Automatic shape expansion with verification to improve 3D retrieval, classification and matching

J. Knopp[1] and M. Prasad[2] and L. V. Gool[1,2]

[1]KU Leuven, Belgium
[2] ETH Zurich, Switzerland

Abstract
The goal of this paper is to retrieve 3D object models from a database, that are similar to a single 3D object model, given as a query. The system has no prior models of any object class and is class-generic. The approach is fully automated and unsupervised. The main contribution of the paper is to improve the quality of such 3D shape retrieval, through query verification and query expansion. These are part of a cascaded, two-stage system:
(i) Verification: after a first inexpensive and coarse retrieval step that uses a standard Bag-of-Words (BoW) over quantized local features, a fast but effective spatial layout verification of those words is used to prune the initial search results.
(ii) Expansion: a new BoW query is issued on the basis of an expanded set of query shapes that, next to the original query, also includes the positively verified results of (i).
We perform comprehensive evaluation and show improved performance. As an additional novelty, we show the usefulness of the query expansion on shape classification with limited training data and shape matching, domains in which it has not been used before. The experiments were performed on a variety of state-of-the-art datasets.

1. Introduction

The amount of available 3D data is rapidly growing thanks to portals such as Google Warehouse [Goo], as well as capturing systems such as Arc3D [VG06] or Photosynth [SSS06]. As a result, the need for effective 3D shape retrieval ([FKS*04, TCF09, OLGM11, KPW*10]) is on the rise too. Retrieval is the problem of finding similar shapes - typically of the same object class - given an example 3D shape, i.e. the query. The goal of this paper is to improve such shape retrieval. Our system has no prior models of any object classes and is class-generic. The approach is fully automated and unsupervised.

The field of 3D retrieval has produced a number of benchmark databases that we will also use in this paper, e.g. SHREC [DGA*09], Princeton [SMKF04], and TOSCA [BBGO11]. On these benchmarks we demonstrate substantial improvements.

We start from a traditional Bag-of-Words (BoW) approach [TCF09,OLGM11,SZ03]. BoW finds distinctive features on shapes, and matches them against a vocabulary of 'visual words' (quantized local 3D feature descriptors). The shape is then represented as a histogram of visual word oc-

currences, the BoW vector. Shape similarity is measured in terms of BoW vector distances and a BoW-based query returns shapes from a database in increasing order of BoW distances to the query (exactly or approximately).

Our approach extends this BoW baseline with two steps, that each boost performance. Firstly, we verify the spatial configuration of the visual words for the shapes ranking high in the BoW search. Secondly, the shapes supported by this verification are used to expand the original query. This entire 3D pipeline, except for the BoW initialization, is novel for 3D retrieval. The whole process is fully-automated, unsupervised, and class-generic. We explore these principles for their use with shape class detection and matching as well, and show these applications to also benefit. Using these methods for such tasks is novel too.

The study of shape representation, categorization, recognition and retrieval are intertwined. Ideally the shape representation (features, BoWs, and distance measures) would be learnt to maximize intra-class similarity and inter-class dissimilarity [ETA02,PTTS08], assuming the user is interested in finding other samples of the same class (where the notion of "class" can have a variety of semantic meanings). How-

ever, (i) features are generally built to optimize a generic notion of saliency and repetitiveness [VS01, SOG09], (ii) retrieval is usually performed in a generic rather than class-specific setting and on large datasets [HJ11], (iii) the number of 'classes' and their descriptions are usually unknown, (iv) manual input is expensive and often unfeasible, and (v) computation comes at a premium (even more for 3D models than for text and images). So we want the search to be fast without loss of accuracy. Therefore, in this work we concentrate on improving retrieval performance in an automatic, unsupervised manner, applicable generically across classes.

2. Related work

Our work draws on a variety of research directions mainly found in the area of image-based search so far.

Local features. Except for a few cases like the seminal spin images [JH99], 3D shape search has for a long time been based on the use of global features, i.e. features derived from the entire shape. It is only recently that 2D concepts like local features have made their entrance as a mainstream approach in 3D search [BBGO11, TCF09]. With them came 2D approaches to retrieval, like Bag-of-Words (BoW). Generic off-the-shelf 3D features [KPW*10, SOG09] in BoW based methods [BBGO11, TCF09] have been shown to be robust to noise, deformation, orientation *etc*. We also follow a local feature-based approach. In our case, it is a 3D generalization of on-line available 3D SURF features [KPW*10]. Please note that the particular choice of feature is not crucial to the proposed method though. We start off with BoW as well, but subsequently improve the shape retrieval with important verification and query expansion steps. As a matter of fact, these steps implement a fully automated relevance feedback loop. We also show, that the local verification is important when i.e. non-complete shapes are compared.

Interestingly, though BoW approaches have been used in shape search, relevance feedback has mostly relied on vanilla features [ETA02, AKW07, PTTS08], or their weighted combination [GFSF10].

Cascaded approaches: Computational time and accuracy often have to be played off against each other, when examining the semantic relevance of search. Regardless of the chosen representation, most methods end up improving search relevance in a somewhat cascaded or iterative manner. For example, two (or more) representations (feature or distance wise) for an object may be constructed: one in which search can be performed fast and another in which accuracy can be improved. Improvement of accuracy can be performed in a variety of ways: (i) incorporating user feedback (most popular in 3D [MP96, ETA02, PTTS08, HJ11]), (ii) structural consistency (as in 2D [PSZ08, MPCM10, CMPM11]) and (iii) a variety of heuristics *e.g.* pseudo-relevance feedback, multiple queries *etc*. (see [LJI*03, BC02, PTTS08]). The query

expansion was found as an optimal retrieval method when searching is defined as a classification task [Efr08].

