

# Partial 3D Object Retrieval combining Local Shape Descriptors with Global Fisher Vectors

M. A. Savelonas<sup>1,2</sup>, I. Pratikakis<sup>1,2</sup> and K. Sfikas<sup>2</sup>

<sup>1</sup>ATHENA Research and Innovation Center, Xanthi, Greece

<sup>2</sup>Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece

---

## Abstract

*This work introduces a partial 3D object retrieval method, applicable on both meshes and point clouds, which is based on a hybrid shape matching scheme combining local shape descriptors with global Fisher vectors. The differential fast point feature histogram (dFPFH) is defined so as to extend the well-known FPFH descriptor in order to capture local geometry transitions. Local shape similarity is quantified by averaging the minimum weighted distances associated with pairs of dFPFH values calculated on the partial query and the target object. Global shape similarity is derived by means of a weighted distance of Fisher vectors. Local and global distances are derived for multiple scales and are being combined to obtain a ranked list of the most similar complete 3D objects. Experiments on the large-scale benchmark dataset for partial object retrieval of the shape retrieval contest (SHREC) 2013, as well as on the publicly available Hampson pottery dataset, support improved performance of the proposed method against seven recently evaluated retrieval methods.*

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications—I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—

---

## 1. Introduction

The main challenge inherent in partial 3D object retrieval relates to the difficulty in effectively quantifying the similarity between a partial query and a complete 3D model. Most partial 3D object retrieval methods rely on local shape descriptors calculated over feature points, either dense, or extracted by means of a salient point detector. Such an approach is suited to partial retrieval, considering that a partial query and its originating complete model are intuitively expected to be identical in a local fashion. Local shape descriptors, apart from being used for the estimation of local shape similarity, can also be employed in an orderless fashion within the context of the bag of visual words (BoVW) paradigm, so as to derive global shape signatures. BoVW methods achieve state of the art performance in 3D object retrieval, with several major works appearing recently [CDF\*04], [OD11], [STP13].

Fisher encoding [PD07] improves over the retrieval performance of standard BoVW, by means of difference encoding and subtracting the mean of a Gaussian fit to all observations. The resulting measures comprise the Fisher vector, which facilitates the assessment of global shape similarity by

means of standard distance measures. This encoding can be computed from much smaller vocabularies at a lower computational cost [TGSS14]. It has also been supported in a recent comparative study [CLVZ11], when compared to the basic  $k$ -means/vector quantization or the support vector encoding [ZYZH10].

The proposed partial 3D object retrieval method, which can be applied on both point clouds and meshes, is based on a hybrid shape matching scheme, defined so as to account for both local and global shape similarity, as well as to address the partiality of the query object. We introduce the differential fast point feature histogram (dFPFH), which extends the well-known FPFH descriptor [RBB09] in order to more accurately capture local geometry transitions. Local shape similarity is quantified by averaging the minimum weighted distances associated with pairs of dFPFH values calculated on the partial query and the target object. This strategy aims to discard dissimilar dFPFH pairs, which can intuitively be attributed to parts of the target object which are missing from the partial query. On the other hand, global shape similarity is derived by means of a weighted distance of Fisher vectors. Weighting of both local and global distances is defined so as

to reduce the influence of the most dissimilar pairs, following once more the aforementioned intuition. Overall, local and global distances, which are derived for multiple scales, are being combined to obtain a ranked list of the most similar complete 3D objects.

Experimental evaluation on the large-scale benchmark dataset for partial object retrieval of the shape retrieval contest (SHREC) 2013 [SMB\*13] support the proposed method against five recently evaluated partial 3D object retrieval methods [SMB\*14], with respect to standard retrieval performance measures. Additional experimentation on the publicly available Hampson pottery dataset provides a real-world application scenario in the cultural heritage (CH) domain, along with extra favourable comparisons with two recent partial 3D object retrieval applications [SPS14], [SPK\*14], that have been evaluated on this dataset.

## 2. Related Work

This section provides an overview of local shape descriptors, as well as of state-of-the-art in partial 3D object retrieval.

### 2.1. Local shape descriptors

The spin-image [JH99], is the first major local shape descriptor. A spin-image of an oriented point is a 2D representation of its surrounding surface, which is constructed on a pose-invariant 2D coordinate system by accumulating the coordinates of neighbouring points. Normal aligned radial features (NARF) [SGVB09] combine an interest point extraction method, along with a feature descriptor in 3D range data. Kernel descriptors [BRD11] provide a principled way to turn pixel attributes to patch-level features and are able to generate rich features from various recognition cues. Besides using gradient and local binary patterns in their framework, the authors developed three more depth kernel descriptors, namely size, PCA and spin. Point feature histograms (PFH) [RBMB08], are directly applicable on point clouds, avoiding the need for mesh generation. PFH and its more efficient sibling, fast PFH (FPFH) [RBB09], encode patterns of point distances within a neighbourhood.

