Eurographics Workshop on 3D Object Retrieval (2016)
A. Ferreira, A. Giachetti, and D. Giorgi (Editors)

An evaluation of local feature encodings for shape retrieval

F. P. Tasse and J. Kosinka† and N. A. Dodgson‡

The Computer Laboratory, University of Cambridge, United Kingdom

Abstract

Local features are successfully used in 3D shape retrieval by encoding features descriptors into global shape signatures. Previous 3D retrieval systems use different encoding methods, such as histogram encoding and Fisher encodings, making it difficult to evaluate one encoding technique against another. We perform a comparative analysis of four recent encoding methods when used in shape retrieval. The analysis shows that Vector of Locally Aggregated Descriptors (VLAD) encoding is the best method of the four tested, since it offers the best trade-off between precision and computational cost.

1. Introduction

Bag-of-features shape retrieval typically consists of three steps: *detection* of local features, their *encoding* into a global descriptor, and *comparison* of shapes with a distance metric. We previously considered how saliency-based feature detection, one possible way to do the first step, influences shape retrieval performance [TKD16]. We consider here the second step: how *encoding* affects shape retrieval performance, regardless of how those features are detected.

Histogram encoding of quantized local features is a simple method for encoding local features. Recent shape retrieval systems have used more sophisticated encoding techniques such as Soft Quantization [PCI*08], Fisher vectors [PSM10] and Vector of Locally Aggregated Descriptors (VLAD) [JDSP10]. However, there is no evaluation of how these encoding methods affect retrieval performance. Inspired by an evaluation of encoding techniques in image classification [CLVZ11], we compare encoding methods for shape retrieval. We show that VLAD encoding is the best when both retrieval performance and time complexity are considered.

We provide a fair comparison of these techniques by fixing other factors such as feature detection, local descriptor, and descriptor distance metric. We evaluate retrieval performance on three benchmarks, one of which is based on a challenging benchmark where queries are range scans [GDB*15]. Figure 1 shows top matches given a range scan as query. The illustrated example supports the quantitative results that suggest that for partial shape retrieval, VLAD outperforms other tested encoding methods.

Our contributions are a quantitative analysis of four feature encoding methods, and recommendations on which techniques perform the best given a dataset type.

† Jiří Kosinka is now at the University of Groningen.
‡ Neil Dodgson is now at the Victoria University of Wellington.

2. Related Work

Shape retrieval methods retrieve, from a collection of shapes, 3D models that are most similar to a given 3D query. Encoding of local features is a popular technique borrowed from image and video retrieval [SZ03]. Local features, sparse or dense, are extracted from the whole dataset and each feature is represented with a multi-dimensional descriptor. The resulting local descriptor space is then partitioned into regions. Encoding local features of an input shape typically starts with a quantization of these features using the descriptor space partitioning, followed by combining the quantized vectors to form a global descriptor. Tabia et al. [TLPG14] use Histogram Encoding [SZ03], based on counting quantized features, to encode their novel covariance descriptors. Bronstein et al. [BBGO11] use a more descriptive method, Soft Quantisation [PCI*08], which consists of summing softly-quantised features, to encode sparse HKS-based features.

Another family of encoding techniques records statistics of differences between local features and cluster elements, rather than pooling quantised features. Examples of such approaches are Fisher Vectors [PSM10] and VLAD [JDSP10]. Few 3D retrieval systems use this type of encoding. Savelonas et al. [SPS15] present shape retrieval based on Fisher encoding of novel local descriptors derived from Fast Point Feature Histograms [RBB09]. Guler et al. [GTU14] use VLAD to encode their proposed volumetric features.

After encoding, global descriptors are typically compared using their normalized scalar product (cosine of angle) [TLPG14], or fed into a discriminative classifier such as similarity-sensitive hashing [BBGO11]. A set of encoding methods we do not evaluate here are those based on supervised learning, such as sparse coding [LBBC14]. These methods learn the best discriminative encoding of local descriptors from training examples.