The quest for this, leads us to propose a 2-pass approach for shape verification and query expansion that enrich the query shape. This is implemented as cascade sub-parts, we introduce new proposals and evaluate each.

Structural verification: BoW representations succeed in representing a shape in terms of feature occurrences and being orientation and noise invariant but fail to capture more layout information. An improvement over vanilla search, was the analysis of word co-occurrences [PCI*07, JDS10] and more recently, structural constraints [CMPM11, AZ10]. Actual shape has more complex representation: *e.g.* statistical models, articulated models (among other graphs), templates *etc*. Fitting such models to a given shape can involve working our parametrization and correspondences. An exhaustive and comprehensive matching has been studied [OMMG10, KLF11, KLM*12] too, but leads to be computational expensive for a use when one needs to compute matches between ten shapes in a few seconds. Hence, we extend schemes that leverage feature co-occurrence to take into account their mutual spatial layout. Though a very loose approximation of more complicated shape models, this scheme of weak structural verification proves very effective in improving search performance.

Paper overview: As mentioned in the text above, the approach consists of three parts. In § 3 we overview the basic Vanilla BoW search that forms the core of retrieval algorithms. Scheme for the verification pass to filter relevant results from the initial BoW results is introduced in § 4. § 5 deals with the expansion phase *i.e.* incorporating the verification for a new, augmented, reliable result for the shape query. As we promised, we also investigated more 3D applications for expansion, they are introduced in § 6. We show the performance improvement on shape search against the state of the art methods (§ 7.1) on the task of shape search in § 7.2. We show how our methods can be used to improve weakly supervised classification in § 7.3. The benefits of expansion are explored for shape matching in § 7.4. Finally, we summarize our observations in § 8.

3. Vanilla search

Given a query shape, the goal is to find the most relevant candidates to the query shape, efficiently. The BoW [SZ03] model is the baseline of most retrieval engines owing to its ability to handle noise and compress the object representation for effective querying of large databases. In order to represent a 3D shape, features are computed according to generic criteria of saliency, repeatability, robustness and invariance [JH99, SOG09, KPW*10]. Clustering is performed to condense the large set of generic features into meaningful "visual words" (clusters). Each feature is assigned to a visual word and the object is then represented as a histogram
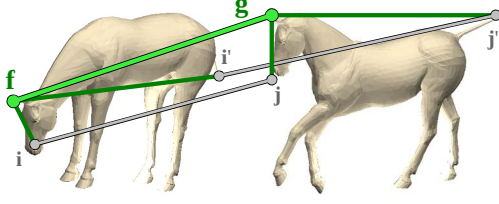
Figure 1: *Illustration of the verification method. Given two shapes and a pair of features* $\{\mathbf{f}, \mathbf{g}\}$*, the distance between them* $\mathrm{dist}(\mathbf{f}, \mathbf{g})$ *signifies if they represent a good correspondence. Instead of computing* $\mathrm{dist}(\mathbf{f}, \mathbf{g})$ *from descriptors, we measure the consistency of correspondences in a neighbourhood. This assumes that if* $\{\mathbf{f}, \mathbf{g}\}$ *is a correct match,* $\{\mathbf{f}, \mathbf{g}\}$ *will be surrounded by correct matches. For example, the candidates* $\{\mathbf{i}, \mathbf{j}\}$ *are in correspondence. Correspondences that are far (i.e.* $\{\mathbf{i}', \mathbf{j}'\}$*) are down-weighted to not be taken into account. This is an illustration, each* $\{\mathbf{f}, \mathbf{g}\}$ *is compared to the hundreds of correspondences.*

of visual word occurrences[†]. The similarity of objects is then measured as a distance between such BoW histograms and shapes are sorted according to that similarity which forms result list.

4. Verification

While visual word co-occurrence based tests are the simplest way of improving on BoW, structural verification *i.e.* including layout in additional to occurrences, is more desirable. Comprehensive object structure recognition is computationally expensive even for the state of the art [FGMR09, OHG11,OMMG10,KLF11] and suited for very small vocabularies [BBGO11], so we devise fast but effective, approximate verification schemes.

Let the descriptor sets of the query shape q and a result shape r (from the result list of § 3), be $\mathcal{F}$ and $\mathcal{G}$ respectively. The goal is to verify whether r is relevant result for q *i.e.* to measure their similarity. We measured the similarity between two shapes by a symmetric Modified Hausdorff Distance firstly introduced by Dubuinsson and Jain [DJ94],

$$d_{MHD}(q,r) = \frac{1}{F}\sum_{\mathbf{f}\in\mathcal{F}}\min_{\mathbf{g}\in\mathcal{G}}\big(\mathrm{dist}(\mathbf{f},\mathbf{g})\big) + \frac{1}{G}\sum_{\mathbf{g}\in\mathcal{G}}\min_{\mathbf{f}\in\mathcal{F}}\big(\mathrm{dist}(\mathbf{g},\mathbf{f})\big). \quad (1)$$

This sums the distance of every feature from the set $\mathcal{F}$ (with F feature tuples) to the most similar feature from the second set $\mathcal{G}$ and vice versa [‡]. Distance $\mathrm{dist}(\mathbf{f}, \mathbf{g})$ represents our belief that feature $\mathbf{f}$ is a correct match for $\mathbf{g}$. This can simply be

Euclidean distance but we devise it as a combination of similarity in the descriptor space with spatial configuration of features. In the following text, we show how $\mathrm{dist}(\cdot,\cdot)$ is computed for structural assumptions to be also robust to topological changes, occlusion, incomplete shapes and to obtain better performance.

