www.cadanda.com
Computer-Aided Design AND APPLICATIONS™

# A Review on 3D Object Retrieval Methodologies Using a Part-Based Representation

Panagiotis Theologou[1], Ioannis Pratikakis[2] and Theoharis Theoharis[3,4]

[1]Democritus University of Thrace, ptheolog@ee.duth.gr
[2]Democritus University of Thrace, ipratika@ee.duth.gr
[3]University of Athens, theotheo@di.uoa.gr
[4]Norwegian University of Science and Technology, theotheo@idi.ntnu.no

## ABSTRACT

A comprehensive overview of 3D Object Retrieval methodologies that use a part-based representation is presented. Taking into account the typical operational pipeline we detail each distinct module and we provide a comparative study between the individual modules as well as the global methodologies. This study relies upon the 3D mesh segmentation scheme used, the feature extraction method chosen, as well as the selected graph matching methodology.

**Keywords:** 3D object retrieval, 3D mesh segmentation, part-based representation.

## 1. INTRODUCTION

The rapid growth in the field of 3D capture technologies has led to the continuous expansion and enlargement of 3D object databases. Therefore, the construction of retrieval algorithms that enable efficient and effective 3D object retrieval from either public or proprietary 3D databases is becoming a necessity. 3D object retrieval is the process which retrieves 3D objects from a database in a ranked order so that the higher the ranking of an object the better the match to a 3D object query, using a measure of similarity. Various methods exist in the literature, which use a global descriptor to represent the object. However, according to Biederman [11] humans tend to recognize objects by analyzing the semantics of their parts. This leads to the hypothesis that two objects are similar, if they consist of similar parts.

In this paper, we strive towards presenting a comprehensive overview of methodologies on 3D Object Retrieval using a part-based representation detailing the key aspects, and discussing advantages and pitfalls of each methodology. In the sequel, the fundamental operational pipeline of retrieval methodologies using a part-based representation is described (Section 2), the existing methodologies in the field are presented (Section 3), and finally a comparative discussion is made (Section 4).

## 2. FUNDAMENTAL OPERATIONAL PIPELINE

The concept of content-based 3D object retrieval involves a query object and a database of 3D objects. The goal is to identify 3D objects in the database that match the query. Practically, database objects will not be identical to the query. Therefore, the retrieval process usually means bringing from the database the $k$ most similar objects to the query. This requires the definition of a similarity/dissimilarity measure between the object in the database and the query. The $k$ objects with the lowest dissimilarity value are the most similar objects to the query, and are retrieved from the database. Representative surveys on the subject are [44],[40],[27].

A typical operational pipeline for 3D object retrieval methodologies using a part-based representation is shown in Fig. 1. First, a 3D mesh segmentation algorithm is applied to all the objects of the database. 3D mesh segmentation is defined as the decomposition of a 3D mesh into meaningful parts, i.e. the identification of the distinct parts that comprise a mesh and constitute a partitioning as close as possible to human perception. Given an input mesh $M$ consisting of a set of vertices $V$, a set of faces $F$, and a set of edges $E$, a mesh segmentation algorithm performs a partition of any of the aforementioned sets into $n$ disjoint subsets, resulting in a division



Taylor & Francis

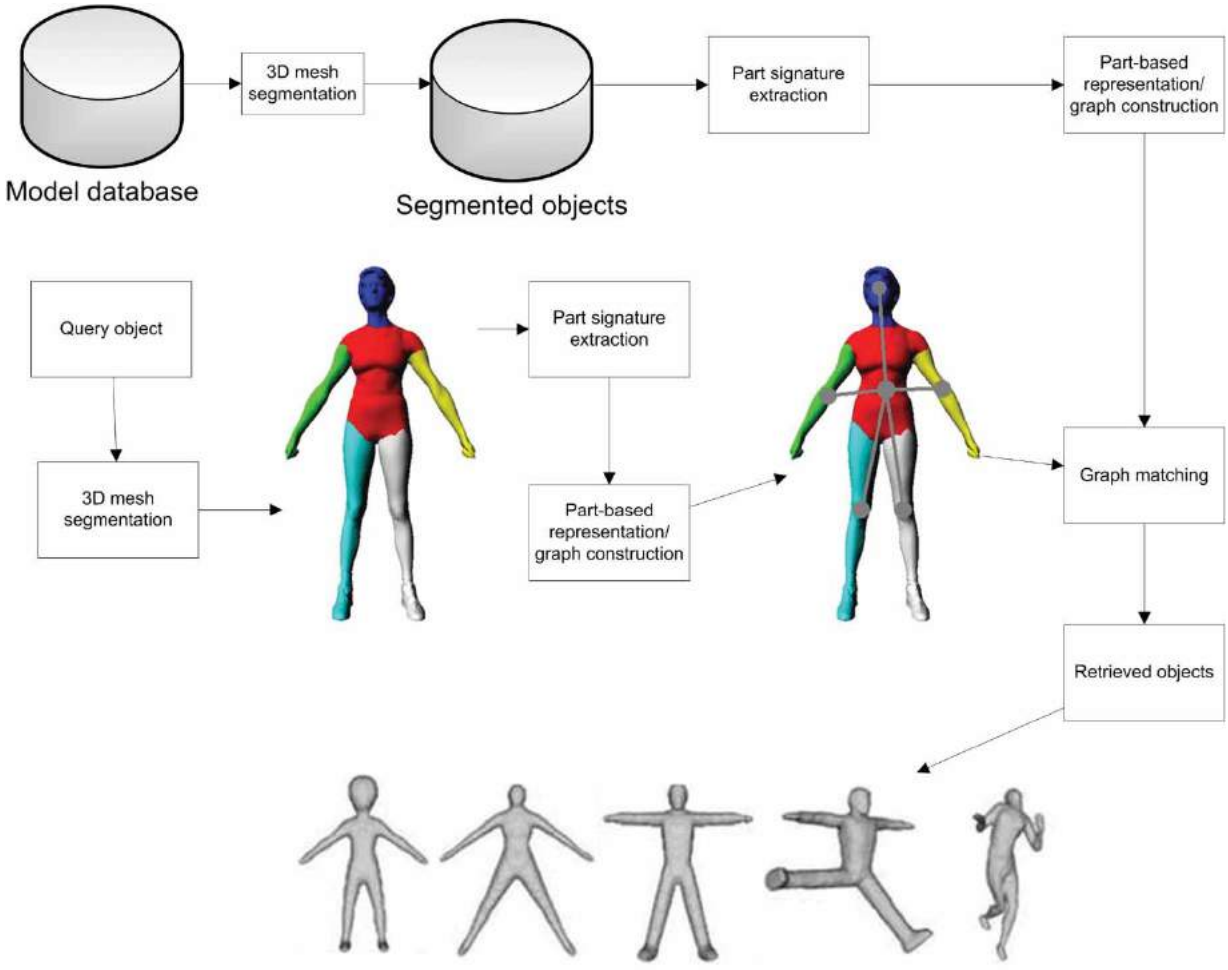


Fig. 1: Typical operational pipeline of part-based shape retrieval methods.

of the object into $n$ parts. For example, a meaningful segmentation of a 'human object' would result in the division of the mesh into head, torso, legs and arms. Recent surveys on 3D mesh segmentation can be found in [2],[34].

The produced partitions along with the adjacency of the parts are the basis for the construction of a graph representation of the object. Each part is represented as a node while edges express the relation between adjacent parts. The next step is feature extraction and is applied to each extracted part. Each node of the graph is associated with a feature vector with attributes of the corresponding segment, which take into account aspects like geometry, topology, etc. Furthermore, each edge of the graph can be attributed to pairwise features, i.e. features that correspond to the boundary of two segments or features that involve two adjacent regions in general.

The final step involves graph matching. The problem can be summarized as follows: given two attributed graphs the goal is to find the minimum distance that represents the dissimilarity of the two objects. Thus, a correspondence has to be found between nodes of two distinct graphs. To measure similarity, a distance measure is defined for two parts. At this point, a choice of the distance metric between vectors is necessary in order to compare the corresponding descriptors of the parts.