DOI: 10.2312/3dor.20161085

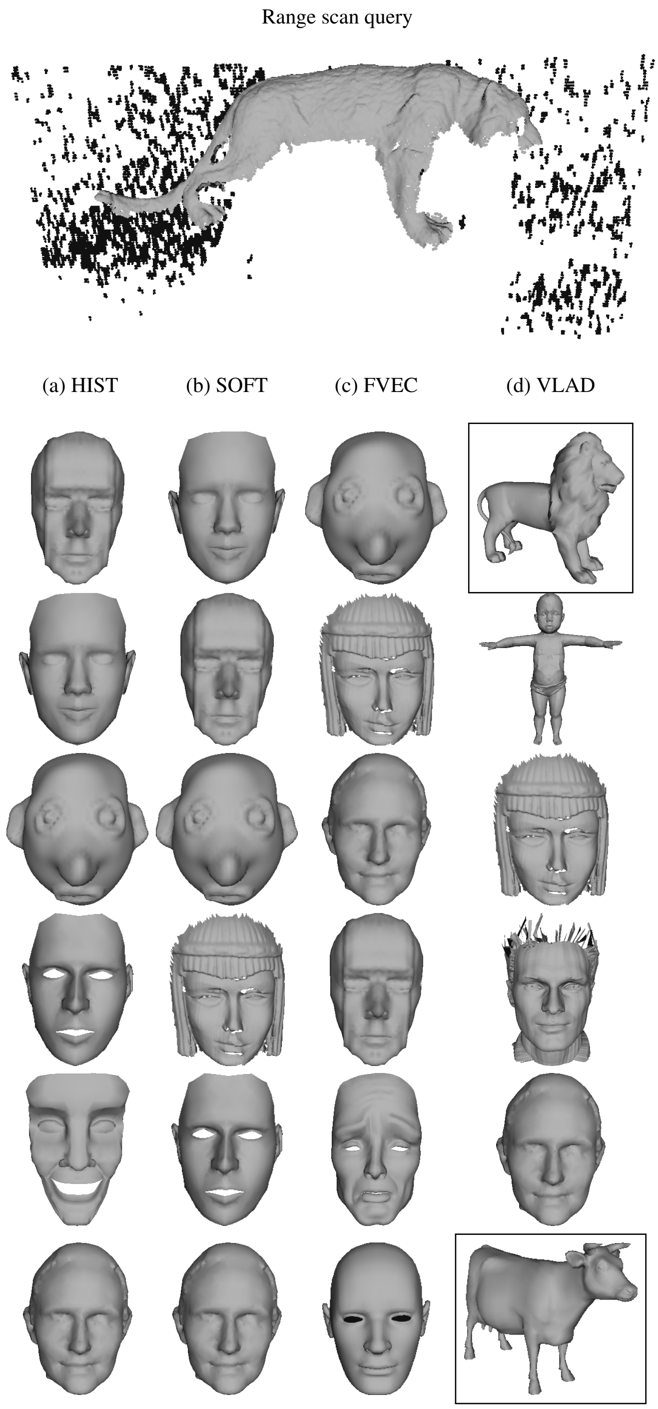

Figure 1: *Top six retrieved models given as query a range scan of lion standing in front of a wall. Retrieval results are shown for the four feature encoding methods: (a) Histogram encoding, (b) Soft Quantization, (c) Fisher Vectors and (d) Vector of Locally Aggregated Descriptors. VLAD-based retrieval produces two correct matches (framed), whereas other methods fail to retrieve relevant shapes. These methods match the query to human faces since, similarly to the query, human faces only contain frontal information. VLAD encoding is less sensitive to missing parts.*

3. Experimental setup

Our focus is on determining the best encoding technique for bag-of-features shape retrieval. For consistency, we use the same datasets and evaluation metrics (outlined below) as in the previous work on effect of salient features on shape retrieval [TKD16]. We evaluate retrieval on three datasets:

- Dataset A comprises 20 classes, each containing 20 different watertight 3D meshes [GBP07].
- Dataset B is the SHREC'15 Non-Rigid Shape Retrieval track [LZC*15] with 1200 models obtained from deforming 60 shapes divided in 50 classes.
- Dataset C is the SHREC'15 Range Scan Shape Retrieval track [GDB*15], comprising 1200 watertight meshes from 60 classes, and 180 query range scans.

