3D Object Retrieval

From Shape Description to Similarity Learning

Ceyhun Burak Akgül, PhD
Vistek ISRA Vision
www.vistek-isravision.com
www.cba-research.com

March 2010

Context

Ceyhun Burak Akgül
PhD Thesis [November 2007]
Density-Based Shape Descriptors and Similarity Learning for 3D Object Retrieval
Available at www.cba-research.com

Prof. Bülent Sankur (Advisor)
Boğaziçi University EE Dept. Dept., Istanbul, Turkey

Prof. Francis Schmitt (Advisor)
Télécom ParisTech (ENST) Signal & Image Proc. Dept., Paris, France.

Assoc. Prof. Yücel Yemez (Co-Advisor)
Koç University Computer Engineering Dept., Istanbul, Turkey

3DOR: Applications

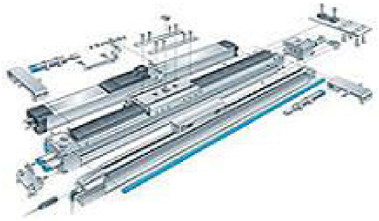
Video Games

Medicine

Product Design

3D data and objects are everywhere!

Engineering



Car Industry

Cultural Heritage

3DOR: Applications

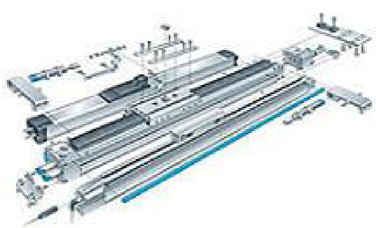
Video Games

Medicine

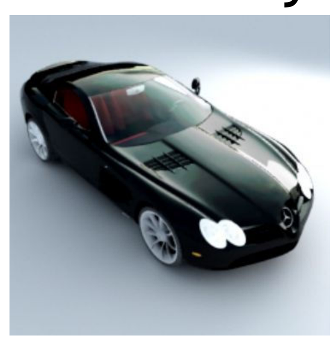
Product Design

How to organize a 3D database?
How to look for an item in a 3D database?

Engineering



Car Industry



Cultural Heritage

3DOR: Why?

Text-based search and manual annotation have limitations:

= “amphora with two handles”

= ?

How many items can one annotate unambiguously?

3DOR: Why?

Text-based search and manual annotation have limitations:

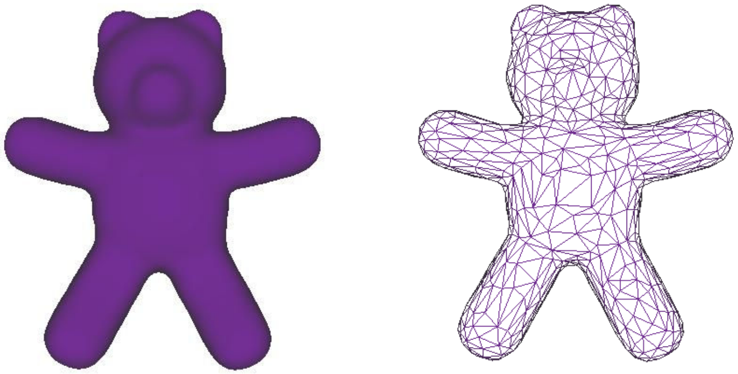
= “amphora with two handles”

= ?

How many items can one annotate unambiguously?

Search and retrieve by content similarity

3DOR: What is a 3D object?

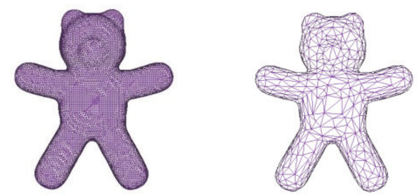


3D Object = A Mesh in 3D

- ✓ Completely given in 3D space
- ✓ No perspective distortion
- ✓ No clutter, no occlusion
- ✓ Single object in an empty background

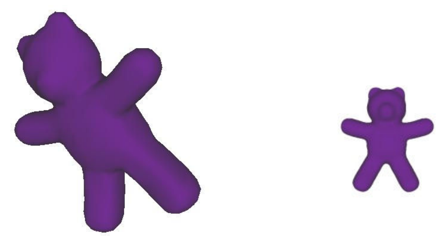
high **low**

- **Arbitrary resolution** →

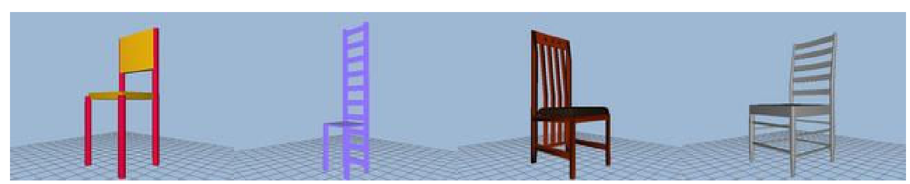


BUT!

- **Arbitrary pose and scale** →



- **High variability** →



Add to these **small shape variations** which don't alter semantics

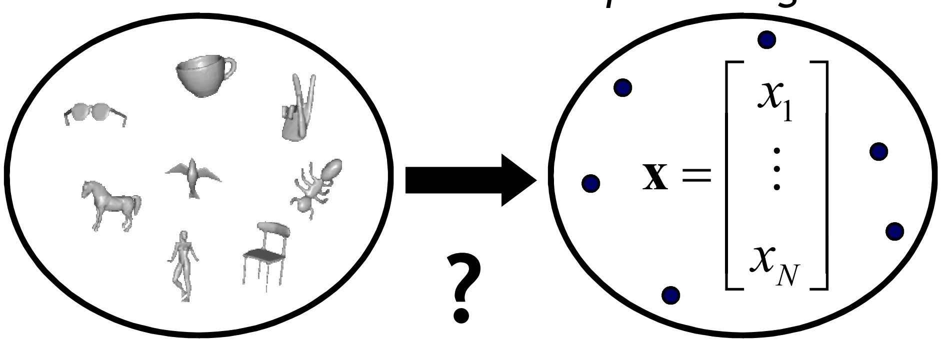
3DOR: Main Research Problems

1 Content Description for 3D Shapes

Shape Space

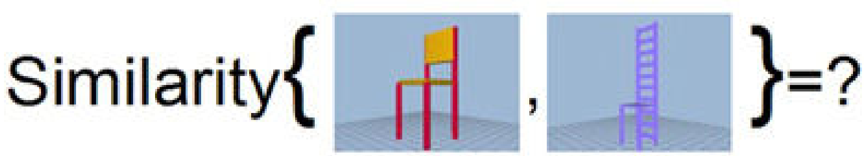
Descriptor Space

Descriptor = High-dim vector



- ***Intrinsic shape characteristics***
- ***Robustness***
- ***Fast computation***
- ***Low storage cost***
- ***Geometric invariance***
- ***High discrimination ability***

2 Similarity Learning for CBR



- ***Similarity models for retrieval***
- ***Statistical learning***
    - *Optimization criteria*
- ***Limited supervision***
    - *Small training set*
- ***Search paradigms***
    - *Relevance feedback*

Shape Description

3DOR: Shape Description

3D Shape Descriptors

Histogram-Based

- Cord and Angle Hist (CAH)
- Shape Distributions (SD)
- Generalized Shape Distr. (GSD)
- Shape Histograms (SECSHELL)
- Extended Gaussian Images (EGI)
- Shape Spectrum Desc. (SSD)
- 3D Hough Transform (3DHT)

Transform-Based

- VOXEL-3DFT
- Radial Cosine Transform
- Angular Radial (ART)
- Spherical Harmonics (SH)
  - Rotation Inv. SH (RiSH)
  - REXT SH
- Spherical Wavelet (SW)
- Concrete Radialized Spherical Projection (CRSP)

2D "Image"-Based

- Light Field Images (LFD)
- Depth Buffer Images (DBI)
- Silhouette (SIL)

Others

- Graph-Based
  - Reeb Graphs
- Moments-Based
  - Zernike Moments
- Spin Images
- Symmetry Descriptors

3DOR: Shape Description

3D Shape Descriptors

Histogram-Based

- Cord and Angle Hist (CAH)
- Shape Distributions (SD)
- Generalized Shape Distr. (GSD)
- Shape Histograms (SECSHELL)
- Extended Gaussian Images (EGI)
- Shape Spectrum Desc. (SSD)
- 3D Hough Transform (3DHT)

Transform-Based

- VOXEL-3DFT
- Radial Cosine Transform
- Angular Radial (ART)
- Spherical Harmonics (SH)
  - Rotation Inv. SH (RiSH)
  - REXT SH
- Spherical Wavelet (SW)
- Concrete Radialized Spherical Projection (CRSP)

2D "Image"-Based

- Light Field Images (LFD)
- Depth Buffer Images (DBI)
- Silhouette (SIL)