## 3. RETRIEVAL METHODS USING A PART-BASED REPRESENTATION

In this section, we present the state-of-the-art methodologies for 3D object retrieval using a part-based representation. They have been classified in terms of the core feature which characterizes each of them. In particular, we present methods based on (i) Attributed Relational Graphs, (ii) Reeb Graphs, (iii) skeletons, (iv) surface-type segmentation, (v) contextual part analogies, and (vi) style-content separation.

### 3.1. Attributed Relational Graph-Based Methods

#### 3.1.1. MPEG-7 Standard: Perceptual 3-D Shape Descriptor

This method [24] uses convexity and graph representation for 2D and 3D object segmentation. The author describes three steps: the initial decomposition stage (IDS), the recursive decomposition stage (RDS), and the iterative merging stage (IMS) (Fig. 2). First the opening operation of mathematical morphology is performed to the binary image representing the object. This operation results in a decomposition of the object by rounding its corners. Multiple decompositions are produced by alternating the parameter $k$ (the radius of the circle) of the ball shaped structuring element $S(k)$, by which the opening is calculated. The best segmentation is selected as the one with the highest weighted convexity value, which is the sum of the convexity value of all parts weighted by their normalized volume. After the initial segmentation, each part is hierarchically decomposed with the same method until a split condition is no longer met. Following the hierarchical segmentation step the merging stage checks for over-segmented parts. This is achieved by calculating the difference between the convexity of the merged part and the weighted convexity of the constituent parts.

The Perceptual 3-D Shape Descriptor is defined in [23]. An Attributed Relational Graph (ARG) representing object features is constructed. Each part represents a node of the graph. Each part is represented by an ellipsoidal blob. There are 4 unary attributes used to describe each node and 3 binary attributes describing edges or relations between nodes. The attributes used are:

- The volume $V$ of the segment
- The convexity $C$ of the segment defined as the ratio of the volume of the object to the volume of its convex hull
- Two eccentricity values $E_1 = \sqrt{1 - c^2/a^2}$ and $E_2 = \sqrt{1 - c^2/b^2}$, where $a$, $b$, and $c$ are the three maximum ranges along the three principal axes.

Edge features include:

- Distances between centers of ellipsoid segments
- The angle between the first principal axes of two adjacent segments
- The angle between the second principal axes of two adjacent segments

The comparison between two graphs is performed using the Double earth mover's distance (EMD).

#### 3.1.2. Retrieval of 3D Articulated Objects Using a Graph-Based Representation

This work [1] decomposes objects using the method described in [3]. Geodesic extrema of an object are considered salient points. The integral geodesic function is used to this end. The core partition is approximated by starting from the minimum of the geodesic function and expanding the partition. When expansion is completed the protrusion parts will have been separated by 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 features are assigned to each node and pairwise features are assigned to each edge. The graph matching of the method is based on EMD of the feature vectors. Unary attributes assigned to the nodes are size, convexity, eccentricities of the ellipsoid approximating the component [23], the spherical harmonic descriptor vector [31]. The binary attributes assigned to the edges of the graph are the distance of the centroids of the segments 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).

#### 3.1.3. Non Rigid 3D Object Retrieval Using Topological Information Guided by Conformal Factors

[33] uses a decomposition of the objects into parts using the conformal factor [6] values of the mesh. This is achieved by quantizing the discrete conformal factor values calculated on each face. The latter leads to the construction of an attributed graph (Fig. 3) with the following features:

- Mean discrete conformal factor value of the segment
- Normalized area of the segment
- Geodesic length between borders of the segment

After the graph construction, graph matching is performed. The matching first locates the *core* node of each graph. All other nodes are used to form strings ending at the core node. The idea is to perform a matching between strings. Consequently, a distance between two strings $p$, $q$ is defined as the sum of the $L_1$ distances of the corresponding nodes. The distance between two objects is the minimum distance of a correspondence between strings of the two different objects. Let $m$ and $n$ denote the cardinality of strings of two objects, then the assignment problem is solved with the Hungarian Algorithm in the $m \times n$ matrix representing the distances between all string of two objects. The matrix becomes square by padding with

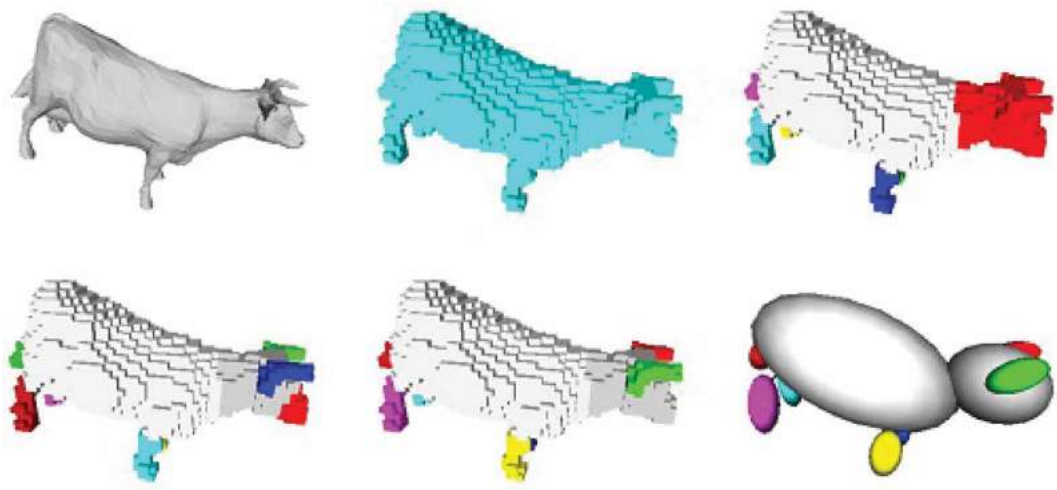

Fig. 2: Segmentation using the [24] method.

the following penalizing factor *PF*:

$$PF = \frac{|m-n|}{m+n} \tag{1}$$

### 3.2. Reeb Graph-Based Methods

The use of Reeb graphs [8] for retrieval by parts is described extensively in [4]. Let $f : M \to \mathbb{R}$ be a real continuous function defined on a mesh $M$. The Reeb graph is the quotient space of the graph of $f$ in $M \times \mathbb{R}$ by the equivalence relation $(X_1, f(X_1)) \sim (X_2, f(X_2))$ if and only if $f(X_1) = f(X_2)$, and $X_1$ and $X_2$ are in the same connected component of $f^{-1}(f(X_1))$. A sample segmentation is depicted in Fig. 4. Representative functions used in the literature for Reeb graphs are the:

- Height function, which defines a point's height on the mesh as 0 (ground) and the remainder get values related to the ground.
- Distance of each point to the center of mass.
- Integral geodesic function, which is defined as the integral of the geodesic distances of each point to all the others. If $M$ is an input mesh, then the integral geodesic of point $v$ is defined as:

$$IG(v) = \int_{p \in M} g(v,p)\,dM \tag{2}$$

#### 3.2.1. *Multiresolution Reeb Graphs*

The Multiresolution Reeb Graph (MRG) is introduced in [20]. The normalized integral geodesic is used to construct the Reeb graph. The method in [7] is an application of the MRG approach used for CAD models. The first step is the construction of the Reeb graph using a normalized approximation of the sum of geodesics as function. Multiple graphs of the same object are constructed with various resolutions (number of nodes). For each node $m$ of the Reeb graph two attributes are defined:

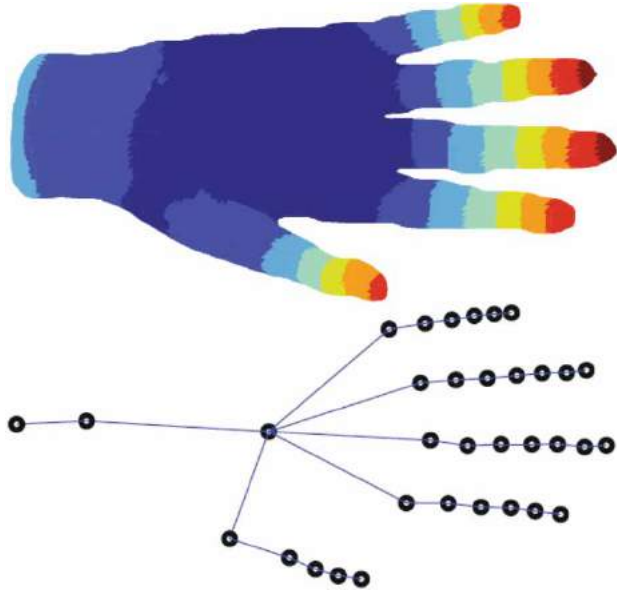

Fig. 3: Segmentation and the corresponding graph using [33].

$$a(m) = \frac{1}{rnum} \cdot \frac{area(m)}{area(M)} \tag{3}$$

$$l(m) = \frac{1}{rnum} \cdot \frac{len(m)}{\sum_n len(n)} \tag{4}$$

where *rnum* is the resolution number of the MRG, $area(M)$ and $area(m)$ denote the area of the object $M$ and the node $m$ respectively. $len(m)$ is the difference of the maximum and the minimum values of

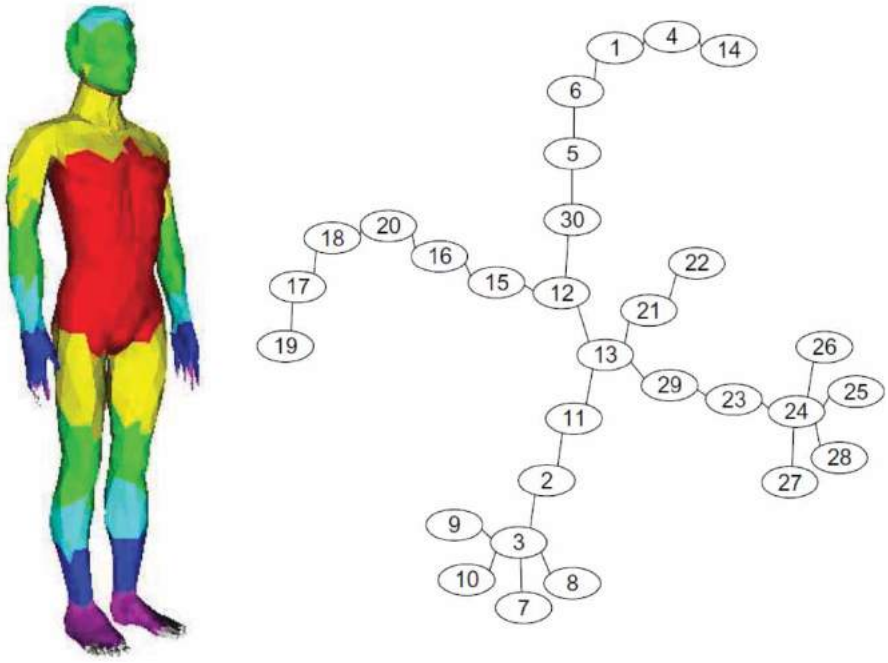


Fig. 4: A human model segmented using Reeb graphs with the integral geodesic function (taken from [4]).

the Reeb function in the segment $m$. The matching of two objects $M_1$, $M_2$ is performed as described below. First a similarity measure between two nodes $m$, $n$, is defined:

$$sim(m,n) = w \cdot \min(a(m), a(n)) + (1-w) \cdot \min(l(m), l(n)) \quad (5)$$

where $w$ is a constant controlling the contribution of the two terms. The graph matching is performed hierarchically, meaning that first nodes at lower resolution are matched and children of the nodes are considered iteratively. When all nodes are matched, the global similarity can be computed as the sum of all matched pairs in $P$:

$$SIM(M_1, M_2) = \sum_{(m,n)\in P} sim(m,n) \quad (6)$$

The Augmented MRG is introduced in [43]. Each node of the MRG is enhanced with attributes so that the matching is more efficient. The graph nodes are represented in spherical coordinates. The radius and the two angles are used as attributes of the node. In addition volume, area, and curvature are used to characterize each node.

### 3.2.2. *Sub-Part Correspondence by Structural Descriptors of 3D Shapes*

The first step of this work [10] is the construction of the Extended Reeb Graph (ERG) of the object. The ERG is directed and acyclic (taking into account function monotony) and is defined on a surface where a finite set of contours is defined. The functions considered are the distance to the center of mass and the integral geodesic. Since the graph is directed, each node is associated with a sub-graph containing all nodes to the leaves. The signature used for each node (sub-graph) is defined in [22] and the result of the spherical harmonic analysis of the corresponding sub-part. A distance between two sub-graphs is defined:

$$d(u_1, u_2) = \frac{w_1 G_s + w_2 St_s + w_3 Sz_s}{3} \quad (7)$$

where $G_s$, $St_s$, $Sz_s$ are the geometric, structural and size distance between the two sub-graphs respectively. In addition $w_1$, $w_2$, and $w_3$ are weights depending on the application. A matching between two graphs $G_1$ and $G_2$ is achieved by defining the following distance measure with respect to the common sub-graph $G$, which is described in [9]:

$$D(G_1, G_2) = 1 - \frac{\sum_{u\in G}(1 - d(\psi_1(u), \psi_2(u)))}{\max(|G_1|, |G_2|)} \quad (8)$$

where $\psi_1$, $\psi_2$ denote sub-graph isomorphisms from $G$ to $G_1$ and from $G$ to $G_2$ respectively. $|G_i|$ denotes the number of nodes of graph $G_i$.

### 3.2.3. *Partial 3D Shape Retrieval by Reeb Pattern Unfolding*

The method in [42] uses Reeb graphs to segment an object. The segments are produced using the method in [41]. 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 center of the chart and the boundaries on the outer circle of the disk. Disk-like charts correspond to the fingers while annulus-like chart corresponds to the palm of a hand object. Let $c_i$ be the disk-like chart of a segment. If $\varphi_i$ is the mapping of $c_i$ to the canonical planar domain $D$, then the unfolding signature $\lambda_{\varphi_i}$ can be defined as follows:

$$\lambda_{\varphi_i}(\rho) = \frac{A_{c_i}(\rho)}{A_{D(\rho)}} = \frac{A_{c_i}(\rho)}{\pi\rho^2} \tag{9}$$

where $\rho$ denotes a subset of the chart, and $A_{c_i}$, $A_D$ 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:

$$\lambda_{\varphi_i}(\rho) = \frac{A_{c_j}(\rho)}{A_{D(\rho)}} = \frac{A_{c_j}(\rho)}{\pi(\rho+1)^2 - \pi} \tag{10}$$

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 number of the disk-like and annulus-like Reeb charts in $P_i$, which are linked by the following equation with $g_{P_i}$ denoting the genus of the Reeb pattern:

$$n_D(P_i) = n_A(P_i) + 1 - 3g_{P_i} \tag{11}$$

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:

$$s(c_{A_i}, c_{B_j}) = 1 - L_{N1}(c_{A_i}, c_{B_j}) \tag{12}$$

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

### 3.3. Skeleton-Based Methods

#### 3.3.1. *Retrieving Articulated 3D Models Using Medial Surfaces*

This method [39] matches objects using medial surfaces. The medial skeleton of an object is extracted using a topology preserving thinning algorithms. The classification of the points lying on the skeleton results in an automatic segmentation of the model (Fig. 5). Nodes are used to construct a graph of the object with edges connecting adjacent nodes. A bipartite graph matching method is employed to find the best match in the matching process while the distance used to measure similarity between nodes is the Euclidean distance of the mean curvature histogram vectors.

