ACM International Conference on Multimedia Information Retrieval, October 30-31, 2008, Vancouver, Canada
*Accepted as an oral paper.*

# Ranking on Semantic Manifold for Shape-Based 3D Model Retrieval

Ryutarou Ohbuchi
University of Yamanashi
4-3-11 Takeda, Kofu-shi
Yamanashi-ken, 400-8511, Japan
+81-55-220-8570
ohbuchiAT yamanashi.ac.jp

Toshiya Shimizu
University of Yamanashi
4-3-11 Takeda, Kofu-shi
Yamanashi-ken, 400-8511, Japan
+81-55-220-8570
shimizutsyAT gmail.com

## ABSTRACT
Semantics associated with 3D shapes are often as important as the shapes themselves in defining "shape similarity" among them. So far, only a small subset of 3D model retrieval methods took semantics into account. Most popular approach to semantic 3D model retrieval is based on Relevance Feedback (RF), an iterative, interactive approach for a system to learn a semantic class that embodies "user intention" for the query. A drawback of a typical RF-based method is its low initial performance as it starts cold without any semantic knowledge. An alternative approach is off-line learning of multiple semantic classes. The approach produces a good retrieval performance without per-query training iterations, but is unable to capture user intention per-query. The method proposed in this paper attempts to combine benefits of the two approaches so that both shared multiple semantic classes and per-query intention can be captured to improve 3D model retrieval. Our method first learns, off-line, the multiple semantic classes by using a semi-supervised manifold learning algorithm to produce a "semantic manifold" of the input features. The RF iteration based on manifold ranking algorithm is then run on the semantic manifold. Our empirical evaluation showed that this method significantly outperforms the manifold ranking run in the original, ambient feature space.

## Categories and Subject Descriptors
H.3.3 [Information Search and Retrieval]: Information filtering. I.3.5 [Computational Geometry and Object Modeling]: Surface based 3D shape models.

## General Terms
Algorithms, Experimentation.

## Keywords
3D polygonal models content-based retrieval, manifold ranking, manifold learning, semi-supervised learning.



## 1. INTRODUCTION
Three-dimensional (3D) models have become ubiquitous. It is used for 3D games running on mobile-phones and on game consoles, for such Web-based applications as Google Earth, for medical diagnostics, and for mechanical or architectural design. The need to organize these 3D models and related contents for effective creation, distribution, and consumption has prompted study into shape-based retrieval of 3D models. A shape-based 3D model retrieval system, by definition, compares shape of 3D models for their retrieval [18, 7]. However, retrieval performances of these systems have been insufficient due to so-called *semantic gap* between shape features and semantics associated with them. For example, 3D models that share a similar meaning may have different shape, or 3D models that have different meanings may share similar shape. Significant amount of work has been performed to narrow the semantic gap, mostly in the field of image and video retrieval, but also in the field of 3D model retrieval.

By far the most popular approach to bridge the semantic gap is *Relevance Feedback* (*RF*). In a RF based retrieval system, the system and the user work interactively and iteratively to capture a within-session semantics of the user. Given a query, the system first presents the user with an initial set of retrieval. The user then tells the system her/his preference on some or all of the models in the initial retrieval. The feedback choice may be *positive*,

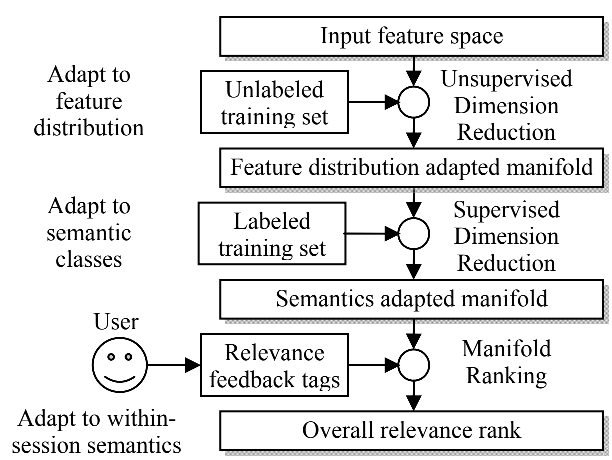



Figure 1: The proposed algorithm runs the manifold ranking-based relevance feedback on a manifold that reflects multiple semantic classes that are learned off-line.

*negative*, or both. The system may allow additional inputs, such as a degree of positive-ness. A RF system would capture the user feedback by using a learning algorithm and tries to improve the next round of retrieval by modifying features, distance metric, query set, etc. After several relevance feedback iterations, the user would be presented with a much improved set of retrieval results. The RF approach has been employed in retrieving 2D images and 2D movies (see, for example, [14, 15, 24, 20, 6]). It is also used in retrieving 3D shape models [3, 8, 9]. Elad et al [3] used Support Vector Machine (SVM), a learning classifier, for the relevance feedback learning. Novotni et al [9] compared several kernel based classifiers, including SVM, Biased Discriminant Analysis (BDA) and its kernelized version kernel-BDA, for their retrieval performance in a RF setting. The method by Leifman, et al [8] tries to make maximum use of small numbers of training samples by switching from the LDA for a small number of training samples (e.g., <20) to the BDA for larger numbers of training samples. Relevance feedback is quite powerful. For example, Novotni et al [9] reports the performance measured in weighted mean precision increased from 30% to over 70% after 10 RF iterations. Despite its success, RF framework has its weakness. A typical within-session RF framework is being able to capture only one semantic class per query session. Furthermore, every query session starts cold so that even a well-established long-term semantic class must be taught time and again for each query session.

Many learning algorithms are designed for single-class learning so that the performance drops sharply if it needs to handle multiple classes. For example, a bundle of SVMs, each one a one-vs.-the other classifier, has been used for multiple class recognition. However, the performance of the bundle is not very good. Also, one often wants to retrieve a 3D model that is not in the "correct" class but is similar in shape (or semantics) to the query. For such a purpose, a hard-decision learning classifier that produces yes/no result is not as desirable. A graded, smooth transition in similarity (or dissimilarity) values is needed.

