CrossMark

2D and 3D shape retrieval using skeleton filling rate

Yahya Sirin[1] · M. Fatih Demirci[1]

Abstract As an increasing number of digital images are generated, a demand for an efficient and effective image retrieval mechanisms grows. In this work, we present a new skeleton-based algorithm for 2D and 3D shape retrieval. The algorithm starts by drawing circles (spheres for 3D) of increasing radius around skeletons. Since each skeleton corresponds to the center of a maximally inscribed circle (sphere), this process results in circles (spheres) that are partially inside the shape. Computing the ratio between pixels that lie within the shape and the total number of pixels allows us to distinguish shapes with similar skeletons. Experimental evaluation of the proposed approach including a comprehensive comparison with the previous techniques demonstrates both effectiveness and robustness of our algorithm for shape retrieval using several 2D and 3D datasets.

Keywords Shape recognition · Shape retrieval · Earth mover's distance · 2D and 3D skeleton

1 Introduction

With recent improvements in computer science and technology, digital images have been popularly used in many applications, such as multimedia entertainment, computer-aided design, digital library, and electronic commerce. As the size of both 2D and 3D shapes increases, their effective and efficient retrieval becomes vital, attracting many researchers

✉ M. Fatih Demirci
mfdemirci@etu.edu.tr

Yahya Sirin
ysirin@etu.edu.tr

[1] Department of Computer Engineering, TOBB University of Economics and Technology, Söğütözü, Ankara 06560, Turkey

from several fields including pattern recognition [64, 101], computer graphics [29, 63], computer vision [50, 71], applied mathematics [57, 91], and event detection [11–13].

It has commonly been assumed that 3D shape retrieval techniques are classified into three categories: geometry-based, view-based and hybrid techniques [91, 101]. This classification is based on whether shape features are extracted from the shape geometry, shape view or both. Geometry-based retrieval techniques only exploit shape geometry and topology information, whereas view-based methods use image features, and spatial maps [91]. However these two classes are not completely disjoint and several hybrid methods have been developed to use both categories [91]. 2D shape retrieval approaches, on the other hand, are classified as either contour-based or region-based, depending on the way features are extracted from the contour or the shape region.

A considerable amount of literature, published on both 2D and 3D shape retrieval, clearly demonstrates that there are still major problems to overcome. Insufficient performance of the shape retrieval techniques, not being invariant to translation, rotation or scaling, and not dealing properly with part matching in which the goal is to find whether the shape to be matched is a part of another shape or vice versa are only a few examples for these problems. Here, insufficient performance of the matching algorithm has been faced especially in the case of interactive user query requests, where the priorities of shape similarities change depending on the user preference and application behavior [49]. Although major differences between input shapes are mainly used in shape retrieval approaches, minor differences may be more important for some matching scenarios. For example, the number of small holes on a CAD model carries important structure for technical drawing databases. Thus, finding a model with the same number of holes is more important than the overall appearance similarity, increasing the difficulty of shape retrieval.

As observed from prior studies [57, 91], 2D and 3D shape retrieval approaches are not directly applicable to each other. While viewing angle, illumination, and image acquisition play a critical role for 2D shape retrieval, they have little or no effect for 3D. On the other hand, when 2D shape retrieval approaches are modified to apply for 3D, their computational complexity increases due to the larger size of 3D models [17]. When a 3D shape is projected onto the 2D plane, one dimension of the shape and its physical information are lost. Figure 1 presents an example, where a 3D model is shown differently in 2D depending on the viewing angle. As a result of these problems, researchers mostly develop their shape retrieval algorithms to apply for either 2D or 3D only. In contrast, we propose a 2D shape retrieval algorithm, which can easily be extended for 3D in this paper.

Many 2D and 3D shape representation and description techniques have been proposed in recent years. Skeleton is a geometry-based shape representation technique used in both

Fig. 1 3D shape seen differently in 2D space according to various angles
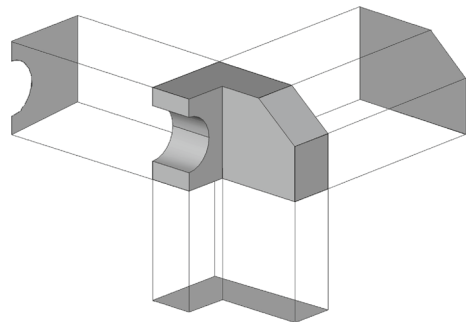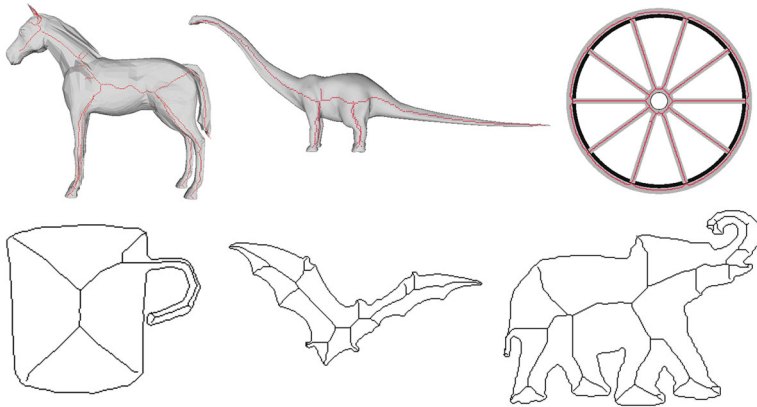
Fig. 2 Sample 3D and 2D skeletons superimposed on their shapes are shown in top and bottom, respectively

2D and 3D [91] and known as the core of the shape [89]. Because of its compact and powerful representation, skeleton has been used in many computer vision applications such as character recognition, image analysis, digital image processing, fingerprint recognition, visual inspection, and binary image compression. Figure 2 shows 2D and 3D skeletons superimposed on their shapes. Skeleton is related to the medial-axis and medial-surface [92] for 2D and 3D shapes, respectively. Specifically, the skeleton of a 2D shape forms the centers of the maximal disks inside the shape boundary, and the radii of these maximal disks represent the thickness of a shape (Fig. 3). Similarly for 3D shapes, skeleton is defined as the centers of the maximal spheres inside the shape surface, and the radii of these maximal spheres represent the thickness of the shape. Given the radius of such discs (or, spheres) associated with each skeleton, 2D and 3D shape can be reconstructed exactly. It is well-known that the skeletonization process is sensitive to minor boundary changes, but this
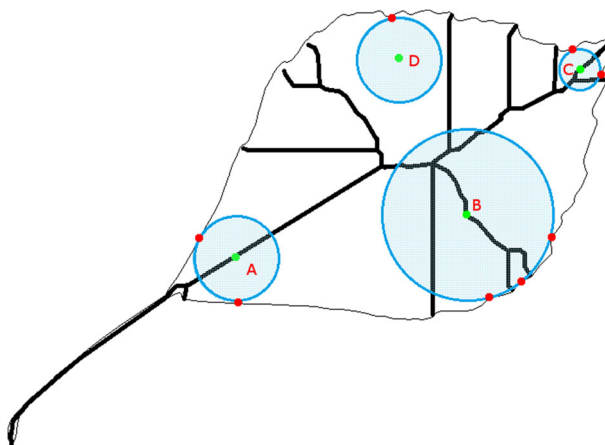
Fig. 3 The skeleton is defined by the loci of centers of maximum discs that touch the shape border in a least two distinct points and that fit entirely within shape. Thus, while points *A*, *B*, and *C* form skeletons, *D* does not

sensitivity is reduced by preprocessing of boundaries or surfaces, or post processing of extracted skeletons [77].

In this paper we propose a novel skeleton-based shape retrieval algorithm, which can be used for both 2D and 3D. The algorithm starts by drawing circles (spheres for 3D) of increasing radius around skeletons. Since each skeleton corresponds to the center of a maximally inscribed circle (sphere), this process results in circles (spheres) that are partially inside the shape. Computing the ratio between pixels that lie within the shape and the total number of pixels in each circle (sphere) enables us to find out the overall shape of the local region. As a result, this process distinguishes regions in which the skeletons are associated with same-sized circles (spheres) but their shapes are different. We use the Earth Mover's Distance (EMD) [70] to compute the similarity between two sets of skeletons associated with this new information. Experimental evaluation of the proposed approach including a comprehensive comparison with alternative techniques in 2D and 3D demonstrates the effectiveness and robustness of the proposed algorithm for shape retrieval using several datasets.

The main contributions of the proposed work are summarized as follows. First, the skeleton representation of a shape in 2D and 3D is enhanced by drawing circles (spheres) of increasing radius. To the best of our knowledge, the proposed framework presents the first technique that enriches the skeleton representation of a shape in this fashion. Second, we show how to use this representation effectively for retrieving similar database shapes. Third, we perform extensive experimentation using five 2D and two 3D datasets of different modality and order of magnitude. The proposed approach outperforms most of the existing techniques and yields comparable results to the best performing algorithms on these datasets. Our preliminary work on 2D shape retrieval is presented in [86].

The remainder of the paper is organized as follows: After providing a summary of the related work in Section 2, we discuss the details of the proposed algorithm in Section 3. Section 4 presents the experimental results for 2D and 3D shape retrieval. We present discussions regarding the overall framework in Section 5, and draw conclusions in Section 6.

2 Related work

3D shape retrieval techniques are broadly classified into three categories: geometry-based, view-based, and hybrid, while 2D shape retrieval techniques are grouped into contour-based and region-based methods. In this section, we review some of the related techniques in these categories. The reader is referred to [36, 101], and [91] for details.

Geometry-based techniques use the distribution of shape features to represent the geometric information. Significant advantages of this class consist of its ability for the partial matching, multi-scale shape representation, and articulated object matching. The feature extraction process adopted by these techniques is usually designed with two goals: having a strong discriminative ability and providing robustness to different geometric representations, such as meshes, volume data, point clouds, and range images. These features can be either local (e.g., probability density-based shape descriptor [1], morphological strategies [35], multi-scale signature [88]), global (e.g., shape distribution [62], skeleton graphs [37, 55], shape histogram [4], reeb graphs (MRG) [33], Extended Gaussian Image (EGI) [34], Complex Extended Gaussian Image(CEGI) [38]) or both (e.g., [10, 46, 72, 85]).

