Pattern Recognition Letters 83 (2016) 368–378

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Shape similarity assessment based on partial feature aggregation and ranking lists☆

Zhenzhong Kuang[a,c], Zongmin Li[a,*], Yujie Liu[a], Changqing Zou[b,d,*]

[a] China University of Petroleum, No. 66 Changjiang West Road, Huangdao District, Qingdao 266580, China
[b] Hunan Provincial Key Laboratory for Technology and Application of Cultural Heritage Digitalization, Hengyang Normal University, Hengyang 421002, China
[c] University of North Carolina, Charlotte, NC 28223, USA
[d] Simon Fraser University, 8888 University Dr, Burnaby, BC V5A 1S6, Canada

ARTICLE INFO

Article history:
Available online 5 July 2016

Keywords:
Shape retrieval
Isometric
Partial aggregation
Distance mapping
RKNN

ABSTRACT

In this paper, we focus on the problem of similarity assessment of isometric 3D shapes, which is of great relevance in improving the effectiveness of retrieval tasks. We first present an effective shape representation technique by proposing a partial aggregation model based on the bag-of-words paradigm. This technique can effectively encode our multiscale local features and has a good discriminatory ability. We then develop a parameter-free distance mapping approach to re-evaluate the similarity results based on intrinsic analysis of a well organized reciprocal k-nearest neighborhood graph. Different from the existing methods which determine k manually and globally, the proposed method can automatically adjust k to a reliable local domain, which therefore ensures a more accurate similarity measurement. We fully study our shape representation technique and evaluate the performance of the proposed distance mapping approach on several popular public shape benchmarks. Experiment results have demonstrated the state-of-the-art performance of our approach.

1. Introduction

The research and development of 3D modeling has resulted in an increasing amount of 3D models in multiple fields including multimedia, graphics, entertainment, design, manufacturing, and so on. The content-based similarity assessment of 3D objects from different classes has been being used in a number of established and emerging fields. To distinguish inter-class shapes, a common feature of existing methods is to employ descriptors that capture the major characteristic of 3D objects.

The similarity assessment of isometric non-rigid 3D shapes is a challenging problem and it has attracted extensive attention from the researchers. This challenge usually becomes much harder when there exist intra-shape deformations caused by the factors such as shape scaling and noises [7,24,25]. In the past few years, there has been considerable research on global and local shape descriptors, such as the global distance feature [9,12,18,31], part-based feature [1,41,44] and the keypoint based feature [2,10,28,38,42].

The keypoint and part based shape descriptors have recently attracted more attention due to their flexibility for partial shape expression, such as depth image [26,42], manifold geometry [17,23,40], and so on. The well-known bags-of-word (BOW) framework [19,20,22,29,37,48], built on a collection of keypoint or patch based features, is usually employed to represent a shape. In particular, the BOW model with soft assignment strategy is more preferred and has demonstrated its advantages in many 2D and 3D shape retrieval tasks [4,7,48]. The spatial information is a critical issue in improving the effectiveness of local descriptors and many versions of the spatial BOW model (e.g. the Hybrid BOW [20]) have been proposed based on the standard BOW model. Lately, Li et al. [23] presented a multiscale shape context (MSC) feature combined with a scale sensitive BOW model for the shape retrieval. The BOW voting schemes in existing 3D shape retrieval methods usually weight all vocabulary words undiscriminatingly and evenly when encoding each local feature, which would accumulate noises from cross-class objects and lead to lower accuracy.

Recently, many context-sensitive methods have been proposed to re-evaluate the similarity ranking [15,45–47]. This is because the initial similarity measurement usually suffers from noise due to inappropriate features or distances (e.g. L_1 or L_2), which would leads to inaccurate ranking results. Kontschieder et al. [16] proposed to use a modified mutual-KNN (mKNN) graph for shape retrieval and

☆ This paper has been recommended for acceptance by Xiang Bai.

* Corresponding author.
E-mail addresses: lizongmin@upc.edu.cn (Z. Li), aaronzou1125@gmail.com (C. Zou).

http://dx.doi.org/10.1016/j.patrec.2016.05.026
0167-8655/

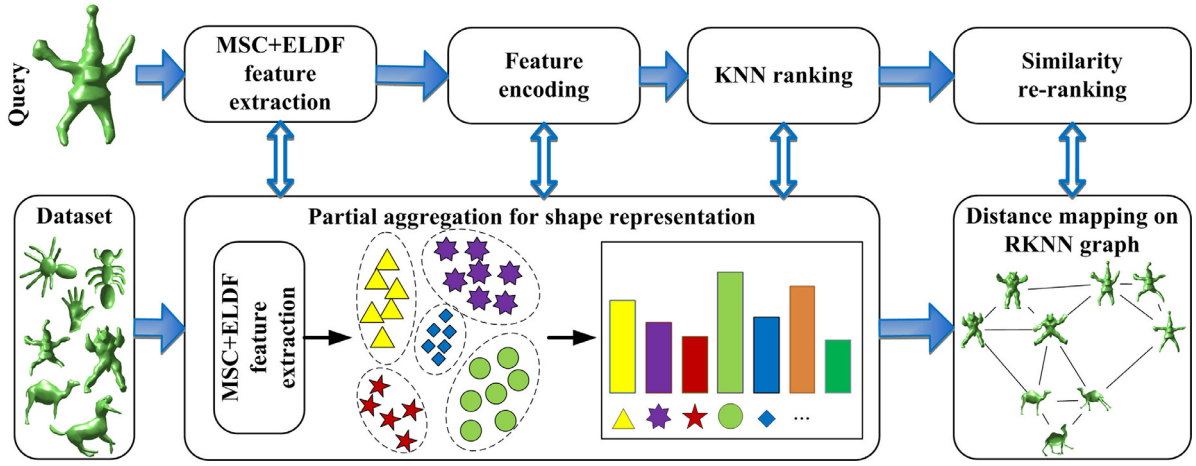

Fig. 1. The pipeline of this proposed approach. Our contribution mainly consist of two parts: (1) a shape representation framework built on improved partial aggregation model and MSC+ELDF feature, and (2) a distance mapping based similarity re-ranking algorithm.

clustering. Yang and Bai et al. studied a flow of algorithms for similarity learning [5,49,50], including locally constraint diffusion process (LCDP), graph transduction (GT), and tensor product graph (TPG). Among these algorithms, TPG [50] integrates the relations of higher order than pairwise affinities into the diffusion process and obtain a good retrieval performance. However, it requires high computation and storage cost. In [13], the authors carried out a detailed comparison of the diffusion related similarity re-ranking methods and, on this basis, they proposed a new diffusion process (DP) to propagate affinity on a k-nearest neighborhood (KNN) graph. In [34] and [30], the authors developed a new re-ranking method (RLSim) by only using the rank list of each query and they achieved a promising performance in improving retrieval accuracy. Lately, Bai et al. [3] suggested employing the neighbor set similarity (NSS) for similarity re-ranking and Li et al.[23] presented a metric mapping method for the re-ranking task. Existing similarity re-ranking algorithms benefit from proper modeling and a well organized KNN graph. Although prior works have attained some promising results, they can be further improved since they use a predefined k in the KNN graph and therefore inevitably introduce noises into the KNN list which further limits their performance.

In this paper, we improve the problem of shape similarity assessment by two parts: an effective shape representation approach and a novel re-ranking technique. The flowchart of this paper is illustrated in Fig. 1. On the part of shape representation, we develop a partial aggregation (PA) model on the works of Li et al. [23] and Bronstein et al. [7]. Unlike previous encoding methods for 3D shape representation, our model aggregates local multiscale features by considering both the scale and position information, which has shown a great performance improvement in experiment. The novelty and advantage of PA is that it alleviates noise and enhances the spatial-sensitivity of local features. Moreover, we also extend MSC [23] and propose a new local descriptor that better fits the problem of shape similarity assessment.

On the part of re-ranking technique, we design a parameter-free distance mapping method to discover the intra-class shapes based on a reciprocal KNN graph. Specifically, our algorithm automatically decides a local graph parameter k and therefore reduces the effect from the noises, which has not been addressed by previous work. We have fully evaluated our methods on different benchmarks. The results show that our approach has achieved state-of-the-art performance.

The rest of this paper is organized as follows. Section 2 presents our feature extraction and partial aggregation model for shape representation. Section 3 describes our distance mapping method for re-ranking. Section 4 shows the results of our experiments, and lastly Section 5 draws conclusions.

2. Partial aggregation for shape representation

Following the flowchart in Fig. 1, we separate the proposed shape representation method into two components: local feature extraction and feature encoding. For the first component, we generate keypoints and multiscale shape context (MSC) domain following [23] and create a more effective local feature. For the second component, we present a distinguished partial aggregation model for feature encoding.