In this paper, we propose a 3D model retrieval method that takes into account both a semantic class that reflects user intention for the session, and shared and well-established multiple semantic classes for an improved retrieval performance (Figure 1). To capture multiple semantic classes, the method utilizes the *Semi-Supervised Dimension Reduction* (*SSDR*) algorithm of Ohbuchi et al. [11] that combines two stages of locality preserving manifold learning algorithms to produce "*semantic*" sub-manifold that captured both a distribution of unlabeled features and multiple semantic classes. To adapt to per-query-session user intention, a relevance feedback framework based on the *Manifold Ranking* (*Mrnk*) algorithm [25, 6] is run on the semantic manifold produced by the SSDR. The manifold ranking is one of the most powerful RF algorithms; according to He, et al [6], Mrnk outperformed both *Support Vector Machine* (*SVM*) [24] and $SVM_{acitive}$ [20, 21] in relevance feedback image retrieval experiments. To our knowledge, the proposed method is the first successful attempt to perform relevance feedback on a manifold that reflects multiple semantic classes.

Our experimental evaluation of the proposed algorithm showed the effectiveness of the proposed algorithm in capturing both long-term, shared semantics and within-session semantics (user intention). We used the *Surflet-Pair Relation Histograms* (SPRH) [23] as the shape feature, which has the *Mean First Tier Highly Relevant* (*FT-HR*) score of 27% by itself if evaluated by using the SHREC 2006 benchmark [21]. The score increased to 58% with the multiresolution feature fusion and adaptation to multiple semantic classes by using the SSDR algorithm [11]. By learning both long-term and within-session semantics, our proposed method produced FT-HR=83% after only three RF iterations if top 20 models are allowed to be tagged for their relevance. If top 50 retrievals are allowed to be tagged for their relevance per RF iteration, FT-HR>90% is achieved after only three RF iterations. These figures are significantly better than the FT-HR=59.4% produced by using the same manifold ranking-based RF algorithm run in the original features space.

The contribution of this paper can be summarized as follows;

- Proposal of a novel 3D model retrieval method that combines short-term, per-query-session user intention and established multiple semantics classes for an improved retrieval performance.
- Experimental evaluation of the proposed method by using a standard 3D model retrieval benchmark. Evaluation showed that manifold ranking on the semantic manifold significantly outperforms the same in the original feature space.

The following of this paper is structured as follows. The next section will describe our proposed method, followed by Section 3 on experiments and their results. We will summarize the paper in Section 4.

# 2. METHOD

Proposed method is novel in that it combines *manifold learning*-based semi-supervised dimension reduction algorithm with a *manifold ranking*-based relevance feedback algorithm in order to capture both a set of long-term, shared, multiple semantic classes and a per-query-session (1-class) user intention. Multiple (e.g., 30 to 100) semantic classes are learned by using off-line semi-supervised dimension reduction algorithm by Ohbuchi, et al [11]. It maps the original, ambient feature space onto a lower dimensional "*semantic manifold*". Then, a short-lived, per-session semantic class is captured by using relevance feedback based on Manifold Ranking (Mrnk) algorithm [25, 6] run on the semantic manifold. The use of semantic-manifold instead of the original, ambient feature space for the manifold ranking improves the effectiveness of the manifold ranking algorithm significantly.

## 2.1 Multiresolution Feature Fusion

To capture multi-scale shape features, our method uses the multiresolution shape comparison approach of Ohbuchi et al, al [11] based on a mathematical morphology-like *Multi-Resolution* (MR) representation. Given a 3D model and a set of predetermined scale values, or *alpha*, the approach first creates a set of 3D MR shape models by using the *3D alpha shapes* algorithm [2]. Once the MR set of 3D models is obtained, a (single resolution) feature extraction algorithm is applied to the model at each resolution level.

The SSDR and the manifold ranking based relevance feedback may be applied to the MR set of features in several different ways. The proposed algorithm performs the SSDR and the manifold ranking based relevance feedback separately at each resolution level (Figure 2). The $L$ relevance ranks produced by the manifold ranking are combined into an overall relevance rank among the pair of models by using a uniform weight linear combination.

## 2.2 Semi-Supervised Dimension Reduction

The Semi-Supervised Dimension Reduction (SSDR) algorithm consists of two steps. The first step is adaptation to the feature distribution by using an *Unsupervised Dimension Reduction* (UDR) algorithm. We call the step *Feature set Adaptation* (*FA*), and the manifold produced by the FA is a *FA-manifold.* For the experiments that follow, we use the *Locally Linear Embedding* (*LLE*) [13] for the UDR. The LLE is a globally non-linear but locally linear dimension reduction algorithm that learns a mapping from the input space onto a subspace, or a manifold, unsupervised. As the LLE produces a mapping defined only at the training samples, the mapping is smoothly and continuously approximated by using the RBF-network [4]. A new feature in the original ambient space is projected onto the approximated manifold to transform it into a point on a FA-manifold.

The second step adapts the FA-manifold to the multiple, predefined semantic classes by using a *Supervised Dimension Reduction* (SDR) algorithm. We call the process *Semantics Adaptation* (*SA*), and the manifold produced by the SA is a *SA-manifold*. For the experiments that follow, we used *Supervised Locality Preserving Projections* (*SLPP*) algorithm by Xhaofei He, et al. [5] and the *Local Fisher Discriminant Analysis* (*LFDA*) by Sugiyama, et al. [17]. Both SLPP and LFDA are linear, and try to decrease distance among samples in the same class while trying to increase distances among different classes. The difference is that the LFDA allows multimodality in a semantic class. Using the SDR, features in the FA-manifold are mapped onto the SA-manifold, which is adapted to both feature distribution and multiple semantic classes. Note that the dimensionality of the semantic manifold is significantly smaller than the input feature space. For example, a 625-dimensional SPRH feature has been reduced down to a 50-dimensional vector on the SA-manifold.