Among these approaches, shape distribution [62] is a powerful method using various geometric properties, e.g., distances between surface points, angles of random surface triples

to measure the similarity between 3D shapes. 3D Zernike moments [59], which are considered as the extensions of spherical harmonics based descriptors [27], form another powerful method used for 3D shape retrieval. These moments are computed as a projection of the function defining the shape onto a set of orthonormal functions within the unit ball. Ben-Chen and Gotsman [8] propose a 3D shape descriptor based on conformal geometry, which computes the amount of local work required to transform an object into a sphere. Shih and Chen [79] suggest an Angular Radial Transformation-based Elevation Descriptor (ART-ED), which identifies both external and internal information of a 3D shape. Furthermore, Bustos et al. [10] present a method based on combining the descriptors of different global and local features to improve the retrieval precision.

Graphs are powerful and flexible data structures allowing for the modeling of different data types. Graph-based approaches also belong to geometry-based techniques. These techniques use the graph topology for shape representation and employ some graph matching algorithm to calculate the similarity between a pair of graphs [17, 20, 33, 37, 89]. Although graphs are widely employed in different domains, efficient and effective graph matching tools remain to be developed for many modern applications.

In contrast to extracting the shape features directly, view-based techniques obtain a set of 2D views for a given 3D model for representation. A typical view-based technique is the Elevation Descriptor (ED) [80], which computes the minimum distance between six corresponding views (elevations) of two models, where each view is described by a set of concentric circular areas. A similar framework, Light Field Descriptor (LFD), is presented by Chen et al. [15, 78]. The approach considers two 3D models to be similar, if they look similar from all viewing angles. Thus, LFD compares silhouettes of the 3D shape obtained from ten viewing angles distributed evenly on the viewing sphere. The silhouettes are encoded by Zernike moments and Fourier descriptors. Frejlichowski [26] expands this technique by rendering several two-dimensional projections of a 3D model and by using the 2D Polar-Fourier transform for obtained projections. The resulting method is shown to be more effective than LFD for 3D shape retrieval. Dense sampling and fast encoding for 3D model retrieval [28] extracts SIFT features [56] to generate the bag-of-visual-features. The 3D model is then described using the distribution of visual words. Eitz et al. [25] propose a view-based descriptor for an efficient multi-sketch shape matching approach. The authors generate a set of 2D sketch-like drawings for each object. Axenopoulos et al. [6] also present a view-based approach for the accurate alignment of 3D objects using their 2D binary (black/white) silhouettes.

Hybrid techniques rely on both visual and geometric features of 3D shapes. DESIRE [94] is one such hybrid shape descriptor comprising three different descriptors: silhouette-based, depth buffer-based, and ray-based. DESIRE achieves superior performance over several famous view-based or geometry-based techniques such as Light Field [15] and Spherical harmonics [41]. Papadakis et al. [65] propose a generic strategy to improve the effectiveness of 3D retrieval systems by combining multiple shape descriptors. Another hybrid 3D shape descriptor named PANORAMA is presented in [66], which uses a set of 2D views of a 3D shape. The method obtains a panoramic view of a 3D shape by projecting it to three axis-aligned cylinders and unfolding the projection images into 2D images. The features for the panoramic views are extracted using Fourier and wavelet transforms. Leng and Xiong [45] show a hybrid shape descriptor named TUGE, which combines the depth buffer-based shape descriptor [95] and the GEDT shape descriptor [27]. The performance of TUGE is reported to be slightly better than DESIRE.

Depending on the way shape features are extracted from the contour or from the whole shape region, 2D shape representation and description techniques are classified as either contour-based or region-based. Although a number of such techniques in each group have been presented in the literature, these algorithms do not yet provide entirely satisfactory results [18]. In general, these algorithms are expected to meet the following principles; good retrieval performance, low computation complexity, robustness to noise, general applicability, and invariance to similarity transformation [42].

Shape context (SC) [7] is one of the most powerful 2D shape descriptors based on counter points. The algorithm starts by generating n discrete counter points. The set of $n-1$ vectors originating from point p_i to all other points represents the overall shape relative to p_i. This information is encoded in the spatial histogram called shape context. Although this algorithm has a good descriptive power, it does not perform well for shapes with articulation parts. To deal with this problem, inner distance shape context algorithm [53] has been proposed. Given a set of boundary points on the shape, this algorithm considers the shortest path between each point pair. The approach has been shown to be robust to shape deformation, however, its main shortcoming is its sensitivity to the number of boundary points, e.g., with low number of such points its performance drops dramatically [32].

Fourier descriptor is another most extensively used shape description method employed by many approaches in the literature, e.g., [14, 39, 93]. Fourier descriptor-based techniques describe both closed curve and partial shapes [51]. Granlund presents the Fourier invariants to describe the rotational symmetry of shapes [31]. Arbter et al. [5] develop the affine-invariant Fourier descriptor to extract affine invariant features. Although this descriptor possesses computational efficiency and attractive invariance properties, it can not be used with shapes consisting of several components. To deal with such shapes, Li et al. recently propose affine invariant ring Fourier descriptor, which computes a set of affine invariant closed curves from the object [48]. Rouber and Steiger-Garcao introduce UNL Fourier descriptor to describe disjointed or articulated shapes in the domain of handwritten character recognition [69]. Although powerful, its sensitivity for occlusion and the high dimensionality of the feature vector obtained by this technique are its two main shortcomings.

Guocheng et al. proposed a shape descriptor called Shape Filling Rate (SFR), which deals with the problems of describing part structure and articulation for shape recognition [32]. Specifically, SFR first performs equal space sampling on the shape contour points. For each contour point p, the algorithm draws circles centered at p and computes the ratio of the number of pixels that lie within the shape to the total number of pixels. The distance between two shapes is computed using χ^2 test statistics without taking into consideration the point coordinates. To improve the retrieval rate, the algorithm is combined with shape context and dynamic programming. Since the approach takes the same number of contour points for each database shape, its retrieval performance drops when some shapes are more complex than others. Our work is inspired by SFR. However, due to the improved performance of shape similarity techniques based on skeleton matching, we employ skeleton shape representation in our approach. In addition, the proposed framework does not require the same number of skeletons for each shape. Thus, shapes are not assumed to have similar complexities as in the case of SFR. Moreover, not only is the proposed approach applicable to 2D shape retrieval only, it can also be used for 3D shape retrieval as shown in this paper.

There is a rich literature for skeleton-based shape matching. To name a few, Liu and Geiger [54] presented a framework in which shape axis tree defined by the locus of midpoints for optimally corresponding boundary points are matched. Using graph topology changing operations may result in matches that do not preserve the coherence of the shape.

Sharvit et al. proposes a shape recognition framework based on shock graphs [74]. Although this method has shown promising results, the errors of fundamental flows can break the hierarchical relations among parts of the shape. In [83], shape recognition is achieved by solving the problem of subgraph isomorphism. Shock graphs are converted into rooted trees, which in turn are matched using a tree matching algorithm. Selection of the root plays an important role in matching, and thus different root selections may lead to improper node correspondences. In [73], the distance between two shapes is computed using the least action path deforming one shape to another. Since finding the optimal alignment of the shock graph has high complexity, the authors adopt a dynamic programming method.

3 Skeleton-based shape recognition

Our framework can be divided into two stages. The first stage associates each skeleton with information reflecting how the local region around the skeleton changes, while the second stage matches skeletons armed with this new information. Figure 4 presents the overview of the proposed approach. Our motivation for using skeletons rather than contour points in 2D comes from the studies showing that skeleton-based shape similarity descriptors perform better than contour-based ones even in the case of partial occlusions [73, 99]. Similarly, skeleton based 3D shape similarity descriptors achieve better retrieval performance than the techniques employing both topological and geometric features together [49, 90].

3.1 Enhanced skeleton representation

The proposed approach starts by computing the skeleton for a given shape. Each skeleton s_i is associated with its coordinate (x_i, y_i) (for 3D (x_i, y_i, z_i)) and the radius r_i of the maximal inscribed circle (sphere) that completely lies inside the shape, i.e., the radius of the maximal inscribed circle centered at s_i equals the minimal distance from s_i to the boundary (surface). For each skeleton s_i, we draw circles (spheres) centered at (x_i, y_i) (for 3D (x_i, y_i, z_i)) with
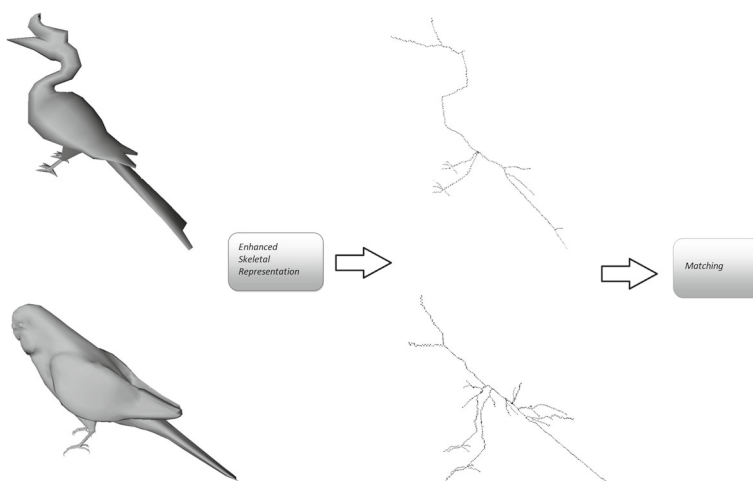
Fig. 4 Overview of the algorithm. After the input shapes are represented as a set of skeletons with new information, we compute the matching between them using EMD

increasing radius starting from r_i to the maximum radius of the skeletons inside the shape. This process results in circles (spheres), which are partially inside the shape. The portion of the pixels that are inside the shape allows us to learn the overall region around the skeleton and forms the basis of our framework (Fig. 5).

