Multimed Tools Appl
DOI 10.1007/s11042-014-2267-9

An overview of partial 3D object retrieval methodologies

**Michalis A. Savelonas · Ioannis Pratikakis ·
Konstantinos Sfikas**

Received: 23 November 2013 / Revised: 15 July 2014 / Accepted: 8 September 2014

Abstract This work offers an overview of the state-of-the-art on the emerging area of 3D object retrieval based on partial queries. This research area is associated with several application domains, including face recognition and digital libraries of cultural heritage objects. The existing partial 3D object retrieval methods can be mainly classified as: i) view-based, ii) part-based, iii) bag of visual words (BoVW)-based, and iv) hybrid methods combining these three main paradigms or methods which cannot be straightforwardly classified. Several methodological aspects are identified, including the use of interest points and the exploitation of 2.5D projections, whereas the available evaluation datasets and campaigns are addressed. A thorough discussion follows, identifying advantages and limitations.

Keywords Partial 3D object retrieval · Interest point detection · View-based retrieval · Part-based retrieval · Bag of visual words

1 Introduction

Partial 3D object retrieval addresses the search of 3D models which are similar to a query, when the available information for the latter is not complete, as it is the case with range scans. For each partial query, a partial 3D object retrieval method is required to return a list of complete objects, retrieved from a database and ranked according to their similarity with the query. The similarity assessed is partial and can be distinguished from global similarity in that it implies a matching of only a part of the complete object with the query.

M. A. Savelonas (✉) · I. Pratikakis · K. Sfikas
ATHENA Research and Innovation Center, GR 67100 Xanthi, Greece
e-mail: msavel@di.uoa.gr

I. Pratikakis
e-mail: ipratika@ee.duth.gr

K. Sfikas
e-mail: ksfikas@di.uoa.gr

M. A. Savelonas · I. Pratikakis
Department of Electrical and Computer Engineering, Democritus University of Thrace, GR 67100 Xanthi, Greece

Published online: 28 September 2014

The interest for partial retrieval algorithms has been significantly boosted by the wide availability of 3D scanners, as well as by progress in 3D graphics technologies. This interest has been further amplified by the advent of several application domains, such as face recognition and digital libraries of cultural heritage (CH) objects, which require partial 3D object retrieval capabilities. In this context, two milestone challenges exist: (i) scanned queries can be rough and noisy; (ii) it is not straightforward to effectively match a partial query against a complete 3D model, since there is a gap between their representations. This representation gap complicates the extraction of a signature that will enable a matching between a complete 3D model and its partial counterpart.

This work aims to study the main properties of state-of-the-art partial 3D object retrieval methods, as well as their retrieval performance. The methods under study can be roughly classified as: i) view-based, ii) part-based, iii) bag of visual words (BoVW)-based, and iv) hybrid methods combining the three main paradigms or methods which cannot be straightforwardly classified. Starting from the methodologies and the experimental results presented, advantages and limitations of each method are outlined, which can serve as a reference for future research. Table 1 provides the definitions of several acronyms used throughout the text. Table 2 summarizes the state-of-the-art methods on partial 3D object retrieval, with respect to the methodological paradigms incorporated and the types of data addressed. Table 3 provides an overview of pre-existing generic retrieval benchmarks, with information on their size, number of classes, availability, as well as with references on related publications. Table 4 summarizes the experimental comparisons reported in the works to be reviewed. The used benchmark datasets are provided along with the methods under comparison and the results obtained in terms of the performance measures used in each work. The content of Tables 2, 3 and 4 will be elaborated in the sections to follow.

Table 1 Definitions of acronyms appearing in the text

Acronym	Definition	Acronym	Definition
ARG	Attributed Relation Graph	MAP	Maximum Average Precision
BoVW	Bag of Visual Words	MDS	Multi-dimensional Scaling
CAD	Computer-Aided Design	MSB	McGill Shape Benchmark
CH	Cultural Heritage	NDCG	Normalized Discounted Cumulative Gain
DCG	Discounted Cumulative Gain	NN	Nearest Neighbor
DGI	Depth Gradient Image	P-R	Precision-Recall
DSIFT	Dense Scale Invariant Feature Transform	PSB	Princeton Shape Benchmark
EMD	Earth Mover's Distance	PSO	Particle Swarm Optimization
ERC	Extremely Randomized Clustering	ROC	Receiver Operating Characteristic
ESB	Engineering Shape Benchmark	RR	Retrieval Rate
FT	First Tier	RT	Retrieval Time
GPU	Graphics Processing Unit	SHREC	Shape Retrieval Contest
HKS	Heat Kernel Signature	SIFT	Scale Invariant Feature Transform
ISDB	International Standard for Digital Broadcasting	ST	Second Tier
LSP	Local Surface Patch	ZDFR	Zernike-Depth-Fourier-Ray

Table 2 Summary of state-of-the-art methods related to partial 3D object retrieval

Class	Method	View-based	Part-based	BoVW	Interest points	Methodology	Data type
View-based	Chaouch et al.[11] (2006)	Yes	No	No	Yes	Depth-buffer and silhouette relevance indices	Mesh
	Daras and Axenopoulos [18] (2009)	Yes	No	No	No	Fourier, Zernike, Krawtchouk	Mesh
	Stavropoulos et al. [68] (2010)	Yes	No	No	Yes/method of Hoffman and Singh (1997)	Hierarchical search of camera parameters	Range image
	Adan et al. [1] (2011)	Yes	No	No	Yes	Depth gradient images for *scene* recognition	Mesh
	Sfikas et al. [64] (2013)	Yes	No	No	Yes/DSIFT	DSIFT interest points on panoramic views	Mesh
Part-based	Chen and Bhanu [13] (2007)	No	Yes	No	Yes/shape index of Dorai and Jain (1997)	Local surface patches based on shape indices	Range image
	Tierny et al. [70] (2009)	No	Yes	No	No	Reeb Graph	Mesh
	Agathos et al. [2] (2010)	No	Yes	No	Yes	ARG, EMD	Mesh
BoVW-based	Lavoué [46] (2012)	No	No	Yes	Yes/dense points	Fourier spectra of local patches	Mesh
	Bronstein et al. [10] (2011)	No	No	Yes	Yes/dense points/ Harris/ MeshDoG	HKS	Mesh, point cloud
	Li et al. [48] (2013)	No	No	Yes	Yes	Curvature-based and geodesic based features/PSO	Mesh
Other-methods	Hetzel et al. [39] (2001)	No	No	No	No	Multi-dim. histogr. of depth, normals and curvatures	Mesh
	Furuya and Ohbuchi [24] (2009)	Yes	No	Yes	Yes, dense points	SIFT descriptor	Mesh

Table 3 Overview of benchmark datasets that have been used in the context of partial 3D object retrieval

Benchmark	Availability	Related publication	Size (#samples)	Number of classes
ESB	https://engineering.purdue.edu/PRECISE/shrec08	[41]	867	44
ITI	http://vcl.iti.gr/3d-object-retrieval/	[17, 18]	544	13
Hampson pottery	http://www.ipet.gr/~akoutsou/benchmark/	[64]	384	16
MSB	http://www.cim.mcgill.ca/~shape/benchMark/	[66]	445	19
PSB	http://shape.cs.princeton.edu/benchmark/	[65]	1814	90
SHREC'07 watertight	http://watertight.ge.imati.cnr.it/	[36]	400	–
SHREC'07 partial	http://partial.ge.imati.cnr.it/	[52]	400	–
SHREC'09 partial	http://www.itl.nist.gov/iad/vug/sharp/benchmark/shrecPartial/	[21]	720	40
SHREC'10 large scale	http://give-lab.cs.uu.nl/SHREC/ULS3SRB/2010/	[72]	10000	54
SHREC'10 range scan	http://www.itl.nist.gov/iad/vug/sharp/contest/2010/RangeScans/	[22]	800	40
SHREC'11 range scan	http://www.itl.nist.gov/iad/vug/sharp/contest/2011/RangeScans/	–	1000	50
SHREC'13 partial	http://dataset.dcc.uchile.cl/	[67]	7200	20

The remainder of this paper is organized as follows: in Section 2 we explore view-based partial 3D object retrieval methods. Part-based and BoVW-based methods are explored in Sections 3 and 4, respectively. Section 5 presents partial 3D object retrieval methods which are either hybrid or hard to classify. Finally, Section 6 provides an in depth discussion on the partial 3D object retrieval methods presented.

2 View-based methods

View-based 3D object retrieval methods are particularly relevant to the context of partial retrieval, since an object view is closely associated with a range scan, one of the primary forms of partial 3D object representation. In this light, even those view-based methods which have been originally formulated for generic 3D object retrieval are of potential interest in a partial retrieval context, assuming some modifications. In that respect, the recent extensive research in generic view-based retrieval methods [26–34, 47, 75–77] is particularly relevant. Such generic view-based schemes can be considered, following the idea of Daras' and Axenopoulos' method (subsection 2.2), which has been proposed for both global and partial retrieval. In the first case, a similarity metric derived from the sum of distances between all possible query-target view pairs is used, whereas in the latter case similarity is determined by the minimum of such distances.

