

Retrieval of 3D Articulated Objects using a graph-based representation

A. Agathos^{1,2}, I. Pratikakis¹, P. Papadakis¹, S. Perantonis¹, P. Azariadis² and N. Sapidis²

¹Computational Intelligence Laboratory, Institute of Informatics and Telecommunications,
National Center for Scientific Research “Demokritos”, Agia Paraskevi, Athens 15310, Greece

²Department of Product and Systems Design Engineering,
University of the Aegean, Ermoupolis, Syros GR-84100, Greece

Abstract

Most of the approaches which address the problem of 3D object retrieval, use global descriptors of the objects which fail to consistently compensate for the intra-class variability of articulated objects. In this paper, a retrieval methodology is presented which is based upon a graph-based object representation. This is composed of a meaningful new mesh segmentation along with a graph matching between the graph of the query object and each of the graphs that correspond to the objects of the 3D object database. The graph matching algorithm is based on the Earth Mover’s Distance (EMD) similarity measure which is calculated using a new ground distance assignment. The superior performance of the proposed methodology is shown after an extensive experimentation comprising alternative descriptors for the constituent components of the 3D object as well as comparison with state of the art retrieval algorithms.

Categories and Subject Descriptors (according to ACM CCS): Pattern Recognition [I.5.4]: Computer Vision—, Information Storage and Retrieval [H.3.3]: Information Search and Retrieval—

1. Introduction

3D object retrieval is the process which retrieves from a database of 3D objects those that match best a 3D object query using a measure of similarity. Most of the approaches which address this problem use descriptors which express the object’s global shape [BKS*04, FS06, GSCO07, JZ07, KFR03, OOFB08, PPPT07, PPT*08, PTK07, Vra04, ZDA*07]. However, most of these approaches fail to consistently compensate for the intra-class variability of articulated objects. This occurs because it is not evident how a global descriptor will become invariant to non-rigid transformations like bending or stretching, thus, resulting in an erroneous matching. In this paper, a retrieval methodology is presented which is based upon a graph-based representation that is built after a 3D mesh segmentation. The motivation of this approach originates from object recognition where the object is described in terms of its components that are attributed with geometric characteristics and relational connections with each other. This description is referred to as the *structural*

description of the object [Bie87]. In order to recognize an object, its structural description is compared with the structural descriptions of already classified objects and the object is classified to the class of the best match.

This recognition process can be naturally adopted for 3D object retrieval. Meaningful components of the object can be extracted using a segmentation algorithm. The structural description of the object is created by using the *Attributed Relational Graph* (ARG) concept, i.e. the components of the object are represented as the nodes of a graph and the relationship of the components with each other are represented as the edges of the graph. To each node *unary* attributes are assigned which describe the geometric characteristics of the component and to each edge *binary* attributes are assigned which describe the relationship of the connected nodes.

Eventually, the problem of matching a query object with the objects stored in the database is transformed into the problem of matching their ARGs [KPYL04, MDA*08, TZ06].

The proposed graph matching algorithm is based on the Earth Mover’s Distance (EMD) similarity measure.

In the proposed methodology the contribution is twofold:

- i. A new mesh segmentation algorithm is presented that partitions the 3D object into its meaningful constituent components;
- ii. A new ground distance assignment is introduced for the calculation of the EMD;

This paper is organized as follows. Section 2 discusses the related work. Section 3 is dedicated to the detailed description of the proposed methodology. In Section 4, the experimental evaluation is discussed while in Section 5 conclusions are drawn.

2. Related Work

Among the existing 3D object retrieval methods, we can distinguish two main categories:

- i. Methods with global shape representations.
- ii. Methods with graph-based shape representations.

Methods that belong to the first category may be further classified according to the spatial dimensionality of the information used for matching the objects in the retrieval process, namely 2D, 3D and their combination (2D/3D).

Those methods that use 2D information utilize descriptors which are generated from image-projections that may be contours, silhouettes, depth buffers or other kinds of 2D information. Thus, similarity is measured using 2D matching techniques [OOFB08, PTK07, Vra04, ZDA*07].

Those methods that use 3D information extract their descriptors from the 3D shape geometry of the 3D object and the shape similarity is measured using appropriate representations either in the spatial domain or in the spectral domain [GSCO07, JZ07, PPPT07, PPT*08, Vra04, KFR03].

Methods that combine both 2D and 3D-based features have also been developed, in order to improve the overall retrieval performance [BKS*04, FS06, Vra04].

In the second category, the structural description of the object is extracted which describes the relations of the object's components using a graph structure. In [HSKK01], a multiresolution Reeb graph is used in order describe the structure of the object. The query's reeb graph is matched hierarchically with each of the objects reeb graphs stored in a database. In [KPYL04], the object is first voxelized and then segmented using a morphological structure. The extracted components create an Attributed Relational Graph. The query's ARG is matched against the ARGs stored in the database using an EMD-based approach. In [TZ06] the mesh is decomposed into its meaningful components and the ARG of the object is constructed based on their decomposition. Retrieval is achieved by matching the query's ARG with the ARGs of the objects stored in the database using an error correcting graph isomorphism algorithm. In [MDA*08] the medial surface of the 3D object

is constructed which is used to segment it into meaningful parts. The parts are approximated by superellipsoids which are used to construct distance field descriptors. These descriptors along with the superellipsoid approximations are used to construct an attributed graph of the object. Retrieval is achieved by graph-matching. In [SSDZ99, SSGD03] the skeletons of the objects are extracted using a volumetric approach that set the base for the generation of shock graphs which are used for the matching process. In [MSF07], the structural description of the object in the form of a graph is also used. Their methodology comprises two steps: first, they compute a common subgraph for each class of the database and then they define a set of editing operations based on the subgraph. These two steps allow them to construct a prototype for each class to which the query object is matched.