4.1. N part weak shape model (WN):

Firstly, a set $\mathcal{M}$ of pairs of corresponding features between q and r based on the similarity of their descriptors is constructed. For every feature on the first shape, we find four most similar $\mathcal{M}^{[§]}$ features from the other shape, and similarly for the second shape. So, $\mathcal{M}$ stores pairs of possibly corresponding features, *e.g.* $\{\mathbf{i}, \mathbf{j}\}$. For scale, rotation and completeness invariance, distance between two features $\mathbf{f}$ and $\mathbf{f}'$ on the same shape is defined as: $\overline{\mathbf{f}\mathbf{f}'} = \|\mathbf{f}_x - \mathbf{f}'_x\|/\sigma_{\mathbf{f}}$, $\mathbf{f}_x$ references to the position of the $\mathbf{f}$, $\sigma_{\mathbf{f}}$ is the local scale of the point $\mathbf{f}$, and it is computed as the median of distances to 10 closest points from the correspondence set $\mathcal{M}$.

Then, the similarity of the correspondence $\mathrm{dist}(\mathbf{f}, \mathbf{g})$ now depends on the matching quality of other correspondences that are in the vicinity of $\mathbf{f}$ and $\mathbf{g}$. It means that the correspondence is correct if neighborhood correspondences are correct too, see Fig. 1. Formally,

$$\mathrm{dist}(\mathbf{f},\mathbf{g}) = 1 - \frac{1}{M}\sum_{\{\mathbf{i},\mathbf{j}\}\in\mathcal{M}}\exp\Big(-\min(\overline{\mathbf{f}\mathbf{i}},\overline{\mathbf{g}\mathbf{j}})^2/\sigma_\alpha^2\Big)\cdot$$
$$\cdot \exp\Big(-(\overline{\mathbf{f}\mathbf{i}}-\overline{\mathbf{g}\mathbf{j}})^2/\sigma_\beta^2\Big), \quad (2)$$

sum factor measures the configuration of $\{\mathbf{f}, \mathbf{g}\}$ to every correspondence from the set $\mathcal{M}$. The *sum* has two parts: first one weights $\{\mathbf{i}, \mathbf{j}\}$ correspondence high, if it is close enough to affect $\{\mathbf{f}, \mathbf{g}\}$. Second part weights $\{\mathbf{i}, \mathbf{j}\}$ high if it is in the correct configuration to $\{\mathbf{f}, \mathbf{g}\}$. So that, the distance $\mathrm{dist}(\mathbf{f}, \mathbf{g})$ will be low if correspondences that are close are in the correct configuration, see Fig. 1. Vicinity of $\{\mathbf{f}, \mathbf{g}\}$ is controlled by σ_α which is set to $5 \cdot \sigma_{\mathbf{f}}$ and $\sigma_\beta = \sigma_{\mathbf{f}}$ controls the error in the configuration of matches.

5. Query expansion

We now present schemes to re-issue a new query using the relevant (and irrelevant) subsets that result from the verification procedure, for a reliable, augmented result list.

[†] Tf-idf weighting [SZ03,PSZ08,BBGO11,KPW*10] is also used to weight high shape-specific words and to down-weight less specific words (*i.e.* commonly occurring ones). A visual word vocabulary size of 10% of the number features in the database has been found optimal (see [SZ03,KPW*10,CMPM11,PSZ08,JHS07]).

[‡] In practice, for $\mathbf{f}$, we found that it is enough to search only in the subset of first $N = 4$ most promising features from $\mathcal{G}$ based on the similarity in the descriptor space. We observed that the performance for $N = 1$ is low while the difference between $N = 4$ and higher is not significant.

[§] To create $\mathcal{M}$, we measured the similarity between features on the original descriptors instead of quantized visual words, they worked better here. As the computation cost of $\mathcal{M}$ could to be expensive, we used approx. nearest neighbor search [ML09]. Then, the computation time of d_{MHD} is less than 1sec in Matlab.

5.1. Average expansion (AE)

The most popular strategy to re-issue a new query is called average query expansion [CMPM11, J.71]. The mean of the BoW vectors associated with the top (10) shapes from the *verified, retrieved shapes* is used to construct a new query. For reliable verification, the expansion threshold from the first set of verified results must be set carefully to avoid inclusion of incorrect documents as it will destroy the query BoW vector (validation set is usually used here, we tuned the threshold for one dataset that has a validation data, then the threshold was kept for the rest of experiments).