### 2.2. Partial 3D object retrieval

Most partial 3D object retrieval methods can be roughly classified as: (i) view-based, with prominent examples in [SMM\*10], [DA09] and [LMM13], (ii) part-based [TVD09], [APP\*10], (iii) BoVW-based [BBGO11], [Lav12], [LGJ14], and finally (iv) hybrid methods combining these three main paradigms [SPK\*14], [FO09]. Five recent partial 3D object retrieval methods, encompassing elements of these categories, were recently compared on the large-scale benchmark dataset of SHREC 2013:

- two methods for range scan-based 3D model retrieval by incorporating 2D-3D alignment [LJ], as well as by entropy-based adaptive view clustering [LLJ13]. We refer to these methods as ‘SBR2D-3D’ and ‘SBR-VC’, respectively,
- two methods using data-aware partitioning [SBS13] and BoVW [SMB\*14]. We refer to these methods as ‘Data-aware’ and ‘S-BoVW’, respectively,
- a method proposed in [SMB\*13], which uses spin images and signature quadratic form distance. We refer to this method as ‘SQFD’.

In addition, two recent partial 3D object retrieval methods have been applied on the publicly available Hampson pottery dataset:

- a panoramic, view-based method, proposed in [SPK\*14]. We refer to this method as ‘Panoramic’,
- a method proposed in [SPS14], which addresses partial retrieval by means of Fisher encoding in a purely global fashion. We refer to this method as ‘Global Fisher’.

## 3. Shape Representation

In its original form, PFH is computed as follows: i) for each point  $p$ , all of its neighbours enclosed in the sphere of a given radius  $r$  are selected ( $r$ -neighbourhood), ii) for every pair of points  $p_i$  and  $p_j$  ( $i \neq j$ ) in the  $r$ -neighborhood of  $p$  and their PCA-estimated normals  $n_i$  and  $n_j$  ( $p_i$  being the point with a smaller angle between its associated normal and the line connecting the points) [Rus09], a Darboux  $u\nu n$  frame ( $u = n_i$ ,  $\nu = (p_i - p_j) \times u$ ,  $n = u \times \nu$ ) is defined and the angular variations of  $n_i$  and  $n_j$  are computed as follows:  $\alpha = u \cdot n_j$ ,  $\phi = u \cdot (p_j - p_i) / \|p_j - p_i\|$ ,  $\theta = \arctan(w \cdot n_j, u \cdot n_j)$ . The histograms which constitute the PFH descriptor have  $b$  binning subdivisions for each one of  $\alpha$ ,  $\phi$  and  $\theta$  angle, where  $b$  is implementation-dependent. This leads to a histogram size equal to  $3b$ .

Fast point feature histogram (FPFH) [RBB09] has been proposed in order to accelerate PFH computations by employing a subset of neighbouring points for histogram calculation. For a given query point  $p_q$ , its single point feature histogram (SPFH) values are first estimated by creating pairs between itself and its  $r$ -neighbours. This is repeated for all points in the dataset, followed by re-weighting of the SPFH values using the SPFH values of  $r$ -neighbours, in order to create the FPFH for  $p_q$ .

We extend FPFH in order to capture local geometric transitions by measuring the differences in feature histograms, associated with concentric spheres. Our FPFH extension, namely differential FPFH (dFPFH) has  $6b$  bins (the standard  $3b$  bins associated with a concentric sphere of radius  $r$  plus the histogram of  $3b$  bins, which quantifies the transitions of FPFH in a local ribbon around  $r$ ):

$$dFPFH(q_p, r) = [FPFH(q_p, r) \Delta FPFH(q_p, r)] \quad (1)$$

where  $\Delta\text{FPFH}(q_p, r) = \text{FPFH}(q_p, r_{\text{outer}}) - \text{FPFH}(r_{\text{inner}})$ . Figure 1 provides an intuitive explanation of dFPFH. In the case of the smooth surface of the vessel illustrated in Fig. 1a, the FPFH histograms of the two concentric spheres are rather similar, resulting in histogram differences approximating zero. On the other hand, the irregularity of the vessel surface in Fig. 1b is reflected in much larger differences of the FPFH histograms.

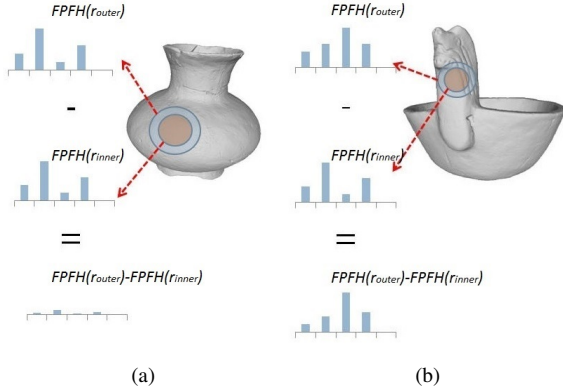


Figure 1: A schematic representation of dFPFH: (a) smooth surfaces result in similar FPFH histograms for the concentric spheres ( $\text{FPFH}(r_{\text{outer}}) \approx \text{FPFH}(r_{\text{inner}})$ ) and histogram differences approximating zero, (b) irregular surfaces result in much larger differences of the FPFH histograms.