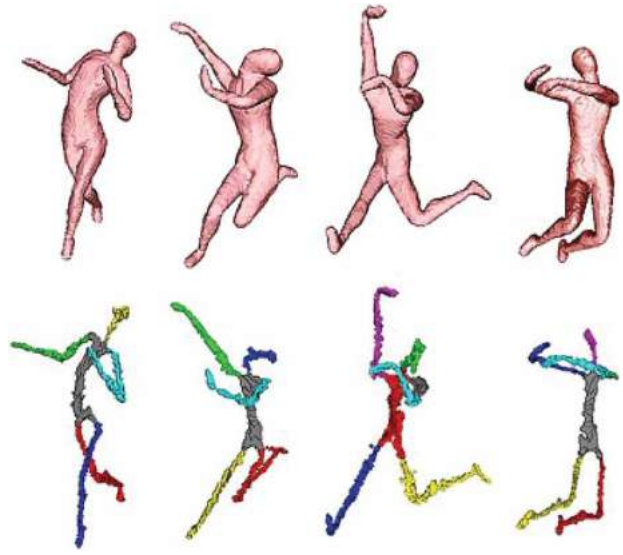

Fig. 5: Segmentation using the medial surface approach (taken from [39]).

#### 3.3.2. *Shape Retrieval Using Shock Graphs*

The work in [13] introduces a new indexing structure for retrieval, using the shock-graph feature matching method in [14]. 2-D silhouettes of an object are required. The medial axis of a silhouette forms the skeleton of the object. Nodes on the skeleton (shocks) are processed. Each shock point is characterized by the radius of the maximal bitangent circle centered at the point. Edges are weighted by the Euclidean distance between adjacent points. Graph embedding techniques are used to transform the shock graphs and finally the EMD is used for the graph matching (Fig. 6).

### 3.4. Methods Using Surface-Type Segmentation

#### 3.4.1. *Salient Geometric Features for Partial Shape Matching*

The first step of this method [16] is the segmentation of the model into patches approximated by quadric surfaces. The goal is to find salient features, i.e. regions of the surface where big difference in curvature exists. A saliency score is defined to this end for each patch $F$:

$$S = \sum_{d \in F} W_1 Area(d) Curv(d)^3 + W_2 N(F) Var(F) \tag{13}$$

where $Area(d)$ denotes the area of the triangle $d$, $Curv(d)$ the corresponding Gaussian curvature, $N(F)$, $Var(F)$ the curvature variance in the patch. The $W_1$ and $W_2$ weights control the contribution of the two terms. Top ten percent of the patches, or the ones that surpass some threshold are considered salient geometric features. Salient features are indexed and geometric hashing is employed to locate them in a retrieval application (Fig. 7).

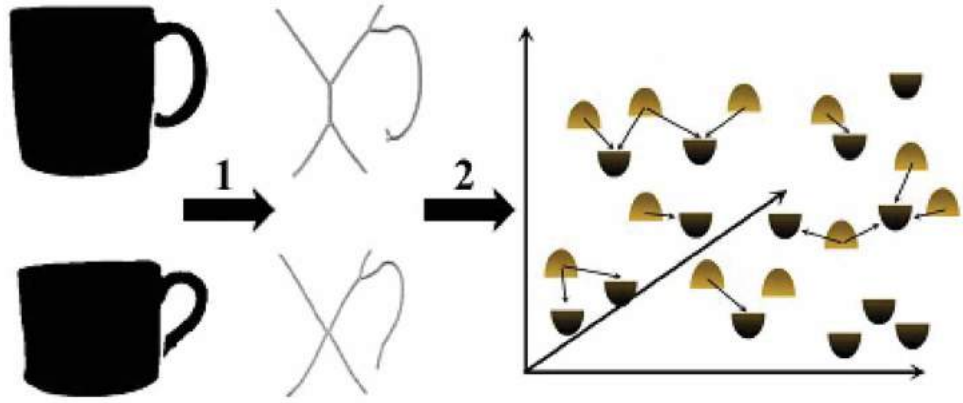


Fig. 6: Overview of the shock graph based method which includes skeleton extraction and EMD matching (taken from [13]).

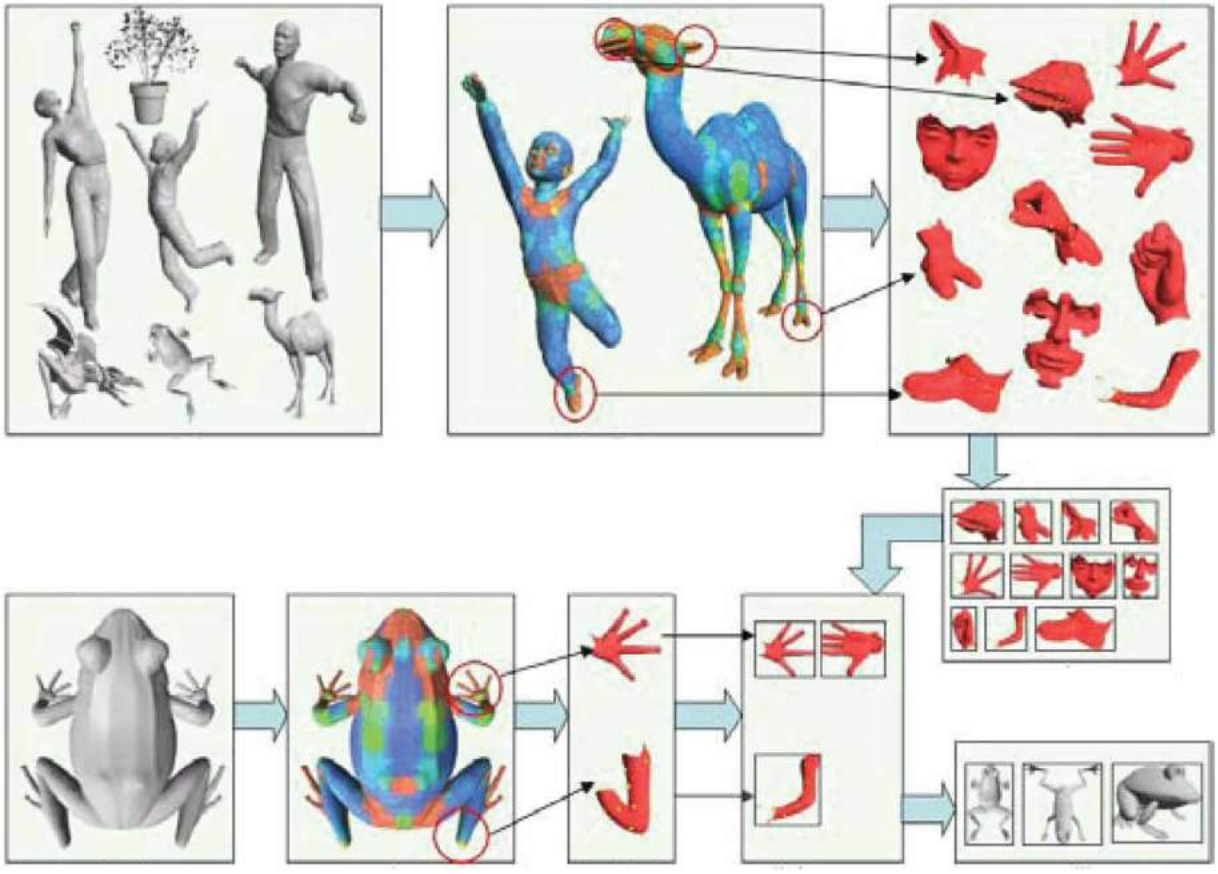

Fig. 7: Salient geometric feature extraction and partial matching (taken from [16]).