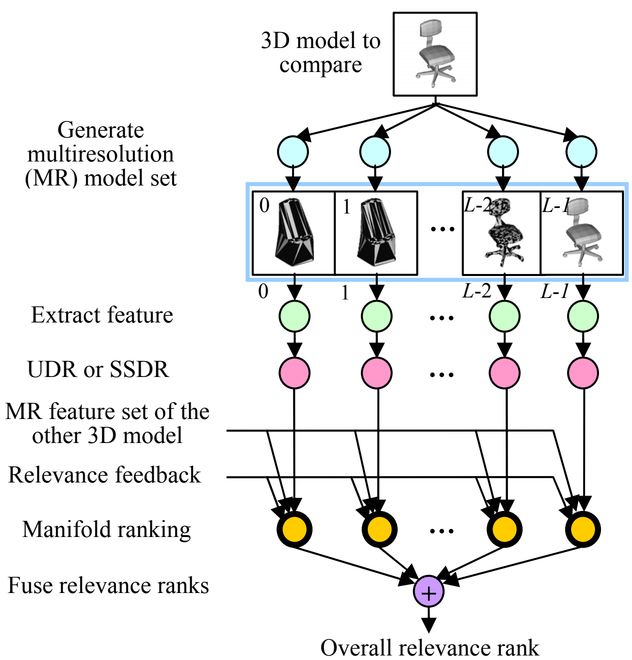



Figure 2: The semi-supervised dimension reduction and the manifold ranking steps are performed per resolution level, and relevance ranks from multiple resolution levels are fused into an overall relevance.

The SSDR algorithm can be summarized as follows, whose details can be found in the paper by Ohbuchi, et al [11].

**(1) Extract feature:** Extract a $k$-dimensional feature from each model in the unlabeled set of 3D models $M_{UL}$ of size $p$ for the UDR. For the experiments that follow, we used the Surflet Pair Relation Histograms (SPRH) by Wahl, et al. [23].

**(2) Unsupervised Learning of Feature Distribution:** Given the set of $p$ unlabeled features, an UDR algorithm learns the $l$-dimensional subspace $S_{UDR}$ spanned by the set. The resulting $S_{UDR}$ maps an input $k$-dimensional feature to the interim, $l$-dimensional FA-manifold in which $l < k$. We used the LLE for the UDR. For the LLE, we must approximate the $S_{UDR}$ so that it is defined everywhere in the input space. We used the RBF-network [4] for the approximation.

**(3) Off-Line Supervised Learning of Semantic Classes:** Extract $k$-dimensional features from the set of labeled 3D models $M_{SL}$ for the SDR. Note that the size $q$ of the $M_{SL}$ is typically much smaller than the size $p$ of the $M_{UL}$. A SDR algorithm then learns categories from the $q$ labeled features in a batch and encodes the knowledge into the $m$-dimensional subspace $S_{SDR}$ to be used for later SDRs. The manifold maps an interim $l$-dimensional feature onto the salient, $m$-dimensional feature used for retrieval. We used the SLPP [5] and the LFDA [17] for the SDR.

**(4) DB pre-processing:** For each $k$-dimensional input feature of all the models in the database $M_D$, employ the UDR and SDR in succession to produce $m$-dimensional salient feature that incorporates semantic concepts learned from the labeled models in the training set $M_{SL}$. Store the salient feature together with the corresponding 3D model for later retrieval. As the dimension $m$ of a salient feature is much smaller than the dimension $k$ of the corresponding input feature, cost of distance computation and feature storage are significantly reduced.

When a query session starts, the feature of the query model is processed by the UDR and then the SDR described above so that it also mapped onto the SA-manifold. The user presents the system with the query feature mapped onto the SA-manifold. The system performs the manifold ranking in the unsupervised mode to retrieve initial retrieval set. Then, the Relevance Feedback (RF) iteration starts. At the $i$th RF iteration, the user is presented with top $S$ retrieval results for feedback. We refer to the size $S$ as *Relevance Feedback scope*, or *RF-scope*. To provide feedback, our method simply tags the positive samples, a method that fits most naturally with the manifold ranking algorithm. The detailed steps for a retrieval session are as follows;

**(1) Query Feature Extraction and Processing:** Extract the input feature of the query model $m_q \in M_Q$ and perform the UDR and SDR in succession to produce a feature in the SA-manifold having dimension $m$.

**(2) Initial ranking and retrieval:** Apply the manifold ranking algorithm [25, 6] in the unsupervised mode to the features in the database and to the (initial) query to compute the initial ranking of the models in the database against the query. Retrieve and present top-ranked *S* models for the relevance feedback tagging.

**(3) Relevance feedback tagging:** Perform relevance feedback on the *S* models retrieved by tagging relevant models with the "positive", and the others with none.

**(4) Re-ranking and retrieval:** Expand the query set by considering the database models having "positive" tag as the part of the expanded query set. Perform manifold ranking using the expanded query set. Retrieve highest ranking models for relevance feedback tagging. If not satisfied with the retrieved set, go to step (3) above.

## 2.3 Manifold Ranking

Our method uses the *manifold ranking* [25] algorithm for relevance feedback learning. Manifold ranking is a graph-based learning algorithm that can be used in unsupervised, supervised, or semi-supervised mode. We used the manifold ranking algorithm by Zhou, et al. [25], among others. The manifold ranking algorithm tries to rank distances (or similarity) of points from a given sample on the manifold of all the samples, taking global distribution of the samples into account.

Let $\chi = \{x_1, \cdots, x_q, x_{q+1}, \cdots, x_n\}$ be a set of features in a *m*-dimensional feature space $\mathbb{R}^m$, in which first $q$ points from $x_1$ to $x_q$ are the queries and the rest are the points we want to rank according to their similarity to the queries. Let $d : \chi \times \chi \to \mathbb{R}$ denote a distance metric on $\chi$, e.g., L1 norm or Cosine distance, that assigns a pair of points $x_i$ and $x_i$ a distance $d(x_i, x_j)$. Let $f : \chi \to \mathbb{R}$ be a ranking function that assigns each $x_i$ a ranking score $f_i$, forming a rank vector $f = [f_1, \cdots, f_n]^T$. Let the *n*-dimensional vector $y = [y_1, \cdots, y_n]^T$ be a label vector.