Each object may well be scanned using various types of scanning equipment, from varying distances or with different settings, resulting in varying point cloud densities. Aiming to alleviate the effects of this variability on retrieval performance, we introduce two extra filtering steps: (i) the input object is downsampled with voxelized grid filtering, in which all points within a voxel are approximated by their centroid. Multiscale information can be derived by considering multiple voxel sizes, (ii) the neighbourhood radius  $r$  considered in dFPFH calculations is adaptively estimated for each point cloud as a linear function of the mean point distance over all  $r$ -neighbourhoods.

## 4. Shape Matching

For shape matching, we incorporate the result of two parallel processes: (i) local shape similarity assessment by averaging the minimum weighted distances associated with pairs of dFPFH values calculated on the partial query and the target object, (ii) global shape similarity assessment by means of a weighted distance of Fisher vectors.

### 4.1. Local shape similarity assessment

Aiming to assess the local similarity between the partial query object  $Q$  and each complete object  $T$  from the repository, we define the mean-minimum distance  $dm^2$  as follows:

$$dm^2(Q, T) = \text{mean}_{q_p \in Q}(\min_{t_p \in T}(L_{d1}(q_p, t_p))) \quad (2)$$

where  $q_p$  is a point of  $Q$ ,  $t_p$  is a point of  $T$ ,  $N$  and  $M$  denote the number of points of  $Q$  and  $T$ , respectively, whereas  $L_{d1}$  is the Manhattan distance  $L_1$  of the dFPFH histograms of  $q_p$  and  $t_p$ . This strategy is justified by considering that the similarity of the partial query  $Q$  with  $T$ , is not associated with the distance of histograms of all possible pairs of points ( $q_p, t_p$ ), but only with the distance of pairs of histograms of similar points. The average of this distance forms  $dm^2$ . We selected  $L_1$  over other distance alternatives (e.g.  $L_2$ ) based on experimentation.

In addition, considering that in Eq. 2 the minimum of  $L_{d1}$  for each  $q_p$  depends on a single pair of points, we introduce the weighted mean-minimum distance  $dm_w^2$ , in which  $L_{d1}$  is replaced by a weighted average of the  $k$  smaller distances:

$$dm_w^2(Q, T) = \text{mean}_{q_p \in Q}[(1/k) \sum_{i=1,2,\dots,k} w_i L_{d1}(q_p, t_p(i))] \quad (3)$$

where  $t_p(i)$ ,  $i = 1, 2, \dots, k$  are the first  $k$  points of object  $T$ , when all points of  $T$  are sorted in increasing order with respect to their distance from  $q_p$ . The weights  $w_i = (1 - (i/k))$  are linearly decreasing, starting from the pair with the smaller distance ( $i=1$ ). This weighting amplifies the influence of the more similar pairs of points, among the selected  $k$  pairs, whereas it ensures a smooth transition to zero, which is the weight associated to those points which are not among the  $k$  selected. We selected linearly decreasing weighting over other alternatives (e.g. quadratic decrease), since it is associated with less calculations and leads to comparable results, as found in our preliminary experimentation.

### 4.2. Global shape similarity assessment

Aiming to assess the global similarity between the partial query object  $Q$  and each object  $T$  from the repository, we employ Fisher encoding, extending the purely global Fisher approach that has been proposed in [SPS14]. The use of Fisher encoding instead of standard BoVW approaches has been experimentally supported in a recent comparative study [CLVZ11]. A Gaussian mixture model (GMM) is estimated from local shape descriptors by means of an expectation maximization algorithm. The resulting GMM defines the visual codebook used [PD07], [SPMV13].

Given a set of  $N$  dFPFH descriptors  $\mathbf{x}_1, \dots, \mathbf{x}_N \in R^D$ , which are used for training, a GMM  $p(\mathbf{x}|\theta)$  is the probability density on  $R^D$  given by

$$p(\mathbf{x}|\theta) = \sum_{k=1}^K p(\mathbf{x}|\mu_k, \Sigma_k) \pi_k \quad (4)$$

$$p(\mathbf{x}|\mu_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^D \det \Sigma_k}} e^{-\frac{1}{2}(\mathbf{x}-\mu_k)^T \Sigma_k^{-1} (\mathbf{x}-\mu_k)} \quad (5)$$

where  $K$  is the number of Gaussian components used,  $\theta = (\pi_1, \mu_1, \Sigma_1, \dots, \pi_K, \mu_K, \Sigma_K)$  is the vector of model parameters, including the prior probability values  $\pi_k \in R_+$  (which sum to one), the means  $\mu_k \in R^D$ , and the positive definite covariance matrices  $\Sigma_k \in R^{D \times D}$  of each Gaussian component. The covariance matrices are assumed to be diagonal, so that the GMM is fully specified by  $(2D+1)K$  scalar parameters. Soft data-to-cluster assignments are defined as

$$q_{ki} = \frac{p(\mathbf{x}_i|\mu_k, \Sigma_k)\pi_k}{\sum_{j=1}^K p(\mathbf{x}_i|\mu_j, \Sigma_j)\pi_j}, k = 1, \dots, K \quad (6)$$