2.1. Local feature extraction

The multiscale property of MSC has enabled it to grasp different co-occurrence information in each keypoint domain, which makes the feature spatial sensitive and informative. Given shape X, a keypoint set $\mathcal{P} \subseteq X$ is detected and a MSC feature is defined for each keypoint $x \in \mathcal{P}$

$$\mathcal{MSC}(x) = (v^l(x), r^l)_{l=1}^{\tau} \tag{1}$$

where τ denotes the number of scales and a domain is assigned to scale l centered at x with radius r^l. $v^l(x)$ represents the feature vector at scale l, which is defined as the histogram frequency of the distances between x and the vertices in the ball domain. Then, the resulting local features are used to represent shape X

$$\mathcal{M}(X) = \{\mathcal{MSC}(x), x \in \mathcal{P}\}. \tag{2}$$

Although some promising results were presented in [23], the retrieval accuracy of MSC is still unsatisfactory due to the limited performance of the adopted local distribution feature (LDF). As an improvement, we propose an extended LDF (ELDF) feature to describe the information at each scale

$$\tilde{v}^l(x) = (v^l(x), \xi^l(x)) \tag{3}$$

where $\xi^l(x)$ is defined as the distribution histogram of the heat diffusion function $\delta(x, \cdot)$ to compensate for the information loss of $v^l(x)$. We use B_1, B_2 and $B = B_1 + B_2$ to represent the length of $v^l(x)$, $\xi^l(x)$ and $\tilde{v}^l(x)$, respectively.

In the recent years, heat diffusion exhibits promising results for shape deformation analysis [7,38,40] and the heat kernel [21] is quite popular used to define shape features

$$h(x, y, t) = \sum_i e^{-\lambda_i t}\phi_i(x)\phi_i(y), \qquad x, y \in X. \tag{4}$$

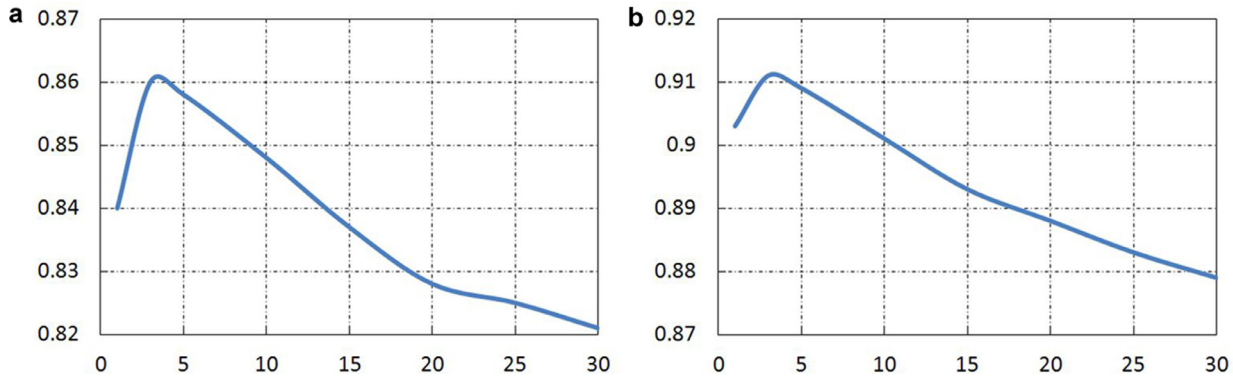

Fig. 2. Retrieval precision of partial aggregation towards varying m values (the horizontal axis) on two different measures: (a) FT, (b) mAP.

However, it is a difficult task to determine the uniform and ideal scale parameter t for $h(x, y, t)$ automatically for different keypoints and shapes. To deal with this issue, we propose to integrate it to achieve a balance among different time scales

$$\delta(x,y) = \int_t h(x,y,t)dt = \sum_i \frac{\phi_i(x)\phi_i(y)}{\lambda_i} \tag{5}$$

where ϕ_i and λ_i are the eigenfunction and eigenvalue of Laplace–Beltrami operator (LBO) which satisfies heat equation $\Delta_X u(x,t) = -\frac{\partial}{\partial t}u(x,t)$ on shape X [21,40]. As with many previous works built on diffusion manifold, it is easy to prove that our ELDF feature is equipped with the property of translation, rotation and scaling invariance [9,38,40].

2.2. Partial aggregation (PA)

For a 3D model, the descriptor of a single vertex could not provide sufficient information, but we are able to tell different shapes apart by combining various local descriptors. To aggregate the multiscale local features of each 3D shape, we extend the BOW scheme by considering how to (1) effectively encode the non-aligned multiscale features without supervision, (2) encode the scale information among local descriptors, and (3) strengthen the discriminative ability.

To avoid mismatching, we do not encode the multiscale feature of each shape directly because there exist unknown scale-offsets for multiscale features extracted from different shapes and it is difficult to perform scale alignment. Instead, we solve the problem by separating $\mathcal{MSC}(x)$ into several scale-based vectors and the feature set of shape X is redefined as

$$\mathcal{M}(X) = \left\{\bar{v}^l(x) \mid 1 \le l \le \tau, x \in \mathcal{P}\right\}, \tag{6}$$

$$\bar{v}^l(x) = (\bar{v}^l_1(x), \bar{v}^l_2(x), \ldots, \bar{v}^l_k(x), \ldots, \bar{v}^l_B(x)), \tag{7}$$

where $1 \le k \le B$. On shape dataset $\mathbb{D}$, we obtain a feature collection $\mathbf{M}_{|\mathbb{D}|\times B} = \{\bar{v}^l(x) \in \mathcal{M}(X) | \forall X \in \mathbb{D}\}$. A scale-based visual word codebook $C = \{C_1, \ldots, C_j, \ldots, C_V\}$ is generated to encode local features across multiple scales by the k-means clustering on $\mathbf{M}_{|\mathbb{D}|\times B}$, where each of the visual word C_j encodes local features of similar scales implicitly. To improve the clustering stability, we initialize the centers by three steps: (a) Find a column $\mathbf{M}_{|\mathbb{D}|\times B}(\cdot, k)$ that has the largest variance; (b) Sort $\mathbf{M}_{|\mathbb{D}|\times B}$ in descending order according to its kth column; (c) Choose V equally-spaced vectors from $\mathbf{M}_{|\mathbb{D}|\times B}$ as the initial centers. The Manhattan metric is used for feature comparison.

For feature encoding, the effectiveness of soft assignment has been evaluated in many previous works [4,7,23]. But, they still performs poor in spatial-sensitive representation, computational efficiency and noise resistance. To overcome the defects, we design a partial aggregation scheme to encode the proposed multiscale feature with both implicit and explicit constrains on scale and position. Given feature point $x \in \mathcal{P} \subseteq X$, we redefine its feature distribution at the lth scale (i.e. $\bar{v}^l(\cdot) \to \chi^l(\cdot)$) as

$$\chi^l(x) = (\chi^l_1(x), \chi^l_2(x), \ldots, \chi^l_j(x), \ldots, \chi^l_V(x)) \tag{8}$$

$$\chi^l_j(x) = \eta^l_j(x) * s(l) \tag{9}$$

where

$$\eta^l_j(x) = \begin{cases} e^{(\vartheta(m+1)-\vartheta(\kappa))} * g(\kappa) & \kappa <= m \\ 0 & \kappa > m \end{cases} \tag{10}$$

represents the scale-based visual word score and

$$s(l) = log(1 + l/\tau) \tag{11}$$

is the explicit scale importance factor. The function $\vartheta = \{\vartheta(1), \ldots, \vartheta(V)\}$, $\vartheta(j) \le \vartheta(j+1)$ ranks $\theta = \{\theta(1), \ldots, \theta(V)\}$ $(\theta(j) = \|v^l(x) - C_j\|_{L_1})$ in ascending order, κ indicates the position of C_j in ϑ. Actually, only part of the scale-based visual words are correlated with $v^l(x)$ due to 1-dimensional scale similarity and visual word reliability, and it would lead to overfitting if all visual words are considered. Thus, we only weight the top m most similar items (see Eq. (10)), which could improve the precision and reduce noise as well (see the influential curves in Fig. 2 and $m = 5$ is adopted). To highlight the importance of visual words at different positions, we design a ranking position (κ) based weighting scheme with a gaussian falloff function

$$g(\kappa) = e^{\epsilon^{\kappa-1}}, 1 \le \kappa \le V. \tag{12}$$

where $\epsilon(= 0.5)$ is a decay parameter. Then, we represent shape X as a $1 \times V$ vector by aggregating its local features

$$F(X) = (F_1(X), F_2(X), \ldots, F_j(X), \ldots, F_V(X)) \tag{13}$$

where

$$F_j(X) = \sum_{x\in\mathcal{P}}\sum_{l=1}^{\tau} \chi^l_j(x). \tag{14}$$

To further improve the discriminative ability, we highlight the visual words that just belong to a specific shape class but are rare in the other classes by using a soft tf-idf frequency score

$$\bar{F}(X) = (\bar{F}_1(X), \bar{F}_2(X), \ldots, \bar{F}_V(X)) \tag{15}$$

where

$$\bar{F}_j(X) = \frac{F_j(X)}{\sum_{j=1}^{V} F_j(X)} log\left(\frac{\sum_X \sum_{j=1}^{V} F_j(X)}{\sum_X F_j(X)}\right). \tag{16}$$

Then, the similarity score between two shapes X and Y are converted to comparison of their features $\bar{F}(X)$ and $\bar{F}(Y)$

$$d(X,Y) = \|\bar{F}(X) - \bar{F}(Y)\|_{L_1}. \tag{17}$$

Compared with the traditional methods for 3D shape retrieval [7,19,20,23,42,44], our method incorporates more spatial and weighting constrains for multiscale feature encoding, which could improve the effectiveness of shape representation. The local encoding schemes in image retrieval and classification problem such as [27] and [35] simply used the top m visual words to encode the flat features, while PA considers a different topic on encoding the non-aligned multiscale features of 3D shapes by using distinct encoding steps. Besides, PA has some other considerations on the scale and position information which are not discussed in [27,35] or other works on the deformable 3D shape retrieval problem.