5.2. Average expansion with negatives (AEneg)

We want to employ information from both the positive and negatively verified samples. In addition to augmenting with mean of the positive examples (similar to the original AE described above), we decreased new BoW vector $\mathbf{b}'$ by the mean of BoWs of shapes that were verified to be negative,

$$\mathbf{b}' = f\left(\frac{1}{P+1}\left(\mathbf{b}_q + \sum_{\mathbf{b}\in\mathcal{P}}\mathbf{b}\right) - \frac{1}{N}\sum_{\mathbf{b}\in\mathcal{N}}\mathbf{b}\right) \qquad (3)$$

$$f(\mathbf{b}_i) = \begin{cases} \mathbf{b}_i & \text{if } \mathbf{b}_i > 0 \\ 0 & \text{if } \mathbf{b}_i \le 0 \end{cases} \qquad (4)$$

where $\mathcal{P}$ is a set of BoWs of positive results shapes, $\mathcal{N}$ negatives, $\mathbf{b}_q$ is BoW of the query shape and $f(\cdot)$ avoids BoW negative values where the effect of negative samples is bigger than positives. Note that normalization is included later in the tf-idf weighting of $\mathbf{b}'$.

5.3. Pairwise coupling (PC)

In § (5.1-5.2), positives and negative verification were used to learn an implicit, modified distance measure between the various shapes. A more explicit procedure would allow for greater transparency. To achieve this, we used Pairwise coupling [WN05] (implemented in libsvm [CL01]) to learn a score-based classifier between the two sets (positive and negative sets are formed as in AEneg).

6. Applications of 3D verification and expansion

The use of the proposed approach in 3D search was already discussed in § 5. We now discuss the application of this method to other related applications namely classification and matching, where the results can be easily improved.

Classification with one training example. Classification ([LLS08, TCF09, KPW*10, BN10]) methods use a large corpus of **labeled** data for supervised learning. The corpus is used to train the model of the class which is later used to recognize the query shape. The process of manual labeling is time consuming and not always reliable. We therefore, explore the role of verification and expansion in reducing the need for annotation, whether retrieval and categorization can

mutually benefit from each other. Here, every class is represented by one shape for the training. The rest of the database (potentially extending to other databases) is used to **automatically** enrich training data for better performance by using verification (see § 4) and expansion (see § 5). So that, every shape is described by a new BoW vector that takes into account verified shape results.

Shape matching. An important role of matching is in registration and establishing denser correspondences for a variety of purposes (*e.g.* warping, deformation *etc.*). We explore the leverage provided by our expansion scheme in improving correspondence estimation. In practice, the proposed method can be used to improve any kind of matching algorithm (ranging from simple Hungarian to more complex [KLF11, OMMG10, OHG11]) as we **improve the shape's feature set** by adding features from relevant shapes.

The method to add relevant features into a shape is introduced now. We take unlabeled shapes from a database and we run shape search with verification to find relevant results. Feature matching [MS05] is performed between query and every positively verified result to establish correspondences. Denser correspondences can be obtained by propagating correspondences between the set of verified shapes in the result.

7. Evaluation

We have introduced schemes of retrieval verification in § 4 and schemes of expansion in § 5. We describe (§ 7.1) and compare (§ 7.2) several state-of-the-art (SOTA) methods with combinations of our verification and expansion proposals on the task of shape search. We then build on our understanding to explore how these ideas in verification and expansion can improve the classification performance in the presence of limited training data and shape matching (§ 7.3 and § 7.4).

7.1. Competitors

We first introduce three popular shape enhancement competitors for our approach. The subsection describes three methods of vanilla search improvement and two modifications of proposed WN.

7.1.1. Latent Dirichlet allocation (LDA)

LDA is a popular method for text and image search where a document is represented by a mixture of topics (see more [PSZ08]), and every topic is a distribution of words. We run standard LDA [XHCT] on all our datasets as a competitor.

7.1.2. *K*-reciprocal nearest neighbors (KRN).

This is a simple yet effective method [QGB*11] for shape search reordering where vanilla search results are used to define a different nearest neighbor metric based purely on

method	TOSCA				Princeton				SHREC'09			
	verifi.	AE	AEneg	PC	verifi.	AE	AEneg	PC	verifi.	AE	AEneg	PC
W2	0.624	0.635	0.649	0.634	0.283	0.337	0.339	0.326	0.262	0.366	0.368	0.290
WN-global	0.635	0.713	**0.718**	0.671	0.284	0.335	0.332	0.321	0.267	0.389	0.389	0.296
WN	0.629	0.624	0.678	0.672	0.284	0.336	0.337	0.323	0.278	0.393	**0.395**	0.313
KRN [QGB*11]	0.626	0.667	0.684	0.668	0.283	0.322	0.322	0.308	0.275	0.338	0.340	0.283
LDA [PSZ08]	0.627				**0.344**				0.328			
DMLLM [PSZ10]	0.632				0.272				0.262			
Vanilla [KPW*10]	0.622				0.282				0.277			

Table 1: *Shape search results. Performance is measured as the average area under PR curve. Th proposed verification and expansion outperform basic search. While incorporating negatives into average expansion improve the performance, it is not very significant. WN-global and W2 are variants of the proposed methods and are discussed in the text.*
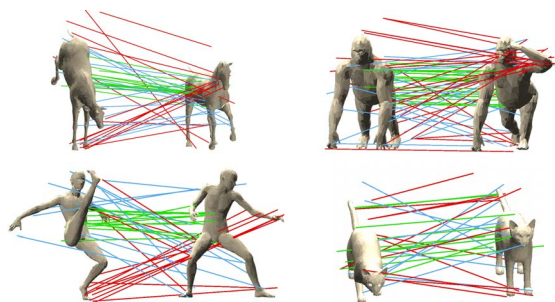
Figure 2: *Features that were selected for DMLL training. Example of four pairs of shapes that were verified to be relevant to each other. Three groups of features are highlighted: 1) red for negatives, where features are matches after nn-matching but they have high WN distance; 2) green for positives, so they have low WN distance; 3) blue shows the random set. Some features are outside the shape due to properties of 3D SURF [KPW*10] as i.e. "U"-like structures will have a feature in the middle.*

symmetricity of mutual BoW similarities. In this method, i-th result shape (s_i) in query's retrieved list is considered verified if the query appeared in the first K results of the query using s_i. For this method, we ran experiments for different K and selected the one with the highest performance (see Qin *et al.* [QGB*11]). The specialty of this method is its speed and ability to use any even global descriptor.