Despite the plethora of 3D object retrieval methods belonging to the first category, there are only a few that deal with the problem of retrieving articulated objects [JZ07, OOFB08, GSCO07, BCG08]. On the contrary, 3D object retrieval methods that belong to the second category exhibit higher efficiency in retrieving articulated objects. In some cases though complicated graph structures are used that decrease the time efficiency of retrieval or are susceptible to the presence of geometrical or topological noise.

The proposed retrieval methodology belongs to the second category of 3D object retrieval algorithms. It is composed of a meaningful new mesh segmentation algorithm along with a graph matching between the graphs of the query object and each of the graphs that correspond to the objects of the 3D object database.

3. Proposed Methodology

Our objective is to create the ARG of a query 3D object and compare it against the ARGs of the 3D objects stored in a database. To this end, a methodology is proposed which consists of the following consecutive steps:

- i. The meaningful components of the object are extracted using the proposed segmentation algorithm;
- ii. An ARG will be constructed based on the segmentation result;
- iii. An EMD-based matching algorithm which compares two ARGs is performed.

The detailed description of the aforementioned steps is given in the sequel.

3.1. 3D Mesh Segmentation Algorithm

The proposed segmentation algorithm partitions closed manifolds and is based on the premise that a 3D object consists of a core body and its constituent protrusible components.

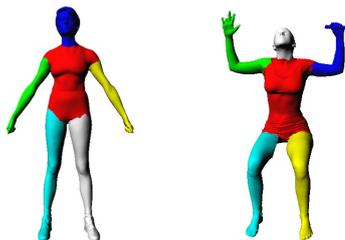


Figure 1: Segmentation of human figures into their main body and their articulated components

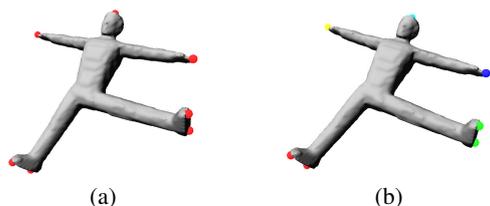


Figure 2: Example of the ‘human’ 3D mesh with its corresponding salient points at the (a) extraction stage (red dots) and (b) grouping stage - each color represents a different group

For example, in Figure 1 two human objects are shown in different poses, that have been segmented into their main body along with their protrusible (articulated) components, i.e. head, hands, legs.

The novelty of the segmentation scheme is twofold:

- i. A novel methodology to trace the partitioning boundaries of the 3D object using closed boundaries is presented;
- ii. A novel algorithm for the core approximation of the 3D object is introduced.

The segmentation algorithm consists of the following steps:

- i. Extraction of salient mesh points which characterize its protrusions. These points are further grouped according to their geodesic proximity where each group represents a main component of the object;
- ii. Approximation of the core (main body) using the grouped salient points from step (i);
- iii. Detection of the partitioning boundaries using closed boundaries constructed by a distance function along the protrusions;
- iv. Refinement of the detected partitioning boundaries.

3.1.1. Salient Points Extraction and Grouping

In order to realize the salient point extraction process, a function first introduced in [HSKK01] is used, which has the property of having low values at the center of the 3D object and high values at its protrusions. This function is called *protrusion function*, $pf()$ [APP*07].

Formally, for each point v of the surface S of a 3D object, the protrusion function is defined as:

$$pf(v) = \int_{p \in S} g(v, p) dS \quad (1)$$

where $g(v, p)$ denotes the geodesic distance between v, p . Since the protrusion function has the aforementioned property and salient points of the mesh reside at the extrema of the mesh it is natural to search for them at the local maxima of the protrusion function, i.e. for each vertex $v \in S$ a neighborhood of vertices N_v is considered, v is a salient point of the mesh if: $pf(v) > pf(v_i), \forall v_i \in N_v$.

In our implementation, N_v is considered to be a geodesic neighborhood adopting the same radius $\sqrt{5 \cdot 10^{-3} \cdot \text{area}(S)}$ as proposed in [LLL07]. The salient points are further filtered according to their protrusion function to ensure that they would be far from the center of the mesh.

The salient points of a ‘human’ object are shown in Figure 2(a).

The segmentation algorithm aims to extract the meaningful components of the object. It often happens that the extracted salient points belong to sub-components of the objects. For example, in Figure 2(a) there exist salient points on the legs of the ‘human’ object that represent its sub-components (toes and the rest of the legs). In order to achieve a single component for the legs of the object those salient points should be grouped.