3. Similarity re-ranking

For shape retrieval, ranking lists are usually employed to find the most similar results, but there always exist lots of noise (i.e. inter-class objects). Thus, a similarity re-ranking process is desired for denoising. The work of Donoser and Bischof [13], Bai et al. [3] and Li et al. [23] focus on recovering the latent similarity relationships without supervision and they give us lots of inspiration. Different from these works, we propose a parameter-free distance mapping (PDM) method based on reciprocal KNN (RKNN) graph.

Suppose the dataset objects $O = (O_i)_{i=1}^n$ have good clustering properties, most of the query results fall into a same cluster and the distances between the query and the intra-cluster objects are much smaller than the other objects. Numerically, the problem becomes the discovery of the optimal paths on the KNN graph, which leads to updating of the distance between the dataset objects and the query. In the following subsections, we introduce our offline process to discover new similarity measures

$$d(O_i, O_j) \to D(O_i, O_j). \tag{18}$$

3.1. Distance mapping based on RKNN graph

Our method consists of three steps: (a) build a local constrained KNN graph, (b) obtain new coordinates of the dataset objects by spectral embedding and (c) compute new distance measures for pairwise objects.

Given a graph $G = (O, E)$ and the edge set E, we define the weight of each edge by

$$W(O_i, O_j) = \begin{cases} e^{-(d(O_i,O_j)/\sigma_{ij})^2} & \text{if } O_j \in \mathcal{R}(O_i) \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where $\mathcal{R}(O_i)$ is the KNN ranking list of O_i, the length of $\mathcal{R}(O_i)$ is denoted as k (which is determined automatically in Section 3.2), σ_{ij} is a parameter used for normalization. For convenience, we use $W(i, j)$ to represent $W(O_i, O_j)$ in the following sections.

To decrease the noise in G, we convert it to a RKNN graph based on a reciprocal restrain which requires to satisfy the symmetry ranking criteria: $O_j \in \mathcal{R}(O_i)$ and $O_i \in \mathcal{R}(O_j)$, where $W(i, j)$ is set to zero if it violates the criteria.

It is important to note that the RKNN graph contains lots of curved paths for intra-class shapes and can be seen as a specific shape. Therefore it is interesting to model the distance mapping problem as the understanding of geometry [21]. The numerous investigations of the intrinsic analysis [11,36,38] show that one can describe the complicated structure of a graph in intrinsic space or spectral space. We start by defining the following Laplacian matrix

$$L = A^{-1}(D - W) \tag{20}$$

where $A = D/\Sigma D$ is a diagonal matrix for normalization and

$$D(i,i) = \sum_{O_j \in O} W(i,j). \tag{21}$$

Then, we embed the above RKNN graph G into the spectral space $S = (\Phi, \Lambda)$ based on the generalized eigensystem

$$L\Phi = \Lambda\Phi \tag{22}$$

$$\Lambda = diag(\mu_i)_{i=1}^{\theta}, \Phi = [\psi_i]_{i=1}^{\theta} \tag{23}$$

where μ_i and ψ_i are eigenvalue and eigenfunctions of L, and the smallest θ eigenpairs are used to resist noise. To get a real solution, we recover the symmetry property of W by setting

$$W(i,j) = (W(i,j)^2 + W(j,i)^2)/2, W(j,i) = W(i,j) \tag{24}$$

Finally, we design a normalized distance function

$$D(O_i, O_j) = \hat{d}(p_i, p_j)/\delta \tag{25}$$

as the new similarity measures between O_i and O_j, where $\delta = max(\hat{d}(\cdot,\cdot))$ and $\hat{d}(p_i, p_j) = \|p_i - p_j\|_{L_2}$ is the L_2 distance between p_i and p_j which are the new coordinates of O_i and O_j, respectively, in spectral space [23].

Compared with previous work [3,13,23], our method has several advantages: (1) The PDM model is based on intrinsic analysis and therefore it can naturally capture the latent pairwise similarity; (2) The information in RKNN graph is more compact and effective with less noise; (3) Unlike existing works that choose a uniform k for all the ranking lists, PDM adapts k automatically by the method in Section 3.2. Although our RKNN graph is similar to the graph adopted in [16] and [34], they are applied to quite different modeling frameworks for similarity re-ranking. Moreover, the PDM model has inherited the robustness property of spectral analysis [9,21,38,40], which indicates that it can still work well even if there exist some noise in the RKNN graph.

3.2. Automatically determined parameters for RKNN graph

Although the fixed sized KNN lists work in some cases [3,13,16,34], they have problems in dealing with the unevenly distributed datasets. Given a query list $\mathcal{R}(q)$, it is reasonable to set its length k as the number of intra-cluster objects. As shown in Algorithm 1, an optimal k is obtained by analyzing the dataset

Algorithm 1: KNN parameter estimation.

Input: $q \in O$, $d(O_i, O_j)$, $O = (O_i)_{i=1}^n$
(1) Pre-Classification: $k \leftarrow S_t$, S_{min}, S_{mean}, S_{max}
(2) Compute belief score: $bf(q, O_j)$, $j = S_t \to S_{max}$
(3) Estimate reliable candidates: RC
for $j = S_t \to S_{max}$ **do**
  Count intra/inter-class candidate: $c(bf = 1)$, $c(bf = 0)$
  if $c(bf = 1) > c(bf = 0)$, $bf(j) = 1$ **then**
    $RC \leftarrow (j - S_t, v)$, $v = c(bf = 1) - c(bf = 0)$
  end if
end for
Output: $k = k + argmax_l\{\, v \mid \forall l, (l, v) \in RC \,\}$

ranking lists in three steps:

(1) Pre-classification. We divide the dataset into several groups by automatic clustering algorithm (e.g. AP clustering, [14]) and initialize k by $S_t = S_{mean} - \lfloor S_{min}/2 \rfloor$, where S_{min}, S_{mean} and S_{max} denote the min, mean and max group size.

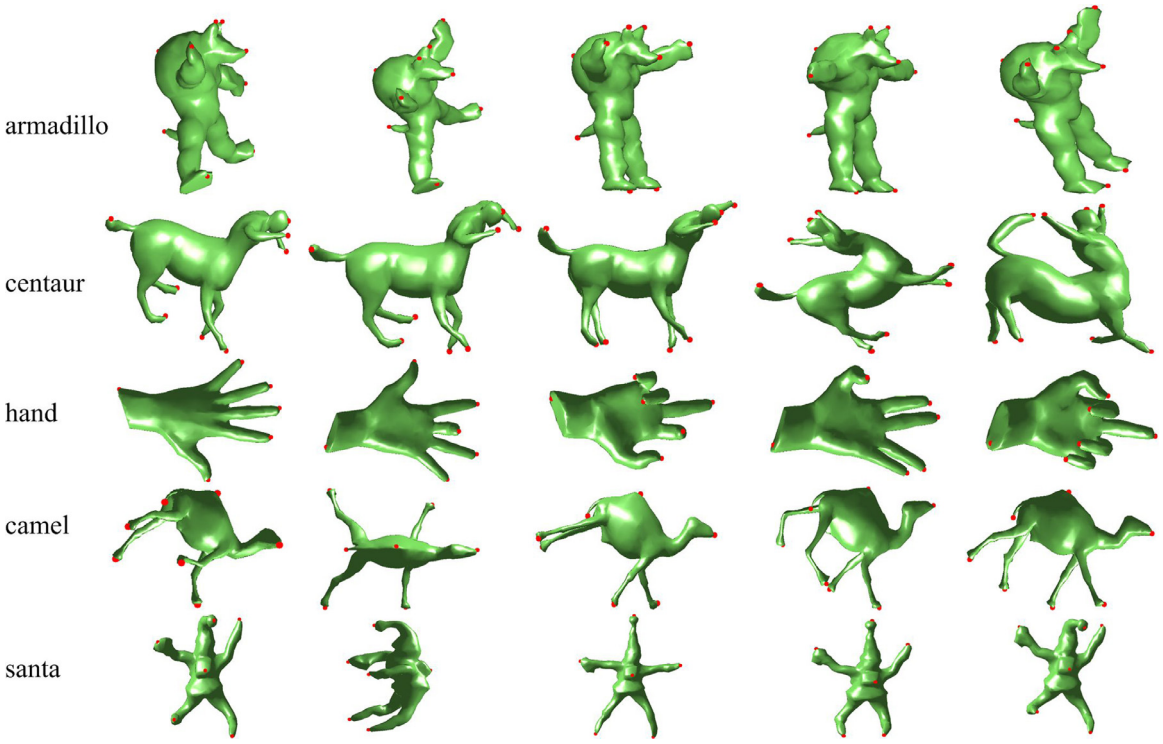

Fig. 3. Some isometric 3D models and their corresponding feature points (marked in red dot) found by our approach.