The same manifold ranking algorithm may be used in two modes, unsupervised and supervised. In the unsupervised mode, the goal is to compute the similarity ranks of the 3D models in the database to the (single) query. So we set $y_i = 1$ for the query model only and $y_i = 0$ for the rest, i.e., all the models in the database. In the supervised 1-class learning mode under the relevance feedback framework, the goal is to compute the similarity ranks of the 3D models in the database to the *expanded query set* containing more than one points. The expanded query set consists of the (initial) query model and all the models in the database that are tagged as "positive" up to current relevance feedback iteration. Thus, we set $y_i = 1$ for the query model and the models that are tagged "relevant" during the RF iterations. The initial value $y_i = 0$ is given to the other models in the database without the "relevant" tag.

Intuitively, the process of manifold ranking resembles to the process of solving a diffusion equation on an irregular mesh. The mesh, or the manifold, is generated by connecting feature points in the high dimensional input features space based on their mutual proximity. The proximity during mesh generation is determined by using a distance measure, e.g., L1-norm, in the ambient feature space. The manifold ranking algorithm iteratively diffuses the initial value of $y_i = 1$ given to the (expanded) query set to its neighbors on the manifold. At the equibilium, the higher the diffused value, the higher the similarity rank of the points in the database to the (expanded) query set. As the diffusion occurs along the manifold, similarity rank thus computed are better than those computed directly in the input feature space.

Create the affinity matrix $\mathbf{W}$ where $\mathbf{W}_{ij}$ indicates the similarity between samples $x_i$ and $x_j$

$$\mathbf{W}_{ij} = \begin{cases} \exp\left(-\dfrac{d(x_i, x_j)}{\sigma}\right) & if \quad i \neq j \\ 0 & \text{otherwise} \end{cases}$$

The distance metric $d(x_i, x_i)$ used for the affinity matrix affects the ranking. We will compare *L1-norm*, *L2-norm*, *L0.5*-norm as well as *Cosine* distance for their ranking performance. The positive parameter $\sigma$ defines the radius of influence. Note that $\mathbf{W}_{ii} = 0$ since there is no ark connecting a point with itself. The matrix $\mathbf{W}$ is positive symmetric. We then form a normalized *graph Laplacian* $\mathbf{L}$,

$$\mathbf{L} = \mathbf{D}^{-\frac{1}{2}}(\mathbf{D} - \mathbf{W})\mathbf{D}^{-\frac{1}{2}}$$

Where $\mathbf{D}$ is a diagonal matrix in which $\mathbf{D}_{ij}$ equals to the sum of the *i*-th row of $\mathbf{W}$, that is, $\mathbf{D}_{ij} = \sum_j \mathbf{W}_{ij}$ Then the ranking vector $f = [f_1, \cdots, f_n]^T$ can be estimated by iterating the following until convergence;

$$f^{(t+1)} = \frac{1}{1+\mu}(\mathbf{I} - \mathbf{L}) f^{(t)} + \frac{1}{1+\mu}\mathbf{Y}$$

The parameter $\mu > 0$ is a regularization parameter, and affects retrieval performance and the convergence of the iteration above. Let $f^*$ be the limit of the above iteration. Rank each point $x_i$ as a label $y_i = \arg\max_{j \le c} f_{ij}^*$. In the case above, $f^*$ has a close form solution;

$$f^* = \left(\mathbf{I} + \frac{1}{\mu}\mathbf{L}\right)^{-1}\mathbf{Y}$$

We tried both iterative and the closed-form solutions on *MatLab*, and mostly used the latter, closed-form solution, for it was faster or our set of parameters when implemented in the *MatLab*.

Most of the compute-intensive work for the SSDR is in the pre-processing stage. The MR model set generation, feature extraction, vector projection for dimension reduction, and manifold ranking is needed per query. The manifold ranking in particular must run once for every RF iteration, whose cost is dominated by the cost of computing $f^*$. The cost increases with the size of matrix $\mathbf{L}$, that is, the number of features *n*. On a contemporary PC, even for a dense matrix, it takes a few seconds to obtain $f^*$ for the SHREC 2006 database having 1,814 models. Further reduction in computational cost is possible if the affinity matrix $\mathbf{W}_{ij}$ is made sparse by limiting the entries to the feature pairs having a distance less than a certain threshold.

## 2.4 Distance Measures

We compared several distance measures, including $L^k$-norm for *k*=0.5, 1.0, and 2.0, as well as the *cosine* measure for distance computation for retrieval, and for forming the affinity matrix **W** to be used for manifold ranking. Let $\mathbf{x}=(x_i)$, $\mathbf{y}=(y_i)$ are the feature vectors and $n$ is the dimension of the vectors. Depending on the value of $k$, $d_k(\mathbf{x},\mathbf{y})$ is the Manhattan distance if $k$=1.0, or the Euclidian distance if $k$=2.0.

$$d_k(\mathbf{x},\mathbf{y}) = \left[\sum_i^n \left(\|x_i - y_i\|^k\right)\right]^{1/k}$$

According to Aggarwal, et al [1], $k$<1.0, e.g., $k$=0.5 is expected to perform better for higher dimensional features. As the cosine measure is a measure of similarity in the range $[0,1]$, we converted it to a distance using the following equation;

$$d_{\cos}(\mathbf{x},\mathbf{y}) = 1 - \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|}$$

Another distance measure we employed is the *Kullback-Leibler divergence* (*KLD*). The KLD is sometimes referred to as *information divergence*, or *relative entropy*, and is not a distance *metric*, for it is not symmetric;

$$d_{KLD}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n}(y_i - x_i)\ln\frac{y_i}{x_i}$$