We detect local features on a shape by extracting local maxima of a random saliency map, similar to how Tasse et al. [TKD15] extract keypoints from saliency maps. Let RK denote this detection method. Using RK, 248 points, on average, are generated per shape in Dataset A. The choice of a random saliency map is motivated by a recent study of saliency-based features for shape retrieval [TKD16] which shows that keypoints based on random saliency outperform all other tested saliency models, including ground-truth.

We now discuss the other factors that influence shape retrieval: local descriptors and feature encoding methods.

3.1. Local feature descriptors

Local descriptors help describe the local neighbourhood, $\mathcal{N}$, of a point. $\mathcal{N}(p)$ is the set of neighbours q of p with $\|p-q\| < r$, where r is the radius of the neighbourhood often referred to as *support radius*. To support any surface representation, we focus on point-based descriptors. We set $r = R/10$, where R is the underlying shape radius. We use the Point Feature Histogram (PFH) [RBB09], a robust local descriptor. We chose PFH because it has been proven to produce better retrieval performance compared to other popular descriptors such as Fast Point Feature Histograms [TKD16].

3.2. Feature encoding methods

Before describing the four tested encoding methods, we present the partitioning techniques on which they are based. Recall that a common step in most feature encoding methods is clustering the set of all local features extracted from the dataset. This descriptor space partitioning is typically done with a hard or soft clustering approach. The main clustering techniques are:

- **K-means clustering:** A set of $D-$dimensional local feature descriptors $\{\mathbf{x}_1,..,\mathbf{x}_N\}$ is partitioned into K clusters with centers $\{\boldsymbol{\mu}_1,...,\boldsymbol{\mu}_K\}$ such that the within-cluster variance is minimized. K-means is a hard clustering method.
- **K-medians clustering:** An alternative type of hard clustering replaces the L2 norm used to compute within-cluster variance with the L1 norm, which is less sensitive to noise. We use K-medians as the default hard clustering method because it produces better retrieval performance, as shown in Section 5.1.

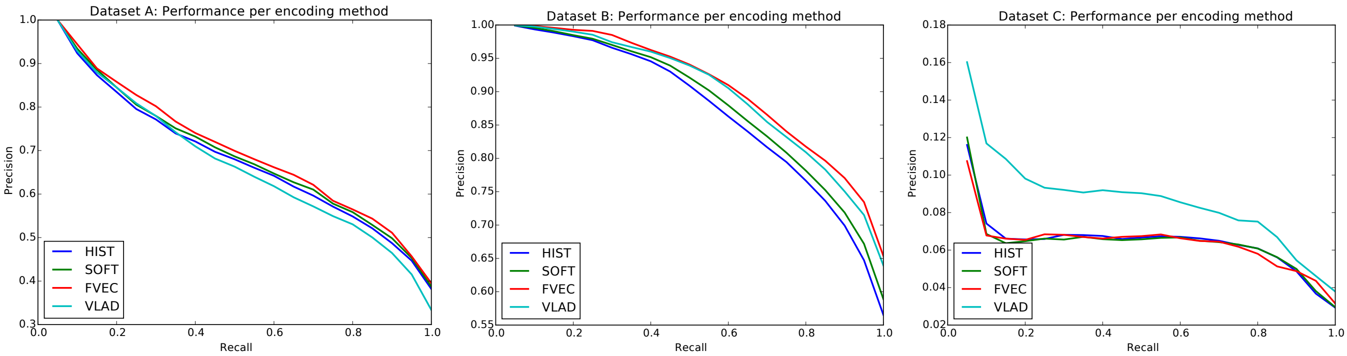

Figure 2: *Precision-recall curves for retrieval performance based on selected encoding methods, using K-medians as the clustering type.*