(2) Compute belief score. We then compute a belief score for each candidate object in $\mathcal{R}(q)$ by

$$bf(q, O_i) = \begin{cases} 1 & \rho(q, O_i) > \alpha \\ 0 & \text{else} \end{cases}, \tag{26}$$

where $\alpha(=0.9)$ is a correlation threshold and $\rho(q, O_i)$ represents the similarity score between the ranking list of q and $O_i \in \mathcal{R}(q)$

$$\rho(q, O_i) = \max_{k \in [S_t, S_{max}]} \frac{|\mathcal{R}(q) \cap \mathcal{R}(O_i)|}{|\mathcal{R}(q) \cup \mathcal{R}(O_i)|}. \tag{27}$$

(3) Estimate reliable candidates. We adjust k by searching the most reliable candidate objects in range $[S_t, S_{max}]$. By counting the number of intra-class and inter-class objects, we obtain a set of reliable candidates RC that could provide positive gains for cutting-off $\mathcal{R}(q)$ to preserve the information of the ranking list. By maximizing it, we obtain the final estimation of k. Although Algorithm 1 cannot always obtain an optimal k due to noise effect, however, an output near the optimal value can still improve the result with our distance mapping method.

4. Experiment

In this section, we validate our method on several well-known shape retrieval benchmarks and experimentally evaluate its performance against state-of-the-art approaches.

Dataset. To illustrate the effectiveness and generalization capability of the propose approach, we perform experiments on four benchmarks, namely, the non-rigid world dataset (NRW) [8,32], McGill dataset [42], SHREC Non-Rigid 3D Models 2010 (SHREC10) dataset [25] and SHREC 2011 Non Rigid 3D Watertight Meshes (SHREC11) dataset [24]. NRW dataset is a basic non-rigid dataset for shape recognition, which contains 148 shapes unevenly categorized into 12 classes. McGill dataset contains 255 objects unevenly divided into 10 classes, while SHREC10 dataset contains 200 shapes evenly distributed into 10 classes with 20 shapes in each and SHREC11 dataset consists of 600 watertight triangle mesh shapes evenly classified into 30 categories. All these datasets have been publicly admitted for shape recognition because they are characterized by a high degree of non-rigid deformations. Besides, some deformations of the shapes are artificially generated, which would lead to misleading recognition. Considering the computation speed, all the shapes are down sampled to 2000 faces which still preserve the main features of each shape. Some typical models are shown in Fig. 3.

Evaluation. Following the work of Osada et al., Lian et al., [24,25,31] and Tabia et al. [41], we perform a leave-one-out KNN retrieval experiment for easy comparison, where each deformed shape is queried against the remaining models in the dataset and matches are regarded correct between different deformations of the same shape (or class). To comprehensively and quantitatively assess the retrieval results, the following popular evaluation criteria [39] are employed: Precision-Recall (PR) diagram, Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), Discounted Cumulative Gain (DCG) and mean Average Precision (mAP).

4.1. Brief overview of the keypoints

We first present the clustering property of the detected keypoints at different scales in Fig. 3, where the keypoints on each deformed shape belong to several classes according to their locations (e.g. leg). By the 2D embedding results plotted in Fig. 4 (e), we see that different shape parts are well separated with large margins. According to Fig. 4 (a), the discriminative ability of single scale is distinct for different parts and shapes. It is clear that the different shape parts become more separable by increasing the number of the scales. By Fig. 5, we observe that large scale has better discriminative ability. But it is not always true for the keypoints located in the same or nearby components, such as the ear and nose in Fig. 5 (e). By combing multiple scales, we obtain the boosted result in the first line of Fig. 4.

Because we only detect a few salient points on each shape, the complexity of our method is quite low. This enables us to find

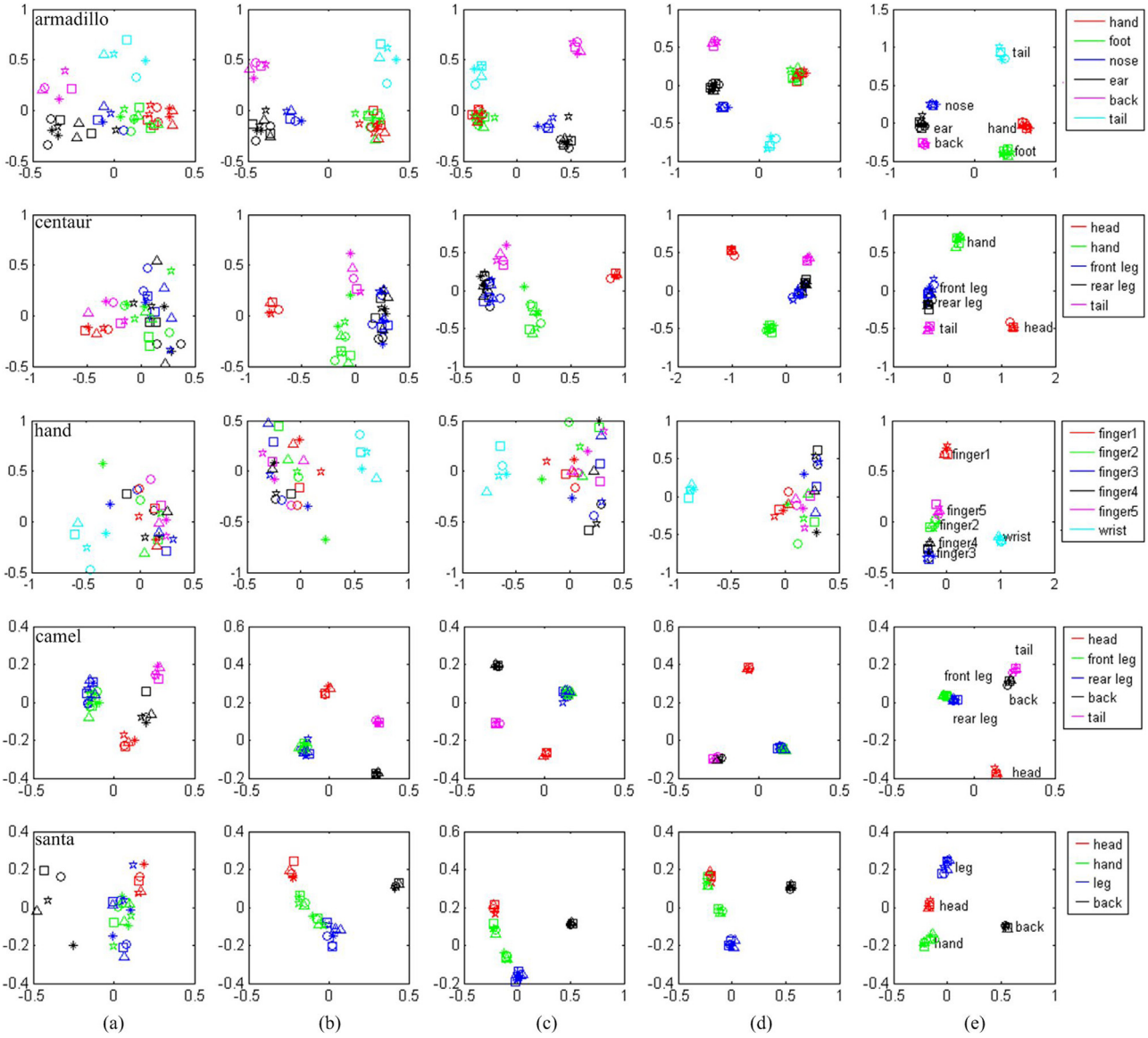

Fig. 4. The ISOMAP embedding [43] of the keypoints in Fig. 3, where $\kappa = 5$ with interval 0.02 is adopted for each keypoint: (a) $S = 0.01$; (b) $S = [0.01, 0.0.09]$; (c) $S = [0.01, 0.19]$; (d) $S = [0.01, 0.39]$; (e) $S = [0.01, 0.59]$. Note that the fingers of the hand shape are sequentially indexed from the thumb to the little finger.