2.1 3D object retrieval based on depth-buffer and silhouette relevance indices

Chaouch and Verroust-Blondet [11] have proposed a 2.5D object retrieval method which uses range images as input. Their method is based on *relevant indices* derived from silhouette or depth-buffers. As a relevance index which depends on the outer object silhouette, they use two

Table 4 Summary of experimental comparisons reported in the literature

Related publication	Benchmark	Comparison results
Chaouch and Verroust-Blondet [11]	PSB	Enhanced depth-buffer-based method [11]: NN 55.5, FT 29.6, ST 40.0, DCG 57.3. Depth-buffer-based method of Vranic [73]: NN 52.8, FT 28.5, ST 38.8, DCG 56.3. Enhanced silhouette-based method [11]: NN 60.5, FT 34.3, ST 44.2, DCG 60.2. Silhouette-based method of Vranic [73]: NN 59.2, FT 32.9, ST 41.8, DCG 58.9.
Daras and Axenopoulos [18]	ITI, PSB, ESB	Daras and Axenopoulos [18] outperforms the methods of Vranic [73] and Ohbuchi et al. [56], with respect to P-R.
Stavropoulos et al. [68]	SHREC'07 watertight, PSB	Stavropoulos et al. [68] outperforms the method of Germann et al. [35], with respect to P-R.
Adan et al. [1]	MSB	Adan et al. [1]: RR 96.8, RT 3 min. Johnson and Hebert [42]: RR 87.8, RT 400 min.
Chen and Bhanu [13]	Ohio State University	Chen and Bhanu [13]: RR 100.0, RT 89.42 s. Johnson and Hebert [42]: RR 100.0, RT 162.07 s. Correa and Shapiro [16]: RR 100.0, RT 150.57 s.
Tierny et al. [70]	SHREC'07 partial	Tierny et al. [70] outperform Biasotti et al. [7] and Cornea et al. [15], with an average NDCG which is 14.1 % and 40.9 % higher, respectively.
Agathos et al. [70]	MSB, ISDB	Agathos et al. [70]: NN 97.6/100.0, FT 74.1/89.5, ST 91.1/95.8, DCG 93.3/97.2 (MSB/ISDB). Papadakis et al. [44]: NN 92.5/84.9, FT 55.7/54.1, ST 69.8/68.5, DCG 85.0/79.9 (MSB/ISDB). Kim et al. [44]: NN 91.8/87.7, FT 65.2/69.9, ST 78.3/84.8, DCG 89.1/88.1 (MSB/ISDB). Also, Agathos et al. [70] outperforms the other two methods, with respect to P-R.
Lavoué [46]	SHREC'07 partial	Lavoué [46] outperforms the methods of Tierny et al. [70] and Toldo et al. [71], with respect to NDCG.
Bronstein et al. [10]	SHREC'10 large scale	Bronstein et al. [10] outperform the methods of Toldo et al. [71] and Lian et al. [50], with respect to MAP (87.4 as opposed to 1.4 and 44.8, respectively) and ROC.
Li et al. [48]	SHREC'07 partial	Li et al. [48] outperforms the methods of Tierny et al. [70] and Toldo et al. [71], with respect to NDCG.
Furuya and Ohbuchi [24]	PSB, MSB, ESB	Furuya and Ohbuchi [24]: R-precision 55.8/76.4/42.5 (PSB/MSB/ESB) Chen et al. [14]: R-precision 45.9/56.9/34.7 (PSB/MSB/ESB) Kazhdan et al. [43]: R-precision 40.5/56.7/34.6 (PSB/MSB/ESB)
Sipiran et al. [67]	SHREC'13 partial	Li-Lu-Johan: NN 34.0, FT 21.06, ST 16.69, MAP 22.3 Sipiran-Bustos: NN 34.76, FT 20.86, ST 13.34, MAP 20.34

alternatives: the first (R_α) is standard and involves the computation of the number of non-null pixels on the image, i.e. the area of the projected surface of the 3D model on the corresponding face of the bounding box:

$$R_a = card\{S_{ab}|S_{ab} = 1, 0 \leq a, b \leq N-1\}$$

where s_{ab} is the pixel value of the image at position (α,b) and N is the image size. To moderate the influence of the area which in some cases may affect the retrieval performance, they consider the square root of the relevance defined in the above equation. As a second alternative for silhouette-based relevance index (R_c), they use the average cord of a 2D mesh, i.e. the average length of all possible cords connecting two contour points:

$$R_c = \sum_{a=0^{N-1}} \sum_{b=0^{N-1}} \sum_{p=0^{N-1}} \sum_{q=0^{N-1}} \frac{\delta_{c_{ab} \cdot c_{pq}}}{L(L-1)} \sqrt{|a-p|^2 + |b-q|^2}$$
$$\delta_{x.y} = 1 \text{ if} x = y = 1 \text{ and } \delta_{x.y} = 0 \text{ otherwise}$$

where c_{ab} is the pixel value at position (a,b) and L is the contour length. The pixels where c_{ab}= 1 define the outer contour of the silhouette.

Several sampling strategies have been considered for selecting the contour points used in this calculation. These strategies include using all points of the outer contour of the silhouette, points of high curvature or a subset of interest points, as in the work of Lowe [51]. However, the authors note that this type of information can well represent the relevance for some particular cases but can be much less efficient for many 3D models due to the unstable behavior of such key points.

Chaouch and Verroust-Blondet also proposed two methods to compute the relevance indices of depth-buffer images. The first one introduces the depth by taking the sum of all values of the non-null pixels of the depth-buffer image, thus computing the volume enclosed between the visible parts of the 3D object and the opposite plane of the bounding box:

$$R_d = \sum_{a=0^{N-1}} \sum_{b=0^{N-1}} u_{ab}$$

where u_{ab} is the pixel value of the depth-buffer image at position (α,b).

The second relevance index proposed for depth-buffer images is the sum of the distances between the center of mass of the 3D model and all its visible points:

$$R_g = \frac{1}{2w} \sum_{a=0^{N-1}} \sum_{b=0^{N-1}} d_{ab}$$
$$d_{ab} = \sqrt{\left|a - \left|N/2\right|^2 + \left|b - \left|N/2\right|^2 + 2w\left|u_{ab} - 1/2\right|^2\right.\right.}$$

where $2w$ is the length of the sides of the extended enclosing bounding box.

Experiments were conducted on range images artificially acquired from the well-known Princeton shape benchmark (PSB) [65]. The obtained results have been compared with the silhouette and depth-buffer methods of Vranic [73] on the same database. Chaouch and Verroust-Blondet report enhanced retrieval quality for both silhouette and depth-buffer-based variants. In quantitative terms, the enhanced silhouette-based variant obtains NN 55.5, FT 29.6, ST 40.0, DCG 57.3, outperforming the silhouette-based method of Vranic, which obtains NN 52.8, FT 28.5, ST 38.8, DCG 56.3. In a similar fashion, the enhanced depth-buffer variant obtains NN 60.5, FT 34.3, ST 44.2, DCG 60.2, outperforming the depth-buffer-based method of Vranic, which obtains NN 59.2, FT 32.9, ST 41.8, DCG 58.9. In addition, the overall computational cost has not been increased.

2.2 Fourier/Zernike/Krawtchouk-based method

Daras and Axenopoulos [17, 18] proposed a view-based 3D object retrieval method using feature vectors comprising polar Fourier coefficient, Zernike and Krawtchouk moments. These features are calculated either on binary images or on depth images extracted from 18 2D views, which are taken from the vertices of a bounding 32-hedron. A pose estimation step based on the work of Daras et al. [19] or Pu et al. [63] precedes view extraction. This step involves the translation, scaling and rotation of the 3D model. The model is translated so that the center of mass coincides with the center of the coordinate system and scaled in order to lie within a bounding sphere of radius equal to one. Rotation estimation involves combining principal component analysis with the visual contact area described in the work of Pu et al. [63], in order to derive the three principal axes of the model. The actual viewpoints employed in the calculations are determined by rotating 24 times in 90 degrees intervals around the three principal axes. The method of Daras and Axenopoulos is summarized in Fig. 1.

Daras and Axenopoulos' method can be applied for 3D object retrieval with complete or partial queries. In the first case, the utilized dissimilarity measure sums the distances of associated 2D views (Fig. 2), whereas in the case of partial retrieval, the minimum distance between the query model and each 2D view is used. Experiments are performed on three

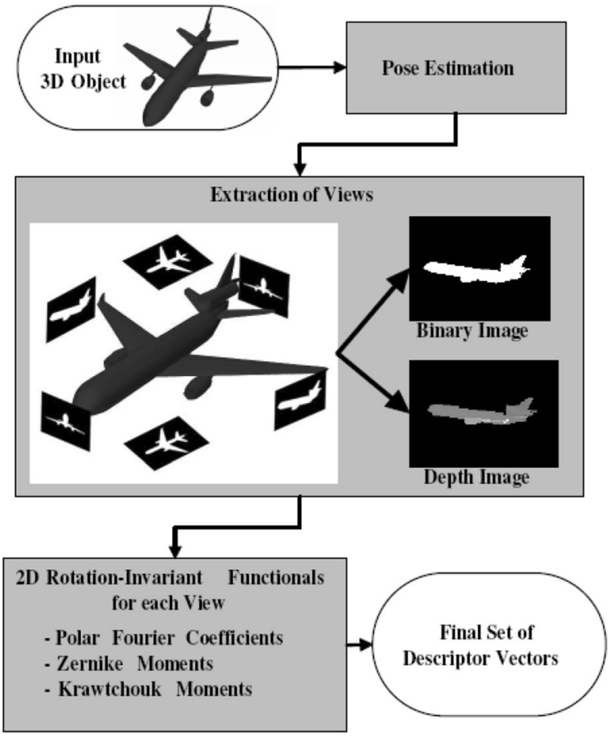


Fig. 1 Block diagram of the method of Daras and Axenopoulos [17]

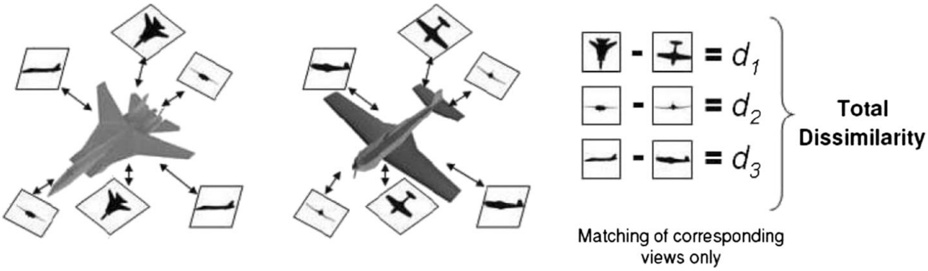


Fig. 2 In the case of complete queries, the total dissimilarity between two 3D objects is the sum of the dissimilarities of the corresponding views [17]

databases: (i) a database of 544 3D models classified into 13 classes, as compiled by the Informatics and Telematics Institute (ITI), which is available online at http://vcl.iti.gr/3d-object-retrieval/, (ii) the engineering shape benchmark (ESB) [41] which contains 867 3D computer-aided design (CAD) models from the mechanical engineering domain, classified into 44 classes and (iii) the well-known PSB. Daras and Axenopoulos' method outperforms the methods of Chen et al. [14], Vranic [73], as well as the BoVW-based method of Ohbuchi et al. [56]. In the more complete journal version of their work, Daras and Axenopoulos [18] combined their feature vector with the spherical trace transform of Zarpalas et al. [78], resulting in enhanced retrieval performance. This method has participated in three tracks of the shape retrieval contest (SHREC)'09, namely the tracks for: (i) structural shape retrieval [38], (ii) a new generic shape benchmark [4] and (iii) partial 3D object retrieval [6].