- **Gaussian Mixture Model (GMM) clustering:** Rather than assigning each descriptor to one cluster, a descriptor belongs to each cluster with a certain probability. The approach assumes that each cluster has a Gaussian distribution, and thus the descriptor space has a mixture distribution parameterized by weights $w_k \in \mathbb{R}_+$, means $\boldsymbol{\mu}_k \in \mathbb{R}^D$ and covariance matrices $\boldsymbol{\Sigma_k}$ [DLR77]. It is commonly assumed that the covariance matrices are diagonal, denoted here by $\boldsymbol{\sigma}_k \in \mathbb{R}^D$.

Feature encoding methods mainly differentiate themselves by the clustering technique, and how local features are transformed into global descriptors. Let $q_k(\mathbf{x}_i)$ be the extent to which a descriptor $\mathbf{x}_i$ belongs to cluster k. In hard clustering, $q_k(\mathbf{x}_i)$ is 0 or 1. In the case of GMM clustering, $q_k(\mathbf{x}_i)$ is commonly referred to as the *posterior probability*.

We evaluate four of the encoding methods mentioned in Section 2 and described below.

3.2.1. Histogram encoding (HIST)

The simplest encoding method is a histogram of quantised local features [SZ03]. Assuming that the descriptor space is partitioned with a hard clustering technique like K-means, and a set of local descriptors $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ extracted from a shape, the shape descriptor is $f_{hist} \in \mathbb{R}^K$ with $[f_{hist}]_k = |\mathbf{x}_i : q_k(\mathbf{x}_i) = 1|$. The hard quantization used in this type of encoding retains little information from the original local features.

3.2.2. Soft Quantization (SOFT)

Rather than reducing each local feature to a single cluster mean, SOFT uses distances to all cluster means to quantize a descriptor [PCI*08]. More specifically, a local descriptor $\mathbf{x}_i$ is quantised to $f_{soft}(\mathbf{x}_i) \in \mathbb{R}^K$ given by

$$[f_{soft}(\mathbf{x}_i)]_k = \frac{g_k(\mathbf{x}_i)}{\sum_{j=1}^K g_j(\mathbf{x}_i)}, \quad g_k(\mathbf{x}_i) = \exp\left(-\frac{\gamma}{2}\|\mathbf{x}_i - \boldsymbol{\mu}_k|\right),$$

where γ controls the importance of distance to a cluster mean and is chosen so that weighting only applies to a small number of clusters. We set $\gamma = 10$ when we evaluate performance of shape retrieval based on SOFT encoding. Given the above quantisation, the global descriptor of a shape with local descriptors $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ is

$$f_{soft} = \sum_{i=1}^n f_{soft}(\mathbf{x}_i).$$

Although SOFT retains more feature information than HIST, it is still based on a hard clustering method.

3.2.3. Fisher Vectors (FVEC)

This method encodes the differences between a set of local descriptors on a shape and GMM clusters of the descriptor space [PSM10]. Given GMM parameters $(w_1,\boldsymbol{\mu}_1,\boldsymbol{\sigma}_1,\ldots,w_K,\boldsymbol{\mu}_K,\boldsymbol{\sigma}_K)$ and n local descriptors extracted from the input shape, D-dimensional vectors $\mathbf{u}_k$ and $\mathbf{v}_k$, representing mean and variance respectively, are computed for each cluster:

$$\begin{aligned}\mathbf{u}_k &= \frac{1}{n\sqrt{w_k}}\sum_{i=1}^n q_k(\mathbf{x}_i)\frac{\mathbf{x}_i-\boldsymbol{\mu}_k}{\boldsymbol{\sigma}_k},\\ \mathbf{v}_k &= \frac{1}{n\sqrt{2w_k}}\sum_{i=1}^n q_k(\mathbf{x}_i)\left[\left(\frac{\mathbf{x}_i-\boldsymbol{\mu}_k}{\boldsymbol{\sigma}_k}\right)^2-1\right].\end{aligned}$$

These means and variances are concatenated to form the final global descriptor $f_{fvec} = [\mathbf{u}_1,\mathbf{v}_1,\ldots,\mathbf{u}_K,\mathbf{v}_K] \in \mathbb{R}^{2DK}$.