More specifically, for skeleton s_i of shape S, we enrich its representational power by adding a new component e_i defined as

$$e_i = \sum_{k=r_i}^{R+\epsilon} \frac{n_i^k}{N_i^k}, \tag{1}$$

where n_i^k is the number of pixels that lie inside the shape, N_i^k is the total number of pixels for circle (sphere) of radius k, R is the largest radius of the skeletons in shape S, and ϵ is a small positive constant. Using the proposed method, 2D and 3D skeletons are represented respectively as 4-dimensional (x, y, r, e) and 5-dimensional (x, y, z, r, e) vectors, forming a more discriminative shape representation than the original one. This new representation allows us to distinguish shapes associated with the skeletons of the same radius but with different e values. Figure 6 presents an example for 3D. Although each skeleton in each shape has the same radius, the value of e in each skeleton is different. Precisely, the radius of the skeleton at the horse leg, sand glass, and hammer are the same. However, when a sphere with larger radius around each skeleton is drawn, the amount of pixels that lie inside the shape differs, offering a more powerful representation. Figure 7 presents a similar scenario for 2D.

3.2 Skeleton matching

Given two shapes P and Q represented using the proposed method, i.e., for 2D

$$\begin{aligned} P &= \{(x_1, y_1, r_1, e_1), \dots, (x_n, y_n, r_n, e_n)\}, \\ Q &= \{(x_1, y_1, r_1, e_1), \dots, (x_m, y_m, r_m, e_m)\}, \end{aligned} \tag{2}$$

and similarly for 3D,

$$\begin{aligned} P &= \{(x_1, y_1, z_1, r_1, e_1), \dots, (x_n, y_n, z_n, r_n, e_n)\}, \\ Q &= \{(x_1, y_1, z_1, r_1, e_1), \dots, (x_m, y_m, z_m, r_m, e_m)\} \end{aligned} \tag{3}$$

we now proceed with the second stage of our algorithm, finding the distance between them. We compute the distance using the Earth Movers' Distance (EMD) algorithm [70], which has been successfully used in several applications, e.g., [2, 19–21, 47, 96, 100]. EMD finds
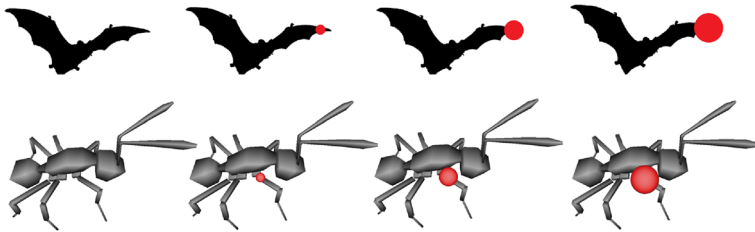
Fig. 5 After finding the largest radius of the skeletons of the input shape, we draw *circles* (*spheres* for 3D) of increasing radius around each skeleton. The portion of the pixels that are inside the shape allows us to learn the overall region around the corresponding skeleton. *Top* and *bottom* parts present this process for 2D and 3D shapes, respectively. The *left most image* shows the original shape, while *right images* show the *circles* (*spheres*) with increasing radii
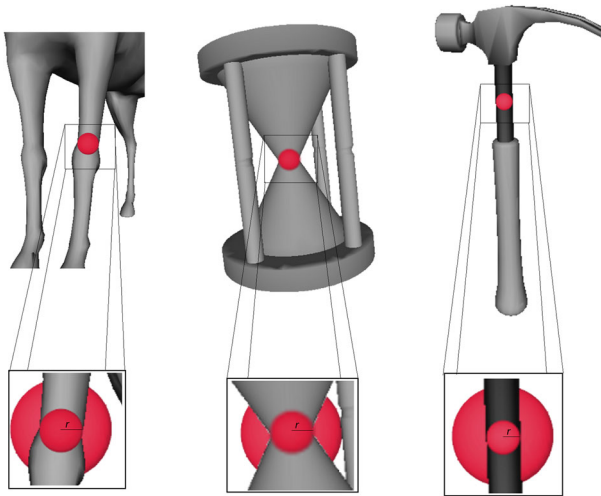
Fig. 6 The radius of the maximal inscribed sphere associated with each skeleton in each shape is the same. However, the skeletons have different e-values

the optimum match between two sets by computing the minimum amount of work needed to transform the first set into the other. In the case that two input sets have unequal total weight, EMD performs partial matching such that the minimum amount of work required to cover the mass in the lower-weight set with the mass in the higher-weight set.

Formally, let P and Q be the first and second sets with n and m points as defined before. Let $D = [d_{ij}]$ be the ground distance matrix, where d_{ij} is the ground distance between points $p_i \in P$ and $q_j \in Q$. Based on the work of Eberly who finds the distance between two points in the scale-space [23], we define ground distances $d_{2D}(p_i, q_j)$ for 2D and $d_{3D}(p_i, q_j)$ for 3D as follows:

$$d_{2D}(p_i, q_j) = \sqrt{\alpha_1(x_i - x_j)^2 + \alpha_1(y_i - y_j)^2 + \alpha_2(e_i - e_j)^2}, \qquad (4)$$

$$d_{3D}(p_i, q_j) = \sqrt{\alpha_1(x_i - x_j)^2 + \alpha_1(y_i - y_j)^2 + \alpha_1(z_i - z_j)^2 + \alpha_2(e_i - e_j)^2}, \quad (5)$$

where α_1 and α_2 are two parameters associated to point coordinates and weights, and have the role to control these elements during the matching. Since rotation and translation of a shape changes the positions of its skeletons but does not change the skeleton weights, in the experiments we set α_1 and α_2 to 0.1 and 0.9, respectively.
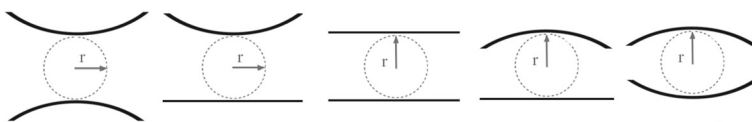
Fig. 7 Although each skeleton located at the center of each circle is associated with the same radius, drawing *circles* with increasing radius allows us to distinguish the skeletons

The objective of EMD is to compute a flow matrix $F = [f_{ij}]$, with f_{ij} being the flow between p_i and q_j, minimizing the overall distance:

$$\text{Work}(P, Q, F) = \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} d_{ij} \tag{6}$$

subject to:

$$f_{ij} \geq 0, \; 1 \leq i \leq m, \; 1 \leq j \leq n$$

$$\sum_{j=1}^{n} f_{ij} \leq w_{p_i}, \; 1 \leq i \leq m$$

$$\sum_{i=1}^{m} f_{ij} \leq w_{q_j}, \; 1 \leq j \leq n \tag{7}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = \min \left(\sum_{i=1}^{m} w_{p_i}, \sum_{j=1}^{n} w_{q_j} \right).$$

The w associated with each point is the weight, and we take this as the e-value in our framework. These constraints ensure that flows from P to Q are non-negative and unidirectional, the total amount of flow sent by each element in P and the total amount of flow received by each element in Q are limited to their weights, and the amount of flow sent from P to Q is maximum. Since the computation of EMD is exponential in the number of features per input set, performing shape retrieval on large datasets becomes a critical issue. Pele and Werman present the fast-EMD, which makes it possible to compute EMD on large databases using thresholded ground distances [68]. Their experiments show that using the fast-EMD the running time decreases by an order of magnitude for image retrieval. We used this EMD version in our experiments.

The original definition of EMD was extended in [16] to account for transformation. When the transformation T is applied to the second set, distances d_{ij}^T are defined as $d_{ij}^T = d(p_i, T(q_j))$, and thus, the objective function becomes $\text{Work}(P, Q, F, T) = \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} d_{ij}^T$. Using this extended version of EMD, the proposed approach is applicable to cases where shapes have undergone transformation.

4 Experiments

This part shows the experiments that we have performed to evaluate the proposed framework. First, we present our tests on the 2D shape retrieval and demonstrate the robustness of the proposed framework against noise and occlusion. We then present the 3D retrieval experiments that we have conducted to illustrate the effectiveness of our approach. Finally, we present the efficiency experiments measuring the time it takes to compute the distance between a pair of shapes in 2D and 3D. All experiments are conducted using Matlab Builder JA and Java on Windows-7.

4.1 2D shape retrieval

We used several 2D datasets for the experiments and employed different performance measures to assess the quality of the proposed shape retrieval technique along with alternative approaches. We first consider two standard databases for 2D shape retrieval, Kimia-99 and Kimia-216. Kimia-99 database consists of 99 shapes, grouped in 9 classes with 11 shapes
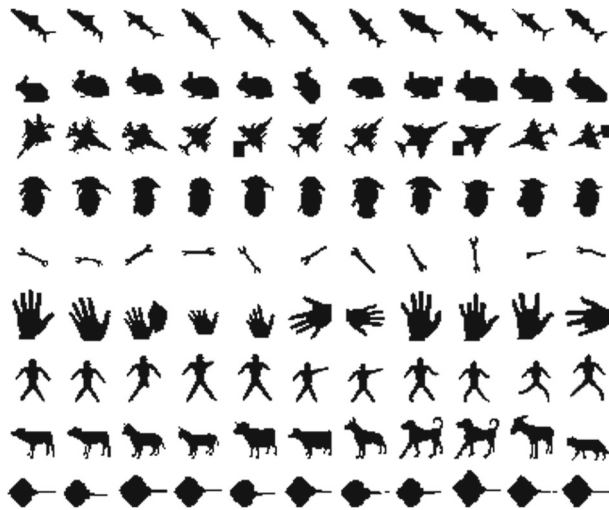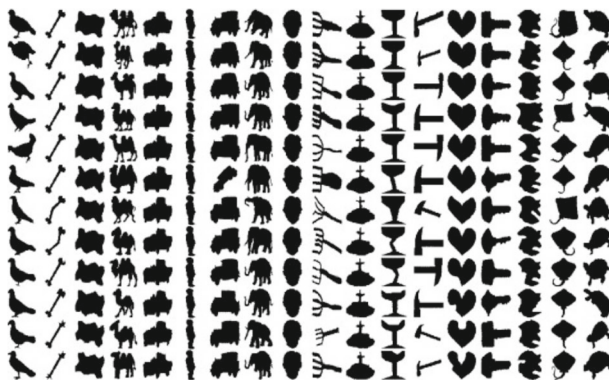
Fig. 8 Kimia-99 database, consisting of 9 classes with 11 shapes per class. Each row shows a different class

per class, whereas Kimia-216 database consists of 18 classes with 12 shapes per class. Figures 8 and 9 present Kimia-99 and Kimia-216 datasets, respectively.