7.1.3. Distance metric learning for large margin (DMLLM)

We used Philbin *et al.* [PSZ10] for use in unsupervised 3D shape distance learning. After vanilla search and WNverification (§ 4.1) , we performed feature matching which leads to these three sets of features: 1) correct (high WNsimilarity, **green** in Fig. 2); 2) not correct (**red** in Fig. 2) and 3) features of random negative pairs (to decrease the confusion between negative and positive training features **blue** in Fig. 2). Then, we follow the work of Philbin *et al.* [PSZ10] and we learn a mapping of descriptors to separate these three sets from each other. When the mapping is learnt, we project every

shape's feature descriptor to the new space and the shape search pipeline was recomputed to obtain new results.

7.1.4. Modifications of WN

We also evaluated two modifications of WN:

WN-global: uses a global point scale. Herein, $\sigma_{\mathbf{f}}$ corresponds to the size of the shape. So that, the parameter is estimated globally and it is constant for all shape's points.

W2: is *2 part weak shape model* modification with the most simple spatial layout a' la ISM [LS03, KPW*10]. The method assumes that the observation of any one feature (part 1, feature) in a specific configuration w.r.t. to the object centre (part 2, anchor) is sufficient to verify the shape. Herein, $dist(\mathbf{f}, \mathbf{g})$ in Eq. 1 is low when the relative distances of $\mathbf{f}$ and $\mathbf{g}$ to the shape's centers are similar. Formally, $dist(\mathbf{f}, \mathbf{g}) = \left(\overline{\mathbf{o}_q\mathbf{f}} - \overline{\mathbf{o}_r\mathbf{g}}\right)^2 / \sigma_{W2}^2$, where $\mathbf{o}$ is the shape centre and σ_{W2} corresponds to 1% of unit shape size and it controls the maximal error between matches $\mathbf{f}$ and $\mathbf{g}$.

7.2. Expansion in shape retrieval

Having introduced the competitors in § 7.1 we pitch combinations of our verification and expansion against them. We test this on the established datasets: *Tosca* [BBK08] of 120 hand-made shapes of 9 classes, *Princeton* [SMKF04] of 1.8K shapes and *SHREC'09* [SHA] with 20 challenging partial queries for 720 shapes. We also run our methods on SHREC'10 [BBGO11], where results of vanilla search are visually similar to BoW of Bronstein *et al.* [BBGO11], but ground-truth to evaluate results wasn't sent to us.