The salient points that are required to be grouped are those which are close to each other in terms of geodesic distance based on a fixed threshold which is half of the mean of the geodesic distances between the salient points. Once the salient points are grouped, the salient point with the largest protrusion value is chosen as the representative for each group, thus called the *representative salient point*.

In Figure 2(b), the result of the grouping of the salient points in the ‘human’ object is shown. As it can be observed, each group represents a unique protrusible component of the object.

3.1.2. Core Approximation

An algorithm which approximates the main body of the object is the one that can acquire all the elements (vertices or faces) of the mesh except those that belong to the protrusion of the mesh. These elements should separate all the protrusions from each other.

In our case, the core approximation is addressed by using the minimum cost paths between the representative salient points. Let P be the set of all minimum cost paths between the representative salient points. The core approximation algorithm expands a set of vertices in ascending order of protrusion function value until the expanded set contains a fixed percentage of all elements of P (15%). Stopping the expansion of our core approximation when this percentage is reached separates all the protrusions from each other.

Figure 3 illustrates the core approximation of the ‘human’



Figure 3: Example of core approximation for the ‘human’ 3D object. The vertices representing the core are coloured in yellow

object. As it can be observed, this methodology produces consistent approximation of the core and its boundaries identify the initial approximation of the partitioning boundaries.

3.1.3. Partitioning Boundary Detection

In this section, we detail the stage of the segmentation algorithm which detects the partitioning boundary, that is the boundary between a protrusion and the main body of the mesh. It is considered that in the area that is enclosed by the desired boundary between the protrusion and the main body, an abrupt change should occur in the volume of the 3D object. Thus, our goal is to detect this change because it signifies the existence of a partitioning boundary. In our case, the abrupt change is measured in terms of closed boundaries perimeter which are placed in an area of the mesh which contains the partitioning boundary. The closed boundaries are defined by a distance function D associated to a salient representative of the group which represents the protrusion. For each representative salient point \hat{s} , the distance function D is defined for each vertex v of the mesh as the shortest distance between v and \hat{s} . The shortest distance is computed using the Dijkstra algorithm with source \hat{s} and cost for each edge (u, v) of the mesh denoted as :

$$\text{cost}(u, v) = \delta \frac{\text{length}(u, v)}{\text{avg_length}} + (1 - \delta) \frac{\text{prot}(u, v)}{\text{avg_prot}} \quad (2)$$

where $\text{prot}(u, v) = |pf(u) - pf(v)|$ and avg_length , avg_prot denote the average values of the length and protrusion difference of the edges of the mesh, respectively. This distance function was introduced in [LLL07]. In our implementation, we set δ equal to 0.4.

The closed boundaries are finally constructed by interpolating on the edges of the mesh the iso-contour generated by setting a constant value D_c on D .

Taking advantage of the fact that the proposed core approximation has its boundaries near the distinct components of the object, the aforementioned area which contains the partitioning boundary is defined as the part of the mesh whose values of D lie in the interval $[(1 - d_1)D_{\text{coremin}}, (1 + d_2)D_{\text{coremin}}]$. D_{coremin} denotes the value of the distance function between the nearest point of the core approximation and the representative \hat{s} , while d_1 , d_2 denote the extent of

the interval ($0 < d_1 < 1$, $d_2 > 0$). In this paper, we set $d_1 = 0.1$, $d_2 = 0.4$. This area is swept with the closed boundaries in fixed steps and the sweeping is terminated when the change of the perimeter between successive closed boundaries are greater than a threshold $r_{\text{max}} = 1.3$. In this way, the aforementioned abrupt change in the volume of the object is detected at the boundary between its main body and the protrusion. The closed boundary in which this abrupt change occurs is the proposed approximation of the partitioning boundary. Figure 4(a) shows the closed boundaries constructed using the distance function, while the approximation of the partitioning boundary is shown in blue.

The choice of the salient representative as the source of the distance function D may lead to the construction of skewed closed boundaries. This choice is refined by properly selecting as source the point that has the minimum protrusion value on an area enclosing the salient points of the group. This source point leads to the creation of closed boundaries that are positioned near to the true partitioning boundary.

3.1.4. Partitioning boundary refinement

The partitioning boundary detected in section 3.1.3 is an iso-contour of the distance function D approximating the true partitioning boundary. In most of the cases this approximation is rough and has to be refined and placed according to Hoffman and Richards [HR84] at the concavities of the object.

In order to achieve this, the minimum-cut methodology of Katz and Tal [KT03] is used. Specifically, a flow network graph is constructed using the dual graph of the mesh [APP*07]. In order to construct the network of [KT03], three regions are defined, as illustrated in the ‘human’ object of Figure 4(b). Specifically, a region **A** is defined containing the triangles of the protrusion of the mesh (yellow triangles), a region **C** is defined containing the partitioning boundary (red triangles) and a region **B** is defined containing the faces of the rest of the mesh (green triangles).

Region **C** is constructed as follows. First the average geodesic distance is found, denoted as AvgGeodDist , between the partitioning boundary extracted in section 3.1.3 and the refined representative calculated as it is shown in Section 3.1.3. Then, region **C** is defined as the triangles of the mesh whose vertices geodesic distance from the refined representative lies in the interval $[0.9 \cdot \text{AvgGeodDist}, 1.1 \cdot \text{AvgGeodDist}]$.