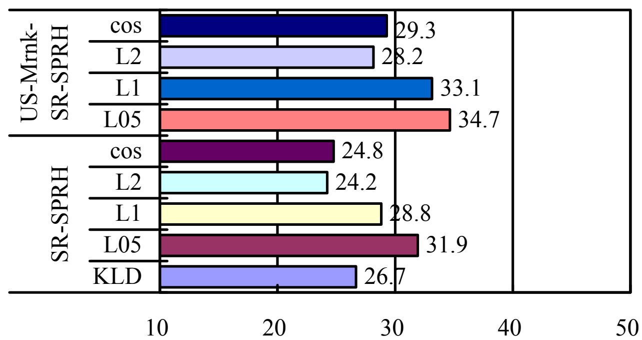
# 3. EXPERIMENTS AND RESULTS

Evaluation of relevance-feedback based retrieval algorithms for 3D models is not standardized. While the SHREC 2007 [22] list of track proposal included the "Relevance feedback" track, the track did not materialize as it saw no entrant. We thus experimentally evaluated the proposed algorithm by simulating relevance feedback on SHREC 2006 [21] database and query set, as it is used widely by many researchers. The database in the SHREC 2006 is the union of the test and train set of the Princeton Shape Benchmark [16], which is a 1,814 model collection of polygon soup models. The SHREC 2006 uses 30 query models that are not in the database. There are many different performance measures used in the SHREC 2006. Among them, we used the *Mean First Tier Highly Relevant* (FT-HR). First Tier is a ratio, in percentile, of the models retrieved from the desired class (i.e., the same class as the query) in the top $k$ retrievals, in which $k$ is the size of the class. The Mean First Tier is a mean over the 30 query models of the SHREC 2006. The designation "highly relevant" comes from the fact that it uses the "correct" ground truth class, instead of the "approximately correct" ground truth class.

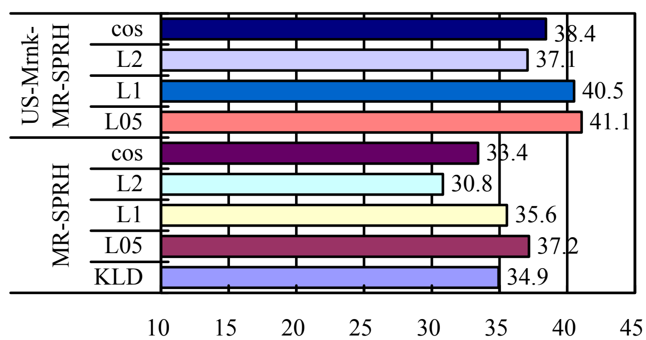
Performing relevance feedback experiment manually is not very viable. We simulated human feedback using the SHREC 2006 ground truth class. At each one of the RF iterations, top $S$ retrieval (i.e., relevance feedback scope) is scanned, and if a model in the ground truth class for the query model is found in the top $S$ models, it is tagged as "relevant". We assumed the user is error free so that relevance feedback is always correct. (Some of the studies on relevance feedback algorithms inject errors into the feedback tags to evaluate robustness.)

## 3.1 Distance Measures and Retrieval Performance

In the first set of experiments, we compared four fixed distance measures, L0.5-norm, L1-norm, L2-norm, and KLD, with the Mrnk in unsupervised mode. Among the fixed distance metrics, for both single resolution (Figure 3a) and multi-resolution fusion (Figure 3b) cases, the L0.5-norm performs the best, as suggested in [1]. It is followed by the L1-norm and then the KLD. In the case of the single-resolution (SR-) SPRH feature, its performance improved from FT-HR=26.7% to FT-HR=31.9% by simply switching from the KLD to L0.5-norm. For the unsupervised Mrnk that adapts to the distribution of features, the L0.5-norm performed the best again, followed by the L1-norm. The performance of the original SR-SPRH using the KLD is FT-HR=26.7%, while the unsupervised manifold ranking that employed the L0.5-norm produced FR-HR=34.7%, an improvement of 8.0%.




(a) Single-resolution SPRH feature.




(b) Fusing multi-resolution SPRH features

Figure 3: Comparison of distance measures using the single-resolution and multi-resolution SPRH features.

## 3.2 Ranking on Semantic Manifold

In the second set of experiments, we compared the performance of the manifold ranking algorithm applied to three different manifolds; (1) the original feature, (2) the *FA-manifold* adapted to feature set distribution produced by the UDR, (3) the *SA-manifold* adapted to multiple semantic classes produced by the SSDR, that is, the UDR followed by the SDR. In this set of experiments, we used the only the SPRH feature in the multiresolution feature fusion framework.

For the SSDR we used the best performing parameters and dimension reduction algorithms according to Ohbuchi, et al [11]. We used the MR-SPRH feature, and applied the LLE [13] for the UDR, followed by the LFDA [17] for the SDR. Compared to the original dimension of 625 for the SPRH feature, the feature on the FA-manifold is 300, and the feature on the SA-manifold is 50.

The UDR algorithm is trained by using 10,000 unlabeled models quasi-randomly sampled from the NTU database [10]. The size of the neighbourhood for the LLE is chosen to be 2.0% of the size of the database, that is, 200, and the spread of the RBF network kernel is 0.9. For the manifold ranking, we used $\mu = 100$ and $\sigma = 0.001$. To train the SDR, we used the SHREC 2006 database, with its 1,814 models and 30 classes.

Figure 4 shows the summary comparison of performances of various methods. In the figure, manifold ranking is applied to the following three feature spaces;

- **SA**: Manifold adapted to feature distribution and multiple semantic classes by using the Semi-Supervised Dimension Reduction [11].
- **FA**: Manifold adapted to feature distribution by using the Unsupervised Dimension Reduction [11].
- **Orig.**: Original feature space.