Fisher encoding [SPMV13] captures the average first and second order differences between the local descriptors and the centres of a GMM, which can be thought of as a soft visual codebook. For the  $k$ -th GMM, where  $k = 1, \dots, K$ , the following vectors are defined

$$\mathbf{u}_k = \frac{1}{N\sqrt{\pi_k}} \sum_{i=1}^N q_{ik} \Sigma_k^{-1/2} (\mathbf{x}_i - \mu_k) \quad (7)$$

$$\mathbf{v}_k = \frac{1}{N\sqrt{2\pi_k}} \sum_{i=1}^N q_{ik} [(\mathbf{x}_i - \mu_k) \Sigma_k^{-1} (\mathbf{x}_i - \mu_k) - 1] \quad (8)$$

The Fisher encoding of the set of local feature vectors is then given by the concatenation of

$$\mathbf{f} = [\mathbf{u}_1^T, \mathbf{v}_1^T, \dots, \mathbf{u}_K^T, \mathbf{v}_K^T] \quad (9)$$

Considering that both vectors,  $\mathbf{u}_k$  and  $\mathbf{v}_k$ , have size equal to the size of the local feature vector, i.e.  $6b = 66$ , considering that  $b = 11$  binning subdivisions are used, it can be derived from Eq. (6) that the resulting Fisher vector  $\mathbf{f}$  has size equal to  $2 \times 66 \times K = 132 \times K$ .

Considering that the originating complete object of  $Q$  and its most similar complete object  $T$  are intuitively expected to have similar Fisher vectors, it is natural to assume that the most dissimilar pairs of Fisher components between  $Q$  and  $T$ , are associated with the GMMs that are over-represented in those parts of  $T$  that are missing from  $Q$ . Starting from this consideration, we define the weighted Fisher vector distance  $dF_w^2$ , in a similar fashion to  $dm_w^2$ , as:

$$dF_w^2(Q, T) = (1/K) \sum_{j=1,2,\dots,K} w_{fj} L_{f1}(Q(j), T(j)) \quad (10)$$

where  $L_{f1}(Q(j), T(j))$  is the  $L_1$  distance of the respective

Fisher vectors  $L_1(\mathbf{f}_Q(j), \mathbf{f}_T(j))$ . The pairs  $(\mathbf{f}_Q(j), \mathbf{f}_T(j))$  are sorted in increasing order with respect to their distance. The weights  $w_{fj} = (1 - (j/K))$  are linearly decreasing, starting from the pair with the smaller distance. This weighting reduces the influence of the more dissimilar pairs of Fisher components in the distance calculation. As is the case with Eq. 2 and Eq. 3, the utilization in Eq. 10 of both  $L_1$  and linearly decreasing weighting is supported by preliminary experimentation.

### 4.3. Hybrid shape similarity assessment

For each voxel size  $v_s$  considered in the voxelized gridding filtering step described in Section 3 (where  $s = 1, 2, \dots, S$ , with  $S$  the number of voxel sizes considered), the hybrid distance  $d_{hybrid}(Q, T, s)$  is a weighted sum of  $dm_w^2(s)$  and  $dF_w^2(s)$ , defined according to Eq. 2 and Eq. 10 by substituting  $dm_w^2$  and  $dF_w^2$  with  $dm_w^2(s)$  and  $dF_w^2(s)$ , respectively:

$$d_{hybrid}^2(Q, T, s) = w_o dm_w^2(s) + dF_w^2(s) \quad (11)$$

where  $w_o$  adjusts the relative influence of local and global shape matching distances. The overall multiscale distance  $d_{multiscale}$ , which is used to obtain a ranked list of complete 3D objects, is a weighted sum:

$$d_{multiscale}^2(Q, T) = \sum_{s=1,2,\dots,S} w_s d_{hybrid}^2(Q, T, s) \quad (12)$$

where the weights  $w_s$  adjust the relative influence of each scale  $s$  considered.

Figure 2 shows the distinct components of the proposed pipeline for partial 3D object retrieval.

## 5. Evaluation

Experiments are performed on two publicly available benchmark datasets. The first dataset has been used in SHREC 2013 track for large scale partial 3D object retrieval [SMB\*13]. The target set has been created from 360 shapes, organized into 20 classes of 18 objects per class. On the other hand, the process of range scan acquisition from the objects of the target set has been simulated in order to obtain a set of partial views. This process results in 7200 queries, associated with varying levels of partiality. Figure 3 shows some samples from the target set of SHREC 2013. Recently, a more extensive comparison of five state-of-the-art methods has been performed on the same dataset [SMB\*14].