We use leave-one-out cross validation to these datasets. Namely, the first shape from the database is removed and used as a query for the remaining database shapes. After the query shape is put back in the database, the procedure is repeated with the second shape from the database, etc., until all database shapes have been used as queries. Given a query view of an object class, when the recognition algorithm returns another view of the same class as its nearest neighbor, we classify this as a correct object recognition.

We compare the performance of the proposed framework against the original skeleton representations, where each skeleton is associated with its coordinate and the radius of the maximal inscribed circle. During the EMD stage, the weight of each point is taken as this radius, whereas for our algorithm the weight is computed as the $e-$value. In addition, we have implemented the SFR approach [32] according to its description. Based on


Fig. 9 Kimia-216 database, consisting of 18 classes with 12 shapes per class. Different views of the same objects appear in columns

the overall matching statistics, we observe that the original skeleton representation, SFR, and the proposed method obtain 85.3, 87.2, and 95.1 % recognition rates, respectively for Kimia-99. Repeating the same experiment for Kimia-216, our framework achieves 94.5 %, while the original skeleton representation and SFR results in 83.7 and 84.2 % recognition rates. The recognition scores of these methods along with some previous shape retrieval techniques, which use these datasets are reported in Fig. 10. One may observe that the proposed method outperforms both the skeleton method and SFR and is competitive to best performing approaches on these datasets.

As mentioned before, the main limitation of SFR comes from the fact that the algorithm takes the same number of contour points for each database shape. When sufficient points are selected for simple shapes only, more complex shapes are represented poorly, decreasing the retrieval rate. To show this more clearly, we have created a subset of MPEG-7 dataset, which consists of 1400 silhouette shapes of 70 classes with 20 shapes in each class. Our subset contains 10 classes with 10 shapes per class (see Fig. 11). For this dataset, the recognition scores for the original skeleton representation and SFR are obtained as 87.0 and 79.0 %, while the proposed framework yields 96.0 % recognition score. We should note that in the experimental evaluation of SFR, the authors combine their approach with shape context and dynamic programing to improve its performance. In our experiments, we have not combined these approaches, as we have implemented SFR according to its description only. However, we expect that performing such a unification with some matching algorithm for both original skeleton representation and the proposed method will improve the retrieval rates as well.

In our next experiment we use the whole MPEG-7 dataset, which has been used extensively to assess the performance of shape retrieval algorithms. For this dataset, the usual
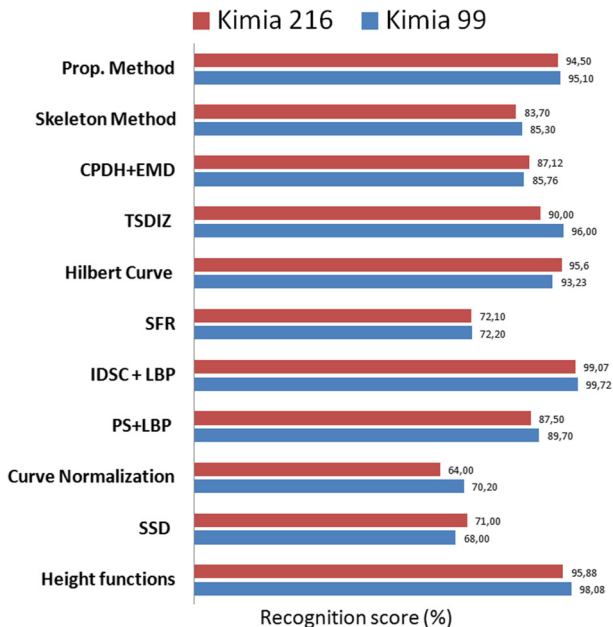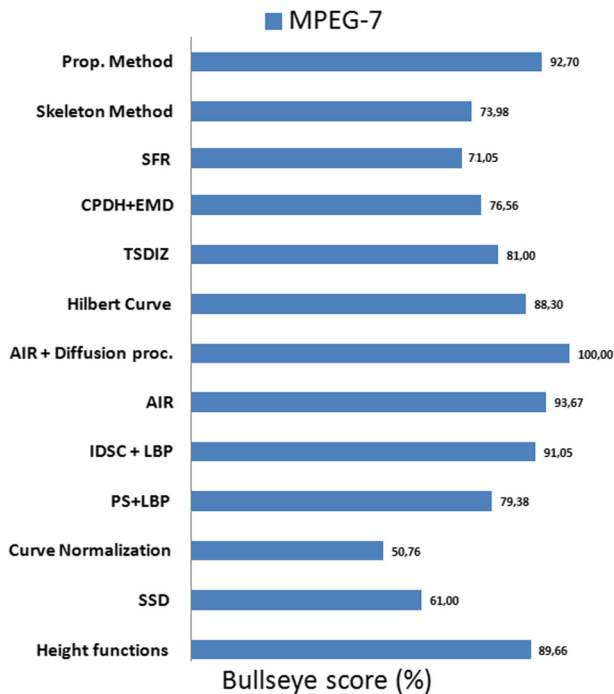

Fig. 10 Recognition scores for Kimia 99 and Kimia 216 databases of several methods: Height functions [97], SSD [67], Curve Normalization [44], PS+LBP [75], IDSC + LBP [76], SFR [32], Hilbert Curve [24], TSDIZ [3], CPDH+EMD [82], and Curve Normalization [44]
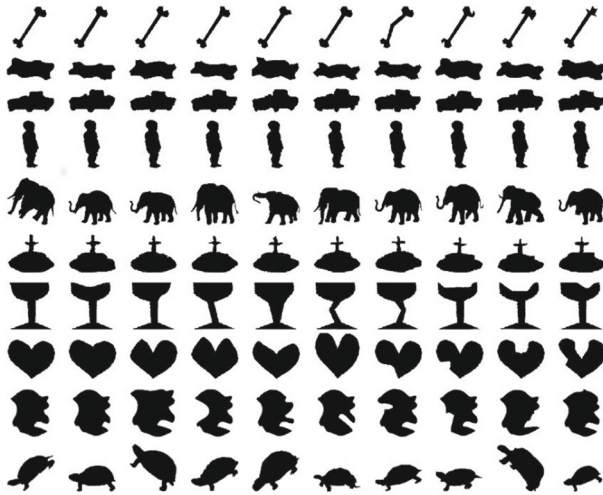
Fig. 11 A subset of MPEG-7 database, consisting of 10 classes with 10 shapes per class

metric employed by many previous works is the bullseye score, which rates how many of the 20 examples of the correct class appear in the top 40 matches. The result of the


Fig. 12 Bullseye scores for MPEG-7 of several methods: Height functions [97], SSD [67], Curve Normalization [44], PS+LBP [75], IDSC + LBP [76], AIR [30], AIR+Diffusion process [22], Hilbert Curve [24], TSDIZ [3], CPDH+EMD [82], and and SFR [32]

Fig. 13 The fifteen different leaf species in the Swedish leaf data set

proposed framework along with other approaches is shown in Fig. 12 and reveal that our approach is comparable to the state-of-the-art methods and even outperform most of them. The method [22] reports the best possible bullseye score by uniting the advantages of different works, where a local affinity matrix with K-nearest neighbors is used. However, selecting a reasonable K value is quite important to obtain good performance, limiting its practical applications [43]. On the other hand, the simplicity of our approach is an important advantage, which makes it suitable to be used on a range of different applications, in which shapes can be represented as a set of skeletons.

In addition to Kimia and MPEG-7 datasets, we also measure the efficacy of the proposed framework on two real datasets, Swedish leaf [87] and tools [9]. The Swedish leaf dataset
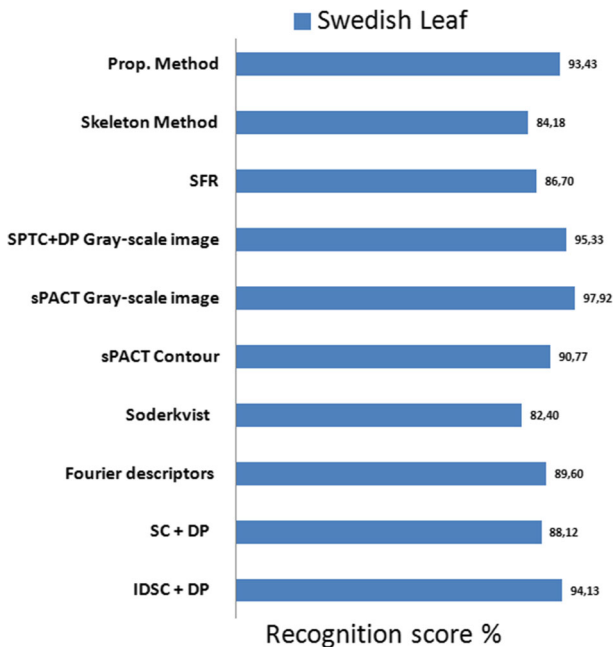
Fig. 14 Recognition scores for the Swedish leaf database of several methods: IDSC+DP [52], SC+DP [52], Fourier descriptors [52], Soderkvist [87], sPACT [98], SPTC+DP [53], and SFR [32]

has collected from 15 different Swedish tree species with 75 leaves per species. Figure 13 presents examples from these 15 leaf species. The result of the leave-one-out experiment of the proposed approach along with some previous shape retrieval techniques is shown in Fig. 14. The tools dataset, on the other hand, contains 35 articulated shapes: 10 scissors, 15 pliers, 5 knives, and 5 pincers. The whole dataset is shown in Fig. 15. We report the bullseye score of our approach and previous methods in Fig. 16. The results of the experiments on these two datasets are consistent with the previous shape retrieval experiments. While for the Swedish leaf dataset our method yields one of the best scores, it outperforms the existing techniques for the tools database. We should note that in the Swedish leaf dataset, the best three performing approaches are either use the grayscale leaf pictures or combine two techniques into one framework. Since grayscale images carry more powerful distinctive features than silhouettes, the recognition rate for sPACT [98] using contour is significantly improved when it uses grayscale. On the other hand, when only silhouettes are available, the proposed framework obtains the second top score.