The effect of verification and expansion is shown in Fig. 3. Results are shown in Tab. 1 and Fig. 4. We can conclude that verification with expansion significantly improves vanilla results. LDA gives significant improvement (and is even winner for the *Princeton* dataset), while it has no spatial information. On the other hand, LDA doesn't perform well on *TOSCA* and has significantly decreased performance when small vocabularies are used (not shown here). WN-global gives retrieval improvement on complete, clean   shapes
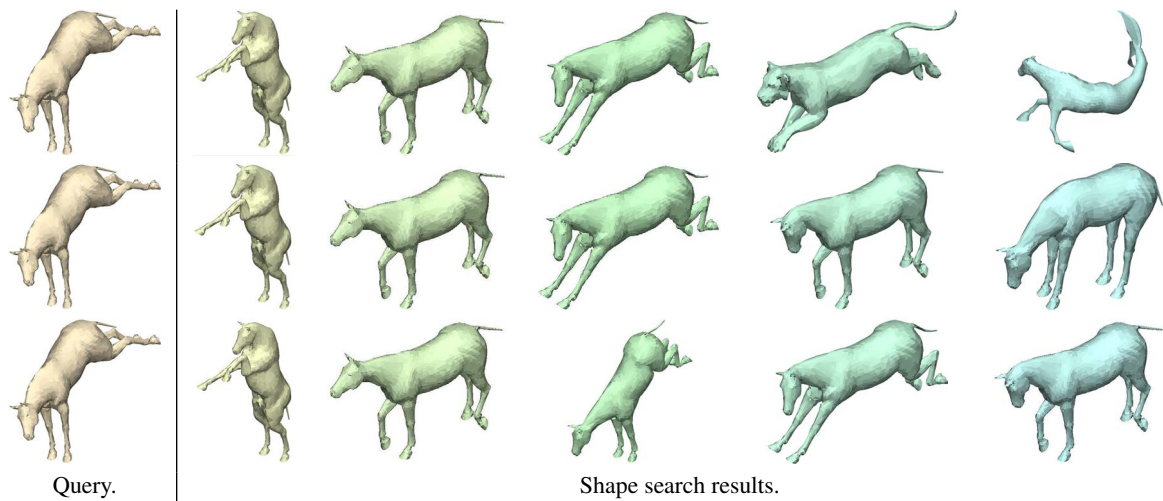
Figure 3: *Shape search results after each step. Note that in the first row –results after vanilla search– several shapes are not from the same category but they have similar shape/pose. Verification method WN in the second row improves results and when expansion is used (third row), we also see the diversity in the result-list.*
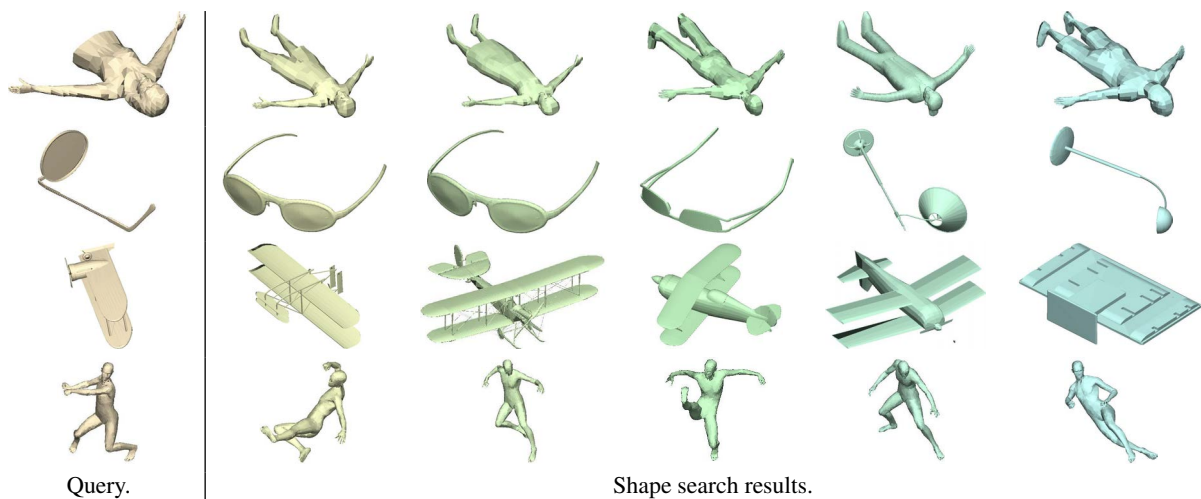

Figure 4: *Shape search results. Every row starts with a query shape and then shapes sorted by the rank continues. Note that query expansion adds diversity to the result list, as shapes have to be similar to all shapes that were used for expansion.*

(*Tosca*), while local information matters when only partial queries are given such as in *SHREC'09*. Note that this method is still local and only the parameter for feature's neighborhood is estimated globally. While W2 improves Vanilla search, the performance is mostly bellow the proposed WN. In conclusion, we note that local verification is important. A number of shape search results after expansion are presented in Fig. 4. Though performance can depend on the specific dataset, verification with expansion generally leads to improved shape search. This demonstrates the utility of incorporating spatial layout and using class-specific information (that could be missing in the original query).

7.3. Expansion for classification

Here, we present expansion and verification to improve classification when few training examples are available. In Tab. 2 results are shown for when: 1) standard classification (ST) is performed given a large amount of labeled training data, 2) We randomly selected *one* shape for training our one-example classification (OE). As expected, the performance drops considerably. 3) We used the rest of training data for expansion to expand the training set as well as the test set (OE+QE), see § 6 for the details of the method.

We used the same datasets as in § 7.2 using the classifiers:
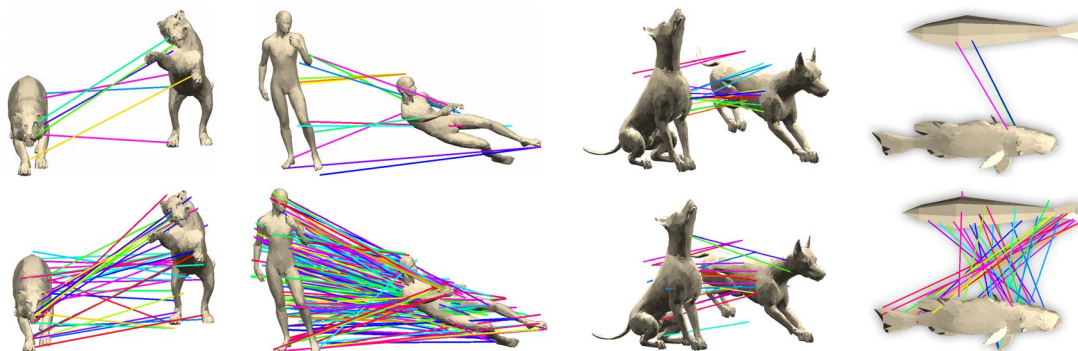
Figure 5: *Matching. First row shows results of matching with original features, second shows the same matching algorithm and parameters on original features plus expanded.*

method	*TOSCA*		*SHREC'09*	
	knn	SVM	knn	SVM
ST: standard	89.4%	79.0%	50.0%	60.0%
OE: one-example	47.3%	54.6%	45.0%	40.0%
OE+QE: one-example+QE	63.2%	63.2%	45.0%	65.0%

Table 2: Performance of the classification.

SVM [TCF09] and k-nearest neighbor [Bis06] . Tab. 2 shows that expanding data for classification improve OE. OE+QE even outperformed ST in one case, (probably because the queries, often partial shapes from SHREC'09, benefit from the expansion exploiting relevant structure).