The second benchmark dataset used for evaluation is related to the CH domain and consists of 3D pottery models originating from the Virtual Hampson Museum collection (<http://hampson.cast.uark.edu>). It is publicly available and has already been used for the evaluation of two state-of-the-art methods [SPS14], [SPK\*14]. The dataset consists

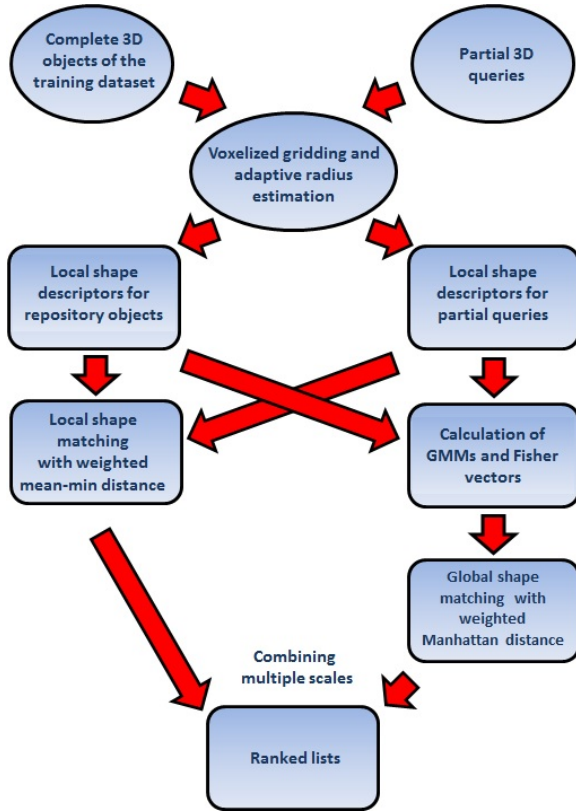


Figure 2: The pipeline of the proposed method.

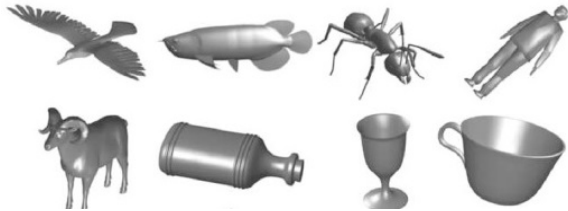


Figure 3: Samples of the SHREC 2013 benchmark dataset [SMB\*14].

of 384 models classified to 23 distinct geometrically defined classes. 21 partial queries have been artificially created by slicing and cap filling complete 3D models. The partial queries used in our experiments have a reduced surface compared to the original 3D object, which is associated with 25% partiality. Figure 4 shows some examples of pottery models used in this dataset.

Experimental evaluation is based on precision-recall (P-R) plots and five quantitative measures: nearest neighbour (NN), first tier (FT), second tier (ST), discounted cumulative gain (DCG) and mean average precision (MAP). More details on these measures can be found in [SMB\*13], [SPK\*14].

Figure 4: Example 3D models of the pottery dataset used (<http://hampson.cast.uark.edu>).

The proposed method has been developed on a hybrid Matlab/C++ architecture. The experiments have been performed on an Intel Core i7 workstation, operating at 3.5 GHz with 16 GB of RAM.

Parameter settings have been experimentally determined as follows: the linear coefficients adaptively associating the radii of the concentric spheres of dFPFH to the mean point distance are  $r = 2.7$ ,  $r_{outer} = 13.6$  and  $r_{inner} = 13.1$ , respectively (Eq. 1). In addition,  $k = 3$  (Eq. 3) and  $K = 10$  GMMs were found to be sufficient for the construction of the visual codebook, leading to Fisher vectors of  $2 \times 66 \times 10 = 1320$  components.  $k$ -means pre-clustering by means of Lloyd's variant [Llo82] has been used to initialize GMM construction. The signed square root function has been applied to the resulting Fisher vectors, followed by  $L_2$  normalization. The weight  $w_o$  (Eq. 11) has been set to 0.4. Finally, the number of scales considered is  $S = 3$ , with respective voxel sizes equal to 0.1, 0.3 and 0.5 and respective weights  $w_s = 0.4, 1.0, 0.4$  (Eq. 12).

Interestingly, it has been observed that by separately employing local and global shape similarity, the retrieval performance in SHREC 2013 is significantly lower (FT approximately equal to 15% and 18%, respectively) than the one obtained by the proposed hybrid approach (FT 28%), verifying that complementary information is derived from these parallel processes. Moreover, weighting as defined in Eq. 3 and Eq. 10 allows a performance boost of approximately 1.5-2.5% with respect to FT, when compared with uniformly weighted  $L_1$ -based retrieval. Finally, our preliminary experimentation showed that for values of  $w_o$  and  $w_s$  in the range [0.1,1.0], the retrieval performance may change up to approximately 2%, with respect to FT.