To evaluate the sensitivity of our algorithm to noise, we perturbed each query in the Kimia-216 dataset by randomly deleting its skeletons whose size was chosen randomly to fall between 1 % and k % of the total number of skeletons. The value of k was systematically increased from 1 to 99 %. The same experimental setup was then used with the perturbed queries for this database. For $k = 20$ %, $k = 50$ %, and $k = 80$ %, the nearest neighbor retrieval score of our approach was dropped around 3.4, 14.5, and 24.4 %, respectively. The results for all values of k shown in Fig. 17 present the algorithm's robustness to missing data.
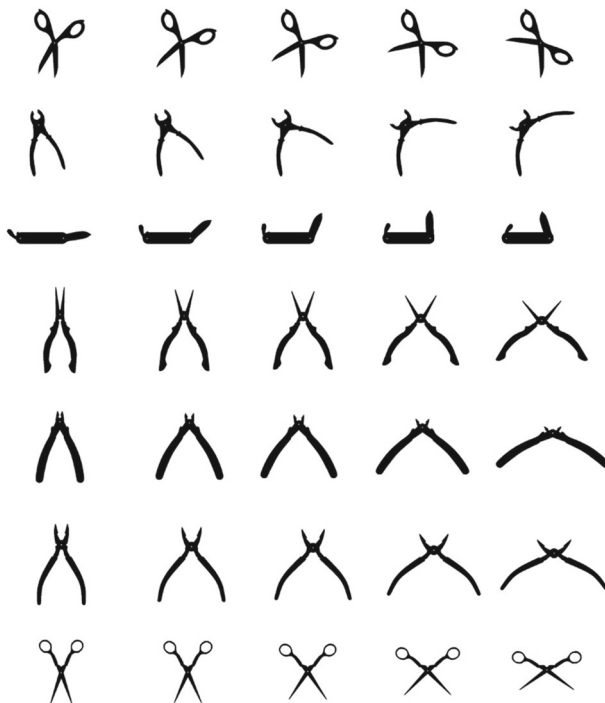

Fig. 15 The tools dataset contains a total of 35 articulated shapes from 4 different classes

Finally, to evaluate the fitness of our approach for dealing with occlusion, we generated a set of occluded scenes, each with two shapes selected from different classes in the Kimia-216 dataset. Each of these occluded scenes was used as a query against the complete Kimia-216 dataset. We define our recognition schema to be effective if one of the query shapes appears in the nearest neighbors. The results are presented in Fig. 18, where the left column shows the occluded query and the top nine nearest neighbors from left to right appear in each row. The correct retrievals are shown inside rectangles. We should point out that our recognition framework accounts for topological similarities. Thus, shapes from different classes but with similar topologies may be retrieved as nearest neighbors. For instance, while the hammer and bone, and camel and elephant are topologically similar, they are from different classes. Shapes belonging to different classes, therefore, may be ranked high in the nearest neighbor list.

4.2 3D shape retrieval

The 3D shape retrieval experiments are made using two well-known 3D databases: the Princeton Shape Benchmark (PSB) [81] and McGill Shape Benchmark (MSB) [84]. PSB contains a total of 1814 shapes, which is subdivided into a training set and a test set with equal size. Shapes belonging to the training set are classified into 90 classes, whereas those in the test set are grouped into 92 classes. MSB consists of 420 objects classified into 19 classes. Different from PSB, which only contains rigid shapes, MSB also has non-rigid shapes such as humans, snakes, pliers, octopuses, and spiders. Figures 19 and 20 present examples of 3D shapes for these datasets.
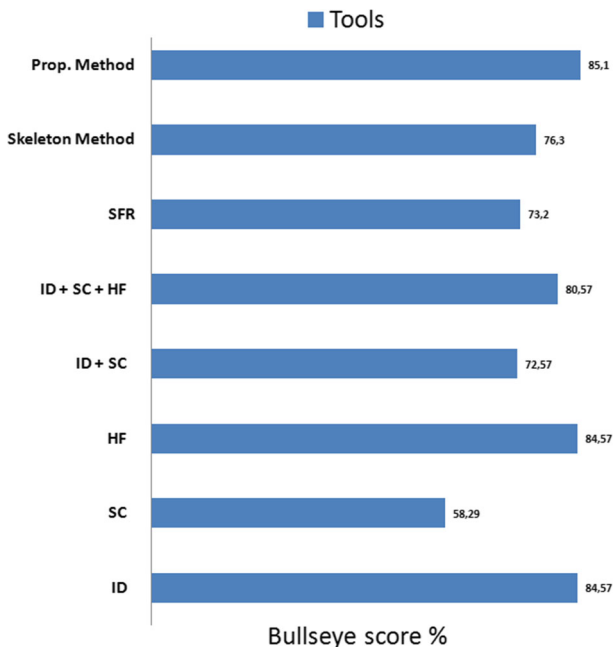

Fig. 16 Bullseye scores for the tools dataset of several methods: ID [53], SC [7], HF [97], ID+SC [58], ID+SC+HF [58], and SFR [32]
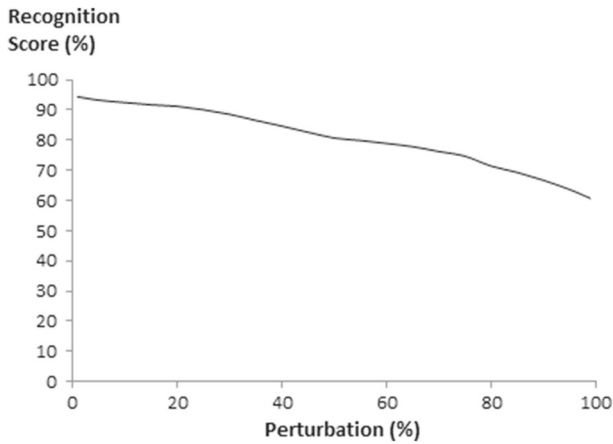
Fig. 17 Nearest neighbor recognition scores decrease gracefully as a function of increased perturbation

To quantify the performance for these datasets, we employ R-precision, which is given as a ratio of the shapes retrieved from the same class as the query in R closest matchings, where R is the size of the query class. In other words, R-precision is the recall for the top R retrievals. Similar to the 2D shape retrieval experiments, we first compare the performance of the proposed framework against the original skeleton representations, where each skeleton is associated with its coordinate in 3D and the radius of the maximal inscribed sphere. The R-precision scores of the proposed framework along with original skeleton representations and other approaches using the same datasets are shown in Fig. 21. The results for the MSB dataset demonstrate that the proposed framework outperforms all approaches except
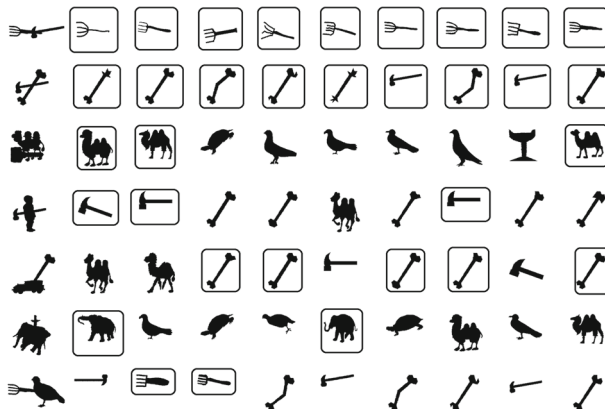
Fig. 18 Results of occlusion experiments where two images in each occluded scene are selected from different classes of the Kimia-216 dataset. The *leftmost column* shows the occluded query scenes for each row, while the top nine ranked nearest neighbors are shown on the right. A *shape inside a box* indicates that it belongs to the class of one of the query shapes. Shapes belonging to different classes but are topologically similar to the query are ranked high in the nearest neighbor list
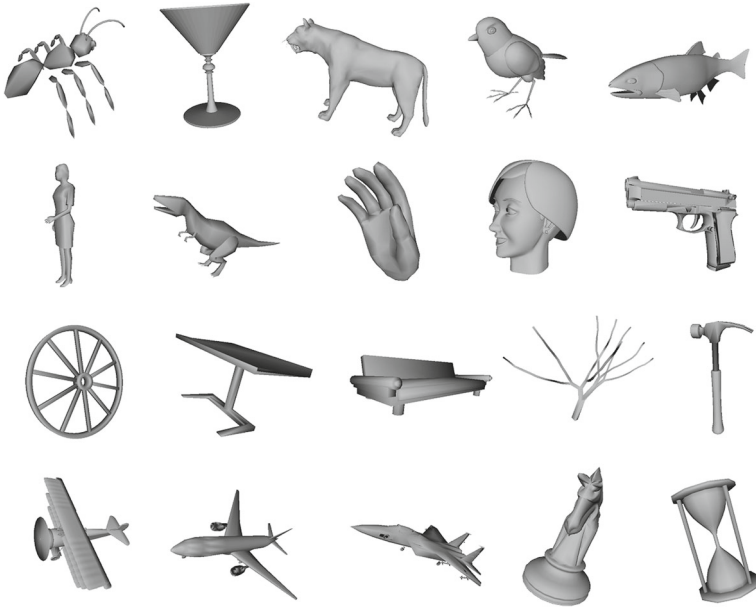
Fig. 19 Views of sample objects from the Princeton Shape Benchmark (PSB)

