

## Improving 3D Shape Retrieval Methods based on Bag-of-Feature Approach by using Local Codebooks

El Wardani Dadi<sup>1,\*</sup>, El Mostafa Daoudi<sup>1</sup> and Claude Tadonki<sup>2</sup>

<sup>1</sup>University Mohammed First, Faculty of Sciences, LaRi Laboratory  
Oujda (Morocco)

<sup>2</sup>Mines ParisTech, Laboratoire de Recherche en Informatique Mathématiques et  
Systèmes, Fontainebleau (France)

wrd.dadi@gmail.com, m.daoudi@fso.ump.ma, claude.tadonki@mines-paristech.fr

### Abstract

Recent investigations illustrate that view-based methods, with pose normalization pre-processing get better performances in retrieving rigid models than other approaches and still the most popular and practical methods in the field of 3D shape retrieval [1, 2, 3, 4, 5]. In this paper we present an improvement of 3D shape retrieval methods based on bag-of-features approach. These methods use this approach to integrate a set of features extracted from 2D views of the 3D objects using the SIFT (Scale Invariant Feature Transform [6]) algorithm into histograms using vector quantization which is based on a global visual codebook. In order to improve the retrieval performances, we propose to associate to each 3D object its local visual codebook instead of a unique global codebook. The experimental results obtained on the Princeton Shape Benchmark database [6], for the BF-SIFT method proposed by Ohbuchi, et al., [2] and CM-BOF proposed by Zhouhui, et al., [3], show that the proposed approach performs better than the original approach.

**Keywords:** 3D-Content-based Shape Retrieval; Bag-Of-Features; SIFT; Vector Quantization; Codebook

### 1. Introduction

Currently, there are an increasing number of 3D objects on the web, including large databases, thanks to recent digitizing and modeling technologies. The need of efficient methods for 3D shape-content based retrieval, in order to ease navigation into related large databases, and also to structure, organize and manage this new multimedia type of data, has become an active topic in various research communities such as computer vision, computer graphics, mechanical CAD, and pattern recognition.

One major challenge in 3D objects indexation is to design an efficient canonical characterization of the objects. In the literature, this characterization is referred to as a descriptor or a signature. Since the descriptor serves as a key in the search process, it is a critical kernel with a strong influence on the searching performances (i.e. computational efficiency and relevance of the results).

Various 3D shape description methods have been proposed in the literature. The reader may refer to a very good survey in [1] and a comparative study of 3D retrieval algorithms [7, 8, 9]. Those algorithms can be clustered into two main families: 2D/3D approaches and 3D/3D approaches. For 2D/3D approaches, the description model is obtained through

---

\* Supported by the "Excellence Grant of Moroccan Ministry of Higher Education. Grant No. G08/004".

different 2D projections of the 3D shape, whereas for the 3D/3D approaches, the description model is obtained from the 3D information directly extracted from the 3D shape. Recent investigations illustrate that view-based methods with pose normalization pre-processing get better performance in retrieving rigid models than other approaches and still the most popular and practical methods in the field of 3D shape retrieval [1, 2, 3, 4, 5].

Our work presented in this paper is inspired by the BF-SIFT method (Ohbuchi, et al., [2]), which is based on a global codebook (visual dictionary) used to describe each 3D objects in the database. We propose an improvement of the method by using local codebooks, since we think that the use of a unique global codebook badly influences the retrieval performance. For this, we propose to associate to each 3D object its local codebook instead of a unique global codebook.

To show the efficiency of our proposed improvement, we have selected two 3D shape retrieval methods from the literature based on bag-of-features approach; the first one is the BF-SIFT (Bag of Features SIFT) proposed by Ohbuchi, et al., [2], the second one is CM-BOF (Clock Matching Bag-Of-Features) proposed by Zhouhui Lian, et al., [3]. These methods use one global and unique codebook. Experimental are performed on the Princeton Shape Benchmark (PSB) [7] that contains various shapes with more geometric details. The obtained results show that our technique provides more accurate results.