Others

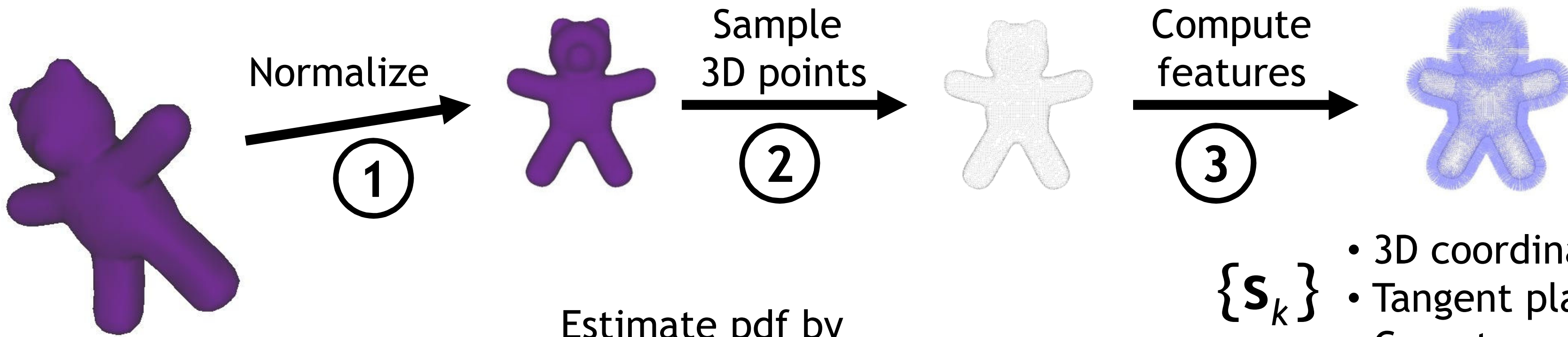
- Graph-Based
  - Reeb Graphs
- Moments-Based
  - Zernike Moments
- Spin Images
- Symmetry Descriptors

Density-Based Framework (DBF)

- ***Unifying framework***
- ***A multivariate extension***
- ***Kernel Density Estimation (KDE) = Parzen windows***
  - smoother than histogram
  - less sensitive to feature measurement errors than histogram

3DOR: Shape Description

DBF: Overview



$\{\mathbf{s}_k\}$

- 3D coordinates
- Tangent plane
- Curvature related

Multivariate local characterization

4

Estimate pdf by
<u>**Kernel Density Estimation**</u>

$$f(\mathbf{t}_n) = C\sum_k exp\left(-\frac{1}{2}\left\|\mathbf{H}^{-1}(\mathbf{t}_n - \mathbf{s}_k)\right\|^2\right)$$

$\mathbf{H}$: Smoothing parameter

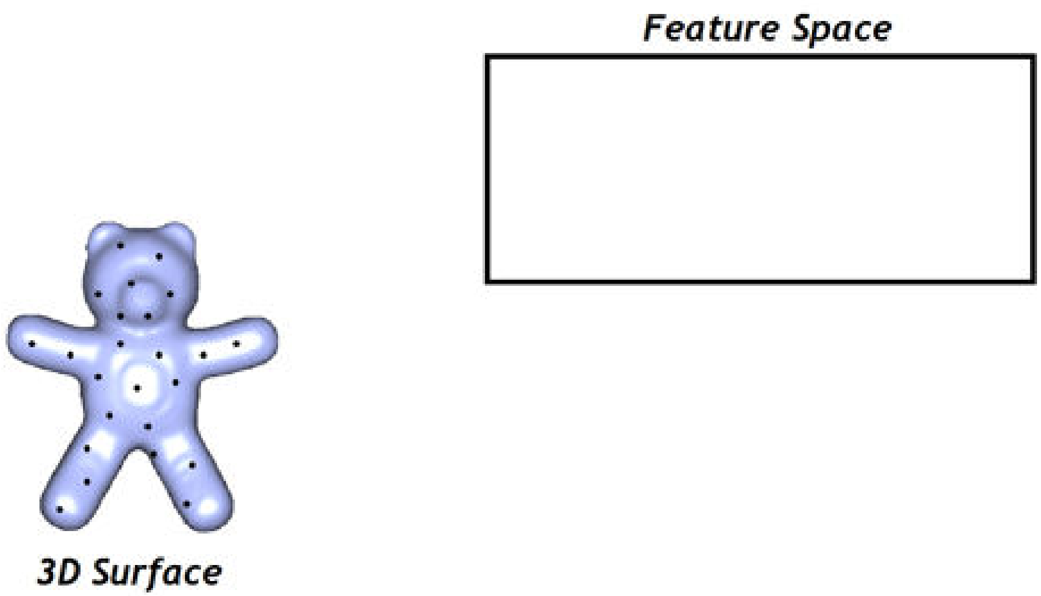
$$\mathbf{f} = \left[f(\mathbf{t}_1) \quad \cdots \quad f(\mathbf{t}_n) \quad \cdots \quad f(\mathbf{t}_N)\right]$$

Global 3D shape descriptor

3DOR: Shape Description

DBF: Robustness

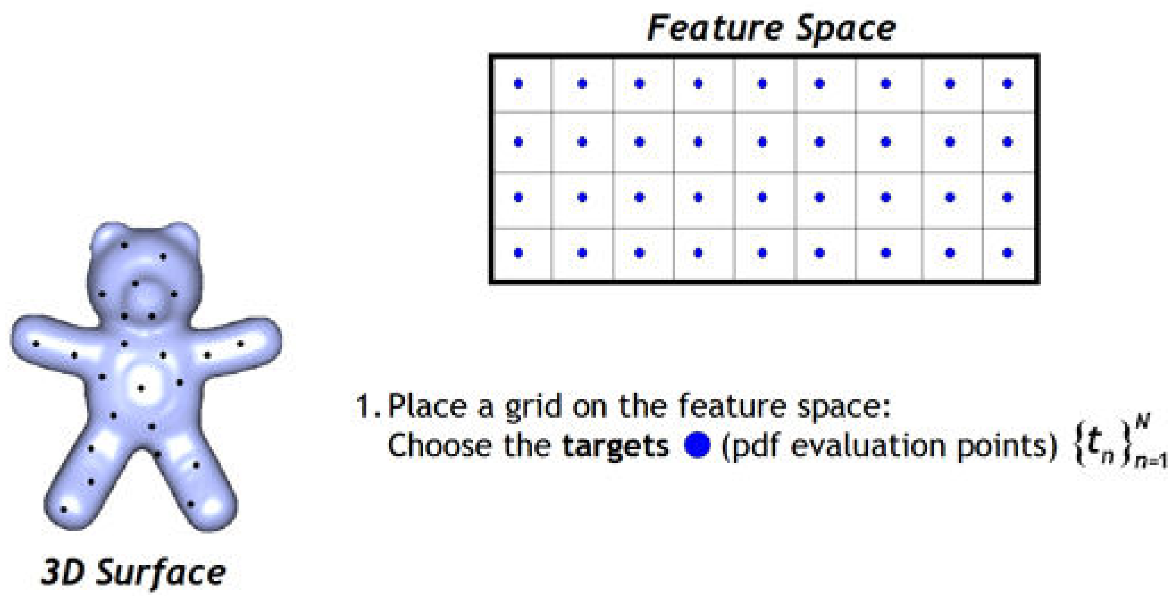
✓ Insensitive to small shape variations & errors



3DOR: Shape Description

DBF: Robustness

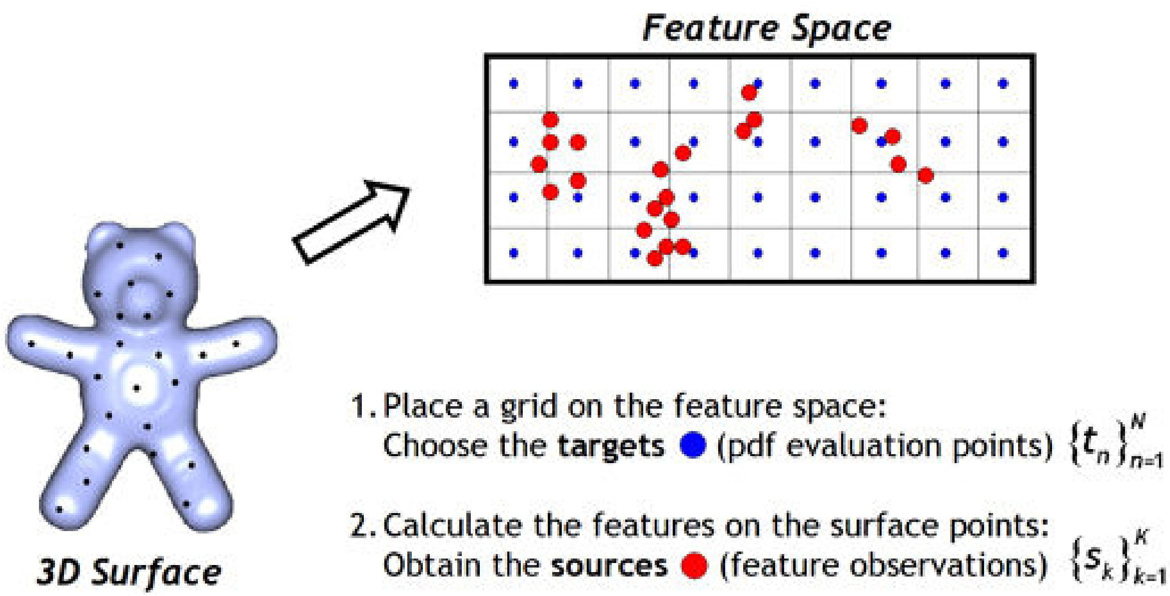
✓ Insensitive to small shape variations & errors



3DOR: Shape Description

DBF: Robustness

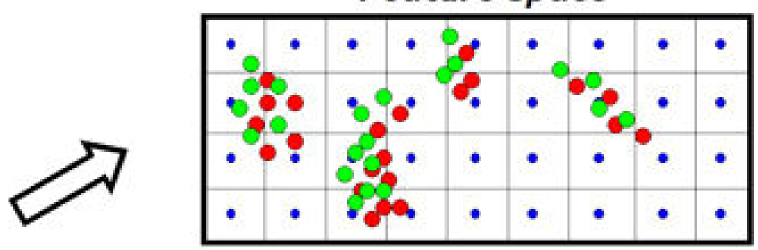
✓ Insensitive to small shape variations & errors