### 3.4.2. *Bag-of-Words Descriptors for Partial Shape Retrieval*

The method in [26] first segments a mesh into patches using Lloyd's algorithm [29]. The latter first divides the mesh surface into patches and iteratively moves the centers of the patches until convergence. A feature point is associated with every patch, for which a descriptor is computed. For each feature point the descriptor is the spectral amplitude vector $c^i = [c^i_1, \ldots, c^i_{n_c}]$, where $c^i_k$ the $k$th spectral coefficient amplitude of the patch $p_i$. The first $n_c$ coefficients are chosen. For each patch individually the Laplace Beltrami operator is computed, the eigendecomposition of which is used in the descriptor. The $k$th spectral coefficient amplitude corresponding to the $k$th vertex of a patch containing $m$ vertices is defined as:

$$c_k = \sqrt{x_k^2 + y_k^2 + z_k^2} \tag{14}$$

where $x_k = \sum_{i=1}^{m} \chi_i D_{i,i} H_i^k$ and $y_k$ and $z_k$ are defined in the same way. $\chi_i$ denotes the $x$ coordinate of the $i$th vertex, $H_i^k$ is the $i$th value of the $k$th eigenvector. $D$ denotes the Lumped Mass matrix. A dictionary is constructed with a simple k-means clustering ($n_w$ clusters) on a large dataset. Then two Bag-of-Words (BoWs) descriptors are defined. The standard BoW $b^M$ of a mesh $M$ is a histogram of size $n_w$ containing the distribution of the visual words. Each patch is assigned the word with the minimum $L_2$ distance (Fig. 8). The spatial BoW $B^M$ is a $n_w \times n_w$ matrix which takes into account adjacency between patches. The distance between two models $M_1$, $M_2$ is defined using both the spatial and standard BoWs:

$$d(M_1, M_2) = 6 \times \left| b^{M_1} - b^{M_2} \right| + \left| B^{M_1} - B^{M_2} \right| \qquad (15)$$

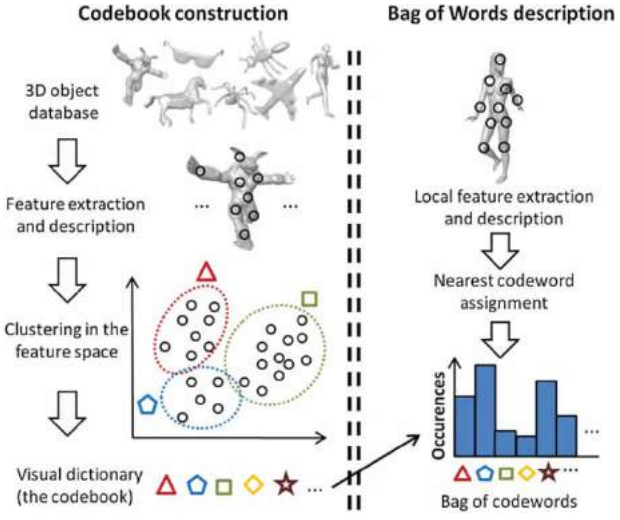


Fig. 8: The BoW approach (taken from [26]).

### 3.4.3. *Thesaurus-Based 3D Object Retrieval with Part-in-Whole Matching*

[15] employs the Hierarchical Fitting Primitives (HFP) [5] algorithm for the segmentation step. The HFP is a hierarchical clustering method, which starts with every face in a separate segment and iteratively merges regions until all faces are in the same cluster. The criterion to merge adjacent regions is proximity to a primitive shape (plane, cylinder, sphere). All the above multilevel segmentation steps are stored for each shape in the database, which is called hierarchically segmented mesh (HSM). For each segment of the HSM tree in the database the spherical harmonic descriptor (SH) [22] is computed. The Euclidean Distance between two segments' SH is used to measure similarity between objects. The level of segmentation in the HSM selected for each object is determined by the number of similar parts among all the objects of the database. The last step results in a large pool of shape segments, which is indexed using k-means clustering creating a thesaurus. The query object is matched with a term of the thesaurus and the corresponding object is retrieved.

## 3.5. Contextual Part Analogies in 3D Objects

This work [37] uses the Shape Diameter Function (SDF) [36] to segment objects. Shape Diameter is defined as the local diameter on each face. The histogram of the SDF values is processed using a Gaussian Mixture Model in combination with the Expectation Maximization algorithm to create parts. The segmentation is performed in a hierarchical way. A $k$ is selected as the expected levels of segmentation. The first level segmentation is performed by selecting the maximum Gaussian mean as the core of the object and the second mean corresponds to the second level parts. For each level of segmentation a new mean Gaussian value is selected and the parts that correspond to that value are separated from the rest of the mesh. The next step is assigning an attribute vector to each part. The following features are considered for the part signature:

- HSDF: the normalized histogram of SDF values of the segment.
- SD(D1,D2,D3,A3): Shape-distribution signatures which include distances among multiple random points picked on the surface of the mesh (D1,D2,D3) and angles (A3) taken from [30].
- CG: The conformal factor values from [6].

The local HSDF measure between two segments $p$, $p'$ is defined below respecting both size and SDF values of the segment. The $L_1$ norm is chosen for the normalized histogram. A bipartite graph is constructed to compare two segments of two distinct objects, with nodes representing part hierarchies and edges connecting part that do not belong to the same object. A capacity value is assigned to each edge $(q, q')$ of the graph:

$$capacity(q, q') = \frac{1}{d(q, q') + \varepsilon} - 1 \qquad (16)$$

where $d(q, q')$ is the feature vector with geometric properties as described previously. Two nodes are added to the graph, a sink T and a source S with capacity set to $1,5 \times capacity(p, p')$. The measure that is used to calculate the distance between the two parts is a context-aware distance, which takes into account hierarchy (Fig. 9), and is defined with respect to the maximum flow (*flow (G)*) of the bipartite graph $G$.

$$D(p, p') = \frac{1}{flow(G) + 1} \qquad (17)$$

## 3.6. Style-Content Separation by Anisotropic Part Scales

[45] introduces a new signature for 3D objects, the Anisotropic Part Scales (APS) signature. First, all


Fig. 9: Contextual part analogies in 3D objects: (a) The SDF values on an octopus, (b) The segmented object, (c) Analogous parts in other objects (taken from [37]).

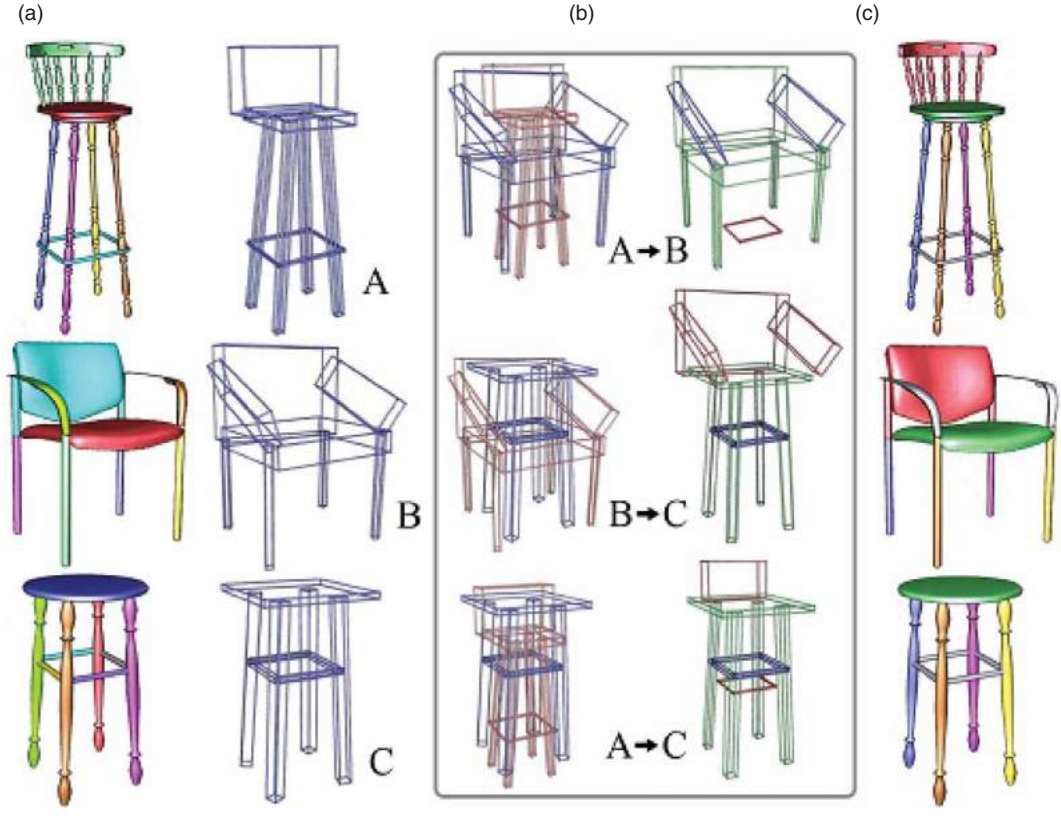


Fig. 10: Part correspondence using the deformed OBB approach: (a) three chairs and the corresponding OBBs, (b) matching of the OBBs, (c) final part correspondence (taken from [45]).