The paper is organized as follows. Section 2 describes the BF-SIFT and CM-BOF algorithms. Our proposed improvement is presented in Section 3. Experimental results are provided and analyzed in Section 4. Section 5 concludes the paper and outlines some perspectives.

## 2. Presentation of BF-SIFT and CM-BOF Methods

### 2.1 The BF-SIFT Method

The BF-SIFT (Bag-of-Features - Scale Invariant Feature Transform) method proposed by Ohbuchi, et al., [2] compares 3D shapes using thousands of local visual features per model. A 3D model is rendered into a set of depth images, and from each image, local visual features are extracted by using the Scale Invariant Feature Transform (SIFT) algorithm of Lowe [6]. To efficiently compare among a large set of local visual features, the algorithm uses bag-of-features (BoF) approach in order to integrate, for each model, the local features into a vector of features. The BoF approach vector quantifies (or encodes) the SIFT features into a representative vector (or “visual word”), using a global codebook. The global codebook is generated with thousands of features extracted from a set of models in the retrieval database. In the following, we present an overview of the BF-SIFT algorithm:

- *Pose normalization (position and scale):* The BF-SIFT performs pose normalization only for position and scale, so that the model is rendered with an appropriate size in each of the multiple-view images. Pose normalization is not performed for rotation.
- *Multi-view rendering:* in this step, a set of depth-buffer views of a 3D object are captured uniformly in all directions in order to catch up all symmetries.
- *SIFT feature extraction:* a 3D object can be approximately represented by a set of depth-buffers from which salient SIFT descriptors are extracted using the SIFT algorithm [6].
- *Vector quantization:* each 3D model is associated with thousands of local features. Each SIFT feature extracted from 3D models is quantified as a *vector* or *visual word* by using

a global visual codebook. The vector quantification is to find frequencies of visual words (local features) generated from a model in the visual codebook which is learned, by using a clustering algorithm type (e.g. *k-means*, *kd-tree*, *ERC-tree*, and *Locality sensitive hashing*).

- *Histogram generation*: quantified local features or “visual words” are accumulated into a histogram with  $N_v$  bins ( $N_v$  is considered as the size of the codebook). The histogram becomes the feature vector of the corresponding 3D model.
- *Distance computation*: Dissimilarity among pairs of feature vectors (the histograms) is computed by using Kullback-Leibler Divergence (KLD).

$$D(x, y) = \sum_{j=1}^{N_v} (y_j - x_j) \ln \frac{y_j}{x_j}$$

Where  $x = (x_i)$  and  $y = (y_i)$  are the features vectors and  $N_v$  the dimension of the vectors.

## 2.2 CM-BOF Method

Zhouhui, et al., [3] have proposed a new visual similarity based 3D shape retrieval approach, which uses Bag-of-Features approach. The method is inspired by the work of Ohbuchi, et al., [2]. This method called CMBOF (Clock Matching Bag-Of-Features) [3] follows the same steps of BF-SIFT to describe a 3D object and to compute the similarity between two 3D objects. The main difference with BF-SIFT is that, in BF-SIFT we compute one descriptor of a given 3D object, whereas the CM-BOF approach describes each 2D view, captured around the 3D object, as a word histogram. In this case a 3D object is represented by more than one descriptor, and then employs an efficient multi-view shape matching scheme (called Clock Matching) to measure the dissimilarity between two 3D objects by computing the minimum distance between all of their possible matching pairs (24), which corresponds to all possible poses of the 3D object in canonical axis. Practically, when we compare two 3D objects, one of them will be placed in the original orientation while the other one may appear in 24 different poses.