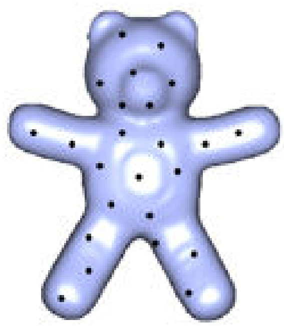


3DOR: Shape Description

DBF: Robustness

✓ **Insensitive to small shape variations & errors**





1. Place a grid on the feature space:
   Choose the **targets** (pdf evaluation points) $\{t_n\}_{n=1}^{N}$

2. Calculate the features on the surface points:
   Obtain the **sources** (feature observations) $\{s_k\}_{k=1}^{K}$

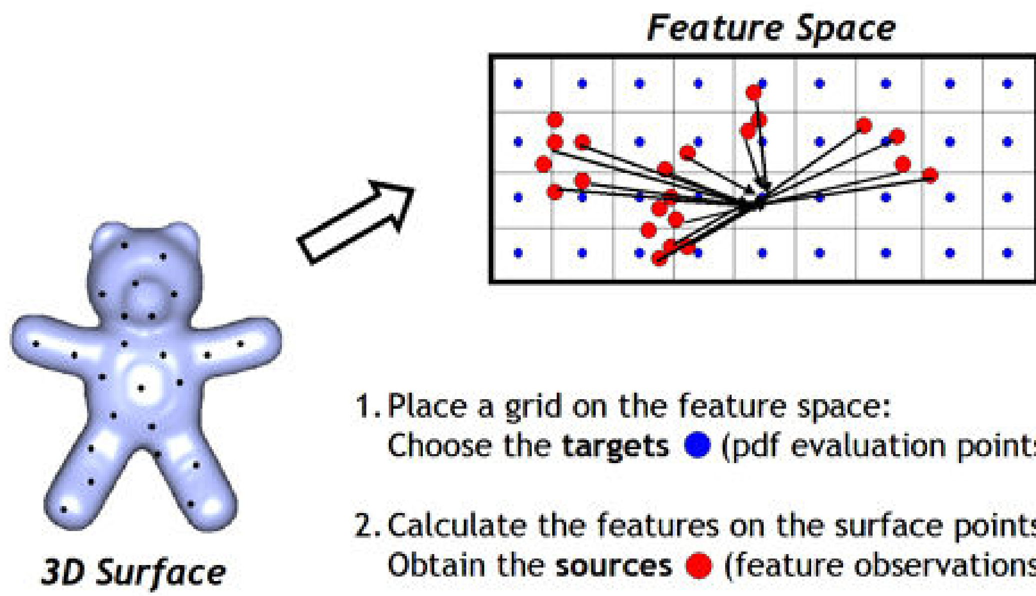
ERRORS
- Scaling
- Rotation
- Translation
- Noise
- Small shape variation
- Mesh degeneracy
- ...

We can use the histogram as the pdf estimate, but…

3DOR: Shape Description

DBF: Robustness

✓ Insensitive to small shape variations & errors

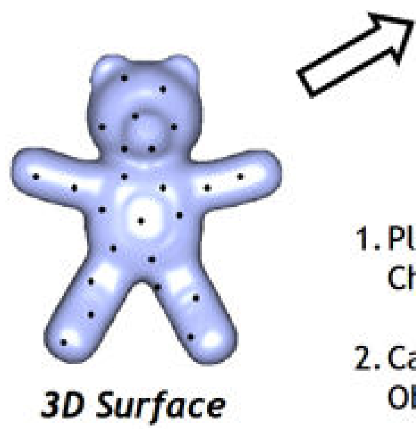


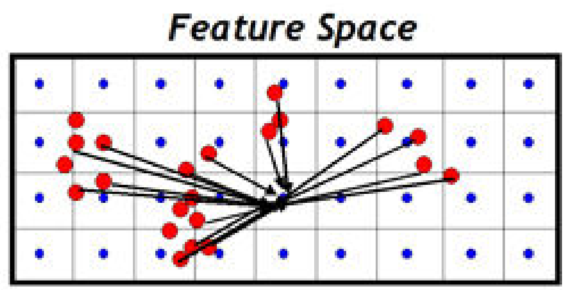
1. Place a grid on the feature space:
   Choose the **targets** ● (pdf evaluation points) $\{t_n\}_{n=1}^{N}$
2. Calculate the features on the surface points:
   Obtain the **sources** ● (feature observations) $\{s_k\}_{k=1}^{K}$
3. Estimate the pdf by **Kernel Density Estimation (KDE):**
   - **Copes with measurement uncertainties**
   - **Soft assignment strategy**
   - **Insensitive to the placement of the grid**
   - **Smoother than the histogram**

3DOR: Shape Description

DBF: Robustness

✓ Insensitive to small shape variations & errors





1. Place a grid on the feature space:
   Choose the **targets** ● (pdf evaluation points) $\{t_n\}_{n=1}^{N}$
2. Calculate the features on the surface points:
   Obtain the **sources** ● (feature observations) $\{s_k\}_{k=1}^{K}$
3. Estimate the pdf by **Kernel Density Estimation (KDE):**

$$f(t_n) = \sum_{k=1}^{K} w_k |H|^{-1} \mathcal{K}\left(H^{-1}(t_n - s_k)\right)$$

Density-Based Descriptor → $[f(t_1), f(t_2), \ldots, f(t_N)]$

3DOR: Shape Description

DBF: Descriptor Computation

✓ **Fast computation:** less than 1 sec per 3D object on average

$$f(t_n) = \sum_{k=1}^{K} w_k \, |H|^{-1} \mathcal{K}\left(H^{-1}(t_n - s_k)\right)$$

- Direct Evaluation $\longrightarrow O(KN)$
- When **the kernel** $\mathcal{K}$ is **Gaussian** $\longrightarrow O(K+N)$
  Fast Gauss Transform (FGT)
  [Greengard and Strain, 1991; Yang et al., 2003]

Example: $K = 130000$ and $N = 1024$

- **Direct** $\longrightarrow$ 125 secs
- **FGT** $\longrightarrow$ 2.5 secs

3DOR: Shape Description

DBF: Geometric Invariance

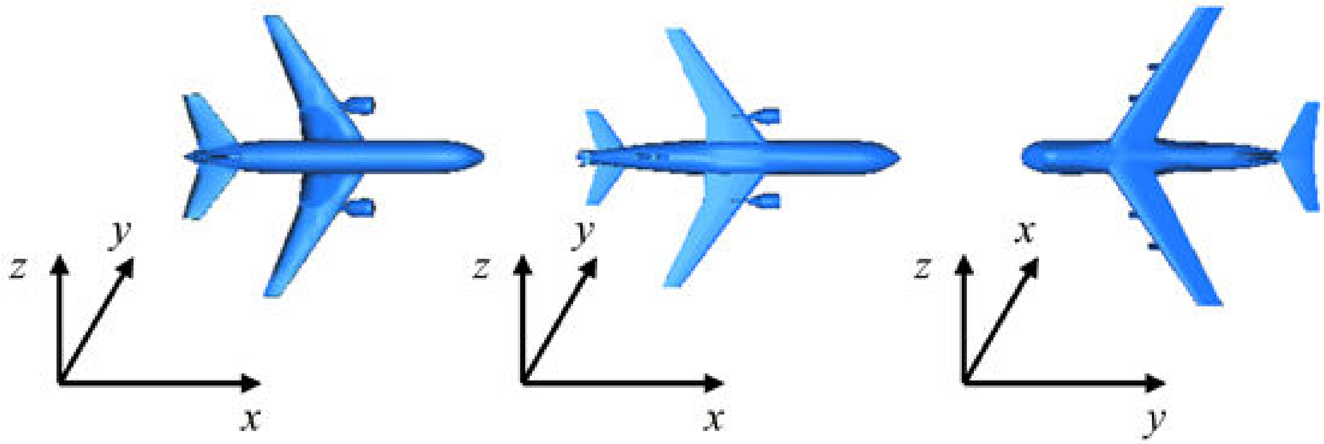
✓ **Correspondence-free shape alignment**

1. **Invariance by feature design**: might loose shape information
2. **Invariance by pre-normalization**: not always stable
3. **Invariance at matching**
   → Evaluate the similarity under all possible transformations and take the minimum
   ✕ Costly if descriptor should be computed for every possible transformation

3DOR: Shape Description

DBF: Geometric Invariance

✓ **Correspondence-free shape alignment**



$$3! = 6 \text{ axis relabelings}$$
$$2^3 = 8 \text{ polarity assignments}$$
$$\Rightarrow 6 \times 8 = 48 \text{ axis configurations}$$

The complexity of invariant matching is 48 times the complexity of vector-to-vector distance computation.