3.2.4. Vector of Locally Aggregated Descriptors (VLAD)

VLAD is similar to FVEC in the sense that it also captures the statistics of differences between the descriptor space clusters and local descriptors on a shape, but it uses only cluster means [JDSP10]. VLAD assigns each local descriptor to its closest cluster, and then computes vector differences between local descriptors and their assigned cluster mean. The vector differences are referred to as “residuals”. For a cluster with index k, the sum of D-dimensional residuals associated with it are accumulated as follows:

$$\mathbf{r}_k = \sum_{i=1}^n q_k(\mathbf{x}_i)(\mathbf{x}_i - \boldsymbol{\mu}_k).$$

Concatenating these sums for each cluster produces a KD-dimensional vector $f_{vlad} = [\mathbf{r}_1,\ldots,\mathbf{r}_K]$. We use the VLFeat library [VF10] to compute FVEC and VLAD.

Inspired by prior work on image classification [PSM10], we apply power normalization followed by L2 normalizations to FVEC and VLAD encodings. We also apply *term frequency-inverse document frequency* (*tf-idf*) weighting to HIST and SOFT vectors, to penalise frequently occurring features [SZ03]. We compute the similarity between two shapes using the standard cosine angle between their corresponding encodings.

3.3. Evaluation metrics

Given a benchmark of 3D models divided into classes, a shape retrieved based on a query is relevant if both target and query belong to the same class. This interpretation of relevance is standard in shape retrieval benchmarks [GBP07, GDB*15, LZC*15]. To evaluate retrieval performance, we use four standard metrics: Precision-Recall (PR) curve, First Tier (FT), Second Tier (ST), and normalized Discounted Cumulative Gain (DCG). Each metric is averaged over all queries to produce overall scores and 95% confidence interval means. Finally, we use the Wilcoxon rank-sum test [Wil45], a non-parametric alternative to the two-samples t-test, at a 95% significance level to report statistically significant differences between AP performances of competing methods.

4. Evaluation of feature encoding methods

Figure 2 shows PR curves for retrieval performance based on feature encoding. Tables 1–3 show performance across all metrics per encoding and clustering method. Note that FVEC uses either K-means or K-medians to initialize GMM clustering. This section analyses the effect of local feature encoding.

Dataset A There is no statistically significant difference between encoding methods when using K-medians (see Table 1). This suggests that HIST with K-medians is sufficient to achieve good retrieval performance on this dataset, and thus time and space consuming methods such as FVEC can be avoided. When K-means is used, HIST, FVEC and VLAD outperform SOFT, indicating that the latter is sensitive to outliers in the descriptor space.

Dataset B All pairwise differences between feature encoding methods were statistically significant, with the exception of the top-performing methods FVEC and VLAD (see Table 2).

Dataset C Although the mean *AP* of VLAD-based retrieval is larger than the alternative encoding methods (see Table 3), there are no statistically significant differences. This is due to the poor performance in retrieval performance of range scans in the query set. VLAD PR curve in Figure 2 (right) shows that it has better precision at low recall, compared to other methods. This suggests that VLAD is a better encoding method for datasets of 3D scans.