The set **A** is constructed by performing a breadth first search starting from the refined representative of the protrusion until region **C** is reached.

Using these three regions the flow network of [KT03] is constructed and the application of the minimum cut algorithm on this network creates a partitioning boundary that

passes through the concave region of the mesh (Figure 4(c)).

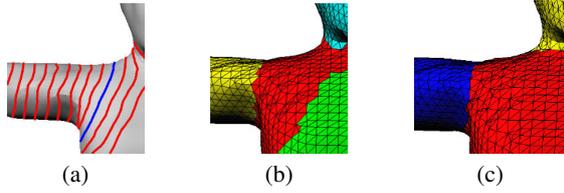


Figure 4: (a) closed boundaries constructed using the distance function, the approximation of the partitioning boundary is shown in blue; (b) region **A** is shown with yellow, region **B** is shown with green and region **C** is shown with red; (c) The final segmentation of the protrusion from its main body after the application of the minimum cut algorithm

3.2. EMD-based Matching

As mentioned in Section 3.1 the proposed segmentation algorithm is capable of segmenting an articulated object into a core element and its articulated components. The segmented components are the nodes of the ARG and each of the nodes representing an articulated component is connected with the node representing the core of the object, forming the edges of the ARG. Unary and binary attributes will be assigned to the nodes and edges respectively.

The proposed algorithm that is used to match two ARGs is based on the EMD similarity measure. In general the EMD measure is used to efficiently express the similarity of two signatures belonging to two different distributions in a feature space [RTG00].

The signatures consist of two sets of nodes $\{v_i\}_{i=1}^n, \{u_j\}_{j=1}^m$ of size n and m , respectively and to each of the set of nodes weights $\{w_i^v\}_{i=1}^n, \{w_j^u\}_{j=1}^m$ are assigned, respectively. Intuitively, the set of weights $\{w_i^v\}_{i=1}^n$ can be considered as piles of earth that needs to be transferred to the holes that the other set of weights create in the feature space. The EMD measures the minimum amount of work required to transfer the piles of earth to the holes. Each unit of earth is transferred from pile i to hole j with cost $d(v_i, u_j)$ (called *ground distance*) and the total amount of earth that is transferred from pile i to hole j is denoted as $f^{(i,j)}$, that is called the *flow of weight*. The transportation problem is solved with a linear programming optimization approach that finds the optimal flow of weight [RTG00] between the two distributions.

The optimal cost of the optimization process is the EMD that expresses the degree of similarity between the two signatures. It is defined as follows :

$$\frac{\sum_{i=1}^n \sum_{j=1}^m f^{(i,j)} d(v_i, u_j)}{\sum_{i=1}^n \sum_{j=1}^m f^{(i,j)}} \quad (3)$$

In this context, the distributions are the two ARGs that need to be matched and the signatures are their nodes to which proper weights are given. The ground distances are defined by the unary and binary attributes of the ARGs. The similarity of the ARGs is expressed by the EMD.

Let $G = (V, E, \mathbf{U}, \mathbf{B})$, $\hat{G} = (\hat{V}, \hat{E}, \hat{\mathbf{U}}, \hat{\mathbf{B}})$ be the attributed relational graphs that need to be matched, where $V = \{v_i\}_{i=1}^n$, $\hat{V} = \{\hat{v}_j\}_{j=1}^m$ are the nodes (v_1, \hat{v}_1 represent the core component of the two objects, respectively), $E = \{r_{1i}\}_{i=2}^n$, $\hat{E} = \{\hat{r}_{1j}\}_{j=2}^m$ are the edges, $\mathbf{U} = \{\mathbf{u}_i\}_{i=1}^n$, $\hat{\mathbf{U}} = \{\hat{\mathbf{u}}_j\}_{j=1}^m$ are the unary attributes of the nodes and $\mathbf{B} = \{\mathbf{b}_i\}_{i=2}^n$, $\hat{\mathbf{B}} = \{\hat{\mathbf{b}}_j\}_{j=2}^m$ are the binary attributes of the edges of the two graphs, respectively. We assume also that $n \geq m$. Weights $\{w_i^v\}_{i=1}^n, \{w_j^u\}_{j=1}^m$ are assigned to the nodes of the graphs and each of them is equal to $\frac{1}{n}$ so as to attain a uniform distribution. As already mentioned, it is assumed that the nodes v_1, \hat{v}_1 represent the core component of the two objects, respectively. These nodes are considered as *fixed* and are always matched in our matching algorithm. Also additional $n - m$ nodes are inserted in \hat{G} which are called in this work *delete* nodes. The reason for doing this is to penalize the $n - m$ nodes of G that are not mapped to any of the nodes of \hat{G} from within the EMD-based matching process. By this way we take into consideration partial matching. Unary attributes $\hat{\mathbf{U}}_d = \{\hat{\mathbf{u}}_{d_j}\}_{j=m+1}^n$ are assigned to the delete nodes that correspond to components with no information. Weights also equal to $\frac{1}{n}$ are assigned to the delete nodes. All other nodes representing the protrusions of the objects are considered as *normal*.