## 3. Improvements based on Local Codebooks

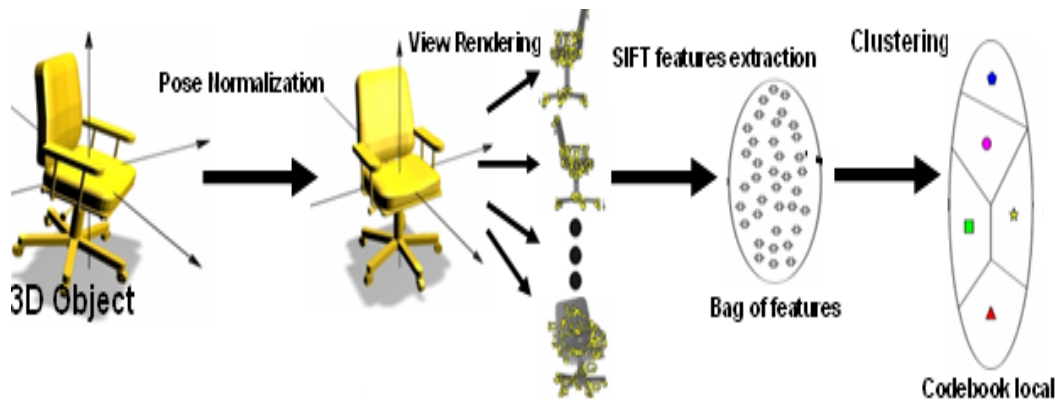
The construction of the visual codebook is one of the sensitive stages. Indeed, the descriptor of each object in the database will be calculated using the visual words in the codebook. For that, it is important to generate codebooks that are as representative as possible. In our variant, we propose to use local codebooks instead of the global codebook used in the two methods described previously, by associating each 3D object in the database with its codebook. In this case, the vector quantification is based on local codebooks instead of a unique global codebook as in the original methods. The different steps for the retrieval process, based on our approach, are similar to the original algorithm but with some differences during the following three steps: Generation of the codebook, Computation of the descriptors and 3D-Shape matching. For our approach (approach based on local codebook), these steps are performed as follows:

- *Generation of the codebook:* we associate to each 3D mode in the database, its local codebook. This local codebook is learned from the features extracted from the target 3D model using a clustering algorithm.
- *Computation of the descriptors:* the descriptor of a given 3D object in the database is computed by using the local codebook associated to the target 3D object.
- *3D-Shape matching:* to compare a 3D object query with a given 3D Object in the database, the descriptor of the query is computed by using the local codebook of the target 3D Object. In this case the descriptor of the query is computed at each comparison.

While for the original approach (based on global codebook), these steps are performed as follows:

- *Generation of the codebook:* The codebook is clustered from the SIFT features extracted from all 3D objects in the database. In this case, to all 3D-objects in the database is associated a unique codebook.
- *Computation of the descriptors:* the descriptor of a given 3D object in the database is computed by using the global codebook.
- *3D-Shape matching:* To compare a 3D object query with a given 3D Object in the database, the descriptor of the query is computed by using the global codebook. In this case the descriptor of the query is not computed at each comparison.

Figure 1 presents the different steps of local codebook generation



**Figure 1. Learning of Local Codebook**

In Figures 2 and 3, an example of different steps of comparing two 3D objects using the BF-SIFT method based on a local codebook.