objects of the database are segmented using the method of [19]. For each part generated by the segmentation method the oriented bounding box (OBB) is computed. Then, the APS signature can be defined for each part composition of a given shape. Defining the actual signature requires the computation of three Laplacian spectra along three directions of the OBB graph. The complete APS signature contains $6n$ eigenvalues, where $n$ is the number of parts. The APS signature is calculated with respect to the OBB representation of the object (Fig. 10). Three Laplacian spectra contribute for scaling along the three dimensions, while three more contribute for transformation, each one for linearity, planarity and sphericity. The matching is performed using a deform-to-fit approach. Segments are deformed to find the best match. The OBB graph of the objects is used. Each OBB is matched to a type: linear, planar, or spherical. The node with the most edges of the first object is selected to be compared to the segments of the second object of the same type. Unary attributes of each segment are used for the comparison, as well as pairwise ones. Two adjacent nodes have a valid relation if their types can be rotated or transformed to fit. An APS style distance is defined between two shapes:

$$D(p,q) = \min_{n \in I_{APS}, |c_p|=|c_q|=n} \frac{\left\| APS(c_p) - APS(c_q) \right\|}{\sqrt[a]{n}} \quad (18)$$

where $c_p$ and $c_q$ are part compositions of shapes $p$ and $q$. $|c|$ denotes the size of a composition while $APS(c)$ denotes the APS signature of $c$. The distance

$\|\cdot\|$ between two signatures is their Euclidean distance. The clustering of the styles is performed with standard spectral k-means clustering. Within the style a co-segmentation is performed and the final content classification is achieved by an exhaustive tree search, using the Euclidean distance of the light field descriptor [12] of the corresponding parts.

## 4. DISCUSSION

A structural snapshot of 3D object retrieval methodologies using a part-based representation is shown in Tab. 1. Typically, the major challenges in the operational pipeline are related to aspects as in the following:

- 3D mesh segmentation
- Node features
- Edge features
- Distance between parts
- Graph matching

The first step of the retrieval procedure partitions each of the 3D objects in the database as well as the query object. The techniques involved can vary significantly and play a vital role in the outcome. 3D mesh segmentation is an open problem, and the segmentation technique used is crucial. An implicit way to address this problem is the use of Reeb graphs. This is advocated by the fact that they combine geometric and structural information. The function mostly used for the construction of the Reeb graph is the integral of geodesic distances (Section 3.2), which is a pose invariant function and can partition protrusions of an object. The method in [33] uses the conformal factor for segmentation. Other state-of-art 3D mesh segmentation methods, involve hierarchical clustering methods [15], morphology-based [23], patch-based segmentation [26],[16], skeletons [39],[13] etc.

The following abbreviations are used in Tab. 1: **IG**: integral geodesic function, **SDF**: shape diameter function, **A**: area, **V**: volume, **C**: curvature, **Conv**: convexity, **CF**: discrete conformal factor, **EE**: eccentricity values of ellipsoid representation, **SH**: spherical harmonic descriptor, **Dbary**: distance from mesh barycenter, **DS**: distance between segment centers, **AS**: angles between segments' principal axes, **GD**: geodesic distance between boundaries.

The result of the mesh segmentation step is a graph representation of the object, where distinct segments are represented by nodes. Each part of the object is assigned a set of geometrical attributes, referred as the feature vector or descriptor. A descriptor is expected to capture and quantify the part's essence. Typical functions used as part descriptors are the total surface area, convexity, curvature, volume, and geodesic distance. In [20], [7] and [43] the geodesic integral is used not only in the segmentation step, but also averaged over the surface of the segment and used as a feature. The average discrete conformal factor defined in [6] is used in [33] and [37]. In [37] the shape diameter function is used, a measure that calculates the thickness of the segment. Apart from the features that describe each node of the graph, namely the unary features, some methods include binary features, i.e. geometric features that correspond to the boundary of two segments or features that involve two adjacent regions and express the relation between them. Example binary features are the Euclidean distance between adjacent segment centers and angles between the segments' principal axes are used in [1] and [23]. Binary features have only been used in three methodologies out of the total number of methodologies reviewed in this paper. Unary features are without doubt more important than binary, since they describe the geometry of the part.

A critical aspect in the retrieval process is the similarity estimation of two parts. Given two descriptor vectors, a distance needs to be defined, so that the similarity of two parts can be evaluated. Any Minkowski distance can be effective, with $L_1$ and $L_2$ distances being most common. The distance measure used to compare two part descriptors seems independent to the rest of the aspects of the method.

In addition to the geometrical comparison between two objects the problem of the topological correspondence of the parts also needs to be addressed. The choice of graph matching algorithm depends on the complexity that one is willing to tolerate. A costly solution would be to exhaustively search for the best correspondence between nodes of the graph. The methods in [1] and [23] use the Earth Mover's Distance (EMD), which was introduced in [32] and reflects the minimal amount of work that must be performed to transform one distribution into the other, with the two distributions corresponding to the two ARGs to be compared. The method in [45] uses exhaustive tree-based search with pruning [46], and selects the node with the most edges as root in order to reduce complexity. The methods in [20], [7] and [43] utilize the multiresolution, hierarchical representation of the graphs, by matching nodes first in the coarse representation and then expanding to finer ones. The partial matching methods in [42] and [37] use maximum bipartite matching algorithms to find similar parts. [33] automatically matches the core nodes of the two graphs, and then uses the Hungarian algorithm to match the branches.

The analysis of the methods comes to conclude that the most critical elements in part-based retrieval are the 3D mesh segmentation and the selection of the most appropriate descriptor. The graph matching problem is mainly a complexity problem, while most distance measures yield satisfactory results. The segmentation approach is closer to human perception than the global approach since it can capture the essence of an object. The key to good retrieval results is the combination of the proper segmentation

| Method | Year | Dataset | Segmentation | Unary Features | Binary Features | Global/Partial Retrieval | Distance Measure | Graph Matching |
|---|---|---|---|---|---|---|---|---|
| Hilaga et al [20] | 2001 | Stanford dataset, 3dcafe | Reeb_graph {IG} | A,IG | – | global | weighted min value | Multiresolution approach |
| Bespalov et al [7] | 2003 | Primitive shapes, CAD, lego | Reeb_graph {IG} | A,IG | – | global | weighted min value | Multiresolution approach |
| Tung and Schmitt [43] | 2005 | Generic Models downloaded from the internet | Reeb_graph {IG/height/ Dbary} | A,IG, Dbary, C,V | – | global, partial | any Minkowski distance | Multiresolution approach |
| Kim et al [23] | 2005 | The MPEG-7 set | Morphology-based | V, C, EE | DS, AS | global | Euclidean Distance | EMD |
| Biasotti et al [10] | 2006 | Drexel University, AIM@SHAPE, Princeton [38], Image-based 3D models Archive, McGill [47] | Reeb_graph {IG/Dbary} | SH, geodesic, node degree | – | global, partial | weighted graph-based measure | Tree search |
| Gal and Cohen-Or [16] | 2006 | Princeton [38] | Quadric fitting | C | – | partial | – | Geometric hashing |
| Siddiqi et al [39] | 2008 | McGill [47] | Medial surfaces | C | – | global | Euclidean Distance | Bipartite graph matching |
| Agathos et al [1] | 2009 | McGill [47] | Protrusion oriented | V, Conv, EE, SH | DS, AS | global | Euclidean Distance | EMD |
| Tierny et al [42] | 2009 | SHREC 2007 [18] | Reeb_graph {geodesic to seeds, C} | area of the Reeb chart | – | partial | $L_1$ | Bipartite graph matching |
| Demirci [13] | 2010 | Silhouettes of 9 objects | Shock graphs | Circle radius | Euclidean distance | global | Euclidean distance | EMD |

(*Continued*)

Tab. 1: Methods overview.