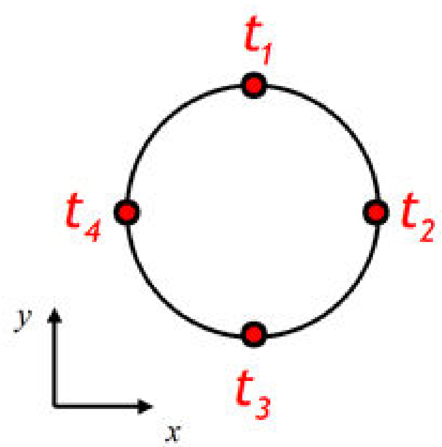
3DOR: Shape Description

DBF: Geometric Invariance

✓ **Correspondence-free shape alignment**

In DBF, for certain class of transformations, there is no need to recompute the descriptor:

→ **Just permute the vector entries!**



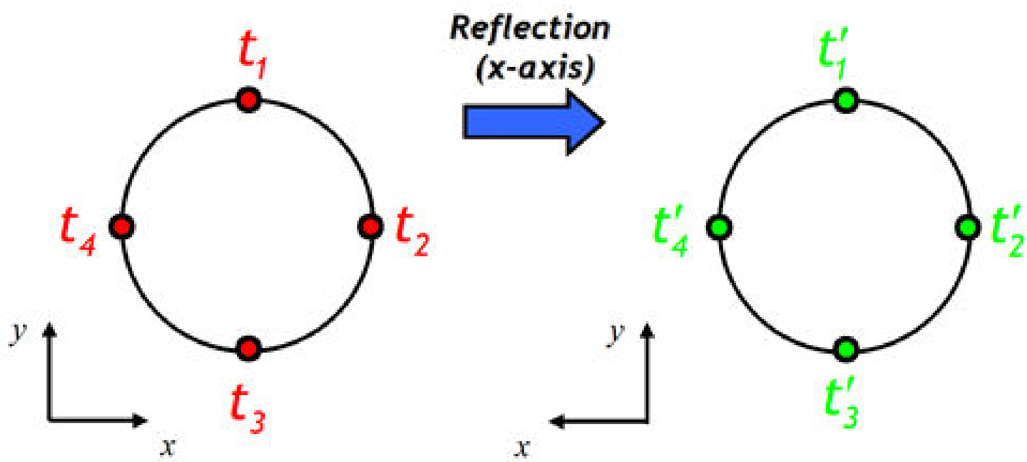
3DOR: Shape Description

DBF: Geometric Invariance

✓ **Correspondence-free shape alignment**

In DBF, for certain class of transformations,
there is no need to recompute the descriptor:

→ **Just permute the vector entries!**



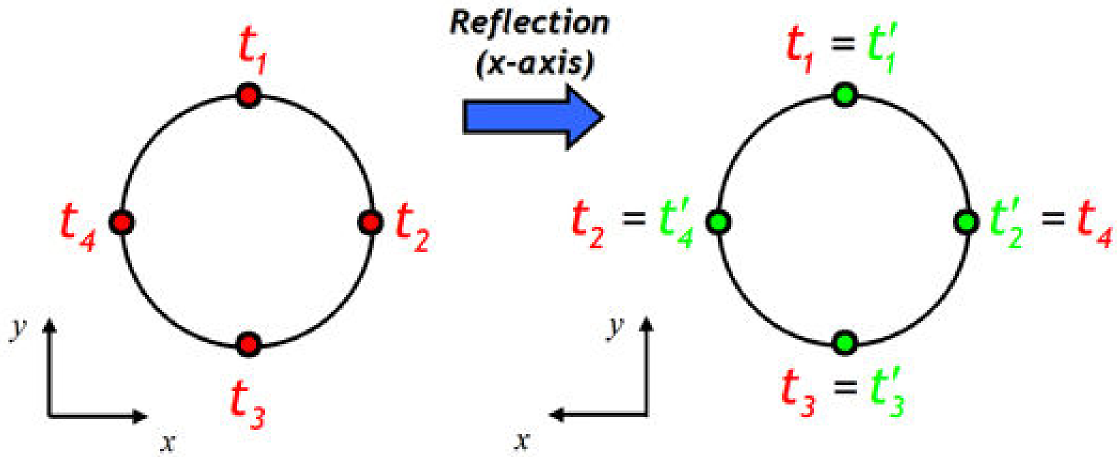
3DOR: Shape Description

DBF: Geometric Invariance

✓ **Correspondence-free shape alignment**

In DBF, for certain class of transformations,
there is no need to recompute the descriptor:

→ **Just permute the vector entries!**

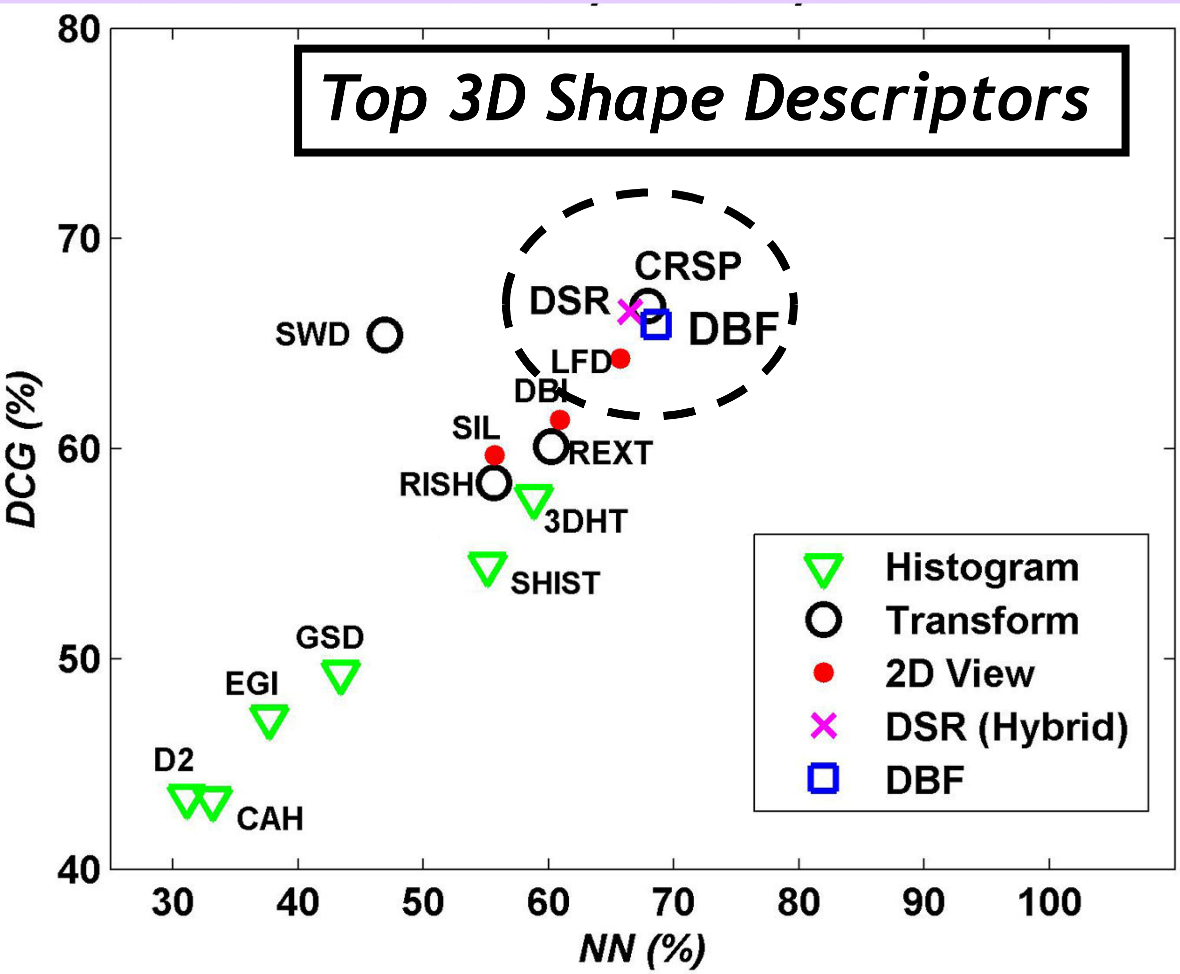


*The target set should be **closed** under the action of the transformation*

3DOR: Shape Description

DBF: Retrieval Performance

✓ Highly discriminative

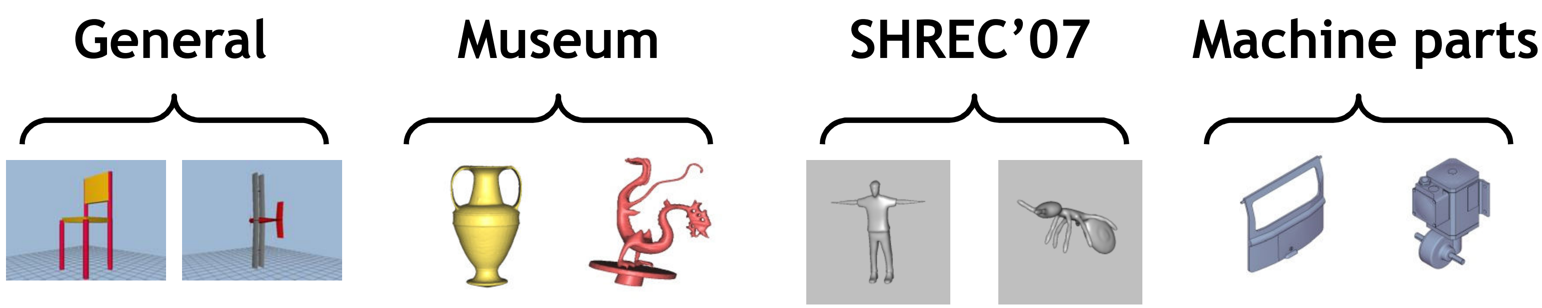


DBF: Density-Based Framework
CRSP: Concrete Radialized Spherical Projection
DSR: Hybrid DBI-SIL-REXT
LFD: Light-Field Descriptor
DBI: Depth-Buffer Images
SIL: Silhouette
REXT: Radialized Extent Function
SWD: Spherical Wavelet Descriptor
RISH: Rotation-Invariant Spherical Harmonics
3DHT: 3D Hough Transform
SHIST: Shape Histograms
GSD: Generalized Shape Distributions
EGI: Extended Gaussian Images
D2: Shape Distributions
CAH: Cord and Angle Histograms