for the multi-scale version of Bag of Feature on Geodesic (BoFoG-M) [40], which produces a slightly better score. BoFoG-M combines the local geometric feature with its spatial context in a multi-scale notion. As mentioned before, performing a similar unification in our algorithm is likely to improve its performance. On the other hand, for PSB containing
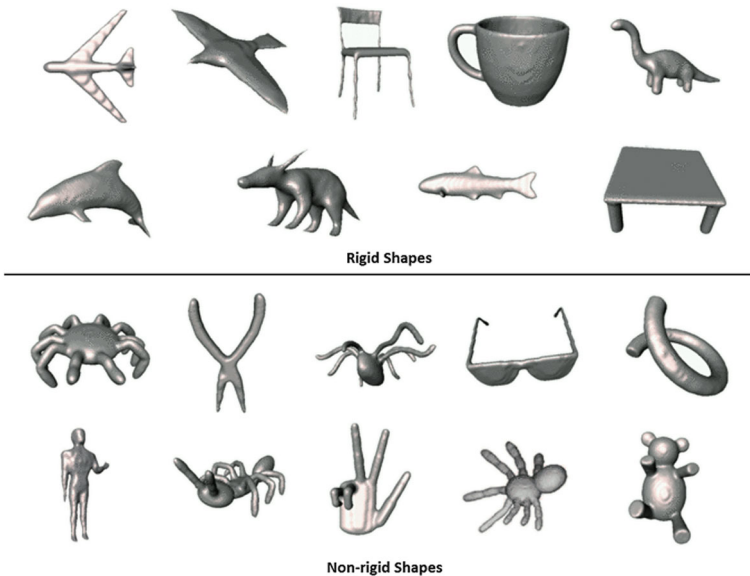
**Fig. 20** Sample views of rigid and non-rigid shapes from the McGill Shape Benchmark (MSB)
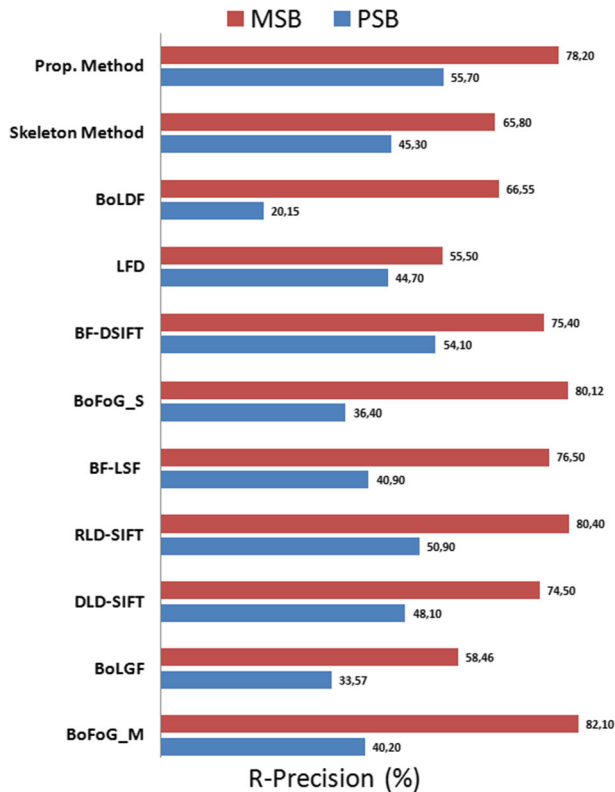
Fig. 21 R-precision scores for PSB and MSB datasets of several methods: Multi-scale version of Bag of Feature on Geodesic (BoFoG-M) [40], Bag of Local Geometry Feature (BoLGF) [40], Dense LD-SIFT (DLD-SIFT) [60], Resampled LD-SIFT (RLD-SIFT) [60], Linear Combination Local Statistical Features (BF-LSF) [61], Single-scale version of Bag of Feature on Geodesic (BoFoG-S) [40], Light Field Descriptor (LFD) [15], and Bag of Local Distance Feature (BoLDF) [40]

only rigid shapes, our framework yields the best R-precision ration. The descriptive power of enhanced skeleton representations and effectiveness of EMD for partial and complete matching are two main reasons for achieving the best or one of the best scores for these datasets.

In our next experiment, we perform a more comprehensive evaluation of the proposed method along with the alternative approaches. Instead of concentrating the recall for top R retrievals as before, the system's performance is evaluated by computing the total number of retrieved shapes that is necessary to retrieve the entire query class. To measure this, we report precision-recall curves for these datasets. Figures 22 and 23 present these curves for MSB and PSB datasets, respectively. As in the previous experiments, the proposed method is competitive to the best performing techniques for MSB and yields the best result for PSB. Upon investigation as to why our algorithm yields poorer performance on MSB, we found that the high similarity between different classes of non-rigid shapes (ants, crabs, spiders, octopuses) poses a negative effect on the quality of the results. We believe that enforcing spatial constraints during the matching stage or taking the advan-
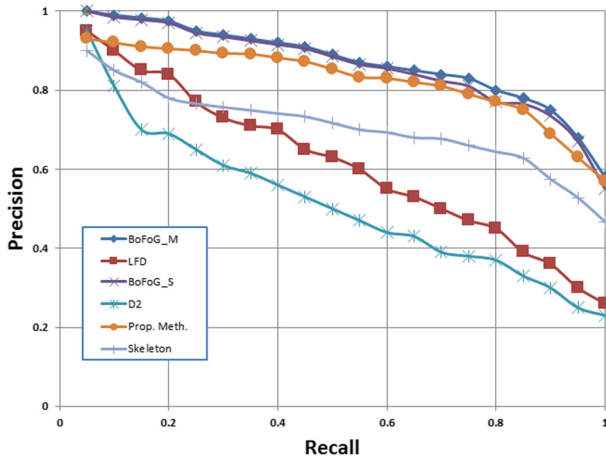
Fig. 22 Precision-recall plot for MSB of several methods: Multi-scale version of Bag of Feature on Geodesic (BoFoG-M), Single-scale version of Bag of Feature on Geodesic (BoFoG-S) [40], Light Field Descriptor (LFD) [15], D2 [63]

tage of graph representations would improve the performance of proposed framework in such cases.

4.3 Efficiency experiments

Our next set of experiments computes the overall efficiency of the proposed work applied for 2D and 3D shape retrieval. To do this, we first measure how much time it takes for our approach to find the distance between two shapes. On an Intel(R) Core(TM) i5-3317U CPU @ 1.70GHz computer with 4GB RAM, it takes about 0.7 sec to match a pair of shapes,
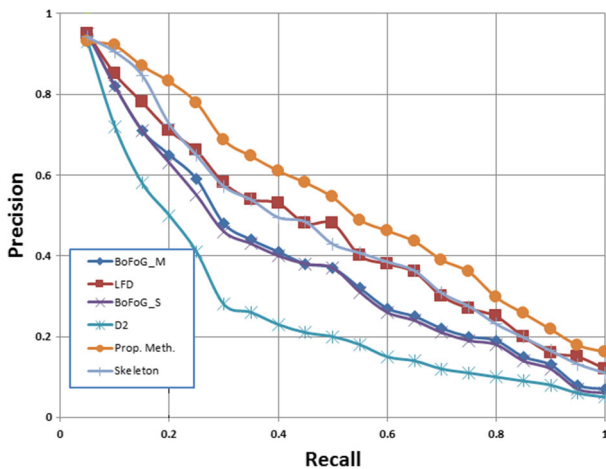

Fig. 23 Precision-recall plot for PSB of several methods: Multi-scale version of Bag of Feature on Geodesic (BoFoG-M), Single-scale version of Bag of Feature on Geodesic (BoFoG-S) [40], Light Field Descriptor (LFD) [15], D2 [63]

each with around 1000 skeletons. The preprocessing time spent for creating the $e-$value for one shape is recorded to be less than 0.05 sec. Since 3D shapes are represented with one extra dimension, the matching time and $e-$value creation time in 2D and 3D are almost the same. Our implementation of the SFR approach [32] and original skeleton representation also takes the same amount of time for matching a pair of shapes on the same computer. As mentioned before, we use the fast-EMD [68] in our experiments. To compute the improvement offered by the fast-EMD, we repeat the previous experiment with the original EMD [70]. According to the results, finding the distance between shapes in 2D and 3D increases to 11-15 seconds, yielding about 20 times slower matching.

5 Discussion

Our objective in this paper was to show the feasibility of using skeletons associated with new information for shape retrieval. Although a skeleton is not a good shape descriptor for all 2D and 3D shapes, it provides an accurate representation for many. This is evident by the retrieval scores of original skeletons as they are comparable to those of the existing methods. This performance is improved further by our algorithm, yielding one of the best shape retrieval scores on both 2D and 3D datasets used in the experiments.

Since the proposed method works with skeletons, input shapes are required to be correctly segmented from the background. Because correct segmentation is an unresolved problem for natural scenes, this requirement limits the applicability of our algorithm. However, there has been attempts in the literature for generating skeletons from natural images. Thus, the proposed framework may be extended to work with shapes in such images, posing an interesting research direction for the future.

We should note that EMD is a global distance that accounts for all the points of the input shapes. Although we have shown that our framework is robust to perturbation of the shapes in terms of missing skeletons, the method is still global. Thus, in case a point representing an occluder with large weight exists in an input shape, its presence will have a negative effect on shape retrieval results since the algorithm cannot selectively exclude the point. On the other hand, if there are some unique attributes shared by input points, they can be used in EMD to improve the part matching ability of our approach.

6 Conclusion

In this paper, we have presented a new skeleton-based 2D and 3D shape recognition algorithm. The algorithm works by drawing circles (spheres) of increasing radius around each skeleton starting from the radius of the maximal inscribed circle (sphere) which lies within the shape. Computing the ratio between pixels that lie within the shape and the total number of pixels in each circle (sphere) enables us to find out the overall shape of the local region. To the best of our knowledge, the proposed framework is the first framework that enriches the skeleton representation of a shape in this fashion. We have experimentally evaluated the method by using the Kimia-99, Kimia-216, a subset of MPEG-7, whole MPEG-7, Swedish leaf, and tools datasets for 2D and the McGill Shape Benchmark (MSB) and Princeton Shape Benchmark (PSB) for 3D shape retrievals. Experimental evaluation of the proposed approach in both 2D and 3D presents that adding a new component to each skeleton improves its representational power over the original skeleton representation and yields better retrieval scores than many existing techniques on several datasets.