On these three spaces, two modes of manifold ranking algorithms, the US-Mrnk for the unsupervised mode and RF-Mrnk for the relevance feedback mode, are applied. These manifold ranking algorithms used L0.5-norm to form the affinity matrix **W**. These two modes of manifold ranking algorithms are compared against No-Mrnk that employed *Cosine* distance directly on these manifolds. For the RF-Mrnk, relevance feedback scope size S of 20 and 50 are chosen; that is, top *S* retrievals were subjected to RF tagging. The performance indices in the Figure 4 are obtained after 4 RF iterations (plus the initial retrieval without feedback.)

As the baseline, the original SPRH feature (without multiresolution feature fusion) using KLD for distance computation had the FT-HR=26.7%. By using the multiresolution (MR) feature extraction with 6 levels, Cosine distance, and linear combination of the MR distances, the performance went up to FT-HR=37.2% (The bar at the bottom of the Figure 5.). The combination of the MR and the Feature set Adaptation (FA) produced FT-HR=45.4%, and the multiple semantic class learning pushed it further to FT-HT=58.1%.

- **Unsupervised Manifold Ranking in the original feature space:** Unsupervised Manifold Ranking (*US-Mrnk*) run in the ambient space (original feature space) produced the FT-HR=41.0%, which is significantly better than the FT-HR=37.2% of the original MR-SPRH feature.

- **Unsupervised Manifold Ranking in the dimension-reduced feature spaces:** US-Mrnk, however, did not improve retrieval performance for the feature sets that are already dimension reduced by using manifold learning algorithms described in [11]. When the US-Mrnk is applied to the FA-feature or SA-feature, their retrieval performances did not change significantly. This outcome is to some extent expected since the manifold ranking algorithm by itself tries to find its own manifold by connecting the features in ambient space before running a diffusion process on the manifold. The methods to find manifold used by the manifold learning and manifold ranking algorithms are quite similar.

- **Supervised Manifold Ranking using Relevance Feedback:** On-line supervised manifold ranking using the Relevance Feedback framework (*RF-Mrnk*) did significantly improved the scores for the original feature. RF-Mrnk also improved performances if applied to the dimension reduced features, that are, features in the FA-manifold and the SA-manifold (Figure 5).

  The RF-Mrnk run on the SA-manifold performed the best, followed by the RF-Mrnk run on the FA-manifold. With only 4 RF iterations using the RF scope size *S*=50, the RF-Mrnk run on the SA-manifold reached FT-HR=93.3%, while the RF-Mrnk run on the FA-manifold reached FT-HR=76.9%. after 4 RF iterations. Compared to these two, the RF-Mrnk run in the original feature space produced significantly lower score of FT-HR=64.3%.

From these results, it can be concluded that the RF-Mrnk benefits from both unsupervised and semi-supervised dimension reduction applied to the original features. The manifold ranking algorithm by itself tries to find a manifold spanned by the features in the ambient space before applying relevance score diffusion. However, RF-Mrnk appears to perform better if it is applied on the lower dimensional FA-manifold adapted to feature distribution and SA-manifold adapted to both feature distribution and multiple semantic classes. The best retrieval performance is produced by the RF-Mrnk run on the SA-manifold.

It is also interesting to note that the semi-supervised dimension reduction algorithm [11] that learns multiple classes at once (FT-HR=58.1%) performed nearly as well as the single class learning using RF with S=20 (FT-HR=59.4%).

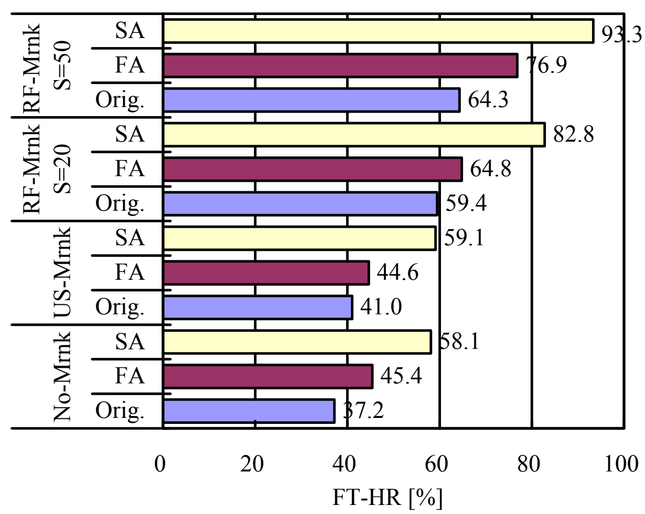
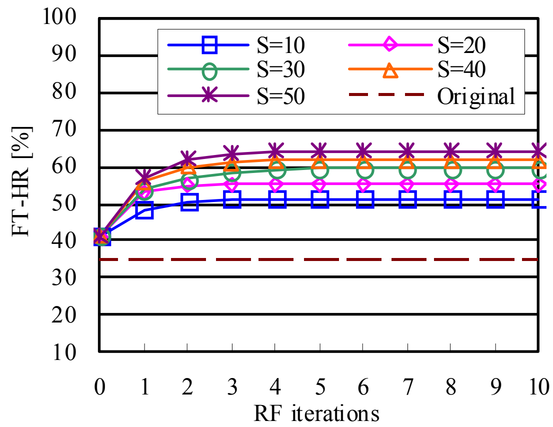


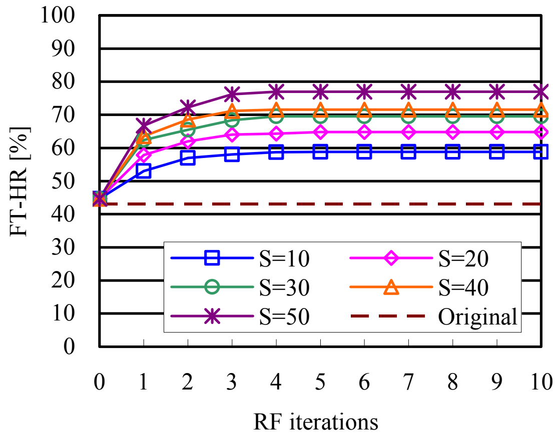

Figure 4: Retrieval performances of various methods based on multi-resolution SPRH feature.