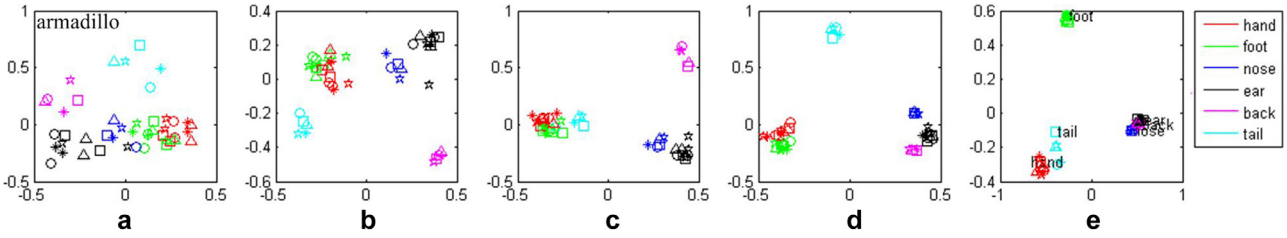

Fig. 5. The ISOMAP embedding [43] of the single scale based keypoints for the armadillo shape in Fig. 3: (a) $s = 0.01$; (b) $s = 0.09$; (b) $s = 0.19$; (b) $s = 0.39$; (b) $s = 0.59$.

shape correspondences with a low cost. But, for shape recognition, we prefer to use the BOW paradigm to summarize the content of the shape other than the shape correspondences because it is still a challenging topic to match shapes appropriately due to the noise and errors in the low-level feature space. Besides, it is also critical to note that, the points at different shape parts require distinct number of scales to separate them apart (e.g. 'hand', 'foot' and 'nose' in the first row of Fig. 4), which brings more difficulty and which also requires to do scale selection. Fortunately, the devised PA model could deal with the problem well. In the next sections, we evaluate the detailed shape retrieval performance together with comparisons against state-of-the-art (Note that the contrast results are from the original papers or implemented according to them).

4.2. Retrieval on NRW dataset

For convenience, we denote our method using with MSC, ELDF and PA as MSC+ELDF+PA. To demonstrate the performance of our method, we first present some retrieval results on this basic shape benchmark and the following five methods are employed

Table 1
Retrieval performances comparison (%) on NRW dataset.

Feature	NN	FT	ST	DCG	mAP
MSC+ELDF+PA	**94.6**	**84.5**	**96.9**	**94.8**	**92.3**
HKS	92.5	72.7	87.8	91.1	82.2
SIHKS	94.6	66.8	84.1	89.0	78.3
WKS	93.9	73.7	88.9	91.7	84.0
POCSVM	75.5	45.0	61.0	–	–
MSC+LDF+SA	90.5	72.0	87.3	89.3	84.6

Table 2
Retrieval performances comparison (%) on McGill dataset.

Method	NN	FT	ST	DCG	mAP
MSC+ELDF+PA	**99.2**	**82.9**	90.9	**95.9**	**87.7**
SIHKS	57.3	43.9	71.6	74.3	51.9
Hybrid 2D/3D	92.5	55.7	69.8	85.0	–
Hybrid BoW	95.7	63.5	79.0	88.6	–
BOW	94.5	62.9	77.6	88.1	–
Graph-based	97.6	74.1	**91.1**	93.3	–
PCA-based VLAT	96.9	65.8	78.1	89.4	–
Covariance method	97.7	73.2	81.8	93.7	–
MSC+LDF+SA	98.0	81.6	91.0	95.3	87.1

for comparison: HKS [7,40], SIHKS [10], WKS [2], Panorama-OCSVM (POCSVM) [32] and MSC+LDF+SA [23].

In Table 1, we see that our method apparently outperforms all of the compared classical methods HKS, SIHKS and WKS on all the evaluation criteria. MSC+ELDF+PA has improved MSC+LDF+SA a lot by 4.1% in NN, 12.5% in FT, 9.6% in ST, 5.5% in DCG and 7.7% in mAP. It is also interesting to see that MSC+ELDF+PA performs much better than POCSVM with significant performance gains: NN(19.1%), FT(39.5%) and ST(36.9%). The results on this dataset show that our method is promising in lifting shape retrieval accuracy.

4.3. Retrieval on McGill dataset

For contrast, we compare with several state-of-the-art methods, including SIHKS [7,10], Hybrid 2D/3D [33], Hybrid BoW [20], BOW [19], the graph-based approach [1], PCA-based VLAT [42], Covariance method [41] and MSC+LDF+SA [23]. For the details of these methods, we refer the reader to their original papers.

Based on the results shown in Table 2, we find that our method outperforms both the classical methods (e.g. SIHKS and BOW) and the recent published works (e.g. PCA-based VLAT, Covariance method and MSC+LDF+SA). MSC+ELDF+PA leads to significant improvement of 1.2% in NN, 1.3% in FT, 0.6% in DCG and 0.6% in mAP compared with MSC+LDF+SA, and exceeds Covariance method by 1.5% in NN, 9.7% in FT, 9.1 in ST and 2.2% in DCG.

4.4. Retrieval on SHREC10 dataset

On the SHREC10 dataset, the retrieval scores resulting from other state-of-the-art approaches and our method are reported in Table 3. Besides SIHKS [10], POCSVM [32] and MSC+LDF+SA [23], five methods from SHREC 2010 contest [25] (BF-DSIFT-E, DMEVD_run1, SD-GDM [12] or DMEVD_run2, DMEVD_run3 and CF) and HIST [9] are employed for contrast.

Although MSC+ELDF+PA method performs slightly worse (0.3%) than MSC+LDF+SA in ST, it outperforms MSC+LDF+SA by 2.4% in FT, 0.2% in DCG and 0.9% in mAP. And our method has produced significant improvement compared with the classical distance related features SD-GDM (e.g. 7% in FT) and HIST (e.g. 15.9% in FT). For NN, FT, DCG and mAP, our method has achieved the best performance.

Table 3
Retrieval performances comparison (%) on SHREC10 dataset.

Method	NN	FT	ST	DCG	mAP
MSC+ELDF+PA	**99.5**	**85.8**	93.4	**96.5**	**90.9**
SIHKS	66.0	49.8	69.6	77.4	57.8
WKS	83.5	49.6	66.6	79.4	58.5
BF-DSIFT-E	98.0	76.6	89.2	94.1	86.3
SD-GDM	99.5	78.8	**94.4**	96.1	90.0
DMEVD_run3	96.0	71.9	85.1	92.0	82.7
CF	92.0	63.5	78.0	87.8	75.2
POCSVM	97.0	62.0	71.5	–	–
HIST	95.0	69.9	81.5	91.1	78.5
MSC+LDF+SA	99.5	83.4	93.7	96.3	90.0

Table 4
Retrieval performances comparison (%) on the SHREC11 dataset.

Method	NN	FT	ST	DCG	mAP
MSC+ELDF+PA	99.8	**98.4**	**99.4**	**99.7**	**99.1**
SIHKS	97.2	76.4	86.1	92.5	82.7
WKS	95.7	75.7	83.1	91.3	80.9
SD-GDM	**100**	96.2	98.4	99.4	98.0
MDS-CM-BOF	99.5	91.3	96.9	98.2	95.0
meshSIFT	99.5	88.4	96.2	98.0	93.8
BOGH	99.3	81.1	88.4	94.9	86.8
KERG	**100**	88.6	95.2	97.5	–
POCSVM	95.5	61.5	71.5	–	–
Hybrid	99.8	93.5	98.5	98.9	96.7
HIST	99.5	94.3	97.7	98.7	96.7
MSC+LDF+SA	99.5	92.4	97.3	98.4	95.7

4.5. Retrieval on SHREC11 dataset

In Table 4, retrieval results of several state-of-the-art methods are provided: SIHKS [10], WKS [2], four methods from SHREC 2011 contest [24] (SD-GDM, MDS-CM-BOF, meshSIFT, BOGH), KERG [6], POCSVM [32], Hybrid [22], HIST [9] and MSC+LDF+SA [23].

We first note that our method has lifted the performance of MSC+LDF+SA greatly by 0.3% in NN, 6.0% in FT, 2.1% in ST, 1.3% in DCG and 3.4% in mAP. Compared with the best performed SD-GDM in the work of [24], our method has slight accuracy decrease (0.2%) in NN, but works superior in FT, ST, DCG and mAP. In accordance with the performances on the other datasets, these results again validate the effectiveness of our approach.

4.6. Precision recall curves

In addition to the listed quantitative results, we plot the PR curves further to intuitively show the performance of the proposed method. According to the curves plotted in Fig. 6, MSC+ELDF+PA outperforms the classical SIHKS method significantly on all datasets and, compared to MSC+LDF+SA, our method presents much higher performance on NRW and SHREC11 datasets. In all, the PR curves of MSC+ELDF+PA almost stay above all the contrast methods at the same recall rate, which suggests the superior performance of our approach.

4.7. Performance of similarity re-ranking

In this part, we explore the performance of the proposed PDM algorithm and we use MSC+ELDF+PA+PDM to denote that PDM is employed to improve the retrieval accuracy of MSC+ELDF+PA. Different kinds of well performed re-ranking methods are adopted for comparison: NSS [3], DP [13], RLSim [30,34], TPG [50], GT [5], mKNN [16], and LCDP [49].