2.3 Use of salient features in range images

Stavropoulos et al. [68] introduced a method for identifying the correspondence between a range image and a full 3D model by searching for the camera viewpoint, orientation, scale and internal geometry that would generate an image similar to the query, as illustrated in Fig. 3. Instead of attempting to match the entire image, only spatial distributions of salient points are compared. The salient points are extracted following the theory of salience, as introduced by Hoffman and Singh [40]. A coarse-to-fine, hierarchical approach is adopted for searching in the parameter space in order to bypass exhaustive searching.

The framework of Stavropoulos et al. is experimentally tested on the dataset used in the "watertight" track of SHREC'07 [36], as well as in the PSB. The obtained retrieval results show that this framework outperforms the method of Germann et al. [35], including cases of noise-infused or occluded 3D models. Moreover, for standard Intel-based workstations the time dedicated for off-line preprocessing and on-line partial matching is only 1 sec and 20 msec, respectively.

2.4 Depth gradient images (DGIs)

Adan et al. [1] introduced and analyzed a 3D object retrieval strategy *for scenes*, based on depth gradient image (DGI) representation. DGI synthesizes both surface and contour information, aiming to avoid restrictions on the layout and visibility of each object in the scene. Figure 4 summarizes this strategy. Let v be an arbitrary view on

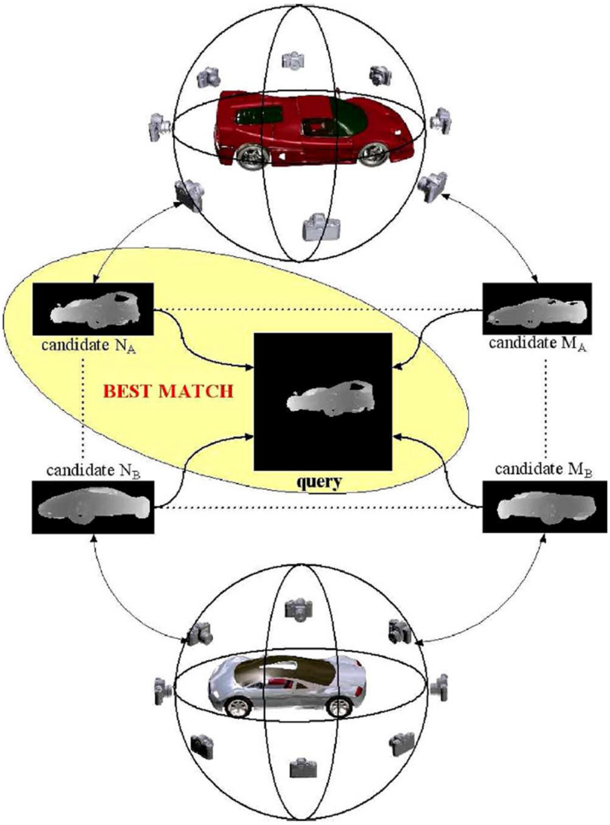
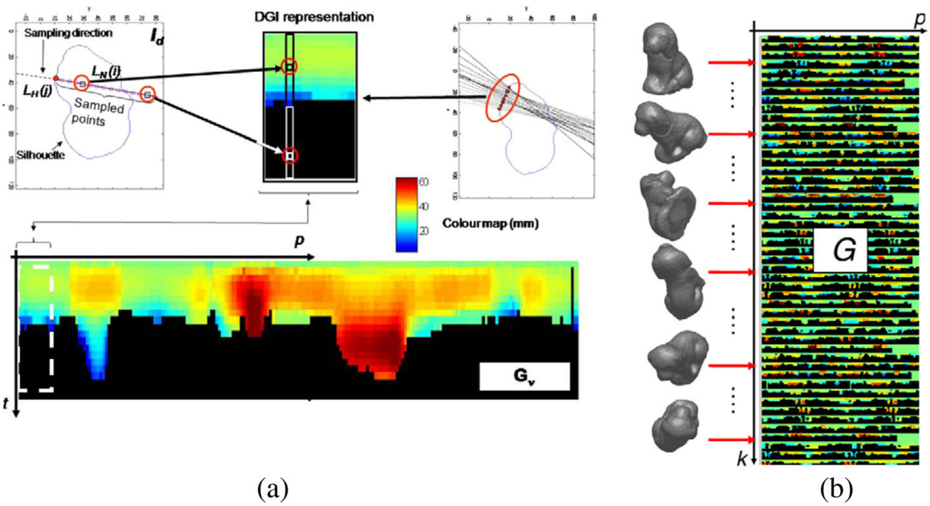


Fig. 3 The algorithm of Stavropoulos et al. searches for the best match in parameter space that consists of all possible positions and orientations of the camera [68]

the object, L be the set of pixels of depth image I_d which represent the object and L_H be the set of pixels corresponding to the object contour. In addition, let p=ord(L_H) be the number of pixel of the contour and t=dim(diag(I_d)) be the dimension of the diagonal of the depth image. The DGI from viewpoint v, G_v, is a $t \times p$ matrix:

$$G_u(i,j) = I_d(L_N(i,L_H(j)))-I_d(L_H(j))$$

where $L_N(i,L_H(j)),i=1,2,\ldots,t$ is the set of pixels that are in the normal direction to the contour at the point $L_H(j)$, which are sorted from $L_H(j)$ towards the object interior. L_N can be sub-sampled in order to reduce the size of the DGI representation and limit memory requirements. The depth values of a set of equally-spaced pixels in the normal direction can thus be taken in the above equation. It is clear that G_υ is invariant to changes in the observer distance. G_υ is a small image where the jth column contains the set of depth gradients in the normal direction in $L_H(j)$, whereas the ith row stores the gradients for points that are equidistant in the image to the contour in their corresponding normal directions. Figure 4(a) shows an example of



(a) (b)

Fig. 4 Building of the DGI representation: (**a**) depth gradients generated over a sampling direction when a contour point is selected. DGI values inside and outside of the object are included, as well as a section of DGI corresponding to a part of the contour and its integration in the DGI, (**b**) the global DGI model [1]

how depth-gradient values for one contour-point (on the left) and consecutive contour-points along with their associated normal directions (on the right) are calculated.

DGI representation is suitable for characterizing both partial views and the complete object. The global DGI model can be defined as follows:

$$G(j,i) = G_u\left(i^{'},j^{'}\right), i^{'} = MOD(i,t), j^{'} = j, \mu = DIV(i,t)$$

where G_μ is the partial DGI obtained from the viewpoint μ . Thus, the global DGI model consists of an image of dimension (*k*/*t*,*p*), which is duplicated in practice, so as to carry out an efficient partial-global DGI matching. Figure 4(b) shows the global DGI for one object. Note that this is a single image of 1M pixels, which synthesizes the surface information of the complete object.

Promising retrieval performance (RR 96.8) is obtained by applying DGI on various scene types, which include occlusion, injected noise and highly complex cluttered scenes. Moreover, DGI outperforms spin-images [42] (RR 87.8) on Mian's public dataset whereas it obtains comparable performance to Mian's tensor-shape method (RR 96.6) [54].

2.5 Use of panoramic views

Sfikas et al. [64] recently proposed a view-based method, extending PANORAMA, a generic 3D object retrieval method introduced by Papadakis et al. [61]. The method of Sfikas et al. uses a panoramic view representation that is able to encode the 3D surface characteristics of target objects onto a 2D image map. For this, a number of panoramic views of each object are extracted on viewpoint axes that are defined by a dodecahedron, thus extending PANORAMA to multiple viewpoint axes. Each axis defines three panoramic view cylinders (one for the axis itself and two more for any two axes, so as to make up an orthonormal basis, along with the first one). To obtain a panoramic view, the object is projected to the lateral surface of a cylinder centered at the origin, with its axis parallel to one of the coordinate axes. For each cross section of the cylinder, rays are cast to the cylinder center, and the respective distances crossed are assigned to the points of the cross section. The cylindrical projection can be viewed as a 2D image map. Figure 5 illustrates an example 3D object and its cylindrical projection. In the same spirit, range image representations of query objects are mapped to the panoramic views, and the algorithm identifies object matches by means of interest point correspondence. The interest points are compared with the use of dense scale invariant feature transform (DSIFT) descriptor [9].

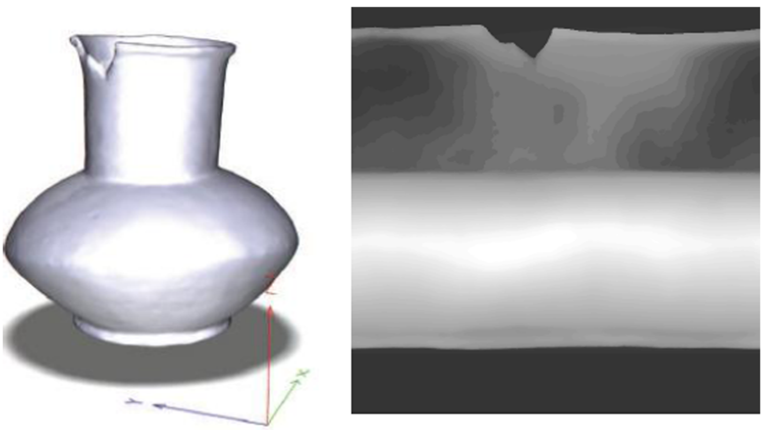
Sfikas et al. experimented on a dataset of pottery models obtained from the Hampson Archeological Museum collection of historical 3D objects [45]. They used models of low resolution that were downloaded from the website of the museum, along with associated metadata information, as a test bed for retrieval experiments. They artificially created a set of 20 partial queries by slicing and cap filling a corresponding amount of complete 3D objects. The experimental results showed that their method is able to handle quite well the problem of partial retrieval (NN 85.7, DCG 68.3) and illustrates stability in its performance with respect to the difficulty of the problem. Figure 6 illustrates example retrieval results obtained by Sfikas et al.