The ground distance assignment is as follows:

$$d(v_i, \hat{v}_j) = \begin{cases} 3 \frac{D_{normal}(v_i, \hat{v}_j)}{1 + D_{normal}(v_i, \hat{v}_j)} & \text{if } v_i, \hat{v}_j \text{ normal} \\ 3 \frac{D_{fixed}(v_i, \hat{v}_j)}{1 + D_{fixed}(v_i, \hat{v}_j)} & \text{if } v_i, \hat{v}_j \text{ fixed} \\ 5 \frac{D_{delete}(v_i, \hat{v}_j)}{0.1 + D_{delete}(v_i, \hat{v}_j)} & \text{if } v_i \text{ normal,} \\ & \hat{v}_j \text{ delete} \\ \infty & \text{otherwise} \end{cases} \quad (4)$$

where,

$$\begin{aligned} D_{normal}(v_i, \hat{v}_j) &= \sqrt{\|\mathbf{u}_i - \hat{\mathbf{u}}_j\|^2 + \|\mathbf{b}_i - \hat{\mathbf{b}}_j\|^2} \\ D_{fixed}(v_i, \hat{v}_j) &= \sqrt{\|\mathbf{u}_i - \hat{\mathbf{u}}_j\|^2} \\ D_{delete}(v_i, \hat{v}_j) &= \sqrt{\|\mathbf{u}_i - \hat{\mathbf{u}}_{d_j}\|^2} \end{aligned} \quad (5)$$

The motivation for choosing the particular values for $d(v_i, \hat{v}_j)$ was to amplify the dissimilarity cases.

As can be seen in equation (5), the binary attributes are considered only in the normal nodes since we want to exploit the relation that they have with the fixed node (core).

When the fixed nodes are matched, only the unary attributes are considered since the core relation with the other nodes is already considered when the normal nodes are matched. Note also that with the selected ground distance the fixed nodes are always going to be matched.

In this paper the following attributes are used :

- i. The unary and binary attributes of Kim *et. al.* [KPYL04]. The purpose of this assignment is to compare the proposed matching methodology with that used in [KPYL04] in order to show the efficiency of our segmentation and matching algorithms;
- ii. Unary attributes defined by Papadakis *et. al.* [PPPT07] descriptor. In this case we do not use any binary attributes.

Considering Kim *et. al.* [KPYL04] attribute assignment, the unary attributes that are assigned to the nodes of the ARG representing the object components are the relative size (rs) of the component, the convexity (c) of the component and the eccentricities (e_1, e_2) of the ellipsoid approximating the component. The binary attributes that are assigned to the edges of the ARG are the distance (l) of the centroids of the components connected by an edge of the graph and the angles (a_1, a_2) that the two most significant principal axes of the connected components create with each other. All of the attributes are normalized in the interval $[0, 1]$. By this way, the vector $[rs, c, e_1, e_2]$ is assigned to the normal and fixed nodes and the vector $[l, a_1, a_2]$ is assigned to the edges of the graphs. All delete nodes are assigned the vector $[0, 1, 1, 1]$. In equation (5), the norm $\|\cdot\|$ denotes the L_2 norm of the attribute vectors.

Considering Papadakis *et. al.* [PPPT07] attribute assignment, we set to the normal and fixed nodes their spherical harmonic descriptor vector and to the delete nodes the vector with zero entries whose dimension is the same as their descriptor. Please note that in this case we do not assign any binary attributes to the graphs, thus in equation (5) there exist no binary term and the norm $\|\cdot\|$ denotes the L_1 norm of the spherical harmonic vectors which is defined as in [PPPT07]. For the EMD computation the implementation of Rubner *et. al.* [RTG00] is used (<http://www.cs.duke.edu/~tomasi/software/emd.htm>).

4. Experimental Results

In this paper, the 3D object database used is the McGill Database (<http://www.cim.mcgill.ca/~shape/benchMark/>) which contains a rich variety of articulated objects. The database contains 255 objects which are divided into ten categories, namely, ‘Ants’, ‘Crabs’, ‘Spectacles’, ‘Hands’, ‘Humans’, ‘Octopuses’, ‘Pliers’, ‘Snakes’, ‘Spiders’ and ‘Teddy-bears’. Since the proposed segmentation algorithm requires that the objects should be manifolds, all of the objects of the database have been transformed into closed manifolds.

In this paper, the goal of the experimental work is twofold. First, we demonstrate the superior performance of the proposed retrieval methodology against two other state of the art retrieval methodologies, namely Kim *et. al.* [KPYL04] and Papadakis *et. al.* [PPT*08]. Second, we demonstrate the efficiency of the proposed mesh segmentation against Kim *et. al.* [KPYL04] segmentation in terms of retrieval accuracy. The latter goal is achieved by accommodating the ARG of the 3D object produced by the proposed segmentation algorithm to the MPEG7 [KPYL04] matching process.