3DOR: Shape Description

DBF: Retrieval Performance

✓ Highly discriminative



	Princeton	Sculpteur	Watertight	ESB
DBF	65.9%	**78.3%**	**86.7%**	**75.7%**
DSR	**66.5%**	76.6%	83.2%	74.1%

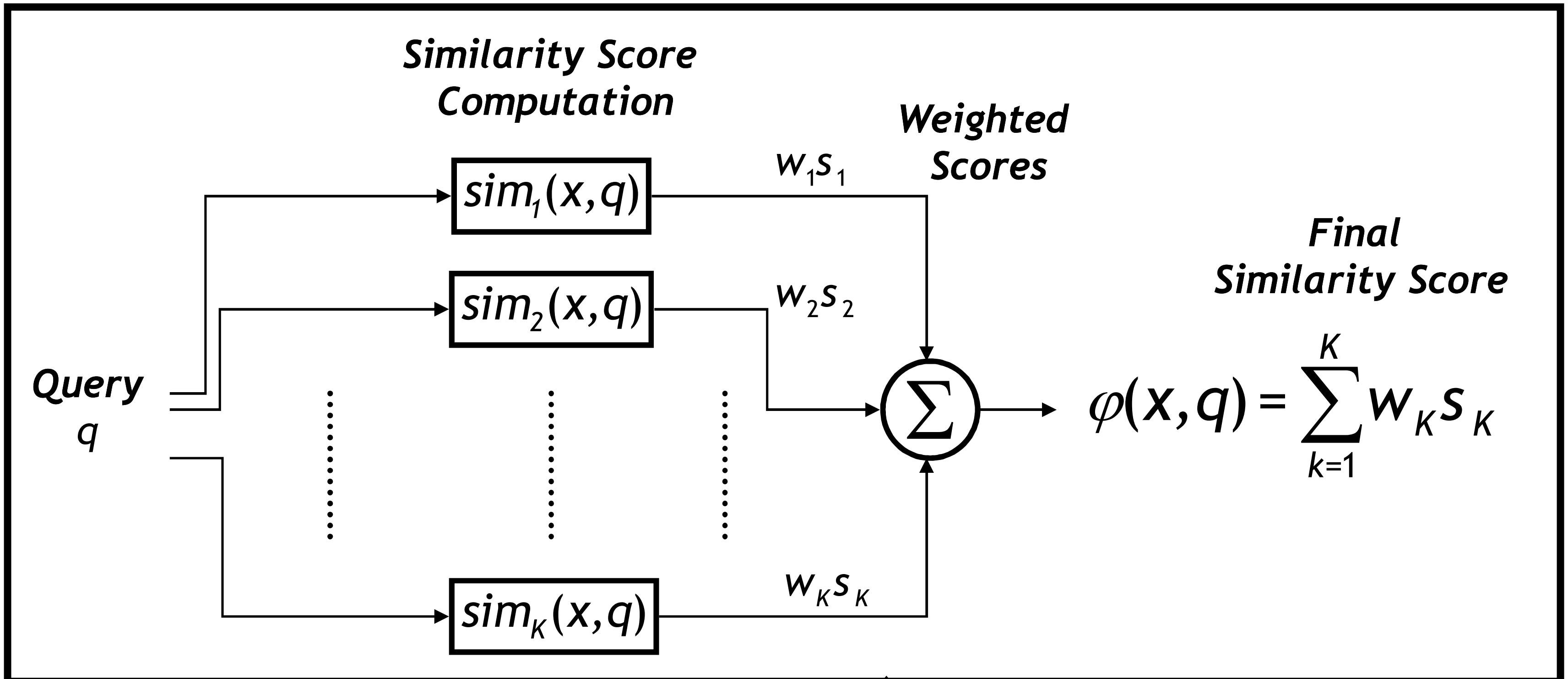
DSR: Hybrid descriptor proposed in D. Vranic's PhD Thesis, 2004.

NOTE: Displayed are **% DCG** values, one of the most popular retrieval statistics.

Similarity Learning

3DOR: Similarity Learning

Score Fusion: Overview



$\{s_k\}_{k=1}^{K}$ ***Similarity scores from different descriptors !***

$\{w_k\}_{k=1}^{K}$ ***Weights?***

3DOR: Similarity Learning

Score Fusion: Contributions

- ✓ **Linear similarity model:** intuitively appealing
- ✓ **Original: Ranking Risk Minimization**
  - no prior work in visual retrieval domain
- ✓ **Flexible:** can be applied to broader CBR domains
- ✓ **Fast computation and convergence**
- ✓ ~10% performance increase using **relevance feedback**

3DOR: Similarity Learning

Score Fusion: Approach

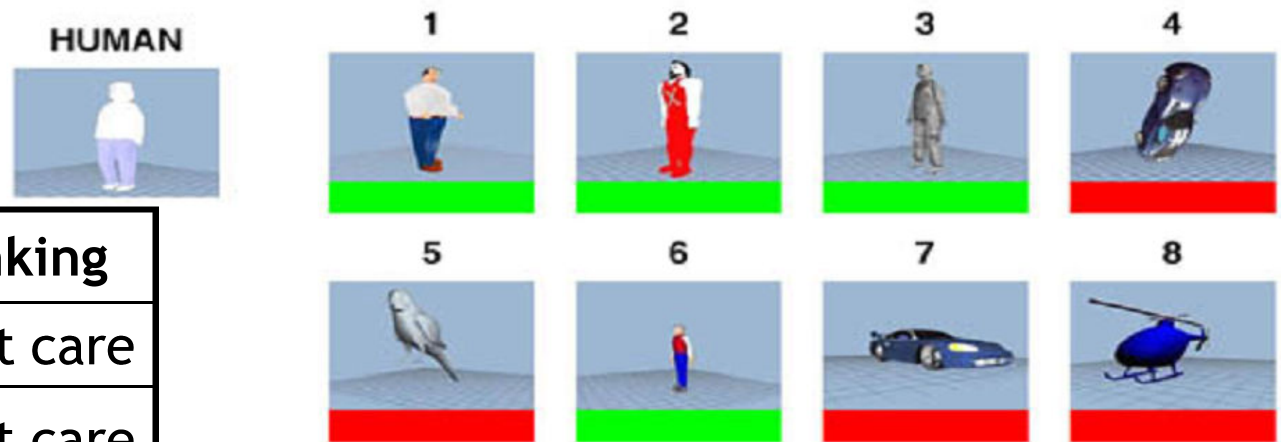
Estimation of the optimal set of weights

1. What kind of criterion to optimize?
   - **Empirical Ranking Risk (ERR)**
     - Given a query, *relevant* entities should be ranked higher than *less relevant* and/or *not-relevant* ones
     - ERR is the **count of misordered pairs of entities**

2. How to incorporate supervision?
   - Ontology-driven search
   - Relevance feedback

3DOR: Similarity Learning

Score Fusion: Empirical Ranking Risk



Pairs	Ranking
(1,2)	Don't care
(1,3)	Don't care
(1,4)	✓
…	
(4,6)	×
…	
(5,6)	×
…	

Number of misrankings = 2

Desideratum

Minimize the number of such errors to learn the optimal set of weights

3DOR: Similarity Learning

Score Fusion: Ranking Risk Minimization

Similarity function $\varphi(x,q) = \sum_k w_k s_k = \langle \mathbf{w}, \mathbf{s} \rangle$ should satisfy :

$$\begin{array}{ll} \varphi(x,q) > \varphi(x',q) & \text{if } x \text{ is more relevant to } q \text{ than } x', \\ \varphi(x,q) < \varphi(x',q) & \text{otherwise.} \end{array}$$

Let y encode the relevance of x to q :

$$y = \begin{cases} +1 & \text{if } x \text{ is relevant to } q, \\ -1 & \text{if not.} \end{cases}$$

Then, we can write :

$$\begin{array}{ll} \langle \mathbf{w}, \mathbf{s} \rangle > \langle \mathbf{w}, \mathbf{s}' \rangle & \text{if } y - y' > 0, \\ \langle \mathbf{w}, \mathbf{s} \rangle < \langle \mathbf{w}, \mathbf{s}' \rangle & \text{if } y - y' < 0. \end{array} \quad \Rightarrow \quad \begin{array}{ll} \langle \mathbf{w}, \mathbf{s} - \mathbf{s}' \rangle > 0 & \text{if } y - y' > 0, \\ \langle \mathbf{w}, \mathbf{s} - \mathbf{s}' \rangle < 0 & \text{if } y - y' < 0. \end{array}$$

3DOR: Similarity Learning

Score Fusion: Ranking Risk Minimization

Let $z \triangleq sign(y - y')$ and $\mathbf{v} \triangleq \mathbf{s} - \mathbf{s}'$, then:

$$\Rightarrow \boxed{\begin{array}{ll} \langle \mathbf{w}, \mathbf{s} - \mathbf{s}' \rangle > 0 & \text{if } y - y' > 0, \\ \langle \mathbf{w}, \mathbf{s} - \mathbf{s}' \rangle < 0 & \text{if } y - y' < 0. \end{array}} \Rightarrow \boxed{\begin{array}{ll} \langle \mathbf{w}, \mathbf{v} \rangle > 0 & \text{if } z = +1, \\ \langle \mathbf{w}, \mathbf{v} \rangle < 0 & \text{if } z = -1. \end{array}}$$