Despite being simple and easily implemented, promising results have been obtained by the proposed algorithm. In the future, we would like to explore the approach further for a better retrieval performance. Specifically, we plan to improve the performance by imposing spatial constraints in the framework, taking the $e-$value as a histogram around each skeleton, using graph representations, computing hierarchical skeletons at different resolution, and employing different distribution-based matching algorithms. In addition, we also plan to explore methods to solve the issue of occlusion so that complex occluded shapes may be retrieved effectively.

References

1. Akgul CB, Sankur B, Yemez Y, Schmitt F (2009) 3d model retrieval using probability density-based shape descriptors. IEEE Trans Pattern Anal Mach Intell 31(6):1117–1133
2. Akimaliev M, Demirci MF (2015) Improving skeletal shape abstraction using multiple optimal solutions. Pattern Recogn 48(11):3504–3515
3. Andaló FA, Miranda PAV, Torres RDS, Falcão AX (2010) Shape feature extraction and description based on tensor scale. Pattern Recogn 43(1):26–36
4. Ankerst M, Kastenmüller G, Kriegel H-P, Seidl T (1999) 3d shape histograms for similarity search and classification in spatial databases. In: Advances in spatial databases. Springer, pp 207–226
5. Arbter K, Snyder WE, Burhardt H, Hirzinger G (1990) Application of affine-invariant fourier descriptors to recognition of 3-d objects. IEEE Trans Pattern Anal Mach Intell 12(7):640–647
6. Axenopoulos A, Litos G, Daras P (2011) 3d model retrieval using accurate pose estimation and view-based similarity. In: Proceedings of the 1st ACM international conference on multimedia retrieval. ACM, p 41
7. Belongie Serge, Malik Jitendra, Puzicha Jan (2002) Shape matching and object recognition using shape contexts. IEEE Trans Pattern Anal Mach Intell 24(4):509–522
8. Ben-Chen M, Gotsman C (2008) Characterizing shape using conformal factors. In: 3DOR, pp 1–8
9. Bronstein AM, Bronstein MM, Bruckstein AM, Kimmel R (2008) Analysis of two-dimensional non-rigid shapes. Int J Comput Vis 78(1):67–88
10. Bustos B, Schreck T, Walter M, Barrios JM, Schaefer M, Keim D (2012) Improving 3d similarity search by enhancing and combining 3d descriptors. Multimed Tools Appl 58(1):81–108
11. Chang X, Yang Y, Hauptmann AG, Xing EP, Yu Y-L (2015) Semantic concept discovery for large-scale zero-shot event detection. In: Proceedings of IJCAI
12. Chang X, Yang Y, Xing E, Yu Y (2015) Complex event detection using semantic saliency and nearly-isotonic svm. In: Proceedings of the 32nd international conference on machine learning (ICML-15), pp 1348–1357
13. Chang X, Yu Y-L, Yang Y, Hauptmann AG (2015) Searching persuasively: Joint event detection and evidence recounting with limited supervision. In: Proceedings of the 23rd Annual ACM Conference on Multimedia Conference. ACM, pp 581–590
14. Chellappa R, Bagdazian R (1984) Fourier coding of image boundaries. IEEE Trans Pattern Anal Mach Intell 6(1):102–105
15. Chen D-Y, Tian X-P, Shen Y-T, Ouhyoung M (2003) On visual similarity based 3d model retrieval. In: Computer graphics forum, vol 22. Wiley online library, pp 223–232
16. Cohen S, Guibas L (1999) The earth mover's distance under transformation sets. In: Computer vision, 1999. The proceedings of the seventh IEEE international conference on, vol 2. IEEE, pp 1076–1083
17. Cornea ND, Demirci MF, Silver D, Shokoufandeh A, Dickinson SJ, Kantor PB (2005) 3d object retrieval using many-to-many matching of curve skeletons. In: Shape Modeling and Applications, 2005 International Conference. IEEE, pp 366–371
18. Daliri MR, Torre V (2008) Robust symbolic representation for shape recognition and retrieval. Pattern Recogn 41(5):1782–1798
19. Demirci MF, Shokoufandeh A, Keselman Y, Dickinson S, Bretzner L (2003) Many-to-many matching of scale-space feature hierarchies using metric embedding. In: Griffin LD, Lillholm M (eds) Scale Space Methods in Computer Vision, volume 2695 of Lecture Notes in Computer Science. Springer, Berlin Heidelberg, pp 17–32
20. Demirci MF, Platel B, Shokoufandeh A, Florack L, Dickinson S (2009) The representation and matching of images using top points. J Math Imaging Vis 35(2):103–116

21. Demirci MF, Osmanlioglu Y, Shokoufandeh A, Dickinson S (2011) Efficient many-to-many feature matching under the ℓ_1 norm. Comput Vis Image Underst 115(7):976–983
22. Donoser M, Bischof H (2013) Diffusion processes for retrieval revisited. In: 2013 IEEE conference on Computer vision and pattern recognition (CVPR). IEEE, pp 1320–1327
23. Eberly D (1994) A differential geometric approach to anisotropic diffusion. In: Bart M, Romeny TH (eds) Geometry-Driven Diffusion in Computer Vision, volume 1 of Computational Imaging and Vision. Springer, Netherlands, pp 371–392
24. Ebrahim Y, Ahmed M, Abdelsalam W, Chau S-C (2009) Shape representation and description using the hilbert curve. Pattern Recogn Lett 30(4):348–358
25. Eitz M, Richter R, Boubekeur T, Hildebrand K, Alexa M (2012) Sketch-based shape retrieval. ACM Trans Graph 31(4):31
26. Frejlichowski D (2011) A three-dimensional shape description algorithm based on polar-fourier transform for 3d model retrieval. In: Heyden A, Kahl F (eds) Image Analysis, volume 6688 of Lecture Notes in Computer Science. Springer, Berlin Heidelberg, pp 457–466
27. Funkhouser T, Min P, Kazhdan M, Chen J, Halderman A, Dobkin D, Jacobs D (2003) A search engine for 3d models. ACM Trans Graph (TOG) 22(1):83–105
28. Furuya T, Ohbuchi R (2009) Dense sampling and fast encoding for 3d model retrieval using bag-of-visual features. In: Proceedings of the ACM international conference on image and video retrieval. ACM, p 26
29. Gal R, Shamir A, Cohen-Or D (2007) Pose-oblivious shape signature. IEEE Trans Vis Comput Graph 13(2):261–271
30. Gopalan R, Turaga P, Chellappa R (2010) Articulation-invariant representation of non-planar shapes. In: Computer vision–ECCV 2010. Springer, pp 286–299
31. Granlund GH (1972) Fourier preprocessing for hand print character recognition. IEEE Trans Comput 21(2):195–201
32. Guocheng A, Fengjun Z, Hong'an W, Guozhong D (2010) Shape filling rate for silhouette representation and recognition. In: 2010 20th international conference on Pattern recognition (ICPR). IAPR, pp 507–510
33. Hilaga M, Shinagawa Y, Kohmura T, Kunii TL (2001) Topology matching for fully automatic similarity estimation of 3d shapes. In: Proceedings of the 28th annual conference on computer graphics and interactive techniques. ACM, pp 203–212
34. Horn BKP (1984) Extended gaussian images. IEEE Proc 72(12):1671–1686
35. Hu R-X, Jia W, Zhao Ya, Gui J (2012) Perceptually motivated morphological strategies for shape retrieval. Pattern Recogn 45(9):3222–3230
36. Iyer N, Jayanti S, Lou K, Kalyanaraman Y, Ramani K (2005) Three-dimensional shape searching: state-of-the-art review and future trends. Comput Aided Des 37(5):509–530
37. Iyer N, Kalyanaraman Y, Lou K, Jayanti S, Ramani K (2003) A reconfigurable 3d engineering shape search system: Part i-shape representation. In: ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, pp 89–98
38. Kang SB, Ikeuchi K (1991) Determining 3-d object pose using the complex extended gaussian image. In: IEEE computer society conference on Computer vision and pattern recognition, 1991. Proceedings CVPR'91. IEEE, pp 580–585
39. Kauppinen H, Seppänen T, Pietikäinen M (1995) An experimental comparison of autoregressive and fourier-based descriptors in 2d shape classification. IEEE Trans Pattern Anal Mach Intell 17(2):201–207
40. Kawamura S, Usui K, Furuya T, Ohbuchi Rx (2012) Local goemetrical feature with spatial context for shape-based 3d model retrieval. In: 3DOR, pp 55–58
41. Kazhdan M, Funkhouser T, Rusinkiewicz S (2003) Rotation invariant spherical harmonic representation of 3 d shape descriptors. In: Symposium on geometry processing, vol 6
42. Kim H-K, Kim J-D (2000) Region-based shape descriptor invariant to rotation, scale and translation. Sig Proc Image Comm 16(1-2):87–93
43. Kuang Z, Li Z, Jiang X, Liu Y, Li H (2015) Retrieval of non-rigid 3d shapes from multiple aspects. Comput Aided Des 58:13–23
44. Laiche N, Larabi S, Ladraa F, Khadraoui Ax (2014) Curve norMalization for shape retrieval. Signal Process Image Commun 29(4):556–571
45. Leng B, Xiong Z (2011) Modelseek: an effective 3d model retrieval system. Multimed Tools Appl 51(3):935–962
46. Li B, Johan H (2013) 3d model retrieval using hybrid features and class information. Multimed Tools Appl 62(3):821–846