3 Part-based methods

Biederman [8] suggested that humans tend to recognize objects by analyzing the semantics of their parts. This suggestion leads to the part-based paradigm in 3D object retrieval, which is based on the hypothesis that two objects are similar, if they consist of similar parts. The relevance of this approach to partial retrieval is obvious, if we consider that object parts are actually an input form for partial retrieval.

3.1 Local surface patches (LSPs)

Chen and Bhanu [13] introduced LSPs for 3D object representation. Their method starts from extracting feature points in range images and defines LSP descriptors [12] for each feature



(a) (b)

Fig. 5 (**a**) An example 3D object and (**b**) its corresponding cylindrical projection on the z-axis [64]

Fig. 6 Sample queries from the pottery dataset. First column indicates the query model and results are illustrated in ranking order [64]

point with large shape variation, as measured by the *shape index* [20] (Fig. 7). Each LSP is defined as the region consisting of a feature point and its neighbors.

For each patch, Chen and Bhanu calculate local surface properties, including a 2D histogram, the surface type and the centroid. The 2D histogram (Fig. 8) consists of shape indices and angles between the normal of the feature point and that of its neighbors. The surface of a patch is classified into different types based on the mean and Gaussian curvatures of the feature point. For every LSP, the mean and standard deviation of shape indices are computed and used as indices to a hash table (Fig. 9). Potential associations between LSPs and candidate models are hypothesized by comparing LSPs of a query and LSPs of a full 3D model, followed by casting votes for those models containing similar surface descriptors. Finally, a rigid transformation is estimated based on the corresponding LSPs, so as to enable the calculation of the match quality between the hypothesized 3D model and the query.

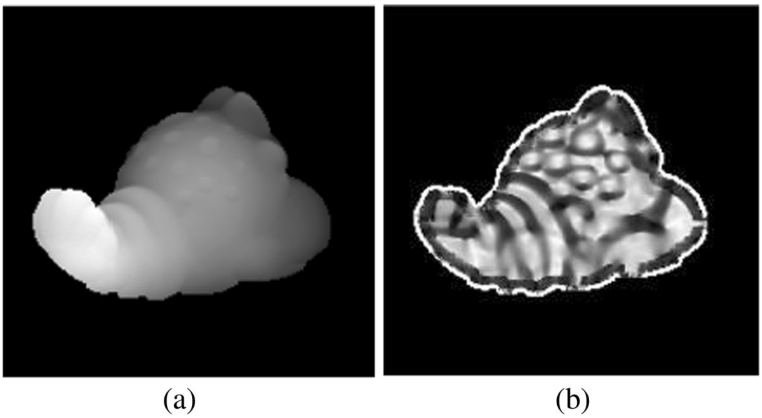


Fig. 7 (**a**) A range image and (**b**) the image of its shape index. In (a), the darker pixels are further away from the camera and the lighter ones are closer. In (b), the darker pixels correspond to concave surfaces and the lighter ones correspond to convex surfaces [13]

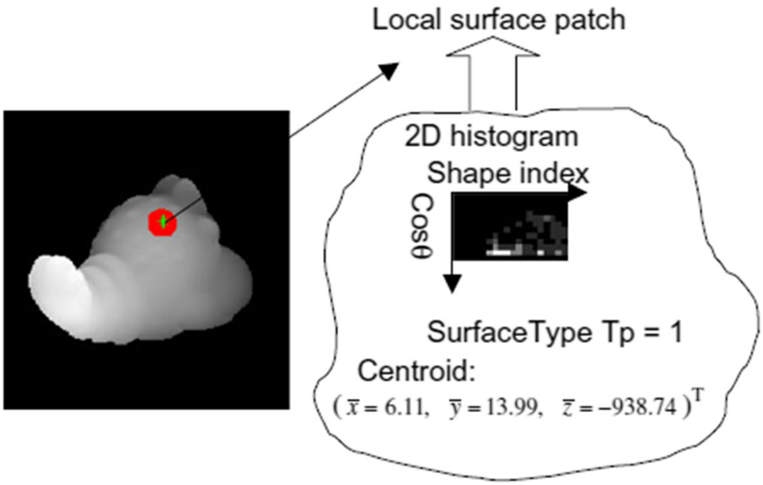


Fig. 8 Illustration of a local surface patch (LSP). Feature point P is indicated in green and its neighbors N are indicated in red [13]

Experiments were performed on a database of real range data, collected by Ohio State University. Comparisons with the spin-image [42] and the spherical spin-image [16] representations on real range data have shown that LSPs are as effective for the matching of 3D objects as these two representations (all methods obtain RR 100.0), but are *more efficient* in finding corresponding parts between a model-query pair, since LSPs require 89.4 s, spin-images require 162.1 s and spherical spin-mages require 150.6 s, with the implementations and hardware configurations used by the authors in their experiments. Unfortunately, the link of the database, as referenced by the authors, is no longer valid and we did not manage to identify a new URL.

3.2 Partial 3D shape retrieval by Reeb pattern unfolding

Tierny et al. [70] proposed a Reeb graph-based partial 3D object retrieval method, using their earlier mesh segmentation method [69]. For each segment of the object a signature is computed using its Reeb chart. Disk-like and annulus-like charts are considered. Disk-like charts correspond to one local maximum of the graph with the local maximum located in the

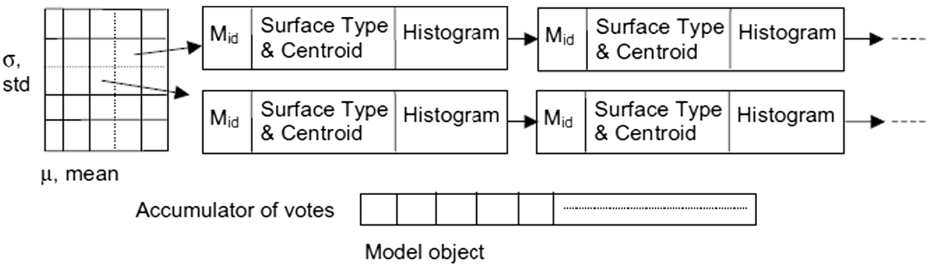


Fig. 9 Structure of the hash table. Every entry in the hash table has a linked list which saves information about the model LSPs and the accumulator records the number of votes received by each model [13]

center of the chart and the boundaries on the outer circle of the disk. Disk-like charts correspond to the fingers, whereas the annulus-like chart corresponds to the palm of a hand object. Let c_i be the disk-like chart of a segment. If φ_i is the mapping of c_i to the canonical planar domain D, then the unfolding signature $l_{\varphi i}$ can be defined as follows:

$$l_{\varphi i}(r) = \frac{A_{ci}(r)}{A_{D(r)}} = \frac{A_{ci}(r)}{\pi r^2}$$

where r denotes a subset of the chart, and A_{ci}, A_D denote the total area of the subset in each of the two domains. Let now c_j be the annulus-like chart of the object. The signature can be computed as follows:

$$l_{\varphi j}(r) = \frac{A_{cj}(r)}{A_{D(r)}} = \frac{A_{cj}(r)}{p(r+1)^2 - p}$$

The Reeb graph matching is performed using the above signature. A Reeb pattern is a part of the Reeb graph which contains protrusion areas. The structural signature of a Reeb pattern P_i is the couple $(n_D(P_i), n_A(P_i))$, where $n_D(P_i)$ and $n_A(P_i)$ are the numbers of the disk-like and annulus-like Reeb charts in P_i, which are linked by the following equation with g_{pi} denoting the genus of the Reeb pattern:

$$n_D(P_i) = n_A(P_i) + 1 - 3g_{Pi}$$

Making use of the structural signature, the maximal common sub-graph is identified. The final step of the method is matching of the Reeb patterns using the following similarity function and a bipartite graph matching algorithm:

$$s(c_{Ai}, c_{Bj}) = 1 - L_{N1}(c_{Ai}, c_{Bj})$$

where L_{N1} is the normalized L_1 distance between the unfolding signatures of the set of matched disk charts c_{Ai} and c_{Bj}.

Experiments were performed on the partial retrieval track of SHREC'07 [52] benchmark database, demonstrating that the method of Tierny el al. outperforms the methods of Biasotti et al. [7] and Cornea et al. [15] in terms of retrieval accuracy, with an average NDCG which is 14.1 and 40.9 % higher, respectively.

3.3 Retrieval of 3D articulated objects using a graph-based representation

Agathos et al. [2] proposed a graph-based representation method that decomposes objects using the mesh segmentation method previously introduced by some of the authors [3]. Geodesic extrema of an object are considered as salient points identified by means of the *protrusion function*, which depends on geodesic distances of all pairs of points in a neighborhood and reaches its local maxima at the tips of mesh protrusions. The core partition is approximated by starting from the minimum of the geodesic function and expanding the partition. When the expansion is completed, the protrusion parts are separated from the core. Boundaries are refined with a minimum cut algorithm to form the final segmentation.

After the segmentation step, each segment of the object is represented as a graph node and adjacent segments are connected in the graph with an edge. Unary and pairwise features are assigned to each node and edge, respectively. The graph matching is based on the earth mover's distance (EMD) of the feature vectors. Unary attributes assigned to the nodes include size, convexity, eccentricities of the ellipsoid approximating the component [3] and the spherical harmonic descriptor vector [59]. The pairwise features assigned to graph edges are the distance of the segment centroids and the angles that the two most significant principal axes of the connected components form with each other. Before the matching of two graphs, penalty nodes are inserted in the graph with the smaller number of nodes (equal to their difference of cardinality).