*Computer-Aided Design & Applications*, 11(6), 2014, 670-684, http://dx.doi.org/10.1080/16864360.2014.914385

| Method | Year | Dataset | Segmentation | Unary Features | Binary Features | Global/Partial Retrieval | Distance Measure | Graph Matching |
|---|---|---|---|---|---|---|---|---|
| Ferreira et al [15] | 2010 | Engineering Shape Benchmark [21] | Hierarchical clustering | SH | – | partial | Euclidean Distance | – |
| Shapira et al [37] | 2010 | SHREC 2007 [18], Princeton [38] | SDF | SDF, CF, Shape Distribution (D1,D2,D3,A3) | – | partial | $L_1$ | Bipartite graph matching |
| Xu et al [45] | 2010 | Princeton [38], iWires [17] | Hierarchical clustering | APS signature, Light field descriptor | APS signature | global | Euclidean Distance | Tree search |
| Sfikas et al [33] | 2012 | TOSCA, SHREC 2007 [18], SHREC 2010 [28], SHREC 2011 [27] | CF quantization | CF, A,GD | – | global | $L_1$ | Hungarian algorithm |
| Lavoué [26] | 2012 | McGill [47], SHREC 2007 [18] | k-means clustering | Spectral coefficients of the Laplacian | – | global, partial | $L_1$ | – |

Tab. 1: Continued.



*Computer-Aided Design & Applications*, 11(6), 2014, 670-684, http://dx.doi.org/10.1080/16864360.2014.914385

method and a meaningful part signature. A milestone for the evolution of retrieval methods is the establishment of the SHape REtrieval Contest (SHREC) [35], and most methods since then use the SHREC Dataset for evaluation purposes.

Since part retrieval implies a part-based representation, it is evident that part retrieval methodologies will be more effective in datasets that comprise objects with distinct parts, eg. McGill dataset for articulated objects. This is the reason that there exists an emphasis at the experimental level on this type of datasets in contrast to generic datasets e.g. Princeton, for which 3D object segmentation is not trivial and part distinction cannot easily be achieved. For the sake of clarity, the datasets shown in Tab. 1, indicate that datasets including objects with distinct parts are by far favored by the presented methods.

In the future the part-based representation is highly likely to play an essential role in content-based 3D object retrieval. Part descriptors are becoming more efficient and the field of 3D mesh segmentation is currently evolving. As better segmentation methods emerge, they are likely to be employed for retrieval purposes more frequently. New tendencies for segmentation include the application of multiple partitioning methods on the same object although this would increase the computational cost. Furthermore, hierarchical approaches seem to be appealing for retrieval purposes since they take into account that an object has different levels of detail and parts are retrieved based on their shape as well as the hierarchical level that produce them.

## ACKNOWLEDGEMENTS

This research has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: **THALES**-3DOR(MIS 379516). Investing in knowledge society through the European Social Fund.

## REFERENCES

[1] Agathos, A.; Pratikakis, I.; Papadakis, P.; Perantonis, S.; Azariadis, P.; Sapidis, S.: 3D articulated object retrieval using a graph-based representation, The Visual Computer, 26(10), 2010, 1301–1319. http://dx.doi.org/10.1007/s00371-010-0523-1

[2] Agathos, A.; Pratikakis, I.; Perantonis, S.; Sapidis, N.; Azariadis, P.: 3D mesh segmentation methodologies for CAD applications, Computer-Aided Design and Applications, 4(6), 2007, 827–841. http://dx.doi.org/10.3722/cadaps.2007.827-841

[3] Agathos, A.; Pratikakis, I.; Perantonis, S.; Sapidis, N.: Protrusion-oriented 3D mesh segmentation, The Visual Computer, 26(1), 2009, 63–81. http://dx.doi.org/10.1007/s00371-009-0383-8

[4] Antini, G.; Berretti, S.; Del Bimbo, A.; Pala, P.: 3D mesh partitioning for retrieval by parts applications, 2005 IEEE International Conference on Multimedia and Expo, 2005, 1210–1213. http://dx.doi.org/10.1109/ICME.2005.1521645

[5] Attene, M.; Falcidieno, B.; Spagnuolo, M.: Hierarchical mesh segmentation based on fitting primitives, The Visual Computer 22(3), 2006, 181–193. http://dx.doi.org/10.1007/s00371-006-0375-x

[6] Ben Chen, M.; Gotsman, C.: Characterizing shape using conformal factors, Eurographics 2008 Workshop on 3D Object Retrieval, 2008, 1–8. http://dx.doi.org/10.2312/3DOR/3DOR08/001-008

[7] Bespalov, D.; Regil, W.; Shokoufandeh, A.: Reeb graph based shape retrieval for CAD, Proceedings of the ASME DETC, 2003, 229–238. http://dx.doi.org/10.1115/DETC2003/CIE-48194

[8] Biasotti, S.; Giorgi, D.; Spagnuolo, M.; Falcidieno, B.: Reeb graphs for shape analysis and applications, Theoretical Computer Science, 392(1), 2008, 5–22. http://dx.doi.org/10.1016/j.tcs.2007.10.018

[9] Biasotti, S.; Marini, S.; Mortara, M.; Patanè, G.; Spagnuolo, M.; Falcidieno, B.: 3D shape matching through topological structures, Lecture Notes in Computer Science 2886, 2003, 194–203. http://dx.doi.org/10.1007/978-3-540-39966-7_18

[10] Biasotti, S.; Marini, S.; Spagnuolo, M.; Falcidieno, B.: Sub-part correspondence by structural descriptors of 3D shapes, Computer-Aided Design, 38(9), 2006, 1002–1019. http://dx.doi.org/10.1016/j.cad.2006.07.003

[11] Biederman, I.: Recognition-by-components: A theory of human image understanding, Psychological Review, 94(2), 1987, 115–147. http://dx.doi.org/10.1037/0033-295X.94.2.115

[12] Chen, D.-Y.; Tian, X.-P.; Shen, Y.-T.; Ouhyoung, M.: On visual similarity based 3D model retrieval, Computer Graphics Forum, 22(3), 2003, 223–232. http://dx.doi.org/10.1111/1467-8659.00669

[13] Demirci, M.F: Efficient shape retrieval under partial matching, $20^{th}$ International Conference on Pattern Recognition, 2010, 3057–3060. http://dx.doi.org/10.1109/ICPR.2010.749

[14] Demirci, M.F.; Shokoufandeh, A.; Keselman, Y.; Bretzner, L.; Dickinson, S.: Object recognition as many-to-many feature matching, International Journal of Computer Vision, 69(2), 2006, 203–222. http://dx.doi.org/10.1007/s11263-006-6993-y

[15] Ferreira A.; Marini, S.; Attene, M.; Fonseca; M.; Spagnuolo, M.; Jorge, J; Falcidieno, B.: Thesaurus-based 3D object retrieval with

part-in-whole matching, International Journal of Computer Vision, 89(2), 2010, 327–347. http://dx.doi.org/10.1007/s11263-009-0257-6

[16] Gal, R.; Cohen-Or, D.: Salient Geometric Features for Partial Shape Matching and Similarity, ACM Transactions on Graphics, 25(1), 2006, 130–150. http://dx.doi.org/10.1145/1122501.1122507

[17] Gal, R.; Sorkine, O.; Mitra, N. J.; Cohen-Or, D.: iWIRES: an analyze-and-edit approach to shape manipulation, ACM Transactions on Graphics, 28(3), 2009, 1–10. http://dx.doi.org/10.1145/1531326.1531339

[18] Giorgi, D.; Biasotti, S.; Baraboshchi, L.: Shape retrieval contest 2007: watertight models track, Techical Report UU-CS-2007-015, 2007.