Table 1 presents the retrieval performance, as quantified by NN, FT, ST and MAP, which was obtained by the proposed method and five state-of-the-art methods on SHREC 2013 benchmark dataset. It can be observed that the proposed method achieves the highest performance with respect to all metrics. Figure 5 illustrates the average P-R scores for all retrieval methods. It should be noted that the results presented for the state-of-the-art methods are the ones presented in [SMB\*14].

Table 2 presents the retrieval performance, as quanti-

Method	NN	FT	ST	MAP
Proposed method	<b>0.3856</b>	<b>0.2772</b>	<b>0.2135</b>	<b>0.2851</b>
SBR-2D-3D	0.3535	0.2290	0.1808	0.2455
SBR-VC	0.3218	0.2065	0.1638	0.2199
Data-aware	0.3457	0.2495	0.2088	0.2836
Polar spin images	0.0931	0.0809	0.0768	0.0968
SQFD	0.3108	0.2043	0.1576	0.1978

Table 1: The results of the proposed method, along with 5 state-of-the-art methods on SHREC 2013 benchmark dataset.

fied by NN, FT, ST and DCG, which was obtained by the proposed method and two state-of-the-art methods on the Hampson pottery dataset. It should be noted that in this case we use DCG instead of MAP, since this measure was used for the evaluation of the Panoramic [SPK\*14] and Global Fisher [SPS14] methods. In addition, an accuracy of three decimal digits is maintained, as in these works. Finally, we present results obtained on queries associated with 25% partiality. The proposed method achieves the highest retrieval performance with respect to all measures considered. This is verified in Fig. 6, which illustrates the average P-R scores for all retrieval methods.

Method	NN	FT	ST	DCG
Proposed method	<b>0.952</b>	<b>0.460</b>	<b>0.642</b>	<b>0.778</b>
Global Fisher	<b>0.952</b>	0.320	0.461	0.694
Panoramic	0.619	0.416	0.626	0.721

Table 2: The results of the proposed method, along with two state-of-the-art methods on the Hampson pottery dataset.

Figure 7 illustrates example ranked lists obtained in the case of the Hampson pottery dataset.

It should be stressed that unlike the methods applied in SHREC 2013 and the Panoramic-based method, which are mesh-based, the proposed method requires only raw point cloud information.

## 6. Conclusions

This work presents a partial 3D object retrieval method, applicable on both meshes and point clouds, which is based on a hybrid shape matching scheme, incorporating both local and global shape similarity for multiple scales. The main contributions of the proposed methodology involve both local shape descriptor, as well as partial retrieval aspects:

- the definition of dFPFH, which extends the well-known FPFH descriptor, in order to capture local geometric transitions,



Figure 7: Example ranked lists obtained by the proposed 3D object retrieval method in the case of the Hampson pottery dataset. Examples of partial queries are shown in the upper row, whereas the respective top-6 objects retrieved are shown below.

- the use of a hybrid shape matching scheme, which incorporates local information directly derived from local shape descriptors, as well as global shape information derived from Fisher vectors,
- the definition of a weighted mean-minimum distance, as well as of a weighted Fisher vector distance, both addressing the partiality of the 3D object query.

The experimental evaluation of the proposed method on the large-scale benchmark dataset for partial object retrieval of SHREC 2013, as well as on the Hampson pottery dataset leads to the following conclusions:

- local and global shape similarity information, derived in multiple scales, act in a complementary fashion, maximizing the achieved retrieval performance when combined,
- the proposed partial 3D object retrieval method achieves state-of-the-art performance, when applied in SHREC 2013 dataset,
- the proposed method outperforms two recent CH applications of partial 3D object retrieval, when applied in the Hampson pottery dataset.

Although the proposed method achieves state-of-the-art retrieval performance, the results remain far from perfect

in absolute numbers. In this respect, the remark in the comparative study presented in [SMB\*14], that the problem of partial 3D object retrieval is very challenging and open to future solutions, is still valid. Hybrid retrieval methods combining multiple techniques for shape similarity assessment provide a promising direction to the partial 3D object retrieval problem.

### Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 600533 PRE-SIOUS.