Experiments are performed on the McGill shape benchmark (MSB) [66] database, which consists of highly articulated shapes, and ISDB [25] database, showing that the retrieval method of Agathos et al. outperforms the part-based method of Kim et al. [44] and the hybrid descriptor of Papadakis et al. [60], in terms of retrieval accuracy. In quantitative terms, the method of Agathos et al. obtains NN 97.6/100.0, FT 74.1/89.5, ST 91.1/95.8, DCG 93.3/97.2 in MSB/ISDB datasets, as opposed to NN 92.5/84.9, FT 55.7/54.1, ST 69.8/68.5, DCG 85.0/79.9 for the method of Papadakis et al. and NN 91.8/87.7, FT 65.2/69.9, ST 78.3/84.8, DCG 89.1/88.1 for the method of Kim et al. Also, Agathos et al. outperform the other two methods, with respect to P-R.

4 Bag of visual words-based methods

The past decade has seen the rise of the bag of visual words (BoVW) approach in computer vision. In relevant literature, BoVW can also be found as *bag of words*, *bag of features* or *bag of visual features*. BoVW methods have been applied for image classification, object detection, image retrieval and even visual localization for robots. In visual information retrieval, the BoVW approach defines that each sample contains a number of local visual features. Since every visual feature, or collection of similar visual features, may appear with different frequencies on each sample, matching the visual feature frequencies of two samples achieves correspondence.

Figure 10 provides a visual abstraction of the BoVW procedure, which can be summarized as follows: (i) *build vocabulary*: extract features from all samples in a training set. Vector

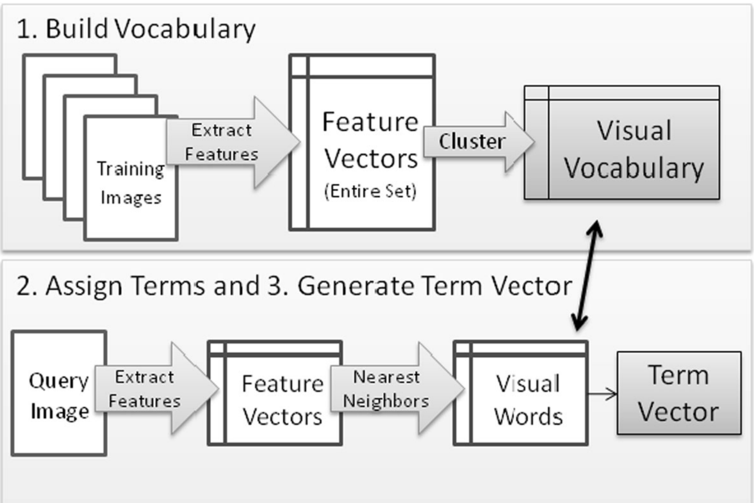


Fig. 10 Process for BoVW image representation [55]

quantize, or cluster, these features into a “visual vocabulary,” where each cluster represents a “visual word” or “term.” In some works, the vocabulary is called the “visual codebook.” Terms in the vocabulary are the codes in the codebook, (ii) *assign terms*: extract features from a test sample. Use nearest neighbors or a related strategy to assign the features to the closest terms in the vocabulary, (iii) *generate term vector*: record the counts of each term that appears in the image to create a normalized histogram representing a “term vector” [55]. This term vector is the BoVW representation of the sample.

Shape Google, as proposed by Bronstein et al. [10], is the one major 3D model object retrieval method in the literature. Other recent notable contributions along the same line include the methods of Furuya and Ohbuchi [24], Lavoué [46], Ohkita et al. [58], Li et al. [49], Atmosukarto and Shapiro [5] and Li et al. [48]. It should be noted that from these BoVW-based methods, those of Bronstein, Furuya and Ohbuchi, Lavoué, as well as of Li et al. [48], have already been applied for partial retrieval. However, all methods feature ideas which are potentially useful in a partial retrieval context.

4.1 Spatially sensitive BoVW methods

Lavoué [46] has presented an alternative 3D object retrieval method which also combines standard BoVW and spatially-sensitive BoVW. His method relies on uniform sampling of feature points based on Lloyd’s relaxation iterations. Each feature point is associated to a descriptor defined as the Fourier spectra of a local patch, which is computed by projecting the geometry onto the eigenvectors of the Laplace–Beltrami operator, so as to speed-up computations and enhance discriminative capability.

The experimental evaluation of this method for partial retrieval has been performed on SHREC’07 partial retrieval benchmark [52]. Each of the query models is composed of sub-parts from two or three models from the testing set. A ground-truth classification of each model of the testing set as highly relevant, marginally relevant or non-relevant is provided for each query. Lavoué’s method is shown to outperform the 3D object retrieval methods of Tierny et al. [70] and Toldo et al. [71], as illustrated in the NDGC curves provided in Lavoué’s paper. The author explains this on the basis that his method discards most of the structural information, hence the topological changes due to the sub-part merging do not significantly affect the BoVW. Moreover, Lavoué [46] has shown that standard and spatially-sensitive BoVW methods are complementary since their combination provides a significant gain with regards to their individual performances. In quantitative terms, the hybrid BoVW method obtains NN 91.8, FT 60.0, ST 74.0, DGC 84.7, as opposed to NN 89.7, FT 56.7, ST 71.5, DGC 83.3 obtained by the standalone spatial variation of BoVW, and NN 90.2, FT 59.0, ST 73.4, DGC 84.1 obtained by the standalone standard variation of BoVW, in experiments conducted on SHREC’07.

A current limitation of Lavoué’s method, as the author admits, is that although it correctly retrieves a model from a partial query, it does not perform the precise matching between the corresponding sub-parts. A solution to perform this matching could have been to construct a graphical structure over the set of feature points, i.e. a graph representation reflecting the attributes and the relations of primitive elements (i.e. points, lines, polyhedra etc.) constituting the object, in an organized fashion, and apply some kind of fast approximate subgraph isomorphism.

4.2 Shape google

Shape Google [10] is a 3D object retrieval method applicable on both meshes and point clouds, which focuses on the retrieval of *non-rigid objects*. It starts by calculating a feature detector and descriptor based on heat kernels of the Laplace-Beltrami operator. The descriptors derived are used to construct a BoVW vocabulary. This representation is invariant to isometric deformations, robust under a wide class of perturbations, and allows one to compare shapes undergoing different deformations. Bronstein et al. [10] take into consideration the spatial relations between features in an approach similar to commute graphs, which has been shown to enhance the retrieval performance. Finally, they adopt metric learning techniques, widely used in the computer vision community and represent shapes as compact binary codes that can be efficiently indexed and compared using the Hamming distance.

Figure 11 provides an overview of the Shape Google pipeline. The shape is represented as a collection of local feature descriptors, either dense or computed at a set of stable points, following an optional stage of feature detection. The descriptors are then represented by "geometric words" from a "geometric vocabulary" using vector quantization, which produces a shape representation as a BoVW or pairs of words, i.e. "expressions". Finally, similarity sensitive hashing is applied on the BoVW. Figure 12 visualizes the discriminative capability of the employed heat kernel-based BoVW. Shape Google can be adopted as a framework with different descriptors and detectors, depending on the application demands.

Shape Google has been tested on the large scale track of SHREC'10 [72], in which the queries include multiple modifications and transformations of the same shape, such as rescaling, induction of noise, sampling and *partial queries*. In the latter case, Shape Google outperformed the methods of Toldo et al. [71] and Lian et al. [50], with a MAP of 87.4 for Shape Google, as opposed to 1.4 for the method of Toldo et al. and 44.8 for the method of Lian et al.

4.3 Global and local features combined with particle swarm optimization

Li et al. [48] proposed a partial 3D object retrieval method which utilizes both geodesic distance-based global features and curvature-based local features, together with the BoVW framework, in order to develop a generic 3D shape retrieval algorithm that can be used for both non-rigid and partial 3D model retrieval. Curvature-based local features, geodesic distance-based global features without multidimensional scaling (MDS), and MDS-based curvature-based ZDFR features (where ZDFR stands for Zernike-Depth-Fourier-Ray-based features) show different properties and retrieval performances in recognizing non-rigid and partial 3D models. To automatically combine these three features, a meta similarity generation algorithm based on particle swarm optimization (PSO) [23] has been proposed to fuse their distance matrices (Fig. 13). This framework can be extended to integrate different or more features, so as to develop other similar unified retrieval algorithms for both non-rigid and partial 3D model retrieval.

Li et al. applied their framework on the partial retrieval track of SHREC'07 [36]. The results of their experiments show that their framework outperforms the methods participating

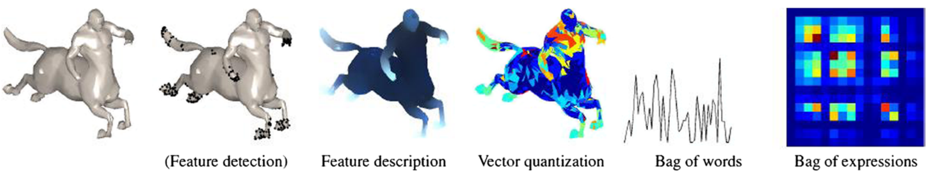


Fig. 11 Overview of the Shape Google pipeline [10]