In the sequel, we will use the following abbreviations:

- The graph-based retrieval methodology that uses the proposed segmentation algorithm and EMD-based matching using Papadakis *et. al.* [PPPT07] attributes is denoted as **EMD-PPPT**.
- The graph-based retrieval methodology that uses the proposed segmentation algorithm and EMD-based matching using Kim *et. al.* [KPYL04] attributes is denoted as **EMD-MPEG7**.
- The graph-based retrieval methodology that uses the proposed segmentation algorithm and the graph matching of Kim *et. al.* [KPYL04] is denoted as **SMPEG7**.
- The graph-based retrieval methodology that uses the segmentation and matching of Kim *et. al.* [KPYL04] is denoted as **MPEG7**.
- The retrieval methodology of Papadakis *et. al.* [PPT*08] that uses a global shape representation is denoted as **Hybrid**.

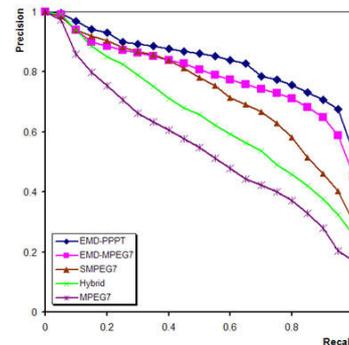


Figure 5: PR plot of the examined retrieval methodologies

The evaluation of the retrieval results is based upon the Precision-Recall (PR) plots (higher Precision indicates better performance) as well as the quantification measures in the following:

- Nearest Neighbor (NN): The percentage of queries where the closest match belongs to the query’s class.
- First Tier (FT): The recall for the ($k-1$) closest matches, where k is the cardinality of the query’s class.

- Second Tier (ST): The recall for the $2(k-1)$ closest matches, where k is the cardinality of the query's class.
- Discounted Cumulative Gain (DCG): A statistic that weights correct results near the front of the list more than correct results later in the ranked list under the assumption that a user is less likely to consider elements near the end of the list.

The above measures range from 0% to 100% and higher values indicate better performance.

In Figure 5 the PR plot is shown for the whole database. As can be seen, EMD-PPPT performs best. It can also be observed that SMPEG7 is significantly better than MPEG7. This shows that our segmentation algorithm is performing better in this database. Also the Hybrid algorithm is inferior to the proposed retrieval methodology.

In Figure 6 the PR plots for some representative classes of the McGill database are shown.

In Table 1, the corresponding scores for each of the retrieval methodologies for each class of the database as well as the average scores for the complete McGill database are shown. As can be observed the EMD-PPPT methodology performs better in total and in most of the classes of the database.

In Figure 7, representative retrieval results are shown for a 'hands' object using the EMD-PPPT and MPEG7 retrieval methodologies.

5. Conclusions

In this paper a graph-based retrieval methodology is proposed. The method builds the structural description of the object using the proposed segmentation algorithm that produces meaningful results. This structural description is represented by an attributed relational graph (ARG). A query retrieval is performed by matching the query's ARG with the ARGs of the database objects using our EMD-based matching algorithm.

Our methodology is very efficient in retrieving articulated objects and was shown to perform significantly better in our extensive evaluation against the compared state of the art retrieval algorithms in the McGill Database of articulated objects.

References

- [APP*07] 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.
- [BCG08] BEN-CHEN M., GOTSMAN C.: Characterizing shape using conformal factors. In *Eurographics Workshop on 3D Object Retrieval* (2008), pp. 1–8.
- [Bie87] BIEDERMAN I.: Recognition-by-components: A theory of human image understanding. *Psychological Review* 94, 2 (1987), 115–147.