### References

- [APP\*10] AGATHOS A., PRATIKAKIS I., PAPADAKIS P., PERANTONIS S., AZARIADIS P., SAPIDIS S.: 3D articulated object retrieval using a graph-based representation. *The Visual Computer* 26, 10 (2010), 1301–1319. 2
- [BBGO11] BRONSTEIN A., BRONSTEIN M., GUIBAS L., OVSJANIKOV M.: Shape google: geometric words and expressions for invariant shape retrieval. *ACM Transactions on Graphics* 30, 1 (2011), 1–20. 2
- [BRD11] BO L., REN X., D.FOX: Depth kernel descriptors for object recognition. In *Proc. IROS* (2011), pp. 821–826. 2
- [CDF\*04] CSURKA G., DANCE C. R., FAN L., WILLAMOWSKI J., BRAY C.: Visual categorization with bags of keypoints. In *Proc. Workshop on Statistical Learning in Computer Vision, ECCV* (2004), pp. 1–22. 1
- [CLVZ11] CHATFIELD K., LEMPITSKY V., VEDALDI A., ZISSERMAN A.: The devil is in the details: an evaluation of recent feature encoding methods. In *Proc. BMVC* (2011), pp. 1–12. 1, 3
- [DA09] DARAS P., AXENOPOULOS A.: A 3D shape retrieval framework supporting multimodal queries. *International Journal of Computer Vision* 89 (2009), 229–247. 2
- [FO09] FURUYA T., OHBUCHI R.: Dense sampling and fast encoding for 3D model retrieval using bag-of-visual features. In *Proc ACM ICIVR* (2009). 2
- [JH99] JOHNSON A., HEBERT M.: Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 5 (1999), 433–449. 2
- [Lav12] LAVOUÉ G.: Combination of bag-of-words descriptors for robust partial shape retrieval. *The Visual Computer* 28, 9 (2012), 931–942. 2
- [LGJ14] LI B., GODIL A., JOHAN H.: Hybrid shape descriptor and meta similarity generation for non-rigid and partial 3D model retrieval. *Multimedia Tools Appl.* 72, 2 (2014), 1531–1560. 2
- [LJ] LI B., JOHAN H.: Sketch-based 3D model retrieval by incorporating 2D-3D alignment. *Multimedia Tools and Applications* 65, 3, 363–385. 2
- [LLJ13] LI B., LU Y., JOHAN H.: Sketch-based 3D model retrieval by viewpoint entropy-based adaptive view clustering. In *Proc. 3DOR* (2013), pp. 49–56. 2
- [Llo82] LLOYD S.: Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28, 2 (1982), 129–136. 5
- [LMM13] LI P., MA H., MING A.: Combining topological and view-based features for 3D model retrieval. *Multimedia Tools and Applications* 65 (2013), 335–361. 2
- [OD11] O'HARA S., DRAPER B.: Introduction to the bag of features paradigm for image classification and retrieval. *Computing Research Repository* (2011). 1
- [PD07] PERRONIN F., DANCE C.: Fisher kernels on visual vocabularies for image categorization. In *Proc. CVPR* (2007). 1, 3
- [RBB09] RUSU R. B., BLODOW N., BEETZ M.: Fast point feature histograms (FPFH) for 3D registration. In *Proc. ICRA* (2009), pp. 3212–3217. 1, 2
- [RBMB08] RUSU R. B., BLODOW N., MARTON Z., BEETZ M.: Aligning point cloud views using persistent feature histograms. In *Proc. IROS* (2008), pp. 3384–3391. 2
- [Rus09] RUSU R. B.: *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments*. PhD thesis, Computer Science Department, Technische Universitaet Muenchen, Germany, 2009. 2
- [SBS13] SIPIRAN I., BUSTOS B., SCHRECK T.: Data-aware 3D partitioning for generic shape retrieval. *Computer Graphics* 37, 5 (2013), 460–472. 2
- [SGVB09] STEDER B., GRISETTI G., VAN M., BURGARD L. W.: Robust online model-based object detection from range images. In *Proc. IROS* (2009), pp. 4739–4744. 2
- [SMB\*13] SIPIRAN I., MERUANE R., BUSTOS B., SCHRECK T., JOHAN H., LI B., LU Y.: SHREC'13 track: Large-scale partial shape retrieval using simulated range images. In *Proc. 3DOR* (2013). 2, 4, 5
- [SMB\*14] SIPIRAN I., MEROUANE R., BUSTOS B., SCHRECK T., LI B., LU Y., JOHAN H.: A benchmark of simulated range images for partial shape retrieval. *The Visual Computer* (2014). 2, 4, 5, 7
- [SMM\*10] STAVROPOULOS G., MOSCHONAS P., MOUSTAKAS K., TZOVARAS D., STRINTZIS M.: 3D model search and retrieval from range images using salient features. *IEEE Transactions on Multimedia* 12, 7 (2010), 692–704. 2
- [SPK\*14] SFIKAS K., PRATIKAKIS I., KOUTSOUDIS A., SAVELONAS M., THEOHARIS T.: Partial matching of 3D cultural heritage objects using panoramic views. *Multimedia Tools and Applications* (2014). 2, 4, 5, 6
- [SPMV13] SÁNCHEZ J., PERRONIN F., MENSINK T., VERBEEK J.: Image classification with the fisher vector: Theory and practice. *International Journal of Computer Vision* 105, 3 (2013), 222–245. 3, 4
- [SPS14] SAVELONAS M., PRATIKAKIS I., SFIKAS K.: Fisher encoding of adaptive fast persistent feature histograms for partial retrieval of 3D pottery objects. In *Proc. 3DOR* (2014), pp. 61–68. 2, 3, 4, 6
- [STP13] SFIKAS K., THEOHARIS T., PRATIKAKIS I.: 3D object retrieval via range image queries in a bag-of-visual-words context. *The Visual Computer* 29 (2013), 1351–1361. 1
- [TGSS14] TAO R., GAVVES E., SNOEK C. G. M., SMEULDERS A. W. M.: Locality in generic instance search from one example. In *Proc. CVPR* (2014), pp. 2099–2106. 1
- [TVD09] TIERNY J., VANDEBORRE J., DAOUDI M.: Partial 3D shape retrieval by reeb pattern unfolding. *Computer Graphics Forum* 28, 1 (2009), 41–55. 2
- [ZYZH10] ZHOU X., YU K., ZHANG T., HUANG T.: Image classification using super-vector coding of local image descriptors. In *Proc. ECCV* (2010). 1