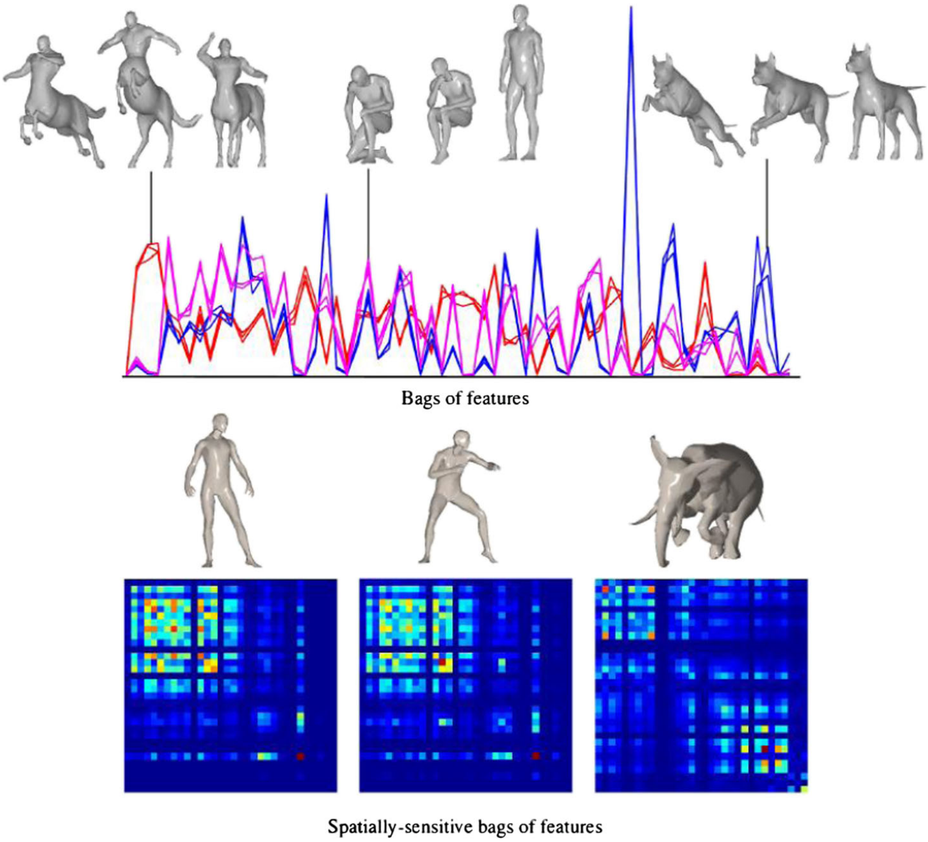


Fig. 12 Top row: examples of BoVW computed for different deformations of centaur (*red*), dog (*blue*), and human (magenta). Note the similarity of BoVW of different transformations and dissimilarity of BoVW of different shapes. Also note the overlap between the centaur and human BoVW, due to partial similarity of these shapes. Bottom row: examples of spatially-sensitive BoVW computed for different deformations of human (*left and center*) and elephant (*right*) [10]

in SHREC'07, as well as the methods of Tierny et al. [70], Biasotti et al. [7], Cornea et al. [15] and Toldo et al. [71], as illustrated in the NDGC curves provided in the paper of Li et al.

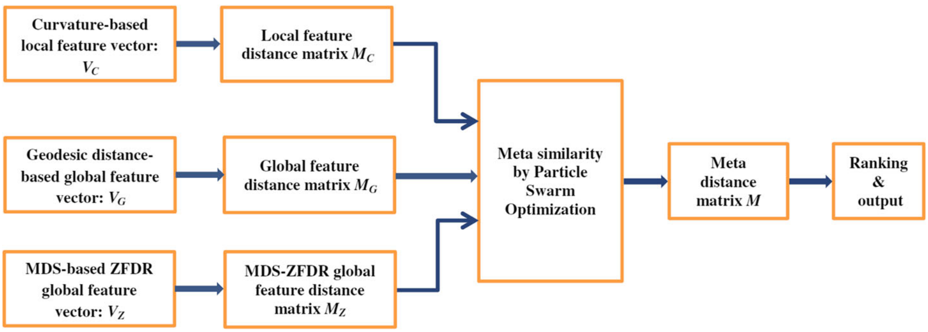


Fig. 13 Overview of the method of Li et al.[48]

5 Other methods

This subsection presents partial 3D object retrieval methods that cannot be solely grouped in one of the previous three categories.

5.1 Local feature histograms for range image classification

Hetzel et al. [39] explore a view-based method for the recognition of free-form objects in range images. They combine a set of local features, pixel depth, surface normal and curvature metrics in a multidimensional histogram, aiming to classify range scans, with pose-estimation as a byproduct. For the representation of surface curvature, they use the shape index of Dorai and Jain [20], which is also used in the previously presented local surface patch (LSP) method of Chen and Bhanu [13].

For histogram comparison, Hetzel et al. employ both χ^2-test-based histogram matching and maximum a-posteriori probability estimation. On ideal test images, both methods produce comparable results, with NN 93.58 for the χ^2-test-based method and NN 92.36 for the probabilistic method. However, the latter method is more capable to deal with partial occlusions, with NN 11.46 for the χ^2-test-based method, as opposed to NN 31.88 for the probabilistic method, in the case of 80 % occlusions. In addition, the latter method provides a measure of confidence for the obtained classification result. Experiments were performed on a collection of synthetic range images, taken from high-resolution polygonal models available on the authors' web site. However, Hetzel et al. do not perform comparisons with other methods in the presence of occlusions.

5.2 3D search and retrieval from range images using salient features

Furuya and Ohbuchi [24] proposed an enhancement of an earlier 3D object retrieval method, published by Ohbuchi et al. [57], which is based on BoVW and uses dense sampling for interest point extraction, followed by the calculation of SIFT descriptors [51]. Their earlier work outperformed state-of-the-art 3D object retrieval methods on MSB database, which contains highly articulated but geometrically simple objects. However, it only equaled the retrieval accuracy of state-of-the-art methods on PSB database, which contains rigid, detailed models.

Furuya and Ohbuchi [24] identified that the aforementioned limitations of the earlier work of Ohbuchi et al. were connected with the quality of the SIFT-based interest point extraction of that method. Figure 14(a) illustrates this issue by depicting the interest points extracted by the method of Ohbuchi et al. on a 3D model example. Furuya and Ohbuchi noted that the depth image of the potted plant produced a large number of small-scale features near the leaves. These features, being scale invariant, could match local geometrical features of other models that are similar in shape, yet completely different in scale. Consequently, the potted plant could potentially match models having completely different overall shape. On the other hand, an important large scale feature, in this case a large trapezoidal shape of the pot, is underrepresented. For simpler, less detailed shapes, e.g. those of MSB database, the original SIFT-based interest point detector worked very well. However, for the PSB database, which contains models having considerably more detail, the salient points cannot provide a balanced representation of model features.

In order to cope with the limitations of the SIFT-based interest point extractor, Furuya and Ohbuchi employed dense random sampling. For each image in the multi-scale image

(a) (b) (c)

Fig. 14 Example of feature points using: (**a**) SIFT-based interest point detector, (**b**) dense interest points, (**c**) a grid of interest points [24]

pyramid of the SIFT algorithm, the pixels to be sampled are drawn randomly from pixels having non-zero intensity value. A range image is rendered with zero pixel value as its background, and images in the SIFT pyramid are blurred according to their scale in the pyramid. Thus, a pixel from an image in the SIFT pyramid that has non-zero value is located on or near but not far from the image of the 3D model. As non-zero pixels are different for each image in the SIFT image pyramid, the positions of samples are different across scales in the SIFT pyramid. For comparisons, Furuya and Ohbuchi also implemented another sampling strategy, which samples the image at regular grid points. Figure 13b and c illustrate the interest points extracted by dense random sampling and grid sampling, respectively. It can be observed that dense sampling located more samples near the pot, whereas grid sampling is uniform, regardless of the image features across image scales. Apart from dense sampling, Furuya and Obhuchi's method adopted the SIFT feature descriptors of Lowe [51], the k-means clustering approach for codebook learning and the computationally efficient extremely randomized clustering (ERC)-tree of Guerts et al. [37] for vector quantization. Computational cost is further reduced by utilizing a graphics processing unit (GPU) implementation of SIFT feature extraction.

In their experiments, Furuya and Obhuchi used the PSB, MSB and ESB databases to comparatively investigate various aspects of their BoVW-based 3D object retrieval method, including codebook learning and encoding, sampling strategy and vocabulary size. Their main conclusions are: (i) although slow (approximately 230 s are required to learn a codebook of 1000 words), *k*-means clustering is preferable, since codebook learning is needed only once, (ii) ERC-tree [37] is much more efficient than *k*-means (at least 40 times faster) for nearest neighbor search, with slightly worse retrieval accuracy (approximately 1.5 %), (iii) the dense sampling approach is much less sensitive to vocabulary size than SIFT-based sampling, as illustrated in P-R curves provided in the paper of Furuya and Ohbuchi, (iv) Ohbuchi and Furuya's method outperforms, among others, the methods of Ohbuchi et al. [57], Chen et al. [14], Wahl et al. [74] and Kazhdan et al. [43] in all datasets. In quantitative terms, Furuya and Ohbuchi obtain R-precision 55.8/76.4/42.5 in PSB/MSB/ESB datasets, respectively, whereas Chen et al. obtain R-precision 45.9/56.9/34.7, Wahl et al. obtain R-precision 37.3/53.9/34.7 and Kazhdan et al. obtain R-precision 40.5/56.7/34.6, for the same retrieval performance measures and the same datasets.

6 Discussion

Table 2 summarizes the state-of-the-art methods on partial 3D object retrieval. It can be observed that the majority of the related literature is devoted on structured data represented by 3D meshes. These methods are also potentially applicable on unstructured data represented as point clouds, since there is an intense research activity on 3D mesh generation algorithms, which convert point clouds to meshes [53, 79].

Table 2 also reveals that *six out of thirteen partial 3D object retrieval methods are actually view-based*, essentially working in 2.5D. *This approach suits partial retrieval* in the sense that a 2D query model can be compared with all 2D projections of each target model in order to keep the minimum distance associated with the most similar projection, as suggested by Daras and Axenopoulos [17, 18]. It can also be recalled that the earlier comparative study of Shilane et al. [65] concluded that the 2.5D object retrieval methods, which were available at the time, outperformed 3D-based methods in terms of retrieval accuracy.

Another conclusion drawn from Table 2 is that *most partial 3D object retrieval methods are based on descriptors calculated over interest points*, either dense, or extracted by means of a salient point detector. Furuya and Ohbuchi [24] and Ohkita et al. [58] argue in favor of using dense points, although *the appropriate strategy might be representation/application-dependent*.