Binary classification

- ❖ The original problem is transformed into one of **binary classification**
- → Any binary classifier can be used.
- ❖ Use **Support Vector Machines (SVM)**

3DOR: Similarity Learning

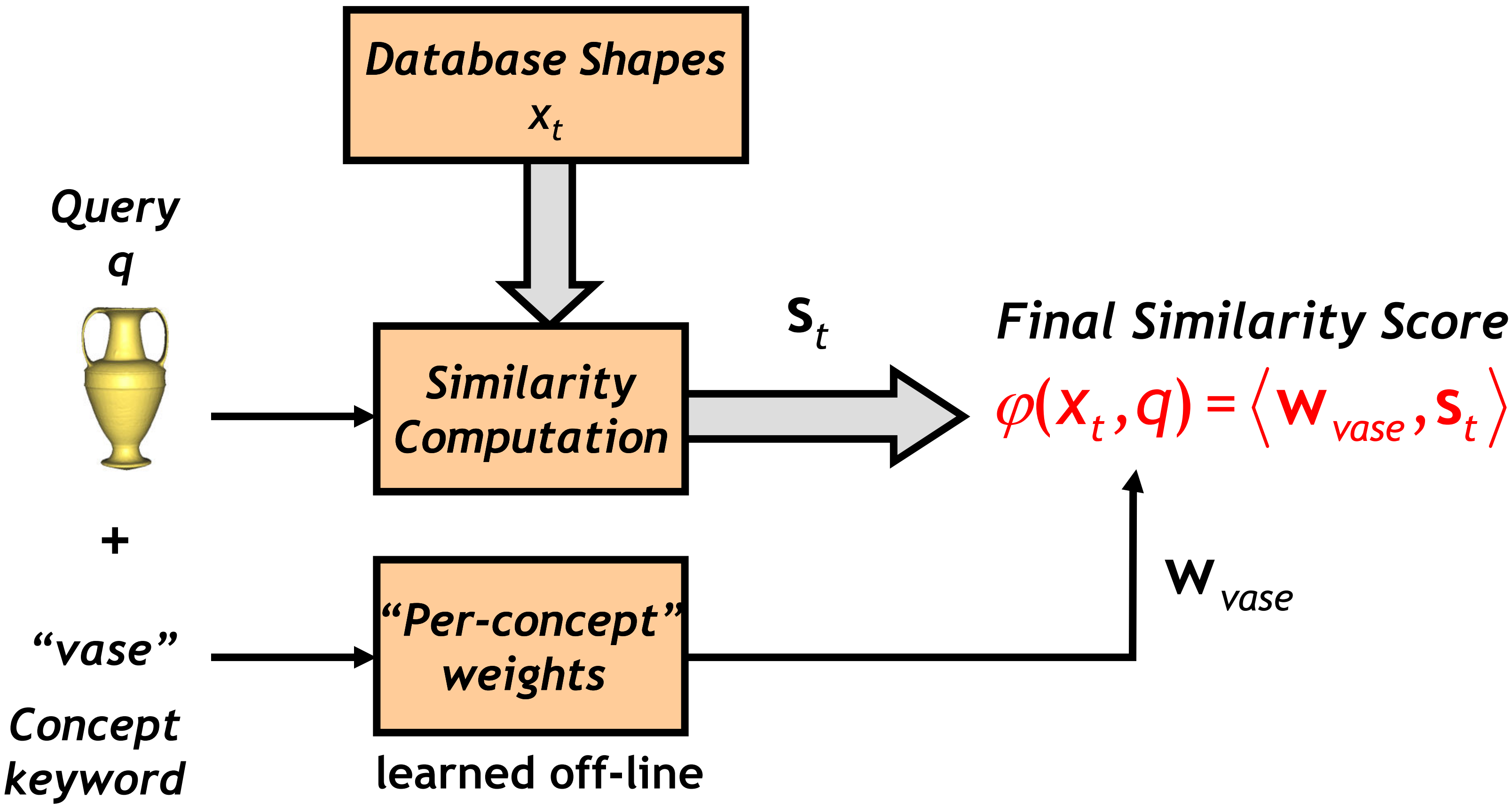
Score Fusion: Retrieval Protocols

1. **Bimodal**
   - Query: shape + concept keyword
   - *Offline* learning of **concept-specific** weights
   - Concepts ⇔ Ontologies

2. **Two-Round**
   - User is active: Relevance feedback
   - *Online* or *offline* learning of **query-specific** weights