In Table 5, we compare PDM with the other methods for similarity re-ranking on the test benchmarks. In comparison with the results of MSC+ELDF+PA shown in Table 1–4, MSC+ELDF+PA+PDM

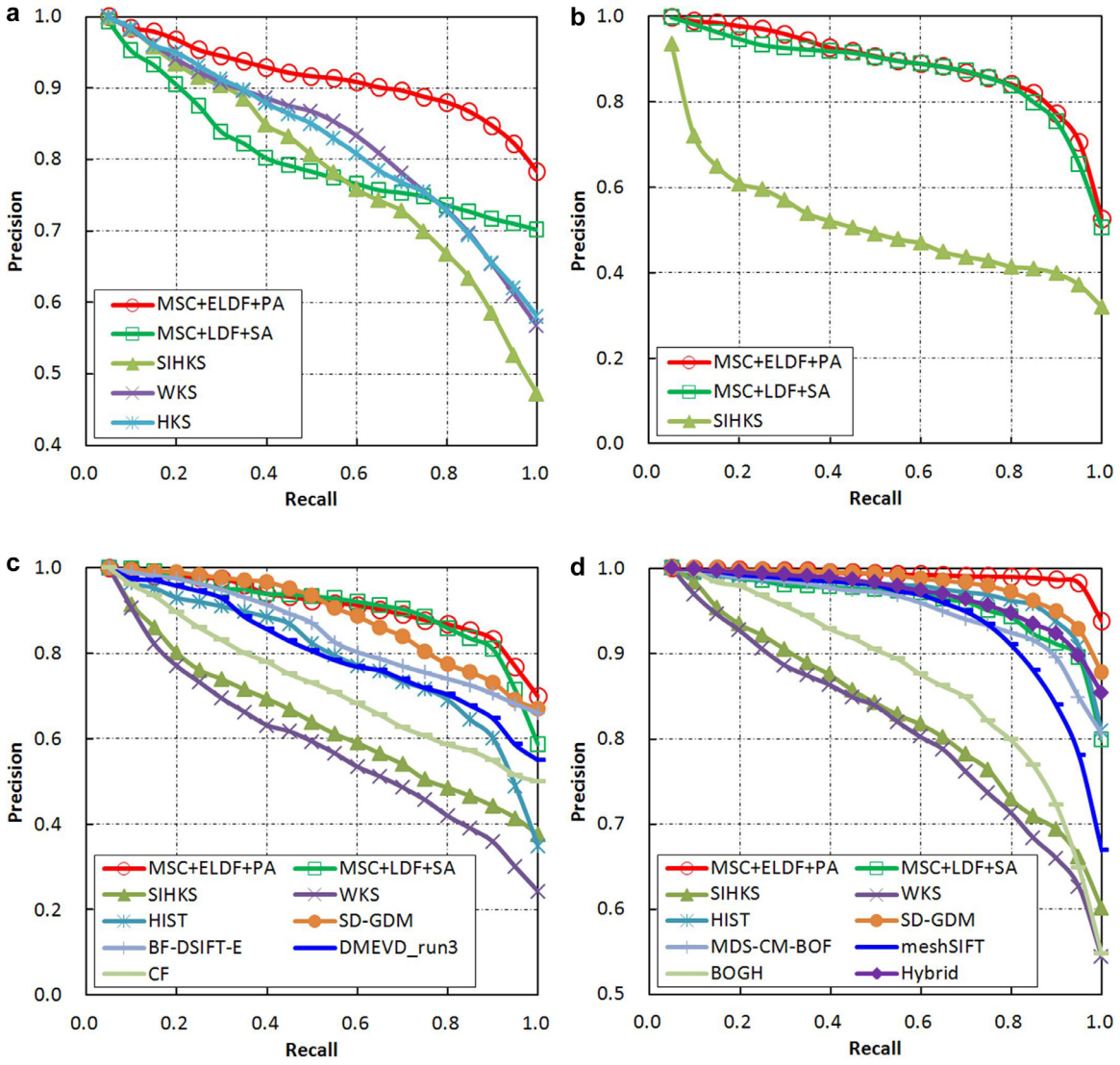

Fig. 6. PR curves of all the comparison methods on the four test benchmarks: (a) NRW dataset, (b) McGill dataset, (c) SHREC10 dataset and (d) SHREC11 dataset.

has resulted in remarkable performance gains (e.g. on McGill dataset: 12.8% in FT, 7.0% in ST, 2.6% in DCG and 9.0% in mAP). On the evenly distributed datasets SHREC10 and SHREC11, all the compared methods works positively well, especially that MSC+ELDF+PA+PDM outperforms all the other methods. However, on the unevenly distributed dataset McGill and NRW, not all these methods produce good results (e.g. mKNN 92.0% v.s. 92.3% baseline), but MSC+ELDF+PA+PDM still obtains quite positive results. While NSS, DP, RLSim and TPG generate good results on some datasets (e.g. DP on SHREC11), they failed to work equally well on all these datasets as with our PDM method did. Similar to 2D shape retrieval, we also present the bullseye score [13] in Table 6. Although our PDM method ranks after the TPG approach on SHREC10 dataset, it has achieved the best accuracy on the other four datasets (e.g. 100% for SHREC11). In Fig. 7, we plot the category based mAP scores of MSC+ELDF+PA and MSC+ELDF+PA+PDM. It is easy to observe that our method has upgraded the retrieval performance of each shape class greatly.

Further, we discuss the scalability of our PDM approach and present more results in Table 7, where two representative distance matrices MPEG7 [13] and MDS-CM-BOF [26] are used (Note that some of the MPEG7 results are shown in Table 6 and Fig. 7). On MPEG7 dataset, our method has reached the highest accuracy on all five measures with significant improvement: 8.4% in NN, 11.8% in FT, 7.0% in ST, 4.5% in DCG and 10.2% in mAP. By using MDS-CM-BOF on SHREC11 dataset, PDM has lifted the precision of FT in 4.3%, ST in 1.8% and mAP in 1.9% with a decrease of 1.7% in NN. By comparing with the other methods, PDM obtain comparable results with RLSim and mKNN. The demonstrated results again display the effectiveness of the proposed method.

4.8. Discussion and analysis

According to the above experiments on various datasets, we discuss the results. First, the multiscale features under the framework of MSC+ELDF has helped to summarize different levels of the co-occurrence information, which favors shape recognition a lot. By using PA model, we are able to organize these information effectively. Despite of the fact that MSC+ELDF+PA produces salient improvements towards state-of-the-art results, MSC+ELDF+PA+PDM lifts the retrieval performance further. Besides, our PDM algorithm can flexibly determine the length of ranking lists, which is quite different from existing works. Additionally, we need to note that the NRW and McGill datasets are unevenly distributed, which is more difficult for shape recognition compared with the SHREC10 and SHREC11 datasets. But, MSC+ELDF+PA and MSC+ELDF+PA+PDM

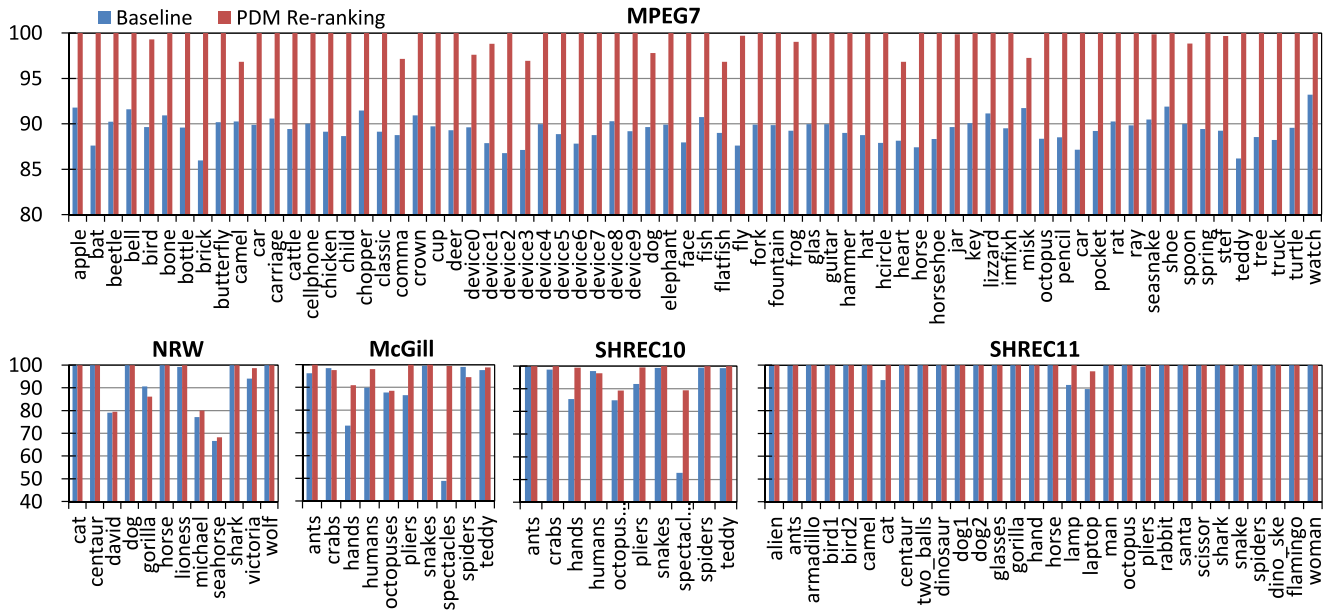

Fig. 7. mAP precision (the vertical axis) of our PDM method on various shape categories of test benchmarks.

still achieve good results on all of them. By summarizing all the experiment results, we are able to confirm that the proposed method has achieved state-of-the-art performance for isometric 3D shape recognition.