Figure 5a to Figure 5c show, for the three feature spaces on which the RF-Mrnk took place, the relationship between the number of RF iterations and the retrieval performance. The three feature spaces are the Orig. (Figure 5a), FA (Figure 5b), and SA (Figure 5c) described above. The FT-HR values in Figure 4 corresponds to those in Figure 5 at RF iterations = 4.

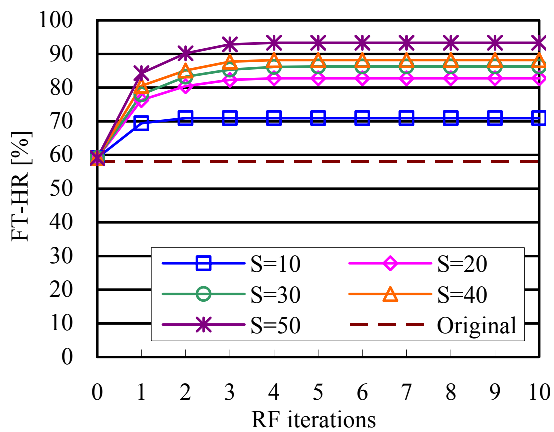
For all of the three feature spaces, the performance reached its saturation point after about 4 RF iterations excluding the initial retrieval ($0^{th}$ iteration). Also, as expected, for all of the three feature spaces, retrieval performance at saturation point improved as the RF scope size *S* is increased.




(a) Original feature space + RF-Mrank.




(b) FA-manifold + RF-Mrank.




(c) SA-manifold + RF-Mrank.

Figure 5: RF based on manifold ranking on three feature spaces, the original feature space (a), the FA-manifold (b), and the SA-manifold (c).

Figure 6 shows a set of retrieval examples for the SHREC 2006 query 20 (the "hand", top left). Relevance feedback scope size S=50 and 4 RF iterations (+ 1 initial retrieval) are used for all the RF cases. In the example, manifold ranking-based RF works significantly better if it is run on the feature-adapted manifold (RF-Mrnk-FA) or on the semantic adapted manifold (RF-Mrnk-SA) than in the ambient feature space (RF-Mrnk-Orig.).

## 4. CONCLUSION AND FUTURE WORK

In this paper, we proposed a novel 3D model retrieval algorithm that takes into account both long-lived semantics learned from multiple semantic classes and a short-lived, within-session user intention learned by using relevance feedback. The former is accounted for by learning multiple (e.g., 30 to 100) semantic classes at once, off-line, by using a Semi-Supervised Dimension Reduction (SSDR) algorithm [11]. The latter is accounted for by using a relevance-feedback framework based on the manifold ranking algorithm [25]. The method performs manifold ranking-based relevance feedback on the semantic-manifold that embodies both feature distribution and multiple semantic classes.

Experimental evaluation showed that proposed approach is quite effective. The use of semantic-manifold instead of the original, ambient feature space improved the effectiveness of the manifold ranking algorithm significantly. The manifold ranking-based RF run on the original features achieved the First Tier Highly Relevant (FT-HR) score of FT-HR=64%, if the feedback scope *S*=50 is used. The same run on the semantic-adapted manifold produced the FT-HR=93% after only 3 RF iterations with feedback scope *S*=50.

## 5. ACKNOWLEDGMENTS

The authors would like to thank Takahiko Furuya, Masaki Tezuka, and Shun Kawamura for performing some of the experiments. This research has been funded in part by the Ministry of Education, Culture, Sports, Sciences, and Technology of Japan (No. 20500089).

## 6. REFERENCES

[1] C.C. Aggarwal, A. Hinneburg, D.A. Keim, On the surprising behavior of distance metrics in high dimensional space, *Proc. International Conf. On Database Theory* (2001).

[2] H. Edelsbrunner, E.P. Mücke, Three-dimensional Alpha Shapes, *ACM TOG*, **13**(1), pp. 43-72 (1994).

[3] M. Elad, A. Tal, S. Ar., Content based retrieval of vrml objects – an iterative and interactive approach, *Proc. EG Multimedia* 39, pp. 97-108 (2001).

[4] S. Haykin, *Neural network a comprehensive foundation*, Second Edition, Prentice Hall, 842 pages (1999).

[5] X. He, P. Niyogi, Locality Preserving Projections, *Advances in Neural Information Processing Systems*, **16**, Vancouver, Canada (2003). http://people.cs.uchicago.edu/~xiaofei/LPP.html

[6] J. He, M. Li, H-J Zhang, H. Tong, C. Zhang, Manifold-Ranking Based Image Retrieval, *Proc. ACM Multimedia* 2004, pp. 9-16 (2004).

[7] M. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, K. Ramani, Three Dimensional Shape Searching: State-of-the-art Review and Future Trends, *Computer Aided Design*, **5**(15), pp. 509-530, (2005).

[8] G. Leifman, R. Meir, A. Tal, Semantic-oriented 3d shape retrieval using relevance feedback, *The Visual Computer*, **21**(8-10), pp. 865-875, (2005).

[9] M. Novotni, G.-J. Park, R. Wessel, R. Klein, Evaluation of Kernel Based Methods for Relevance Feedback in 3D Shape Retrieval, *Proc. The Fourth International Workshop on Content-Based Multimedia Indexing* (*CBMI'05*) (2005).

[10] NTU 3D Model Database, ver.1, http://3d.csie.ntu.edu.tw/

[11] R. Ohbuchi, A. Yamamoto, J. Kobayashi, Learning semantic categories for 3D Model Retrieval, *Proc. ACM MIR 2007*, pp. 31-40 (2007).

[12] J. Pu, K. Lou, K. Ramani, A 2D Sketch-Based User Interface for 3D CAD Model Retrieval, *Computer Aided Design and Application*, **2**(6), pp.717-727 (2005).

[13] S.T. Roweis, L.K. Saul, Nonlinear Dimensionality Reduction by Locally Linear Embedding, *Science*, **290**(5500), pp. 2323-2326 (2000).