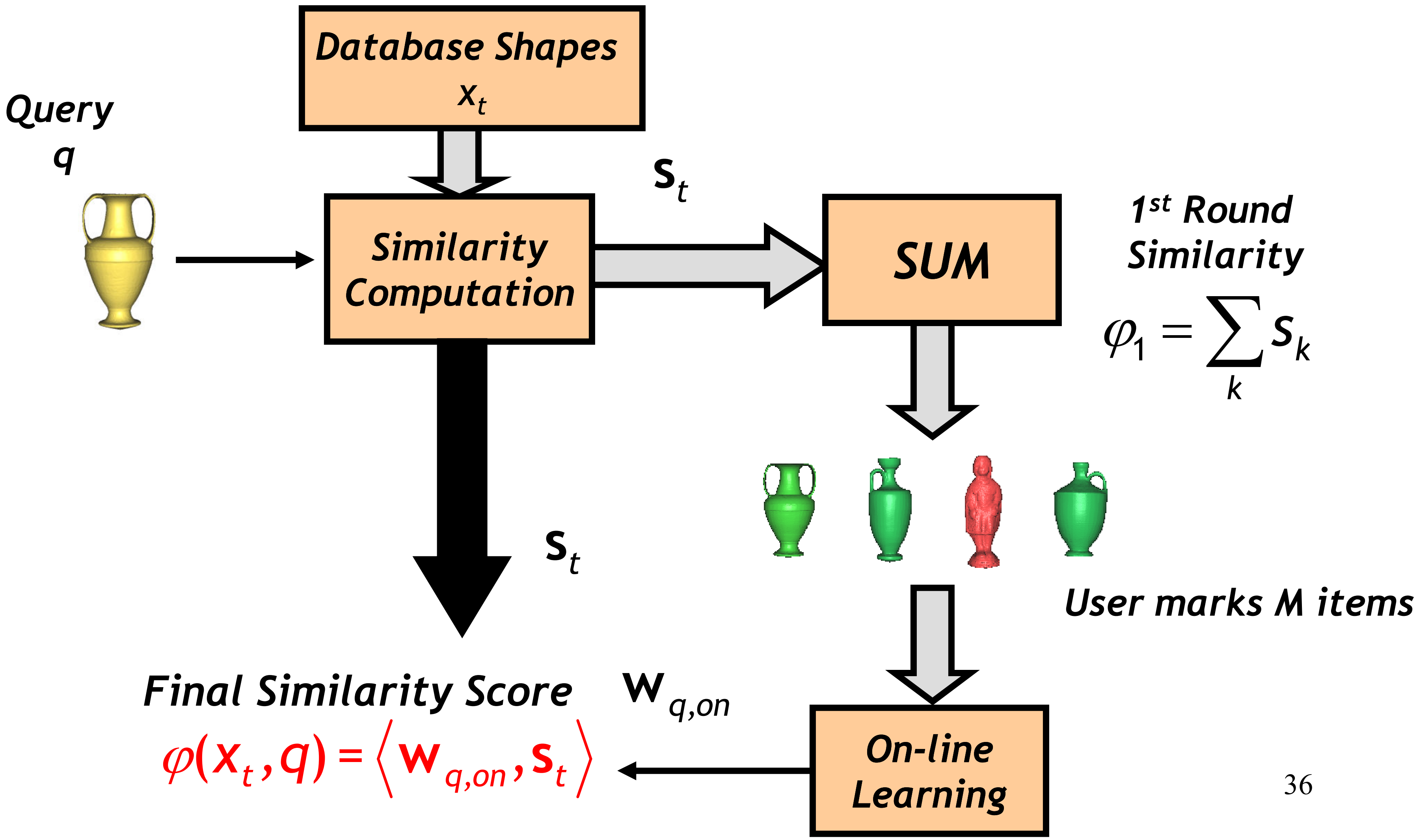
3DOR: Similarity Learning

Retrieval Protocols: Bimodal



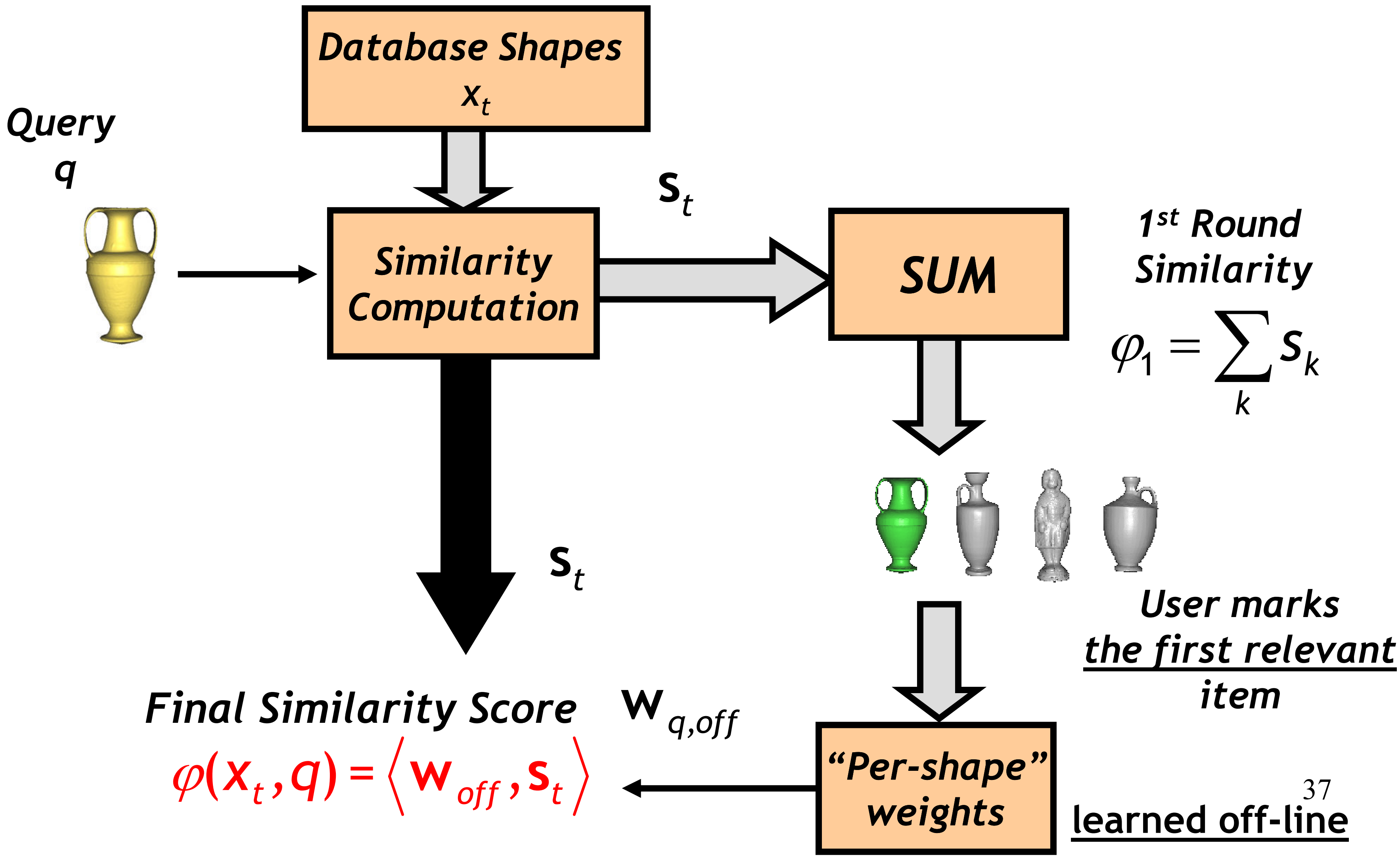
3DOR: Similarity Learning

Retrieval Protocols: Two-Round On-line



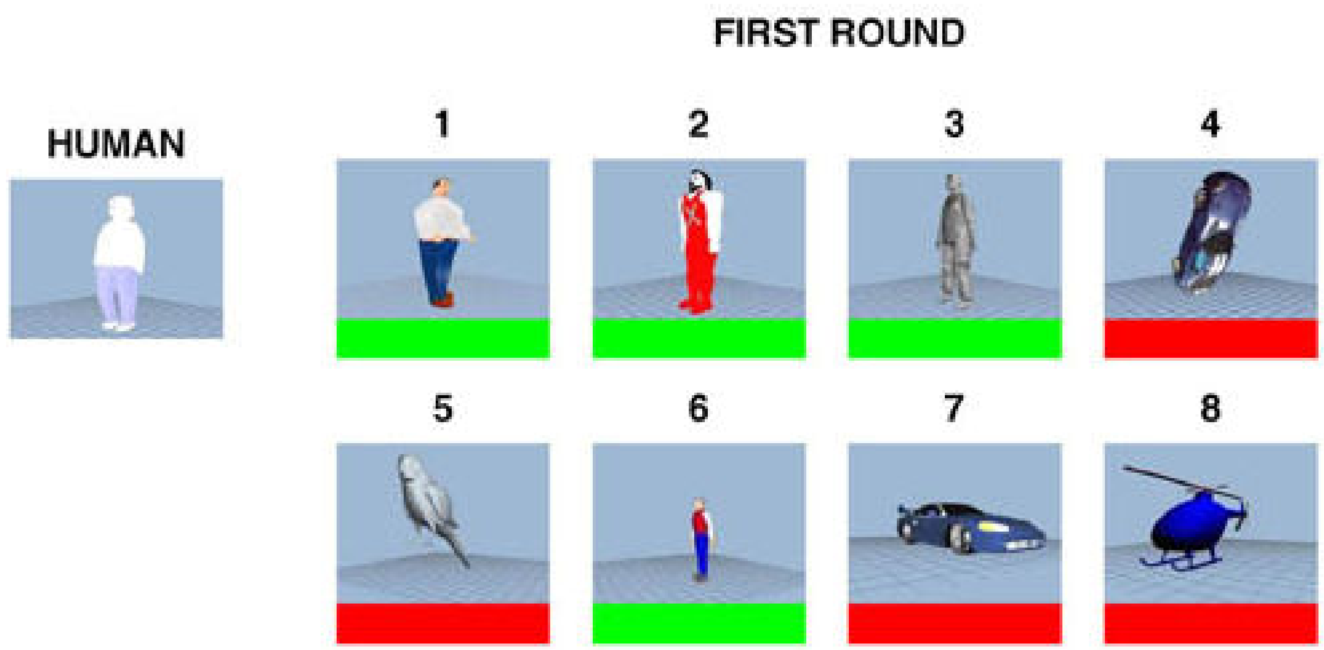
3DOR: Similarity Learning

Retrieval Protocols: Two-Round Off-line



3DOR: Similarity Learning

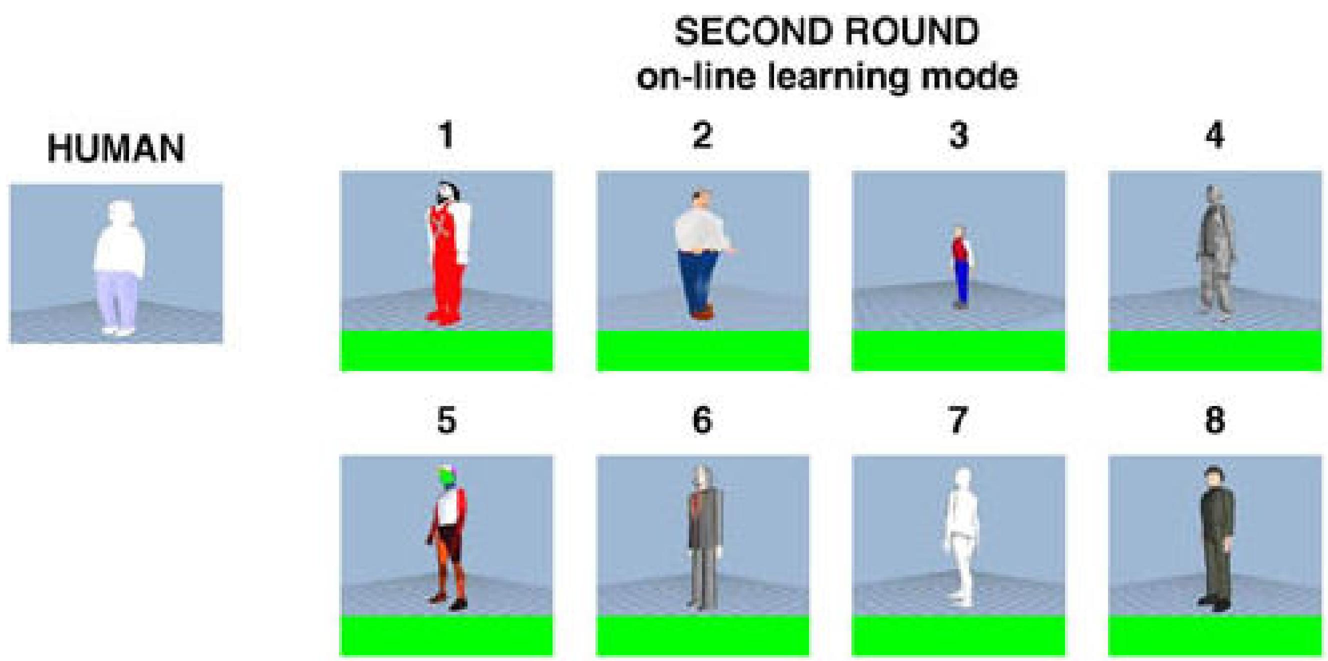
Two-Round On-line Example: “Human”