7.4. Expansion for denser shape matching

We present an example of using expansion to improve shape matching. Fig. 5 shows results of matching shapes with original and expanded features, as described in § 6. This attempt at matching, ignores the more sophisticated conditions of matching (as in [PKVG11,KLF11,OMMG10], but estimates correspondences in less than $0.5s$ (compared to tens of mins for [PKVG11,KLF11,OMMG10] or minutes in Hungarian). The initial attempt shows promise for deeper application in this field.

8. Conclusion

As they say *the proof of the pudding is in the eating*, and while improving on the *vanilla* flavor of shape retrieval was relatively easy, we hope to have convinced the reader about the power and range of applications afforded by shape verification and expansion mechanisms.

As promised, we improve upon state-of-the-art 3D shape retrieval with a simple yet reliable, cascade method of verification and expansion, relying on weak structural verification. This allows us to have significant improvement while using an automatic, unsupervised and fast scheme.

While primarily developed for retrieval, the generality of presented method and the interplay between the retrieval and the other areas of visual learning, provides us with a rich playground of problems to improve upon.

Though 3D databases are expanding rapidly, a variety of their clientele (such as animators) are interested in some categories more than others, thus making the role of categorization in retrieval particularly important. What's more, preliminary attempts also provide us with promise in the fields of registration and dense matching.

References

[AKW07] AKBAR S., KÜNG J., WAGNER R.: Multi-feature integration with relevance feedback on 3d model similarity retrieval. *J. Mob. Multimed. 3* (September 2007), 235–254. 2

[AZ10] ARANDJELOVIĆ R., ZISSERMAN A.: Efficient image retrieval for 3D structures. In *Proc. BMVC.* (2010). 2

[BBGO11] BRONSTEIN A. M., BRONSTEIN M. M., GUIBAS L. J., OVSJANIKOV M.: Shape Google: Geometric words and expressions for invariant shape retrieval. *ACM ToG* (2011). 1, 2, 3, 5

[BBK08] BRONSTEIN A., BRONSTEIN M., KIMMEL R.: *Numerical Geometry of Non-Rigid Shapes*. Springer Publishing Company, Incorporated, 2008. 5

[BC02] BANG H. Y., CHEN T.: Feature space warping: an approach to relevance feedback. In *Intl. Conf. Image Proc.* (2002), vol. 1, pp. I–968 – I–971 vol.1. 2

[Bis06] BISHOP C.: *Pattern Recognition and Machine Learning*. Springer, New York, 2006. 7

[BN10] BARIYA P., NISHINO K.: Scale-hierarchical 3d object recognition in cluttered scenes. In *CVPR* (2010). 4

[CL01] CHANG C.-C., LIN C.-J.: *LIBSVM: a library for support vector machines*, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm. 4

[CMPM11] CHUM O., MIKULIK A., PERDOCH M., MATAS J.: Total recall ii: Query expansion revisited. In *CVPR* (2011). 2, 3, 4

[DGA*09] DUTAGACI H., GODIL A., AXENOPOULOS A., DARAS P., FURUYA T., R. OHBUCHI R.: Shrec 2009 - shape retrieval contest of partial 3d models. 1

[DJ94] DUBUISSON M.-P., JAIN A.: Modified hausdorff distance for object matching. In *ICPR* (1994). 3

[Efr08] EFRON M.: Query expansion and dimensionality reduction: Notions of optimality in rocchio relevance feedback and latent semantic indexing. *Inf. Pr. Manage.* (2008), 163–180. 2

[ETA02] ELAD M., TAL A., AR S.: Content based retrieval of vrml objects: an iterative and interactive approach. In *Eurographics WS on Multimedia* (2002), Springer-Verlag. 1, 2

[FGMR09] FELZENSZWALB P. F., GIRSHICK R. B., MCALLESTER D., RAMANAN D.: Object detection with discriminatively trained part based models. *PAMI* (2009). 3

[FKS*04] FUNKHOUSER T. A., KAZHDAN M. M., SHILANE P., MIN P., KIEFER W., TAL A., RUSINKIEWICZ S., DOBKIN D. P.: Modeling by example. *ACM Trans. Graph.* (2004). 1

[GFSF10] GIORGI D., FROSINI P., SPAGNUOLO M., FALCIDIENO B.: 3d relevance feedback via multilevel relevance judgements. *The Visual Computer* (2010). 2

[Goo] http://sketchup.google.com/3dwarehouse/. 1

[HJ11] HOI S. C. H., JIN R.: Active multiple kernel learning for interactive 3d object retrieval systems. *ACM Trans. Interact. Intell. Syst. 1* (2011), 3:1–3:27. 2

[J.71] J. R.: Relevance feedback in in information retrieval. In *The SMART retrieval system - Experimetns in Automatic Document Processinig* (1971). 4

[JDS10] JÉGOU H., DOUZE M., SCHMID C.: Improving bag-of-features for large scale image search. *IJCV* (2010). 2

[JH99] JOHNSON A. E., HEBERT M.: Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE PAMI 21*, 5 (1999), 433–449. 2

[JHS07] JÉGOU H., HARZALLAH H., SCHMID C.: A contextual dissimilarity measure for accurate and efficient image search. In *CVPR* (2007). 3