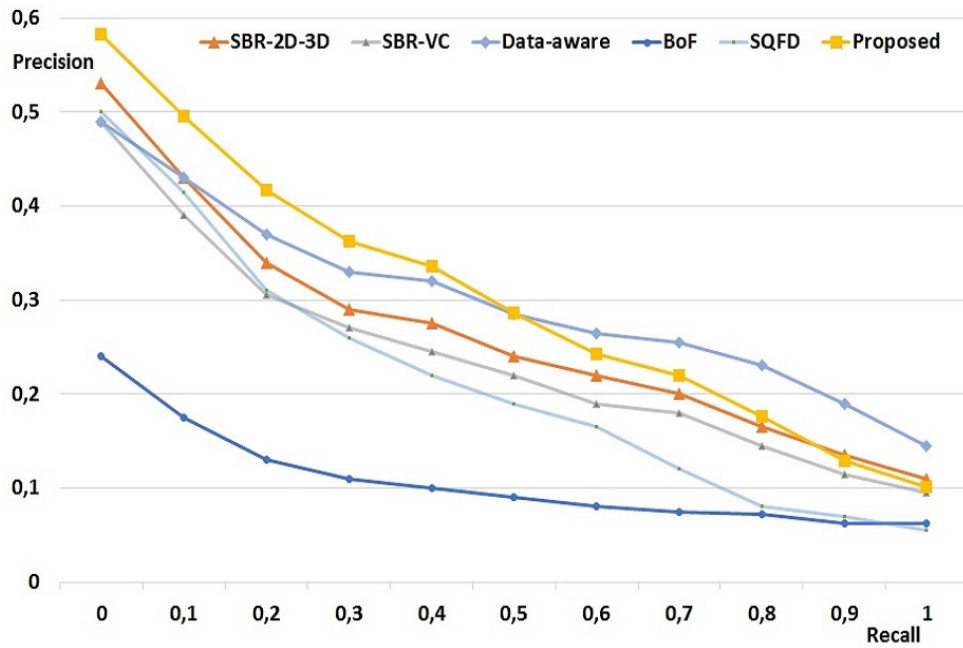


Figure 5: Average P-R for all retrieval methods applied on SHREC 2013.

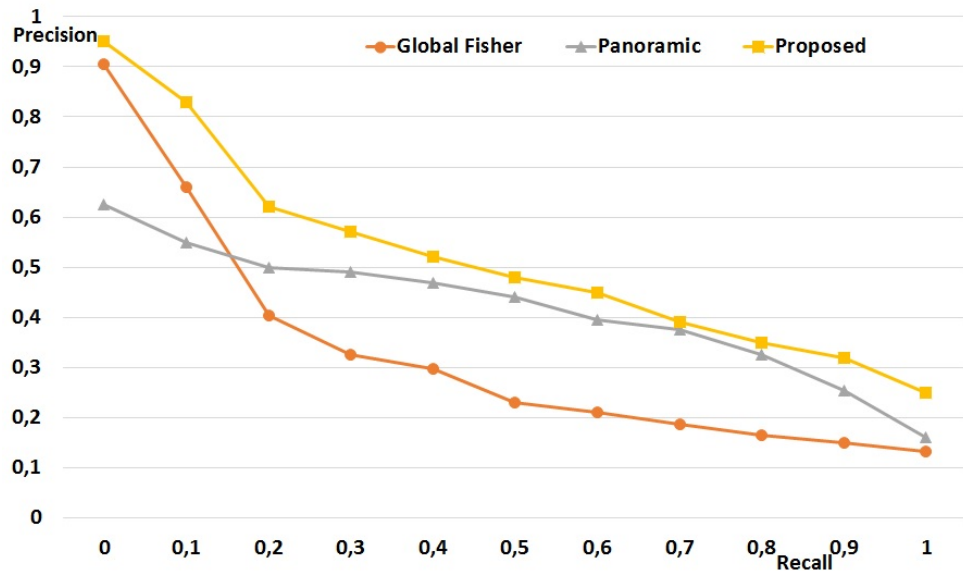


Figure 6: Average P-R for all retrieval methods applied on the Hampson pottery dataset, for queries of 25% partiality.