[19] Golovinskiy, A.; Funkhouser, T.: Consistent segmentation of 3D models, Computers and Graphics, 33(3), 2009, 262–269. http://dx.doi.org/10.1016/j.cag.2009.03.010

[20] Hilaga, M.; Shinagawa, Y.; Kohmura, T.; Kunii, T.L.: Topology matching for fully automatic similarity estimation of 3D shapes, in Proceedings of the 28th annual conference on computer graphics and interactive techniques, 2001, 203–212. http://dx.doi.org/10.1145/383259.383282

[21] Jayanti, S.; Kalyanaraman, Y.; Iyer, N.; Ramani, K.: Developing an engineering shape benchmark for cad models, Computer Aided Design, 39(9), 2006, 939–953. http://dx.doi.org/10.1016/j.cad.2006.06.007

[22] Kazhdan, M.; Funkhouser, T.; Rusinkiewicz, S.: Rotation invariant spherical representation of 3D shape descriptors, In Symposium on Geometry Processing, 2003, 156–164.

[23] Kim, D.H.; Park, I.; Yun, I.; Lee, S.: A new MPEG-7 standard: perceptual 3-D shape descriptor, Lecture Notes in Computer Science 3332, 2005, 238–245. http://dx.doi.org/10.1007/978-3-540-30542-2_30

[24] Kim, D.H.; Yun, I.D.; Lee, S.U.: A new shape decomposition scheme for graph-based representation, Pattern Recognition, 38(5), 2005, 673–689. http://dx.doi.org/10.1016/j.patcog.2004.10.003

[25] Kuhn, H.W.: The Hungarian method for the assignment problem, Naval Research Logistics Quarterly 2(1-2), 2006, 83–97. http://dx.doi.org/10.1002/nav.3800020109

[26] Lavoué, G.: Combination of bag-of-words descriptors for robust partial shape retrieval, The Visual Computer, 28(9), 2012, 931–942. http://dx.doi.org/10.1007/s00371-012-0724-x

[27] Lian, Z.; Godil, A.; Bustos, B.; Daoudi, M.; Hermans, J.; Kawamura, S.; Kurita, Y.; Lavoué, G.; Van Nguyen, H.; Ohbuchi, R.; Ohkita, Y.; Ohishi, Y;, Porikli, F.; Reuter, M.; Sipiran, I.; Smeets, D.; Suetens, P.; Tabia, H.; Vandermeulen, D.: A comparison of methods for non-rigid 3D shape retrieval, Pattern Recognition, 46(1), 2013, 449–461. http://dx.doi.org/10.1016/j.patcog.2012.07.014

[28] Lian, Z.; Godil, A., Fabry, T.; Furuya, T., Hermans, J.; Ohbuchi, R.; Shu, C.; Smeets, D.; Suetens, P., Vandermuelen, D., Wuhrer, S.: SHREC'10 track: non-rigid 3D shape retrieval, Eurographics Workshop on 3D Object Retrieval, 2010, 1–8.

[29] Lloyd, S.: Least squares quantization in PCM, IEEE Transactions on Information Theory, 28(2), 1982, 129–137. http://dx.doi.org/10.1109/TIT.1982.1056489

[30] Osada, R.; Funkhouser, T; Chazelle, B.; Dobkin, D.: Shape distributions, ACM Transactions on Graphics, 21(4), 2002, 807–832. http://dx.doi.org/10.1145/571647.571648

[31] Papadakis, P.; Pratikakis, I.; Perantonis, S.; Theoharis, T.; Efficient 3D shape matching and retrieval using a concrete radicalized spherical projection representation, Pattern Recognition, 40(9), 2007, 2437–2452. http://dx.doi.org/10.1016/j.patcog.2006.12.026

[32] Rubner, Y.; Tomasi, C.; Guibas, L.J.: The Earth Mover's Distance as a Metric for Image Retrieval, International Journal of Computer Vision, 40, 2000, 99–121. http://dx.doi.org/10.1023/A:1026543900054

[33] Sfikas, K.; Theoharis, T.; Pratikakis, I.: Non-rigid 3D object retrieval using topological information guided by conformal factors, The Visual Computer, 28(9), 2012, 943–955. http://dx.doi.org/10.1007/s00371-012-0714-z

[34] Shamir, A.: A survey on Mesh Segmentation Techniques, Computer Graphics Forum, 27(6), 2008, 1539–1556. http://dx.doi.org/10.1111/j.1467-8659.2007.01103.x

[35] Shape Retrieval Contest (SHREC), http://www.aimatshape.net/event/SHREC.

[36] Shapira, L.; Shalom, S.; Shamir, A.; Cohen-Or, D.: Consistent mesh partitioning and skeletonisation using the shape diameter function, The Visual Computer 24, 2008, 249–259. http://dx.doi.org/10.1007/s00371-007-0197-5

[37] Shapira, L.; Shalom, S.; Shamir, A.; Cohen-Or, D.; Zhang, H.: Contextual part analogies in 3D objects, International Journal of Computer Vision, 89(2), 2010, 309–326. http://dx.doi.org/10.1007/s11263-009-0279-0

[38] Shilane, P.; Min, P.; Kazhdan, M.; Funkhouser, T.: The princeton shape benchmark, International Conference on Shape Modeling and Applications, 2004, 167–178. http://dx.doi.org/10.1109/SMI.2004.1314504

[39] Siddiqi, K.; Zhang, J.; Macrini, D.; Shokoufandeh, A.; Bouix, S.; Dickinson, S.: Retrieving

articulated 3-D models using medial surfaces, Machine Vision and Applications, 19(4), 2008, 261–275. http://dx.doi.org/10.1007/s00138-007-0097-8

[40] Tangelder, J.; Veltkamp, R.: A survey of content based 3D shape retrieval methods, Multimedia Tools and Applications, 39(3), 2008, 441–471. http://dx.doi.org/10.1007/s11042-007-0181-0

[41] Tierny, J.; Vandeborre, J.P.; Daoudi, M.: Enhancing 3D mesh topological skeletons with discrete contour constrictions, The Visual Computer, 24(3), 2008, 155–172. http://dx.doi.org/10.1007/s00371-007-0181-0

[42] Tierny, J.; Vandeborre, J.; Daoudi, M.: Partial 3D shape retrieval by reeb pattern unfolding, Computer Graphics Forum, 28(1), 2009, 41–55. http://dx.doi.org/10.1111/j.1467-8659.2008.01190.x

[43] Tung, T.; Schmitt, F.: The augmented multiresolution reeb graph approach for content-based retrieval of 3D shapes, International Journal of Shape Modeling, 11(1), 2004, 91–120. http://dx.doi.org/10.1142/S0218654305000748

[44] Van Kaick, O.; Zhang, H.; Hamarneh, G.; Cohen-Or, D.: A survey on shape correspondence, Computer Graphics Forum, 30(6), 2011, 1681–1707. http://dx.doi.org/10.1111/j.1467-8659.2011.01884.x

[45] Xu, K.; Li, H.; Zhang, H.; Cohen-Or, D.; Xiong, Y.; Cheng, Z.: Style-content separation by anisotropic part scales, ACM Transactions on Graphics, 29(6), 2010, 1–9. http://dx.doi.org/10.1145/1882261.1866206

[46] Zhang, H.; Sheffer, A.; Cohen-Or, D; Zhou, Q.; Van Kaick, O.; Tagliasacchi, A.: Deformation-driven shape correspondence, Computer Graphics Forum, 27(5), 2008, 1431–1439. http://dx.doi.org/10.1111/j.1467-8659.2008.01283.x

[47] Zhang, J.; Kaplow, R.; Chen, R.; Siddiqi, K.: The McGill Shape Benchmark, 2005. www.cim.mcgill.ca/~shape/benchMark/