[14] J.J. Roccio, Relevance feedback in information retrieval, *The SMART Retrieval System*, pp. 313-323, Prentice-Hall (1971).

[15] Y. Rui, T.S. Huang, M. Ortega, S. Mehrotra, Relevance feedback: a power tool for interactive content-based image retrieval, *IEEE Trans. Circuits and Systems for Video Technology*, p. 59-66 (1998).

[16] P. Shilane, P. Min, M. Kazhdan, T. Funkhouser, The Princeton Shape Benchmark, *Proc. SMI '04*, pp. 167-178 (2004). http://shape.cs.princeton.edu/search.html

[17] M. Sugiyama, Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis, *Journal of Machine Learning Research*, **8**, pp.1027-1061 (2007).

[18] J. Tangelder, R. C. Veltkamp, A Survey of Content Based 3D Shape Retrieval Methods, *Proc. SMI '04*, pp. 145-156 (2004).

[19] J. B. Tanenbaum, V. de Silva, J.C. Langford, A Global Geometric Framework for Nonlinear Dimensionality Reduction, *Science*, **290**(5500), pp. 2319-2323 (2000).

[20] S. Tong, E. Chang, Support vector machine active learning for image retrieval, *Proc. ACM Multimedia* 2001, pp. 107-118 (2001).

[21] R. C. Veltkamp, R. Ruijsenaars, M. Spagnuolo, Roelof van Zwol, SHREC2006 3D Shape Retrieval Contest, *Dept of Info and Comp. Sci., Utrecht University*, *Technical Report* UU-CS-2006-030 (ISSN: 0924-3275).(2006).

[22] R.C. Veltkamp, F.B. ter Harr, SHREC 2007 3D Shape Retrieval Contest, *Dept of Info and Comp. Sci., Utrecht University, Technical Report* UU-CS-2007-015 (2007).

[23] E. Wahl, U. Hillenbrand, G. Hirzinger, Surflet-Pair-Relation Histograms: A Statistical 3D-Shape Representation for Rapid Classification, *Proc. 3DIM 2003*, pp. 474-481 (2003).

[24] L. Zhang, F. Lin, B. Zhang, Support vector machine learning for image retrieval, *Proc. IEEE ICIP 2001*, vol. 2, pp. 721-724 (2001).

[25] D. Zhou, O. Bousquet, T.N. Lal, J. Weston, B. Schölkopf, Learning with Local and Global Consistency, *Proc. NIPS 2003* (2003).

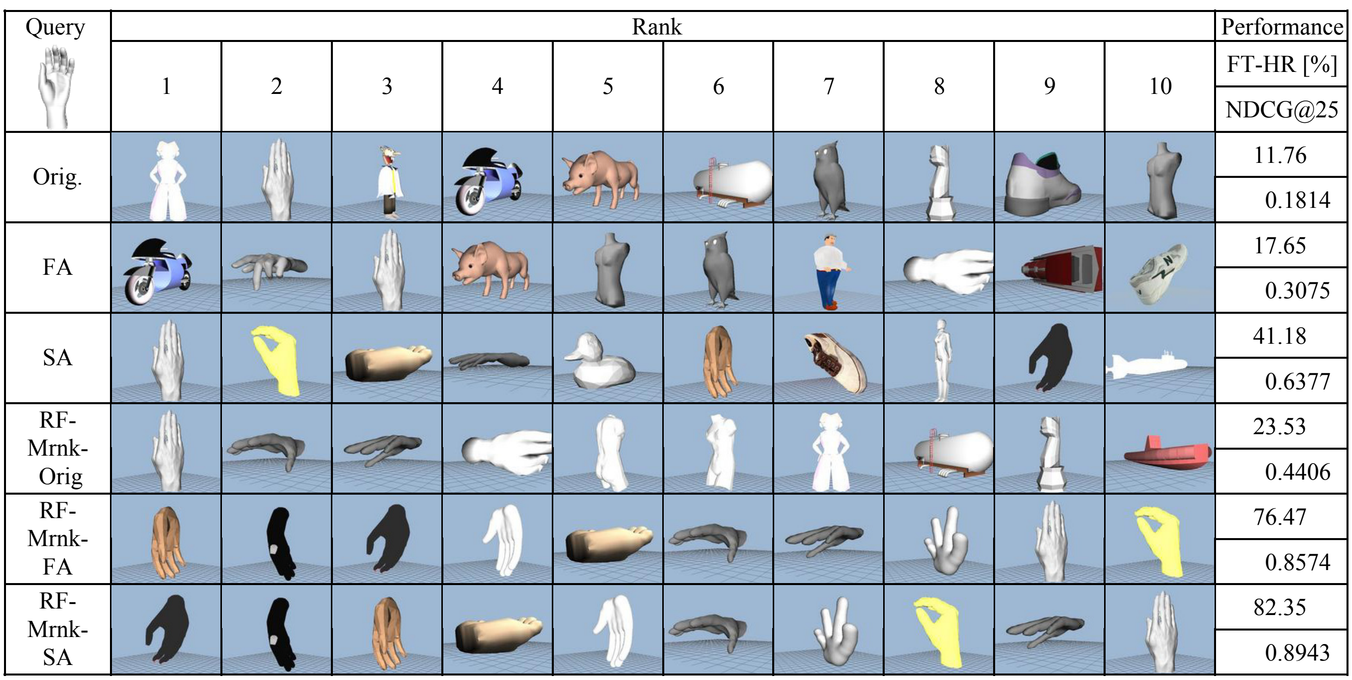



Figure 6: Retrieving the query20 (the "hand", top left) from the SHREC 2006. Manifold-ranking based Relevance Feedback (RF-Mrnk) run on the Semantic Adapted manifold (RF-Mrnk-SA) performed the best, followed by the RF-Mrnk run on the Feature Adapted manifold (RF-Mrnk-FA). (RF scope size *S*=50. RF iterations=4.)