Summary FVEC has a larger mean performance than both HIST and SOFT. This difference is significant on Dataset B for both K-means and K-medians clustering. On the other hand, FVEC and VLAD achieve similar performance on all datasets, but VLAD-based retrieval has better precision at low recall. Combined with the fact that VLAD produces global descriptors half the size of FVEC descriptors and has a lower computational cost (see Section 6), VLAD is a good choice for feature encoding. We explain the better performance of Fisher and VLAD by the fact that they both encode more information on differences between the distribution of local features on a given shape and the general distribution of these features across the whole dataset.

Table 1: *Dataset A: Performance per encoding method. Parameters: detector=RK ($n = 248 \pm 11$), descriptor=PFH, $K = 100$.*

	FT	ST	DCG	AP
HIST/K-medians	0.61 ± 0.03	0.73 ± 0.03	0.83 ± 0.02	0.68 ± 0.03
HIST/K-means	0.59 ± 0.03	0.71 ± 0.03	0.82 ± 0.02	0.66 ± 0.03
SOFT/K-medians	0.61 ± 0.03	$\mathbf{0.73 \pm 0.03}$	0.84 ± 0.02	0.68 ± 0.03
SOFT/K-means	0.54 ± 0.03	0.67 ± 0.03	0.80 ± 0.02	0.61 ± 0.03
FVEC/K-medians	$\mathbf{0.63 \pm 0.03}$	0.73 ± 0.03	$\mathbf{0.84 \pm 0.02}$	$\mathbf{0.70 \pm 0.03}$
FVEC/K-means	0.62 ± 0.03	0.73 ± 0.03	0.84 ± 0.02	0.69 ± 0.03
VLAD/K-medians	0.59 ± 0.03	0.71 ± 0.03	0.83 ± 0.02	0.66 ± 0.03
VLAD/K-means	0.59 ± 0.03	0.71 ± 0.03	0.82 ± 0.02	0.66 ± 0.03

Table 2: *Dataset B: Performance per encoding method. Parameters: detector=RK ($n = 312 \pm 8$), descriptor=PFH, $K = 100$.*

	FT	ST	DCG	AP
HIST/K-medians	0.81 ± 0.01	0.88 ± 0.01	0.95 ± 0.01	0.86 ± 0.01
HIST/K-means	0.80 ± 0.01	0.87 ± 0.01	0.94 ± 0.01	0.85 ± 0.01
SOFT/K-medians	0.83 ± 0.01	0.89 ± 0.01	0.95 ± 0.01	0.87 ± 0.01
SOFT/K-means	0.65 ± 0.01	0.76 ± 0.01	0.89 ± 0.01	0.72 ± 0.01
FVEC/K-medians	$\mathbf{0.86 \pm 0.01}$	0.91 ± 0.01	0.96 ± 0.00	0.90 ± 0.01
FVEC/K-means	0.85 ± 0.01	0.90 ± 0.01	0.96 ± 0.00	0.89 ± 0.01
VLAD/K-medians	0.85 ± 0.01	0.91 ± 0.01	0.96 ± 0.00	0.89 ± 0.01
VLAD/K-means	0.86 ± 0.01	$\mathbf{0.93 \pm 0.01}$	$\mathbf{0.97 \pm 0.00}$	$\mathbf{0.90 \pm 0.01}$

Table 3: *Dataset C: Performance per encoding method. Parameters: detector=RK ($n = 134 \pm 8$), descriptor=PFH, $K = 100$.*

	FT	ST	DCG	AP
HIST/K-medians	0.04 ± 0.01	0.10 ± 0.03	0.35 ± 0.02	0.06 ± 0.01
HIST/K-means	0.06 ± 0.02	0.11 ± 0.03	0.36 ± 0.02	0.07 ± 0.02
SOFT/K-medians	0.04 ± 0.01	0.10 ± 0.03	0.35 ± 0.02	0.06 ± 0.01
SOFT/K-means	0.05 ± 0.02	0.09 ± 0.03	0.36 ± 0.02	0.06 ± 0.01
FVEC/K-medians	0.05 ± 0.01	0.10 ± 0.03	0.35 ± 0.02	0.06 ± 0.01
FVEC/K-means	0.05 ± 0.02	0.10 ± 0.03	0.35 ± 0.02	0.07 ± 0.01
VLAD/K-medians	0.07 ± 0.02	0.12 ± 0.03	0.38 ± 0.02	$\mathbf{0.09 \pm 0.02}$
VLAD/K-means	$\mathbf{0.08 \pm 0.02}$	$\mathbf{0.14 \pm 0.03}$	$\mathbf{0.38 \pm 0.02}$	0.08 ± 0.02