Table 1: Quantitative measure scores of the examined retrieval methodologies

| Class | Method | NN(%) | FT(%) | ST(%) | DCG(%) |
|--------------------|-----------|-------------|-------------|-------------|-------------|
| Complete McGill db | EMD-PPPT | 97.6 | 74.1 | 91.1 | 93.3 |
| | EMD-MPEG7 | 93.3 | 69.2 | 88.9 | 90.8 |
| | SMPEG7 | 91.8 | 65.2 | 78.3 | 89.1 |
| | Hybrid | 92.5 | 55.7 | 69.8 | 85.0 |
| | MPEG7 | 83.9 | 47.5 | 63.2 | 79.2 |
| Ants | EMD-PPPT | 96.7 | 54.9 | 79.7 | 88.4 |
| | EMD-MPEG7 | 96.7 | 58.5 | 79.9 | 87.5 |
| | SMPEG7 | 80.0 | 57.1 | 75.6 | 86.7 |
| | Hybrid | 100 | 73.6 | 89.2 | 94.8 |
| | MPEG7 | 90.0 | 62.1 | 75.5 | 87.1 |
| Crabs | EMD-PPPT | 100 | 98.2 | 99.8 | 99.9 |
| | EMD-MPEG7 | 100 | 89.8 | 98.2 | 99.2 |
| | SMPEG7 | 100 | 72.9 | 90.3 | 95.9 |
| | Hybrid | 100 | 55.2 | 71.8 | 88.7 |
| | MPEG7 | 90.0 | 45.9 | 65.5 | 82.2 |
| Spectacles | EMD-PPPT | 100 | 70.3 | 99.8 | 94.0 |
| | EMD-MPEG7 | 96.0 | 63.7 | 94.3 | 89.2 |
| | SMPEG7 | 96.0 | 55.8 | 63.7 | 82.7 |
| | Hybrid | 96.0 | 53.5 | 63.3 | 85.9 |
| | MPEG7 | 84.0 | 37.8 | 50.8 | 73.6 |
| Hands | EMD-PPPT | 95.0 | 83.9 | 88.9 | 95.2 |
| | EMD-MPEG7 | 95.0 | 79.7 | 88.2 | 93.4 |
| | SMPEG7 | 95.0 | 78.7 | 87.9 | 93.0 |
| | Hybrid | 90.0 | 43.4 | 57.6 | 77.8 |
| | MPEG7 | 60.0 | 30.0 | 41.3 | 63.1 |
| Humans | EMD-PPPT | 96.6 | 93.5 | 96.4 | 98.1 |
| | EMD-MPEG7 | 96.6 | 86.8 | 99.3 | 97.4 |
| | SMPEG7 | 96.6 | 84.5 | 98.0 | 97.3 |
| | Hybrid | 100 | 47.0 | 63.8 | 83.1 |
| | MPEG7 | 79.3 | 40.5 | 59.1 | 77.9 |
| Octopuses | EMD-PPPT | 88.0 | 58.8 | 81.8 | 88.1 |
| | EMD-MPEG7 | 80.0 | 45.2 | 73.2 | 79.1 |
| | SMPEG7 | 84.0 | 42.0 | 63.0 | 80.5 |
| | Hybrid | 56.0 | 29.5 | 45.0 | 68.9 |
| | MPEG7 | 72.0 | 46.8 | 76.2 | 77.8 |
| Pliers | EMD-PPPT | 100 | 100 | 100 | 100 |
| | EMD-MPEG7 | 100 | 85.5 | 100 | 98.6 |
| | SMPEG7 | 100 | 86.1 | 95.5 | 97.8 |
| | Hybrid | 100 | 71.6 | 87.9 | 94.6 |
| | MPEG7 | 95.0 | 65.5 | 77.9 | 89.5 |
| Snakes | EMD-PPPT | 100 | 43.2 | 95.2 | 84.7 |
| | EMD-MPEG7 | 80.0 | 46.2 | 85.8 | 83.4 |
| | SMPEG7 | 80.0 | 44.2 | 48.0 | 76.6 |
| | Hybrid | 80.0 | 23.7 | 28.7 | 62.4 |
| | MPEG7 | 76.0 | 36.8 | 40.7 | 69.3 |
| Spiders | EMD-PPPT | 100 | 87.2 | 100 | 98.4 |
| | EMD-MPEG7 | 100 | 85.7 | 97.3 | 97.5 |
| | SMPEG7 | 96.8 | 74.8 | 86.6 | 93.9 |
| | Hybrid | 100 | 71.5 | 91.0 | 93.7 |
| | MPEG7 | 90.3 | 37.3 | 61.8 | 77.8 |
| Teddy-bears | EMD-PPPT | 100 | 45.3 | 63.2 | 83.9 |
| | EMD-MPEG7 | 85.0 | 42.6 | 66.3 | 78.8 |
| | SMPEG7 | 90.0 | 55.8 | 70.8 | 84.6 |
| | Hybrid | 100 | 90.3 | 98.4 | 99.1 |
| | MPEG7 | 100 | 79.2 | 84.5 | 93.4 |

- [BKS*04] BUSTOS B., KEIM D., SAUPE D., SCHRECK T., VRANIC D.: Automatic selection and combination of descriptors for effective 3D similarity search. In *IEEE Sixth Int. Symp. on Multimedia Software Engineering* (2004), pp. 514–521.
- [FS06] FUNKHOUSER T., SHILANE P.: Partial matching of 3D shapes with priority-driven search. In *Fourth Eurographics symposium on Geometry processing* (2006), pp. 131–142.
- [GSCO07] GAL R., SHAMIR A., COHEN-OR D.: Pose oblivious shape signature. *IEEE Transactions of Visualization and Computer Graphics* 13, 2 (2007), 261–271.
- [HR84] HOFFMAN D., RICHARDS W.: Parts of recognition. *Cognition* 18 (1984), 65–96.
- [HSKK01] HILAGA M., SHINAGAWA Y., KOMURA T., KUNII T. L.: Topology matching for full automatic similarity estimation of 3d. In *ACM SIGGRAPH* (2001), pp. 203–212.
- [JZ07] JAIN V., ZHANG H.: A spectral approach to shape-based