User marks 4 relevant and 4 non-relevant models among the first 8

3DOR: Similarity Learning

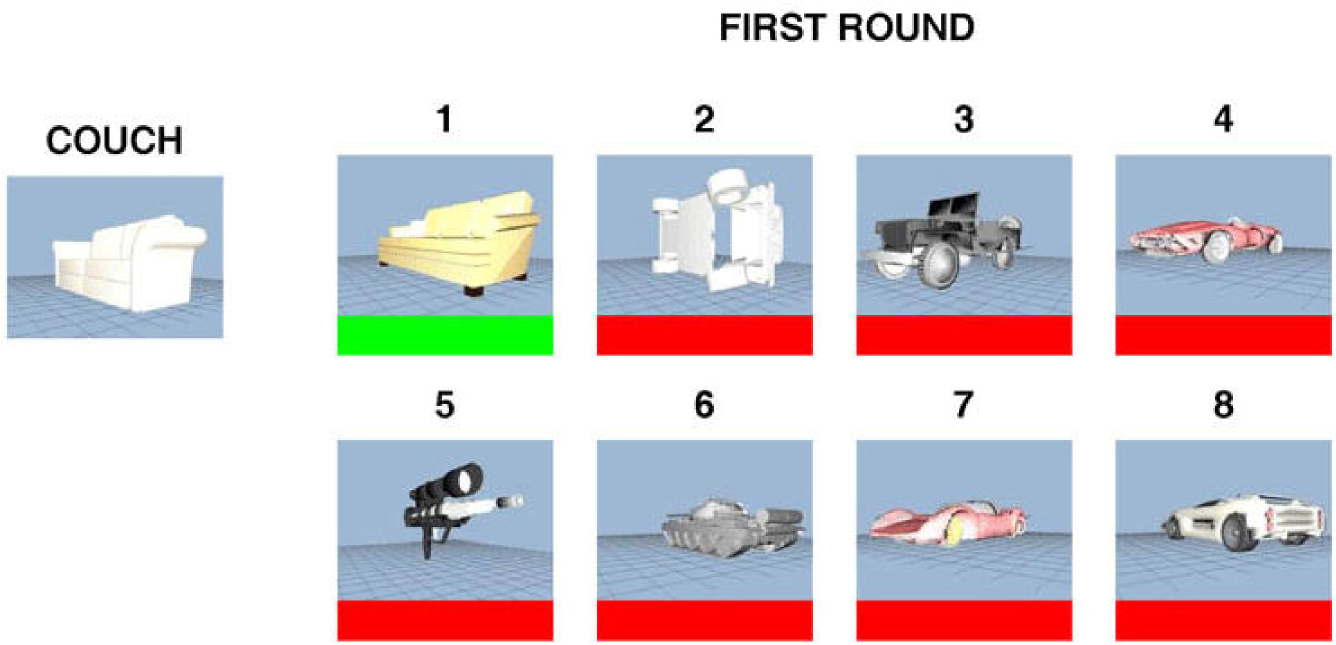
Two-Round On-line Example: “Human”



All the retrieved models are relevant after the second round

3DOR: Similarity Learning

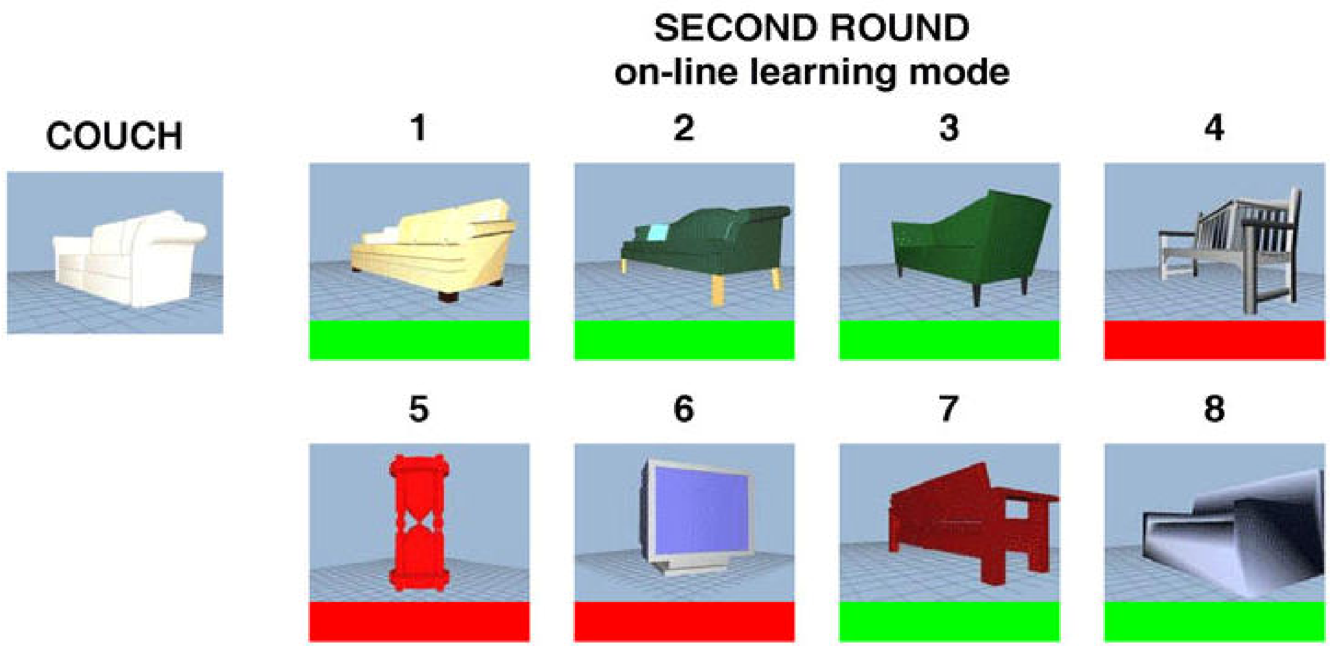
Two-Round On-line Example: “Couch”



User marks 1 relevant and 7 non-relevant models among the first 8

3DOR: Similarity Learning

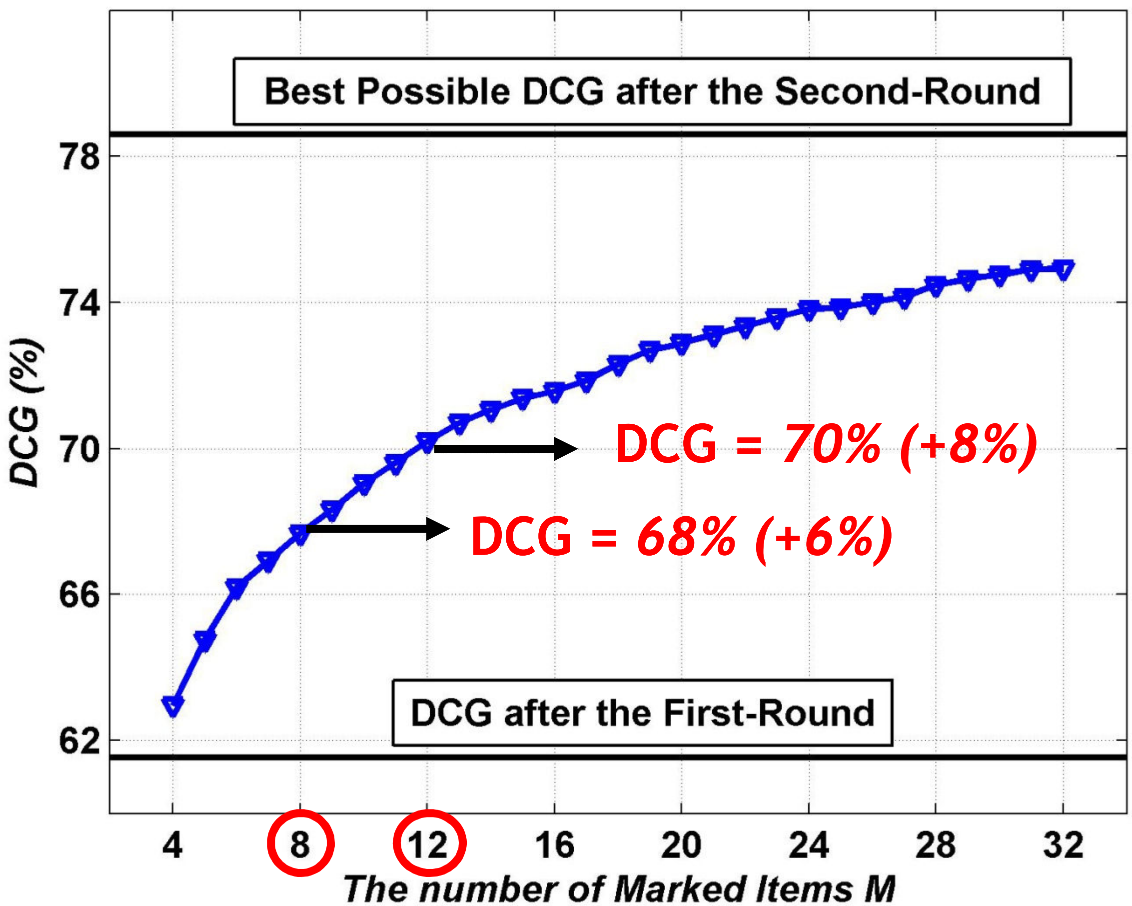
Two-Round On-line Example: "Couch"



5 models are relevant after the second round

3DOR: Similarity Learning

Two-Round On-line: Convergence



3DOR: Similarity Learning

Score Fusion: Performance Summary

Additive DCG Gain in all Protocols (%)

Two-Round On-line (M=8)	**Two-Round On-line** (M=12)	**Two-Round Offline**	**Bimodal**
6.0	8.0	5.0	2.0-4.0

Reminder

Performance of top descriptors differ only by 1-2 %.

Observation

Bimodal: Harder task than **Two-Round**

References

3D SHAPE DESCRIPTION

C. B. Akgül, B. Sankur, Y. Yemez, F. Schmitt.
3D Model Retrieval using Probability Density Based Shape Descriptors.
IEEE Trans on Pattern Analysis and Machine Intelligence, June 2009.

SIMILARITY LEARNING

C. B. Akgül, B. Sankur, Y. Yemez, F. Schmitt.
Similarity Learning for 3D Object Retrieval using Relevance Feedback and Risk Minimization.
Int. Journal of Computer Vision, Special Issue on 3D Object Retrieval, to appear, 2010.