5. Effect of clustering

One of the key parameters in feature encoding is the clustering step. This section examines the influence of the clustering algorithm and cluster size on retrieval.

5.1. K-means *vs* K-medians clustering

We investigate the difference between K-means and K-medians, two hard clustering methods that may be used in HIST, SOFT and VLAD encoding, or used to initialize GMM clustering for FVEC encoding. To study their effect on retrieval, we evaluate retrieval performance when they are used in each encoding. Tables 1–3 summarize the results. Statistical analysis shows that K-medians produces significantly better retrieval performance than K-means when using SOFT in Datasets A and B. There is no significant influence of clustering method on other encoding approaches, suggesting that SOFT is more sensitive to outliers since the L1 norm used in K-medians is more robust to outliers compared to L2.

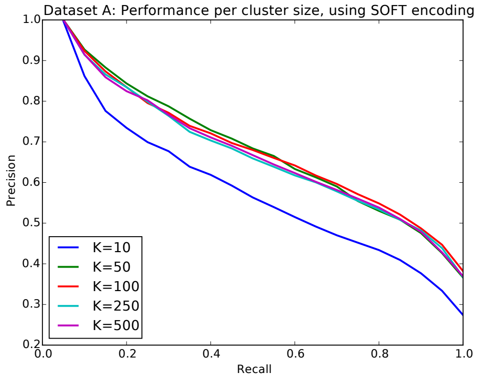

Figure 3: *Performance on Dataset A for various cluster sizes.*

Table 4: *Dataset A: Timings in secs. Parameters: detector = RK, descriptor = PFH, $K = 100$, clustering = K-medians.*

	Descriptors	Clustering	Encoding	Retrieval	Total
HIST	170	21	18	1	210
SOFT	170	21	18	1	210
FVEC	170	41	33	68	306
VLAD	170	21	26	28	244

5.2. Number of clusters

By default we set the cluster size $K = 100$, after exploring a range of values for K. Figure 3 shows retrieval performance for different K when SOFT encoding with K-medians is used. Setting $K = 10$ produces significantly worse performance than larger values of K. This statistically significant difference is not present when FVEC or VLAD are used, and thus their results are not reported here. These encodings are less affected by cluster size, since they retain more information about the original shape feature descriptors.

6. Computational cost

Table 4 shows computation time spent on training a retrieval system (computing descriptors and encodings) and testing it (computing similarities between pairs of shapes in the database). We record computation times for retrieval based on feature encoding method. Results show that FVEC is more computationally expensive than other encoding methods, followed by VLAD.

7. Conclusion

We compared recent feature encoding methods for shape retrieval. We saw that on the tested datasets, SOFT is the most sensitive to the clustering method, in contrast with other encoding techniques. Using K-medians as the default clustering method, retrieval based on FVEC or VLAD had better performance than other approaches. Moreover on a query dataset of range scans, VLAD-based retrieval had better precision at low recall. Given the low computational cost of VLAD compared to FVEC, we recommend VLAD as the default feature encoding for shape retrieval.

Tasse was supported by a Google European Doctoral Fellowship and an IDB Cambridge International Scholarship. Kosinka was supported by the Engineering and Physical Sciences Research Council [EP/H030115/1].