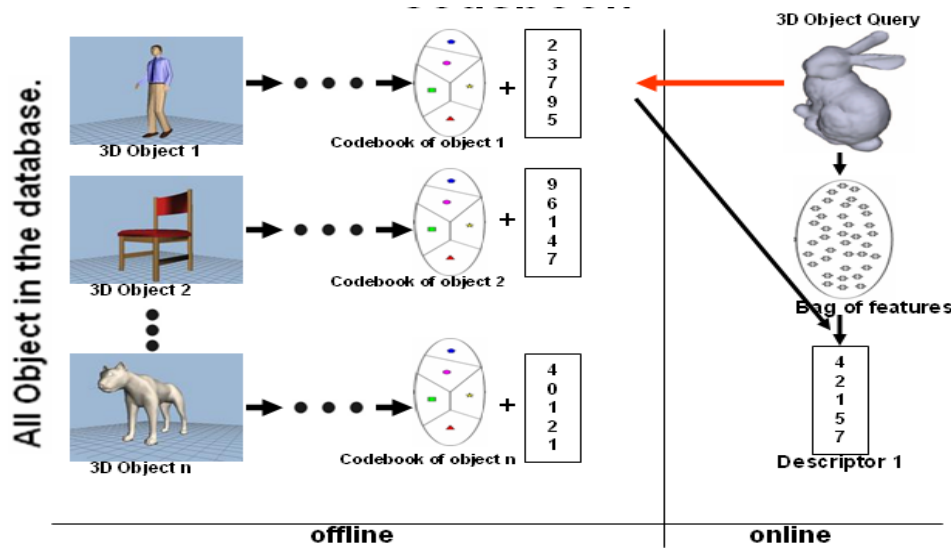


Figure 2. Processing of Comparison between a 3D Object-query and the First 3D Model in Database based on Local Codebook

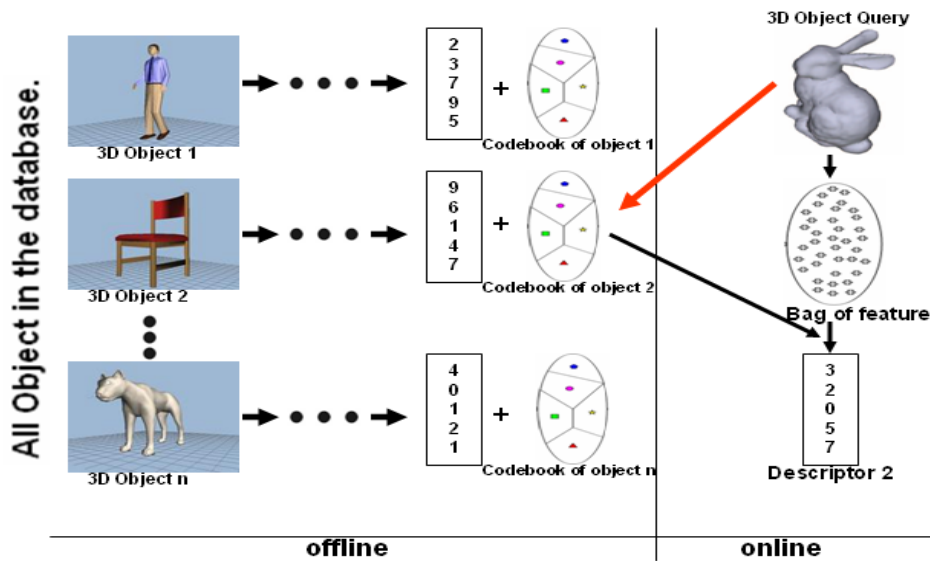


Figure 3. Processing of Comparison between a 3D Object-query and the Second 3D Model in Database based on Local Codebook

#### 4. Experiments and Results

Our tests are made on the Princeton 3D Shape Benchmark database [7] with a set of various rigid shapes. For implementation, we proceed as follows:

- For view rendering in both methods BF-SIFT and CM-BOF, we capture 66 depth buffer 2D views around a given 3D object using an OpenGL executable program of Zhouhui Lian [3].

