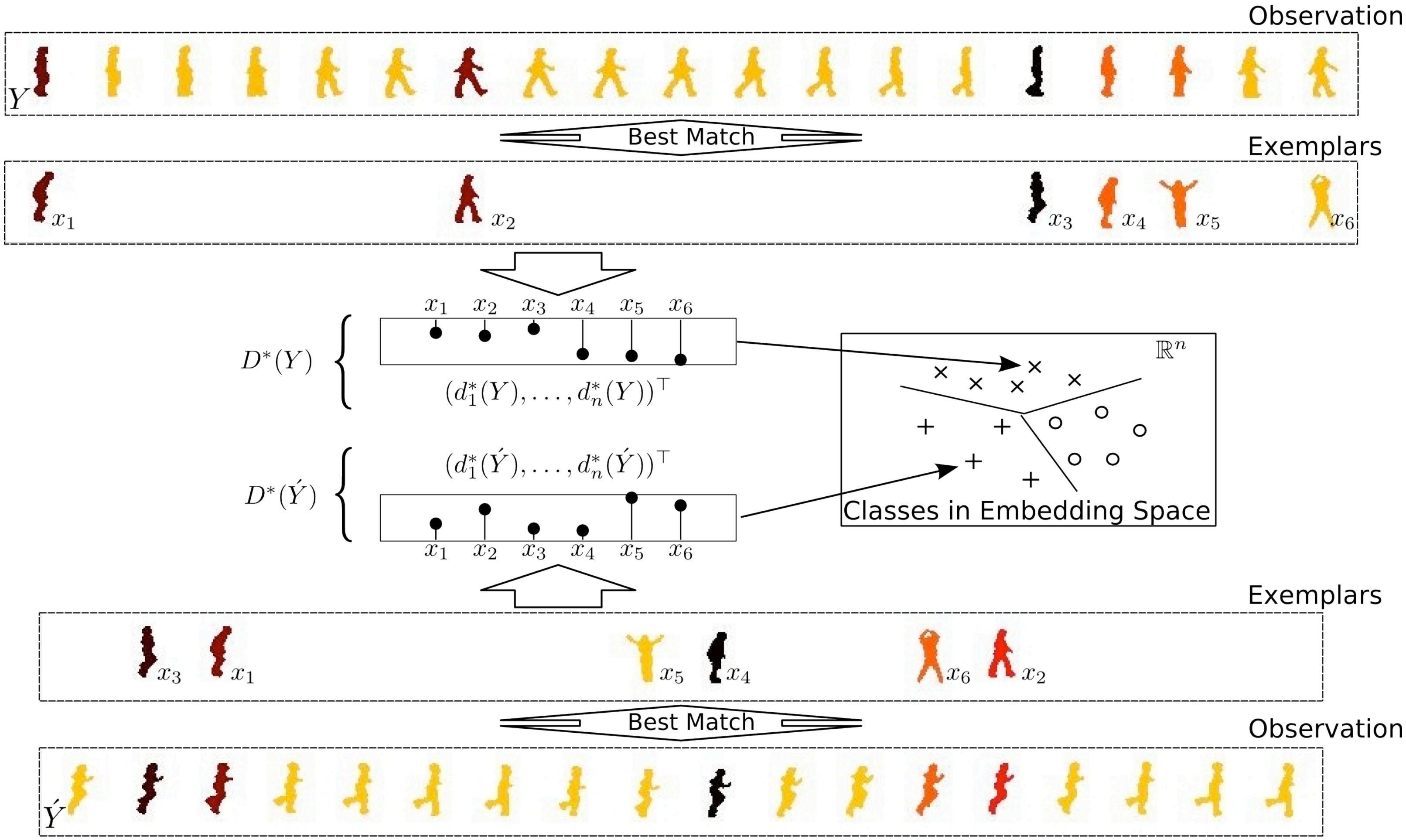
Action Modeling: Dynamics

- Traditional representations explicitly or implicitly model dynamics.

Weizmann-Dataset [Blank05]

► **Is dynamic required to identify an action ?**

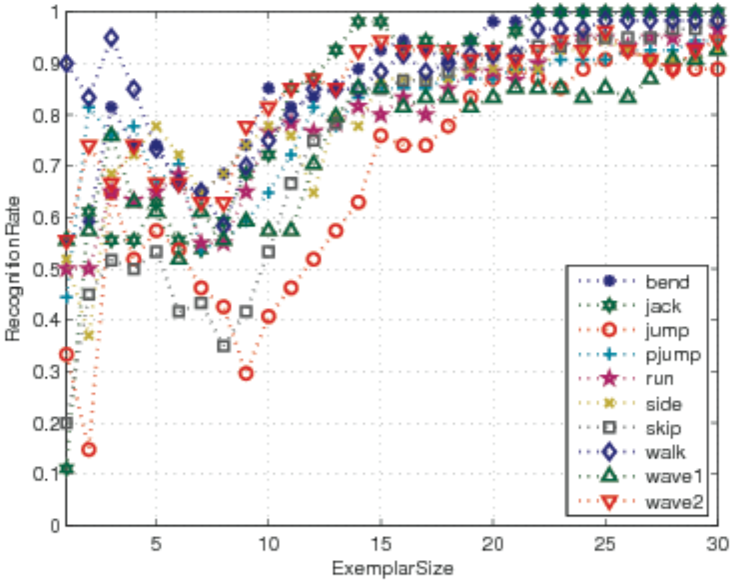
Exemplar-based Embedding



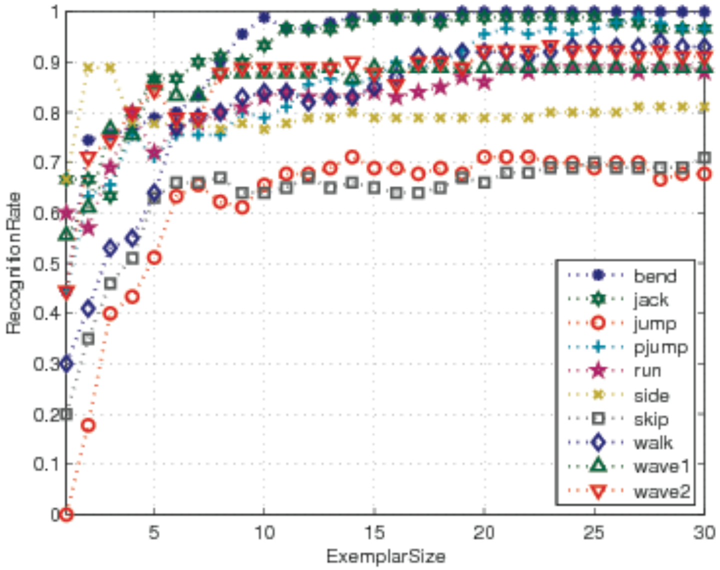
- Intuition: similar sequences will yield similar proximities to discriminative exemplars. Hence their point representation in R^n should be close.

Results

- ►Weizmann dataset : 10 actions, 9 actors, 2D silhouettes



Silhouettes



edges

[CVPR'08]

Conclusion

- Shape Modeling :
  - Industrial step (numerous start-ups and industrial projects such as 3D TV)
  - Difficulties still to be solved: appearance, temporal coherence are challenging.
- Motion Modeling:
  - Markerless approaches not really robust.
  - Difficulties: adequation model->observations hence holistic approaches ?!

Conclusion

- Action Modeling:
  - Simple actions only yet, applications to games.
  - Difficulties: the semantic interpretation is complicated due tothe dimension of the space considered. To reduce such dimension, use context ?