References

[BBGO11] Bronstein A. M., Bronstein M. M., Guibas L. J., Ovsjanikov M.: Shape Google: Geometric words and expressions for invariant shape retrieval. *ACM Trans. Graph. 30*, 1 (Feb. 2011), 1:1–1:20. 1

[CLVZ11] Chatfield K., Lempitsky V., Vedaldi A., Zisserman A.: The devil is in the details: an evaluation of recent feature encoding methods. In *British Machine Vision Conference* (2011). 1

[DLR77] Dempster A. P., Laird N. M., Rubin D. B.: Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B 39* (1977), 1–38. 3

[GBP07] Giorgi D., Biasotti S., Paraboschi L.: Shape retrieval contest 2007: Watertight models track, 2007. 2, 4

[GDB*15] Godil A., Dutagaci H., Bustos B., Choi S., Dong S., Furuya T., Li H., Link N., Moriyama A., Meruane R., Ohbuchi R., Paulus D., Schreck T., Seib V., Sipiran I., Yin H., Zhang C.: Range Scans based 3D Shape Retrieval. In *Eurographics Workshop on 3D Object Retrieval* (2015). 1, 2, 4

[GTU14] Guler R., Tari S., Unal G.: Screened poisson hyperfields for shape coding. *SIAM J. Imaging Sciences 7*, 4 (2014), 2558–2590. 1

[JDSP10] Jégou H., Douze M., Schmid C., Pérez P.: Aggregating local descriptors into a compact image representation. In *CVPR* (Jun 2010), pp. 3304–3311. 1, 3

[LBBC14] Litman R., Bronstein A., Bronstein M., Castellani U.: Supervised learning of bag-of-features shape descriptors using sparse coding. *Computer Graphics Forum* (2014). 1

[LZC*15] Lian Z., Zhang J., Choi S., ElNaghy H., El-Sana J., Furuya T., Giachetti A., Guler R. A., Lai L., Li C., Li H., Limberger F. A., Martin R., Nakanishi R. U., Neto A. P., Nonato L. G., Ohbuchi R., Pevzner K., Pickup D., Rosin P., Sharf A., Sun L., Sun X., Tari S., Unal G., Wilson R. C.: Non-rigid 3D Shape Retrieval. In *Eurographics Workshop on 3D Object Retrieval* (2015). 2, 4

[PCI*08] Philbin J., Chum O., Isard M., Sivic J., Zisserman A.: Lost in quantization: Improving particular object retrieval in large scale image databases. In *CVPR'08* (June 2008), pp. 1–8. 1, 3

[PSM10] Perronnin F., Sánchez J., Mensink T.: Improving the fisher kernel for large-scale image classification. In *ECCV'10* (Berlin, Heidelberg, 2010), Springer-Verlag, pp. 143–156. 1, 3

[RBB09] Rusu R. B., Blodow N., Beetz M.: Fast point feature histograms (FPFH) for 3D registration. In *ICRA* (2009), pp. 1848–1853. 1, 2

[SPS15] Savelonas M. A., Pratikakis I., Sfikas K.: Partial 3d object retrieval combining local shape descriptors with global fisher vectors. In *Eurographics Workshop on 3D Object Retrieval* (2015). 1

[SZ03] Sivic J., Zisserman A.: Video Google: A text retrieval approach to object matching in videos. In *ICCV* (2003). 1, 3, 4

[TKD15] Tasse F., Kosinka J., Dodgson N.: Cluster-based point set saliency. In *ICCV 2015* (2015). 2

[TKD16] Tasse F., Kosinka J., Dodgson N.: How well do saliency-based features perform for shape retrieval?, 2016. Under review. 1, 2

[TLPG14] Tabia H., Laga H., Picard D., Gosselin P.-H.: Covariance descriptors for 3D shape matching and retrieval. In *CVPR* (Jun 2014). 1

[VF10] Vedaldi A., Fulkerson B.: VLFeat—an open and portable library of computer vision algorithms. In *ACM Int'l Conf. on Multimedia* (2010). 3

[Wil45] Wilcoxon F.: Individual comparisons by ranking methods. *Biometrics Bulletin 1*, 6 (1945), 80–83. 4