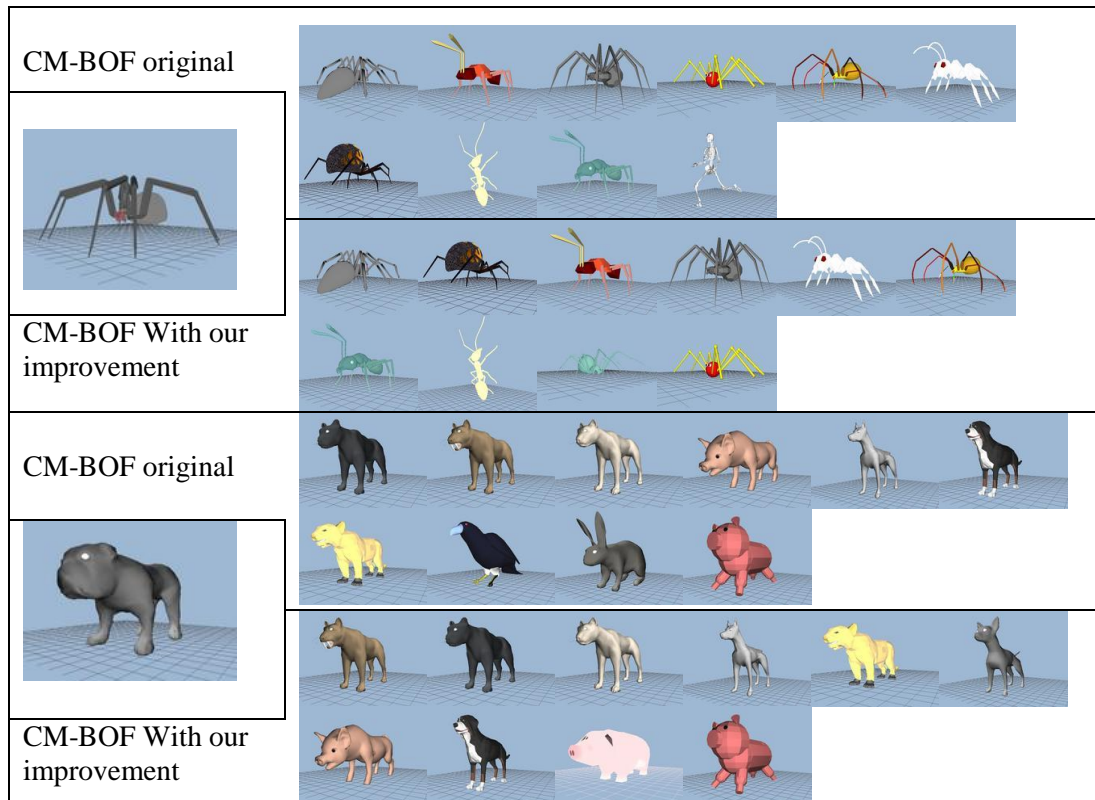
- To extract local feature from a depth-buffer view, the SIFT is implemented using the VLFeat MATLAB source of Veldaldi [10].
- To learn the codebook, we use the k-means function also from the VLFeat MATLAB source of Veldaldi [10], in order to cluster the set of local features by setting  $N_v$  to the size of vocabulary.
- For vector quantification, we use the MATLAB implementation of the linear k-nearest neighbor (KNN) search.

In the first experimentations, we compare the retrieval results of the BF-SIFT method with and without our improvement. Figure 3 shows that, for our approach (local codebooks) all top 6 retrieved objects are similar to the query, while for the original methods (methods based on global codebook), there are retrieved objects from the top 6 are not similar to the query (the retrieved objects number 2 for the queries 1 and 2 and the retrieved object number 4 for the query 3).



**Figure 3. The 6 Top 3D Objects Retrieved from Different “Class” query, using the BF-SIFT and our Method**

In the following second experimentations, we compare the retrieval results of the CM-BOF method with and without our improvement. Figure 4 shows that, for our approach (local codebooks) all top 10 retrieved objects are similar to the query, while for the original methods (methods based on global codebook), there are retrieved objects from the top 10 are not similar to the query (the retrieved objects number 10 for the query 1 and the retrieved object number 8 for the query 2).



**Figure 4. The 10 Top 3D Objects Retrieved from Different “Class” query, using the CM-BOF with and without our Improvement**

We think that, using our approach, the similarity distance between two similar objects is minimal; since the local codebook is generated by using features of the target 3D-object, then it is close to this object. For a 3D-object query, we first compute its descriptor using the local codebook. If the query is similar to the target object then its features are similar to the features of the target object, therefore the query object is also close to the local codebook (it contains the features of the target object).

Experimental results corroborate with the theoretical idea, and show that our improvement (using local codebooks) applied to the BF-SIFT and CM-BOF methods performs better than the original ones (BF-SIFT and CM-BOF based on global codebook) but at the expense of more computational cost.