The BoVW paradigm is represented with three recent works. In the case of the work of Lavoué [46], his BoVW variations employ spatially sensitive information, which has been shown to enhance retrieval performance. Li et al. [48] used BoVW in the case of their curvature-based local features. It should be pointed out that local features are an intrinsic element of BoVW. Ohkita et al. [58] suggested that a shape matching which is based on all pairs of interest points of a query and a target model performs admirably well and only suffers in terms of computational complexity. It is tempting to speculate that, despite its combinatorial complexity, *this approach could perhaps be used for matching instead of BoVW, when the number of interest points is small enough to allow calculations for all respective pairs*.

Several partial 3D object retrieval methods have been evaluated in pre-existing generic retrieval benchmarks, such as the ESB, PSB and ISDB. An overview on such benchmarks, with information on their size, number of classes, availability, as well as with references on related publications, is provided in Table 3. It can be observed that SHREC tracks offer the most recent benchmarks, which range from databases of range images and watertight models to large scale databases. In 2007, SHREC has included a track specialized in partial retrieval. The accompanying datasets have been used for evaluation in several works presented in this survey. In 2009 and 2013, SHREC also included partial retrieval tracks [21, 67] whereas in 2010–2011, the same contest included tracks devoted to range scans [22]. However, in 2011 due to lack of participants, no results were published.

Table 4 summarizes the experimental comparisons reported in the reviewed works. The used benchmark datasets are provided along with the methods under comparison and the results obtained in terms of the performance measures used in each work. In cases of comparisons performed by means of P-R or NDCG curves, which cannot be condensed in a table, the means of performance evaluation and the "winner" method are apposed and the reader is referred to the original papers for details. Very often in the experimental comparisons, a partial 3D object retrieval method is shown to outperform a generic retrieval method, such as the methods of Vranic [73], Papadakis et al. [60] and Johnson and Hebert [42]. However, it should be kept in mind that some state-of-the-art partial 3D object retrieval methods aim to address special object classes, as is the case with the method of Agathos et al. [2], which focuses on the retrieval of articulated objects, and the methods of Lavoué [46] and Bronstein et al. [10], which focus on the retrieval of non-rigid objects.

Starting from the above discussion, several challenges associated with partial 3D object retrieval can be identified, pinpointing research directions and future work:

i. there is a need for systematic comparative studies, since as can be observed in Table 4, several state-of-the-art methods have not yet been compared with each other,
ii. the accuracy obtained by state-of-the-art partial retrieval methods is not yet sufficient for several practical applications. For example, the competing methods in SHREC'13 merely surpassed 20 % in FT,
iii. partial retrieval may well be used to cope with the difficulty and inefficiency of the 3D digitization process by means of a *predictive scanning* system, which, starting from a partially scanned query, may result in an accurate prediction of the complete object. Predictive scanning opens a whole new range of possibilities for decreased acquisition times, effort and cost. In addition it might be used to complement the information derived from low-cost scanning devices, such as Kinect. The work of Pauly et al. [62] was a step towards this direction,
iv. the formulation of hybrid methods combining multiple methodological paradigms, as is the case with the work of Li et al. [48], which uses both local and global descriptors, or with the work of Furuya and Ohbuchi [24], which is both view-based and BoVW-based,
v. research on bridging the gap between geometrical attributes and high-level semantic information is still rather limited. The use of prior knowledge might be inevitable to advance in this direction.

Acknowledgments This study received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 600533 PRESIOUS.

References

1. Adan A, Merchan P, Salamanca S (2011) 3D scene retrieval and recognition with depth gradient images. Patt Rec Lett 32(9):1337–1353
2. Agathos A, Pratikakis I, Papadakis P, Perantonis S, Azariadis P, Sapidis S (2010) 3D articulated object retrieval using a graph-based representation. Vis Comp 26(10):1301–1319
3. Agathos A, Pratikakis I, Perantonis S, Sapidis N (2009) Protrusion-oriented 3D mesh segmentation. Vis Comp 26(1):63–81
4. Akgul C, Axenopoulos A, Bustos B, Chaouch M, Daras P, Dutagaci H, Furuya T, Godil A, Kreft S, Lian Z, Napoleon T, Mademlis A, Ohbuchi R, Rosin PL, Sankur B, Schreck T, Sun X, Tezuka M, Yemez Y, Verroust-Blondet A, Walter M (2009) SHREC 2009 track: generic shape retrieval contest. Proc. Eurographics, 3DOR workshop
5. Atmosukarto I, Shapiro LG (2013) 3D object retrieval using salient views. Int J Mult Inform Retr 2:103–115
6. Axenopoulos A, Daras P, Dutagaci H, Furuya T, Godil A, Ohbuchi R (2009) SHREC 2009: shape retrieval contest of partial 3D models. Proc. Eurographics, 3DOR workshop
7. Biasotti S, Marini S, Spagnuolo M, Falcidieno B (2006) Sub-part correspondence by structural descriptors of 3D shapes. Comput-Aid Design 38(9):1002–1019
8. Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94(2):115–147
9. Bosch A, Zisserman A, Munoz X (2007) Image graphiclassification using random forests and ferns. Proc ICCV: 1–8

10. Bronstein AM, Bronstein MM, Guibas LJ, Ovsjanikov M (2011) Shape google: geometric words and expressions for invariant shape retrieval. ACM TOG 30(1):1–20
11. Chaouch M, Verroust-Blondet A (2006) Enhanced 2D/3D approaches based on relevance index for 3D-shape retrieval. Proc SMI: 36
12. Chen H, Bhanu B (2004) 3D free-form object recognition in range images using local surface patches. Proc ICPR 3:136–139
13. Chen H, Bhanu B (2007) 3D free-form object recognition in range images using local surface patches. Patt Rec Lett 28(10):1252–1262
14. Chen D-Y, Ouhyoung M, Tian X-P, Shen Y-T, Ouhyoung M (2003) On visual similarity based 3d model retrieval. Proc. Eurographics, 3DOR workshop: 223–232
15. Cornea ND, Demirci MF, Silver D, Shokoufandeh A, Dickinson S, Kantor PB (2005) 3D object retrieval using many-to-many matching of curve skeletons. Proc IEEE SMI: 366–371
16. Correa S, Shapiro L (2001) A new signature-based method for efficient 3-D object recognition. Proc. IEEE CVPR 1:769–776
17. Daras P, Axenopoulos A (2009) A compact multi-view descriptor for 3D object retrieval. Proc. CBMI, 115–119
18. Daras P, Axenopoulos A (2009) A 3D shape retrieval framework supporting multimodal queries. Int J Comp Vis 89:229–247
19. Daras P, Zarpalas D, Tzovaras D, Strintzis MG (2006) Efficient 3-d model search and retrieval using generalized 3-d radon transforms. IEEE Trans Multimedia 8(1):101–114
20. Dorai C, Jain A (1997) COSMOS-A representation scheme for 3D free-form objects. IEEE Trans Patt Anal Mach Intell 19(10):1115–1130
21. Dutagaci H, Godil A, Axenopoulos A, Daras P, Furuya T, Ohbuchi R (2009) SHREC 2009 track: querying with partial models. Proc. Eurographics, 3DOR workshop: 69–76
22. Dutagaci H, Godil A, Cheung CP, Furuya T, Hillenbrand U, Ohbuchi R (2010) SHREC 2010 track: range scan retrieval. Proc. Eurographics, 3DOR workshop: 109–115
23. Eberhart RC, Hu X (1999) Human tremor analysis using Particle Swarm Optimization. Proc Congress on Evol Comput: 1927–1930
24. Furuya T, Ohbuchi R (2009) Dense sampling and fast encoding for 3D model retrieval using bag-of-visual features. Proc ACM Int Conf Imag Vid Retriev
25. Gal R, Shamir A, Cohen-Or D (2007) Pose oblivious shape signature. IEEE Trans Vis Comput Graph 13(2): 261–271
26. Gao Y, Dai Q, Wang M, Zhang N (2011) 3D model retrieval using weighted bipartite graph matching. Signal Process: Image Commun 26:39–47
27. Gao Y, Dai Q, Zhang N-Y (2010) 3D Model comparison using spatial structure circular descriptor. Pattern Recogn 43(3):1142–1151
28. Gao Y, Tang J, Hong R, Yan S, Dai Q, Zhang N, Chua T-S (2012) Camera constraint-free view-based 3-D object retrieval. IEEE Trans Image Process 21(4):2269–2281
29. Gao Y, Tang J, Li H, Dai Q, Zhang N (2010) View-based 3D model retrieval with probabilistic graph model. Neurocomputing 73(10–12):1900–1905
30. Gao Y, Wang M, Ji R, Wu X, Dai Q (2014) 3-D object retrieval with Hausdorff distance learning. IEEE Trans Ind Electron 61(4):2088–2098
31. Gao Y, Wang M, Shen J, Dai Q, Zhang N (2010) Intelligent query: open another door to 3D object retrieval. Proc ACM Multimedia: 1711–1714
32. Gao Y, Wang M, Tao D, Ji R, Dai Q (2012) 3-D object retrieval and recognition with hypergraph analysis. IEEE Trans Image Process 21(9):4290–4303
33. Gao Y, Wang M, Zha Z-J, Tian Q, Dai Q, Zhang N (2011) Less is more: efficient 3-D object retrieval with query view selection. IEEE Trans Multimedia 13(5):1007–1018
34. Gao Y, Yang Y, Dai Q, Zhang N (2010) 3D object retrieval with bag-of-region-words. Proc ACM Multimedia: 955–958
35. Germann M, Breirenstein MD, Park IK, Pfister H (2007) Automatic pose estimation for range images on the GPU. Proc 3D Digital Imaging and Modeling: 81–90
36. Giorgi D, Biasotti S, Paraboschi L (2007) SHREC 2007 track: shape retrieval-watertight models, Proc. Eurographics, 3DOR workshop
37. Guerts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. Mach Learn 36(1):3–42
38. Hartveldt J, Spagnuolo M, Axenopoulos A, Biasotti S, Daras P, Dutagaci H, Furuya T, Godil A, Li X, Mademlis A, Marini S, Napoleon T, Ohbuchi R, Tezuka M (2009) SHREC 2009 track: structural shape retrieval on watertight models. Proc. Eurographics, 3DOR workshop
39. Hetzel G, Keibe B, Levi P, Schiele B (2001) 3D object recognition from range images using local feature histograms. Proc IEEE CVPR 2:394–399
40. Hoffman DD, Singh M (1997) Salience of visual parts. Cognition 63(1):29–78