Similar to most of previous work, the proposed descriptor focuses on statistical information. Thus, it confronts more errors when dealing with the inter-class shapes that are different only in local parts (e.g. different human classes), especially when noise occur. And one solution is to do further justification after the first round of recognition, which is left for the future work.

Finally, we analyze the major time consumption of the proposed approach: the overall complexity of feature extraction (MSC+ELDF) for shape X with N vertices is $O(N^2)$; the complexity of PA encoding for each shape is $O(V^2|\mathcal{P}|)$; the complexity of the offline distance mapping with n nodes is $O(n^2)$. Then, on an Intel(R) Xeon(R) Dual CPU W3530@3GHz with 3GB memory, we report the average online processing time of each shape by using the MATLAB software in Table 8.

5. Conclusion

In this paper, we propose a novel approach for shape similarity assessment. Based on the framework of MSC, the proposed ELDF feature could summarize different levels of shape information in detail. By the experiment data, we find that the developed PA model has helped to extract the discriminative information with less noise. Further, our similarity re-ranking method has worked effectively in measuring the pairwise similarity relationships. Finally, we have quantitatively and qualitatively evaluated the effectiveness of our approach by experiment. Our method has achieved the best performance over all the compared methods for shape retrieval on multiple popular shape benchmarks. The code is available for free download at https://github.com/kuangzz/MSC.

Table 5
Similarity re-ranking performance of comparison methods on test benchmarks.

MSC+ELDF+PA+PDM	NN	FT	ST	DCG	mAP
NRW	93.2	86.4	97.5	94.2	92.7
McGill	98.4	95.4	98.2	98.6	96.8
SHREC10	98.5	95.5	98.4	98.6	97.4
SHREC11	100	99.8	99.8	100	99.9
MSC+ELDF+PA+NSS	NN	FT	ST	DCG	mAP
NRW	94.6	85.8	96.2	94.4	91.4
McGill	98.0	87.7	97.4	97.1	92.7
SHREC10	98.5	90.4	97.5	97.7	94.8
SHREC11	99.8	99.4	99.9	99.8	99.6
MSC+ELDF+PA+DP	NN	FT	ST	DCG	mAP
NRW	87.8	74.5	88.3	90.9	88.3
McGill	91.4	90.0	98.2	95.7	93.1
SHREC10	89.5	88.6	97.3	94.7	92.8
SHREC11	99.8	99.8	99.8	99.9	99.9
MSC+ELDF+PA+RLSim	NN	FT	ST	DCG	mAP
NRW	93.9	86.3	96.8	95.0	92.7
McGill	98.8	90.5	97.0	97.5	93.8
SHREC10	99.0	93.4	97.5	98.2	95.8
SHREC11	99.7	99.1	99.8	99.7	99.4
MSC+ELDF+PA+TPG	NN	FT	ST	DCG	mAP
NRW	83.1	81.5	93.3	90.1	87.8
McGill	98.8	91.1	98.8	98.2	94.7
SHREC10	99.5	93.5	98.6	98.3	96.2
SHREC11	99.8	99.8	99.8	99.9	99.9
MSC+ELDF+PA+GT	NN	FT	ST	DCG	mAP
NRW	92.6	83.1	94.4	92.8	89.5
McGill	97.6	87.6	95.7	96.0	90.6
SHREC10	98.5	92.5	97.3	97.4	94.5
SHREC11	99.2	99.1	99.6	99.6	99.2
MSC+ELDF+PA+mkNN	NN	FT	ST	DCG	mAP
NRW	95.9	83.1	97.1	94.5	92.0
McGill	99.2	83.7	95.9	96.4	89.8
SHREC10	99.0	90.6	98.0	97.7	94.5
SHREC11	99.8	99.0	99.6	99.7	99.4

Acknowledgments

This work is partly supported by the Fundamental Research Funds for the Central Universities (No. 15CX06017A), National Natural Science Foundation of China (No. 61379106), the Young Scientists Fund of the National Natural Science Foundation of China (No. 61502153), Natural Science Foundation of Hunan Province of China (No. 2016JJ3031), the Shandong Provincial Natural Science Foundation (No. ZR2013FM036), the Scholarship from China Scholarship Council (No. 201506450033) and the open fund project of Hunan Provincial Key Laboratory for Technology and Application of Cultural Heritage Digitalization. The authors would like to thank the anonymous reviewers and editors for their insightful comments and valuable suggestions which have led to substantial improvements of the paper.

Table 6
Comparison of the bullseye score (i.e. *top 40* accuracy) on the test datasets, where the data in brackets represent the gains against the baseline.

	baseline	PDM	NSS	DP	RLSim	TPG	GT	mkNN	LCDP
NRW	99.2	**99.9(0.7)**	**99.9(0.7)**	90.2(−9.0)	99.8(0.6)	99.4(0.2)	98.2(−1.0)	99.0(−0.2)	–
McGill	89.9	**98.2(8.3)**	95.4(5.5)	96.8(6.9)	96.5(6.6)	97.2(7.3)	94.5(4.6)	93.9(4.0)	–
SHREC10	93.9	98.5(4.6)	97.8(3.9)	97.8(3.9)	97.7(3.8)	**98.8(4.9)**	97.6(3.7)	98.2(4.3)	–
SHREC11	99.5	**100(0.5)**	**100(0.5)**	99.8(0.3)	99.8(0.3)	99.8(0.3)	99.6(0.1)	99.7(0.2)	–
MPEG7	93.6	**100(6.4)**	**100(6.4)**	**100(6.4)**	99.9(6.3)	99.9(6.3)	99.9(6.3)	99.7(6.1)	96.0(2.4)

Table 7
Experiment results of similarity re-ranking on MPEG7 dataset and SHREC11 dataset.

	NN	FT	ST	DCG	mAP
MPEG7(AIR+PDM)	99.5	99.3	100	99.7	99.6
MPEG7(AIR+NSS)	97.6	94.7	99.9	98.8	98.2
MPEG7(AIR+DP)	97.6	94.8	99.9	98.8	98.2
MPEG7(AIR+RLSim)	96.1	93.2	98.9	97.8	96.2
MPEG7(AIR+TPG)	96.5	94.0	99.8	98.4	97.4
MPEG7(AIR+GT)	95.6	93.4	99.8	97.5	95.8
MPEG7(AIR+mKNN)	95.4	89.0	99.6	96.9	93.8
Baseline	91.1	87.5	93.0	95.2	89.4
MDS-CM-BOF+PDM	97.8	95.6	98.7	98.3	96.9
MDS-CM-BOF+NSS	99.3	94.0	98.2	98.6	96.4
MDS-CM-BOF+DP	97.8	94.6	97.9	98.2	96.4
MDS-CM-BOF+RLSim	99.2	96.3	97.7	98.9	97.3
MDS-CM-BOF+TPG	98.5	93.1	98.0	98.0	95.8
MDS-CM-BOF+GT	99.8	94.9	98.5	98.8	96.9
MDS-CM-BOF+mKNN	99.7	95.1	99.2	98.9	97.2
Baseline	99.5	91.3	96.9	98.2	95.0

Table 8
Average time consumption of the proposed approach (in seconds).

	MSC+ELDF	PA voting	Sum
NRW	0.851	0.568	1.419
McGill	0.736	0.479	1.215
SHREC10	0.774	0.454	1.228
SHREC11	0.792	0.421	1.213

References

[1] A. Agathos, I. Pratikakis, P. Papadakis, S. Perantonis, P. Azariadis, N. Sapidis, Retrieval of 3d articulated objects using a graph-based representation, in: Proceedings of the 2nd Eurographics Conference on 3D Object Retrieval, Eurographics Association, 2009, pp. 29–36.

[2] M. Aubry, U. Schlickewei, D. Cremers, The wave kernel signature: A quantum mechanical approach to shape analysis, in: Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV Workshops), IEEE, 2011, pp. 1626–1633.

[3] X. Bai, S. Bai, X. Wang, Beyond diffusion process: Neighbor set similarity for fast re-ranking, Inf. Sci. 325 (2015) 342–354.

[4] X. Bai, S. Bai, Z. Zhu, L.J. Latecki, 3d shape matching via two layer coding, Pattern Anal. Mach. Intell. IEEE Trans. 37 (12) (2015) 2361–2373.

[5] X. Bai, X. Yang, L.J. Latecki, W. Liu, Z. Tu, Learning context-sensitive shape similarity by graph transduction, Pattern Anal. Mach. Intell. IEEE Trans. 32 (5) (2010) 861–874.