47. Li P, Wang Q, Zhang L (2013) A novel earth mover's distance methodology for image matching with gaussian mixture models ICCV
48. Li S-S, Huang Y-D, Yang J-W (2013) Affine invariant ring fourier descriptors. In: International conference on wavelet analysis and pattern recognition, pp 62–66
49. Lian Z, Godil A, Bustos B, Daoudi M, Hermans J, Kawamura S, Kurita Y, Lavoué G, Nguyen HV, Ohbuchi R et al (2013) A comparison of methods for non-rigid 3d shape retrieval. Pattern Recogn 46(1):449–461
50. Lian Z, Rosin PL, Sun X (2010) Rectilinearity of 3d meshes. Int J Comput Vis 89(2-3):130–151
51. Lin CC, Chellappa R (1987) Classification of partial 2d shapes using fourier descriptors. IEEE Trans Pattern Anal Mach Intell 9(5):686–690
52. Ling H, Jacobs DW (2005) Using the inner-distance for classification of articulated shapes. In: IEEE computer society conference on Computer vision and pattern recognition, 2005. CVPR 2005, vol 2. IEEE, pp 719–726
53. Ling H, Jacobs DW (2007) Shape classification using the inner-distance. IEEE Trans Pattern Anal Mach Intell 29(2):286–299
54. Liu T-L, Geiger D (1999) Approximate tree matching and shape similarity. In: The proceedings of the seventh IEEE international conference on Computer vision, 1999, vol 1. IEEE, pp 456–462
55. Lou K, Jayanti S, Iyer N, Kalyanaraman Y, Prabhakar S, Ramani K (2003) A reconfigurable 3d engineering shape search system: Part ii-database indexing, retrieval, and clustering. In: ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, pp 169–178
56. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110
57. Mémoli F, Sapiro G (2005) A theoretical and computational framework for isometry invariant recognition of point cloud data. Found Comput Math 5(3):313–347
58. Nanni L, Brahnam S, Lumini Ax (2012) Local phase quantization descriptor for improving shape retrieval/classification. Pattern Recogn Lett 33(16):2254–2260
59. Novotni M, Klein R (2004) Shape retrieval using 3d zernike descriptors. Comput Aided Des 36(11):1047–1062
60. Ohishi Y, Ohbuchi R (2013) Densely sampled local visual features on 3d mesh for retrieval. In: 2013 14th international workshop on Image analysis for multimedia interactive services (WIAMIS). IEEE, pp 1–4
61. Ohkita Y, Ohishi Y, Furuya T, Ohbuchi R (2012) Non-rigid 3d model retrieval using set of local statistical features. In: 2012 IEEE international conference on Multimedia and expo workshops (ICMEW). IEEE, pp 593–598
62. Osada R, Funkhouser T, Chazelle B, Dobkin D (2001) Matching 3d models with shape distributions. In: SMI 2001 international conference on Shape modeling and applications. IEEE, pp 154–166
63. Osada R, Funkhouser T, Chazelle B, Dobkin D (2002) Shape distributions. ACM Trans Graph (TOG) 21(4):807–832
64. Papadakis P, Pratikakis I, Perantonis S, Theoharis T (2007) Efficient 3d shape matching and retrieval using a concrete radialized spherical projection representation. Pattern Recogn 40(9):2437–2452
65. Papadakis P, Pratikakis I, Theoharis T, Passalis G, Perantonis S (2008) 3d object retrieval using an efficient and compact hybrid shape descriptor. In: Eurographics workshop on 3d object retrieval
66. Papadakis P, Pratikakis I, Theoharis T, Perantonis S (2010) Panorama: A 3d shape descriptor based on panoramic views for unsupervised 3d object retrieval. Int J Comput Vis 89(2-3):177–192
67. Pedrosa GV, Batista MA, Barcelos CAZ (2013) Image feature descriptor based on shape salience points. Neurocomputing 120:156–163
68. Pele O, Werman M (2009) Fast and robust earth mover's distances. In: ICCV
69. Rauber TW, Steiger-Garcao AS (1992) Shape description by unl fourier features-an application to handwritten character recognition. In: 11Th IAPR international conference on pattern recognition, pp 466–469
70. Rubner Y, Tomasi C, Guibas LJ (2000) The earth mover's distance as a metric for image retrieval. Int J Comput Vis 40(2):99–121
71. Ruggeri MR, Patanè G, Spagnuolo M, Saupe D (2010) Spectral-driven isometry-invariant matching of 3d shapes. Int J Comput Vis 89(2-3):248–265
72. Schreck T, Scherer M, Walter M, Bustos B, Yoon SM, Kuijper A (2012) Graph-based combinations of fragment descriptors for improved 3d object retrieval. In: Proceedings of the 3rd multimedia systems conference. ACM, pp 23–28

73. Sebastian TB, Klein PN, Kimia BB (2004) Recognition of shapes by editing their shock graphs. IEEE Trans Pattern Anal Mach Intell 26(5):550–571

74. Sharvit D, Chan J, Tek H, Kimia B (1998) Symmetry-based indexing of image databases. In: 1998. Proceedings. IEEE workshop on Content-based access of image and video libraries. IEEE, pp 56–62

75. Shekar BH, Pilar B (2014) Shape representation and classification through pattern spectrum and local binary pattern–a decision level fusion approach. In: 2014 fifth international conference on Signal and image processing (ICSIP). IEEE, pp 218–224

76. Shekar BH, Pilar B, Kittler J (2015) An unification of inner distance shape context and local binary pattern for shape representation and classification. In: Proceedings of the 2nd international conference on perception and machine intelligence. ACM, pp 46–55

77. Shen W, Bai X, Hu R, Wang H, Latecki LJ (2011) Skeleton growing and pruning with bending potential ratio. Pattern Recogn 44(2):196–209

78. Shen Y-T, Chen D-Y, Tian X-P, Ouhyoung M (2003) 3D model search engine based on lightfield descriptors. In: Eurographics

79. Shih J-L, Chen H-Y (2009) A 3d model retrieval approach using the interior and exterior 3d shape information. Multimed Tools Appl 43(1):45–62

80. Shih J-L, Lee C-H, Wang JTa (2007) A new 3d model retrieval approach based on the elevation descriptor. Pattern Recogn 40(1):283–295

81. Shilane P, Min P, Kazhdan M, Funkhouser T (2004) The princeton shape benchmark. In: Shape modeling applications, 2004. proceedings. IEEE, pp 167–178

82. Shu X, Wu X-J (2011) A novel contour descriptor for 2d shape matching and its application to image retrieval. Image Vis Comput 29(4):286–294

83. Siddiqi K, Bouix S, Tannenbaum A, Zucker SW (2002) Hamilton-jacobi skeletons. Int J Comput Vis 48(3):215–231

84. Siddiqi K, Zhang J, Macrini D, Shokoufandeh A, Bouix S, Dickinson S (2008) Retrieving articulated 3-d models using medial surfaces. Mach Vis Appl 19(4):261–275

85. Sipiran I, Bustos B, Schreck T (2013) Data-aware 3d partitioning for generic shape retrieval. Comput Graph 37(5):460–472

86. Sirin Y, Demirci MF (2014) Skeleton filling rate for shape recognition. In: 2014 22nd international conference on Pattern recognition (ICPR). IAPR, pp 4005–4009

87. Söderkvist O (2001) Computer vision classification of leaves from swedish trees

88. Sun J, Ovsjanikov M, Guibas L (2009) A concise and provably informative multi-scale signature based on heat diffusion. In: Computer graphics forum, vol 28. Wiley online library, pp 1383–1392

89. Sundar H, Silver D, Gagvani N, Dickinson S (2003) Skeleton based shape matching and retrieval. In: Shape modeling international, 2003. IEEE, pp 130–139

90. Tam GKL, Lau RWH (2007) Deformable model retrieval based on topological and geometric signatures. IEEE Trans Vis Comput Graph 13(3):470–482

91. Tangelder JWH, Veltkamp RC (2008) A survey of content based 3d shape retrieval methods. Multimed Tools Appl 39(3):441–471

92. Van der Zwan M, Meiburg Y, Telea A (2013) A dense medial descriptor for image analysis. In: VISAPP (1), pp 285–293

93. Van Otterloo PJ (1991) A Contour-oriented Approach to Shape Analysis. Prentice Hall International (UK) Ltd., Hertfordshire, UK

94. Vranic DV (2005) Desire: a composite 3d-shape descriptor. In: IEEE international conference on Multimedia and expo, 2005. ICME 2005. IEEE, pp 4–pp

95. Vranić DV, Saupe D (2004) 3d model retrieval. In: Proc SCCG 2000, pp 3–6

96. Wang Fan, Guibas LJ (2012) Supervised earth mover's distance learning and its computer vision applications. In: Computer vision–ECCV 2012. Springer, pp 442–455

97. Wang J, Bai X, You X, Liu W, Latecki LJ (2012) Shape matching and classification using height functions. Pattern Recogn Lett 33(2):134–143

98. Wu J, Rehg JM (2008) Where am i: Place instance and category recognition using spatial pact. In: 2008. CVPR 2008. IEEE conference on Computer vision and pattern recognition. IEEE, pp 1–8

99. Xie J, Heng P-A, Shah M (2008) Shape matching and modeling using skeletal context. Pattern Recogn 41(5):1756–1767

100. Xu J, Zhang Z, Tung AK, Yu G (2012) Efficient and effective similarity search over probabilistic data based on earth mover's distance. VLDB J Int J Very Large Data Bases 21(4):535–559

101. Zhang D, Lu G (2004) Review of shape representation and description techniques. Pattern Recogn 37(1):1–19

Yahya Sirin received his BS and MS both in electrical and electronics engineering from Ankara University. He is currently performing his PhD in computer engineering from TOBB University of Economics and Technology. His research interests include shape retrieval and object recognition in computer vision.

M. Fatih Demirci received his PhD in computer science from Drexel University in 2005. After working as a postdoctoral research scientist at Utrecht University, The Netherlands, he joined TOBB University of Economics and Technology, Turkey as a faculy member at computer engineering department. His research interests include structural pattern recognition in computer vision, shape indexing, and applied graph theory. Dr.Demirci was the recipient of the best PhD dissertation award in engineering and physical sciences, Drexel University in 2006 and TUBITAK Career project in 2010.