41. Jayanti S, Kalyanaraman Y, Iyer N, Ramani K (2006) Developing an engineering shape benchmark for cad models. Comput Aided Des 38:939–953
42. Johnson AE, Hebert M (1999) Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Trans Patt Anal Mach Intell 21(5):433–449
43. Kazhdan M, Funkhouser T, Rusinkiewicz S (2003) Rotation invariant spherical harmonic representation of 3D shape descriptors. Proc Eurographics/ACM SIGGRAPH Symp Geom Process: 156–164
44. Kim DH, Park IK, Yun ID, Lee SU (2004) A new MPEG-7 standard: perceptual 3D shape descriptor. Proc Pacific Rim Conf Multimedia: 238–245
45. Koutsoudis A (2013) A 3D pottery content based retrieval benchmark dataset. http://www.ipet.gr/~akoutsou/benchmark
46. Lavoué G (2012) Combination of bag-of-words descriptors for robust partial shape retrieval. Vis Comput 28(9):931–942
47. Li F, Dai Q, Xu W, Er G (2010) Statistical modeling and many-to-many matching for view-based 3D object retrieval. Signal Process Image Commun 25(1):18–27
48. Li B, Godil A, Joham H (2013) Hybrid shape descriptor and meta similarity generation for non-rigid and partial 3D model retrieval. Accepted to Multimed Tools Appl
49. Li P, Ma H, Ming A (2013) Combining topological and view-based features for 3D model retrieval. Multimed Tools Appl 65:335–361
50. Lian Z, Rosin PL, Sun X (2010) Rectilinearity of 3D meshes. Int J Comput Vis 89(2):130–151
51. Lowe D (2004) Distinctive image features from scale-invariant keypoint. IJCV 60(2):91–110
52. Marini S, Paraboschi L, Biasotti S (2007) Proc IEEE SMI, SHREC 2007 track: partial matching: 13–16
53. Marton ZC, Rusu RB, Beetz M (2009) On fast surface reconstruction methods for large and noisy datasets. Proc IEEE ICRA: 2829–2834
54. Mian AS, Bennamoun M, Owens R (2006) Three-dimensional model-based object recognition and segmentation in cluttered scenes. IEEE Trans Patt Anal Mach Intell 28(10):1584–1601
55. O'Hara S, Draper BA (2011) Introduction to the bag of features paradigm form image classification and retrieval, arXiv. doi 1101.3354
56. Ohbuchi R, Osada K, Furuya T, Banno T (2008) Salient local visual features for shape-based 3D model retrieval, Proc. IEEE SMI: 93–102
57. Ohbuchi R, Osada K, Furuya T, Banno T (2008) Salient local visual features for shape-based 3D model retrieval. Proc IEEE SMI: 93–102
58. Ohkita Y, Ohishi Y, Furuya T, Ohbuchi R (2012) Non-rigid 3D model retrieval using set of local statistical features. Proc IEEE Multim Expo Workshops: 593–598
59. Papadakis P, Pratikakis I, Perantonis S, Theoharis T (2007) Efficient 3D shape matching and retrieval using a concrete radicalized spherical projection representation. Patt Rec 40(9):2437–2452
60. Papadakis P, Pratikakis I, Theoharis T, Passalis G, Perantonis S (2008) 3D object retrieval using an efficient and compact hybrid shape descriptor. Proc Eurographics, 3DOR workshop
61. Papadakis P, Pratikakis I, Theoharis T, Perantonis SJ (2010) PANORAMA: A 3D shape descriptor based on panoramic views for unsupervised 3D object retrieval. Int J Comp Vis 89(2–3):177–192
62. Pauly M, Mitra NJ, Giesen J, Guibas L, Gross M (2005) Example-based 3D scan completion. Proc. Symp Geometry Processing: 23–32
63. Pu J, Ramani K (2005) An approach to drawing-like view generation from 3D models. Proc. IDETC/CIE, ASME
64. Sfikas K, Pratikakis I, Koutsoudis A, Savelonas M, Theoharis T (2013) 3D object partial matching using panoramic views. Proc ICIAP, MMC4H workshop, Naples, Italy
65. Shilane P, Min P, Kazhdan M, Funkhouser T (2004) The Princeton shape benchmark. Shape Modeling International
66. Siddiqi K, Zhang J, Macrini D, Shokoufandeh A, Bouix S, Dickinson S (2008) Retrieving articulated 3D models using medial surfaces. Mach Vis Appl 19(4):261–274
67. Sipiran I, Meruane R, Bustos B, Schreck T, Johan H, Li B, Lu Y (2013) SHREC 2013 track: large-scale partial shape retrieval using simulated range images. Proc. Eurographics, 3DOR workshop: 81–88
68. Stavropoulos G, Moschonas P, Moustakas K, Tzovaras D, Strintzis MG (2010) 3D model search and retrieval from range images using salient features. IEEE Trans Multim 12(7):692–704
69. Tierny J, Vandeborre JP, Daoudi M (2008) Enhancing 3D mesh topological skeletons with discrete contour constrictions. Vis Comput 24(3):155–172
70. Tierny J, Vandeborre J, Daoudi M (2009) Partial 3D shape retrieval by Reeb pattern unfolding. Comput Graph For 28(1):41–55

71. Toldo R, Castellani U, Fusiello A (2009) Visual vocabulary signature for 3D object retrieval and partial matching. Proc Eurographics, 3DOR workshop. 21–28
72. Veltkamp RC, Giezeman G-J, Bast H, Baumbach T, Furuya T, Giesen J, Godil A, Lian Z, Ohbuchi R, Saleem W (2010) SHREC 2010 Track: Large Scale Retrieval. Proc. Eurographics, 3DOR workshop. 63–69
73. Vranic D (2004) 3D model retrieval. PhD thesis, University of Leipzig, Germany
74. Wahl E, Hillenbrand U, Hirzinger G (2003). Surflet-pair-relation histograms: a statistical 3D-shape representation for rapid classification, Proc 3DIM. 474–481
75. Wang M, Gao Y, Lu K, Rui Y (2013) View-based discriminative probabilistic modeling for 3D object retrieval and recognition. IEEE Trans Image Process 22(4):1395–1407
76. Wang F, Li F, Dai Q, Er G (2008) View-based 3D object retrieval and recognition using tangent subspace analysis. Proc. SPIE 6822, Visual Communications and Image Processing
77. Wen Y, Gao Y, Hong R, Luan H, Liu Q, Shen J, Ji R (2012) View-based 3D object retrieval by bipartite graph matching. Research Collection School of Information Systems. 897–900
78. Zarpalas D, Daras P, Axenopoulos A, Tzovaras D, Strintzis MG (2007) 3D model search and retrieval using the spherical trace transform. EURASIP J Adv Sig Proc
79. Zhou K, Gong M, Huang X, Guo B (2011) Data-parallel octrees for surface reconstruction. IEEE Trans Vis Graph 17(5):669–681

Michalis A. Savelonas received the B.Sc. degree in Physics in 1998, the M.Sc. degree in Cybernetics in 2000 with honors, and the Ph.D. degree in the area of Image Analysis in 2008 with honors, all from the University of Athens, Greece. For his Ph.D. research he received a scholarship by the Greek General Secretariat of Research and Technology (25 %) and the European Social Fund (75 %). He is currently a research fellow in Athena Research and Innovation Center, branch of Xanthi, as well as in the Dept. of Electrical and Computer Engineering in the Democritus University of Thrace. In addition, he is a regular research associate of the Dept. of Informatics and Telecommunications of the University of Athens, Greece, as well as of the Center for Technological Research of Central Greece. In the past, he has served in various academic positions in the University of Athens, University of Houston, TX, USA, Hellenic Air Force Academy and Technological Educational Institute of Lamia. In 2002–2004 he has been working in software industry for the development of real-time systems within the context of various projects. Dr. Savelonas has co-authored more than 50 research articles in peer-reviewed international journals, conferences and book chapters, whereas he has been actively involved in several EU, US and Greek R&D projects. He is a reviewer in prestigious international journals including Pattern Recognition and IEEE Trans. Image Processing. His research interests include image analysis, pattern recognition, 3D retrieval and biomedical applications.

Ioannis Pratikakis is an Assistant Professor at the Department of Electrical and Computer Engineering of Democritus University of Thrace in Xanthi, Greece. He received the Ph.D. degree in 3D Image analysis from the Electronics engineering and Informatics department at Vrije Universiteit Brussel, Belgium, in January 1999. From March 1999 to March 2000, he joined IRISA/ViSTA group, Rennes, France as an INRIA postdoctoral fellow. From January 2003 to June 2010, he was working as Adjunct Researcher at the Institute of Informatics and Telecommunications in the National Centre for Scientific Research "Demokritos", Athens, Greece. His research interests lie in image processing, pattern recognition, vision and graphics, and more specifically, in document image analysis and recognition, medical image analysis as well as multimedia content analysis, search and retrieval with a particular focus on visual content. He has served as co-chair of the Eurographics Workshop on 3D object retrieval (3DOR) in 2008 and 2009 as well as Guest Editor for the Special issue on 3D object retrieval at the International Journal of Computer Vision. He is Senior Member of the IEEE, member of the Board of the Hellenic Artificial Intelligence Society for the period 2010–2014 and member of the European Association for Computer Graphics (Eurographics).

Konstantinos Sfikas is currently a Ph.D. candidate at the department of Informatics and Telecommunications, National & Kapodistrian University of Athens. He received his B.Sc. in Informatics from the Technological Educational Institute of Athens in 2006 and his M.Sc. in Advanced Informatics Systems from the National and Kapodistrian University of Athens in 2008. His main research interests are in the field of computer graphics, with special focus on both rigid and non-rigid 3D object retrieval.