[KLF11] KIM V. G., LIPMAN Y., FUNKHOUSER T.: Blended intrinsic maps. *Trans. on Graphics* (2011). 2, 3, 4, 7

[KLM*12] KIM V. G., LI W., MITRA N., DIVERDI S., FUNKHOUSER T.: Exploring collections of 3d models using fuzzy correspondences. *Trans. on Graphics* (2012). 2

[KPW*10] KNOPP J., PRASAD M., WILLEMS G., TIMOFTE R., VAN GOOL L.: Hough transform and 3d surf for robust three dimensional classification. In *ECCV* (2010). 1, 2, 3, 4, 5

[LJI*03] LOU K., JAYANTI S., IYER N., KALYANARAMAN Y., PRABHAKAR S., RAMANI K.: A reconfigurable 3d engineering shape search system: Part ii — database indexing, retrieval, and clustering. *ASME* (2003). 2

[LLS08] LEIBE B., LEONARDIS A., SCHIELE B.: Robust object detection with interleaved categorization and segmentation. *IJCV 77* (2008), 259–289. 4

[LS03] LEIBE B., SCHIELE B.: Interleaved object categorization and segmentation. In *BMVC03* (2003), vol. 2, pp. 264–271. 5

[ML09] MUJA M., LOWE D.: Fast approximate nearest neighbors with automatic algorithm configuration. In *VISAPP* (2009). 3

[MP96] MINKA T. P., PICARD R. W.: Interactive learning with a "society of models". In *Proc. CVPR* (Washington, DC, USA, 1996), CVPR '96, IEEE Computer Society, pp. 447–. 2

[MPCM10] MIKULIK A., PERDOCH M., CHUM O., MATAS J.: Learning a fine vocabulary. In *ECCV* (2010). 2

[MS05] MIKOLAJCZYK K., SCHMID C.: A performance evaluation of local descriptors. *PAMI* (2005). 4

[OHG11] OVSJANIKOV M., HUANG Q.-X., GUIBAS L. J.: A condition number for non-rigid shape matching. *Comput. Graph. Forum* (2011). 3, 4

[OLGM11] OVSJANIKOV M., LI W., GUIBAS L., MITRA N. J.: Exploration of continuous variability in collections of 3d shapes. *ACM Transactions on Graphics* (2011). 1

[OMMG10] OVSJANIKOV M., MÉRIGOT Q., MÉMOLI F., GUIBAS L.: One point isometric matching with the heat kernel. In *Eurographics Symposium on Geometry Processing (SGP)* (2010). 2, 3, 4, 7

[PCI*07] PHILBIN J., CHUM O., ISARD M., SIVIC J., ZISSERMAN A.: Object retrieval with large vocabularies and fast spatial matching. In *Proc. CVPR* (2007). 2

[PKVG11] PRASAD M., KNOPP J., VAN GOOL L.: Class-specific 3D Localization using Constellations of Object Parts. In *Proc. BMVC.* (2011). 7

[PSZ08] PHILBIN J., SIVIC J., ZISSERMAN A.: Geometric LDA: A generative model for particular object discovery. In *Proc. BMVC.* (2008). 2, 3, 4, 5

[PSZ10] PHILBIN J., SIVIC J., ZISSERMAN A.: Descriptor learning for efficient retrieval. In *ECCV* (2010). 5

[PTTS08] PAPADAKIS I. P. P., T. T., T. T., S. P.: Relevance feedback in content-based 3d object retrieval a comparative study. In *Computer-Aided Design & Applications* (2008). 1, 2

[QGB*11] QIN D., GAMMETER S., BOSSARD L., QUACK T., GOOL L. J. V.: Hello neighbor: Accurate object retrieval with k-reciprocal nearest neighbors. In *CVPR* (2011). 4, 5

[SHA] SHARP: Shrec 2009 - shape retrieval contest of partial 3d models. http://www.itl.nist.gov/iad/vug/sharp/benchmark/shrecPartial/. 5

[SMKF04] SHILANE P., MIN P., KAZHDAN M., FUNKHOUSER T.: The princeton shape benchmark. In *SMI* (2004). 1, 5

[SOG09] SUN J., OVSJANIKOV M., GUIBAS L.: A concise and provably informative multi-scale signature based on heat diffusion. In *SGP* (2009), pp. 1383–1392. 2

[SSS06] SNAVELY N., SEITZ S. M., SZELISKI R.: Photo tourism: Exploring photo collections in 3d. In *SIGGRAPH Conference Proceedings* (2006), pp. 835–846. 1

[SZ03] SIVIC J., ZISSERMAN A.: Video Google: A text retrieval approach to object matching in videos. In *ICCV* (2003). 1, 2, 3

[TCF09] TOLDO R., CASTELLANI U., FUSIELLO A.: A bag of words approach for 3d object categorization. In *MIRAGE* (2009). 1, 2, 4, 7

[VG06] VERGAUWEN M., GOOL L. V.: Web-based 3d reconstruction service. *Mach. Vision Appl. 17*, 6 (2006), 411–426. 1

[VS01] VRANIC D. V., SAUPE D.: 3d shape descriptor based on 3d fourier transform. In *ECMCS* (2001). 2

[WN05] WU B., NEVATIA R.: Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors. In *Proc. ICCV* (2005), vol. 1. 4

[XHCT] XUAN-HIEU P., CAM-TU N.: Gibbslda++. http://gibbslda.sourceforge.net/. 4