## 5. Conclusion and Perspectives

In this paper we have proposed a new approach to improve retrieval performance of 3D shape retrieval method based on bag of features. The key idea is to use local codebooks to vector quantization of salient local features, extracted from a given 3D object, basing on its associated codebook instead of using a global unique codebook used in the original methods. To validate our approach, we have compared the performances of two methods; the BF-SIFT and CM-BOF with and without using our improvement. For the experimental tests, we have used the Princeton 3D Shape Benchmark database with a set of different and rigid shapes. The top k retrieval results show that, using the local codebooks performs better than the

original approach based on global codebook. To reduce the computation costs, we are working for exploiting the new computing platforms (GPU, multi Core).

## Acknowledgements

We thank T. Furuya and L. Zhouhui for their precious description of the BF-SIFT and CM-BOF methods and for the use of the OpenGL executable program of Zhouhui.

## References

- [1] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods", *Multimedia Tools and Applications*, vol. 39, no. 3, (2008) September, pp. 441–471.
- [2] R. Ohbuchi, K. Osada, T. Furuya and T. Banno, "Salient local visual features for shape-based 3D model retrieval", *Proc. IEEE Shape Modeling International (SMI)*, (2008), pp. 93-102.
- [3] Z. Lian, A. Godil and X. Sun, "Visual Similarity based 3D Shape Retrieval Using Bag-of-Features", *IEEE International Conference On Shape Modeling and Applications (SMI)*, (2010).
- [4] G. Passalis, T. Theoharis and I. A. Kakadiaris, "PTK: A novel depth buffer-based shape descriptor for three dimensional object retrieval", *The Visual Computer*, vol. 23, no. 1, (2007), pp. 5–14.
- [5] J. Shih, C. Hsing and J. Wang, "A new 3D model retrieval approach based on the elevation descriptor", *Pattern Recognition*, vol. 40, no. 1, (2007), pp. 283–295.
- [6] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *Int'l Journal of Computer Vision*, vol. 60, no. 2, (2004) November.
- [7] P. Shilane, P. Min, M. Kazhdan and T. Funkhouser, "The Princeton shape benchmark," in *Shape Modeling and Applications Conference, SMI'2004*, Genova, Italy, (2004) June, IEEE, pp. 167–178.
- [8] T. Zaharia and F. Preteux, "3D versus 2D/3D shape descriptors: A comparative study", in *SPIE Conf. on Image Processing: Algorithms and Systems III - IS & T/ SPIE Symposium on Electronic Imaging, Science and Technology '03*, San Jose, CA, (2004) January, vol. 5298.
- [9] B. Bustos, D. A. Keim, T. Schreck and D. Vranic, "An experimental comparison of feature-based 3D retrieval methods", in *2nd Int. Symp. on 3D Data Processing, Visualization, and Transmission (3DPVT'04)*, Thessaloniki, Greece, (2004) September.
- [10] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms", <http://www.vlfeat.org/>, (2008).

## Authors



**El Wardani DADI**

He is currently a PhD. Student, in Computer Science, at Faculty of Sciences of University of Mohammed First-Morocco. His research interests include 3D shape retrieval, application in high performance computing.



**El Mostafa DAOUDI**

He received the Ph.D. degree in Parallel Computing from Institut National Polytechnique of Grenoble-France in 1989 and the PhD degree in Computer Sciences from Faculté Polytechnique of Mons-Belgium in 1991. He is currently Professor at Faculty of Sciences, University of Mohammed First-Morocco. His research interests include High Performance Computing, Parallel Algorithms and Complexity, 3D-Imaging, Parallel Scheduling.





**Claude M. TADONKI**

He received the PhD degree in Computer Science in 2001 and is now a senior researcher and lecturer at the Mines ParisTech institute (Paris/France) since 2011. His main research topics included High Performance Computing, Operation Research, Matrix Computation, Combinatorial Algorithm and Complexity, Scientific and Technical Programming, Automatic Code Transformations. He has worked at several laboratories and universities. He has initiated various scientific projects and national/international collaborations, and has given significant number of CS courses in different contexts including industries. He is an active member of well established scientific corporations and reviewer of international journals and conferences. Personal web page: [www.omegacomputer.com/staff/tadonki](http://www.omegacomputer.com/staff/tadonki).