[6] V. Barra, S. Biasotti, 3d shape retrieval using kernels on extended reeb graphs, Pattern Recognit. 46 (11) (2013) 2985–2999.

[7] A.M. Bronstein, M.M. Bronstein, L.J. Guibas, M. Ovsjanikov, Shape google: Geometric words and expressions for invariant shape retrieval, ACM Trans. Graph. (TOG) 30 (1) (2011) 1.

[8] A.M. Bronstein, M.M. Bronstein, R. Kimmel, Efficient computation of isometry-invariant distances between surfaces, SIAM J. Sci. Comput. 28 (5) (2006) 1812–1836.

[9] M.M. Bronstein, A.M. Bronstein, Shape recognition with spectral distances, Pattern Anal. Mach. Intell. IEEE Trans. 33 (5) (2011) 1065–1071.

[10] M.M. Bronstein, I. Kokkinos, Scale-invariant heat kernel signatures for non–rigid shape recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2010, pp. 1704–1711.

[11] R.R. Coifman, S. Lafon, Diffusion maps, Appl Comput. Harmon. Anal. 21 (1) (2006) 5–30.

[12] S. Dirk, H. Jeroen, V. Dirk, S. Paul, Isometric deformation invariant 3d shape recognition, Pattern Recognit. 45 (2012) 2817–2831.

[13] M. Donoser, H. Bischof, Diffusion processes for retrieval revisited, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1320–1327.

[14] B.J. Frey, D. Dueck, Clustering by passing messages between data points, Science 315 (5814) (2007) 972–976.

[15] J. Jiang, B. Wang, Z. Tu, Unsupervised metric learning by self-smoothing operator, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), IEEE, 2011, pp. 794–801.

[16] P. Kontschieder, M. Donoser, H. Bischof, Beyond pairwise shape similarity analysis, in: Computer Vision–ACCV 2009, Springer, 2010, pp. 655–666.

[17] Z. Kuang, Z. Li, X. Jiang, Y. Liu, H. Li, Retrieval of non-rigid 3d shapes from multiple aspects, Comput. Aided Des. 58 (2015) 13–23.

[18] Z. Kuang, Z. Li, Q. Lv, T. Weiwei, Y. Liu, Modal function transformation for isometric 3d shape representation, Comput. Graph. 46 (2015) 209–220.

[19] H. Laga, T. Schreck, A. Ferreira, A. Godil, I. Pratikakis, et al., Bag of words and local spectral descriptor for 3d partial shape retrieval, in: Proceedings of the Eurographics Workshop on 3D Object Retrieval (3DOR11), Citeseer, 2011, pp. 41–48.

[20] G. Lavoué, Combination of bag-of-words descriptors for robust partial shape retrieval, Vis. Comput. 28 (9) (2012) 931–942.

[21] B. Lévy, Laplace–Beltrami eigenfunctions towards an algorithm that "understands" geometry, in: Proceedings of the IEEE International Conference on Shape Modeling and Applications, SMI, IEEE, 2006. 13–13

[22] B. Li, A. Godil, H. Johan, Hybrid shape descriptor and meta similarity generation for non-rigid and partial 3d model retrieval, Multimed. Tools Appl. (2013) 1–30.

[23] Z. Li, Z. Kuang, Y. Liu, J. Wang, Multiscale shape context and re-ranking for deformable shape retrieval, Comput. Graph. 54 (2016) 8–17.

[24] Z. Lian, A. Godil, B. Bustos, M. Daoudi, J. Hermans, S. Kawamura, Y. Kurita, G. Lavoué, H. Van Nguyen, R. Ohbuchi, et al., A comparison of methods for non-rigid 3d shape retrieval, Pattern Recognit. 46 (1) (2013) 449–461.

[25] Z. Lian, A. Godil, T. Fabry, T. Furuya, J. Hermans, R. Ohbuchi, C. Shu, D. Smeets, P. Suetens, D. Vandermeulen, et al., Shrec10 track: non-rigid 3d shape retrieval, in: Proceedings of the Eurographics Workshop on 3D Object Retrieval, 2010, pp. 1–8.

[26] Z. Lian, A. Godil, X. Sun, H. Zhang, Non-rigid 3d shape retrieval using multidimensional scaling and bag-of-features, in: Proceedings of the 17th IEEE International Conference on Image Processing (ICIP), IEEE, 2010, pp. 3181–3184.

[27] L. Liu, L. Wang, X. Liu, In defense of soft-assignment coding, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), IEEE, 2011, pp. 2486–2493.

[28] C. Maes, T. Fabry, J. Keustermans, D. Smeets, P. Suetens, D. Vandermeulen, Feature detection on 3d face surfaces for pose normalisation and recognition, in: Proceedings of the Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS), 2010, pp. 1–6.

[29] L. Nanni, A. Lumini, S. Brahnam, Weighted fusion of shape descriptor for robust shape classification, Horiz. Comput. Sci. Res. 11 (2015).

[30] C.Y. Okada, D.C.G. Pedronette, R. da S Torres, Unsupervised distance learning by rank correlation measures for image retrieval, in: Proceedings of the 5th ACM on International Conference on Multimedia Retrieval, ACM, 2015, pp. 331–338.

[31] R. Osada, T. Funkhouser, B. Chazelle, D. Dobkin, Shape distributions, ACM Trans. Graph. (TOG) 21 (4) (2002) 807–832.

[32] P. Papadakis, Enhanced pose normalization and matching of non-rigid objects based on support vector machine modelling, Pattern Recognit. 47 (1) (2014) 216–227.

[33] P. Papadakis, I. Pratikakis, T. Theoharis, G. Passalis, S. Perantonis, et al., 3d object retrieval using an efficient and compact hybrid shape descriptor, in: Proceedings of the Eurographics Workshop on 3D Object retrieval, 2008.

[34] D.C.G. Pedronette, R.d. S. Torres, Image re-ranking and rank aggregation based on similarity of ranked lists, Pattern Recognit. 46 (8) (2013) 2350–2360.

[35] J. Philbin, O. Chum, M. Isard, J. Sivic, A. Zisserman, Lost in quantization: Improving particular object retrieval in large scale image databases, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR, IEEE, 2008, pp. 1–8.

[36] H. Qiu, E.R. Hancock, Clustering and embedding using commute times, Pattern Anal. Mach. Intell. IEEE Trans. 29 (11) (2007) 1873–1890.

[37] B. Ramesh, C. Xiang, T.H. Lee, Shape classification using invariant features and contextual information in the bag-of-words model, Pattern Recognit. 48 (3) (2015) 894–906.

[38] R.M. Rustamov, Laplace–Beltrami eigenfunctions for deformation invariant shape representation, in: Proceedings of the fifth Eurographics symposium on Geometry processing, Eurographics Association, 2007, pp. 225–233.

[39] P. Shilane, P. Min, M. Kazhdan, T. Funkhouser, The princeton shape benchmark, in: Proceedings of the Shape modeling applications., IEEE, 2004, pp. 167–178.
[40] J. Sun, M. Ovsjanikov, L. Guibas, A concise and provably informative multi-scale signature based on heat diffusion, Comput. Graph. Forum 28 (5) (2009) 1383–1392.
[41] H. Tabia, H. Laga, D. Picard, P.-H. Gosselin, Covariance descriptors for 3d shape matching and retrieval, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2014, pp. 4185–4192.
[42] H. Tabia, D. Picard, H. Laga, P.-H. Gosselin, Compact vectors of locally aggregated tensors for 3d shape retrieval, in: Proceedings of the Sixth Eurographics Workshop on 3D Object Retrieval, Eurographics Association, 2013, pp. 17–24.
[43] J.B. Tenenbaum, V. De Silva, J.C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (5500) (2000) 2319–2323.
[44] R. Toldo, U. Castellani, A. Fusiello, The bag of words approach for retrieval and categorization of 3d objects, Vis. Comput. 26 (10) (2010) 1257–1268.
[45] B. Wang, J. Jiang, W. Wang, Z.-H. Zhou, Z. Tu, Unsupervised metric fusion by cross diffusion, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2012, pp. 2997–3004.
[46] B. Wang, Z. Tu, Affinity learning via self-diffusion for image segmentation and clustering, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2012, pp. 2312–2319.
[47] J. Wang, Y. Li, X. Bai, Y. Zhang, C. Wang, N. Tang, Learning context-sensitive similarity by shortest path propagation, Pattern Recognit. 44 (10) (2011) 2367–2374.
[48] X. Wang, B. Feng, X. Bai, W. Liu, L.J. Latecki, Bag of contour fragments for robust shape classification, Pattern Recognit. 47 (6) (2014) 2116–2125.
[49] X. Yang, S. Koknar-Tezel, L.J. Latecki, Locally constrained diffusion process on locally densified distance spaces with applications to shape retrieval, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2009, pp. 357–364.
[50] X. Yang, L. Prasad, L.J. Latecki, Affinity learning with diffusion on tensor product graph, Pattern Anal. Mach. Intell. IEEE Trans. 35 (1) (2013) 28–38.