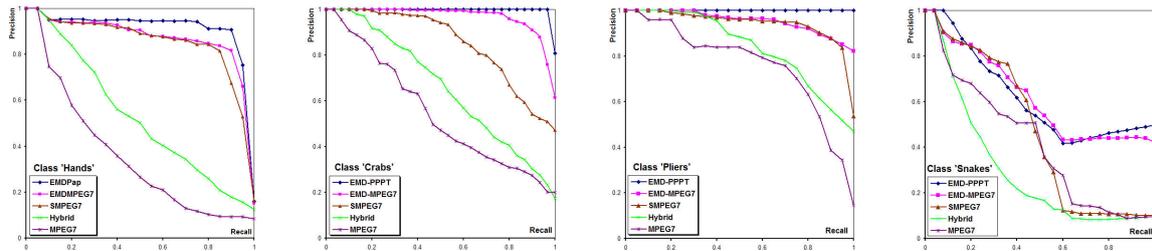


Figure 6: PR plots of McGill database representative classes

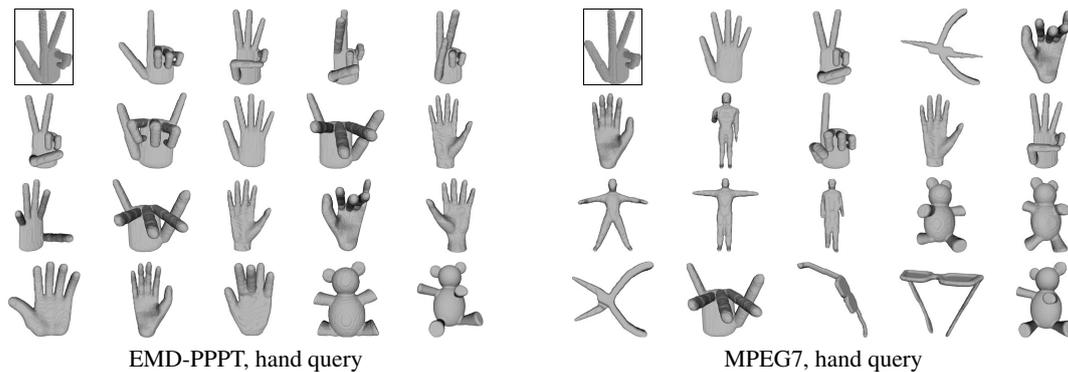


Figure 7: Retrieval results for a 'hands' query object. The query object is on the top left side and the ranking order goes from left to right.

retrieval of articulated 3D models. *Comp. Aided Des.* 39, 5 (2007), 398–407.

[KFR03] KAZHDAN M., FUNKHOUSER T., RUSINKIEWICZ S.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In *Eurographics/ACM SIGGRAPH symposium on Geometry processing* (2003), pp. 156–164.

[KPYL04] KIM D. H., PARK I. K., YUN I. D., LEE S. U.: A new mpeg-7 standard: Perceptual 3D shape descriptor. In *5th Pacific Rim Conference on Multimedia* (2004), pp. 238–245.

[KT03] KATZ S., TAL A.: Hierarchical mesh decomposition using fuzzy clustering and cuts. *ACM TOG* 22, 3 (2003), 954–961.

[LLL07] LIN H. S., LIAO H. M., LIN J.: Visual salience-guided mesh decomposition. *IEEE Transactions On Multimedia* 9, 1 (2007), 46–57.

[MDA*08] MADEMLIS A., DARAS P., AXENOPOULOS A., TZOVARAS D., STRINTZIS M. G.: Combining topological and geometrical features for global and partial 3-D shape retrieval. *IEEE Transactions On Multimedia* 10, 5 (2008), 819–831.

[MSF07] MARINI S., SPAGNUOLO M., FALCIDIENO B.: Structural shape prototypes for the automatic classification of 3D objects. *Computer Graphics and Applications* 27, 4 (2007), 28–37.

[OOFB08] OHBUCHI R., OSADA K., FURUYA T., BANNO T.: Salient local visual features for shape-based 3d model retrieval. In *IEEE Int. Conf. on Shape Modeling and Applications* (2008), pp. 93–102.

[PPPT07] PAPADAKIS P., PRATIKAKIS I., PERANTONIS S., THEOHARIS T.: Efficient 3D shape matching and retrieval using a concrete radialized spherical projection representation. *Pattern Recognition* 40, 9 (2007), 2437–2452.

[PPT*08] PAPADAKIS P., PRATIKAKIS I., THEOHARIS T., PASSALIS G., PERANTONIS S.: 3D object retrieval using an efficient and compact hybrid shape descriptor. In *Eurographics Workshop on 3D Object Retrieval* (2008), pp. 9–16.

[PTK07] PASSALIS G., THEOHARIS T., KAKADIARIS I. A.: Ptk: A novel depth buffer-based shape descriptor for three-dimensional object retrieval. *Visual Computer* 23, 1 (2007), 5–14.

[RTG00] RUBNER Y., TOMASI C., GUIBAS L. J.: The earth mover’s distance as a metric for image retrieval. *International Journal of Computer Vision* 40, 2 (2000), 99–121.

[SSDZ99] SIDDIQI K., SHOKOUFANDEH A., DICKINSON S. J., ZUCKER S. W.: Shock graphs and shape matching. *Int. J. Comput. Vision* 35, 1 (1999), 13–32.

[SSGD03] SUNDAR H., SILVER D., GAGVANI N., DICKINSON S.: Skeleton based shape matching and retrieval. In *Shape modeling International* (2003), pp. 130–139.

[TZ06] TAL A., ZUCKERBERGER E.: Mesh retrieval by components. In *International Conference on Computer Graphics Theory and Applications* (2006), pp. 142–149.

[Vra04] VRANIC D. V.: *3D Model Retrieval*. PhD thesis, University of Leipzig, 2004.

[ZDA*07] ZARPALAS D., DARAS P., AXENOPOULOS A., TZOVARAS D., STRINTZIS M. G.: 3D model search and retrieval using the spherical trace transform. *EURASIP Journal on Advances in Signal Processing* (2007), Article ID 23912, 14 pages.